A Short Course on Spectral Theory

William Arveson

Springer

Graduate Texts in Mathematics 209

Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

This page intentionally left blank

William Arveson

A Short Course on Spectral Theory

Springer

William Arveson
Department of Mathematics
University of California, Berkeley
Berkeley, CA 94720-0001
USA
arveson@math.berkeley.edu

Mathematics Subject Classification (2000): 46-01, 46Hxx, 46Lxx, 47Axx, 58C40

Library of Congress Cataloging-in-Publication Data
Arveson, William.
A short course on spectral theory/William Arveson.
p. cm. — (Graduate texts in mathematics; 209)
Includes bibliographical references and index.
ISBN 0-387-95300-0 (alk. paper)
1. Spectral theory (Mathematics) I. Title. II. Series.
QA320 .A83 2001
515′.7222–dc21 2001032836

Printed on acid-free paper.

Production managed by Francine McNeill; manufacturing supervised by Jacqui Ashri.
Photocomposed copy prepared from the author's AMSLaTeX files.
Printed and bound by Maple-Vail Book Manufacturing Group, York, PA.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-95300-0 SPIN 10838691

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

To Lee

This page intentionally left blank

Preface

This book presents the basic tools of modern analysis within the context of what might be called the fundamental problem of operator theory: to calculate spectra of specific operators on infinite-dimensional spaces, especially operators on Hilbert spaces. The tools are diverse, and they provide the basis for more refined methods that allow one to approach problems that go well beyond the computation of spectra; the mathematical foundations of quantum physics, noncommutative K-theory, and the classification of simple C^*-algebras being three areas of current research activity that require mastery of the material presented here.

The notion of spectrum of an operator is based on the more abstract notion of the spectrum of an element of a complex Banach algebra. After working out these fundamentals we turn to more concrete problems of computing spectra of operators of various types. For normal operators, this amounts to a treatment of the spectral theorem. Integral operators require the development of the Riesz theory of compact operators and the ideal $\mathcal{L}^2$ of Hilbert–Schmidt operators. Toeplitz operators require several important tools; in order to calculate the spectra of Toeplitz operators with continuous symbol one needs to know the theory of Fredholm operators and index, the structure of the Toeplitz C^*-algebra and its connection with the topology of curves, and the index theorem for continuous symbols.

I have given these lectures several times in a fifteen-week course at Berkeley (Mathematics 206), which is normally taken by first- or second-year graduate students with a foundation in measure theory and elementary functional analysis. It is a pleasure to teach that course because many deep and important ideas emerge in natural ways. My lectures have evolved significantly over the years, but have always focused on the notion of spectrum and the role of Banach algebras as the appropriate modern foundation for such considerations. For a serious student of modern analysis, this material is the essential beginning.

Berkeley, California — William Arveson
July 2001

This page intentionally left blank

Contents

CHAPTER 1

Spectral Theory and Banach Algebras

The spectrum of a bounded operator on a Banach space is best studied within the context of Banach algebras, and most of this chapter is devoted to the theory of Banach algebras. However, one should keep in mind that it is the spectral theory of operators that we want to understand. Many examples are discussed in varying detail. While the general theory is elegant and concise, it depends on its power to simplify and illuminate important examples such as those that gave it life in the first place.

1.1. Origins of Spectral Theory

The idea of the spectrum of an operator grew out of attempts to understand concrete problems of linear algebra involving the solution of linear equations and their infinite-dimensional generalizations.

The fundamental problem of linear algebra over the complex numbers is the solution of systems of linear equations. One is given

(a) an $n \times n$ matrix (a_{ij}) of complex numbers,
(b) an n-tuple $g = (g_1, g_2, \dots, g_n)$ of complex numbers,

and one attempts to solve the system of linear equations

$$\begin{aligned} a_{11}f_1 + \cdots + a_{1n}f_n &= g_1, \\ &\cdots \\ a_{n1}f_1 + \cdots + a_{nn}f_n &= g_n \end{aligned} \tag{1.1}$$

for $f = (f_1, \dots, f_n) \in \mathbb{C}^n$. More precisely, one wants to determine if the system (1.1) has solutions and to find all solutions when they exist.

Elementary courses on linear algebra emphasize that the left side of (1.1) defines a linear operator $f \mapsto Af$ on the n-dimensional vector space $\mathbb{C}^n$. The existence of solutions of (1.1) for any choice of g is equivalent to surjectivity of A; uniqueness of solutions is equivalent to injectivity of A. Thus the system of equations (1.1) is uniquely solvable for all choices of g if and only if the linear operator A is invertible. This ties the idea of invertibility to the problem of solving (1.1), and in this finite-dimensional case there is a simple criterion: The operator A is invertible precisely when the determinant of the matrix (a_{ij}) is nonzero.

However elegant it may appear, this criterion is of limited practical value, since the determinants of large matrices can be prohibitively hard to compute. In infinite dimensions the difficulty lies deeper than that, because for most operators on an infinite-dimensional Banach space there is no meaningful concept of determinant. Indeed, there is no numerical invariant for operators that determines invertibility in infinite dimensions as the determinant does in finite dimensions.

In addition to the idea of invertibility, the second general principle behind solving (1.1) involves the notion of eigenvalues. And in finite dimensions, spectral theory reduces to the theory of eigenvalues. More precisely, eigenvalues and eigenvectors for an operator A occur in pairs (λ, f), where $Af = \lambda f$. Here, f is a nonzero vector in $\mathbb{C}^n$ and λ is a complex number. If we fix a complex number λ and consider the set $V_\lambda \subseteq \mathbb{C}^n$ of all vectors f for which $Af = \lambda f$, we find that V_λ is always a linear subspace of $\mathbb{C}^n$, and for most choices of λ it is the trivial subspace $\{0\}$. V_λ is nontrivial if and only if the operator $A - \lambda\mathbf{1}$ has nontrivial kernel: equivalently, if and only if $A - \lambda\mathbf{1}$ is *not* invertible. The spectrum $\sigma(A)$ of A is defined as the set of all such $\lambda \in \mathbb{C}$, and it is a *nonempty* set of complex numbers containing no more than n elements.

Assuming that A is invertible, let us now recall how to actually calculate the solution of (1.1) in terms of the given vector g. Whether or not A is invertible, the eigenspaces $\{V_\lambda\colon \lambda \in \sigma(A)\}$ frequently do not span the ambient space $\mathbb{C}^n$ (in order for the eigenspaces to span it is necessary for A to be diagonalizable). But when they do span, the problem of solving (1.1) is reduced as follows. One may decompose g into a linear combination

$$g = g_1 + g_2 + \cdots + g_k,$$

where $g_j \in V_{\lambda_j}$, $\lambda_1, \dots, \lambda_k$ being eigenvalues of A. Then the solution of (1.1) is given by

$$f = \lambda_1^{-1} g_1 + \lambda_2^{-1} g_2 + \cdots + \lambda_k^{-1} g_k.$$

Notice that $\lambda_j \neq 0$ for every j because A is invertible. When the spectral subspaces V_λ fail to span the problem is somewhat more involved, but the role of the spectrum remains fundamental.

Remark 1.1.1. We have alluded to the fact that the spectrum of any operator on $\mathbb{C}^n$ is nonempty. Perhaps the most familiar proof involves the function $f(\lambda) = \det(A - \lambda\mathbf{1})$. One notes that f is a nonconstant polynomial with complex coefficients whose zeros are the points of $\sigma(A)$, and then appeals to the fundamental theorem of algebra. For a proof that avoids determinants see [**5**].

The fact that the complex number field is algebraically closed is central to the proof that $\sigma(A) \neq \emptyset$, and in fact an operator acting on a real vector space need not have any eigenvalues at all: consider a 90 degree rotation about the origin as an operator on $\mathbb{R}^2$. For this reason, spectral theory concerns complex linear operators on complex vector spaces and their infinite-dimensional generalizations.

We now say something about the extension of these results to infinite dimensions. For example, if one replaces the sums in (1.1) with integrals, one obtains a class of problems about integral equations. Rather than attempt a general definition of that term, let us simply look at a few examples in a somewhat formal way, though it would not be very hard to make the following discussion completely rigorous. Here are some early examples of integral equations.

EXAMPLE 1.1.2. This example is due to Niels Henrik Abel (ca 1823), whose name is attached to abelian groups, abelian functions, abelian von Neumann algebras, and the like. Abel considered the following problem. Fix a number α in the open unit interval and let g be a suitably smooth function on the interval $(0,1)$ satisfying $g(\alpha) = 0$. Abel was led to seek a function f for which

$$\int_\alpha^x \frac{1}{(x-y)^\alpha} f(y)\,dy = g(x)$$

on the interval $\alpha < x < 1$, and he wrote down the following "solution":

$$f(y) = \frac{\sin \pi\alpha}{\pi} \int_\alpha^y \frac{g'(x)}{(y-x)^{2-\alpha}}\,dx.$$

EXAMPLE 1.1.3. Given a function $g \in L^2(\mathbb{R})$, find a function f such that

$$\int_{-\infty}^{\infty} e^{ixy} f(y)\,dy = g(x), \qquad x \in \mathbb{R}. \tag{1.2}$$

The "solution" of this problem is the following:

$$f(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixy} g(x)\,dx.$$

In fact, one has to be careful about the meaning of these two integrals. But in an appropriate sense the solution f is uniquely determined, it belongs to $L^2(\mathbb{R})$, and the Fourier transform operator defined by the left side of (1.2) is an invertible operator on L^2. Indeed, it is a scalar multiple of an invertible isometry whose inverse is exhibited above. This is the essential statement of the Plancherel theorem [**15**].

EXAMPLE 1.1.4. This family of examples goes back to Vito Volterra (ca 1900). Given a continuous complex-valued function $k(x,y)$ defined on the triangle $0 \le y \le x \le 1$ and given $g \in C[0,1]$, find a function f such that

$$\int_0^x k(x,y) f(y)\,dy = g(x), \qquad 0 \le x \le 1. \tag{1.3}$$

This is often called a Volterra equation of the first kind. A Volterra equation of the second kind involves a given complex parameter λ as well as a function $g \in C[0,1]$, and asks whether or not the equation

$$\int_0^x k(x,y) f(y)\,dy - \lambda f(x) = g(x), \qquad 0 \le x \le 1 \tag{1.4}$$

can be solved for f.

We will develop powerful methods that are effective for a broad class of problems including those of Example 1.1.4. For example, we will see that the spectrum of the operator $f \mapsto Kf$ defined on the Banach space $C[0,1]$ by the left side of (1.3) satisfies $\sigma(K) = \{0\}$. One deduces that for every $\lambda \neq 0$ and every $g \in C[0,1]$, the equation (1.4) has a unique solution $f \in C[0,1]$. Significantly, there are no "formulas" for these solution functions, as we had in Examples 1.1.2 and 1.1.3.

Exercises. The first two exercises illustrate the problems that arise when one attempts to develop a determinant theory for operators on an infinite-dimensional Banach space. We consider the simple case of diagonal operators acting on the Hilbert space $\ell^2 = \ell^2(\mathbb{N})$ of all square summable sequences of complex numbers. Fix a sequence of positive numbers $a_1, a_2, \dots$ satisfying $0 < \epsilon \leq a_n \leq M < \infty$ and consider the operator A defined on ℓ^2 by

$$(Ax)_n = a_n x_n, \qquad n = 1, 2, \dots, \qquad x \in \ell^2. \tag{1.4}$$

(1) Show that A is a bounded operator on ℓ^2, and exhibit a bounded operator B on ℓ^2 such that $AB = BA = \mathbf{1}$ where $\mathbf{1}$ is the identity operator.

One would like to have a notion of determinant with at least these two properties: $D(\mathbf{1}) = 1$ and $D(ST) = D(S)D(T)$ for operators S, T on ℓ^2. It follows that such a "determinant" will satisfy $D(A) \neq 0$ for the operators A of (1.4). It is also reasonable to expect that for these operators we should have

$$D(A) = \lim_{n \to \infty} a_1 a_2 \cdots a_n. \tag{1.5}$$

(2) Let $a_1, a_2, \dots$ be a bounded monotone increasing sequence of positive numbers and let $D_n = a_1 a_2 \cdots a_n$. Show that the sequence D_n converges to a nonzero limit $D(A)$ iff

$$\sum_{n=1}^{\infty} (1 - a_n) < \infty.$$

Thus, this attempt to define a reasonable notion of determinant fails, even for invertible diagonal operators of the form (1.4) with sequences such as $a_n = n/(n+1)$, $n = 1, 2, \dots$. On the other hand, it is possible to develop a determinant theory for *certain* invertible operators, namely operators $A = \mathbf{1} + T$, where T is a "trace-class" operator; for diagonal operators defined by a sequence as in (1.4) this requirement is that

$$\sum_{n=1}^{\infty} |1 - a_n| < \infty.$$

The following exercises relate to Volterra operators on the Banach space $C[0,1]$ of continuous complex-valued functions f on the unit interval, with sup norm

$$\|f\| = \sup_{0\leq x\leq 1} |f(x)|.$$

Exercise (3) implies that Volterra operators are bounded, and the result of Exercise (5) implies that they are in fact compact operators.

(3) Let $k(x,y)$ be a Volterra kernel as in Example (1.1.4), and let $f \in C[0,1]$. Show that the function g defined on the unit interval by equation (1.3) is continuous, and that the linear map $K : f \to g$ defines a bounded operator on $C[0,1]$.

(4) For the kernel $k(x,y) = 1$ for $0 \leq y \leq x \leq 1$ consider the corresponding Volterra operator $V : C[0,1] \to C[0,1]$, namely

$$Vf(x) = \int_0^x f(y)\,dy, \qquad f \in C[0,1].$$

Given a function $g \in C[0,1]$, show that the equation $Vf = g$ has a solution $f \in C[0,1]$ iff g is continuously differentiable and $g(0) = 0$.

(5) Let $k(x,y)$, $0 \leq x,y \leq 1$, be a continuous function defined on the unit square, and consider the bounded operator K defined on $C[0,1]$ by

$$Kf(x) = \int_0^1 k(x,y)f(y)\,dy, \qquad 0 \leq x \leq 1.$$

Let $B_1 = \{f \in C[0,1] : \|f\| \leq 1\}$ be the closed unit ball in $C[0,1]$. Show that K is a *compact* operator in the sense that the norm closure of the image KB_1 of B_1 under K is a compact subset of $C[0,1]$. Hint: Show that there is a positive constant M such that for every $g \in KB_1$ and every $x,y \in [0,1]$ we have $|g(x) - g(y)| \leq M \cdot |x-y|$.

1.2. The Spectrum of an Operator

Throughout this section, E will denote a complex Banach space. By an *operator* on E we mean a bounded linear transformation $T : E \to E$; $\mathcal{B}(E)$ will denote the space of all operators on E. $\mathcal{B}(E)$ is itself a complex Banach space with respect to the operator norm. We may compose two operators $A, B \in \mathcal{B}(E)$ to obtain an operator product $AB \in \mathcal{B}(E)$, and this defines an associative multiplication satisfying both distributive laws $A(B+C) = AB + AC$ and $(A+B)C = AB + BC$. We write $\mathbf{1}$ for the identity operator.

THEOREM 1.2.1. *For every $A \in \mathcal{B}(E)$, the following are equivalent.*

(1) *For every $y \in E$ there is a unique $x \in E$ such that $Ax = y$.*

(2) *There is an operator $B \in \mathcal{B}(E)$ such that $AB = BA = \mathbf{1}$.*

PROOF. We prove the nontrivial implication (1) $\implies$ (2). The hypothesis (1) implies that A is invertible as a linear transformation on the vector space E, and we may consider its inverse $B : E \to E$. As a subset of $E \oplus E$, the graph of B is related to the graph of A as follows:

$$\Gamma(B) = \{(x, Bx) : x \in E\} = \{(Ay, y) : y \in E\}.$$

The space on the right is closed in $E \oplus E$ because A is continuous. Hence the graph of B is closed, and the closed graph theorem implies $B \in \mathcal{B}(E)$. □

DEFINITION 1.2.2. Let $A \in \mathcal{B}(E)$.

(1) A is said to be invertible if there is an operator $B \in \mathcal{B}(E)$ such that $AB = BA = \mathbf{1}$.
(2) The spectrum $\sigma(A)$ of A is the set of all complex numbers λ for which $A - \lambda\mathbf{1}$ is not invertible.
(3) The resolvent set $\rho(A)$ of A is the complement $\rho(A) = \mathbb{C} \setminus \sigma(A)$.

In Examples (1.1.2)–(1.1.4) of the previous section, we were presented with an operator, and various assertions were made about its spectrum. For example, in order to determine whether a given operator A is invertible, one has exactly the problem of determining whether or not $0 \in \sigma(A)$. The spectrum is the most important invariant attached to an operator.

REMARK 1.2.3. *Remarks on operator spectra.* We have defined the spectrum of an operator $T \in \mathcal{B}(E)$, but it is often useful to have more precise information about various points of $\sigma(T)$. For example, suppose there is a nonzero vector $x \in E$ for which $Tx = \lambda x$ for some complex number λ. In this case, λ is called an eigenvalue (with associated eigenvector x). Obviously, $T - \lambda\mathbf{1}$ is not invertible, so that $\lambda \in \sigma(T)$. The set of all eigenvalues of T is a subset of $\sigma(T)$ called the *point spectrum* of T (and is written $\sigma_p(T)$). When E is finite dimensional $\sigma(T) = \sigma_p(T)$, but that is not so in general. Indeed, many of the natural operators of analysis have no point spectrum at all.

Another type of spectral point occurs when $T - \lambda$ is one-to-one but not onto. This can happen in two ways: Either the range of $T - \lambda$ is not closed in E, or it is closed but not all of E. Terminology has been invented to classify such behavior (*compression spectrum*, *residual spectrum*), but we will not use it, since it is better to look at a good example. Consider the Volterra operator V acting on $C[0, 1]$ as follows:

$$Vf(x) = \int_0^x f(t)\, dt, \qquad 0 \le x \le 1.$$

This operator is not invertible; in fact, we will see later that its spectrum is exactly $\{0\}$. On the other hand, one may easily check that V is one-to-one. The result of Exercise (4) in section 1 implies that its range is not closed and the closure of its range is a subspace of codimension one in $C[0, 1]$.

Exercises.

(1) Give explicit examples of bounded operators A, B on $\ell^2(\mathbb{N})$ such that $AB = \mathbf{1}$ and BA is the projection onto a closed infinite-dimensional subspace of infinite codimension.

(2) Let A and B be the operators defined on $\ell^2(\mathbb{N})$ by

$$A(x_1, x_2, \dots) = (0, x_1, x_2, \dots),$$
$$B(x_1, x_2, \dots) = (x_2, x_3, x_4, \dots),$$

for $x = (x_1, x_2, \dots) \in \ell^2(\mathbb{N})$. Show that $\|A\| = \|B\| = 1$, and compute both BA and AB. Deduce that A is injective but not surjective, B is surjective but not injective, and that $\sigma(AB) \neq \sigma(BA)$.

(3) Let E be a Banach space and let A and B be bounded operators on E. Show that $\mathbf{1} - AB$ is invertible if and only if $\mathbf{1} - BA$ is invertible. Hint: Think about how to relate the *formal* Neumann series for $(\mathbf{1} - AB)^{-1}$,

$$(\mathbf{1} - AB)^{-1} = \mathbf{1} + AB + (AB)^2 + (AB)^3 + \dots,$$

to that for $(\mathbf{1} - BA)^{-1}$ and turn your idea into a rigorous proof.

(4) Use the result of the preceding exercise to show that for any two bounded operators A, B acting on a Banach space, $\sigma(AB)$ and $\sigma(BA)$ agree except perhaps for 0: $\sigma(AB) \setminus \{0\} = \sigma(BA) \setminus \{0\}$.

1.3. Banach Algebras: Examples

We have pointed out that spectral theory is useful when the underlying field of scalars is the complex numbers, and in the sequel this will always be the case.

Definition 1.3.1 (Complex algebra). By an algebra over $\mathbb{C}$ we mean a complex vector space A together with a binary operation representing multiplication $x, y \in A \mapsto xy \in A$ satisfying

(1) Bilinearity: For $\alpha, \beta \in \mathbb{C}$ and $x, y, z \in A$ we have

$$(\alpha \cdot x + \beta \cdot y)z = \alpha \cdot xz + \beta \cdot yz,$$
$$x(\alpha \cdot y + \beta \cdot z) = \alpha \cdot xy + \beta \cdot xz.$$

(2) Associativity: $x(yz) = (xy)z$.

A complex algebra may or may not have a multiplicative identity. As a rather extreme example of one that does not, let A be any complex vector space and define multiplication in A by $xy = 0$ for all x, y. When an algebra does have an identity then it is uniquely determined, and we denote it by $\mathbf{1}$. The identity is also called the *unit*, and an algebra with unit is called a unital algebra. A commutative algebra is one in which $xy = yx$ for every x, y.

Definition 1.3.2 (Normed algebras, Banach algebras). A normed algebra is a pair $A, \|\cdot\|$ consisting of an algebra A together with a norm $\|\cdot\| : A \to [0, \infty)$ which is related to the multiplication as follows:

$$\|xy\| \leq \|x\| \cdot \|y\|, \qquad x, y \in A.$$

A Banach algebra is a normed algebra that is a (complete) Banach space relative to its given norm.

Remark 1.3.3. We recall a useful criterion for completeness: A normed linear space E is a Banach space iff every absolutely convergent series converges. More explicitly, E is complete iff for every sequence of elements $x_n \in E$ satisfying $\sum_n \|x_n\| < \infty$, there is an element $y \in E$ such that

$$\lim_{n\to\infty} \|y - (x_1 + \cdots + x_n)\| = 0;$$

see Exercise (1) below.

The following examples of Banach algebras illustrate the diversity of the concept.

Example 1.3.4. Let E be any Banach space and let A be the algebra $\mathcal{B}(E)$ of all bounded operators on E, $x \cdot y$ denoting the operator product. This is a unital Banach algebra in which the identity satisfies $\|\mathbf{1}\| = 1$. It is complete because E is complete.

Example 1.3.5. $C(X)$. Let X be a compact Hausdorff space and consider the unital algebra $C(X)$ of all complex valued continuous functions defined on X, the multiplication and addition being defined pointwise, $fg(x) = f(x)g(x)$, $(f+g)(x) = f(x)+g(x)$. Relative to the sup norm, $C(X)$ becomes a commutative Banach algebra with unit.

Example 1.3.6. *The disk algebra.* Let $D = \{z \in \mathbb{C} : |z| \leq 1\}$ be the closed unit disk in the complex plane and let A denote the subspace of $C(D)$ consisting of all complex functions f whose restrictions to the interior $\{z : |z| < 1\}$ are analytic. A is obviously a unital subalgebra of $C(D)$. To see that it is closed (and therefore a commutative Banach algebra in its own right) notice that if f_n is any sequence in A that converges to f in the norm of $C(D)$, then the restriction of f to the interior of D is the uniform limit on compact sets of the restrictions f_n and hence is analytic there.

This example is the simplest nontrivial example of a *function algebra.* Function algebras are subalgebras of $C(X)$ that exhibit nontrivial aspects of analyticity. They underwent spirited development during the 1960s and 1970s but have now fallen out of favor, due partly to the development of better technology for the theory of several complex variables.

Example 1.3.7. $\ell^1(\mathbb{Z})$. Consider the Banach space $\ell^1(\mathbb{Z})$ of all doubly infinite sequences of complex numbers $x = (x_n)$ with norm

$$\|x\| = \sum_{n=-\infty}^{\infty} |x_n|.$$

Multiplication in $A = \ell^1(\mathbb{Z})$ is defined by convolution:

$$(x * y)_n = \sum_{k=-\infty}^{\infty} x_k y_{n-k}, \qquad x, y \in A.$$

This is another example of a commutative unital Banach algebra, one that is rather different from any of the previous examples. It is called the Wiener algebra (after Norbert Wiener), and plays an important role in many questions involving Fourier series and harmonic analysis. It is discussed in more detail in Section 1.10.

EXAMPLE 1.3.8. $L^1(\mathbb{R})$. Consider the Banach space $L^1(\mathbb{R})$ of all integrable functions on the real line, where as usual we identify functions that agree almost everywhere. The multiplication here is defined by convolution:

$$f * g(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\,dt, \qquad f, g \in L^1(\mathbb{R}),$$

and for this example, it is somewhat more delicate to check that all the axioms for a commutative Banach algebra are satisfied. For example, by Fubini's theorem we have

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |f(t)||g(x-t)|\,dt \right) dx = \int_{\mathbb{R}^2} |f(t)||g(x-t)|\,dx\,dt = \|f\| \cdot \|g\|,$$

and from the latter, one readily deduces that $\|f * g\| \leq \|f\| \cdot \|g\|$.

Notice that this Banach algebra has no unit. However, it has a normalized approximate unit in the sense that there is a sequence of functions $e_n \in L^1(\mathbb{R})$ satisfying $\|e_n\| = 1$ for all n with the property

$$\lim_{n\to\infty} \|e_n * f - f\| = \lim_{n\to\infty} \|f * e_n - f\| = 0, \qquad f \in L^1(\mathbb{R}).$$

One obtains such a sequence by taking e_n to be any nonnegative function supported in the interval $[-1/n, 1/n]$ that has integral 1 (see the exercises at the end of the section).

Helson's book [**15**] is an excellent reference for harmonic analysis on $\mathbb{R}$ and $\mathbb{Z}$.

EXAMPLE 1.3.9. *An extremely nonunital one.* Banach algebras may not have even approximate units in general. More generally, a Banach algebra A need not be the closed linear span of the set $A^2 = \{xy : x, y \in A\}$ of all of its products. As an extreme example of this misbehavior, let A be any Banach space and make it into a Banach algebra using the trivial multiplication $xy = 0$, $x, y \in A$.

EXAMPLE 1.3.10. *Matrix algebras.* The algebra $M_n = M_n(\mathbb{C})$ of all complex $n \times n$ matrices is a unital algebra, and there are many norms that make it into a finite-dimensional Banach algebra. For example, with respect to the norm

$$\|(a_{ij})\| = \sum_{i,j=1}^{n} |a_{ij}|,$$

M_n becomes a Banach algebra in which the identity has norm n. Other Banach algebra norms on M_n arise as in Example 1.3.4, by realizing M_n as $\mathcal{B}(E)$ where E is an n-dimensional Banach space. For these norms on M_n, the identity has norm 1.

EXAMPLE 1.3.11. *Noncommutative group algebras.* Let G be a locally compact group. More precisely, G is a group as well as a topological space, endowed with a locally compact Hausdorff topology that is compatible with the group operations in that the maps $(x, y) \in G \times G \mapsto xy \in G$ and $x \mapsto x^{-1}$ are continuous.

A simple example is the "$ax+b$" group, the group generated by dilations and translations of the real line. This group is isomorphic to the group of all 2×2 matrices of the form $\begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}$ where $a, b \in \mathbb{R}$, $a > 0$, with the obvious topology. A related class of examples consists of the groups $\mathrm{SL}(n, \mathbb{R})$ of all invertible $n \times n$ matrices of real numbers having determinant 1.

In order to define the group algebra of G we have to say a few words about Haar measure. Let $\mathcal{B}$ denote the sigma algebra generated by the topology of G (sets in $\mathcal{B}$ are called Borel sets). A Radon measure is a Borel measure $\mu : \mathcal{B} \to [0, +\infty]$ having the following two additional properties:

(1) (Local finiteness) $\mu(K)$ is finite for every compact set K.
(2) (Regularity) For every $E \in \mathcal{B}$, we have

$$\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ is compact}\}.$$

A discussion of Radon measures can be found in [**3**]. The fundamental result of A. Haar asserts essentially the following:

THEOREM 1.3.12. *For any locally compact group G there is a nonzero Radon measure μ on G that is invariant under left translations in the sense that $\mu(x \cdot E) = \mu(E)$ for every Borel set E and every $x \in G$. If ν is another such measure, then there is a positive constant c such that $\nu(E) = c \cdot \mu(E)$ for every Borel set E.*

See Hewitt and Ross [**16**] for the computation of Haar measure for specific examples such as the $ax + b$ group and the groups $\mathrm{SL}(n, \mathbb{R})$. A proof of the existence of Haar measure can be found in Loomis [**17**] or Hewitt and Ross [**16**].

We will write dx for $d\mu(x)$, where μ is a left Haar measure on a locally compact group G. The group algebra of G is the space $L^1(G)$ of all integrable functions $f : G \to \mathbb{C}$ with norm

$$\|f\| = \int_G |f(x)|\, dx,$$

and multiplication is defined by convolution:

$$f * g(x) = \int_G f(t)g(t^{-1}x)\, dt, \qquad x \in G.$$

The basic facts about the group algebra $L^1(G)$ are similar to the commutative cases $L^1(\mathbb{Z})$ and $L^1(\mathbb{R})$) we have already encountered:

(1) For $f, g \in L^1(G)$, $f * g \in L^1(G)$ and we have $\|f * g\| \leq \|f\| \cdot \|g\|$.
(2) $L^1(G)$ is a Banach algebra.
(3) $L^1(G)$ is commutative iff G is a commutative group.
(4) $L^1(G)$ has a unit iff G is a discrete group.

Many significant properties of groups are reflected in their group algebra, (3) and (4) being the simplest examples of this phenomenon. Group algebras are the subject of continuing research today, and are of fundamental importance in many fields of mathematics.

Exercises.

(1) Let E be a normed linear space. Show that E is a Banach space iff for every sequence of elements $x_n \in X$ satisfying $\sum_n \|x_n\| < \infty$, there is an element $y \in X$ such that

$$\lim_{n\to\infty} \|y - (x_1 + \cdots + x_n)\| = 0.$$

(2) Prove that the convolution algebra $L^1(\mathbb{R})$ does not have an identity.
(3) For every $n = 1, 2, \ldots$ let ϕ_n be a nonnegative function in $L^1(\mathbb{R})$ such that ϕ_n vanishes outside the interval $[-1/n, 1/n]$ and

$$\int_{-\infty}^{\infty} \phi_n(t)\, dt = 1.$$

Show that $\phi_1, \phi_2, \ldots$ is an approximate identity for the convolution algebra $L^1(\mathbb{R})$ in the sense that

$$\lim_{n\to\infty} \|f * \phi_n - f\|_1 = 0$$

for every $f \in L^1(\mathbb{R})$.
(4) Let $f \in L^1(\mathbb{R})$. The Fourier transform of f is defined as follows:

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{it\xi} f(t)\, dt, \qquad \xi \in \mathbb{R}.$$

Show that $\hat{f}$ belongs to the algebra $C_\infty(\mathbb{R})$ of all continuous functions on $\mathbb{R}$ that vanish at ∞.
(5) Show that the Fourier transform is a homomorphism of the convolution algebra $L^1(\mathbb{R})$ onto a subalgebra $\mathcal{A}$ of $C_\infty(\mathbb{R})$ which is closed under complex conjugation and separates points of $\mathbb{R}$.

1.4. The Regular Representation

Let A be a Banach algebra. Notice first that multiplication is *jointly continuous* in the sense that for any $x_0, y_0 \in A$,

$$\lim_{(x,y)\to(x_0,y_0)} \|xy - x_0 y_0\| = 0.$$

Indeed, this is rather obvious from the estimate

$$\|xy - x_0y_0\| = \|(x - x_0)y + x_0(y - y_0)\| \le \|x - x_0\|\|y\| + \|x_0\|\|y - y_0\|.$$

We now show how more general structures lead to Banach algebras, after they are renormed with an equivalent norm. Let A be a complex algebra, which is also a Banach space relative to some given norm, in such a way that multiplication is *separately* continuous in the sense that for each $x_0 \in A$ there is a constant M (depending on x_0) such that for every $x \in A$ we have

$$\|xx_0\| \le M \cdot \|x\| \quad \text{and} \quad \|x_0x\| \le M \cdot \|x\|. \tag{1.6}$$

LEMMA 1.4.1. *Under the conditions (1.6), there is a constant $c > 0$ such that*

$$\|xy\| \le c \cdot \|x\|\|y\|, \qquad x, y \in A.$$

PROOF. For every $x \in A$ define a linear transformation $L_x : A \to A$ by $L_x(z) = xz$. By the second inequality of (1.6), $\|L_x\|$ must be bounded. Consider the family of all operators $\{L_x : \|x\| \le 1\}$. This is is a set of bounded operators on A which, by the first inequality of (1.6), is pointwise bounded:

$$\sup_{\|x\|\le 1} \|L_x(z)\| < \infty, \qquad \text{for all } z \in A.$$

The Banach–Steinhaus theorem implies that this family of operators is uniformly bounded in norm, and the existence of c follows. ☐

Notice that the proof uses the completeness of A in an essential way. We now show that if A also contains a unit e, it can be renormed with an equivalent norm so as to make it into a Banach algebra in which the unit has the "correct" norm $\|e\| = 1$.

THEOREM 1.4.2. *Let A be a complex algebra with unit e that is also a Banach space with respect to which multiplication is separately continuous. Then the map $x \in A \mapsto L_x \in \mathcal{B}(A)$ defines an isomorphism of the algebraic structure of A onto a closed subalgebra of $\mathcal{B}(A)$ such that*

(1) $L_e = \mathbf{1}$.
(2) *For every $x \in A$, we have*

$$\|e\|^{-1}\|x\| \le \|L_x\| \le c\|e\|\|x\|,$$

where c is a positive constant.

In particular, $\|x\|_1 = \|L_x\|$ defines an equivalent norm on A that is a Banach algebra norm for which $\|e\|_1 = 1$.

PROOF. The map $x \mapsto L_x$ is clearly a homomorphism of algebras for which $L_e = \mathbf{1}$. By Lemma 1.4.1, we have

$$\|L_xy\| = \|xy\| \le c \cdot \|x\|\|y\|,$$

and hence $\|L_x\| \le c\|x\|$. Writing

$$\|L_x\| \ge \|L_x(e/\|e\|)\| = \frac{\|x\|}{\|e\|},$$

we see that $\|L_x\| \geq \|x\|/\|e\|$, establishing the inequality of (2).

Since the operator norm $\|x\|_1 = \|L_x\|$ is equivalent to the norm on A and since A is complete, it follows that $\{L_x : x \in A\}$ is a complete, and therefore closed, subalgebra of $\mathcal{B}(A)$. The remaining assertions follow. $\square$

The map $x \in A \mapsto L_x \in \mathcal{B}(A)$ is called the left regular representation, or simply the regular representation of A. We emphasize that if A is a nonunital Banach algebra, then the regular representation need not be one-to-one. Indeed, for the Banach algebras of Example 1.3.9, the regular representation is the zero map.

Exercises. Let E and F be normed linear spaces and let $\mathcal{B}(E, F)$ denote the normed vector space of all bounded linear operators from E to F, with norm

$$\|A\| = \sup\{\|Ax\| : x \in E, \quad \|x\| \leq 1\}.$$

We write $\mathcal{B}(E)$ for the algebra $\mathcal{B}(E, E)$ of all bounded operators on a normed linear space E. An operator $A \in \mathcal{B}(E)$ is called *compact* if the norm-closure of $\{Ax : \|x\| \leq 1\}$, the image of the unit ball under A, is a compact subset of E. Since compact subsets of E must be norm-bounded, it follows that *compact operators are bounded.*

(1) Let E and F be normed linear spaces with $E \neq \{0\}$. Show that $\mathcal{B}(E, F)$ is a Banach space iff F is a Banach space.

(2) The rank of an operator $A \in \mathcal{B}(E)$ is the dimension of the vector space AE. Let $A \in \mathcal{B}(E)$ be an operator with the property that there is a sequence of finite-rank operators $A_1, A_2, \dots$ such that $\|A - A_n\| \to 0$ as $n \to \infty$. Show that A is a compact operator.

(3) Let $a_1, a_2, \dots$ be a bounded sequence of complex numbers and let A be the corresponding diagonal operator on the Hilbert space $\ell^2 = \ell^2(\mathbb{N})$,

$$Af(n) = a_n f(n), \qquad n = 1, 2, \dots, f \in \ell^2.$$

Show that A is compact iff $\lim_{n\to\infty} a_n = 0$.

Let k be a continuous complex-valued function defined on the unit square $[0, 1] \times [0, 1]$. A simple argument shows that for every $f \in C[0, 1]$ the function Af defined on $[0, 1]$ by

$$Af(x) = \int_0^1 k(x, y) f(y)\, dy, \qquad 0 \leq x \leq 1, \tag{1.7}$$

is continuous (you may assume this in the following two exercises).

(4) Show that the operator A of (1.7) is bounded and its norm satisfies $\|A\| \leq \|k\|_\infty$, $\|\cdot\|_\infty$ denoting the sup norm in $C([0, 1] \times [0, 1])$.

(5) Show that for the operator A of (1.7), there is a sequence of finite-rank operators A_n, $n = 1, 2, \dots$, such that $\|A - A_n\| \to 0$ as $n \to \infty$ and deduce that A is compact. Hint: Start by looking at the case $k(x, y) = u(x)v(y)$ with $u, v \in C[0, 1]$.

1.5. The General Linear Group of A

Let A be a Banach algebra with unit $\mathbf{1}$, which, by the results of the previous section, we may assume satisfies $\|\mathbf{1}\| = \mathbf{1}$ after renorming A appropriately. An element $x \in A$ is said to be *invertible* if there is an element $y \in A$ such that $xy = yx = \mathbf{1}$.

Remark 1.5.1. If x is an element of A that is both left and right invertible in the sense that there are elements $y_1, y_2 \in A$ with $xy_1 = y_2x = \mathbf{1}$, then x is invertible. Indeed, that is apparent from the string of identities

$$y_2 = y_2 \cdot \mathbf{1} = y_2xy_1 = \mathbf{1} \cdot y_1 = y_1.$$

We will write A^{-1} (and occasionally $\mathrm{GL}(A)$) for the set of all invertible elements of A. It is quite obvious that A^{-1} is a group; this group is sometimes called the general linear group of the unital Banach algebra A.

Theorem 1.5.2. *If x is an element of A satisfying $\|x\| < 1$, then $\mathbf{1} - x$ is invertible, and its inverse is given by the absolutely convergent Neumann series $(\mathbf{1}-x)^{-1} = \mathbf{1}+x+x^2+\dots$. Moreover, we have the following estimates:*

$$\|(\mathbf{1} - x)^{-1}\| \leq \frac{1}{1 - \|x\|}, \tag{1.8}$$

$$\|\mathbf{1} - (\mathbf{1} - x)^{-1}\| \leq \frac{\|x\|}{1 - \|x\|}. \tag{1.9}$$

Proof. Since $\|x^n\| \leq \|x\|^n$ for every $n = 1, 2, \dots$, we can define an element $z \in A$ as the sum of the absolutely convergent series

$$z = \sum_{n=0}^{\infty} x^n.$$

We have

$$z(\mathbf{1} - x) = (\mathbf{1} - x)z = \lim_{N\to\infty} (\mathbf{1} - x) \sum_{k=1}^{N} x^k = \lim_{N\to\infty} (\mathbf{1} - x^{N+1}) = \mathbf{1};$$

hence $\mathbf{1} - x$ is invertible and its inverse is z. The inequality (1.8) follows from

$$\|z\| \leq \sum_{n=0}^{\infty} \|x^n\| \leq \sum_{n=0}^{\infty} \|x\|^n = \frac{1}{1 - \|x\|}.$$

Since

$$\mathbf{1} - z = -\sum_{n=1}^{\infty} x^n = -xz,$$

we have $\|\mathbf{1} - z\| \leq \|x\| \cdot \|z\|$, thus (1.9) follows from (1.8). ∎

Corollary 1. A^{-1} is an open set in A and $x \mapsto x^{-1}$ is a continuous map of A^{-1} to itself.

PROOF. To see that A^{-1} is open, choose an invertible element x_0 and an arbitrary element $h \in A$. We have $x_0 + h = x_0(\mathbf{1} + x_0^{-1}h)$. So if $\|x_0^{-1}h\| < 1$ then by the preceding theorem $x_0 + h$ is invertible. In particular, if $\|h\| < \|x_0^{-1}\|^{-1}$, then this condition is satisfied, proving that $x_0 + h$ is invertible when $\|h\|$ is sufficiently small.

Supposing that h has been so chosen, we can write

$$(x_0 + h)^{-1} - x_0^{-1} = (x_0(\mathbf{1} + x_0^{-1}h))^{-1} - x_0^{-1} = [(\mathbf{1} + x_0^{-1}h)^{-1} - \mathbf{1}] \cdot x_0^{-1}.$$

Thus for $\|h\| < \|x_0^{-1}\|^{-1}$ we have

$$\|(x_0 + h)^{-1} - x_0^{-1}\| \leq \|(\mathbf{1} + x_0^{-1}h)^{-1} - \mathbf{1}\| \cdot \|x_0^{-1}\| \leq \frac{\|x_0^{-1}h\| \cdot \|x_0^{-1}\|}{1 - \|x_0^{-1}h\|},$$

and the last term obviously tends to zero as $\|h\| \to 0$. $\square$

COROLLARY 2. A^{-1} is a topological group in its relative norm topology; that is,

(1) $(x, y) \in A^{-1} \times A^{-1} \mapsto xy \in A^{-1}$ is continuous, and
(2) $x \in A^{-1} \mapsto x^{-1} \in A^{-1}$ is continuous.

Exercises. Let A be a Banach algebra with unit $\mathbf{1}$ satisfying $\|\mathbf{1}\| = 1$, and let G be the topological group A^{-1}.

(1) Show that for every element $x \in A$ satisfying $\|x\| < 1$, there is a continuous function $f : [0, 1] \to G$ such that $f(0) = \mathbf{1}$ and $f(1) = (\mathbf{1} - x)^{-1}$.
(2) Show that for every element $x \in G$ there is an $\epsilon > 0$ with the following property: For every element $y \in G$ satisfying $\|y - x\| < \epsilon$ there is an arc in G connecting y to x.
(3) Let G_0 be the set of all finite products of elements of G of the form $\mathbf{1} - x$ or $(\mathbf{1} - x)^{-1}$, where $x \in A$ satisfies $\|x\| < 1$. Show that G_0 is the connected component of $\mathbf{1}$ in G. Hint: An open subgroup of G must also be closed.
(4) Deduce that G_0 is a normal subgroup of G and that the quotient topology on G/G_0 makes it into a discrete group.

The group $\Gamma = G/G_0$ is sometimes called the abstract index group of A. It is frequently (but not always) commutative even when G is not, and it is closely related to the K-theoretic group $K_1(A)$. In fact, $K_1(A)$ is in a certain sense an "abelianized" version of Γ.

We have not yet discussed the exponential map $x \in A \mapsto e^x \in A^{-1}$ of a Banach algebra A (see equation (2.2) below), but we should point out here that the connected component of the identity G_0 is also characterized as the set of all finite products of exponentials $e^{x_1}e^{x_2} \cdots e^{x_n}$, $x_1, x_2, \dots, x_n \in A$, $n = 1, 2, \dots$. When A is a commutative Banach algebra, this implies that $G_0 = \{e^x : x \in A\}$ is the range of the exponential map.

1.6. Spectrum of an Element of a Banach Algebra

Throughout this section, A will denote a unital Banach algebra for which $\|\mathbf{1}\| = 1$. One should keep in mind the operator-theoretic setting, in which A is the algebra $\mathcal{B}(E)$ of bounded operators on a complex Banach space E.

Given an element $x \in A$ and a complex number λ, it is convenient to abuse notation somewhat by writing $x - \lambda$ for $x - \lambda\mathbf{1}$.

Definition 1.6.1. For every element $x \in A$, the spectrum of x is defined as the set

$$\sigma(x) = \{\lambda \in \mathbb{C} : x - \lambda \notin A^{-1}\}.$$

We will develop the basic properties of the spectrum, the first being that it is always compact.

Proposition 1.6.2. *For every $x \in A$, $\sigma(x)$ is a closed subset of the disk $\{z \in \mathbb{C} : |z| \leq \|x\|\}$.*

Proof. The complement of the spectrum is given by

$$\mathbb{C} \setminus \sigma(x) = \{\lambda \in \mathbb{C} : x - \lambda \in A^{-1}\}.$$

Since A^{-1} is open and the map $\lambda \in \mathbb{C} \mapsto x - \lambda \in A$ is continuous, the complement of $\sigma(x)$ must be open.

To prove the second assertion, we will show that no complex number λ with $|\lambda| > \|x\|$ can belong to $\sigma(x)$. Indeed, for such a λ the formula

$$x - \lambda = (-\lambda)(\mathbf{1} - \lambda^{-1}x),$$

together with the fact that $\|\lambda^{-1}x\| < 1$, implies that $x - \lambda$ is invertible. $\square$

We now prove a fundamental result of Gelfand.

Theorem 1.6.3. $\sigma(x) \neq \emptyset$ *for every* $x \in A$.

Proof. The idea is to show that if $\sigma(x) = \emptyset$, the A-valued function $f(\lambda) = (x - \lambda)^{-1}$ is a bounded entire function that tends to zero as $\lambda \to \infty$; an appeal to Liouville's theorem yields the desired conclusion. The details are as follows.

For every $\lambda_0 \notin \sigma(x)$, $(x - \lambda)^{-1}$ is defined for all λ sufficiently close to λ_0 because $\sigma(x)$ is closed, and we claim that

$$\lim_{\lambda \to \lambda_0} \frac{1}{\lambda - \lambda_0}[(x - \lambda)^{-1} - (x - \lambda_0)^{-1}] = (x - \lambda_0)^{-2} \tag{1.10}$$

in the norm topology of A. Indeed, we can write

$$\begin{aligned}(x - \lambda)^{-1} - (x - \lambda_0)^{-1} &= (x - \lambda)^{-1}[(x - \lambda_0) - (x - \lambda)](x - \lambda_0)^{-1} \\ &= (\lambda - \lambda_0)(x - \lambda)^{-1}(x - \lambda_0)^{-1}.\end{aligned}$$

Divide by $\lambda - \lambda_0$, and use the fact that $(x - \lambda)^{-1} \to (x - \lambda_0)^{-1}$ as $\lambda \to \lambda_0$ to obtain (1.10).

Contrapositively, assume that $\sigma(x)$ is empty, and choose an arbitrary bounded linear functional ρ on A. The scalar-valued function

$$f(\lambda) = \rho((x - \lambda)^{-1})$$

is defined everywhere in $\mathbb{C}$, and it is clear from (1.10) that f has a complex derivative everywhere satisfying $f'(\lambda) = \rho((x - \lambda)^{-2})$. Thus f is an entire function.

Notice that f is bounded. To see this we need to estimate $\|(x - \lambda)^{-1}\|$ for large λ. Indeed, if $|\lambda| > \|x\|$, then

$$\|(x - \lambda)^{-1}\| = \frac{1}{|\lambda|}\|(\mathbf{1} - \lambda^{-1}x)^{-1}\|.$$

The estimates of Theorem 1.5.2 therefore imply that

$$\|(x - \lambda)^{-1}\| \leq \frac{1}{|\lambda|(1 - \|x\|/|\lambda|)} = \frac{1}{|\lambda| - \|x\|},$$

and the right side clearly tends to zero as $|\lambda| \to \infty$. Thus the function $\lambda \mapsto \|(x - \lambda)^{-1}\|$ vanishes at infinity. It follows that f is a bounded entire function, which, by Liouville's theorem, must be constant. The constant value is 0 because f vanishes at infinity.

We conclude that $\rho((x - \lambda)^{-1}) = 0$ for every $\lambda \in \mathbb{C}$ and every bounded linear functional ρ. The Hahn–Banach theorem implies that $(x - \lambda)^{-1} = 0$ for every $\lambda \in \mathbb{C}$. But this is absurd because $(x - \lambda)^{-1}$ is invertible (and $\mathbf{1} \neq 0$ in A). □

The following application illustrates the power of this result.

Definition 1.6.4. A division algebra (over $\mathbb{C}$) is a complex associative algebra A with unit $\mathbf{1}$ such that every nonzero element in A is invertible.

Definition 1.6.5. An isomorphism of Banach algebras A and B is an isomorphism $\theta : A \to B$ of the underlying algebraic structures that is also a topological isomorphism; thus there are positive constants a, b such that

$$a\|x\| \leq \|\theta(x)\| \leq b\|x\|$$

for every element $x \in A$.

Corollary 1. Any Banach division algebra is isomorphic to the one-dimensional algebra $\mathbb{C}$.

Proof. Define $\theta : \mathbb{C} \to A$ by $\theta(\lambda) = \lambda\mathbf{1}$. θ is clearly an isomorphism of $\mathbb{C}$ onto the Banach subalgebra $\mathbb{C}\mathbf{1}$ of A consisting of all scalar multiples of the identity, and it suffices to show that θ is onto A. But for any element $x \in A$ Gelfand's theorem implies that there is a complex number $\lambda \in \sigma(x)$. Thus $x - \lambda$ is not invertible. Since A is a division algebra, $x - \lambda$ must be 0, hence $x = \theta(\lambda)$, as asserted. □

There are many division algebras in mathematics, especially commutative ones. For example, there is the algebra of all rational functions $r(z) = p(z)/q(z)$ of one complex variable, where p and q are polynomials with $q \neq 0$, or the algebra of all formal Laurent series of the form $\sum_{-\infty}^{\infty} a_n z^n$, where (a_n) is a doubly infinite sequence of complex numbers with $a_n = 0$ for sufficiently large negative n. It is significant that examples such as these cannot be endowed with a norm that makes them into a Banach algebra.

Exercises.

(1) Give an example of a one-dimensional Banach algebra that is not isomorphic to the algebra of complex numbers.

(2) Let X be a compact Hausdorff space and let $A = C(X)$ be the Banach algebra of all complex-valued continuous functions on X. Show that for every $f \in C(X)$, $\sigma(f) = f(X)$.

(3) Let T be the operator defined on $L^2[0,1]$ by $Tf(x) = xf(x)$, $x \in [0,1]$. What is the spectrum of T? Does T have point spectrum?

For the remaining exercises, let $(a_n : n = 1, 2, \dots)$ be a bounded sequence of complex numbers and let H be a complex Hilbert space having an orthonormal basis $e_1, e_2, \dots$.

(4) Show that there is a (necessarily unique) bounded operator $A \in \mathcal{B}(H)$ satisfying $Ae_n = a_n e_{n+1}$ for every $n = 1, 2, \dots$. Such an operator A is called a *unilateral weighted shift* (with weight sequence (a_n)).

A *unitary operator* on a Hilbert space H is an invertible isometry $U \in \mathcal{B}(H)$.

(5) Let $A \in \mathcal{B}(H)$ be a weighted shift as above. Show that for every complex number λ with $|\lambda| = 1$ there is a unitary operator $U = U_\lambda \in \mathcal{B}(H)$ such that $UAU^{-1} = \lambda A$.

(6) Deduce that the spectrum of a weighted shift must be the union of (possibly degenerate) concentric circles about $z = 0$.

(7) Let A be the weighted shift associated with a sequence $(a_n) \in \ell^\infty$.
   (a) Calculate $\|A\|$ in terms of (a_n).
   (b) Assuming that $a_n \to 0$ as $n \to \infty$, show that
$$\lim_{n\to\infty} \|A^n\|^{1/n} = 0.$$

1.7. Spectral Radius

Throughout this section, A denotes a unital Banach algebra with $\|\mathbf{1}\| = 1$. We introduce the concept of spectral radius and prove a useful asymptotic formula due to Gelfand, Mazur, and Beurling.

Definition 1.7.1. For every $x \in A$ the spectral radius of x is defined by
$$r(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}.$$

REMARK 1.7.2. Since the spectrum of x is contained in the central disk of radius $\|x\|$, it follows that $r(x) \leq \|x\|$. Notice too that for every $\lambda \in \mathbb{C}$ we have $r(\lambda x) = |\lambda| r(x)$.

We require the following rudimentary form of the spectral mapping theorem. If x is an element of A and f is a polynomial, then

$$f(\sigma(x)) \subseteq \sigma(f(x)). \tag{1.11}$$

To see why this is so, fix $\lambda \in \sigma(x))$. Since $z \mapsto f(z) - f(\lambda)$ is a polynomial having a zero at $z = \lambda$, there a polynomial g such that

$$f(z) - f(\lambda) = (z - \lambda)g(z).$$

Thus

$$f(x) - f(\lambda)\mathbf{1} = (x - \lambda)g(x) = g(x)(x - \lambda)$$

cannot be invertible: A right (respectively left) inverse of $f(x) - f(\lambda)\mathbf{1}$ gives rise to a right (respectively left) inverse of $x - \lambda$. Hence $f(\lambda) \in \sigma(f(x))$.

As a final observation, we note that for every $x \in A$ one has

$$r(x) \leq \inf_{n \geq 1} \|x^n\|^{1/n}. \tag{1.12}$$

Indeed, for every $\lambda \in \sigma(x)$ (1.11) implies that $\lambda^n \in \sigma(x^n)$; hence

$$|\lambda|^n = |\lambda^n| \leq r(x^n) \leq \|x^n\|,$$

and (1.12) follows after one takes nth roots.

The following formula is normally attributed to Gelfand and Mazur, although special cases were discovered independently by Beurling.

THEOREM 1.7.3. *For every $x \in A$ we have*

$$\lim_{n \to \infty} \|x^n\|^{1/n} = r(x).$$

The assertion here is that the limit exists in general, and has $r(x)$ as its value.

PROOF. From (1.12) we have $r(x) \leq \liminf_n \|x^n\|^{1/n}$, so it suffices to prove that

$$\limsup_{n \to \infty} \|x^n\|^{1/n} \leq r(x). \tag{1.13}$$

We need only consider the case $x \neq 0$. To prove (1.13) choose $\lambda \in \mathbb{C}$ satisfying $|\lambda| < 1/r(x)$ (when $r(x) = 0$, λ may be chosen arbitrarily). We claim that the sequence $\{(\lambda x)^n : n = 1, 2, \dots\}$ is bounded.

Indeed, by the Banach–Steinhaus theorem it suffices to show that for every bounded linear functional ρ on A we have

$$|\rho(x^n)\lambda^n| = |\rho((\lambda x)^n)| \leq M_\rho < \infty, \qquad n = 1, 2, \dots,$$

where M_ρ perhaps depends on ρ. To that end, consider the complex-valued function f defined on the (perhaps infinite) disk $\{z \in \mathbb{C} : |z| < 1/r(x)\}$ by

$$f(z) = \rho\left((\mathbf{1} - zx)^{-1}\right).$$

Note first that f is analytic. Indeed, for $|z| < 1/\|x\|$ we may expand $(\mathbf{1} - zx)^{-1}$ into a convergent series $\mathbf{1} + zx + (zx)^2 + \cdots$ to obtain a power series representation for f:

$$f(z) = \sum_{n=0}^{\infty} \rho(x^n) z^n. \tag{1.14}$$

On the other hand, in the larger region $R = \{z : 0 < |z| < 1/r(x)\}$ we can write

$$f(z) = \frac{1}{z}\rho\left((z^{-1}\mathbf{1} - x)^{-1}\right),$$

and from formula (1.10) it is clear that f is analytic on R. Taken with (1.14), this implies that f is analytic on the disk $\{z : |z| < 1/r(x)\}$.

On the smaller disk $\{z : |z| < 1/\|x\|\}$, (1.14) gives a power series representation for f; but since f is analytic on the larger disk $\{z : |z| < 1/r(x)\}$, it follows that the same series (1.14) must converge to $f(z)$ for all $|z| < 1/r(x)$. Thus we are free to take $z = \lambda$ in (1.14), and the resulting series converges. It follows that $\rho(x^n)\lambda^n$ is a bounded sequence, proving the claim.

Now choose any complex number λ satisfying $0 < |\lambda| < 1/r(x)$. By the claim, there is a constant $M = M_\lambda$ such that $|\lambda|^n \|x\|^n = \|\lambda x\|^n \leq M$ for every $n = 1, 2, \ldots$. after taking nth roots, we find that

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leq \limsup_{n\to\infty} \frac{M^{1/n}}{|\lambda|} = \frac{1}{|\lambda|}.$$

By allowing $|\lambda|$ to increase to $1/r(x)$ we obtain (1.13). □

Definition 1.7.4. An element x of a Banach algebra A (with or without unit) is called *quasinilpotent* if

$$\lim_{n\to\infty} \|x^n\|^{1/n} = 0.$$

Significantly, quasinilpotence is characterized quite simply in spectral terms.

Corollary 1. An element x of a unital Banach algebra A is quasinilpotent iff $\sigma(x) = \{0\}$.

Proof. x is quasinilpotent $\iff r(x) = 0 \iff \sigma(x) = \{0\}$. □

Exercises.

(1) Let $a_1, a_2, \ldots$ be a sequence of complex numbers such that $a_n \to 0$ as $n \to \infty$. Show that the associated weighted shift operator on ℓ^2 (see the Exercises of Section 1.6) has spectrum $\{0\}$.

(2) Consider the simplex $\Delta_n \subset [0,1]^n$ defined by

$$\Delta_n = \{(x_1, \ldots, x_n) \in [0,1]^n : x_1 \leq x_2 \leq \cdots \leq x_n\}.$$

Show that the volume of Δ_n is $1/n!$. Give a decent proof here: For example, you might consider the natural action of the permutation group S_n on the cube $[0,1]^n$ and think about how permutations act on Δ_n.

(3) Let $k(x, y)$ be a Volterra kernel as in Example 1.1.4, and let K be its corresponding integral operator on the Banach space $C[0, 1]$. Estimate the norms $\|K^n\|$ by showing that there is a positive constant M such that for every $f \in C[0, 1]$ and every $n = 1, 2, \dots,$

$$\|K^n f\| \leq \frac{M^n}{n!}\|f\|.$$

(4) Let K be a Volterra operator as in the preceding exercise. Show that for every complex number $\lambda \neq 0$ and every $g \in C[0, 1]$, the Volterra equation of the second kind $Kf - \lambda f = g$ has a unique solution $f \in C[0, 1]$.

1.8. Ideals and Quotients

The purpose of this section is to collect some basic information about ideals in Banach algebras and their quotient algebras. We begin with a complex algebra A.

DEFINITION 1.8.1. An *ideal* in A is linear subspace $I \subseteq A$ that is invariant under both left and right multiplication, $AI + IA \subseteq I$.

There are two trivial ideals, namely $I = \{0\}$ and $I = A$, and A is called *simple* if these are the only ideals. An ideal is *proper* if it is not all of A.

Suppose now that I is a proper ideal of A. Forming the quotient vector space A/I, we have a natural linear map $x \in A \mapsto \dot{x} = x + I \in A/I$ of A *onto* A/I. Since I is a two-sided ideal, one can unambiguously define a multiplication in A/I by

$$(x + I) \cdot (y + I) = xy + I, \qquad x, y \in A.$$

This multiplication makes A/I into a complex algebra, and the natural map $x \mapsto \dot{x}$ becomes a surjective homomorphism of complex algebras having the given ideal I as its kernel.

This information is conveniently summarized in the short exact sequence of complex algebras

$$0 \longrightarrow I \longrightarrow A \longrightarrow A/I \longrightarrow 0, \tag{1.15}$$

the map of I to A being the inclusion map, and the map of A onto A/I being $x \mapsto \dot{x}$. A basic philosophical principle of mathematics is to determine what information about A can be extracted from corresponding information about both the ideal I and its quotient A/I. For example, suppose that A is finite-dimensional as a vector space over $\mathbb{C}$. Then both I and A/I are finite-dimensional vector spaces, and from the observation that (1.15) is an exact sequence of vector spaces and linear maps one finds that the dimension of A is determined by the dimensions of the ideal and its quotient by way of $\dim A = \dim I + \dim A/I$ (see Exercise (1) below). The methods of homological algebra provide refinements of this observation that allow the computation of more subtle invariants of algebras (such as K-theoretic invariants), which have appropriate generalizations to the category of Banach algebras.

PROPOSITION 1.8.2. *Let A be a Banach algebra with normalized unit $\mathbf{1}$ and let I be a proper ideal in A. Then for every $z \in I$ we have $\|\mathbf{1} + z\| \geq 1$. In particular, the closure of a proper ideal is a proper ideal.*

PROOF. If there is an element $z \in I$ with $\|\mathbf{1} + z\| < 1$, then by Theorem 1.5.2 z must be invertible in A; hence $\mathbf{1} = z^{-1}z \in I$, which implies that I cannot be a proper ideal. The second assertion follows from the continuity of the norm; if $\|\mathbf{1} + z\| \geq 1$ for all $z \in I$, then $\|\mathbf{1} + z\| \geq 1$ persists for all z in the closure of I. □

REMARK 1.8.3. If I is a proper closed ideal in a Banach algebra A with normalized unit $\mathbf{1}$, then the unit of A/I satisfies

$$\|\dot{\mathbf{1}}\| = \inf_{z \in I} \|\mathbf{1} + z\| = 1;$$

hence the unit of A/I is also normalized. More significantly, it follows that a unital Banach algebra A with normalized unit is simple iff it is *topologically simple* (i.e., A has no nontrivial *closed* ideals; see the corollary of Theorem 1.8.5 below). That assertion is false for nonunital Banach algebras. For example, in the Banach algebra $\mathcal{K}$ of all compact operators on the Hilbert space ℓ^2, the set of finite-rank operators is a proper ideal that is dense in $\mathcal{K}$. Indeed, $\mathcal{K}$ contains many proper ideals, such as the ideal $\mathcal{L}^2$ of Hilbert–Schmidt operators that we will encounter later on. Nevertheless, $\mathcal{K}$ is topologically simple (for example, see [**2**], Corollary 1 of Theorem 1.4.2).

More generally, let I be a closed ideal in an arbitrary Banach algebra A (with or without unit). Then A/I is a Banach space; it is also a complex algebra relative to the multiplication defined above, and in fact it is a Banach algebra since for any $x, y \in A$,

$$\begin{aligned}\|\dot{x}\dot{y}\| &= \inf_{z \in I} \|xy + z\| \leq \inf_{z_1, z_2 \in I} \|xy + \underbrace{xz_2 + z_1 y + z_1 z_2}_{\in I}\| \\ &= \inf_{z_1, z_2 \in I} \|(x + z_1)(x + z_2)\| \leq \|\dot{x}\|\|\dot{y}\|.\end{aligned}$$

Notice, too, that (1.15) becomes an exact sequence of Banach algebras and continuous homomorphisms. If $\pi : A \to A/I$ denotes the natural surjective homomorphism, then we obviously have $\|\pi\| \leq 1$ in general, and $\|\pi\| = 1$ when A is unital with normalized unit.

The sequence (1.15) gives rise to a natural factorization of homomorphisms as follows. Let A, B be Banach algebras and let $\omega : A \to B$ be a homomorphism of Banach algebras (a bounded homomorphism of the underlying algebraic structures). Then $\ker \omega$ is a closed ideal in A, and there is a unique homomorphism $\dot{\omega} : A/\ker \omega \to B$ such that for all $x \in A$ we have $\omega(x) = \dot{\omega}(x + \ker \omega)$. The properties of this promotion of ω to $\dot{\omega}$ are summarized as follows:

PROPOSITION 1.8.4. *Every bounded homomorphism of Banach algebras* $\omega : A \to B$ *has a unique factorization* $\omega = \dot{\omega} \circ \pi$, *where* $\dot{\omega}$ *is an injective homomorphism of* $A/\ker\omega$ *to* B *and* $\pi : A \to A/\ker\omega$ *is the natural projection. One has* $\|\dot{\omega}\| = \|\omega\|$.

PROOF. The assertions in the first sentence are straightforward, and we prove $\|\dot{\omega}\| = \|\omega\|$. From the factorization $\omega = \dot{\omega}\circ\pi$ and the fact that $\|\pi\| \leq 1$ we have $\|\omega\| \leq \|\dot{\omega}\|$; the opposite inequality follows from

$$\|\dot{\omega}(\dot{x})\| = \|\omega(x)\| = \|\omega(x+z)\| \leq \|\omega\|\|x+z\|, \qquad z \in \ker\omega,$$

after the infimum is taken over all $z \in \ker\omega$. □

Before introducing maximal ideals, we review some basic principles of set theory. A *partially ordered set* is a pair $(S, \leq)$ consisting of a set S and a binary relation $\leq$ that is transitive ($x \leq y$, $y \leq z \implies x \leq z$) and satisfies $x \leq y \leq x \implies x = y$. An element $x \in S$ is said to be *maximal* if there is no element $y \in S$ satisfying $x \leq y$ and $y \neq x$. A linearly ordered subset of S is a subset $L \subseteq S$ for which any two elements $x, y \in L$ are related by either $x \leq y$ or $y \leq x$. The set $\mathcal{L}$ of all linearly ordered subsets of S is itself partially ordered by set inclusion.

The *Hausdorff maximality principle* makes the assertion that every partially ordered set has a maximal linearly ordered subset; that is, the partially ordered set $\mathcal{L}$ has a maximal element. *Zorn's lemma* makes the assertion that every partially ordered set S that is *inductive*, in the sense that every linearly ordered subset of S has an upper bound in S, must contain a maximal element. While the maximality principle appears to be rather different from Zorn's lemma, they are actually equivalent in any model of set theory that is appropriate for functional analysis. Indeed, both Zorn's lemma and the maximality principle are equivalent to the axiom of choice. Our experience has been that most proofs in functional analysis that require the axiom of choice, or some reformulation of it in terms of transfinite induction, usually run more smoothly (and are simpler) when they are formulated so as to make use of Zorn's lemma. That will be the way such things are handled throughout this book.

An ideal M in a complex algebra A is said to be a *maximal ideal* if it is a maximal element in the partially ordered set of all proper ideals of A. Thus a maximal ideal is a *proper* ideal $M \subseteq A$ with the property that for any ideal $N \subseteq A$,

$$M \subseteq N \implies N = M \quad \text{or} \quad N = A.$$

Maximal ideals are particularly useful objects when one is working with unital Banach algebras.

THEOREM 1.8.5. *Let A be a unital Banach algebra. Then every maximal ideal of A is closed, and every proper ideal of A is contained in some maximal ideal.*

PROOF. For the first assertion, let M be a maximal ideal of A. Remark 1.8.3 implies that the unit $\mathbf{1}$ cannot belong to the closure $\overline{M}$ of M; hence $\overline{M}$ is a proper ideal of A. Since $M \subseteq \overline{M}$, maximality of M implies that $M = \overline{M}$ is closed.

Suppose now that I is some proper ideal of A, and consider the set $\mathcal{P}$ of all proper ideals of A that contain I. The family of sets $\mathcal{P}$ is partially ordered in the natural way by set inclusion, and we claim that it is inductive in the sense that every linearly ordered subset $\mathcal{L} = \{J_\alpha : \alpha \in S\}$ of $\mathcal{P}$ has an upper bound in $\mathcal{P}$. Indeed, the union $\cup_\alpha J_\alpha$ is an ideal in A because it is the union of a linearly ordered family of ideals. It cannot contain the unit $\mathbf{1}$ of A because $\mathbf{1} \notin J_\alpha$ for every $\alpha \in S$. Hence $\cup_\alpha J_\alpha$ is an element of $\mathcal{P}$ as well as an upper bound for $\mathcal{L}$.

Zorn's lemma implies that $\mathcal{P}$ has a maximal element M, and M is a proper ideal that contains I. It is a maximal ideal because if N is any ideal containing M, then N must contain I and hence $N \in \mathcal{P}$. Since M is a maximal element of $\mathcal{P}$, we conclude that $M = N$. $\square$

COROLLARY 1. *A unital Banach algebra is simple iff it is topologically simple.*

Exercises.

(1) *Review of linear algebra.* Let V and W be finite-dimensional vector spaces over $\mathbb{C}$ and let $T : V \to W$ be a linear map satisfying $TV = W$, and having kernel $K = \{x \in V : Tx = 0\}$. Then we have a short exact sequence of vector spaces
$$0 \longrightarrow K \longrightarrow V \longrightarrow W \longrightarrow 0.$$
Show that $\dim V = \dim K + \dim W$. Your proof should proceed from the definition of the dimension of a finite-dimensional vector space as the cardinality of any basis for it.

(2) *More linear algebra.* For $n = 1, 2, \dots$, let $V_1, V_2, \dots, V_n$ be finite-dimensional vector spaces and set $V_0 = V_{n+1} = 0$ (the trivial vector space). Suppose that for each $k = 0, 1, \dots, n$ we have a linear map of V_k to V_{k+1} such that the associated sequence of vector spaces
$$0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow \cdots \longrightarrow V_n \longrightarrow 0$$
is exact. Show that $\sum_{k=1}^{n} (-1)^k \dim V_k = 0$.

(3) Show that every normed linear space E has a basis $\mathcal{B} \subseteq E$ consisting of *unit* vectors, and deduce that every infinite-dimensional normed linear space has a discontinuous linear functional $f : E \to \mathbb{C}$. Recall that a basis for a vector space V is a set of vectors $\mathcal{B}$ with the following two properties: every finite subset of $\mathcal{B}$ is linearly independent, and every vector in V is a finite linear combination of elements of $\mathcal{B}$.

(4) Let A be a complex algebra and let I be a proper ideal of A. Show that I is a maximal ideal iff the quotient algebra A/I is simple.

(5) Let A be a unital Banach algebra, let n be a positive integer, and let $\omega : A \to M_n$ be a homomorphism of complex algebras such that $\omega(A) = M_n$, M_n denoting the algebra of all $n \times n$ matrices over $\mathbb{C}$. Show that ω is continuous (where M_n is topologized in the natural way by $\mathbb{C}^{n^2}$). Deduce that every linear functional $f : A \to \mathbb{C}$ satisfying $f(xy) = f(x)f(y)$, $x, y \in A$, is continuous.

1.9. Commutative Banach Algebras

We now work out Gelfand's generalization of the Fourier transform. Let A be a commutative Banach algebra with unit $\mathbf{1}$ satisfying $\|\mathbf{1}\| = 1$. We consider the set $\hom(A, \mathbb{C})$ of all homomorphisms $\omega : A \to \mathbb{C}$. An element $\omega \in \hom(A, \mathbb{C})$ is a complex linear functional satisfying $\omega(xy) = \omega(x)\omega(y)$ for all $x, y \in A$; notice that we do not assume that ω is continuous, but as we will see momentarily, that will be the case. The *Gelfand spectrum* of A is defined as the set

$$\operatorname{sp}(A) = \{\omega \in \hom(A, \mathbb{C}) : \omega \neq 0\}$$

of all nontrivial complex homomorphisms of A. It is also called the *maximal ideal space* of A, since there is a natural bijection of $\operatorname{sp}(A)$ onto the set of all maximal ideals of A (see Exercise (2) below).

REMARK 1.9.1. Every element $\omega \in \operatorname{sp}(A)$ satisfies $\omega(\mathbf{1}) = 1$. Indeed, for fixed ω the complex number $\lambda = \omega(\mathbf{1})$ satisfies $\lambda\omega(x) = \omega(\mathbf{1} \cdot x) = \omega(x)$ for every $x \in A$. Since the set of complex numbers $\omega(A)$ must contain something other than 0, it follows that $\lambda = 1$.

REMARK 1.9.2. Every element $\omega \in \operatorname{sp}(A)$ is continuous. This is an immediate consequence of the case $n = 1$ of Exercise (5) of the preceding section, but perhaps it is better to supply more detail. Indeed, we claim that $\|\omega\| = 1$. For the proof, note that $\ker\omega$ is an ideal in A with the property that the quotient algebra $A/\ker\omega$ is isomorphic to the field of complex numbers. Hence $\ker\omega$ is a maximal ideal in A. By Theorem 1.8.5, it is closed. Because of the decomposition $\omega = \dot{\omega} \circ \pi$ where π is the natural homomorphism of A onto $A/\ker\omega$ and $\dot{\omega}$ is the linear map between the two one-dimensional Banach spaces $A/\ker\omega$ and $\mathbb{C}$ given by $\dot{\omega}(\lambda\dot{\mathbf{1}}) = \lambda\omega(\mathbf{1}) = \lambda$, we have $\|\dot{\omega}\| = 1$. Hence $\|\omega\| \leq \|\dot{\omega}\|\|\pi\| \leq 1$. The opposite inequality is clear from $\|\omega\| \geq |\omega(\mathbf{1})| = 1$.

With these observations in hand, one can introduce a topology on $\operatorname{sp}(A)$ as follows. We have seen that $\operatorname{sp}(A)$ is a subset of the unit ball of the dual A' of A, and by Alaoglu's theorem the latter is a compact Hausdorff space in its relative weak*-topology. Thus $\operatorname{sp}(A)$ inherits a natural Hausdorff topology as a subspace of a compact Hausdorff space.

PROPOSITION 1.9.3. *In its relative weak*-topology,* $\operatorname{sp}(A)$ *is a compact Hausdorff space.*

PROOF. It suffices to show that $\mathrm{sp}(A)$ is a weak*-closed subset of the unit ball of the dual of A. Notice that a linear functional $f : A \to \mathbb{C}$ belongs to $\mathrm{sp}(A)$ iff $\|f\| \leq 1$, $f(\mathbf{1}) = 1$, and $f(yz) = f(y)f(z)$ for all $y, z \in A$. These conditions obviously define a weak*-closed subset of the unit ball of A'. □

REMARK 1.9.4. *The Gelfand map.* Every element $x \in A$ gives rise to a function $\hat{x} : \mathrm{sp}(A) \to \mathbb{C}$ by way of $\hat{x}(\omega) = \omega(x)$, $\omega \in \mathrm{sp}(A)$; $\hat{x}$ is called the *Gelfand transform* of x, and $x \mapsto \hat{x}$ is called the *Gelfand map.* The functions $\hat{x}$ are continuous by definition of the weak*-topology on $\mathrm{sp}(A)$. For $x, y \in A$ we have

$$\hat{x}(\omega)\hat{y}(\omega) = \omega(x)\omega(y) = \omega(xy) = \widehat{xy}(\omega).$$

Moreover, since every element ω of $\mathrm{sp}(A)$ satisfies $\omega(\mathbf{1}) = 1$, it follows that $\hat{\mathbf{1}}$ is the constant function 1 in $C(\mathrm{sp}(A))$. It follows that *the Gelfand map is a homomorphism of A onto a unital subalgebra of $C(\mathrm{sp}(A))$ that separates points of* $\mathrm{sp}(A)$. The previous remarks also imply that $\|\hat{x}\|_\infty \leq \|x\|$, $x \in A$.

Most significantly, the Gelfand map exhibits spectral information about elements of A in an explicit way.

THEOREM 1.9.5. *Let A be a commutative Banach algebra with unit. For every element $x \in A$, we have*

$$\sigma(x) = \{\hat{x}(p) : p \in \mathrm{sp}(A)\}.$$

PROOF. Since for any $x \in A$ and $\lambda \in \mathbb{C}$, $\widehat{x - \lambda} = \hat{x} - \lambda$ and $\sigma(x - \lambda) = \sigma(x) - \lambda$, it suffices to establish the following assertion: An element $x \in A$ is invertible iff $\hat{x}$ never vanishes.

Indeed, if x is invertible, then there is a $y \in A$ such that $xy = \mathbf{1}$; hence $\hat{x}(\omega)\hat{y}(\omega) = \widehat{xy}(\omega) = 1$, $\omega \in \mathrm{sp}(A)$, so that $\hat{x}$ has no zeros.

Conversely, suppose that x is a noninvertible element of A. We must show that there is an element $\omega \in \mathrm{sp}(A)$ such that $\omega(x) = 0$. For that, consider the set $xA = \{xa : a \in A\} \subseteq A$. This set is an ideal that does not contain $\mathbf{1}$. By Theorem 1.8.5, xA is contained in some maximal ideal $M \subseteq A$, necessarily closed. We will show that there is an element $\omega \in \mathrm{sp}(A)$ such that $M = \ker \omega$. Indeed, A/M is a simple Banach algebra with unit; therefore it has no nontrivial ideals at all. Since A/M is also commutative, this implies that A/M is a field (for any nonzero element $\zeta \in A/M$, $\zeta \cdot A/M$ is a nonzero ideal, which must therefore contain the unit of A/M). By Corollary 1 of Theorem 1.6.3, A/M is isomorphic to $\mathbb{C}$. Choosing an isomorphism $\dot{\omega} : A/M \to \mathbb{C}$, we obtain a complex homomorphism $\omega : A \to \mathbb{C}$ by way of $\omega(x) = \dot{\omega}(x + M)$. It is clear that $\ker \omega = M$, and finally $\hat{x}$ vanishes at ω because $x \in xA \subseteq M$. □

Theorem 1.9.5 provides an effective procedure for computing the spectrum of elements of any unital commutative Banach algebra A. One first identifies the Gelfand spectrum $\mathrm{sp}(A)$ in concrete terms as a topological space and the Gelfand map of A into $C(\mathrm{sp}(A))$. Once these calculations have been carried out, the spectrum of an element $x \in A$ is exhibited as the range of values of $\hat{x}$. In the following section we discuss two important examples that illustrate the method.

Exercises. In the first four exercises, A denotes a commutative Banach algebra with unit.

(1) Show that if A is nontrivial in the sense that $A \neq \{0\}$ (equivalently, $\mathbf{1} \neq 0$), one has $\mathrm{sp}(A) \neq \emptyset$.
(2) Show that the mapping $\omega \in \mathrm{sp}(A) \to \ker \omega$ is a bijection of the Gelfand spectrum onto the set of all maximal ideals in A. For this reason, $\mathrm{sp}(A)$ is often called the *maximal ideal space* of A.
(3) Show that the Gelfand map is an isometry iff $\|x^2\| = \|x\|^2$ for every $x \in A$.
(4) The *radical* of A is defined as the set $\mathrm{rad}(A)$ of all quasinilpotent elements of A,
$$\mathrm{rad}(A) = \left\{ x \in A : \lim_{n \to \infty} \|x^n\|^{1/n} = 0 \right\}.$$
Show that $\mathrm{rad}(A)$ is a closed ideal in A with the property that $A/\mathrm{rad}(A)$ has no nonzero quasinilpotents (such a commutative Banach algebra is called *semisimple*).
(5) Let A and B be commutative unital Banach algebras and let $\theta : A \to B$ be a homomorphism of the complex algebra structures such that $\theta(\mathbf{1}_A) = \mathbf{1}_B$. Do not assume that θ is continuous.
   (a) Show that θ induces a continuous map $\hat{\theta} : \mathrm{sp}(B) \to \mathrm{sp}(A)$ by way of $\hat{\theta}(\omega) = \omega \circ \theta$.
   (b) Assuming that B is semisimple, show that θ is necessarily bounded. Hint: Use the closed graph theorem.
   (c) Deduce that every automorphism of a commutative unital semisimple Banach algebra is a topological automorphism.

1.10. Examples: $C(X)$ and the Wiener Algebra

We now look more closely at two important examples of commutative Banach algebras. Following the program described above, we calculate their maximal ideal spaces, their Gelfand maps, and describe an application of the method to prove a classical theorem of Wiener on absolutely convergent Fourier series.

Example 1.10.1. $C(X)$. The Gelfand spectrum of the Banach algebra $A = C(X)$ of all continuous functions on a compact Hausdorff space X can be identified with X in the following way. Every point $p \in X$ determines a complex homomorphism $\omega_p \in \mathrm{sp}(C(X))$ by evaluation:
$$\omega_p(f) = f(p), \qquad f \in C(X).$$
The map $p \mapsto \omega_p$ is obviously one-to-one, and it is continuous by definition of the weak*-topology on the dual space of $C(X)$. The work amounts to

showing that every $\omega \in \operatorname{sp}(C(X))$ arises in this way from some point of X. The method we use is based on a characterization of positive linear functionals on $C(X)$ in terms of an extremal property of their norm (Lemma 1.10.3). This is a useful technique for other purposes, and we will see it again in Chapter 4.

REMARK 1.10.2. Every compact convex set $K \subseteq \mathbb{C}$ is the intersection of all closed half-spaces that contain it. It is also true that K is the intersection of all closed disks that contain it. Equivalently, if $z_0 \in \mathbb{C}$ is any point not in the closed convex hull of K, then there is a disk $D = D_{a,R} = \{z \in \mathbb{C} : |z - a| \leq R\}$ such that $K \subseteq D$ and $z_0 \neq D$. The reader is encouraged to draw a picture illustrating this geometric fact.

LEMMA 1.10.3. *Let ρ be a linear functional on $C(X)$ satisfying $\|\rho\| = \rho(\mathbf{1}) = 1$. Then, for every $f \in C(X)$,*

$$\rho(f) \in \overline{\operatorname{conv}} f(X),$$

$\overline{\operatorname{conv}} f(X)$ denoting the closed convex hull of the range of f.

In particular, if f^* denotes the complex conjugate of $f \in C(X)$, then we have $\rho(f^*) = \overline{\rho(f)}$.

PROOF. Fix $f \in C(X)$. In view of Remark 1.10.2, to prove the first assertion it suffices to show that every disk $D = \{z \in \mathbb{C} : |z - a| \leq R\}$ that contains $f(X)$ must also contain $\rho(f)$; equivalently,

$$|f(p) - a| \leq R, \quad \forall p \in X \implies |\rho(f) - a| \leq R.$$

But if $|f(p) - a| \leq R$ for every p, then $\|f - a \cdot \mathbf{1}\| \leq R$. Since $\|\rho\| = \rho(\mathbf{1}) = 1$, this implies $|\rho(f) - a| = |\rho(f - a \cdot \mathbf{1})| \leq R$, as required.

For the second assertion, let $f = g + ih \in C(X)$ with g and h real-valued continuous functions. By the preceding paragraph, $\rho(g)$ and $\rho(h)$ are real numbers; hence $\rho(f^*) = \rho(g - ih) = \rho(g) - i\rho(h)$ is the complex conjugate of $\rho(f) = \rho(g) + i\rho(h)$. □

THEOREM 1.10.4. *The map $p \in X \mapsto \omega_p \in \operatorname{sp}(C(X))$ is a homeomorphism of X onto the Gelfand spectrum of $C(X)$. This map identifies X with $\operatorname{sp}(C(X))$ in such a way that the Gelfand map becomes the identity map of $C(X)$ to $C(X)$.*

In particular, the spectrum of $f \in C(X)$ is $f(X)$.

PROOF. In view of the preliminary remarks above, the proof reduces to showing that every complex homomorphism ω is associated with some point $p \in X$, $\omega = \omega_p$. Fixing ω, we have to show that

$$\bigcap_{f \in C(X)} \{p \in X : f(p) = \omega(f)\} \neq \emptyset.$$

The left side is an intersection of compact subsets of X; so if it is empty, then by the finite intersection property there is a finite set of functions

$f_1, \dots, f_n \in C(X)$ such that

$$\bigcap_{k=1}^{n} \{p \in X : f_k(p) = \omega(f)\} = \emptyset.$$

Define $g \in C(X)$ by

$$g(p) = \sum_{k=1}^{n} |f_k(p) - \omega(f_k)|^2, \qquad p \in X.$$

Then g is obviously nonnegative, and by the choice of f_k, it has no zeros on X. Hence there is an $\epsilon > 0$ such that $g(p) \geq \epsilon$, $p \in X$.

Since $\|\omega\| = \omega(\mathbf{1}) = 1$ and $g - \epsilon\mathbf{1} \geq 0$, Lemma 1.10.3 also implies that $\omega(g - \epsilon\mathbf{1}) \geq 0$; hence

$$\omega(g) \geq \epsilon \cdot \omega(\mathbf{1}) = \epsilon > 0.$$

On the other hand, Lemma 1.10.3 also implies that for each k,

$$\begin{aligned} \omega(|f_k - \omega(f_k)\mathbf{1}|^2) =& \omega((f_k - \omega(f_k)\mathbf{1})^*(f_k - \omega(f_k)\mathbf{1})) \\ =& |\omega(f_k - \omega(f_k)\mathbf{1})|^2 = |0|^2 = 0, \end{aligned}$$

and after summing on k we obtain $\omega(g) = 0$, contradicting the preceding inequality. □

EXAMPLE 1.10.5. *The Wiener algebra.* Consider the space $\mathcal{W}$ of all continuous functions on the unit circle whose Fourier series converges absolutely, that is, all functions $f : \mathbb{T} \to \mathbb{C}$ whose Fourier series have the form

$$f(e^{i\theta}) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}, \tag{1.16}$$

where $\sum_n |a_n| < \infty$. One may verify directly that $\mathcal{W}$ is a subalgebra of $C(\mathbb{T})$ (because $\ell^1(\mathbb{Z})$ is a linear space closed under convolution), which obviously contains the constant functions. The algebra of functions $\mathcal{W}$ is called the Wiener algebra.

In connection with his study of Tauberian theorems in the 1930s, Norbert Wiener carried out a deep analysis of the translation-invariant subspaces of the Banach spaces $\ell^1(\mathbb{Z})$ and $L^1(\mathbb{R})$; notice that since both $\mathbb{Z}$ and $\mathbb{R}$ are additive groups, they act naturally as groups of isometric translation operators on their respective L^1 spaces. For example, the kth translate of a sequence $(a_n)_{n\in\mathbb{Z}}$ in $\ell^1(\mathbb{Z})$ is the sequence $(a_{n-k})_{n\in\mathbb{Z}}$. Among other things, Wiener proved that the translates of a sequence $(a_n) \in \ell^1(\mathbb{Z})$ have all of $\ell^1(\mathbb{Z})$ as their closed linear span iff the function f defined in (1.16) never vanishes. He did this by establishing the following key property of the algebra $\mathcal{W}$.

THEOREM 1.10.6. *If $f \in \mathcal{W}$ and f has no zeros on $\mathbb{T}$, then the reciprocal $1/f$ belongs to $\mathcal{W}$.*

Wiener's original proof of Theorem 1.10.6 was a remarkable exercise in hard classical analysis. Subsequently, Gelfand gave an elegant conceptual proof using the elementary theory of Banach algebras, basing the critical step on Theorem 1.9.5. We now describe Gelfand's proof.

Consider the Banach algebra $A = \ell^1(\mathbb{Z})$, with multiplication defined by convolution $*$. The unit of A is the sequence $\mathbf{1} = (e_n)$, where $e_0 = 1$ and $e_n = 0$ for $n \neq 0$. We show first that $\operatorname{sp}(A)$ can be identified with the unit circle $\mathbb{T}$.

Indeed, for every $\lambda \in \mathbb{T}$ we can define a bounded linear functional ω_λ on A by

$$\omega_\lambda(a) = \sum_{n=-\infty}^{\infty} a_n \lambda^n, \qquad a = (a_n) \in \ell^1(\mathbb{Z}).$$

Obviously, $\omega_\lambda(\mathbf{1}) = 1$, and one verifies directly that $\omega_\lambda(a * b) = \omega_\lambda(a)\omega_\lambda(b)$. Hence $\omega_\lambda \in \operatorname{sp}(A)$.

We claim that every $\omega \in \operatorname{sp}(A)$ has the form ω_λ for a unique point $\lambda \in \mathbb{T}$. To see that, fix $\omega \in \operatorname{sp}(A)$ and define a complex number λ by $\lambda = \omega(\zeta)$, where $\zeta = (\zeta_n)$ is the sequence $\zeta_n = 1$ if $n = 1$, and $\zeta_n = 0$ otherwise. Then ζ has unit norm in A, and hence $|\lambda| = |\omega(\zeta)| \leq \|\zeta\| = 1$. Another direct computation shows that ζ is invertible in A, and its inverse is the sequence $\tilde{\zeta} = (\tilde{\zeta}_n)$, where $\tilde{\zeta}_n = 1$ for $n = -1$, and $\tilde{\zeta}_n = 0$ otherwise. Since $\|\tilde{\zeta}\| = 1$ and $|1/\lambda| = |1/\omega(\zeta)| = |\omega(\tilde{\zeta})| \leq \|\tilde{\zeta}\| = 1$, we find that $|\lambda| = 1$. Notice that $\omega = \omega_\lambda$. Indeed, we must have $\omega(\zeta^n) = \lambda^n = \omega_\lambda(\zeta^n)$ for every $n \in \mathbb{Z}$, ζ^n being the unit sequence with a single nonzero component in the nth position. Since the set $\{\zeta^n : n \in \mathbb{Z}\}$ obviously has $\ell^1(\mathbb{Z})$ as its closed linear span, it follows that $\omega = \omega_\lambda$. Then $\lambda = \omega(\zeta)$ is obviously uniquely determined by ω.

These remarks show that the map $\lambda \mapsto \omega_\lambda$ is a bijection of $\mathbb{T}$ on $\operatorname{sp}(A)$. The inverse of this map (given by $\omega \in \operatorname{sp}(A) \mapsto \omega(\zeta) \in \mathbb{T}$) is obviously continuous, so by compactness of $\operatorname{sp}(A)$ it must be a homeomorphism. Thus we have identified $\operatorname{sp}(A)$ with the unit circle $\mathbb{T}$ and the Gelfand map with the Fourier transform, which carries a sequence $a \in \ell^1(\mathbb{Z})$ to the function $\hat{a} \in C(\mathbb{T})$ given by

$$\hat{a}(e^{i\theta}) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

Having computed $\operatorname{sp}(A)$ and the Gelfand map in concrete terms, we observe that the range of the Gelfand map $\{\hat{a} : a \in A\}$ is exactly the Wiener algebra $\mathcal{W}$. The proof of Theorem 1.10.6 can now proceed as follows. Let f be a function in $\mathcal{W}$ having no zeros on $\mathbb{T}$ and let a be the element of $A = \ell^1(\mathbb{Z})$ having Gelfand transform f. By Theorem 1.9.5, there is an element $b \in A$ such that $a * b = \mathbf{1}$; hence $\hat{a}(\lambda)\hat{b}(\lambda) = 1$, $\lambda \in \mathbb{T}$. It follows that $1/f = \hat{b} \in \mathcal{W}$, as asserted.

Exercises. Let $\mathcal{B}$ be the space of all continuous functions f defined on the closed unit disk $\Delta = \{z \in \mathbb{C} : |z| \leq 1\}$, which can be represented there by a convergent power series of the form

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \in \Delta,$$

for some sequence $a_0, a_1, a_2, \ldots$ in $\mathbb{C}$ satisfying $\sum_n |a_n| < \infty$.

(1) Prove the following analogue of Wiener's theorem, Theorem 1.10.6. If $f \in \mathcal{B}$ satisfies $f(z) \neq 0$ for every $z \in \Delta$, then $g = 1/f$ belongs to $\mathcal{B}$.

In the following exercise, $\mathbb{Z}_+$ denotes the additive semigroup of all nonnegative integers.

(2) Let T be the isometric shift operator that acts on $\ell^1(\mathbb{Z}_+)$ by

$$T(x_0, x_1, x_2, \ldots) = (0, x_0, x_1, x_2, \ldots),$$

and let $a = (a_0, a_1, a_2, \ldots) \in \ell^1(\mathbb{Z}_+)$. Show that the set of translates $\{a, Ta, T^2a, \ldots\}$ spans $\ell^1(\mathbb{Z}_+)$ if and only if the power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad |z| \leq 1,$$

has no zeros in the closed unit disk. Hint: Use the previous exercise.

1.11. Spectral Permanence Theorem

Let A be a Banach algebra with (normalized) unit; A is not necessarily commutative. Suppose we also have a Banach subalgebra $B \subseteq A$ of A that contains the unit of A. Then for every element $x \in B$ it makes sense to speak of the spectrum $\sigma_B(x)$ of x relative to B as well as the spectrum $\sigma_A(x)$ of x relative to A. There can be significant differences between these two versions of the spectrum of x, and we now discuss this phenomenon.

Proposition 1.11.1. *Let B be a Banach subalgebra of A that contains the unit of A. For every element $x \in B$ we have $\sigma_A(x) \subseteq \sigma_B(x)$.*

Proof. This is an immediate consequence of the fact that invertible elements of B are invertible elements of A. □

Example 1.11.2. Consider the Banach algebra $A = C(\mathbb{T})$ of continuous functions on the unit circle, and let B be the Banach subalgebra generated by the current variable $\zeta(z) = z$, $z \in \mathbb{T}$. Thus B is the closure (in the sup norm of $\mathbb{T}$) of the algebra of polynomials

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n.$$

Let us compute the two spectra $\sigma_A(\zeta)$ and $\sigma_B(\zeta)$. The discussion of $C(X)$ in the previous section implies that

$$\sigma_A(\zeta) = \zeta(\mathbb{T}) = \mathbb{T}.$$

We now show that $\sigma_B(\zeta)$ is the closed unit disk $\Delta \subseteq \mathbb{C}$. Indeed, the general principles we have developed for computing spectra in commutative

Banach algebras imply that, in order to compute $\sigma_B(\zeta)$, we should first compute the Gelfand spectrum $\mathrm{sp}(B)$. We will identify $\mathrm{sp}(B)$ with Δ. Indeed, for every $z \in \Delta$ the maximum modulus principle implies that

$$|p(z)| \le \sup_{|\lambda|=1} |p(\lambda)|. \tag{1.17}$$

It follows that the linear functional ω_z on B defined on polynomials by $\omega_z(p) = p(z)$ satisfies $\|\omega_z\| \le 1$, and hence extends uniquely to a linear functional on B, which we denote by the same letter ω_z. Obviously, ω_z belongs to $\mathrm{sp}(B)$. The map $z \in \Delta \mapsto \omega_z \in \mathrm{sp}(B)$ is continuous and one-to-one. It is onto because for every $\omega \in \mathrm{sp}(B)$, the complex number $z = \omega(\zeta)$ satisfies $|z| = |\omega(\zeta)| \le \|\zeta\| = 1$, and it has the property that that $\omega(p) = p(z) = \omega_z(p)$ for every polynomial p. Hence $\omega = \omega_z$ on B.

Having identified $\mathrm{sp}(B)$ with Δ and observing that $\hat{\zeta}$ is identified with the current variable $\hat{\zeta}(z) = z$, $z \in \Delta$, we can appeal to Theorem 1.9.5 to conclude that $\sigma_B(\zeta) = \Delta$.

The following result is sometimes called the spectral permanence theorem, since it implies that points in the boundary of $\sigma_B(x)$ cannot be removed by replacing B with a larger algebra.

THEOREM 1.11.3. *Let B be a Banach subalgebra of a unital Banach algebra A which contains the unit of A. Then for every $x \in B$ we have*

$$\partial\sigma_B(x) \subseteq \sigma_A(x).$$

PROOF. It suffices to show that $0 \in \partial\sigma_B(x) \implies 0 \in \sigma_A(x)$. Contrapositively, assume that $0 \ne \sigma_A(x)$ and $0 \in \partial\sigma_B(x)$. Then x is invertible in A and there is a sequence of complex numbers $\lambda_n \to 0$ such that $\lambda_n \notin \sigma_B(x)$. Thus $(x - \lambda_n)^{-1}$ is a sequence of elements of B with the property that, since inversion is continuous in A^{-1}, converges to x^{-1} as $n \to \infty$. It follows that $x^{-1} = \lim_n (x - \lambda_n)^{-1} \in \bar{B} = B$, contradicting the fact that $0 \in \sigma_B(x)$. $\square$

One can reformulate the preceding result into a more precise description of the relation between $\sigma_B(x)$ and $\sigma_A(x)$ as follows. Given a compact set K of complex numbers, a *hole* of K is defined as a *bounded* component of its complement $\mathbb{C} \setminus K$. Let us decompose $\mathbb{C} \setminus \sigma_A(x)$ into its connected components, obtaining an unbounded component Ω_∞ together with a sequence of holes $\Omega_1, \Omega_2, \dots$,

$$\mathbb{C} \setminus \sigma_A(x) = \Omega_\infty \sqcup \Omega_1 \sqcup \Omega_2 \sqcup \cdots .$$

Of course, there may be only a finite number of holes or none at all.

We require an elementary topological fact:

LEMMA 1.11.4. *Let Ω be a connected topological space, and let X be a closed subset of Ω such that $\emptyset \ne X \ne \Omega$. Then $\partial X \ne \emptyset$.*

PROOF. If $\partial X = \emptyset$, then $\Omega = \operatorname{int}(X) \sqcup (\Omega \setminus X)$ is a decomposition of Ω into disjoint open sets; hence either $\operatorname{int}(X) = \emptyset$ or $X = \Omega$, and hence $\operatorname{int}(X) = \emptyset$. But this implies that $X = \operatorname{int}(X) \cup \partial X = \emptyset$, a contradiction. $\square$

COROLLARY 1. Let $\mathbf{1}_A \in B \subseteq A$ be as above, let $x \in A$, and let Ω be a bounded component of $\mathbf{C} \setminus \sigma_A(x)$. Then either $\Omega \cap \sigma_B(x) = \emptyset$ or $\Omega \subseteq \sigma_B(x)$.

PROOF. Let Ω be a hole of $\sigma_A(x)$. Consider $X = \Omega \cap \sigma_B(x)$ as a closed subspace of the topological space Ω. Since Ω is an open set in $\mathbb{C}$, the boundary $\partial_\Omega X$ of X relative to Ω is contained in

$$\partial \sigma_B(x) \subseteq \sigma_A(x) \subseteq \mathbb{C} \setminus \Omega.$$

Hence $\partial_\Omega X = \emptyset$. Lemma 1.11.4 implies that either $X = \emptyset$ or $X = \Omega$, as asserted. $\square$

We deduce the following description of $\sigma_B(x)$ in terms of $\sigma_A(x)$.

COROLLARY 2. Let $x \in B \subseteq A$ be as in the previous theorem. Then $\sigma_B(x)$ is obtained from $\sigma_A(x)$ by adjoining to it some (and perhaps none) of its holes.

For example, if $\sigma_A(x)$ is the unit circle, then the only possibilities for $\sigma_B(x)$ are the unit circle and the closed unit disk.

Exercises.

(1) Let A be a unital Banach algebra, let $x \in A$, and let Ω_∞ be the unbounded component of $\mathbb{C} \setminus \sigma_A(x)$. Show that for every $\lambda \in \Omega_\infty$ there is a sequence of polynomials $p_1, p_2, \dots$ such that

$$\lim_{n \to \infty} \|(x - \lambda \mathbf{1})^{-1} - p_n(x)\| = 0.$$

(2) Let A be a unital Banach algebra that is generated by $\{\mathbf{1}, x\}$ for some $x \in A$. Show that $\sigma_A(x)$ has no holes.

(3) Deduce the following theorem of Runge. Let $X \subseteq \mathbb{C}$ be a compact set whose complement is connected. Show that if $f(z) = p(z)/q(z)$ is a rational function (p, q being polynomials) with $q(z) \neq 0$ for every $z \in X$, then there is a sequence of polynomials $f_1, f_2, \dots$ such that

$$\sup_{z \in X} |f(z) - f_n(z)| \to 0, \quad \text{as } n \to \infty.$$

1.12. Brief on the Analytic Functional Calculus

The analytic functional calculus provides an effective way of forming new operators having specified properties out of given ones, in a very general context. We will not have to make use of the analytic functional calculus in this book. In this section we describe this calculus in some detail, but refer the reader to other sources (such as [**12**]) for a treatment that includes proofs we have omitted.

Let C be a simple closed oriented curve in the complex plane $\mathbb{C}$ that is piecewise continuously differentiable. We refer to such objects simply as oriented curves. Thus, an oriented curve C can be parameterized in different ways by continuous functions $\gamma : [0,1] \to C$ that are piecewise continuously differentiable, one-to-one on $[0,1)$, and periodic $\gamma(0) = \gamma(1)$. Every continuous function f on C can be integrated around C by either forming a limit of appropriate Riemann sums that respect the orientation of C, or alternatively by choosing a parameterization $\gamma : [0,1] \to \Gamma$ consistent with the orientation and setting

$$\int_C f(\lambda)\, d\lambda = \int_0^1 f(\gamma(t))\gamma'(t)\, dt.$$

The notion of integral over C generalizes in a straightforward way to vector-valued functions, namely to continuous functions f defined on C that take values in a Banach space E. Fixing such a function f, one considers finite oriented partitions $\mathcal{P} = \{\gamma_0, \gamma_1, \dots, \gamma_n\}$ of the curve C (that is, partitions of C that are consistent with its orientation). With every such partition there is a corresponding Riemann sum

$$R(f, \mathcal{P}) = \sum_{k=1}^{n} f(\gamma_k)(\gamma_k - \gamma_{k-1}),$$

and the techniques of elementary calculus can be adapted in a straightforward way to show that the limit of these Riemann sums exists (as the norm $\|\mathcal{P}\| = \max_k |\gamma_k - \gamma_{k-1}|$ of the partition $\mathcal{P}$ tends to 0) relative to the norm topology of E. See Exercise (1) below. Thus one can define

$$\int_C f(\lambda)\, d\lambda = \lim_{\|\mathcal{P}\| \to 0} R(f, \mathcal{P}),$$

and one has the estimate

$$\left\| \int_C f(\lambda)\, d\lambda \right\| \le \int_C \|f(\lambda))\|\, d|\lambda| \le \sup_{\lambda \in C} \|f(\lambda)\| \ell(C), \tag{1.18}$$

$\ell(C)$ denoting the length of C. It follows that for every bounded linear functional ρ on E we have

$$\rho\left(\int_C f(\lambda)\, d\lambda \right) = \int_C \rho(f(\lambda))\, d\lambda.$$

Reversing the orientation of C has the effect of replacing $\int_C f(\lambda)\, d\lambda$ with its negative $-\int_C f(\lambda)\, d\lambda$. Thus we have assigned a clear meaning to the integral of a continuous function $f : C \to E$ as an element of E.

We also require a few facts about the classical notion of winding number. Let C be an oriented curve. Then for every λ in $\mathbb{C} \setminus C$ we can define an integer [**1**]

$$n(C, \lambda) = \frac{1}{2\pi i} \int_C \frac{dz}{z - \lambda}.$$

If λ belongs to the bounded component of the complement of C then one has $n(C,\lambda) = 1$ when C is oriented counterclockwise and $n(C,\lambda) = -1$ otherwise. On the other hand, $n(C,\lambda) = 0$ if λ belongs to the unbounded component of $\mathbb{C} \setminus C$, regardless of orientation.

A *cycle* is an element of the abelian group generated by oriented curves, subject to the relation $C + C^* = 0$, where C^* denotes the curve obtained by reversing the orientation of C. To review terminology, let S be a set that is endowed with an involutory map $s \mapsto s^*$, $s \in S$, and let $G(S)$ be the free abelian group generated by S modulo the subgroup generated by $s + s^*$, $s \in S$. In more concrete terms, the free abelian group generated by S can be realized as the abelian group $\mathbb{Z}(S)$ of integer-valued functions $n : S \to \mathbb{Z}$ that satisfy $n(s) = 0$ off some finite subset of S, with the pointwise operations

$$(m + n)(s) = m(s) + n(s), \qquad s \in S.$$

There is a natural notion of linear combinations of elements of $\mathbb{Z}(S)$; for $p, q \in \mathbb{Z}$ and $m, n \in \mathbb{Z}(S)$, $p \cdot m + q \cdot n$ denotes the function $s \mapsto pm(s) + qn(s)$. If we identify elements of S with their image in $\mathbb{Z}(S)$ by way of $s \in S \mapsto \chi_{\{s\}}$, then the elements of $\mathbb{Z}(S)$ are linear combinations of elements of S,

$$p_1 \cdot s_1 + \cdots + p_n \cdot s_n, \qquad p_k \in \mathbb{Z}, \quad s_k \in S.$$

The subgroup $H \subseteq \mathbb{Z}(S)$ generated by elements of the form $s + s^*$ is identified with the subgroup of all functions $n \in \mathbb{Z}(S)$ satisfying $n(s) \in 2\mathbb{Z}$ if $s^* = s$ and $n(s^*) = n(s)$ if $s^* \neq s$ (note that for the example in which S consists of oriented curves, the case $s^* = s$ never occurs). Letting $\dot{s}$ denote the coset $s + H \in G(S)$, then $\dot{s}^* = -\dot{s}$, and the most general element of $G(S)$ is a linear combination

$$p_1 \cdot \dot{s}_1 + \cdots + p_n \cdot \dot{s}_n.$$

The universal property that follows from this construction asserts that every function ϕ from S to an abelian group G that satisfies $\phi(s^*) = -\phi(s)$ for all s can be extended uniquely to a group homomorphism $\hat{\phi} : G(S) \to G$, which acts on elements of $G(S)$ as follows:

$$\hat{\phi}\left(\sum_{k=1}^{n} p_k \cdot \dot{s}_k\right) = \sum_{k=1}^{n} p_k \cdot \phi(s_k).$$

A cycle can be visualized as a conglomerate of several oriented curves, traversed one by one, perhaps several times. Every nonzero cycle Γ can be written as a linear combination $\Gamma = p_1\dot{C}_1 + \cdots + p_n\dot{C}_n$ with nonzero integer coefficients p_k, where the C_k are oriented curves with the property $C_k \notin \{C_j, C_j^*\}$ for $k \neq j$. This expression for Γ is not unique, but the lack of uniqueness is characterized by the simple fact that $p \cdot \dot{s} = -p \cdot \dot{s}^*$, $p \in \mathbb{Z}$, $s \in S$. Thus the union of sets $C_1 \cup \cdots \cup C_n$ (point sets without orientation) is uniquely determined, and we think of this set as the underlying point set of Γ. The empty set is the underlying point set of the zero cycle. Fixing $\lambda \in \mathbb{C}$, the set of all cycles that do not contain λ is a subgroup of the group of all cycles (it is the universal group of cycles generated by all oriented curves that do not contain λ); hence for every such Γ there is a well-defined winding number $n(\Gamma, \lambda) \in \mathbb{Z}$ defined by general principles as above by taking $\phi(C) = n(C, \lambda)$ on oriented curves C. The map $\Gamma \mapsto n(\Gamma, \lambda)$ is a homomorphism of the group of all cycles that do not contain λ into $\mathbb{Z}$.

It is important that cycles, like curves, have well-defined interiors.

Definition 1.12.1. Let Γ be a cycle. The interior of Γ is defined as the set of all points $\lambda \in \mathbb{C} \setminus \Gamma$ such that $n(\Gamma, \lambda) \neq 0$, and it is written $\operatorname{int}(\Gamma)$.

It is a worthwhile exercise to experiment with this definition. For example, consider a cycle Γ consisting of two concentric circles of different radii about the origin. If the outer circle and inner circle have the same orientation, then that cycle has interior consisting of all points within the outer circle that do not belong to the inner circle. If the two circles have opposite orientations, then the interior of the cycle consists of the annular region lying between the two circles.

If we are given an open set $U \subseteq \mathbb{C}$, a Banach space E, and a continuous function $f : U \to E$, then we have seen how to define the integral of f over any oriented curve $C \subseteq U$. The set of all cycles whose underlying point sets are contained in U is also a group with a similar universal property, namely the universal group generated by the oriented curves contained in U. Thus by general principles we have a definition of

$$\int_\Gamma f(\lambda)\, d\lambda \in E$$

for all cycles $\Gamma \subseteq U$, and this integral satisfies

$$\int_{\Gamma_1 + \Gamma_2} f(\lambda)\, d\lambda = \int_{\Gamma_1} f(\lambda)\, d\lambda + \int_{\Gamma_2} f(\lambda)\, d\lambda.$$

Finally, we introduce the algebra of locally analytic functions on a compact subset of $\mathbb{C}$. Let $X \subseteq \mathbb{C}$ be compact. By a locally analytic function on X we mean an analytic function f defined on some open set $U \supseteq X$. Two such functions f (defined on $U \supseteq X$) and g (defined on $V \supseteq X$) are said to be equivalent if there is an open set W such that $X \subset W \subset U \cap V$ and the restrictions of f and g to W agree. The set of equivalence classes of locally analytic functions on X forms a complex algebra, whose unit is the class of the constant function $f(z) = 1$, $z \in \mathbb{C}$. This commutative algebra is denoted by $\mathcal{A}(X)$.

We now have an effective notion of cycle, a notion of the integral of a vector-valued function over a cycle contained in the interior of its domain, and the notion of an algebra of locally analytic functions $\mathcal{A}(X)$ associated with a compact set $X \subseteq \mathbb{C}$. These are the basic constituents of the analytic functional calculus, which we now describe.

Let A be a Banach algebra with normalized unit $\mathbf{1}$ and fix an element $a \in A$ with spectrum $X = \sigma(a)$. Given $f \in \mathcal{A}(X)$ we want to define $f(a)$

in a manner consistent with the Cauchy integral theorem. To do this we choose a cycle Γ with the following properties:

- f is analytic on $\Gamma \cup \mathrm{int}(\Gamma)$.
- $\Gamma \cap X = \emptyset$.
- $n(\Gamma, z) = 1$ for all $z \in X$.

The first and third conditions together imply that there is a representative in the class of f whose domain contains not only X and Γ, but also all points interior to Γ. The third condition asserts that the cycle winds around every point of X exactly once in the positive direction, allowing for cancellations as one moves along the various components of Γ.

For example, if X is the unit circle and f is an analytic function defined on some annular region $U = \{z \in \mathbb{C} : r < |z| < R\}$ where $0 < r < 1 < R < \infty$, one may take Γ to be the union of two circles $\Gamma_k = \{|z| = r_k\}$, $k = 1, 2$, where $r < r_1 < 1 < r_2 < R$, where Γ_2 is oriented in the counterclockwise direction, and Γ_1 is oriented clockwise.

Consider the resolvent function $(\lambda\mathbf{1} - a)^{-1}$. This is certainly defined for all λ in an open set containing Γ, and it is a continuous function with values in A. Thus we can define

$$f(a) = \frac{1}{2\pi i} \int_\Gamma f(\lambda)(\lambda\mathbf{1} - a)^{-1}\, d\lambda. \tag{1.19}$$

The fact is that $f(a)$ depends on neither the particular choice of Γ nor the choice of representative of f (this is an exercise in the use of the Cauchy integral theorem of complex analysis). Moreover, $f \in \mathcal{A}(X) \mapsto f(a)$ is a unital homomorphism of complex algebras that has the following property: For every power series

$$f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$$

converging on some open disk $\{|z| < R\}$ containing X, the corresponding series $c_0\mathbf{1} + c_1 a + c_2 a^2 + \cdots$ is absolutely convergent relative to the norm of A, and we have

$$f(a) = \sum_{n=1}^{\infty} c_n a^n.$$

The reader is referred to pp. 566–577 of [**12**] for further detail.

Exercises.

(1) Let C be an oriented curve in $\mathbb{C}$, let f be a continuous function defined on C taking values in a Banach space E, and consider the set of all finite oriented partitions $\mathcal{P}$ of C.
   (a) Show that for every $\epsilon > 0$ there is a $\delta > 0$ with the property that for every pair of oriented partitions $\mathcal{P}_1$, $\mathcal{P}_2$ satisfying $\|\mathcal{P}_k\| \le \delta$ for $k = 1, 2$, one has $\|R(f, \mathcal{P}_1) - R(f, \mathcal{P}_2)\| \le \epsilon$.
   (b) Verify the estimate (1.18).

Let T be a bounded operator on a Banach space E.

(2) Let $D = \{z \in \mathbb{C} : |z| < R\}$ be an open disk containing $\sigma(T)$. Let $f : D \to \mathbb{C}$ be an analytic function defined on D, with power series

$$f(z) = \sum_{n=0}^{\infty} c_n z^n, \qquad z \in D.$$

Show that the infinite series of operators

$$\sum_{n=0}^{\infty} c_n T^n$$

converges absolutely in the sense that $\sum_n |c_n| \|T^n\| < \infty$.

(3) Give a definition of $\sin T$ and $\cos T$ using power series.

(4) Use your definitions in the preceding exercise to show that

$$(\sin T)^2 + (\cos T)^2 = \mathbf{1}.$$

CHAPTER 2

Operators on Hilbert Space

We now take up the theory of operators on Hilbert space. It is appropriate to develop this subject within the context of C^*-algebras, and the most basic properties of C^*-algebras, their ideals, quotients, and morphisms, are worked out in this chapter. We discuss commutative C^*-algebras in detail, including the characterization of $C(X)$, the functional calculus for normal operators, and the spectral theorem. Unfortunately, the literature of operator theory contains at least three dissimilar statements that are called the spectral theorem. The assertions are that normal operators are associated with multiplication operators, that they are associated with spectral measures, and that they admit a Borel functional calculus. While these statements are all in some sense equivalent, only the first of them is a clear generalization of the idea of diagonalizing a matrix, and that is the one we offer as the proper up-to-date formulation of the spectral theorem.

Throughout this chapter, Hilbert spaces will be assumed separable or finite dimensional. This is an unnecessary restriction, since all the results we discuss have appropriate generalizations to the inseparable cases. But the formulation of the spectral theorem that we use becomes somewhat esoteric for inseparable spaces, and in dealing with traces or Hilbert–Schmidt operators, the fact that orthonormal bases $\{e_\alpha : \alpha \in I\}$ are uncountable while the corresponding sums of numbers $\sum_\alpha \|Ae_\alpha\|^2$ have only countably many nonzero terms can distract attention from the fundamental issues of analysis. In some cases we offer comments to assist the generalizers in carrying out their work.

2.1. Operators and Their C^*-Algebras

In this section, we discuss the operator-theoretic version of the Riesz lemma, we introduce some commonly used terminology, and we discuss the multiplication algebra of a measure space. Throughout, H will denote a Hilbert space with inner product $\langle \xi, \eta \rangle$, linear in ξ and antilinear in η.

The Riesz lemma asserts that every bounded linear functional f on H can be represented uniquely as the inner product with a vector $\eta \in H$,

$$f(\xi) = \langle \xi, \eta \rangle, \qquad \xi \in H;$$

moreover, one has $\|f\| = \|\eta\|$. The Riesz lemma implies that the mapping $f \to \eta$ is an *anti*linear isometry of the dual of H onto H.

Every operator $A \in \mathcal{B}(H)$ gives rise to a complex-valued function of two variables $[\xi, \eta] = \langle A\xi, \eta\rangle$, $\xi, \eta \in H$. Notice that this form is linear in ξ and antilinear in η; such bilinear forms are called *sesquilinear*. The sesquilinear form associated with A is also bounded in the sense that there is a positive constant C such that $|[\xi, \eta]| \leq C\|\xi\|\|\eta\|$ for all $\xi, \eta \in H$, and the smallest such constant is the operator norm $C = \|A\|$. Frequently, the easiest way to define a bounded operator is to specify its sesquilinear form. The following result guarantees the existence of a unique operator in such definitions, and is also called the Riesz lemma.

Proposition 2.1.1. *For every bounded complex-valued sesquilinear form $[\cdot, \cdot]$ on H there is a unique bounded operator A on H such that*

$$[\xi, \eta] = \langle A\xi, \eta\rangle, \qquad \xi, \eta \in H.$$

Proof. Fix a vector $\xi \in H$ and consider the linear functional f defined on H by $f(\eta) = \overline{[\xi, \eta]}$, the bar denoting complex conjugation. Since f is a bounded linear functional, the Riesz lemma in its above form implies that there is a unique vector $A\xi \in H$ satisfying $f(\eta) = \langle \eta, A\xi\rangle$; and after taking the complex conjugate we find that the function $A : H \to H$ that we have defined must satisfy

$$[\xi, \eta] = \langle A\xi, \eta\rangle, \qquad \xi, \eta \in H.$$

It is straightforward to verify that this formula implies that A is a linear transformation. It is bounded because

$$\sup_{\|\xi\|\leq 1} \|A\xi\| = \sup_{\|\xi\|\leq 1, \|\eta\|\leq 1} |[\xi, \eta]| < \infty.$$

The uniqueness of the operator A is evident from the uniqueness assertion of the Riesz lemma for linear functionals. □

Similarly, there is a characterization of bounded operators $A \in \mathcal{B}(H, K)$ from one Hilbert space to another in terms of bounded sesquilinear forms $[\cdot, \cdot] : H \times K \to \mathbb{C}$ by way of the identification $[\xi, \eta] = \langle A\xi, \eta\rangle$, $\xi \in H$, $\eta \in K$. Note that the inner product on the right is that of K, not H.

We immediately deduce the existence of adjoints of bounded operators from one Hilbert space to another. When more than one Hilbert space is involved there might be confusion about the meaning of inner products; when we want to be explicit about which inner product is involved we will write $\langle \xi, \eta\rangle_H$ for the inner product of two vectors $\xi, \eta \in H$.

Corollary 1. Let H, K be Hilbert spaces and let $A \in \mathcal{B}(H, K)$ be a bounded operator from H to K. There is a unique operator $A^* \in \mathcal{B}(K, H)$ satisfying

$$\langle A\xi, \eta\rangle_K = \langle \xi, A^*\eta\rangle_H, \qquad \xi \in H, \quad \eta \in K.$$

PROOF. One simply applies the above results to the bounded sesquilinear form $[\cdot,\cdot]$ defined on $K \times H$ by $[\eta,\xi] = \langle \eta, A\xi\rangle$ to deduce the existence of a unique operator $A^* \in \mathcal{B}(K,H)$ satisfying $\langle A^*\eta,\xi\rangle_H = \langle \eta, A\xi\rangle_K$, and then takes the complex conjugate of both sides. □

The case $H = K$ is of particular importance, since we may deduce that for every $A \in \mathcal{B}(H)$ there is a unique operator $A^* \in \mathcal{B}(H)$ such that $\langle A\xi,\eta\rangle = \langle \xi, A^*\eta\rangle$. The basic properties of the mapping $A \mapsto A^*$ are summarized as follows:

(1) $A^{**} = A$.
(2) $(\lambda A + \mu B)^* = \bar{\lambda}A^* + \bar{\mu}B^*$.
(3) $(AB)^* = B^*A^*$.
(4) $\|A^*A\| = \|A\|^2$.

Properties (1), (2), (3) together define an *involution* in a complex algebra. Property (4) is the critical relation between the norm in $\mathcal{B}(H)$ to the involution. It is the characteristic property of a C^*-algebra (see Definition 2.2.1 below). To verify property (4), note that $\|A^*A\|$ is given by

$$\sup_{\|\xi\|,\|\eta\|\le 1} |\langle A^*A\xi,\eta\rangle| = \sup_{\|\xi\|,\|\eta\|\le 1} |\langle A\xi, A\eta\rangle| \le \sup_{\|\xi\|,\|\eta\|\le 1} \|A\xi\|\|A\eta\| = \|A\|^2$$

while on the other hand,

$$\|A\|^2 = \sup_{\|\xi\|\le 1} \langle A\xi, A\xi\rangle = \sup_{\|\xi\|\le 1} \langle A^*A\xi,\xi\rangle \le \|A^*A\|.$$

We will also make use of standard terminology for various types of operators $A \in \mathcal{B}(H)$. An operator A is called *normal* if it commutes with its adjoint, $A^*A = AA^*$. An operator A on H is an isometry iff $\langle A\xi, A\xi\rangle = \langle \xi,\xi\rangle$ for every $\xi \in H$ and in turn this is equivalent to the equation $A^*A = \mathbf{1}$. An invertible isometry A is characterized by $A^*A = AA^* = \mathbf{1}$ and is called a *unitary* operator. A self-adjoint operator with nonnegative spectrum is called a *positive* operator. It is a nontrivial fact that positivity is characterized by the condition $\langle A\xi,\xi\rangle \ge 0$ for every $\xi \in H$, as we will see. More generally, for two self-adjoint operators A and B one writes $A \le B$ if $B - A$ is positive. Finally, a *projection* is a self-adjoint idempotent: $A^2 = A = A^*$.

The following elementary facts about the geometry of Hilbert spaces will be used freely below:

(1) Every nonempty closed convex set C in a Hilbert space H has a unique element of smallest norm; that is, there is a unique element $x \in C$ such that $\|x\| = \inf\{\|y\| : y \in C\}$.
(2) Let M be a closed linear subspace of H. Then every vector $\xi \in H$ has a unique decomposition $\xi = \xi_1 + \xi_2$ where $\xi_1 \in M$ and $\xi_2 \in M^\perp = \{\eta \in H : \langle \eta, M\rangle = \{0\}\}$.
(3) Let P be any projection in $\mathcal{B}(H)$. Then $M = \{\xi \in H : P\xi = \xi\}$ is a closed subspace of H. Conversely, every closed subspace of H is associated in this way with a unique projection $P \in \mathcal{B}(H)$.

DEFINITION 2.1.2. A *C^*-algebra of operators* is a norm-closed subalgebra $\mathcal{A} \subseteq \mathcal{B}(H)$ of the algebra of all bounded operators on some Hilbert space, which is also closed under the adjoint operation $\mathcal{A}^* = \mathcal{A}$.

There are many examples of such C^*-algebras. For example, let $\mathcal{S} \subseteq \mathcal{B}(H)$ be any nonempty set of operators. The intersection of all C^*-algebras in $\mathcal{B}(H)$ that contain $\mathcal{S}$ is called the C^*-algebra generated by $\mathcal{S}$, often written $C^*(\mathcal{S})$. It can be realized in somewhat more concrete terms as follows. Consider the set $\mathcal{P}$ of all finite products $T_1T_2\cdots T_n$, $n = 1, 2, \dots$, where $T_k \in \mathcal{S} \cup \mathcal{S}^*$. The set of all finite linear combinations of elements of $\mathcal{P}$ is obviously the smallest self-adjoint algebra containing $\mathcal{S}$, and hence its norm-closure is the C^*-algebra generated by $\mathcal{S}$. While this "construction" appears to exhibit the elements of $C^*(\mathcal{S})$ in a systematic way, it is not very useful for obtaining structural information about $C^*(\mathcal{S})$, since the nature of the limits of such linear combinations has not been made explicit.

A substantial amount of current work in noncommutative analysis has gone into determining the properties and structure of the C^*-algebra generated by a finite set of operators that satisfy certain relations.

The norm topology on $\mathcal{B}(H)$ is inappropriate for topological issues that require more flexibility, and $\mathcal{B}(H)$ has several useful and natural topologies that are weaker than the norm topology. We will have to make use of only two of them. In general, a locally convex topology can be defined on a complex vector space V by specifying a family $\mathcal{S}$ of seminorms on V that separates the points of V. Given a finite subset $F = \{|\cdot|_1, \dots, |\cdot|_n\} \subseteq \mathcal{S}$ and a positive ϵ, one associates a corresponding subset of V:

$$U_{F,\epsilon} = \{x \in V : |x|_1 < \epsilon, \dots, |x|_n < \epsilon\}.$$

The set of all such $U_{F,\epsilon}$ is a basic system of neighborhoods of the origin for a unique locally convex Hausdorff topology on V.

For example, the norm topology is defined by the somewhat degenerate family $\mathcal{S} = \{\|\cdot\|\}$, where $\|\cdot\|$ is the operator norm. The *weak operator topology* is defined by the family of seminorms $|A| = |\langle A\xi, \eta\rangle|$, ξ, η ranging over all vectors in H. The *strong operator topology* is defined by the family of seminorms $|A| = \|A\xi\|$, where $\xi \in H$. For example, a net of operators $A_n \in \mathcal{B}(H)$ converges strongly to 0 if and only if for every $\xi \in H$,

$$\lim_{n\to\infty} \|A_n\xi\| = 0.$$

A *von Neumann algebra* is a self-adjoint subalgebra of $\mathcal{B}(H)$ that contains the identity operator and is closed in the weak operator topology. While it is true that von Neumann algebras are C^*-algebras of operators, they have many properties that are not shared by more general C^*-algebras. For example, von Neumann algebras contain enough projections to generate them as C^*-algebras, while more general unital C^*-algebras may contain no projections other than the trivial ones $\mathbf{0}$ and $\mathbf{1}$. The theory of von Neumann algebras has undergone extensive development, and it has a different flavor from that of the general theory of C^*-algebras. It is appropriate to view the theory of von Neumann algebras as a noncommutative generalization of measure theory, and to view the theory of C^*-algebras as a noncommutative generalization of the theory of topological spaces [**8**].

Let $\mathcal{S} \subseteq \mathcal{B}(H)$ be a set of operators. The *commutant* of $\mathcal{S}$ is the set of all operators $T \in \mathcal{B}(H)$ satisfying $ST = TS$ for every $S \in \mathcal{S}$; it is denoted by $\mathcal{S}'$. The commutant of any set of operators is an algebra containing the identity operator, and one may easily check that $\mathcal{S}'$ *is a weakly closed unital subalgebra of* $\mathcal{B}(H)$. If $\mathcal{S} = \mathcal{S}^*$ is closed under the involution of $\mathcal{B}(H)$, then $\mathcal{S}'$ is a von Neumann algebra.

We conclude the section with a discussion of multiplication operators on Hilbert spaces associated with measure spaces. Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space; we suppress explicit reference to the σ-algebra of sets $\mathcal{B}$ unless there is cause for confusion. $L^2(X, \mu)$ is a Hilbert space, which may or may not be separable; the measure space (X, μ) is called *separable* when $L^2(X, \mu)$ is a separable Hilbert space. Every function $f \in L^\infty(X, \mu)$ gives rise to an operator M_f that acts as follows:

$$(M_f\xi)(p) = f(p)\xi(p), \qquad p \in X, \quad \xi \in L^2(X, \mu).$$

$L^\infty(X, \mu)$ is a commutative C^*-algebra with unit relative to its pointwise operations and its essential norm

$$\|f\|_\infty = \operatorname{ess\,sup}\{|f(p)| : p \in X\}.$$

In more detail, the involution in $L^\infty(X, \mu)$ is defined by $f^*(p) = \overline{f(p)}$, $p \in X$; the norm is

$$\|f\|_\infty = \sup\{t > 0 : \mu\{p \in X : |f(p)| > t\} > 0\};$$

and the involution is related to the norm by $\|f^*f\|_\infty = \|f\|_\infty^2$.

THEOREM 2.1.3. *For every $f \in L^\infty(X, \mu)$ M_f is a bounded operator on $L^2(X, \mu)$. The map $f \mapsto M_f$ is an isometric $*$-isomorphism of $L^\infty(X, \mu)$ onto a commutative C^*-algebra of operators $\mathcal{M} \subseteq \mathcal{B}(H)$.*

PROOF. The key assertion here is $\|M_f\| = \|f\|_\infty$. Indeed, the inequality $\leq$ is clear from the fact that $|f(p)| \leq \|f\|_\infty$ for almost every $p \in X$, since this entails $|f \cdot \xi| \leq \|f\|_\infty |\xi|$ pointwise almost everywhere for $\xi \in L^2(X, \mu)$, hence $\|f \cdot \xi\|_2 \leq \|f\|_\infty \|\xi\|_2$. For the opposite inequality, assume $f \neq 0$ and choose a number c, $0 \leq c < \|f\|_\infty$. The set $\{p \in X : |f(p)| > c\}$ has positive measure, so by σ-finiteness we can find a subset $E \subseteq \{p \in X : |f(p)| > c\}$ having finite positive measure. Thus $\chi_E \in L^2(X, \mu)$ and from

$$|f(p) \cdot \chi_E(p)| \geq c\chi_E(p), \qquad p \in X,$$

we obtain $\|M_f \chi_E\|_2 \geq c\|\chi_E\|_2$ after squaring and integrating. Since χ_E is not the zero element of $L^2(X, \mu)$, $\|M_f\| \geq c$. The inequality $\|M_f\| \geq \|f\|_\infty$ follows after one takes the supremum over such c.

Obviously, $f \mapsto M_f$ is a homomorphism of algebras that carries the unit of $L^\infty(X, \mu)$ to $\mathbf{1}$, and one may verify $M_f^* = M_{f^*}$ directly. The set of operators $\{M_f : f \in L^\infty\}$ is norm-closed because L^∞ is a Banach space. $\square$

The set of operators $\mathcal{M} = \{M_f : f \in L^\infty(X, \mu)\}$ is called the *multiplication algebra* of the measure space (X, μ). It is an abelian von Neumann algebra, since it is closed in the weak operator topology, though that is not obvious from what has been said. We will look more closely at multiplication algebras in Chapter 4.

Let us now compute the spectra of multiplication operators. Since an element of $L^\infty(X, \mu)$ is not a function but an equivalence class of functions that agree almost everywhere, the notion of the range of $f \in L^\infty(X, \mu)$ must be approached with some care. Choose a representative in the class of f, which we will call f. We can use f to define a measure m_f on the σ-algebra of Borel sets in $\mathbb{C}$:

$$m_f(S) = \mu\{p \in X : f(p) \in S\}, \qquad S \subseteq \mathbb{C}.$$

It is a straightforward exercise to show that every function g that agrees almost everywhere with f gives rise to the same measure, $m_g = m_f$; hence this measure depends only on the *equivalence class* of f as an element of $L^\infty(X, \mu)$. If μ is a finite measure, then so is m_f. But if μ is only σ-finite, then m_f need not be σ-finite; indeed, in such cases points of $\mathbb{C}$ can have infinite m_f-measure (consider the case of a constant function f). In all cases, however, m_f is a countably additive measure defined on the Borel σ-algebra of the complex plane. As such it has a uniquely defined closed support, defined as follows. By the Lindelöf property, the union G of all open subsets of $\mathbb{C}$ having m_f-measure zero can be reduced to the union of a countable subfamily of open sets of measure zero; hence G satisfies $m_f(G) = 0$. Obviously, G is the largest open set of m_f-measure zero. It follows that the complement $F = \mathbb{C} \setminus G$ is a closed set with the following property: A complex number λ belongs to F if and only if for every $\epsilon > 0$ we have

$$\mu\{p \in X : |f(p) - \lambda| < \epsilon\} > 0. \tag{2.1}$$

Moreover, every point of the complement of F has a neighborhood of m_f-measure zero.

The set F is called the *essential range* of f. To reiterate: λ belongs to the essential range of f if and only if every neighborhood of λ has positive m_f-measure. The essential range of f is a compact set F with the property that

$$\|f\|_\infty = \sup\{|\lambda| : \lambda \in F\}.$$

THEOREM 2.1.4. *For every $f \in L^\infty(X, \mu)$, the spectrum of the multiplication operator M_f is the essential range of f.*

PROOF. If λ does not belong to the essential range of f, then there is an $\epsilon > 0$ such that $\{p \in X : |f(p) - \lambda| < \epsilon\} = 0$, i.e., $|f(p) - \lambda| \geq \epsilon$ almost everywhere $(d\mu)$. It follows that the function

$$g(p) = \frac{1}{f(p) - \lambda}, \qquad p \in X,$$

belongs to $L^\infty(X,\mu)$, and its multiplication operator M_g is a left and right inverse of $M_f - \lambda\mathbf{1}$.

Conversely, suppose λ is a point in the essential range of f. We will exhibit a sequence of unit vectors $\xi_1, \xi_2, \dots \in L^2(X,\mu)$ with the property

$$\lim_{n\to\infty} \|M_f\xi_n - \lambda\xi_n\| = 0,$$

showing that $\lambda \in \sigma(M_f)$. Indeed, $\{p \in X : |f(p) - \lambda| \leq 1/n\}$ is a set of positive measure for every $n = 1, 2, \dots$, and using σ-finiteness of μ we find a subset

$$E_n \subseteq \{p \in X : |f(p) - \lambda| \leq 1/n\}$$

satisfying $0 < \mu(E_n) < \infty$. Letting ξ_n be the unit vector $\mu(E_n)^{-1/2}\chi_{E_n}$ one has

$$|(f(p) - \lambda)\xi_n(p)| \leq n^{-1}|\xi_n(p)|, \qquad p \in X,$$

and hence $\|(f - \lambda)\xi_n\|_{L^2} \leq 1/n$ tends to 0 as $n \to \infty$. □

Exercises.

(1) Let $[\cdot,\cdot] : H \times H \to \mathbb{C}$ be a sesquilinear form defined on a Hilbert space H. Show that $[\cdot,\cdot]$ satisfies the *polarization formula*

$$4[\xi,\eta] = \sum_{k=0}^{3} i^k \left[\xi + i^k\eta, \xi + i^k\eta\right].$$

(2) Let $A \in \mathcal{B}(H)$ be a Hilbert space operator. The *quadratic form* of A is the function $q_A : H \to \mathbb{C}$ defined by $q_A(\xi) = \langle A\xi, \xi\rangle$. The *numerical range* and *numerical radius* of A are defined, respectively, by

$$W(A) = \{q_A(\xi) : \|\xi\| = 1\} \subseteq \mathbb{C},$$
$$w(A) = \sup\{|q_A(\xi)| : \|\xi\| = 1\}.$$

(a) Show that A is self-adjoint iff q_A is real-valued.

(b) Show that $w(A) \leq \|A\| \leq 2w(A)$ and deduce that $q_A = q_B$ only when $A = B$. Hint: Polarize.

(3) Show that the adjoint operation $A \mapsto A^*$ in $\mathcal{B}(H)$ is weakly continuous but not strongly continuous. Hint: Consider the sequence of powers of the unilateral shift $S, S^2, S^3, \dots$.

(4) Show that the only operators that commute with all operators in $\mathcal{B}(H)$ are the scalar multiples of the identity.

(5) Let $\mathcal{C}$ be the closure in the strong operator topology of the set of all unitary operators in $\mathcal{B}(H)$. Show that $\mathcal{C}$ consists of isometries.

(6) Show that the unilateral shift S belongs to $\mathcal{C}$ by exhibiting a sequence of unitary operators $U_1, U_2, \dots$ that converges to S in the strong operator topology. Hint: Consider the matrix of S relative to the obvious basis, and look for unitary matrices that strongly approximate large $n \times n$ blocks of it.

(7) Let (X, μ) be a σ-finite measure space and let $f : X \to \mathbb{C}$ be a bounded complex-valued Borel function. Show that the essential range of f can be characterized as the intersection

$$\bigcap\{\overline{g(X)} : g \sim f\}$$

of the *closed* ranges of all bounded Borel functions $g : X \to \mathbb{C}$ that agree with f almost everywhere $(d\mu)$.

2.2. Commutative C^*-Algebras

DEFINITION 2.2.1. A C^*-algebra is a Banach algebra A that is endowed with an involution $x \mapsto x^*$ satisfying $\|x^*x\| = \|x\|^2$ for every $x \in A$.

More explicitly, the involution $*$ is an antilinear mapping of A into itself that satisfies $(xy)^* = y^*x^*$, $x^{**} = x$, and is related to the norm of A by the asserted formula. C^*-algebras need not contain a unit. Any norm-closed self-adjoint subalgebra $\mathcal{A}$ of $\mathcal{B}(H)$ is a C^*-algebra, as we have seen in the previous section. On the other hand, abstract C^*-algebras are not necessarily associated with operators on any specific Hilbert space.

We now show that every commutative C^*-algebra with unit is isometrically $*$-isomorphic to the algebra $C(X)$ of all complex-valued continuous functions on a compact Hausdorff space X. A similar result holds for nonunital commutative C^*-algebras, provided that one is willing to replace X with a locally compact Hausdorff space and $C(X)$ with the algebra of continuous functions vanishing at infinity. We will confine attention to the unital case here; the nonunital generalization can be found in [**2**], for example.

This C^*-algebraic characterization of spaces has led analysts to think of noncommutative C^*-algebras as noncommutative generalizations of topological spaces, and of problems concerning the classification of these algebras up to $*$-isomorphism as a noncommutative generalization of (algebraic) topology. For example, the K-theory of spaces developed by Grothendieck, Atiyah, Bott, and others during the period 1955–1965 has now been generalized to C^*-algebras in a way that provides effective tools for the computation of these invariants [**8**]. Indeed, contemporary work on the classification of *simple* C^*-algebras has led to the expectation that the most important simple C^*-algebras are completely determined by their K-theory! Since very different topological spaces can have the same K-theory, this is an aspect of "noncommutative topology" that is entirely new and has no counterpart in the classical theory of topological spaces.

We begin with a brief discussion of the exponential map in a (perhaps noncommutative) unital Banach algebra A. For every element $x \in A$ the exponential of x is defined by

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n. \tag{2.2}$$

Notice that this series converges absolutely, since

$$\sum_{n=0}^{\infty} \|\frac{1}{n!}x^n\| \leq \sum_{n=0}^{\infty} \|x\|^n/n! = e^{\|x\|} < \infty,$$

and we have the estimate $\|e^x\| \leq e^{\|x\|}$. Obviously, $e^0 = \mathbf{1}$.

REMARK 2.2.2. *Rearranging products of series.* Let $a_0, a_1, a_2, \dots$ and $b_0, b_1, b_2, \dots$ be two sequences of elements of A such that $\sum_n \|a_n\| < \infty$, $\sum_n \|b_n\| < \infty$, and let $x = \sum_n a_n$, $y = \sum_n b_n$. Then the product xy is given by the series $xy = \sum_n c_n$, where

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0, \qquad n = 0, 1, 2, \dots, \tag{2.3}$$

the series $\sum_n c_n$ being absolutely convergent in the sense that $\sum_n \|c_n\| < \infty$. The proof is an instructive exercise in making estimates, and is left for the reader in Exercise (1) below.

PROPOSITION 2.2.3. *Let x, y be elements of a unital Banach algebra A satisfying $xy = yx$. Then $e^{x+y} = e^x e^y$.*

PROOF. Using formula (2.3), we have

$$e^x e^y = \sum_{p,q=0}^{\infty} \frac{1}{p!}x^p \frac{1}{q!}y^q = \sum_{n=0}^{\infty} \left(\sum_{p+q=n} \frac{1}{p!q!}x^p y^q \right).$$

Since $xy = yx$, the proof of the binomial theorem applies here to give

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = n! \sum_{p+q=n} \frac{1}{p!q!}x^p y^q;$$

hence the right side of the preceding formula becomes

$$\sum_{n=0}^{\infty} \frac{1}{n!}(x+y)^n = e^{x+y}.$$

$\square$

Much of the terminology introduced in the preceding section can be applied to abstract C^*-algebras as well as C^*-algebras of operators. For example, a normal element of a C^*-algebra is an element that commutes with its adjoint, and a unitary element of a unital C^*-algebra is an element u satisfying $u^*u = uu^* = \mathbf{1}$. A unitary element has norm 1, since $\|u\|^2 = \|u^*u\| = \|\mathbf{1}\| = 1$.

THEOREM 2.2.4. *Let A be a commutative C^*-algebra with unit, and let $X = \operatorname{sp}(A)$ be the Gelfand spectrum of A. Then the Gelfand map is an isometric $*$-isomorphism of A onto $C(X)$.*

Proof. We show first that every $\omega \in \mathrm{sp}(A)$ preserves the adjoint in the sense that $\omega(x^*) = \overline{\omega(x)}$, $x \in A$. Since every $x \in A$ can be written uniquely in the form $x = x_1 + ix_2$ where x_1 and x_2 are self-adjoint, it suffices to show that $\omega(x)$ is real for any self-adjoint element $x \in A$. To prove this, fix $x = x^* \in A$, fix $t \in \mathbb{R}$, and consider the exponential

$$u_t = e^{itx} = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} x^n.$$

Notice that

$$\omega(u_t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \omega(x^n) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \omega(x)^n = e^{it\omega(x)}. \tag{2.4}$$

Note, too, that u_t is unitary. Indeed, by inspection of the exponential series (noting that t^n is real and x^n is self-adjoint for every $n \geq 0$), we have $u_t^* = e^{-itx}$, and hence $u_t^* u_t = e^{-itx} e^{itx} = e^0 = \mathbf{1}$ by Proposition 2.2.3. Similarly, $u_t u_t^* = \mathbf{1}$. It follows that $\|u_t\| = 1$, and thus $|\omega(u_t)| \leq \|\omega\| = 1$ for every $t \in \mathbb{R}$. Using formula (2.4) and the fact that $\Re(it\omega(x)) = -t\Im\omega(x)$, we find that

$$e^{-t\Im\omega(x)} = e^{\Re(it\omega(x))} = |e^{it\omega(x)}| = |\omega(u_t)| \leq 1, \qquad t \in \mathbb{R}.$$

Since $t \in \mathbb{R}$ is arbitrary, this implies that the imaginary part of $\omega(x)$ must vanish, proving that $\omega(x)$ is real.

This shows that the Gelfand map of A to $C(X)$ is self-adjoint in the sense that the Gelfand transform of x^* is the complex conjugate of the function $\hat{x}$, for every $x \in A$. It follows that $\{\hat{x} : x \in A\}$ is a self-adjoint subalgebra of $C(X)$ that separates points and contains the constant functions. The Stone–Weierstrass theorem implies that $\{\hat{x} : x \in A\}$ is norm-dense in $C(X)$.

We complete the proof by showing that the Gelfand map is isometric. We claim first that for $x \in A$, $\|x^2\| = \|x\|^2$. Indeed, using the formula $\|z^*z\| = \|z\|^2$ and the fact that x^* commutes with x we have

$$\left\|x^2\right\| = \left\|(x^2)^* x^2\right\|^{1/2} = \|x^* x x^* x\|^{1/2} = \left\|(x^* x)^2\right\|^{1/2} = \|x^* x\| = \|x\|^2.$$

Replacing x with x^2 gives $\|x^4\| = \|x\|^4$, and after further iteration

$$\left\|x^{2^n}\right\| = \|x\|^{2^n}, \qquad n = 1, 2, \dots.$$

The Gelfand–Mazur formula for the spectral radius (Theorem 1.7.3) implies

$$\|x\| = \|x^{2^n}\|^{1/2^n} = \lim_{n \to \infty} \|x^{2^n}\|^{1/2^n} = r(x),$$

while from Theorem 1.9.5, we have

$$r(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\} = \sup\{|\hat{x}(\omega)| : \omega \in \mathrm{sp}(A)\} = \|\hat{x}\|_\infty,$$

and hence the asserted formula $\|x\| = \|\hat{x}\|_\infty$. □

Corollary 1. Let A be a (perhaps noncommutative) unital C^*-algebra. Then the spectrum of any self-adjoint element x of A is real.

PROOF. Choose an element $x = x^*$ of A, and let B be the norm-closure of the set of all polynomials in x. Then B is a commutative C^*-subalgebra of A that contains the unit of A, hence $\sigma_A(x) \subseteq \sigma_B(x)$. On the other hand, Theorem 2.2.4 implies that $\omega(x)$ is real for every $\omega \in \operatorname{sp}(B)$, and hence $\operatorname{sp}_A(x) \subseteq \sigma_B(x) = \{\omega(x) : \omega \in \operatorname{sp}(B)\} \subseteq \mathbb{R}$. $\square$

The following result strengthens the spectral permanence theorem for the category of C^*-algebras:

COROLLARY 2. Let A be a unital C^*-algebra and let $B \subseteq A$ be a C^*-subalgebra of A that contains the unit of A. Then for every $x \in B$ we have $\sigma_B(x) = \sigma_A(x)$. In particular, for every self-adjoint $x \in A$,

$$\|x\| = r(x).$$

PROOF. We know that $\sigma_A(x) \subseteq \sigma_B(x)$ in general, and to prove the opposite inclusion it suffices to show that for any element $x \in B$ which is invertible in A one has $x^{-1} \in B$.

Fix such an x. Then x^*x is a self-adjoint element of B that is also invertible in A. By the preceding corollary, $\sigma_B(x^*x)$ is real. In particular, every point of $\sigma_B(x^*x)$ is a boundary point. By Theorem 1.11.3, $\sigma_B(x^*x) = \partial\sigma_B(x^*x) \subseteq \sigma_A(x^*x)$. Since $0 \notin \sigma_A(x^*x)$, $0 \notin \sigma_B(x^*x)$, and hence x^*x is invertible in B, equivalently, $(x^*x)^{-1} \in B$. Obviously, $(x^*x)^{-1}x^*$ is a left inverse of x; hence $x^{-1} = (x^*x)^{-1}x^*$ must belong to B.

The assertion that $\|x\| = r(x)$ follows after an application of Theorem 2.2.4 to the C^*-subalgebra of A generated by x and $\mathbf{1}$. $\square$

Thus we may compute the spectrum of a Hilbert space operator T relative to any C^*-algebra that contains T and the identity. In particular, we may restrict attention to the unital C^*-algebra generated by T. This is particularly useful in dealing with normal operators, since in those cases the generated C^*-algebra is commutative. We will pursue applications to normal operators in the following section.

Exercises.

(1) Prove the assertions made in Remark 2.2.2.
(2) Let A be a C^*-algebra.
   (a) Show that the involution in A satisfies $\|x^*\| = \|x\|$.
   (b) Show that if A contains a unit $\mathbf{1}$, then $\|\mathbf{1}\| = 1$.

   In the following exercises, X and Y denote compact Hausdorff spaces, and $\theta : C(X) \to C(Y)$ denotes an isomorphism of complex algebras. We do *not* assume continuity of θ:

(3) Let $p \in Y$. Show that there is a unique point $q \in X$ such that
$$\theta f(p) = f(q), \qquad f \in C(X).$$
(4) Show that there is a homeomorphism $\phi : Y \to X$ such that $\theta f = f \circ \phi$. Hint: Think in terms of the Gelfand spectrum.

(5) Conclude that θ is necessarily a self-adjoint linear map in the sense that $\theta(f^*) = \theta(f)^*$, $f \in C(X)$.
(6) Formulate and prove a theorem that characterizes *unital* algebra homomorphisms $\theta : C(X) \to C(Y)$ in terms of certain maps $\phi : Y \to X$. Which maps ϕ give rise to isomorphisms?

In the remaining exercises, let H be a Hilbert space and let $T \in \mathcal{B}(H)^{-1}$ be an invertible operator. Define $\theta : \mathcal{B}(H) \to \mathcal{B}(H)$ by

$$\theta(A) = TAT^{-1}, \qquad A \in \mathcal{B}(H).$$

(7) Show that θ is an automorphism of the Banach algebra structure of $\mathcal{B}(H)$.
(8) Show that the map $\theta : \mathcal{B}(H) \to \mathcal{B}(H)$ of the preceding exercise satisfies $\theta(A^*) = \theta(A)^*$ for all $A \in \mathcal{B}(H)$ if and only if T is a scalar multiple of a unitary operator.

2.3. Continuous Functions of Normal Operators

One can reinterpret Theorem 2.2.4 so as to provide a powerful functional calculus for normal operators. Sometimes this functional calculus is referred to as a weak form of the spectral theorem, or even as the spectral theorem itself; but that is a half-truth at best. The spectral theorem proper will be taken up in Section 2.4.

Throughout this section T will denote a normal operator on a Hilbert space H. The spectrum of T is a compact subset X of the complex plane, and by the Stone–Weierstrass theorem polynomials in z and $\bar{z}$ of the form

$$f(z) = \sum_{m,n=0}^{N} c_{mn} z^m \bar{z}^n, \qquad z \in X, \tag{2.5}$$

form a unital self-adjoint subalgebra of $C(X)$ that is norm-dense in $C(X)$. Given such a function f (or more properly, given the set of coefficients $\{c_{mn} : 0 \leq m, n \leq N\}$), one can write down a corresponding operator

$$f(T) = \sum_{m,n=0}^{N} c_{mn} T^m T^{*n}. \tag{2.6}$$

Notice that this much could have been done even if the operator T were not normal, since we have been explicit about the order of the factors T^m and T^{*n} on the right side of (2.6). However, for nonnormal operators $f \mapsto f(T)$ is not a well-defined map of functions on X to $\mathcal{B}(H)$, even for holomorphic polynomials $f(z) = a_0 + a_1 z + \cdots + a_N z^N$ (one can easily see why this is so by considering the case of nilpotent 2×2 matrices acting as operators on $\mathbb{C}^2$).

But for normal operators, we have:

THEOREM 2.3.1. *Let $T \in \mathcal{B}(H)$ be a normal operator with spectrum $X \subseteq \mathbb{C}$. Then the map that carries polynomials f of the form (2.5) to operators of the form $f(T)$ in (2.6) extends uniquely to an isometric $*$-isomorphism of $C(X)$ onto the C^*-algebra generated by T and $\mathbf{1}$.*

PROOF. Let $\mathcal{A}$ be the C^*-algebra generated by T and $\mathbf{1}$. We apply Theorem 2.2.4 to $\mathcal{A}$ as follows.

We claim first that the map $\omega \in \operatorname{sp}(\mathcal{A}) \mapsto \omega(T) \in \mathbb{C}$ is a homeomorphism of the Gelfand spectrum of $\mathcal{A}$ onto $X = \sigma(T)$. Indeed, this map is obviously a continuous map of $\operatorname{sp}(\mathcal{A})$ into $\mathbb{C}$, and it is injective because if ω_1 and ω_2 are two elements of $\operatorname{sp}(\mathcal{A})$ with $\omega_1(T) = \omega_2(T)$, then by Theorem 2.2.4

$$\omega_1(T^*) = \overline{\omega_1(T)} = \overline{\omega_2(T)} = \omega_2(T^*),$$

and hence ω_1 and ω_2 agree on the linear span of all products $T^m T^{*n}$, a dense subspace of $\mathcal{A}$. By compactness of $\operatorname{sp}(A)$, this map is a homeomorphism of $\operatorname{sp}(A)$ onto the spectrum of T relative to $\mathcal{A}$ which, by Corollary 2 of Theorem 2.2.4, is $X = \sigma(T)$.

These remarks identify $\operatorname{sp}(\mathcal{A})$ with X in such a way that the Gelfand map carries an operator of the form $f(T)$ in (2.6) to a polynomial $f \in C(X)$ of the form $f(z)$ in (2.5).

We conclude from Theorem 2.2.4 that the inverse of the Gelfand map defines an isometric $*$-isomorphism of $C(X)$ onto $\mathcal{A}$ that uniquely extends the map $f \mapsto f(T)$ described above. □

Exercises.

(1) Show that the spectrum of a normal operator $T \in \mathcal{B}(H)$ is connected if and only if the C^*-algebra generated by T and $\mathbf{1}$ contains no projections other than $\mathbf{0}$ and $\mathbf{1}$.

Consider the algebra $\mathcal{C}$ of all continuous functions $f : \mathbb{C} \to \mathbb{C}$. There is no natural norm on $\mathcal{C}$, but for every compact subset $X \subseteq \mathbb{C}$ there is a seminorm

$$\|f\|_X = \sup_{z \in X} |f(z)|.$$

$\mathcal{C}$ is a commutative $*$-algebra with unit.

(2) Given a normal operator $T \in \mathcal{B}(H)$, show that there is a natural extension of the functional calculus to a $*$-homomorphism $f \in \mathcal{C} \to f(T) \in \mathcal{B}(H)$ that satisfies $\|f(T)\| = \|f\|_{\sigma(T)}$.
(3) *Continuity of the functional calculus.* Fix a function $f \in \mathcal{C}$ and let $T_1, T_2, \dots$ be a sequence of normal operators that converges in norm to an operator T, $\lim_n \|T_n - T\| = 0$. Show that $f(T_n)$ converges in norm to $f(T)$.

2.4. The Spectral Theorem and Diagonalization

The spectral theorem is a generalization of the familiar theorem from linear algebra asserting that a self-adjoint $n \times n$ matrix A can be diagonalized; more precisely, there is a diagonal matrix D and a unitary matrix U such that $A = UDU^{-1}$. The diagonal components of D are the eigenvalues of A listed in some order, repeated according to their multiplicity. A similar diagonalization result is valid for normal $n \times n$ complex matrices.

In reading this section one should keep in mind not only the finite-dimensional case, or the infinite-dimensional case of self-adjoint operators having pure point spectrum, but also the case of operators having continuous spectrum and no eigenvalues at all, such as the operator X acting on L^2 of the unit interval $[0, 1]$ by

$$Xf(t) = tf(t), \qquad 0 \leq t \leq 1. \tag{2.7}$$

We assume that we are given a normal operator A acting on a *separable* infinite-dimensional Hilbert space H. There is an appropriate version of the spectral theorem for operators acting on inseparable spaces, which we describe briefly at the end of the section. However, we point out that operators acting on inseparable Hilbert spaces (in particular, normal ones) rarely arise in practice.

In order to properly formulate the spectral theorem we must generalize the notion of an orthonormal basis so as to accommodate "continuous" bases, and we must introduce a precise notion of "diagonalizable" operator relative to this generalized notion of basis.

Consider first the classical notion of orthonormal basis for H. This is a sequence $\mathcal{E} = \{e_1, e_2, \dots\}$ of mutually orthogonal unit vectors in H that have H as their closed linear span. Fixing such an $\mathcal{E}$ we can define a unitary operator $W : \ell^2 \to H$ as follows:

$$W\lambda = \lambda_1 e_1 + \lambda_2 e_2 + \cdots, \qquad \lambda \in \ell^2. \tag{2.8}$$

It is clear that every unitary operator $W : \ell^2 \to H$ arises in this way from a unique orthonormal basis $\mathcal{E}$ for H. We conclude that *specifying a particular orthonormal basis for H is the same as specifying a particular unitary operator from ℓ^2 to H.*

Continuing in this vein, suppose we are also given a normal operator $A \in \mathcal{B}(H)$ that has each of the given basis vectors as an eigenvalue:

$$Ae_k = a_k e_k, \qquad k = 1, 2, \dots. \tag{2.9}$$

It follows that the sequence of eigenvalues (a_k) belongs to ℓ^∞, and for the unitary operator $W : \ell^2 \to H$ of (2.8) we find that the transformed operator $B = W^{-1}AW \in \mathcal{B}(\ell^2)$ is a multiplication operator:

$$(B\lambda)_k = a_k \lambda_k, \qquad \lambda \in \ell^2, \quad k = 1, 2, \dots.$$

Thus, *an operator A acting on H is diagonalized by a given orthonormal basis if and only if the unitary operator associated with the basis implements an equivalence between A and a multiplication operator acting on ℓ^2.*

This notion of diagonalization is inadequate as it stands, since it involves only normal operators having pure point spectrum. However, it can be generalized in a natural way so as to include the possibility of continuous spectrum.

DEFINITION 2.4.1. An operator A acting on a separable Hilbert space H is said to be *diagonalizable* if there is a (necessarily separable) σ-finite measure space (X, μ), a function $f \in L^\infty(X, \mu)$, and a unitary operator $W : L^2(X, \mu) \to H$ such that $WM_f = AW$, M_f denoting multiplication by f:

$$(M_f\xi)(x) = f(x)\xi(x), \qquad x \in X, \quad \xi \in L^2(X, \mu).$$

Notice that a diagonalizable operator is necessarily normal, simply because multiplication operators are normal. Note, too, that the operator X of (2.7) is diagonalizable, since it is already a multiplication operator. Some more subtle examples are described in the exercises. The spectral theorem asserts that conversely, every normal operator is diagonalizable. We have broken the proof into a sequence of three simpler assertions.

LEMMA 2.4.2. *Let $A_1, A_2, \dots$ be a finite or infinite sequence of diagonalizable operators acting on respective Hilbert spaces $H_1, H_2, \dots$, satisfying $\sup_n \|A_n\| < \infty$. Then the direct sum $A_1 \oplus A_2 \oplus \cdots$ is a diagonalizable operator on $H_1 \oplus H_2 \oplus \cdots$.*

PROOF. This assertion follows from the fact that the countable direct sum of σ-finite measure spaces is a σ-finite measure space. In more detail, by hypothesis, we may find σ-finite measure spaces (X_n, μ_n), functions $f_n \in L^\infty(X_n, \mu_n)$, and unitary operators $W_n : L^2(X_n, \mu_n) \to H_n$, $n = 1, 2, \dots$ such that

$$W_n M_{f_n} = A_n W_n, \qquad n = 1, 2, \dots.$$

Since A_n is unitarily equivalent to M_{f_n}, our previous work with multiplication operators implies that the norm of $f_n \in L^\infty(X_n, \mu_n)$ satisfies $\|f_n\|_\infty = \|A_n\|$, hence

$$\sup_n \|f_n\|_\infty = \sup_n \|A_n\| < \infty.$$

Let $X = X_1 \sqcup X_2 \sqcup \cdots$ be the disjoint union of sets with the obvious σ-algebra of subsets and consider the measure μ defined on X by

$$\mu(E) = \mu_1(E \cap X_1) + \mu_2(E \cap X_2) + \cdots$$

for Borel sets $E \subseteq X$. The measure μ is σ-finite because each μ_n is. Moreover, there is a natural identification of $L^2(X, \mu)$ with the direct sum of L^2-spaces $L^2(X_1, \mu_1) \oplus L^2(X_2, \mu_2) \oplus \cdots$. Thus the direct sum of unitary operators $W = W_1 \oplus W_2 \oplus \cdots$ gives rise to a unitary operator from $L^2(X, \mu)$ to $H_1 \oplus H_2 \oplus \cdots$. The unique function $f : X \to \mathbb{C}$ satisfying

$f \restriction_{X_n} = f_n$ belongs to $L^\infty(X, \mu)$, it determines a bounded multiplication operator $M_f \in \mathcal{B}(L^2(X, \mu))$, and the unitary operator W intertwines M_f and $A_1 \oplus A_2 \oplus \cdots$. Hence $A_1 \oplus A_2 \oplus \cdots$ is diagonalizable. $\square$

LEMMA 2.4.3. *Let A be a bounded operator on a separable Hilbert space H and let $\mathcal{A}$ be the complex algebra generated by A, A^*, and the identity. Then there is a (finite or infinite) sequence of nonzero $\mathcal{A}$-invariant subspaces $H_1, H_2, \dots$ such that:*

(1) $H = H_1 \oplus H_2 \oplus \cdots$.
(2) *Each H_n contains a cyclic vector ξ_n for $\mathcal{A}$: $H_n = \overline{\mathcal{A}\xi_n}$, $n = 1, 2, \dots$.*

PROOF. This is a standard exhaustion argument. By Zorn's lemma we can find a family of mutually orthogonal nonzero subspaces $\{H_\alpha : \alpha \in I\}$ of H, each of which is $\mathcal{A}$-invariant, each containing a vector ξ_α such that H_α is spanned by $\mathcal{A}\xi_\alpha$, and that is maximal with respect to these properties. Since H is separable, the index set I must be finite or countable, and we can replace it with a subset of the positive integers if we wish.

It remains only to show that the spaces H_α span H. But if they did not then the orthocomplement K of $\sum_\alpha H_\alpha$ would be a nonzero $\mathcal{A}$-invariant subspace of H (note that since $\mathcal{A}$ is a self-adjoint set of operators, the orthocomplement of an $\mathcal{A}$-invariant subspace is $\mathcal{A}$-invariant). Picking any nonzero vector ξ in K we obtain a nonzero cyclic subspace $K_0 = \overline{\mathcal{A}\xi} \subseteq K$ that can be adjoined to the family $\{H_\alpha\}$ to contradict maximality. $\square$

The key step follows:

LEMMA 2.4.4. *Let A be a normal operator on a Hilbert space H and assume that the $*$-algebra generated by A and the identity has a cyclic vector. Then A is diagonalizable.*

PROOF. The cyclic vector hypothesis means that there is a vector $\xi \in H$ such that the set of vectors $\mathcal{A}\xi$ is dense in H, where $\mathcal{A}$ is the $*$-algebra generated by A and $\mathbf{1}$. Fix such a vector ξ and let $X \subseteq \mathbb{C}$ be the spectrum of A. We will show that there is a finite measure μ on X with the property that A is unitarily equivalent to the multiplication operator $M_\zeta \in \mathcal{B}(L^2(X, \mu))$, $\zeta(z) = z$ ($z \in X$) being the current variable function in $C(X) \subseteq L^\infty(X, \mu)$. Recalling that the functional calculus for normal operators provides a $*$-homomorphism $f \in C(X) \mapsto f(A) \in \mathcal{B}(H)$, we define a linear functional ρ on $C(X)$ by $\rho(f) = \langle f(A)\xi, \xi\rangle$. Since

$$\rho(|f|^2) = \rho(\bar{f}f) = \langle f(A)^* f(A)\xi, \xi\rangle = \|f(A)\xi\|^2 \geq 0,$$

ρ is a positive linear functional; hence the Riesz–Markov theorem provides a unique finite positive Borel measure μ on X such that

$$\int_X f(x)\,d\mu(x) = \langle f(A)\xi, \xi\rangle, \qquad f \in C(X).$$

If we consider $C(X)$ as a subspace of $L^2(X,\mu)$, then $C(X)$ is dense, and for $f, g \in C(X)$ we have

$$\begin{aligned}\langle f(A)\xi, g(A)\xi\rangle &= \langle g(A)^* f(A)\xi, \xi\rangle = \rho(\bar{g}f) \\ &= \int_X f(x)\bar{g}(x)\,d\mu(x) = \langle f, g\rangle_{L^2(X,\mu)}.\end{aligned}$$

Thus the map $f \in C(X) \mapsto f(A)\xi \in H$ is an isometry of the dense subspace $C(X) \subseteq L^2(X,\mu)$ onto the subspace $\{f(A)\xi : f \in C(X)\} \subseteq H$, which is dense in H because ξ is cyclic for the $*$-algebra generated by A and $\mathbf{1}$. The closure of this operator is a unitary operator $W : L^2(X,\mu) \to H$.

It remains to verify that for every $f \in C(X)$ we have $WM_f = f(A)W$ (the assertion of Lemma 2.4.4 being that this formula holds for $f(z) = z$, $z \in X$). For that, fix $f \in C(X)$. Since $C(X)$ is dense in $L^2(X,\mu)$ it is enough to check that

$$WM_f g = f(A)Wg, \qquad g \in C(X).$$

But for fixed g, $WM_f g = W(fg) = (fg)(A)\xi = f(A)g(A)\xi = f(A)Wg$. $\square$

Spectral Theorem 2.4.5. *Every normal operator acting on a separable Hilbert space is diagonalizable.*

Proof. Let $\mathcal{A}$ be the $*$-algebra generated by A and the identity. By Lemma 2.4.3 we can decompose H into a finite or countably infinite direct sum of nonzero subspaces $H_1 \oplus H_2 \oplus \cdots$ such that $\mathcal{A}H_k \subseteq H_k$ and the restriction of $\mathcal{A}$ to H_k has a cyclic vector, $k = 1, 2, \dots$. By Lemma 2.4.4, the restriction A_k of A to H_k is diagonalizable. Since the decomposition

$$A = A_1 \oplus A_2 \oplus \cdots$$

exhibits A as a uniformly bounded orthogonal direct sum of diagonalizable operators, Lemma 2.4.2 above implies that A is diagonalizable. $\square$

Remark 2.4.6. *Comments on inseparability.* If one insists on generalizing this form of the spectral theorem so as to include normal operators acting on inseparable Hilbert spaces, then it is possible to do so but some technical changes are necessary.

The definition of diagonalizable operator must be generalized so as to allow inseparable measure spaces that are not σ-finite. Thus one says that an operator $A \in \mathcal{B}(H)$ is diagonalizable if there is a positive measure space (X,μ), a function $f \in L^\infty(X,\mu)$, and a unitary operator $W : L^2(X,\mu) \to H$ such that $WM_f = AW$. One must replace Lemma 2.4.2 with the assertion that the direct sum of a uniformly bounded family $\{A_\alpha : \alpha \in I\}$ of diagonalizable operators is diagonalizable, where I is an index set of *arbitrary* cardinality. The proof of that result is similar to the one given, except that one has to construct uncountable direct sums of measure spaces. This requires some care but poses no substantial difficulties. No change is required for the key Lemma 2.4.4, but one must replace Lemma 2.4.3 with the assertion that every normal operator is a perhaps uncountable direct sum of normal operators having cyclic vectors. Once these preparations are made, the proof of the spectral theorem can be pushed through in general.

Exercises.

(1) Let X be a Borel space, let f be a bounded complex-valued Borel function defined on X, and let μ and ν be two σ-finite measures on X. The multiplication operator M_f defines bounded operators A on $L^2(X, \mu)$ and B on $L^2(X, \nu)$. Assuming that μ and ν are mutually absolutely continuous, show that there is a unitary operator $W : L^2(X, \mu) \to L^2(X, \nu)$ such that $WA = BW$. Hint: Use the Radon–Nikodym theorem.

(2) Show that every diagonalizable operator on a separable Hilbert space is unitarily equivalent to a multiplication operator M_f acting on $L^2(X, \mu)$ where (X, μ) is a probability space, that is, a measure space for which $\mu(X) = 1$.

The following exercises concern the self-adjoint operator A defined on the Hilbert space of bilateral sequences $H = \ell^2(\mathbb{Z})$ by

$$A\xi_n = \xi_{n+1} + \xi_{n-1}, \qquad n \in \mathbb{Z}, \quad \xi \in \ell^2(\mathbb{Z}).$$

(3) Show that A is diagonalizable by exhibiting an explicit unitary operator $W : L^2(\mathbb{T}, d\theta/2\pi) \to H$ for which $WM_f = AW$, where $f : \mathbb{T} \to \mathbb{R}$ is the function $f(e^{i\theta}) = 2\cos\theta$. Deduce that the spectrum of A is the interval $[-2, 2]$ and that the point spectrum of A is empty.

(4) Let U be the operator defined on $L^2(\mathbb{T}, d\theta/2\pi)$ by

$$Uf(e^{i\theta}) = f(e^{-i\theta}), \qquad 0 \leq \theta \leq 2\pi.$$

Show that U is a unitary operator on $L^2(\mathbb{T}, d\theta/2\pi)$ that satisfies $U^2 = \mathbf{1}$, and which commutes with A.

(5) Let $\mathcal{B}$ the the set of all operators on $L^2(\mathbb{T}, d\theta/2\pi)$ that have the form $M_f + M_g U$ where $f, g \in L^\infty(\mathbb{T}, d\theta/2\pi)$ and U is the unitary operator of the preceding exercise. Show that $\mathcal{B}$ is $*$-isomorphic to the C^*-algebra of all 2×2 matrices of functions $M_2(\mathcal{B}_0)$, where $\mathcal{B}_0$ is the abelian C^*-algebra $L^\infty(X, \mu)$, X being the upper half of the unit circle $X = \mathbb{T} \cap \{z = x + iy \in \mathbb{C} : y \geq 0\}$ and μ being the restriction of the measure $d\sigma = d\theta/2\pi$ to X.

The following exercises ask you to compare the operator A to a related operator B that acts on the Hilbert space $L^2([-2, 2], \nu)$, ν being Lebesgue measure on the interval $[-2, 2]$. The operator B is defined by

$$Bf(x) = xf(x), \qquad x \in [-2, 2], \quad f \in L^2([-2, 2], \nu).$$

(6) Show that B has spectrum $[-2, 2]$, that it has no point spectrum, and deduce that for every $f \in C[-2, 2]$ we have $\|f(A)\| = \|f(B)\|$.

(7) Show that A and B are *not* unitarily equivalent. Hint: What is the commutant of B?

(8) Show that A is unitarily equivalent to $B \oplus B$.

2.5. Representations of Banach $*$-Algebras

We now discuss some basic facts of representation theory that are best formulated in very general terms.

DEFINITION 2.5.1. A Banach $*$-algebra is a Banach algebra A that is endowed with an involution $x \mapsto x^*$ satisfying $\|x^*\| = \|x\|$, $x \in A$.

Every C^*-algebra is, of course, a Banach $*$-algebra; but we will see many examples of Banach $*$-algebras for which the C^*-condition $\|x^*x\| = \|x\|^2$ fails.

DEFINITION 2.5.2. A representation of a Banach $*$-algebra is a homomorphism $\pi : A \to \mathcal{B}(H)$ of A into the $*$-algebra of bounded operators on some Hilbert space satisfying $\pi(x^*) = \pi(x)^*$ for all $x \in A$.

Notice that we have not postulated that representations π are bounded, but merely that they are homomorphisms of the complex $*$-algebra structure. The set of all representations of A on a fixed Hilbert space H is denoted $\operatorname{rep}(A, H)$. The image $\pi(A)$ of A under a $*$-representation is a self-adjoint subalgebra of $\mathcal{B}(H)$ that may or may not be closed in the operator norm. A representation $\pi : A \to \mathcal{B}(H)$ is said to be *nondegenerate* if for every $\xi \in H$,

$$\pi(x)\xi = 0, \quad \forall x \in A \implies \xi = 0.$$

REMARK 2.5.3. A representation $\pi \in \operatorname{rep}(A, H)$ is *nondegenerate* iff $H = [\pi(A)H]$ is the closed linear span of the set of vectors

$$\pi(A)H = \{\pi(x)\xi : x \in A, \quad \xi \in H\}.$$

More generally, letting $N_\pi = \{\xi \in H : \pi(A)\xi = \{0\}\}$ be the null space of the operator algebra $\pi(A)$, H decomposes into an orthogonal direct sum of $\pi(A)$-invariant subspaces:

$$H = N_\pi \oplus [\pi(A)H].$$

See Exercise (1) below. The closed subspace $[\pi(A)H]$ is called the *essential space* of π.

Given two representations $\pi_k \in \operatorname{rep}(A, H_k)$, $k = 1, 2$, there is a natural notion of the *direct sum* of representations $\pi_1 \oplus \pi_2 \in \operatorname{rep}(A, H_1 \oplus H_2)$,

$$\pi_1 \oplus \pi_2(x) = \pi_1(x) \oplus \pi_2(x), \qquad x \in A.$$

A *subrepresentation* of a representation $\pi \in \operatorname{rep}(A, H)$ is a representation $\pi_0 \in \operatorname{rep}(A, H_0)$ obtained from π by restricting to a $\pi(A)$-invariant subspace $H_0 \subseteq H$ as follows:

$$\pi_0(x) = \pi(x) \restriction_{H_0} \in \mathcal{B}(H_0), \qquad x \in A.$$

Finally, two representations $\pi_k \in \operatorname{rep}(A, H_k)$, $k = 1, 2$, are said to be *unitarily equivalent* (or simply *equivalent*) if there is a unitary operator $W : H_1 \to H_2$ such that $W\pi_1(x)W^* = \pi_2(x)$ for every $x \in A$. It is clear that equivalent representations are indistinguishable from each other.

Thus we may paraphrase Remark 2.5.3 as follows: Every representation π of a Banach $*$-algebra on a Hilbert space is equivalent to the direct sum $\pi_e \oplus \pi_0$ of a nondegenerate representation π_e with the zero representation π_0 on some Hilbert space. Thus, the representation theory of Banach $*$-algebras reduces to the theory of nondegenerate representations.

PROPOSITION 2.5.4. *Every nonunital Banach $*$-algebra can be embedded as a maximal ideal of codimension* 1 *in a unital Banach $*$-algebra for which* $\|\mathbf{1}\| = 1$.

PROOF. Let A be a nonunital Banach $*$-algebra. The vector space $A \oplus \mathbb{C}$ can be made into a $*$-algebra $\tilde{A}$ by introducing the operations

$$(a, \lambda) \cdot (b, \mu) = (ab + \lambda b + \mu a, \lambda\mu), \qquad (a, \lambda)^* = (a^*, \bar{\lambda}).$$

The element $\mathbf{1} = (0, 1)$ is a unit for $\tilde{A}$, and we have $(a, \lambda) = a + \lambda\mathbf{1}$. Obviously, A is a maximal ideal of codimension 1 in $\tilde{A}$. $\tilde{A}$ becomes a Banach $*$-algebra by way of the norm $\|(a, \lambda)\| = \|a\| + |\lambda|$, with respect to which the inclusion map of A in $\tilde{A}$ is an isometric $*$-homomorphism. □

The following implies that representations of Banach $*$-algebras are necessarily bounded. There are many applications of this remarkable result.

THEOREM 2.5.5. *Let $\pi \in \operatorname{rep}(A, H)$ be a representation of a Banach $*$-algebra A on a Hilbert space H. Then $\|\pi\| \leq 1$.*

PROOF. By the preceding remarks, it suffices to consider the case in which π is nondegenerate.

We deal first with the case in which A has a unit $\mathbf{1}$. Because of nondegeneracy we have $\pi(\mathbf{1}) = \mathbf{1}$ (see Exercise (2), below). Notice that for every $a \in A$, $\sigma(\pi(a)) \subseteq \sigma(a)$. Indeed, if $\lambda \in \mathbb{C} \setminus \sigma(a)$, then $(a - \lambda)^{-1} \in A$, and since $\pi(\mathbf{1}) = \mathbf{1}$, $\pi((a - \lambda)^{-1})$ is the inverse of $\pi(a) - \lambda$. Hence $\lambda \in \mathbb{C} \setminus \sigma(\pi(a))$.

We show next that $\|\pi(a)\| \leq \|a\|$ for every $a \in A$. To see that, we use the C^*-property of the norm in $\mathcal{B}(H)$ to write

$$\|\pi(a)\|^2 = \|\pi(a)^*\pi(a)\| = \|\pi(a^*a)\|.$$

Since $\pi(a^*a)$ is a self-adjoint element of $\mathcal{B}(H)$, its norm agrees with its spectral radius, so that by the preceding paragraph,

$$\|\pi(a^*a)\| = r(\pi(a^*a)) \leq r(a^*a) \leq \|a^*a\| \leq \|a^*\|\|a\| = \|a\|^2.$$

Hence $\|\pi(a)\| \leq \|a\|$.

Suppose now that A has no unit, and let $\tilde{A}$ be its unital extension discussed in Proposition 2.5.4. The natural extension of π to $\tilde{A}$ is

$$\tilde{\pi}(a + \lambda\mathbf{1}) = \pi(a) + \lambda\mathbf{1},$$

and one readily verifies that $\tilde{\pi}$ is a representation of $\tilde{A}$ on H. By what was just proved we have $\|\tilde{\pi}\| \leq 1$, and since A is isometrically included in $\tilde{A}$, it follows that $\|\pi\| \leq 1$. $\square$

Exercises.

(1) Let $\mathcal{A} = \mathcal{A}^* \subseteq \mathcal{B}(H)$ be a self-adjoint algebra of operators on a Hilbert space, and let
$$N = \{\xi \in H : \mathcal{A}\xi = \{0\}\}$$
be the null space of $\mathcal{A}$. Show that the orthogonal complement of N is the closed linear span of $\mathcal{A}H = \{T\xi : T \in \mathcal{A}, \xi \in H\}$ and that both N and $[\mathcal{A}H]$ are $\mathcal{A}$-invariant subspaces.

(2) Let A be a Banach $*$-algebra with unit $\mathbf{1}$, and let $\pi \in \operatorname{rep}(A, H)$ be a representation of A. Show that π is nondegenerate iff $\pi(\mathbf{1}) = \mathbf{1}_H$.

(3) Let A be a Banach $*$-algebra. A representation $\pi \in \operatorname{rep}(A, H)$ is said to be *cyclic* if there is a vector $\xi \in H$ with the property that the set of vectors $\pi(A)\xi$ is dense in H. Show that a representation $\pi \in \operatorname{rep}(A, H)$ is nondegenerate iff it can be decomposed into a direct sum of cyclic subrepresentations in the following sense: There is a family $H_i \subseteq H$, $i \in I$, of nonzero subspaces of H that are mutually orthogonal, $\pi(A)$-invariant, that sum to H, and such that for each $i \in I$ there is a vector $\xi_i \in H_i$ with $\overline{\pi(A)\xi_i} = H_i$.

(4) Let A be a Banach $*$-algebra. A representation $\pi \in \operatorname{rep}(A, H)$ is said to be *irreducible* if the only closed $\pi(A)$-invariant subspaces of H are the trivial ones $\{0\}$ and H. Show that π is irreducible iff the commutant of $\pi(A)$ consists of scalar multiples of the identity operator.

(5) Let X be a compact Hausdorff space and let π be an irreducible representation of the C^*-algebra $C(X)$ on a Hilbert space H. Show that H is one-dimensional and there is a unique point $p \in X$ such that
$$\pi(f) = f(p)\mathbf{1}, \qquad f \in C(X).$$

2.6. Borel Functions of Normal Operators

Let N be a normal operator acting on a Hilbert space H with spectrum $X \subseteq \mathbb{C}$. We have discussed how to form continuous functions of N of the form $f(N)$, $f \in C(X)$. We now show how this functional calculus can be extended, in a more or less ultimate way, to bounded Borel functions.

Let X be a compact metrizable space. A complex-valued function defined on X is called a *Borel* function if it is measurable with respect to the Borel σ-algebra $\mathcal{B}$ of X, the σ-algebra of subsets of X generated by its topology. The space of all bounded complex-valued Borel functions on X is denoted $B(X)$; it is closed in the sup norm and is a unital commutative C^*-algebra relative to the pointwise operations and the natural involution $f^*(p) = \bar{f}(p)$, $p \in X$. Clearly $C(X) \subseteq B(X)$, but the difference between these two C^*-algebras is significant. Notice, for example, that while $C(X)$ is separable, $B(X)$ is typically inseparable; while $C(X)$ has nontrivial projections only when X fails to be connected, $B(X)$ is always generated by its projections.

We will show that every representation $\pi \in \operatorname{rep}(C(X), H)$ can be extended *in a particular way* to a representation $\tilde{\pi} \in \operatorname{rep}(B(X), H)$.

DEFINITION 2.6.1. A representation $\pi \in \operatorname{rep}(B(X), H)$ is called a σ-representation if it has the following property: For every uniformly bounded sequence $f_1, f_2, \dots \in B(X)$ which converges pointwise to zero in that

$$\lim_{n\to\infty} f_n(p) = 0, \qquad p \in X,$$

the sequence of operators $\pi(f_n)$ converges strongly to 0,

$$\lim_{n\to\infty} \|\pi(f_n)\xi\| = 0, \qquad \xi \in H.$$

REMARK 2.6.2. It is significant that because π is a representation, we can replace strong convergence in the definition above with weak convergence. To see that the two definitions are equivalent, suppose $\pi \in \operatorname{rep}(B(X), H)$ has the property that for every uniformly bounded sequence $f_1, f_2, \dots$ that converges pointwise to 0, $\pi(f_n)$ converges weakly to 0. We claim that π is a σ-representation. Indeed, for fixed $\xi \in H$ we have

$$\|\pi(f)\xi\|^2 = \langle \pi(f)\xi, \pi(f)\xi\rangle = \langle \pi(f)^*\pi(f)\xi, \xi\rangle = \langle \pi(f^*f)\xi, \xi\rangle. \tag{2.10}$$

If $f_1, f_2, \dots$ is a bounded sequence converging pointwise to 0, then $f_n^*f_n(p) = |f_n(p)|^2$, $p \in X$, is also a bounded sequence converging pointwise to 0, and hence $\pi(f_n^*f_n) \to 0$ weakly by hypothesis. The identity (2.10) implies that $\pi(f_n) \to 0$ strongly, as required.

THEOREM 2.6.3. *Let X be a compact metrizable space and let H be a Hilbert space. Every nondegenerate representation $\pi \in \operatorname{rep}(C(X), H)$ extends uniquely to a σ-representation $\tilde{\pi} \in \operatorname{rep}(B(X), H)$.*

PROOF. We deal first with uniqueness, and for that some notation will be useful. Let $\mathcal{B}$ be the σ-algebra of all Borel sets in X and let $M(X)$ be the Banach space of all complex-valued Borel measures $\mu : \mathcal{B} \to \mathbb{C}$. An element of $M(X)$ is a function $\mu : \mathcal{B} \to \mathbb{C}$ satisfying $\mu(\emptyset) = 0$, and for every sequence of mutually disjoint Borel sets $E_1, E_2, \dots$,

$$\mu(E_1 \cup E_2 \cup \cdots) = \sum_{n=1}^{\infty} \mu(E_n),$$

where the right side is interpreted as a convergent series of complex numbers. For every measure $\mu \in M(X)$ there is a smallest positive Borel measure $|\mu|$ satisfying

$$|\mu(S)| \le |\mu|(S), \qquad S \in \mathcal{B},$$

and the norm is given by $\|\mu\| = |\mu|(X) < \infty$.

Given a σ-representation $\tilde{\pi}$ that extends π, fix $\xi, \eta \in H$ and consider the set function $\mu_{\xi,\eta} : \mathcal{B} \to \mathbb{C}$ defined by

$$\mu_{\xi,\eta}(S) = \langle \tilde{\pi}(\chi_S)\xi, \eta\rangle, \qquad S \in \mathcal{B}.$$

It is clear that $\mu_{\xi,\eta}$ is a finitely additive measure because $\tilde{\pi}$ preserves the algebraic operations of multiplication and addition. We claim that, in fact, $\mu_{\xi,\eta}$ is countably additive. To see this, let $E_1, E_2, \dots$ be a sequence of mutually disjoint Borel sets with union $F = \cup_n E_n$. We have to show that

$$\mu_{\xi,\eta}(F) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu_{\xi,\eta}(E_k).$$

Letting $F_n = E_1 \cup \cdots \cup E_n$, we have

$$\mu_{\xi,\eta}(F) - \sum_{k=1}^{n} \mu_{\xi,\eta}(E_k) = \mu_{\xi,\eta}(F \setminus F_n) = \langle \tilde{\pi}(\chi_{F\setminus F_n})\xi, \eta\rangle.$$

Since the sequence of functions $f_n = \chi_{F\setminus F_n}$ is uniformly bounded and tends to zero pointwise, the right side of the preceding formula must tend to zero as $n \to \infty$ because $\tilde{\pi}$ is σ-representation.

We claim next that for every $f \in B(X)$ we have

$$\langle \tilde{\pi}(f)\xi, \eta\rangle = \int_X f\, d\mu_{\xi,\eta}. \tag{2.11}$$

Indeed, (2.11) is true when $f = \chi_E$ is a characteristic function by definition of $\mu_{\xi,\eta}$. By taking linear combinations it follows for simple functions f; it follows in general by an obvious limiting argument, since every function in $B(X)$ can be uniformly approximated by a sequence of simple functions (see Exercise (1) below).

To prove uniqueness, let $\tilde{\pi}$ and $\tilde{\pi}'$ be two σ-representations that extend the same representation π of $C(X)$. It suffices to show that for every $f \in B(X)$ and $\xi, \eta \in H$,

$$\langle \tilde{\pi}(f)\xi, \eta\rangle = \langle \tilde{\pi}'(f)\xi, \eta\rangle. \tag{2.12}$$

Notice that (2.12) holds for all $f \in C(X)$ because $\tilde{\pi}(f) = \tilde{\pi}'(f) = \pi(f)$ in that case. Consider the measure $\mu_{\xi,\eta}$ and its counterpart $\mu'_{\xi,\eta}$ for $\tilde{\pi}'$. Taking $f \in C(X)$, formulas (2.11) and (2.12) together imply that

$$\int_X f\, d\mu_{\xi,\eta} = \int_X f\, d\mu'_{\xi,\eta},$$

and hence $\mu_{\xi,\eta} = \mu'_{\xi,\eta}$ by the uniqueness assertion of the Riesz–Markov theorem on the representation of bounded linear functionals on $C(X)$ in terms of measures. Applying (2.11) we conclude that for all $g \in B(X)$,

$$\langle \tilde{\pi}(g)\xi, \eta\rangle = \int_X g\, d\mu_{\xi,\eta} = \int_X g\, d\mu'_{\xi,\eta} = \langle \tilde{\pi}'(g)\xi, \eta\rangle,$$

and uniqueness is proved.

Turning now to existence one simply reverses the argument as follows. Starting with $\pi \in \operatorname{rep}(C(X), H)$, fix a pair of vectors $\xi, \eta \in H$ and consider the linear functional

$$f \in C(X) \mapsto \langle \pi(f)\xi, \eta \rangle.$$

This is a bounded linear functional of norm at most $\|\xi\|\|\eta\|$. By the Riesz–Markov theorem there is a unique $\mu_{\xi,\eta} \in M(X)$ such that

$$\langle \pi(f)\xi, \eta \rangle = \int_X f \, d\mu_{\xi,\eta}, \qquad f \in C(X), \tag{2.13}$$

and moreover, $\|\mu_{\xi,\eta}\| \leq \|\xi\|\|\eta\|$. Notice, too, that the map $\xi, \eta \mapsto \mu_{\xi,\eta} \in M(X)$ is linear in ξ and antilinear in η.

Fix a function $f \in B(X)$. We define an operator $\tilde{\pi}(f) \in \mathcal{B}(H)$ by appealing to the Riesz lemma for sesquilinear forms as follows: Since

$$\xi, \eta \mapsto \int_X f \, d\mu_{\xi,\eta}$$

is a bounded sesquilinear form of norm at most $\|f\|\|\xi\|\|\eta\|$, there is a unique operator $\tilde{\pi}(f) \in \mathcal{B}(H)$ such that

$$\langle \tilde{\pi}(f)\xi, \eta \rangle = \int_X f \, d\mu_{\xi,\eta}, \qquad \xi, \eta \in H. \tag{2.14}$$

Obviously, the operator mapping $\tilde{\pi} : C(X) \to \mathcal{B}(H)$ is linear and satisfies $\|\tilde{\pi}(f)\| \leq \|f\|$, for $f \in B(X)$. It is also clear from the definition (2.13) of the measures $\mu_{\xi,\eta}$ and the defining formula (2.14) for $\tilde{\pi}$ that $\tilde{\pi}(f) = \pi(f)$ when $f \in C(X)$. A straightforward argument (which we omit) shows that $\tilde{\pi}$ carries real-valued functions to self-adjoint operators, and hence $\tilde{\pi}(f^*) = \tilde{\pi}(f)^*$, $f \in B(X)$.

Thus it remains to show that $\tilde{\pi}$ is multiplicative, $\tilde{\pi}(fg) = \tilde{\pi}(f)\tilde{\pi}(g)$, for $f, g \in B(X)$ and that it satisfies the continuity property of Definition 2.6.1.

To prove the multiplication property, note first that for every $\xi, \eta \in H$ and $g \in C(X)$ we have $g \cdot \mu_{\xi,\eta} = \mu_{\pi(g)\xi,\eta}$. Indeed, this follows from the fact that for every $f \in C(X)$,

$$\int_X fg \, d\mu_{\xi,\eta} = \langle \pi(fg)\xi, \eta \rangle = \langle \pi(f)\pi(g)\xi, \eta \rangle = \int_X f \, d\mu_{\pi(g)\xi,\eta}.$$

We claim next that for $F \in B(X)$, $F \cdot \mu_{\xi,\eta} = \mu_{\xi,\tilde{\pi}(F)^*\eta}$. This is a similar string of identities, where we note that for $g \in C(X)$ we have

$$\begin{aligned}\int_X g \, d(F \cdot \mu_{\xi,\eta}) = \int_X gF \, d\mu_{\xi,\eta} = \int_X F \, d\mu_{\pi(g)\xi,\eta} &= \langle \tilde{\pi}(F)\pi(g)\xi, \eta \rangle \\ = \langle \pi(g)\xi, \tilde{\pi}(F)^*\eta \rangle &= \int_X g \, d\mu_{\xi,\tilde{\pi}(F)^*\eta}.\end{aligned}$$

Finally, we claim that $\tilde{\pi}(FG) = \tilde{\pi}(F)\tilde{\pi}(G)$, for $F, G \in B(X)$. Indeed, fixing F and G and choosing $\xi, \eta \in H$ we have

$$\begin{aligned}\langle \tilde{\pi}(FG)\xi, \eta\rangle &= \int_X FG\, d\mu_{\xi,\eta} = \int_X G\, d\mu_{\xi,\tilde{\pi}(F)^*\eta} \\ &= \langle \tilde{\pi}(G)\xi, \tilde{\pi}(F)^*\eta\rangle = \langle \tilde{\pi}(F)\tilde{\pi}(G)\xi, \eta\rangle.\end{aligned}$$

The proof that $\tilde{\pi}$ is a σ-representation is a straightforward application of the bounded convergence theorem. Let $F_1, F_2, \dots$ be a uniformly bounded sequence in $B(X)$ converging pointwise to 0. For every ξ, η in H we have,

$$|\langle \tilde{\pi}(F_n)\xi, \eta\rangle| = \left|\int_X F_n\, d\mu_{\xi,\eta}\right| \le \int_X |F_n|\, d|\mu_{\xi,\eta}|$$

and the right side tends to 0 as $n \to \infty$ by the bounded convergence theorem, since $|\mu_{\xi,\eta}|$ is a finite positive measure on X and $|F_n|$ is a uniformly bounded sequence of functions tending pointwise to zero. In view of Remark 2.6.2, $\tilde{\pi}$ is a σ representation of $B(X)$. $\square$

Applying these results to a normal operator $N \in \mathcal{B}(H)$ we consider the continuous functional calculus $f \in C(\sigma(N)) \mapsto f(N)$. By Theorem 2.6.3 there is a unique σ-representation of the algebra $B(\sigma(N))$ that extends the original. This map is also written as if we were applying bounded Borel functions $f \in B(\sigma(N))$ to the operator to obtain $f(N)$. The properties of this Borel functional calculus will be exploited in the following section.

Exercises.

(1) Show that for every $f \in B(X)$ and every $\epsilon > 0$, there is a finite linear combination of characteristic functions in $B(X)$ (i.e., a simple function)

$$g = c_1\chi_{E_1} + c_2\chi_{E_2} + \cdots + c_n\chi_{E_n}$$

such that $\|f - g\| \le \epsilon$. Hint: Cover the range $f(X) \subseteq \mathbb{C}$ with a finely meshed grid and "pull back."

(2) Let $(X, \mathcal{B})$ be a Borel space. For every σ-finite measure μ on X let π_μ be the representation of $B(X)$ on $L^2(X,\mu)$ defined by

$$\pi_\mu(f)\xi(p) = f(p)\xi(p), \xi \in L^2(X,\mu).$$

(a) Show that π_μ is a σ-representation of $B(X)$ on $L^2(X,\mu)$. (Notice that the definition of σ-representation makes good sense in this more general context.)

(b) Given two σ-finite measures μ, ν on $(X, \mathcal{B})$, show that π_μ and π_ν are unitarily equivalent iff μ and ν are mutually absolutely continuous.

(c) Deduce that a multiplication operator acting on the L^2 space of a σ-finite measure is unitarily equivalent to a multiplication operator acting on the L^2 space of a finite measure space.

2.7. Spectral Measures

We have formulated the spectral theorem in terms of diagonalizing operators. In this section we present an equivalent formulation of the spectral theorem in terms of spectral measures. While this is the more classical form of the spectral theorem, it suffers from certain defects (mostly aesthetic) that are associated with the somewhat peculiar technology of spectral measures. In the defense of spectral measures we point out that they can provide a very effective tool for dealing with broader issues, such as the multiplicity theory of normal operators. And there are important results that are most clearly formulated in terms of spectral measures. Example: Stone's theorem, that makes the elegant assertion that a strongly continuous one-parameter group of unitary operators is the Fourier transform of a spectral measure on the real line.

Let us first revisit the idea of diagonalizing a normal matrix. Let N be a normal operator acting on a Hilbert space H of finite dimension n. There is an orthonormal basis $e_1, \dots, e_n$ for H consisting of eigenvalues of N,

$$Ne_k = \lambda_k e_k, \qquad k = 1, \dots, n,$$

where $\lambda_1, \dots, \lambda_n$ are complex numbers. There may be repetitions among the λ_k, but the set $\{\lambda_1, \dots, \lambda_n\}$ is exactly the spectrum of N.

This decomposition of H into eigenspaces can be reformulated in a basis-free way as follows. For every $\lambda \in \mathbb{C}$ let H_λ be the eigenspace

$$H_\lambda = \{\xi \in H : N\xi = \lambda\xi\}.$$

The subspaces $\{H_\lambda : \lambda \in \mathbb{C}\}$ are mutually orthogonal, they sum to H, each is invariant under both N and N^*, and H_λ is nonzero iff $\lambda \in \sigma(N)$. These observations can be converted into a structural statement about N as follows. Let E_λ be the projection onto H_λ. The E_λ form a system of mutually orthogonal projections in $\mathcal{B}(H)$, they sum to $\mathbf{1}$, $E_\lambda \neq 0 \iff \lambda \in \sigma(N)$, and we have

$$N = \sum_{\lambda \in \sigma(N)} \lambda \cdot E_\lambda. \tag{2.15}$$

Functions of N can be expressed in a similar way:

$$f(N) = \sum_{\lambda \in \sigma(N)} f(\lambda) \cdot E_\lambda.$$

What is peculiar here is that these sums have a multiplicative property that runs counter to the intuition of numerical sums,

$$\left(\sum_{\lambda \in \sigma(N)} f(\lambda) \cdot E_\lambda\right) \left(\sum_{\lambda \in \sigma(N)} g(\lambda) \cdot E_\lambda\right) = \sum_{\lambda \in \sigma(N)} f(\lambda)g(\lambda) \cdot E_\lambda,$$

a consequence of the fact that the E_λ are projections satisfying $E_\lambda E_\mu = 0$ for $\lambda \neq \mu$.

In any case, formula (2.15) expresses the operator N as a "spectral integral" in which the right side represents the integral of the complex-valued function $f(z) = z$, $z \in \sigma(N)$, against the projection-valued measure

$$E(S) = \sum_{\lambda \in S} E_\lambda, \qquad S \subseteq \mathbb{C}.$$

Despite its somewhat awkward appearance, the projection-valued function $\lambda \in \mathbb{C} \mapsto E_\lambda$ (or the projection-valued measure associated with it) contains critical information about the operator N. For example, $\sigma(N)$ is the set of points λ for which $E_\lambda \neq 0$. More significantly, the multiplicity $m(\lambda)$ of an eigenvalue $\lambda \in \sigma(N)$ is given by

$$m(\lambda) = \operatorname{rank} E_\lambda = \dim H_\lambda. \tag{2.16}$$

The function $m : \sigma(N) \to \mathbb{N}$ is called the *multiplicity function* of the normal operator N. It has these properties: $m(\lambda) > 0$ for every $\lambda \in \sigma(N)$, and

$$\sum_{\lambda \in \sigma(N)} m(\lambda) = \dim H.$$

Once one knows the spectrum and the multiplicity function of a normal operator N on a finite-dimensional Hilbert space, one knows N up to unitary equivalence (see Exercise (1) below). There is a natural generalization of this classification of normal operators to the infinite-dimensional case (see [**2**]), but we are not concerned with that here.

Our goal in this section is to point out how the formula (2.15) can be generalized to normal operators acting on infinite-dimensional Hilbert spaces by simply reformulating the results of the preceding section. Let $\mathcal{B}$ denote the σ-algebra of all Borel sets in $\mathbb{C}$. By a spectral measure (on $\mathbb{C}$) we mean a function $E \in \mathcal{B} \to P(E) \in \mathcal{B}(H)$ taking projections as values, such that $P(\emptyset) = 0$, $P(\mathbb{C}) = \mathbf{1}$, and for every sequence $E_1, E_2, \dots$ of mutually disjoint sets, we have

$$P(E_1 \cup E_2 \cup \cdots) = \sum_{n=1}^{\infty} P(E_n). \tag{2.17}$$

The sum on the right of (2.17) is interpreted as the limit in the strong operator topology of the sequence of partial sums $P(E_1) + \cdots + P(E_n)$. The fact that this limit exists is a consequence of the following observations.

PROPOSITION 2.7.1. *A spectral measure P has the following properties:*

(1) $E_1 \subseteq E_2 \implies P(E_1) \leq P(E_2)$.
(2) $E \cap F = \emptyset \implies P(E) \perp P(F)$.
(3) *For every* $E, F \in \mathcal{B}$, $P(E \cap F) = P(E)P(F)$.

PROOF. The first assertion follows from finite additivity of P, together with the decomposition $F = E \cup (F \setminus E)$ and the fact that $P(F \setminus E) \geq 0$.

For (2), we can write

$$\mathbf{1} = P(E \cup (\mathbb{C} \setminus E)) = P(E) + P(\mathbb{C} \setminus E).$$

Hence by (1), $P(F) \leq P(\mathbb{C} \setminus E) = \mathbf{1} - P(E)$, the latter being the projection onto $P(E)H^{\perp}$.

To deduce (3) from (2), one can write $P(E) = P(E \cap F) + P(E \setminus F)$, $P(F) = P(E \cap F) + P(F \setminus E)$, and observe that because of (2), $P(E \cap F)$, $P(E \setminus F)$, and $P(F \setminus E)$ are mutually orthogonal projections. □

These observations imply that the projections $P(E_1), P(E_2), \ldots$ appearing on the right of (2.17) are mutually orthogonal, so that the infinite sum has a clear meaning.

Starting now with a spectral measure $P : \mathcal{B} \to \mathcal{B}(H)$ and a bounded Borel function $f : \mathbb{C} \to \mathbb{C}$, we want to give meaning to the spectral integral $\int f \, dP$. This is done as follows. For every pair of vectors $\xi, \eta \in H$ we can define a complex-valued measure $\mu_{\xi,\eta}$ on $\mathbb{C}$ by $\mu_{\xi,\eta}(E) = \langle P(E)\xi, \eta \rangle$. Then $\mu_{\xi,\eta}$ is a countably additive complex-valued measure on $\mathcal{B}$ whose total variation is estimated as follows:

$$\|\mu_{\xi,\eta}\| \leq \|\xi\|\|\eta\|, \qquad \xi, \eta \in H.$$

Moreover, the map of $H \times H$ into the space of measures on $\mathbb{C}$ defined by $\xi, \eta \mapsto \mu_{\xi,\eta}$ is linear in ξ and antilinear in η. Thus we can define a bounded sesquilinear form $[\cdot, \cdot]$ on $H \times H$ by simple integration,

$$[\xi, \eta] = \int_{\mathbb{C}} f \, d\mu_{\xi,\eta},$$

and a straightforward estimate shows that

$$|[\xi, \eta]| \leq \sup_{z \in \mathbb{C}} |f(z)| \|\xi\|\|\eta\| = \|f\|_{\infty} \|\xi\|\|\eta\|.$$

By the Riesz lemma, there is a unique operator $\pi(f) \in \mathcal{B}(H)$ satisfying

$$\langle \pi(f)\xi, \eta \rangle = \int_{\mathbb{C}} f \, d\mu_{\xi,\eta}, \qquad \xi, \eta \in H,$$

and one has $\|\pi(f)\| \leq \|f\|_{\infty}$. This defines the operator $\pi(f)$ as a weak integral, and we can now interpret it as $\int f \, dP$.

More precisely, for every spectral measure P defined on $\mathbb{C}$ and taking values in the set of projections of $\mathcal{B}(H)$ and every bounded Borel function $f : \mathbb{C} \to \mathbb{C}$ there is a unique operator $\int f \, dP$ defined by

$$\left\langle \left(\int f \, dP \right) \xi, \eta \right\rangle = \int_{\mathbb{C}} f(z) \, \langle P(dz)\xi, \eta \rangle, \qquad \xi, \eta \in H.$$

We leave it for the reader to verify that $f \mapsto \int f \, dP$ is a σ-representation of the C^*-algebra $B(\mathbb{C})$ of all bounded Borel functions on $\mathbb{C}$, using the methods of the preceding section.

Spectral measures as we have discussed them are more general than required for the discussion of bounded normal operators. However, if a spectral measure P has *compact support* in the sense that there is a compact subset $K \subseteq \mathbb{C}$ with $P(\mathbb{C} \setminus K) = 0$, then P is associated with a bounded operator as follows. Since P is concentrated on K, the function $f(z) = z$ is bounded almost everywhere with respect to P, and hence

$$N = \int_{\mathbb{C}} z \, dP(z) = \int_K z \, dP(z)$$

defines a bounded normal operator with the property that

$$\int_{\mathbb{C}} f(z) \, dP(z) = f(N), \qquad f \in B(\mathbb{C}). \tag{2.18}$$

Thus, spectral integrals are simply another way of looking at the functional calculus for Borel functions.

Indeed, if we turn this around by starting with a bounded normal operator $N \in \mathcal{B}(H)$ and asking how to construct its spectral measure P, then the reply is simply to apply the characteristic functions of Borel sets to N according to the calculus of the preceding section:

$$P(E) = \chi_E(N), \qquad E \in \mathcal{B}.$$

Because $f \in B(\sigma(N)) \mapsto f(N)$ is a σ-representation extending the continuous functional calculus for N, P can be regarded as a spectral measure that is supported on $\sigma(N)$. Again, the preceding formula (2.18) simply provides a reinterpretation of the extended functional calculus as a spectral integral.

Exercises.

(1) Let $N_1 \in \mathcal{B}(H_1)$ and $N_2 \in \mathcal{B}(H_2)$ be two normal operators acting on finite-dimensional Hilbert spaces H_1, H_2. Show that there is a unitary operator $W : H_2 \to H_2$ such that $WN_1W^{-1} = N_2$ iff N_1 and N_2 have the same spectrum and the same multiplicity function.

(2) Calculate the spectral measure of the multiplication operator X defined on $L^2[0, 1]$ by $(X\xi)(t) = t\xi(t)$, $0 \leq t \leq 1$.

(3) A *resolution of the identity* is a function $\lambda \in \mathbb{R} \mapsto P_\lambda \in \mathcal{B}(H)$ from $\mathbb{R}$ to the projections on a Hilbert space with the following properties:
- $\lambda \leq \mu \implies P_\lambda \leq P_\mu$.
- Relative to the strong operator topology,
$$\lim_{\lambda \to -\infty} P_\lambda = 0, \qquad \lim_{\lambda \to +\infty} P_\lambda = \mathbf{1}.$$
- (Right continuity) For every $\lambda \in \mathbb{R}$,
$$\lim_{\mu \to \lambda+} P_\mu = P_\lambda.$$

Early formulations of the spectral theorem made extensive use of resolutions of the identity. It was gradually realized that these objects are equivalent to spectral measures, in much the same way that Stieltjes integrals are equivalent to integrals with respect to a measure. This exercise is related to the bijective correspondence that exists between resolutions of the identity and spectral measures on the real line.

(a) Consider the Borel space $(\mathbb{R}, \mathcal{B})$ of the real line. Given a spectral measure $E : \mathcal{B} \to \mathcal{B}(H)$, show that the function $P_\lambda = E((-\infty, \lambda])$, $\lambda \in \mathbb{R}$, is a resolution of the identity.
(b) Given two spectral measures $E, F : \mathcal{B} \to \mathcal{B}(H)$ that give rise to the same resolution of the identity, show that $E = F$.

2.8. Compact Operators

An operator A on a Hilbert space H is *compact* if the image of the unit ball $\{A\xi : \|\xi\| \leq 1\}$ is totally bounded. There is an enormous literature concerning classes of compact operators acting on Hilbert spaces. In this section we scratch the surface by discussing normal compact operators and Hilbert–Schmidt operators.

Compact normal operators can be diagonalized in the classical sense, in that there is an orthonormal basis consisting of eigenvectors. We base this on the following assertion about "approximate" eigenvectors.

PROPOSITION 2.8.1. *Let N be a normal operator acting on an infinite-dimensional Hilbert space H. For every accumulation point $\lambda \in \sigma(N)$ there is an orthonormal sequence $\xi_1, \xi_2, \dots$ in H such that*

$$\lim_{n\to\infty} \|N\xi_n - \lambda\xi_n\| = 0.$$

PROOF. By the Spectral Theorem we may assume that $H = L^2(X, \mu)$ has been coordinatized by a σ-finite measure space and that $N = M_f$ is multiplication by an L^∞ function. By Theorem 2.1.4 the spectrum of N is the essential range Λ of f.

Since λ is an accumulation point of Λ, we can find a sequence of distinct points $\lambda_n \in \Lambda$ that converges to λ. For each n choose $\epsilon_n > 0$ small enough that $\epsilon_n \to 0$ and the disks $D_n = \{z \in \mathbb{C} : |z - \lambda_n| < \epsilon_n\}$, $n = 1, 2, \dots$, are mutually disjoint. For each n the set $\{p \in X : f(p) \in D_n\}$ has positive measure because λ_n belongs to the essential range of f; and by σ-finiteness there is a subset $E_n \subseteq \{p \in X : f(p) \in D_n\}$ of finite positive measure, $n = 1, 2, \dots$. Considered as elements of $L^2(X, \mu)$, the characteristic functions $\chi_{E_1}, \chi_{E_2}, \dots$ are mutually orthogonal because the sets $E_1, E_2, \dots$ are mutually disjoint. Moreover,

$$|f - \lambda| \cdot \chi_{E_n} \leq (|f - \lambda_n| + |\lambda_n - \lambda|)\chi_{E_n} \leq (\epsilon_n + |\lambda_n - \lambda|)\chi_{E_n}.$$

It follows that

$$\|(N - \lambda)\chi_{E_n}\| \leq (\epsilon_n + |\lambda_n - \lambda|)\|\chi_{E_n}\|_2 = (\epsilon_n + |\lambda_n - \lambda|)\mu(E_n)^{1/2},$$

and the orthonormal sequence can be taken as $\xi_n = \mu(E_n)^{-1/2}\chi_{E_n}$, $n = 1, 2, \dots$. □

We obtain the following description of compact normal operators acting on infinite-dimensional separable Hilbert spaces.

THEOREM 2.8.2. *Let $N \in \mathcal{B}(H)$ be a compact normal operator. Then $0 \in \sigma(N)$, and $\sigma(N)$ is either finite or has the form $\{0, \lambda_1, \lambda_2, \dots\}$, where (λ_n) is a sequence of distinct complex numbers converging to 0. For each $\lambda \neq 0$ in $\sigma(N)$ the space $H_\lambda = \{\xi \in H : N\xi = \lambda\xi\}$ is nonzero and finite-dimensional.*

Let E_k be the projection onto H_{λ_k}. The E_k are mutually orthogonal and we have

$$N = \sum_{k=1}^{\infty} \lambda_k E_k,$$

the partial sums of the series converging in the operator norm to N. In particular, there is an orthonormal basis $e_1, e_2, \dots$ for H consisting of eigenvectors of N.

PROOF. A compact operator on H cannot be invertible; for if it were, then some open ball about 0 would be totally bounded, a clear absurdity as one sees by considering an orthogonal sequence of vectors having the same norm $r > 0$. Hence $0 \in \sigma(N)$.

We claim that $\sigma(N) \setminus \{0\}$ consists of isolated points. Indeed, for every accumulation point $\lambda \in \sigma(N)$, Proposition 2.8.1 implies that there is an orthonormal sequence $e_1, e_2, \dots$ satisfying $\|Ne_n - \lambda e_n\| \to 0$ as $n \to \infty$. Since N is compact, $\|Ne_k\| \to 0$ as $k \to \infty$ (see Exercise (1) below); hence

$$|\lambda| = \lim_{n\to\infty} \|\lambda e_n\| = \lim_{n\to\infty} \|Ne_n - \lambda e_n\| = 0.$$

It follows that $\sigma(N) \setminus \{0\}$ cannot contain accumulation points of $\sigma(N)$.

Thus $\sigma(N)$ is either finite or it consists of 0 together with a sequence $\lambda_1, \lambda_2, \dots$ of distinct isolated points converging to 0. Consider the case where $\sigma(N) = \{0, \lambda_1, \lambda_2, \dots\}$ is infinite. For each $n = 1, 2, \dots$, the characteristic function $u_n = \chi_{\{\lambda_n\}}$ belongs to $C(\sigma(N))$, and we can express the current variable $\zeta(z) = z$, $z \in \sigma(N)$, as an infinite series

$$\zeta = \sum_{k=1}^{\infty} \lambda_k u_k$$

converging *uniformly* in the norm of $C(\sigma(N))$; indeed, we have

$$\|\zeta - \sum_{k=1}^{n} \lambda_k u_k\|_\infty = \| \sum_{k=n+1}^{\infty} \lambda_k u_k\|_\infty = \sup_{k>n} |\lambda_k|,$$

which tends to 0 as $n \to \infty$. By the properties of the continuous functional calculus it follows that

$$\lim_{n\to\infty} \left\| N - \sum_{k=1}^{n} \lambda_k E_k \right\| = \limsup_{k\to\infty} |\lambda_k| = 0,$$

where E_k is the projection $E_k = u_k(N)$. Once one has such a series representation

$$N = \sum_{k=1}^{\infty} \lambda_k E_k$$

of N, one easily identifies the range of E_k as $\{\xi \in H : N\xi = \lambda_k \xi\}$. That completes the proof in the case where $\sigma(N)$ is infinite. The case of finite spectrum will be left for the reader. $\square$

Turning away from normal operators, let us fix an orthonormal basis $e_1, e_2, \dots$ for H. A *Hilbert–Schmidt operator* is an operator A on H with the property that

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 < \infty. \tag{2.19}$$

As we will see, Hilbert–Schmidt operators are not only bounded, but compact. They form an ideal $\mathcal{L}^2$ in the C^*-algebra $\mathcal{K}$ of all compact operators, and $\mathcal{L}^2$ is a Hilbert space in its own right.

Most (but not all) of the integral operators that we have encountered are Hilbert–Schmidt operators, and that is why the theory of Hilbert–Schmidt operators is important for approaching classical problems involving integral equations. While in this book we have concentrated on the idea of solving such equations, Hilbert–Schmidt operators enter into many aspects of operator theory and functional analysis, including the theory of Gaussian stochastic processes, representations of the canonical commutation and anticommutation relations of mathematical physics, and the theory of unitary representations of locally compact groups.

We first rephrase the definition of Hilbert–Schmidt operator so as to emphasize the role of the trace. Recall that an operator A on H is said to be positive if A is self-adjoint and has nonnegative spectrum. This is equivalent to the assertion $\langle A\xi, \xi\rangle \geq 0$ for every $\xi \in H$, as one can see in concrete terms by appealing to the spectral theorem and Exercise (5) below. It follows that the set $\mathcal{B}(H)^+$ of all positive operators on H is a cone, being closed under sums and multiplication by nonnegative scalars. For every positive operator A we can define an extended real number $\operatorname{trace} A \in [0, +\infty]$ as follows:

$$\operatorname{trace} A = \sum_{k=1}^{\infty} \langle Ae_k, e_k\rangle,$$

$e_1, e_2, \dots$ being an orthonormal basis for H, which for the moment we hold fixed as A varies. It is clear that

$$\begin{aligned} \operatorname{trace}(A+B) &= \operatorname{trace} A + \operatorname{trace} B, \\ \operatorname{trace}(\lambda A) &= \lambda \cdot \operatorname{trace} A \end{aligned} \tag{2.20}$$

for $A, B \in \mathcal{B}(H)^+$ and positive scalars λ, with the obvious conventions for handling sums and products of extended numbers in $(0, +\infty]$.

PROPOSITION 2.8.3. *The trace has the following properties:*

(1) $\operatorname{trace} A^*A = \operatorname{trace} AA^*$, *for any* $A \in \mathcal{B}(H)$.
(2) *For* $B \geq 0$ *and* U *unitary,* $\operatorname{trace} UBU^* = \operatorname{trace} B$.
(3) *The trace does not depend on the choice of basis* $\{e_k\}$.

PROOF. For (1), consider the double sequence of nonnegative terms $|\langle Ae_p, e_q\rangle|^2 = |\langle e_p, A^*e_q\rangle|^2$, $p, q = 1, 2, \dots$. Summing first on q and then on p, we obtain

$$\sum_{p=1}^{\infty}\sum_{q=1}^{\infty} |\langle Ae_p, e_q\rangle|^2 = \sum_{p=1}^{\infty} \|Ae_p\|^2,$$

while summing in the opposite order gives

$$\sum_{q=1}^{\infty}\sum_{p=1}^{\infty} |\langle e_p, A^*e_q\rangle|^2 = \sum_{q=1}^{\infty} \|A^*e_q\|^2.$$

Since the sum of a nonnegative double sequence is independent of the order of summation, this proves (1). Assertion (2) follows from it by setting $A = UB^{1/2}$ in (1), noting that $B = A^*A$ and $UBU^* = AA^*$.

To prove (3) let $f_1, f_2, \dots$ be another orthonormal basis and let U be the unique unitary operator on H satisfying $Ue_k = f_k$ for $k = 1, 2, \dots$. Then $f_k = U^*e_k$, and for every positive operator B, (2) implies

$$\sum_{k=1}^{\infty}\langle Bf_k, f_k\rangle = \sum_{k=1}^{\infty}\langle BU^*e_k, U^*e_k\rangle = \operatorname{trace} UBU^* = \operatorname{trace} B,$$

as asserted. □

By (2.20), the set of all positive operators with finite trace is a cone. By analogy with integration theory, we define $\mathcal{L}^1$ to be the linear space spanned by the positive operators having finite trace. Operators in $\mathcal{L}^1$ are called *trace class* operators. Every trace class operator can be written in the form

$$A = P_1 - P_2 + i(P_3 - P_4),$$

where P_k is positive and has finite trace. This decomposition is not unique, but the basic properties (2.20) imply that there is a unique linear functional defined on $\mathcal{L}^1$ by

$$\operatorname{trace} A = \operatorname{trace} P_1 - \operatorname{trace} P_2 + i(\operatorname{trace} P_3 - \operatorname{trace} P_4).$$

Obviously, for every $A \in \mathcal{L}^1$ and every orthonormal basis $e_1, e_2, \dots$ we have

$$\operatorname{trace} A = \sum_{n=1}^{\infty}\langle Ae_n, e_n\rangle,$$

where the series on the right is *absolutely* convergent. The value $\operatorname{trace} A$ of the sum does not depend on the choice of basis.

There is a natural norm on $\mathcal{L}^1$ that makes it into a Banach space (namely $\|A\|_{\mathcal{L}^1} = \operatorname{trace}|A|$), having many important operator-theoretic properties, and we refer the reader to [**19**] for a fuller development. What is important for us here is the relation between $\mathcal{L}^1$ and Hilbert–Schmidt operators, which we now describe.

According to (2.19), A is a Hilbert–Schmidt operator precisely when $\operatorname{trace} A^*A < \infty$, equivalently, when $A^*A \in \mathcal{L}^1$. The set of all Hilbert–Schmidt operators on H is denoted by $\mathcal{L}^2$. It is clear that $\mathcal{L}^2$ is closed under multiplication by scalars, and note that it is closed under addition as well. Indeed, for any two operators A, B we have the "parallelogram law"

$$(A+B)^*(A+B) + (A-B)^*(A-B) = 2A^*A + 2B^*B, \tag{2.21}$$

from which it follows that $0 \leq (A+B)^*(A+B) \leq 2A^*A + 2B^*B$. If both A and B belong to $\mathcal{L}^2$, then

$$\operatorname{trace}(A+B)^*(A+B) \leq 2 \operatorname{trace} A^*A + 2 \operatorname{trace} B^*B < \infty;$$

hence $A + B \in \mathcal{L}^2$.

Thus $\mathcal{L}^2$ is a complex vector space, which by Proposition 2.8.3 (1) is closed under the adjoint operation. That it is a left ideal is an obvious consequence of the defining property (2.19); and since $\mathcal{L}^2$ is self-adjoint, it must be a two-sided ideal.

The operator space $\mathcal{L}^2$ has a natural inner product, defined as follows. Corresponding to the polarization formula for sesquilinear forms on a complex vector space there is a polarization formula for bounded operators $A, B \in \mathcal{B}(H)$:

$$4B^*A = \sum_{k=0}^{3} i^k (A + i^k B)^*(A + i^k B). \tag{2.22}$$

The proof is a similar computation (see Exercise (2) below). If both A and B belong to $\mathcal{L}^2$, then each of the four terms on the right of (2.22) belongs to $\mathcal{L}^1$; hence so does B^*A, and we have

$$4 \operatorname{trace} B^*A = \sum_{k=0}^{3} i^k \operatorname{trace}(A + i^k B)^*(A + i^k B).$$

It follows that one can define an inner product on $\mathcal{L}^2$ as follows:

$$\langle A, B \rangle_2 = \operatorname{trace} B^*A, \qquad A, B \in \mathcal{L}^2. \tag{2.23}$$

It is significant that this inner product space is complete (see Exercise (3) below). $\mathcal{L}^2$ is therefore a Hilbert space.

PROPOSITION 2.8.4. *Every Hilbert–Schmidt operator A is compact, and satisfies $\|A\|^2 \leq \operatorname{trace} A^*A$.*

PROOF. We first prove the inequality $\|A\|^2 \leq \operatorname{trace} A^*A$. Indeed, for every unit vector e we can find an orthonormal basis $e_1, e_2, \dots$ starting with $e_1 = e$. Hence $\|Ae\|^2 \leq \sum_n \|Ae_n\|^2 = \operatorname{trace} A^*A$, and since e is arbitrary we obtain

$$\|A\|^2 = \sup_{\|e\|=1} \|Ae\|^2 \leq \operatorname{trace} A^*A.$$

To see that every Hilbert–Schmidt operator A is compact, fix an orthonormal basis $e_1, e_2, \dots$ and for every $n \geq 1$ let Q_n be the projection onto the subspace spanned by $e_{n+1}, e_{n+2}, \dots$. Obviously, $F_n = A(\mathbf{1} - Q_n)$ is a finite-rank operator, and by the preceding paragraph we have

$$\|A - F_n\|^2 = \|AQ_n\|^2 \leq \operatorname{trace}(Q_n A^* A Q_n) = \sum_{k=n+1}^{\infty} \|Ae_k\|^2.$$

The right side tends to 0 as $n \to \infty$ because $\sum_k \|Ae_k\|^2 < \infty$. Hence $A = \lim_n F_n$ is the norm limit of a sequence of finite-rank operators, and is therefore compact. □

Example 2.8.5. *Hilbert–Schmidt integral operators.* Let (X, μ) be a (separable) σ-finite measure space and let $k \in L^2(X \times X, \mu \times \mu)$ be a square-integrable function of two variables on X. We want to define an integral operator A on $L^2(X, \mu)$ by way of

$$A\xi(x) = \int_X k(x, y)\xi(y)\, d\mu(y), \qquad \xi \in L^2(X, \mu), \tag{2.24}$$

but there are several things that have to be checked.

In the first place, since

$$\int_{X \times X} |k(x, y)|^2\, d\mu(x) d\mu(y) < \infty,$$

the Fubini theorem implies that for almost every $x \in X$ $(d\mu)$ the section $k(x, \cdot)$ belongs to $L^2(X, d\mu)$, and for such x the function $y \mapsto k(x, y)\xi(y)$ belongs to $L^1(X, \mu)$. This implies that the integral in (2.24) is well defined for almost every x, and writing its value as $A\xi(x)$, we have the estimate

$$|A\xi(x)| \leq \int_X |k(x, y)||\xi(y)|\, d\mu(y).$$

Moreover, another application of Fubini's theorem implies that for every $\eta \in L^2(X, \mu)$ we have

$$\int_X |A\xi(x)||\eta(x)|\, d\mu(x) \leq \int_{X \times X} |k(x, y)||\eta(x)||\xi(y)|\, d\mu(x) d\mu(y),$$

which by the Schwarz inequality is dominated by

$$\|k\| \left(\int_{X \times X} |\eta(x)|^2 |\xi(y)|^2\, d\mu(x) d\mu(y) \right)^{1/2} = \|k\| \|\xi\|_2 \|\eta\|_2,$$

where $\|k\|$ denotes the norm of k as an element of $L^2(X \times X, \mu \times \mu)$.

It follows that formula (2.24) defines a linear operator A on $L^2(X, \mu)$ satisfying $|\langle A\xi, \eta\rangle| \leq \|k\| \|\xi\|_2 \|\eta\|_2$ for every $\xi, \eta \in L^2(X, \mu)$, and hence $\|A\| \leq \|k\|$.

Let us now calculate trace A^*A. Choose an orthonormal basis $e_1, e_2, \dots$ for $L^1(X, d\mu)$. For every $m, n = 1, 2, \dots$, we have

$$\begin{aligned}\langle Ae_m, e_n\rangle &= \int_X Ae_m(x)\bar{e}_n(x)\,d\mu(x)\\ &= \int_{X\times X} k(x,y)\bar{e}_n(x)e_m(y)\,d\mu(y)d\mu(x).\end{aligned}$$

Writing $u_{mn}(x,y) = e_n(x)\bar{e}_m(y)$, we find that $\{u_{mn} : m, n = 1, 2, \dots\}$ is an orthonormal basis for $L^2(X \times X, \mu \times \mu)$, and the preceding formula becomes

$$\langle Ae_m, e_n\rangle = \langle k, u_{mn}\rangle,$$

the inner product on the right being that of $L^2(X \times X, \mu \times \mu)$. It follows that trace A^*A is given by

$$\sum_{m=0}^{\infty} \|Ae_m\|^2 = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} |\langle Ae_m, e_n\rangle|^2 = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} |\langle k, u_{mn}\rangle|^2 = \|k\|^2.$$

We summarize the results of this discussion as follows:

PROPOSITION 2.8.6. *Let (X, μ) be a separable σ-finite measure space. For every function $k \in L^2(X \times X, \mu \times \mu)$ there is a unique bounded operator A_k on $L^2(X, \mu)$ satisfying*

$$A_k\xi(x) = \int_X k(x,y)\xi(y)\,d\mu(y), \qquad \xi \in L^2(X, \mu).$$

The map $k \mapsto A_k$ is an isometric isomorphism of the Hilbert space $L^2(X \times X, \mu \times \mu)$ onto the Hilbert space $\mathcal{L}^2$ of all Hilbert–Schmidt operators on $L^2(X, \mu)$.

Exercises.

(1) Let A be a compact operator on a Hilbert space H. Show that for every sequence of mutually orthogonal unit vectors $\xi_1, \xi_2, \dots \in H$ we have
$$\lim_{n\to\infty} \|A\xi_n\| = 0.$$
Hint: Consider the decreasing sequence of projections P_n defined by the decreasing sequence of closed subspaces $[\xi_n, \xi_{n+1}, \xi_{n+2}, \dots]$, $n = 1, 2, \dots$.

(2) Let $e_1, e_2, \dots$ be an orthonormal basis for a Hilbert space H and let $A \in \mathcal{B}(H)$. Show that A is compact iff
$$\lim_{n\to\infty} \|(\mathbf{1} - E_n)A(\mathbf{1} - E_n)\| = 0,$$
where E_n denotes the projection onto span$\{e_1, \dots, e_n\}$.

(3) Verify the polarization formula for bounded operators on a Hilbert space H:
$$4B^*A = \sum_{k=0}^{3} i^k \left(A + i^kB\right)^* \left(A + i^kB\right).$$

(4) Let $\|A\|_2 = \langle A, A\rangle_2^{1/2}$ for every Hilbert–Schmidt operator A.
  (a) Let $A_1, A_2, \dots$ be a sequence in $\mathcal{L}^2$ that satisfies
$$\lim_{m,n\to\infty} \|A_m - A_n\|_2 = 0.$$
  Show that there is an operator $A \in \mathcal{B}(H)$ such that $\|A_n - A\| \to 0$ as $n \to \infty$.
  (b) Show that $\mathcal{L}^2$ is a Hilbert space relative to the inner product (2.23).

(5) Show that a multiplication operator M_f is self-adjoint and has non-negative spectrum iff $\langle M_f\xi, \xi\rangle \geq 0$ for every $\xi \in L^2(X, \mu)$.

2.9. Adjoining a Unit to a C^*-Algebra

We have discussed the procedure of adjoining a unit to a nonunital Banach algebra so as to obtain a unital one. Proposition 2.5.4 describes the corresponding procedure for the category of Banach $*$-algebras. If one applies the latter to a nonunital C^*-algebra such as the compact operators $\mathcal{K} \subseteq \mathcal{B}(H)$, the result is a unital Banach $*$-algebra, but its norm fails to satisfy the C^* condition $\|x^*x\| = \|x\|^2$. Fortunately, one can always renorm this unitalization so that it becomes a C^*-algebra, without changing the norm on the ideal representing the original algebra, in a unique way. The details are as follows.

Let A be a C^*-algebra without unit and let $L : A \to \mathcal{B}(A)$ be the left regular representation of A, in which L_x represents left multiplication by x, $x \in A$. For any Banach algebra, L is a homomorphism of the algebra structure of A such that $\|L_x\| \leq \|x\|$ for every $x \in A$. Let A^e denote the set of operators on A given by
$$A^e = \{L_a + \lambda\mathbf{1} : a \in A, \quad \lambda \in \mathbb{C}\}.$$
Then A^e is a complex algebra with unit, and we may define an involution in A^e by $(L_a + \lambda\mathbf{1})^* = L_{a^*} + \bar\lambda\mathbf{1}$, $a \in A$, $\lambda \in \mathbb{C}$. The operator norm determines a norm on A^e, which makes it into a normed algebra. Moreover, the natural map $\pi : A \to A^e$ defined by $\pi(a) = L_a$ is a $*$-homomorphism satisfying $\pi(a) = 0 \implies a = 0$, $a \in A$. We will show that (a) there is a C^*-algebra norm on A^e and (b) with respect to that norm, π is an isometry.

Remark 2.9.1. Suppose one is given a Banach algebra A that is also endowed with an involution $*$ satisfying $\|x^*x\| \geq \|x\|^2$ for all $x \in A$. Then A is a C^*-algebra: $\|x\|^2 = \|x^*x\|$, $x \in A$. To see this, note that the given inequality implies that $\|x\|^2 \leq \|x^*x\| \leq \|x^*\| \cdot \|x\|$, so that $\|x\| \leq \|x^*\|$ for all $x \in A$. By replacing x with x^* we obtain the opposite inequality; hence $\|x\| = \|x^*\|$. It follows that $\|x^*x\| \leq \|x^*\| \cdot \|x\| = \|x\|^2$, providing the other half of the asserted equality.

Proposition 2.9.2. *The involution in A^e satisfies $\|X^*X\| = \|X\|^2$ for every $X \in A^e$, and A^e is closed in the operator norm of $\mathcal{B}(A)$; hence it is* *a unital C^*-algebra. Moreover, the regular representation is an isometric $*$-isomorphism of A onto a maximal ideal of codimension one in A^e.*

PROOF. Notice first that $\|L_a\| = \|a\|$ for every $a \in A$. Indeed, $\leq$ is true for any Banach algebra, and the opposite inequality follows for an element a of norm 1 because

$$\|L_a\| \geq \|L_a(a^*)\| = \|aa^*\| = \|a^*\|^2 = \|a\|^2 = 1.$$

The set $\{L_a : a \in A\}$ is obviously an ideal in A^e of codimension at most one. If the codimension were zero, then the identity operator would have the form L_f for some element $f \in A$; that would imply f was a unit for A, contrary to hypothesis. Hence $\{L_a : a \in A\}$ has codimension one. Since L is an isometry, this ideal must be closed in the operator norm of $\mathcal{B}(A)$; and since A^e is obtained from this ideal by adjoining the one-dimensional space spanned by $\mathbf{1}$, it follows that A^e must also be norm closed.

It remains to show that the involution in A^e satisfies $\|X^*X\| = \|X\|^2$. By Remark 2.9.1, it is enough to verify the inequality $\|X\|^2 \leq \|X^*X\|$ for $X = L_a + \lambda\mathbf{1}$ in A^e. For such an X, we have

$$\begin{aligned}
\|X\|^2 &= \sup_{\|b\|\leq 1} \|(L_a + \lambda\mathbf{1})(b)\|^2 = \sup_{\|b\|\leq 1} \|ab + \lambda b\|^2 \\
&= \sup_{\|b\|\leq 1} \|(ab + \lambda b)^*(ab + \lambda b)\| = \sup_{\|b\|\leq 1} \|b^*(X^*X(b))\| \\
&\leq \sup_{\|b\|\leq 1} \|X^*X(b)\| \leq \|X^*X\|.
\end{aligned}$$

□

The following result asserts that C^*-algebras have a remarkable property of rigidity that is not shared by other types of Banach $*$-algebras.

PROPOSITION 2.9.3. *Every $*$-homomorphism $\pi : A \to B$ of C^*-algebras has norm at most 1. If π has trivial kernel, then it is an isometry.*

PROOF. Suppose first that A has a unit $\mathbf{1}_A$. By passing from B to the closure of the $*$-subalgebra $\pi(A)$ if necessary, we may assume that $\pi(A)$ is dense in B. In this case, $\pi(\mathbf{1}_A)$ is the unit $\mathbf{1}_B$ of B. Thus we may argue as we did for nondegenerate representations. For example, since π must map invertible elements of A to invertible elements of B, it follows that $\sigma(\pi(x)) \subseteq \sigma(x)$ for every element $x \in A$. Corollary 2 of Theorem 2.2.4 implies that for self-adjoint elements $x \in A$ we have

$$\|\pi(x)\| = r(\pi(x)) \leq r(x) = \|x\|,$$

so that for general elements $z \in A$ we have

$$\|\pi(z)\|^2 = \|\pi(z)^*\pi(x)\| = \|\pi(z^*z)\| \leq \|z^*z\| = \|z\|^2.$$

If, in addition, π has trivial kernel, then we claim that $\|\pi(x)\| = \|x\|$ for every $x \in A$. As above, this reduces to the case where $x = x^*$ is self-adjoint; and by Corollary 2 of Theorem 2.2.4 it is enough to show that x

and $\pi(x)$ have the same spectrum when $x = x^*$. We have already seen that $\sigma(\pi(x)) \subseteq \sigma(x)$. For the opposite inclusion, suppose that λ is a point of $\sigma(x)$ that does not belong to $\sigma(\pi(x))$. There is a continuous function $f : \sigma(x) \to \mathbb{R}$ such that f vanishes on $\sigma(\pi(x))$ and $f(\lambda) \neq 0$. Since $f = 0$ on $\sigma(\pi(x))$, we must have $f(\pi(x)) = 0$. Notice that $f(\pi(x)) = \pi(f(x))$ (this is obvious if f is a polynomial, and it follows for general continuous f by an application of the Weierstrass approximation theorem and the previously established fact that π is a bounded linear map of A to B). But $\pi(f(x)) = 0$ implies that $f(x) = 0$ because π has trivial kernel; in turn, $f(x) = 0$ implies that $f = 0$ on $\sigma(x)$, contradicting the fact that $f(\lambda) \neq 0$.

Now assume that A has no unit and let A^e be the unital extension of A, identifying A with its image in A^e. By adjoining a unit to B if necessary, we may assume that B has a unit $\mathbf{1}_B$. One may verify directly that the map $\tilde{\pi} : A^e \to B$ defined by

$$\tilde{\pi}(a + \lambda \mathbf{1}) = \pi(a) + \lambda \mathbf{1}_B$$

is a $*$-homomorphism carrying the unit of A^e to $\mathbf{1}_B$. The argument above implies that $\|\pi\| \leq \|\tilde{\pi}\| \leq 1$. Finally, assuming that π is one-to-one, we claim that $\tilde{\pi}$ is one-to-one. For if $a \in A$ and $\lambda \neq 0$ is a scalar for which $\pi(a) + \lambda \mathbf{1}_B = \tilde{\pi}(a + \lambda \mathbf{1}) = 0$, set $f = -\lambda^{-1} a \in A$. Since $\pi(f) = \mathbf{1}_B$, $\pi(f)$ is a unit for $\pi(A)$, and since π has trivial kernel, f must be a unit for A, contrary to hypothesis. Thus $\pi(a) + \lambda \mathbf{1}_B = 0 \implies \lambda = 0$ and $a = 0$, and thus $\tilde{\pi}$ is one-to-one as asserted. The preceding paragraphs imply that $\tilde{\pi}$ is isometric; hence π is isometric. □

COROLLARY 1. Let A be a complex algebra with involution. If there is a norm on A that makes it into a C^*-algebra, then that norm is unique.

PROOF. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two (complete) Banach algebra norms on A satisfying $\|x^*x\|_k = \|x\|_k^2$ for $x \in A$, and let A_k be the algebra A considered as a C^*-algebra in each norm respectively, $k = 1, 2$. The identity map of A can be regarded as a $*$-isomorphism of A_1 onto A_2. By Proposition 2.9.3 this map must be an isometry; hence $\|x\|_1 = \|x\|_2$ for all $x \in A$. □

COROLLARY 2. Let A be a nonunital C^*-algebra, let $\pi : A \to A^e$ be the natural map of A into its unitalization, and endow A^e with its C^*-norm. Then π is an isometric $*$-isomorphism of A onto an ideal of codimension 1 in A^e.

Exercises.

(1) Let A be a nonunital C^*-algebra and let $\pi : A \to A^e$ be the natural map of A into its unitalization. Considering A^e as a C^*-algebra, suppose that there is an isometric $*$-homomorphism $\sigma : A \to B$ of A into another unital C^*-algebra B such that $\sigma(A)$ is an ideal of codimension 1 in B. Show that there is a unique isometric $*$-isomorphism $\theta : A^e \to B$ such that $\theta \circ \pi = \sigma$.

(2) Let $\mathcal{K}$ be the C^*-algebra of compact operators on a Hilbert space H. Show that the space of operators $\{\lambda\mathbf{1} + K : \lambda \in \mathbb{C}, K \in \mathcal{K}\}$ is a C^*-algebra $*$-isomorphic to $\mathcal{K}^e$.

(3) Let X be a compact Hausdorff space and let F be a proper closed subset of X. Let A be the ideal of all functions $f \in C(X)$ that vanish throughout F, $f(p) = 0$, $p \in F$. Note that A is a C^*-algebra in its own right.
   (a) Show that A has a unit if and only if F is both closed and open.
   (b) Assuming that F is not open, identify the unitalization of A in concrete terms by exhibiting a compact Hausdorff space Y such that $A^e \cong C(Y)$, describing the precise relationship of Y to X and F.

2.10. Quotients of C^*-Algebras

In order to discuss compact perturbations of operators on a Hilbert space one must bring in the Calkin algebra (the C^*-algebra $\mathcal{B}(H)/\mathcal{K}$ obtained by passing to the quotient modulo compact operators), and that requires some basic results about the formation of quotients of C^*-algebras. We work out the relevant material in this section, in a general setting.

Throughout, A will denote a C^*-algebra that need *not* contain a unit. When no unit is present there is an effective substitute, called an *approximate unit.* More precisely, an approximate unit for A is a net $\{e_\lambda : \lambda \in I\}$ indexed by an increasing directed set I (which need not be the positive integers $\mathbb{N}$ and which need not even be countable) that has the following properties:

(1) $e_\lambda = e_\lambda^*$ and $\sigma(e_\lambda) \subseteq [0, 1]$.
(2) $\lim_{\lambda\to\infty} \|xe_\lambda - x\| = 0$, for every $x \in A$.

The meaning of the second assertion of (1) requires clarification, since our discussion of spectra has so far been limited to unital Banach algebras and unital C^*-algebras. The spectrum of an element x of a nonunital C^*-algebra A is defined by embedding A in its unitilization A^e; $\sigma(x)$ is then well defined by considering x to be an element of A^e. The spectrum of an element of a nonunital C^*-algebra is a compact set of complex numbers which necessarily contains 0.

Significantly, approximate units exist in arbitrary C^*-algebras (see Theorem 1.8.2 of [**2**], for example); but all we require here is the following:

LEMMA 2.10.1. *Let A be a C^*-algebra and let J be a closed left ideal in A. For every element $x \in J$ there is a sequence $e_1, e_2, \dots$ of self-adjoint elements of J such that $\sigma(e_n) \subseteq [0, 1]$ and*

$$\lim_{n\to\infty} \|xe_n - x\| = 0.$$

PROOF. By adjoining a unit to A if necessary, we can assume that A is unital. Suppose first that the given element x is self-adjoint, and define

$$e_n = nx^2(\mathbf{1} + nx^2)^{-1} = f_n(x), \qquad n = 1, 2, \ldots,$$

f_n being the real function

$$f_n(t) = \frac{nt^2}{1 + nt^2}, \qquad t \in \mathbb{R}.$$

Since f_n is continuous and vanishes at the origin, e_n belongs to the closed linear span of the positive powers of x, hence $e_n \in J$. Moreover, since $0 \leq f_n(t) \leq 1$ for all $t \in \mathbb{R}$, we have $\sigma(e_n) \subseteq [0, 1]$.

Writing

$$\|xe_n - x\|^2 = \|x(\mathbf{1} - e_n)\|^2 = \|(\mathbf{1} - e_n)x^2(\mathbf{1} - e_n)\| \leq \|x^2(\mathbf{1} - e_n)\|,$$

and using the fact that $1 - f_n(t) = 1/(1 + nt^2)$, we find that

$$x^2(\mathbf{1} - e_n) = x^2(\mathbf{1} + nx^2)^{-1} = \frac{1}{n}nx^2(\mathbf{1} + nx^2)^{-1}$$

has norm at most $1/n$. Thus $\|x^2(\mathbf{1} - e_n)\| \to 0$ as $n \to \infty$, and (2) is proved for the case $x^* = x$.

In the general case, we apply the preceding paragraph to the self-adjoint element $x^*x \in J$ to find a sequence of self-adjoint elements $e_n \in J$ satisfying $\sigma(e_n) \subseteq [0, 1]$, for which $\|x^*x - x^*xe_n\| \to 0$ as $n \to \infty$. In this case we have

$$\|x - xe_n\|^2 = \|(\mathbf{1} - e_n)x^*x(\mathbf{1} - e_n)\| \leq \|x^*x(\mathbf{1} - e_n)\|,$$

and (2) follows because the right side tends to 0 as $n \to \infty$. □

THEOREM 2.10.2. *Every closed ideal in a C^*-algebra is self-adjoint.*

PROOF. Let J be a closed ideal in a C^*-algebra A and choose an element $x \in J$. We have to show that $x^* \in J$. By Lemma (2.10.1) there is a sequence of self-adjoint elements $e_1, e_2, \ldots$ in J such that xe_n converges in norm to x as $n \to \infty$. Taking adjoints we find that $e_n x^*$ converges to x^*; since $e_n x^* \in J$ it follows that $x^* \in \bar{J} = J$. □

Suppose now that we are given a closed ideal J in a C^*-algebra A. We form the quotient Banach algebra as in Section 1.8. Since $J^* = J$, we can introduce an antilinear mapping on cosets by

$$(x + J)^* = x^* + J, \qquad x \in A,$$

and this defines an involution of the quotient algebra A/J.

THEOREM 2.10.3. *The involution above makes A/J into a C^*-algebra.*

PROOF. It suffices to show that for every element $x \in A$ the coset $\dot{x} = x + J$ satisfies $\|\dot{x}\|^2 \leq \|\dot{x}^*\dot{x}\|$. To prove this, consider the following set of elements of J:

$$E = \{e \in J : e^* = e, \sigma(e) \subseteq [0, 1]\}.$$

We claim that for every $x \in A$,

$$\|\dot{x}\| = \inf_{e \in E} \|x - xe\|. \tag{2.25}$$

Indeed, the inequality $\leq$ is clear from the fact that $xe \in J$ for every $e \in E$. For the opposite inequality, fix an element $k \in J$ and choose a sequence of elements $e_1, e_2, \dots$ satisfying the conditions of Lemma 2.10.1 with $ke_n \to k$. Then

$$(x + k)(\mathbf{1} - e_n) = (x - xe_n) + (k - ke_n).$$

The second term on the right tends to 0 as $n \to \infty$, and since $\|x + k\| \geq \|(x + k)(\mathbf{1} - e_n)\|$ for every n, we have

$$\|x + k\| \geq \liminf_{n \to \infty} \|x - xe_n\| \geq \inf_{e \in E} \|x - xe\|.$$

If we now take the infimum over all $k \in J$, we obtain

$$\|\dot{x}\| = \inf_{k \in J} \|x + k\| \geq \inf_{e \in E} \|x - xe\|,$$

and formula (2.25) is proved.

To see that $\|\dot{x}\|^2 \leq \|\dot{x}^*\dot{x}\|$, fix x and apply (2.25) as follows:

$$\begin{aligned} \|\dot{x}\|^2 &= \inf_{e \in E} \|x - xe\|^2 = \inf_{e \in E} \|(\mathbf{1} - e)x^*x(\mathbf{1} - e)\| \\ &\leq \inf_{e \in E} \|x^*x(\mathbf{1} - e)\| = \|x^*x + J\| = \|\dot{x}^*\dot{x}\|. \end{aligned}$$

□

THEOREM 2.10.4. *Let A and B be C^*-algebras and let $\pi : A \to B$ be a $*$-homomorphism. Then $\pi(A)$ is a C^*-subalgebra of B, and the natural promotion of π,*

$$\dot{\pi} : A/\ker \pi \to B,$$

is an isometric $*$-isomorphism of $A/\ker \pi$ onto $\pi(A)$.

PROOF. The map $\dot{\pi} : A/\ker \pi \to B$ is a $*$-homomorphism having kernel $\{0\}$. Since $A/\ker \pi$ is a C^*-algebra Proposition 2.9.3 implies that $\dot{\pi}$ is isometric. Hence its range $\pi(A) = \dot{\pi}(A/\ker \pi)$ is norm-closed in B. □

Exercises.

(1) Let $\{e_1, e_2, \dots\}$ be an orthonormal basis for a separable Hilbert space H, and let E_n be the projection on the span of $\{e_1, \dots, e_n\}$. Show that an operator $T \in \mathcal{B}(H)$ is compact iff

$$\lim_{n \to \infty} \|T - TE_n\| = 0,$$

and deduce that $\{E_n : n \in \mathbb{N}\}$ is an approximate unit for the C^*-algebra $\mathcal{K}$.

(2) Let U be a unitary operator on a Hilbert space H. Then $\sigma(U) \subseteq \mathbb{T}$, and hence there is a unique representation $\rho \in \operatorname{rep}(C(\mathbb{T}), H)$ satisfying $\rho(f) = f(U)$ for $f \in C(\mathbb{T})$. Identify $\ker \rho$ as an ideal in $C(\mathbb{T})$, identify the quotient $C(\mathbb{T})/\ker \rho$ in concrete terms as a commutative C^*-algebra, and similarly describe the natural factorization $\rho = \dot{\rho} \circ \pi$, where

$$\pi : C(\mathbb{T}) \to C(\mathbb{T})/\ker \rho$$

is the natural map onto the quotient C^*-algebra.

The remaining exercises relate to the Stone–Čech compactification of the real line, and of more general locally compact Hausdorff spaces. Let $C_b(\mathbb{R})$ be the space of all bounded continuous complex-valued functions of a real variable.

(3) Show that there is a compact Hausdorff space $\beta\mathbb{R}$ and an isometric $*$-isomorphism of $C_b(\mathbb{R})$ onto $C(\beta\mathbb{R})$. (Hint: $C_b(\mathbb{R})$ is a unital C^*-algebra. You must be explicit about this isomorphism or you will have trouble later on.)
(4) For every $t \in \mathbb{R}$, show that there is a (*naturally* defined) point $\hat{t} \in \beta\mathbb{R}$, and that the map $t \mapsto \hat{t}$ is a homeomorphism of $\mathbb{R}$ onto a dense subspace of $\beta\mathbb{R}$.

The space $\beta\mathbb{R}$ is called the Stone–Čech compactification of the real line $\mathbb{R}$.

(5) Identifying $\mathbb{R}$ with its image in $\beta\mathbb{R}$, the subspace $\beta\mathbb{R} \setminus \mathbb{R}$ is called the *corona* of $\mathbb{R}$. Show that the corona is closed (and hence, $\mathbb{R}$ is an open subset of $\beta\mathbb{R}$). Hint: For which points $p \in \beta\mathbb{R}$ does evaluation at p vanish on the ideal $C_0(\mathbb{R}) \subseteq C_b(\mathbb{R})$?
(6) Deduce that the quotient C^*-algebra $C_b(\mathbb{R})/C_0(\mathbb{R})$ is isometrically isomorphic to $C(\beta\mathbb{R} \setminus \mathbb{R})$.

A *compactification* of $\mathbb{R}$ is a pair (ϕ, Y) where Y is a compact Hausdorff space and $\phi : \mathbb{R} \to Y$ is a continuous map such that $\phi(\mathbb{R})$ is dense in Y.

(7) Show that $(t \mapsto \hat{t}, \beta\mathbb{R})$ is a universal compactification of $\mathbb{R}$ in the following sense: If (ϕ, Y) is any compactification of $\mathbb{R}$, then there is a unique extension of $\phi : \mathbb{R} \to Y$ to a continuous surjection $\hat{\phi} : \beta\mathbb{R} \to Y$. Hint: The map ϕ induces a $*$-isomorphism of $C(Y)$ onto a unital C^*-subalgebra of $C_b(\mathbb{R})$.

Your proof above extends easily to give a more general theorem, in which $\mathbb{R}$ is replaced by any locally compact noncompact Hausdorff space X (such as $\mathbb{R}^n$, $\mathbb{Z}^n$, or an open manifold), and one obtains a universal compactification βX called the Stone-Čech compactification of X. Formulate this theorem for yourself.

This page intentionally left blank

CHAPTER 3

Asymptotics: Compact Perturbations and Fredholm Theory

Operator theory modulo compact perturbations should be regarded as a study of the "asymptotic" properties of operators. After making this vague notion more precise in the context of Hilbert space operators, we take up the general theory of compact and Fredholm operators acting on Banach spaces and discuss a remarkable asymptotic invariant, the Fredholm index.

3.1. The Calkin Algebra

Let H be a separable Hilbert space and let $\mathcal{K}$ be the C^*-algebra of all compact operators on H. We have seen that $\mathcal{K}$ is a closed ideal in $\mathcal{B}(H)$. The quotient C^*-algebra $\mathcal{C} = \mathcal{B}(H)/\mathcal{K}$ is called the Calkin algebra. The Calkin algebra is important because it is the repository of all asymptotic information about operators on H. The purpose of this section is to discuss this aspect of operator theory in preparation for the more precise results to follow.

Let us begin in a simpler, commutative, context. A bounded sequence $x = (x_1, x_2, \dots)$ of complex numbers is an element of the C^*-algebra ℓ^∞, where addition, scalar multiplication, and multiplication are defined pointwise, and the norm is the usual one:

$$\|x\|_\infty = \sup_{n\geq 1} |x_n|.$$

We want to discuss properties of the sequence x that depend only on the behavior of the sequence at infinity, for example, the notion of a convergent sequence. Such properties can be expressed in terms of certain functions defined on *all* of ℓ^∞, such as

$$\limsup_{n\to\infty} |x_n| = \lim_{n\to\infty} (\sup\{|x_n|, |x_{n+1}|, |x_{n+2}|, \dots\}).$$

Other examples are the limit inferior and the limit superior of the sequence of real parts $\Re x_n$ of the components of x. In particular, a sequence x converges if and only if

$$\limsup_{n\to\infty} \Re x_n = \liminf_{n\to\infty} \Re x_n \quad \text{and} \quad \limsup_{n\to\infty} \Im x_n = \liminf_{n\to\infty} \Im x_n.$$

One can formalize the idea of an asymptotic invariant as follows. Let us say that a function $\phi : \ell^\infty \to \mathbb{C}$ is *asymptotic* if it is continuous relative to the norm topology of ℓ^∞ and has the property that for any two sequences

$x, y \in \ell^\infty$ for which $\lim_{n\to\infty} |x_n - y_n| = 0$, one has $\phi(x) = \phi(y)$. Notice that we do not require that ϕ be a linear functional; in fact, many of the important asymptotic properties of sequences, such as the examples above, are nonlinear.

The proper domain for asymptotic functions is the quotient C^*-algebra ℓ^∞/c_0. More precisely, consider the space c_0 of all sequences x that converge to zero:

$$\lim_{n\to\infty} x_n = 0.$$

Here c_0 is a closed ideal in ℓ^∞, and the quotient ℓ^∞/c_0 is a commutative C^*-algebra, whose Gelfand spectrum is identified with the corona $\beta\mathbb{N} \setminus \mathbb{N}$ of the Čech compactification of $\mathbb{N}$. Notice that by their definition, asymptotic functions $\phi : \ell^\infty \to \mathbb{C}$ promote naturally to continuous functions

$$\dot{\phi} : \ell^\infty/c_0 \to \mathbb{C}$$

by way of $\dot{\phi}(x + c_0) = \phi(x)$. Conversely, every continuous complex-valued function defined on ℓ^∞/c_0 is associated with an asymptotic function defined on ℓ^∞.

These remarks show that the asymptotic properties of sequences are tied to the quotient C^*-algebra ℓ^∞/c_0, or equivalently, to the corona space $\beta\mathbb{N} \setminus \mathbb{N}$. The latter is a very mysterious object: It is a compact Hausdorff space without isolated points, but whose topology is so large that no point p of $\beta\mathbb{N} \setminus \mathbb{N}$ can be approached with a sequence $p_1, p_2, \dots$ of distinct points of $\beta\mathbb{N} \setminus \mathbb{N}$. In particular, it is not possible to realize this space as a subset of any metric space. Thus one does not approach the analysis of asymptotic properties by analyzing $\beta\mathbb{N} \setminus \mathbb{N}$ as a topological space, but rather by dealing directly with concrete properties of the quotient C^*-algebra ℓ^∞/c_0.

Turning now to operator theory, the noncommutative counterpart of ℓ^∞ is the algebra $\mathcal{B}(H)$ of all bounded operators on a separable infinite-dimensional Hilbert space H. Let us introduce coordinates in H by choosing an orthonormal basis $\{e_1, e_2, \dots\}$. Let E_n be the projection of H onto the n-dimensional space spanned by $e_1, \dots, e_n$. The sequence E_n is increasing in the sense that $E_n \leq E_{n+1}$, and we have

$$\lim_{n\to\infty} E_n = \mathbf{1}$$

relative to the strong operator topology of $\mathcal{B}(H)$. Choose an operator $A \in \mathcal{B}(H)$ and consider its matrix (a_{ij}) relative to this basis:

$$a_{ij} = \langle Ae_j, e_i \rangle, \qquad i, j = 1, 2, \dots.$$

Notice that the matrix of $(\mathbf{1} - E_n)A(\mathbf{1} - E_n)$ is obtained from (a_{ij}) by replacing the first n rows and columns of (a_{ij}) with zeros and leaving the remaining entries fixed. Moreover, the result of Exercise (2) of Section 2.8 implies that A is compact iff

$$\lim_{n\to\infty} \|(\mathbf{1} - E_n)A(\mathbf{1} - E_n)\| = 0.$$

Thus the ideal $\mathcal{K}$ of compact operators in $\mathcal{B}(H)$ becomes the noncommutative counterpart of the ideal c_0 of all null-convergent sequences in ℓ^∞.

Similarly, one may consider asymptotic invariants of operators. For example, this could mean a continuous function $\phi : \mathcal{B}(H) \to \mathbb{C}$ with the property that $\phi(A) = \phi(B)$ whenever $A - B$ is compact. As before, such a function promotes naturally to a continuous function on the Calkin algebra

$$\dot{\phi} : \mathcal{B}(H)/\mathcal{K} \to \mathbb{C},$$

and every continuous complex function defined on the Calkin algebra arises in this way from an asymptotic function defined on $\mathcal{B}(H)$.

The most obvious example of an asymptotic invariant of operators $A \in \mathcal{B}(H)$ is their coset norm in the Calkin algebra,

$$\|A + \mathcal{K}\| = \lim_{n\to\infty} \|(\mathbf{1} - E_n)A(\mathbf{1} - E_n)\|,$$

corresponding to the coset norm of sequences $x \in \ell^\infty$,

$$\|x + c_0\| = \lim_{n\to\infty} \|(x_n, x_{n+1}, x_{n+2}, \dots)\|_\infty.$$

Another example is the essential spectrum, or more specifically the essential spectral radius,

$$r_e(A) = \sup\{|\lambda| : \lambda \in \sigma_e(T)\}.$$

Further examples are described in the Exercises.

Exercises. These exercises concern Banach limits and their noncommutative counterparts. Let $\ell^\infty = \ell^\infty(\mathbb{N})$ denote the Banach space of all bounded sequences of complex numbers $a = (a_n : n \geq 1)$, with the sup norm. We regard ℓ^∞ as a commutative C^*-algebra with unit $\mathbf{1} = (1, 1, 1, \dots)$ relative to the pointwise operations. Let T be the linear operator defined on ℓ^∞ by translating one step to the left and discarding the initial component:

$$(Ta)_n = a_{n+1}, \qquad n = 1, 2, \dots.$$

A *Banach limit* is a linear functional Λ on ℓ^∞ satisfying $\|\Lambda\| = \Lambda(\mathbf{1}) = 1$, that is translation invariant in the sense that $\Lambda(Ta) = \Lambda(a)$, $a \in \ell^\infty$. For the following exercises, Λ will denote a Banach limit.

(1) Show that Λ is a positive linear functional in the sense that

$$a_n \geq 0, \quad n = 1, 2, \dots \implies \Lambda(a) \geq 0.$$

(2) Show that for every real-valued sequence $a \in \ell^\infty$,

$$\liminf_{n\geq 1} a_n \leq \Lambda(a) \leq \limsup_{n\geq 1} a_n,$$

and deduce that $\Lambda(a) = \lim_{n\to\infty} a_n$ whenever a is a (complex) convergent sequence in c; in particular, for every $b \in \ell^\infty$ and $k \in c_0$,

$$\Lambda(b + k) = \Lambda(b).$$

(3) For $n = 1, 2, \dots$, let σ_n be the linear functional on ℓ^∞ defined by

$$\sigma_n(a) = \frac{a_1 + a_2 + \cdots + a_n}{n}, \qquad a \in \ell^\infty.$$

Then σ_n obviously satisfies $\|\sigma_n\| = \sigma_n(\mathbf{1}) = 1$. By estimating the norm, show that $\lim_{n\to\infty} \|\sigma_n \circ T - \sigma_n\| = 0$.

(4) (Existence of Banach limits) For every $n = 1, 2, \dots$, let K_n be the closure (in the weak*-topology of the dual of ℓ^∞) of the set of linear functionals $\{\sigma_n, \sigma_{n+1}, \sigma_{n+2}, \dots\}$. Show that $\cap_n K_n \neq \emptyset$, and that every linear functional in this intersection is a Banach limit.

In the remaining exercises, you will consider "noncommutative" Banach limits, as linear functionals on the noncommutative counterpart of ℓ^∞. Let $e_1, e_2, \dots$ be an orthonormal basis for a Hilbert space H, let $S \in \mathcal{B}(H)$ be the unilateral shift associated with this orthonormal basis by requiring $Se_n = e_{n+1}$, $n = 1, 2, \dots$, and let Λ be a Banach limit. Define a bounded linear functional ρ on $\mathcal{B}(H)$ as follows: $\rho(A) = \Lambda(a)$, where $a = (a_n)$ is the sequence $a_n = \langle Ae_n, e_n \rangle$, $n = 1, 2, \dots$. It is obvious that ρ is a positive linear functional in the sense that $A \geq 0 \implies \rho(A) \geq 0$, and of course $\rho(\mathbf{1}) = 1$.

(5) Show that $\rho(K) = 0$ for every compact operator K.

For the last exercise, it may help to compare the matrix of an operator A (relative to a fixed orthonormal basis (e_n)) to the matrix of its kth "translate" $S^{k*}AS^k$, noting that the latter is obtained from the matrix (a_{mn}) of A by deleting the first k rows and columns of (a_{mn}) and repositioning the result. How is the matrix of S^kAS^{k*} related to (a_{mn})?

(6) Show that $\rho(S^*AS) = \rho(A)$ and $\rho(SAS^*) = \rho(A)$, for every operator $A \in \mathcal{B}(H)$.

3.2. Riesz Theory of Compact Operators

Let E be a complex Banach space. An operator $T \in \mathcal{B}(E)$ is said to be *compact* if the image of the unit ball $\{T\xi : \|\xi\| \leq 1\}$ of E has compact closure relative to the *norm* topology of E. The set of all compact operators on E is denoted by $\mathcal{K}(E)$.

Since bounded sets in finite-dimensional Banach spaces are precompact, a finite-rank operator must be compact. The result of Exercise (3) below implies that $\mathcal{K}(E)$ is a norm-closed two-sided ideal in $\mathcal{B}(E)$. In particular, any operator T that can be norm-approximated by a sequence of finite-rank operators F_n, in the sense that $\|T - F_n\| \to 0$ as $n \to \infty$, must be compact. If E is a Hilbert space, then $\mathcal{K}(E)$ is the norm closure of the space of finite-rank operators, and that fact is useful for proving results about compact operators on Hilbert spaces. However, the reader should keep in mind that this convenient approximation property can fail for Banach spaces: $\mathcal{K}(E)$ can be properly larger than the norm closure of the finite-rank operators.

REMARK 3.2.1. *Kernels and cokernels.* We introduce some terminology that will be useful throughout the sequel. Suppose that $A \in \mathcal{B}(E)$ is a bounded operator that, for simplicity, we assume has closed range. There are two natural Banach spaces associated with A, namely, its *kernel* and *cokernel*:

$$\ker A = \{x \in E : Ax = 0\}, \qquad \operatorname{coker} A = E/AE.$$

The notion of cokernel bears some elaboration. An elementary result from the theory of Banach spaces asserts that there is a natural isomorphism between the annihilator $AE^{\perp} \subseteq E'$ of AE and the dual space of E/AE. On the other hand, the annihilator of AE is precisely the kernel of the operator adjoint $A' \in \mathcal{B}(E')$ of A. Thus we conclude that

$$\dim \operatorname{coker} A = \dim \ker A'$$

at least for every operator $A \in \mathcal{B}(E)$ whose range is closed and of finite codimension in E.

For such operators the two integers $\dim \ker A$ and $\dim \operatorname{coker} A$ provide important information about solutions of linear equations of the form

$$Ax = y,$$

where y is given and x is to be found. The number $\dim \ker A$ measures the degree of failure of uniqueness of solutions, and the number $\dim \operatorname{coker} A$ measures the degree of failure of existence of solutions. Much of what follows in this chapter has subtle and important implications for understanding these numerical invariants and their relation to each other.

The purpose of this section is to establish the following two general results about compact operators and their spectra.

THEOREM 3.2.2 (Fredholm alternative). *Let $T \in \mathcal{K}(E)$ and let λ be a nonzero complex number. Then either*

(1) *$\lambda - T$ is invertible, or*
(2) $\ker(\lambda - T) \neq \{0\}$.

Moreover, the kernel of $\lambda - T$ is finite dimensional, the range of $\lambda - T$ is a closed subspace of E of finite codimension, and we have

$$\dim \ker(\lambda - T) = \dim \operatorname{coker} (\lambda - T).$$

THEOREM 3.2.3 (Countability of spectrum). *Let T be a compact operator on an infinite-dimensional Banach space E. Then $0 \in \sigma(T)$, and every nonzero point of $\sigma(T)$ is an isolated point of $\sigma(T)$.*

REMARK 3.2.4. The Fredholm alternative leads to an effective procedure for solving linear equations of the form

$$Tx - \lambda x = y, \tag{3.1}$$

where T is a given compact operator, $\lambda \neq 0$ is a complex number, and y is a given vector in E. One first determines whether or not there are nontrivial eigenvectors with eigenvalue λ, by carrying out an analysis with the specific information one has about T. If there are no nonzero eigenvectors, then equation (3.1) is uniquely solvable for every given $y \in E$. Otherwise, there is a finite linearly independent set of vectors $x_1, \dots, x_n$ that span the eigenspace $\{x \in E : Tx = \lambda x\}$. In this case the equation has a solution iff y belongs to $(\lambda - T)E$; moreover, the general solution x of (3.1) can be determined from any particular solution x_0 as in undergraduate linear algebra and differential equations:

$$x = x_0 + a_1 x_1 + \cdots + a_n x_n,$$

where $a_1, \dots, a_n$ are arbitrary complex numbers.

This begs the issue of whether or not y belongs to the range of $\lambda - T$. To approach that, one first computes the adjoint $T' \in \mathcal{B}(E')$. Noting that the annihilator of $(\lambda - T)E$ is the dual eigenspace $\{g \in E' : T'g = \lambda g\}$, one sees from Theorem 3.2.2 that there is a set of n linearly independent linear functionals $f_1, \dots, f_n \in E'$ which span the space $\{g \in E' : T'g = \lambda g\}$. Once one has computed such a basis $f_1, \dots, f_n$ one may conclude that for a given $y \in E$, (3.1) has a solution iff

$$f_1(y) = \cdots = f_n(y) = 0.$$

Finally, notice that Theorem 3.2.3 implies that when E is infinite dimensional, the spectrum of any compact operator is either just $\{0\}$ (which is, by the Gelfand–Mazur theorem, equivalent to the assertion that T is quasinilpotent), or it consists of 0 and a finite number of nonzero points, or else it has the form

$$\sigma(T) = \{0\} \cup \{\lambda_1, \lambda_2, \dots\},$$

where $\lambda_1, \lambda_2, \dots$ is a sequence of nonzero complex numbers converging to 0.

REMARK 3.2.5. Note first that by replacing T with $\lambda^{-1}T$, we may without loss of generality assume that $\lambda = 1$ in the assertions of Theorem 3.2.2. The kernel of $\mathbf{1} - T$ is finite dimensional. This is an immediate consequence of Exercise (1) below, since T is a compact operator whose restriction to $\ker(\mathbf{1} - T)$ is the identity operator of $\ker(\mathbf{1} - T)$.

Similarly, if R denotes the *closure* of $(\mathbf{1} - T)E$, then R must be of finite codimension in E because the annihilator of R in the dual of E is the kernel of the operator $\mathbf{1} - T'$, and T' is compact by the result of Exercise (4) below.

The proof of the Fredholm alternative (Theorem 3.2.2) involves three steps, which we establish as Lemmas:

LEMMA 3.2.6. *Let $T \in \mathcal{K}(E)$, and let $M \subseteq E$ be a closed subspace of E. Then $(\mathbf{1} - T)M$ is closed.*

PROOF. We first point out that it suffices to prove the assertion for the case where the restriction of $\mathbf{1} - T$ is one-to-one. Indeed, let $F = M \cap \ker(\mathbf{1} - T)$. By Remark 3.2.5, F is a finite-dimensional subspace of M, and thus it must have a complement, a closed subspace $N \subseteq M$ with the property that $N \cap F = \{0\}$ and $N + F = M$ (see Exercise (2) below). It follows that $(\mathbf{1} - T)M = (\mathbf{1} - T)N$, and the restriction of $\mathbf{1} - T$ to N has trivial kernel.

Thus we may assume that $M \cap \ker(\mathbf{1} - T) = \{0\}$. Pick an element y in the closure of $(\mathbf{1} - T)M$. We will show that $y \in (\mathbf{1} - T)M$. To see that, choose a sequence $x_n \in M$ such that $x_n - Tx_n \to y$ as $n \to \infty$. We claim that $\|x_n\|$ is bounded. Indeed, if it is not, then there is a subsequence $x_{n'}$ of x_n such that $\|x_{n'}\| \to \infty$. Set $e_{n'} = \|x_{n'}\|^{-1} x_{n'}$. This defines a sequence of unit vectors of M for which $\|Te_{n'} - e_{n'}\| \to 0$. Since T is a compact operator, there must be a subsequence $e_{n''}$ with the property that $Te_{n''}$ converges in the norm of E. Since $\|e_{n'} - Te_{n'}\| \to 0$, it follows that $e_{n''}$ must converge to some vector f, which must be a unit vector in M because each e_n has these properties and M is closed. Finally, we have $f = Tf$, contradicting the fact that the restriction of $\mathbf{1} - T$ to M is injective.

Thus the sequence $x_1, x_2, \dots$ is bounded. Again, compactness of T implies that there is a subsequence $x_{n'}$ with the property that $Tx_{n'}$ converges in norm to some vector. Since $x_n - Tx_n \to y$, it follows that $x_{n'}$ must itself converge to some vector $x \in M$, and we have

$$x - Tx = \lim_{n' \to \infty} x_{n'} - Tx_{n'} = y,$$

and hence $y \in (\mathbf{1} - T)M$. □

LEMMA 3.2.7. *For every compact operator T on E,*

$$\ker(\mathbf{1} - T) = \{0\} \iff (\mathbf{1} - T)E = E.$$

PROOF. We first prove $\Longrightarrow$. For every $n = 0, 1, 2, \dots$, set

$$M_n = (\mathbf{1} - T)^n E.$$

We have $E = M_0 \supseteq M_1 \supseteq M_2 \supseteq \cdots$, each M_n is T-invariant in that $TM_n \subseteq M_n$, and from Lemma 3.2.6 and an obvious induction it follows that M_n is closed.

We claim that if $(\mathbf{1} - T)E \neq E$, then $M_n \neq M_{n+1}$ for every $n = 0, 1, \dots$. To see this, assume that there is a vector $x_0 \in E$ that fails to belong to $(\mathbf{1} - T)E$, and fix n. We will show that $(\mathbf{1} - T)^n x_0 \notin M_{n+1}$. Indeed, if there were to exist a $y_0 \in E$ such that $(\mathbf{1} - T)^n x_0 = (\mathbf{1} - T)^{n+1} y_0$, then $(\mathbf{1} - T)^n (x_0 - (\mathbf{1} - T)y_0) = 0$. Since we are assuming that $\mathbf{1} - T$ is injective, $(\mathbf{1} - T)^n$ is also injective; hence the previous formula implies $x_0 = (\mathbf{1} - T)y_0 \in (\mathbf{1} - T)E$, contrary to assumption.

Thus, assuming $(\mathbf{1}-T)E \neq E$, it follows that the sequence $M_0, M_1, \dots$ is strictly decreasing. For each $n = 0, 1, 2, \dots$ we choose a unit vector $e_n \in M_n$ such that

$$d(e_n, M_{n+1}) = \inf_{y \in M_{n+1}} \|e_n - y\| \geq \frac{1}{2}. \tag{3.2}$$

Notice that

$$\|Te_n - Te_{n+k}\| \geq \frac{1}{2}, \qquad k \geq 1, \quad n = 0, 1, 2, \dots. \tag{3.3}$$

Indeed, we have

$$Te_n - Te_{n+1} = e_n - [(\mathbf{1} - T)e_n + Te_{n+k})].$$

The bracketed term on the right belongs to M_{n+1}, since $Te_{n+k} \in M_{n+k} \subseteq M_{n+1}$ and $(\mathbf{1} - T)e_n \in (\mathbf{1} - T)M_n \subseteq M_{n+1}$. Hence

$$\|e_n - [(\mathbf{1} - T)e_n + Te_{n+k})]\| \geq d(e_n, M_{n+1}) \geq \frac{1}{2},$$

which proves (3.3). Clearly, (3.3) violates the compactness hypothesis on T, and hence $(\mathbf{1} - T)E = E$.

For the proof of $\Longleftarrow$, consider the adjoint operator $T' \in \mathcal{B}(E')$. The hypothesis $(\mathbf{1}-T)E = E$ implies that $\ker(\mathbf{1}-T') = \{0\}$. Since T' is compact (see Exercise (4)), the argument just given implies that $(\mathbf{1} - T')E' = E'$. In turn, this implies that $\ker(\mathbf{1} - T) = \{0\}$. Indeed, every bounded linear functional f on E has the form $f = g \circ (\mathbf{1} - T)$ by hypothesis; hence for any vector $x \in \ker(\mathbf{1} - T)$ we have $f(x) = g((\mathbf{1} - T)x) = 0$, and $x = 0$ follows from the Hahn–Banach theorem. □

To summarize progress, we have shown that $\ker(\mathbf{1}-T)$ and $\operatorname{coker}(\mathbf{1}-T)$ are both finite dimensional and that $\mathbf{1} - T$ has closed range; and we have the assertion of Lemma 3.2.7. We now extend the result of Lemma 3.2.7, as follows:

LEMMA 3.2.8. *For every compact operator T on E,*

$$\dim \ker(\mathbf{1} - T) = \dim \operatorname{coker}(\mathbf{1} - T).$$

PROOF. Choose a basis $x_1, \dots, x_m$ for $\ker(\mathbf{1} - T)$ and choose vectors $y_1, \dots, y_n \in E$ whose cosets $\dot{y}_1, \dots, \dot{y}_n$ are a basis for the cokernel $E/(\mathbf{1} - T)E$. Notice that the linear span $[y_1, \dots, y_n]$ intersects trivially with $(\mathbf{1} - T)E$. We have to show that $m = n$.

The alternatives $m \leq n$ or $m \geq n$ can be dealt with in turn. Assuming first that $m \leq n$, we choose a closed complement N for $\ker(\mathbf{1} - T)$ and consider the finite-rank operator $F \in \mathcal{B}(E)$ defined as zero on N and so as to map x_k to y_k for $k = 1, 2, \dots, m$. Notice that $\ker(\mathbf{1} - T) \cap \ker F = \{0\}$. The operator $\tilde{T} = T + F$, being a finite-rank perturbation of T, is compact. We claim that $\mathbf{1} - \tilde{T}$ has trivial kernel. Indeed, if $\tilde{T}x = x$, then $x - Tx = Fx \in (\mathbf{1} - T)E \cap [y_1, \dots, y_n] = \{0\}$. Hence $x \in \ker(\mathbf{1} - T) \cap \ker F = \{0\}$, proving the claim. It follows from Lemma 2.4.3 that $E = (\mathbf{1} - \tilde{T})E$. Now, on the one hand,

$$E/(\mathbf{1} - T)E = [\dot{y}_1, \dots, \dot{y}_n],$$

while since $(\mathbf{1} - \tilde{T})E \subseteq (\mathbf{1} - T)E + FE = (\mathbf{1} - T)E + [y_1, \dots, y_m]$, we also have

$$(\mathbf{1} - \tilde{T})E/(\mathbf{1} - T)E \subseteq [\dot{y}_1, \dots, \dot{y}_m].$$

Since $E = (\mathbf{1} - \tilde{T})E$, these relations imply that $[\dot{y}_1, \dots, \dot{y}_n] \subseteq [\dot{y}_1, \dots, \dot{y}_m]$, from which we conclude that $n = m$.

Assuming that $m \geq n$, one can construct a finite-rank operator G mapping $[x_1, \dots, x_m]$ onto $[y_1, \dots, y_n]$. By arguing with the perturbation $T + G$ in a similar way one shows that $\mathbf{1} - (T + G)$ is injective, and argues to the conclusion that m can be no larger than n. The reader is asked to flesh out this argument in Exercise (5) below. □

Proof of Theorem 3.2.2. We deduce Theorem 3.2.2 from the preceding discussion as follows. If $\mathbf{1} - T$ is not invertible, then $\ker(\mathbf{1} - T)$ must be nontrivial, since if the kernel is trivial, then by Lemma 3.2.7, $\mathbf{1} - T$ is onto, hence invertible. The finite dimensionality of $\ker(\mathbf{1} - T)$, and the closure and finite codimensionality of $(\mathbf{1} - T)E$, have also been established, and Lemma 3.2.8 provides the formula relating the dimensions of the kernel and cokernel. □

Proof of Theorem 3.2.3. We show first that $0 \in \sigma(T)$. Indeed, if 0 does not belong to $\sigma(T)$, then T is invertible. Since $\mathcal{K}(E)$ is an ideal, it follows that $\mathbf{1} = T^{-1}T$ is compact. This implies that the unit ball in E is compact, and hence E is finite dimensional (Exercise (1) below).

In order to establish the remaining assertions of Theorem 3.2.3, it suffices to prove the following: If $\lambda_1, \lambda_2, \dots$ is a sequence of distinct nonzero complex numbers in $\sigma(T)$, then

$$\lim_{n\to\infty} \lambda_n = 0. \tag{3.4}$$

To prove this, assume that $\lambda_1, \lambda_2, \dots$ is a sequence of distinct nonzero points in $\sigma(T)$ that does not converge to 0. By passing to a subsequence if necessary, we can assume that there is an $\epsilon > 0$ such that $|\lambda_n| \geq \epsilon$ for every $n = 1, 2, \dots$.

Theorem 3.2.2 implies that $\lambda_n - T$ has nonzero kernel for every n; hence we can find a unit vector e_n such that $Te_n = \lambda_n e_n$ for every n. Notice that the sequence $e_1, e_2, \dots$ is linearly independent. Indeed, for fixed n, $\lambda_1, \dots, \lambda_n$ are distinct complex numbers, so we can find polynomials $p_1, \dots, p_n$ such that $p_i(\lambda_j) = \delta_{ij}$ for $1 \leq i, j \leq n$. If some linear combination of $e_1, \dots, e_n$ vanishes,

$$a_1e_1 + \cdots + a_ne_n = 0,$$

then after applying $p_k(T)$ to this equation and using $p_k(T)e_j = \delta_{kj}e_k$ we obtain

$$a_ke_k = a_1p_k(T)e_1 + \cdots + a_np_k(T)e_n = p_k(T)(a_1e_1 + \cdots + a_ne_n) = 0,$$

and hence $a_k = 0$ for all k.

The subspaces $M_1, M_2, \dots$ defined by $M_n = [e_1, e_2, \dots, e_n]$ are strictly increasing with n; hence we can find unit vectors $u_1, u_2, \dots$ such that $u_k \in M_k$ and $d(u_k, M_{k-1}) \geq \frac{1}{2}$ for every $k = 2, 3, \dots$. Finally, notice that $(T - \lambda_n)M_n \subseteq M_{n-1}$ for every $n \geq 2$, simply because each e_k is an eigenvector of T with eigenvalue λ_k. In particular, $(T - \lambda_n)e_n \in M_{n-1}$ for $n \geq 2$.

It follows that for $1 \leq k < n$ we have

$$Tu_n - Tu_k = \lambda_n u_n + [(T - \lambda_n)u_n - Tu_k].$$

Since the bracketed vector on the right belongs to M_{n-1}, it follows that

$$\|Tu_n - Tu_k\| \geq d(\lambda_n u_n, M_{n-1}) = |\lambda_n| d(u_n, M_{n-1}) \geq \epsilon/2,$$

and the latter inequality contradicts the compactness hypothesis on T. $\square$

Exercises.

(1) (a) Let r be a real number satisfying $0 < r < 1$. Show that an infinite-dimensional Banach space E contains a sequence of unit vectors $e_1, e_2, \dots$ satisfying $\|e_k - e_j\| \geq r$ for all $j \neq k$. Hint: Use induction and elementary properties of the quotient norm in E/F where F is a finite-dimensional subspace of E.

(b) Deduce that the unit ball of a Banach space E is compact iff E is finite dimensional.

(2) (a) Let F be a finite-dimensional subspace of a Banach space E. Show that there is an operator $P \in \mathcal{B}(E)$ satisfying $P^2 = P$ and $PE = F$. Hint: Pick a basis $x_1, \dots, x_n$ for F and find bounded linear functionals $f_1, \dots, f_n$ on E such that $f_i(x_j) = \delta_{ij}$.

(b) Deduce that every finite-dimensional subspace $F \subseteq E$ is complemented in the sense that there is a closed subspace $G \subseteq E$ with $G \cap F = \{0\}$ and $G + F = E$.

(c) Show that every closed subspace $M \subseteq E$ of finite codimension in E is complemented.

(3) Show that for any Banach space E, $\mathcal{K}(E)$ is a norm-closed two-sided ideal in $\mathcal{B}(E)$.

(4) Let T be a compact operator on a Banach space E. Show that the adjoint $T' \in \mathcal{B}(E')$ is compact. Hint: Use Ascoli's theorem.

(5) Supply the missing details to the last paragraph of the proof of Lemma 3.2.8.

3.3. Fredholm Operators

A bounded operator T on a Banach space E is said to be a *Fredholm operator* if $\ker T$ is finite dimensional and TE is a closed subspace of finite codimension in E. More briefly, one says that T has finite-dimensional kernel and cokernel. Notice that the assertion about $\operatorname{coker} T$ is subtle, in that one must verify that the range of T is closed, and of finite codimension. In general, there are nonclosed subspaces of Banach spaces that are of finite codimension, such as the kernels of discontinuous linear functionals. On the other hand, if a linear subspace R of finite codimension in E is the *range* of a bounded linear operator $T \in \mathcal{B}(E)$, then R must be closed (see Exercise (1) below). Thus, one can make a more symmetric linear-algebraic definition of a Fredholm operator as a *bounded* operator on E with the property that both $\ker T$ and $\operatorname{coker} T = E/TE$ are finite dimensional as complex vector spaces.

REMARK 3.3.1. Obviously, invertible operators have the Fredholm property. A noninvertible example is the unilateral shift acting on $\ell^2(\mathbb{N})$: Its range is a closed subspace of codimension 1, and its kernel is $\{0\}$. In the preceding section we have seen that any operator $\lambda + T$, with T compact and λ a nonzero scalar, is a Fredholm operator. In this section we summarize the basic properties of Fredholm operators and establish an important criterion, Atkinson's theorem. These results imply that Fredholmness is an asymptotic property in the sense that it is stable under compact perturbations.

Throughout the section E denotes an infinite-dimensional Banach space, and $\mathcal{K}(E)$ denotes the closed ideal of all compact operators in $\mathcal{B}(E)$. The natural homomorphism of $\mathcal{B}(E)$ onto the quotient Banach algebra $\mathcal{B}(E)/\mathcal{K}(E)$ is denoted by $T \mapsto \dot{T} = T + \mathcal{K}(E)$.

THEOREM 3.3.2 (Atkinson's theorem). *A bounded operator T on E is a Fredholm operator iff $\dot{T}$ is invertible in $\mathcal{B}(E)/\mathcal{K}(E)$.*

Before giving the proof, we collect some of its immediate consequences. Let $\mathcal{F}(E)$ be the set of all Fredholm operators on E.

COROLLARY 1. A bounded operator T belongs to $\mathcal{F}(E)$ iff there is an operator $S \in \mathcal{B}(E)$ such that $\mathbf{1} - ST$ and $\mathbf{1} - TS$ are both compact.

PROOF OF COROLLARY 1. If $\dot{T}$ is invertible in $\mathcal{B}(E)/\mathcal{K}(E)$, then its inverse is an element of the form $\dot{S}$ for some $S \in \mathcal{B}(E)$, and the operators $\mathbf{1} - ST$ and $\mathbf{1} - TS$ must be compact because they map to 0 in the quotient algebra. The converse follows immediately from Atkinson's theorem. □

COROLLARY 2. The set $\mathcal{F}(E)$ of Fredholm operators is open in the norm topology of $\mathcal{B}(E)$, it is stable under compact perturbations, it contains all invertible operators of $\mathcal{B}(E)$, and it is closed under operator multiplication.

PROOF OF COROLLARY 2. Atkinson's theorem implies that $\mathcal{F}(E)$ is the inverse image of the general linear group of $\mathcal{B}(E)/\mathcal{K}(E)$ under the continuous homomorphism $T \mapsto \dot{T}$; hence these assertions all follow from the fact that the set of invertible elements of a unital Banach algebra A forms a group that is open in the norm topology of A. □

The *essential spectrum* $\sigma_e(T)$ of an operator $T \in \mathcal{B}(E)$ is defined as the spectrum of the image $\dot{T}$ of T in $\mathcal{B}(E)/\mathcal{K}(E)$. $\sigma_e(T)$ is a compact subset of

$\sigma(T)$. The following result implies that there are points in the spectrum of T that cannot be removed by perturbing T with compact operators.

COROLLARY 3. Let T be a bounded operator on an infinite-dimensional Banach space E. Then $\sigma_e(T) \neq \emptyset$, and

$$\sigma_e(T) \subseteq \cap\{\sigma(T+K) : K \in \mathcal{K}(E)\}.$$

Perhaps it is overkill to present this Corollary as a consequence of Atkinson's theorem, since it can be readily deduced from more basic considerations (see Exercise (2) below).

PROOF OF THEOREM 3.3.2. For the proof of Atkinson's theorem, suppose first that $\dot{T}$ is invertible, and let $S \in \mathcal{B}(E)$ be an operator such that $\dot{S} = \dot{T}^{-1}$. From the formulas $\dot{S}\dot{T} = \mathbf{1}$ and $\dot{T}\dot{S} = \mathbf{1}$, it follows that there are compact operators K_1, K_2 such that

$$\mathbf{1} - ST = K_1, \qquad \mathbf{1} - TS = K_2.$$

We have to show that $\ker T$ is finite dimensional and that TE is a closed subspace of finite codimension in E.

For the first assertion, we have $ST = \mathbf{1} - K_1$, so that $\ker T \subseteq \ker ST = \ker(\mathbf{1}-K_1)$, and Theorem 3.2.2 implies that $\ker(\mathbf{1}-K_1)$ is finite dimensional. Consider now the range TE. Since $TS = \mathbf{1} - K_2$, we have $TE \supseteq TSE = (\mathbf{1} - K_2)E$, and by Theorem 3.2.2, $(\mathbf{1} - K_2)E$ is a closed subspace of E of finite codimension. Using elementary linear algebra we can make an obvious inductive argument to find a finite set of vectors $v_1, \dots, v_r$ such that

$$TE = (\mathbf{1} - K_2)E + [v_1, \dots, v_r],$$

exhibiting TE as a closed subspace of finite codimension in E.

Conversely, suppose that T is a Fredholm operator on E. Since $\ker T$ is finite dimensional and TE is a closed subspace of finite codimension, there are bounded operators P, Q on E such that $P^2 = P$, $Q^2 = Q$, $PE = \ker T$, and $QE = TE$ (see Exercise (2) of the preceding section). Notice that since P and $\mathbf{1} - Q$ are finite-rank idempotents, it suffices to show that there is a bounded operator S on E such that

$$ST = \mathbf{1} - P, \qquad TS = Q = \mathbf{1} - (\mathbf{1} - Q). \tag{3.5}$$

The formulas (3.5) imply that $\dot{S}\dot{T} = \dot{T}\dot{S} = \mathbf{1}$ in $\mathcal{B}(E)/\mathcal{K}(E)$. The operator S is obtained as follows. Let $N = (\mathbf{1} - P)E$. The restriction T_0 of T to N is an operator with trivial kernel that maps onto TE (since $TP = 0$). By the closed graph theorem T_0 is an invertible operator. Let $S_0 \in \mathcal{B}(TE, N)$ be its inverse. We have $S_0Tx = x$ for all $x \in N$, and $TS_0y = y$ for all $y \in TE$. Letting S be the composition $S = S_0 \circ Q$, one finds that formulas (3.5) are satisfied. □

REMARK 3.3.3. The proof of Atkinson's theorem shows somewhat more than we have asserted, namely that for any bounded operator T on E the following three conditions are equivalent:

(1) T is a Fredholm operator.
(2) There is an operator $S \in \mathcal{B}(E)$ such that $\mathbf{1} - ST$ and $\mathbf{1} - TS$ are compact.
(3) There is an operator $S \in \mathcal{B}(E)$ such that $\mathbf{1} - ST$ and $\mathbf{1} - TS$ are finite-rank operators.

In particular, we have the remarkable conclusion that *invertibility modulo compact operators is the same as invertibility modulo finite-rank operators.*

Exercises.

(1) Let E be a Banach space and let T be a bounded operator on E such that the vector space E/TE is finite dimensional. Show that the range of T is closed.

The Weyl spectrum $\sigma_W(T)$ of a bounded operator T on E is defined as the intersection $\cap\{\sigma(T + K) : K \in \mathcal{K}(E)\}$ of the spectra of all compact perturbations of T. It is empty when E is finite dimensional.

(2) Show that when E is infinite-dimensional the essential spectrum $\sigma_e(T)$ is a nonempty subset of $\sigma_W(T)$. Use elementary properties of Banach algebras and their quotients, but not Atkinson's theorem.

Let S be the unilateral shift, realized on a Hilbert space H with orthonormal basis $e_1, e_2, \dots$ as the unique bounded operator S satisfying $Se_n = e_{n+1}$, $n = 1, 2, \dots$.

(3) Show that the essential spectrum of S is the unit circle
$$\mathbb{T} = \{\lambda \in \mathbb{C} : |\lambda| = 1\}.$$
(4) Show that the Weyl spectrum of S is the closed unit disk.

3.4. The Fredholm Index

We introduce the Fredholm index, develop its basic properties in general, and end the section with a brief discussion of the index in the more concrete setting of operators on a Hilbert space.

Let T be a Fredholm operator on a Banach space E. Both vector spaces $\ker T = \{x \in E : Tx = 0\}$ and $\operatorname{coker} T = E/TE$ are finite-dimensional, and the *index* of T is defined as the difference
$$\operatorname{ind} T = \dim \ker T - \dim \operatorname{coker} T.$$
The Fredholm alternative (Theorem 3.2.2) becomes the assertion that an operator of the form $\lambda + T$, with T compact and λ a nonzero scalar, is a Fredholm operator of index zero. The unilateral shift S is a Fredholm operator with $\operatorname{ind} S = -1$ (see Remark 3.3.1). We have also pointed out in the last section that the dimension of $\operatorname{coker} T$ is the same as $\dim \ker T'$, where $T' \in \mathcal{B}(E')$ is the adjoint of T, so that
$$\operatorname{ind} T = \dim \ker T - \dim \ker T'.$$

This formula is perhaps most useful for operators on Hilbert spaces, where one can replace T' with the Hilbert space adjoint T^*.

Atkinson's theorem implies that the product ST of two Fredholm operators $S, T \in \mathcal{B}(E)$ is a Fredholm operator. The most important property of the index is its logarithmic additivity,

$$\operatorname{ind} ST = \operatorname{ind} S + \operatorname{ind} T, \tag{3.6}$$

which will be proved shortly. Once this formula is established, the remaining properties of the Fredholm index follow easily. Thus it is significant that formula (3.6) is fundamentally a result in infinite-dimensional linear algebra, having nothing to do with the topology of E or $\mathcal{B}(E)$. While it is not operator-theoretic orthodoxy to do so, we have chosen to present the general algebraic result and deduce (3.6) from it. This proof is not only natural from a formal point of view, it is also quite transparent.

For the moment, we shift attention away from the category of Banach spaces with bounded linear operators as maps to the category of complex vector spaces with linear transformations as maps. Let V be a complex vector space. By an operator on V we simply mean a linear transformation $T : V \to V$, and the set of all such is denoted by $\mathcal{L}(V)$, which is a complex algebra with unit. Every operator $T \in \mathcal{L}(V)$ has two vector spaces associated with it, namely, its kernel and cokernel

$$\ker T = \{x \in V : Tx = 0\}, \qquad \operatorname{coker} T = E/TE.$$

T is said to be a *Fredholm operator* if both of these vector spaces are finite dimensional. The set of Fredholm operators on V is denoted by $\mathcal{F}(V)$. Every operator $T \in \mathcal{F}(V)$ has an index, namely,

$$\operatorname{ind} T = \dim \ker T - \dim \operatorname{coker} T.$$

Notice that if E is a complex Banach space and V is its underlying vector space structure, then, as we have already seen, a bounded operator belongs to $\mathcal{F}(E)$ iff it defines an algebraic Fredholm operator on V, that is, $\mathcal{F}(E) = \mathcal{F}(V) \cap \mathcal{B}(E)$. Thus the following result implies the addition formula (3.6) for Fredholm operators on Banach spaces.

THEOREM 3.4.1 (Addition formula). *Let V be a complex vector space and let A, B be Fredholm operators on V. Then AB is a Fredholm operator, and*

$$\operatorname{ind} AB = \operatorname{ind} A + \operatorname{ind} B.$$

We will deduce Theorem 3.4.1 from two more precise formulas, in which both defects

$$\dim \ker A + \dim \ker B - \dim \ker AB$$

and

$$\dim \operatorname{coker} A + \dim \operatorname{coker} B - \dim \operatorname{coker} AB$$

are computed explicitly.

LEMMA 3.4.2. *Let V be a vector space, and let $A, B \in \mathcal{F}(V)$. Then*

$$\dim \ker A + \dim \ker B = \dim \ker AB + \dim(\ker A/(BV \cap \ker A)). \tag{3.7}$$

PROOF. Noting that $\ker B \subseteq \ker AB$ we claim

$$\dim(\ker AB/\ker B) = \dim(BV \cap \ker A). \tag{3.8}$$

To prove this, it is enough to exhibit a linear map L from $\ker AB$ onto $BV \cap \ker A$ whose kernel is exactly $\ker B$. It is defined by $L : x \mapsto Bx$, $x \in \ker AB$. Clearly, $L(\ker AB) \subseteq BV \cap \ker A$, and $Lx = 0$ iff $x \in \ker B$. L is surjective, since if y has the form $y = Bv \in \ker A$ for some $v \in V$, then $ABv = Ay = 0$; hence $v \in \ker AB$ and $Lv = Bv = y$.

We now add $\dim \ker B + \dim(\ker A/(BV \cap \ker A))$ to both sides of (3.8). Since $\dim(\ker AB/\ker B) + \dim \ker B = \dim \ker AB$, the left side becomes

$$\dim \ker AB + \dim(\ker A/(BV \cap \ker A));$$

for a similar reason, the right side becomes

$$\dim \ker B + \dim \ker A,$$

and we obtain the asserted formula. □

LEMMA 3.4.3. *Let V be a vector space, and let $A, B \in \mathcal{F}(V)$. Then*

$$\dim \operatorname{coker} A + \dim \operatorname{coker} B = \dim \operatorname{coker} AB + \dim((BV + \ker A)/BV). \tag{3.9}$$

PROOF. We first establish an elementary formula. If M is a subspace of V of finite codimension, then

$$\dim(V/M) = \dim(AV/AM) + \dim((M + \ker A)/M). \tag{3.10}$$

For the proof, consider the natural linear map $L : V/M \to AV/AM$ defined by $L(v+M) = Av+AM$. The range of L is obviously AV/AM, and we claim that $\ker L = (M + \ker A)/M$. Indeed, a coset $v + M$ belongs to the kernel of L iff $Av + AM = 0$ iff $Av \in AM$ iff there is an element $m \in M$ such that $A(v - m) = 0$, and the latter is equivalent to $v \in M + \ker A$. Formula (3.10) now follows from a familiar identity of finite-dimensional linear algebra:

$$\dim \operatorname{domain} L = \dim \operatorname{ran} L + \dim \ker L.$$

Taking $M = BV$ in (3.10), we obtain

$$\dim(V/BV) = \dim(AV/ABV) + \dim((BV + \ker A)/BV).$$

If we add $\dim V/AV$ to both sides, the left side becomes

$$\dim \operatorname{coker} A + \dim \operatorname{coker} B,$$

while the right side becomes

$$\dim(V/AV) + \dim(AV/ABV) + \dim((BV + \ker A)/BV).$$

Since $ABV \subseteq AV$, the first two terms sum to $\dim V/ABV = \dim \operatorname{coker} AB$, completing the proof. □

PROOF. Turning to the proof of Theorem 3.4.1, Lemma 3.4.2 implies that

$$\dim\ker AB \leq \dim\ker A + \dim\ker B < \infty,$$

while Lemma 3.4.3 implies

$$\dim\operatorname{coker} AB \leq \dim\operatorname{coker} A + \dim\operatorname{coker} B < \infty.$$

Thus $A, B \in \mathcal{F}(V) \implies AB \in \mathcal{F}(V)$. Now, for any two subspaces M, N of a vector space there is an obvious linear map of M onto $(N+M)/N$ with kernel $N \cap M$; hence $M/(N\cap M) \cong (N+M)/N$. It follows that

$$\ker A/BV \cap \ker A \cong (BV + \ker A)/BV,$$

and in particular,

$$\dim(\ker A/BV \cap \ker A) = \dim((BV + \ker A)/BV).$$

We infer from Lemmas 3.4.2 and 3.4.3 that

$$\begin{aligned} \dim\ker A + \dim\ker B - \dim\ker AB &= \\ \dim\operatorname{coker} A + \dim\operatorname{coker} B &- \dim\operatorname{coker} AB, \end{aligned} \tag{3.11}$$

and the required formula $\operatorname{ind} AB = \operatorname{ind} A + \operatorname{ind} B$ follows after one rearranges terms in (3.11). □

Returning now to the setting in which E is an infinite dimensional Banach space, we obtain a fundamental result:

COROLLARY 1. *For any two Fredholm operators A, B on E, the product AB is Fredholm, and*

$$\operatorname{ind} AB = \operatorname{ind} A + \operatorname{ind} B.$$

PROOF. Atkinson's theorem implies that $\mathcal{F}(E)$ is closed under operator multiplication. If we forget the topology of E and apply Theorem 3.4.1, we obtain the asserted formula. □

COROLLARY 2 (Stability of index). *For every Fredholm operator $A \in \mathcal{B}(E)$ and compact operator K,*

$$\operatorname{ind}(A+K) = \operatorname{ind} A.$$

PROOF. By Atkinson's theorem there is a Fredholm operator $B \in \mathcal{B}(E)$ such that $AB = \mathbf{1} + L$ with $L \in \mathcal{K}(E)$. We have $(A+K)B = \mathbf{1} + L'$ where $L' = L - KB \in \mathcal{K}(E)$. As we have already pointed out, the Fredholm alternative implies that $\operatorname{ind}(\mathbf{1}+L) = \operatorname{ind}(\mathbf{1}+L') = 0$; hence $\operatorname{ind} AB = \operatorname{ind}(A+K)B = 0$. Using Corollary 1 one has

$$\operatorname{ind}(A+K) + \operatorname{ind} B = \operatorname{ind}(A+K)B = \operatorname{ind} AB = \operatorname{ind} A + \operatorname{ind} B,$$

and the formula follows after one cancels the integer $\operatorname{ind} B$. □

REMARK 3.4.4. Given a Fredholm operator A and an integer n, one can find finite-rank operators F and F' such that

$$\dim\ker(A+F) > n, \qquad \dim\operatorname{coker}(A+F') > n$$

(it is an instructive exercise to carry this out with A the unilateral shift). In particular, both $\dim\ker(A+F)$ and $\dim\operatorname{coker}(A+F)$ fluctuate in an unbounded way as F varies over the finite-rank operators. It is quite remarkable that these fluctuations cancel each other, so that the difference $\dim\ker(A+F) - \dim\operatorname{coker}(A+F)$ remains at the constant value $\operatorname{ind} A$.

COROLLARY 3 (Continuity of index). Given a Fredholm operator A, let $A_1, A_2, \dots$ be a sequence of bounded operators that converges to A, $\lim_{n\to\infty}\|A_n - A\| = 0$. There is an n_0 such that for $n \geq n_0$, A_n is a Fredholm operator with $\operatorname{ind} A_n = \operatorname{ind} A$.

PROOF. By Atkinson's theorem, $\mathcal{F}(E)$ is open, so that $A_n \in \mathcal{F}(E)$ for sufficiently large n. We can also find a Fredholm operator B such that $AB = \mathbf{1} + K$ with K compact. Writing $A_n = A + T_n$ with $\|T_n\| \to 0$ as $n \to \infty$, we can find n_0 so that, for $n \geq n_0$, $\|T_n B\| < 1$ and hence $\mathbf{1} + T_n B$ is invertible. For such n, we have

$$\operatorname{ind} A_n + \operatorname{ind} B = \operatorname{ind}(A + T_n)B = \operatorname{ind}(\mathbf{1} + T_n B + K).$$

The right side vanishes because $\mathbf{1} + T_n B + K$ is a compact perturbation of an invertible operator (see Exercise (1) below). On the other hand,

$$\operatorname{ind} A + \operatorname{ind} B = \operatorname{ind} AB = \operatorname{ind}(\mathbf{1} + K) = 0$$

by the Fredholm alternative; hence $\operatorname{ind} A_n = -\operatorname{ind} B = \operatorname{ind} A$ for sufficiently large n. □

Finally, let us consider the case of Fredholm operators acting on a Hilbert space H. The unique feature of Hilbert space is the existence of the adjoint operation $A \mapsto A^*$, carrying $\mathcal{B}(H)$ to itself. One cannot identify A^* with the Banach space adjoint $A' \in \mathcal{B}(H')$, as one sees by considering the fact that $A \mapsto A^*$ is an antilinear map, while, for operators A on Banach spaces, $A \mapsto A'$ is a linear map. That is because the identification of H' with H given by the Riesz lemma is not a linear map but an antilinear map. But the difference between A^* and A' is slight; and when one is working with Hilbert spaces it is customary to use A^* rather than A'. Thus for Fredholm operators A acting on a Hilbert space we have $AH^\perp = \ker A^*$; hence $\dim\operatorname{coker} A = \dim\ker A^*$ and

$$\operatorname{ind} A = \dim\ker A - \dim\ker A^*.$$

Exercises.

(1) Let E be an infinite-dimensional Banach space.

(a) Show that a Fredholm operator T on E is a compact perturbation of an invertible operator iff its index vanishes. Hint: If $\operatorname{ind} T = 0$, show how to construct a finite-rank perturbation of T that is one-to-one and onto.

(b) Deduce the following concrete description of the equivalence relation $A \sim B \iff \operatorname{ind} A = \operatorname{ind} B$: Two Fredholm operators A and B on E have the same index iff there is an invertible operator C such that $A - BC$ is compact.

(2) Let S be the unilateral shift acting on a Hilbert space H (see the Exercises of the preceding section).

(a) Show that there is no compact operator K such that $S + K$ is invertible.

(b) Let $T \in \mathcal{F}(H)$ be a Fredholm operator of positive index n. Show that there is an invertible operator $C \in \mathcal{B}(H)$ and a compact operator K such that $T = S^{*n}C + K$.

(3) (a) Let N be a normal Fredholm operator on a Hilbert space H. Show that the index of N vanishes.

(b) Deduce that the unilateral shift S is not a compact perturbation of a normal operator.

(4) With S as in the preceding exercises, let $S \oplus S^* \in \mathcal{B}(H \oplus H)$ be the direct sum of S with its adjoint S^*. Show that $S \oplus S^*$ is a Fredholm operator and calculate its index.

(5) Let U be the bilateral shift, defined on a Hilbert space H by its action on a bilateral orthonormal basis $\{e_n : n \in \mathbb{Z}\}$ for H by $Ue_n = e_{n+1}$, $n \in \mathbb{Z}$. Let P be the projection onto the one-dimensional space spanned by e_0. Show that $U - UP$ is unitarily equivalent to the operator $S \oplus S^*$ of the preceding exercise, and deduce that $S \oplus S^*$ is a compact perturbation of a normal operator.

(6) Show that the spectrum of $S \oplus S^*$ is the closed unit disk, but the Weyl spectrum of $S \oplus S^*$ is the unit circle.

CHAPTER 4

Methods and Applications

In this chapter, a variety of operator-theoretic methods are developed within the context of determining the spectra of Toeplitz operators.

Let $\mathbb{Z}_+$ be the additive semigroup of nonnegative integers, and let A be a bounded operator that acts as follows on the Hilbert space $\ell^2(\mathbb{Z}_+)$:

$$(A\xi)_n = \sum_{k=0}^{\infty} c_{n-k}\xi_k, \qquad n = 0, 1, 2, \ldots, \tag{4.1}$$

where (c_n) is a bilateral sequence of complex numbers. Such an operator A is called a *Toeplitz operator* with associated sequence (c_n). More invariantly, a Toeplitz operator is a bounded operator A on a Hilbert space H with the property that there is an orthonormal basis $e_0, e_1, e_2, \ldots$ for which the matrix (a_{ij}) of A relative to this basis depends only on $i - j$,

$$(a_{ij}) = \begin{pmatrix} c_0 & c_{-1} & c_{-2} & c_{-3} & \ldots \\ c_1 & c_0 & c_{-1} & c_{-2} & \ldots \\ c_2 & c_1 & c_0 & c_{-1} & \ldots \\ c_3 & c_2 & c_1 & c_0 & \ldots \\ & & \ldots & & \end{pmatrix}. \tag{4.2}$$

Toeplitz operators arise in diverse applications, and a great deal of effort has gone into computing their spectra. The results are definitive for Toeplitz operators with "continuous symbol," and these results are presented in Section 4.6. For more general Toeplitz operators the results are incomplete, and this is an area of continuing research.

The results of Section 4.6 require tools that have significance extending well beyond the immediate problem of computing spectra, and we develop these methods in a general context appropriate for broader application. Topics treated in this chapter include a discussion of maximal abelian von Neumann algebras, the characterization of bounded Toeplitz matrices and the notion of symbol, the structure of the Toeplitz C^*-algebra including the identification of its Fredholm operators and their relation to the topology of curves, the elementary theory of the Hardy space H^2, and the index theorem. We conclude the chapter with a discussion of states of C^*-algebras and the Gelfand–Naimark theorem.

4.1. Maximal Abelian von Neumann Algebras

A von Neumann algebra is an algebra $\mathcal{M}$ of operators on a Hilbert space H that contains the identity operator, is self-adjoint in the sense that $\mathcal{M}^* = \mathcal{M}$, and is closed in the weak operator topology of $\mathcal{B}(H)$. We will not have much to say about general von Neumann algebras, but we will look closely at the commutative ones. The set of all commutative self-adjoint operator algebras acting on H is partially ordered with respect to inclusion, and a maximal element of this set is called a maximal abelian self-adjoint algebra. They are commonly denoted by the colorless acronym MASA.

Remark 4.1.1. Since the closure in the weak operator topology of any commutative self-adjoint subalgebra of $\mathcal{B}(H)$ is a commutative self-adjoint algebra, a MASA is a weakly closed subalgebra of $\mathcal{B}(H)$. It must contain the identity operator, since otherwise, it could be enlarged nontrivially by adjoining the identity to it. Hence a MASA is an abelian von Neumann algebra.

Actually, a MASA $\mathcal{M}$ coincides with its commutant $\mathcal{M}' = \{T \in \mathcal{B}(H) : TA = AT, A \in \mathcal{M}\}$. It is clearly a subset of $\mathcal{M}'$ because it is commutative. On the other hand, if $A \in \mathcal{M}'$, then writing $A = X + iY$ with X, Y self-adjoint elements of $\mathcal{M}'$ (here we use the fact that $\mathcal{M}$ is self-adjoint) we find that X must belong to $\mathcal{M}$ because the algebra generated by $\mathcal{M}$ and X is a commutative algebra containing $\mathcal{M}$. Similarly, $Y \in \mathcal{M}$, and hence $\mathcal{M}' = \mathcal{M}$. Finally, A straightforward application of Zorn's lemma shows that every self-adjoint family of commuting operators is contained in some MASA.

Theorem 4.1.2. *Let (X, μ) be a σ-finite measure space. Then the multiplication algebra $\mathcal{M} = \{M_f : f \in L^\infty(X, \mu)\}$ is a maximal abelian von Neumann algebra in $\mathcal{B}(L^2(X, \mu))$.*

Proof. Let $T \neq 0$ be a bounded operator on $L^2(X, \mu)$ that commutes with every operator in $\mathcal{M}$. We have to show that $T \in \mathcal{M}$.

Consider first the case in which μ is a finite measure. The constant function 1 belongs to $L^2(X\mu)$, and we can define a function g in $L^2(X, \mu)$ by $g = T1$. We will show that $g \in L^\infty$, $\|g\|_\infty \leq \|T\|$, and $T = M_g$. Note that for every $f \in L^\infty(X, \mu)$ we have $fg = M_f T1 = TM_f 1 = Tf$. Since $T \neq 0$, it follows that $g \neq 0$, and moreover,

$$\|fg\|_2 \leq \|T\| \cdot \|f\|_2.$$

Choosing $E \subseteq X$ to be a Borel set and taking $f = \chi_E$, we obtain

$$\int_E |g|^2 \, d\mu = \|\chi_E g\|_2^2 \leq \|T\|^2 \|\chi_E\|_2^2 = \|T\|^2 \mu(E). \tag{4.3}$$

This inequality implies that $|g(p)| \leq \|T\|$ almost everywhere. Indeed, if $c \geq 0$ is any number such that $E = \{p \in X : |g(p)| > c\}$ has positive measure, then (4.3) implies

$$c^2\mu(E) \leq \int_E |g|^2 \leq \|T\|^2\mu(E),$$

and hence $c \leq \|T\|$. Since $\|g\|_\infty$ is the supremum of all such c, we conclude that $g \in L^\infty(X,\mu)$ and $\|g\|_\infty \leq \|T\|$.

We have shown that M_g is a bounded operator that satisfies $M_g f = fg = Tf$ for all $f \in L^\infty(X,\mu)$; hence $M_g = T$ because $L^\infty(X,\mu)$ is dense in $L^2(X,\mu)$.

In the general case where μ is σ-finite, we decompose X into a sequence of disjoint Borel sets of finite measure:

$$X = X_1 \sqcup X_2 \sqcup \cdots .$$

Letting μ_n be the restriction of μ to X_n, $\mu_n(E) = \mu(E \cap X_n)$, we find that L^2 decomposes into a direct sum of Hilbert spaces:

$$L^2(X,\mu) = L^2(X_1,\mu_1) \oplus L^2(X_2,\mu_2) \oplus \ldots .$$

Since the projection of $L^2(X,\mu)$ onto $L^2(X_n,\mu_n)$ belongs to $\mathcal{M}$ (it is the operator that multiplies by the characteristic function of X_n), it must commute with T, and we obtain a corresponding decomposition

$$T = T_1 \oplus T_2 \oplus \cdots ,$$

where T_n is the restriction of T to $L^2(X_n,\mu_n)$. Since T_n commutes with the multiplication algebra of $L^2(X_n,\mu_n)$, the argument just given implies that there is a function $f_n \in L^\infty(X_n,\mu_n)$ such that $T_n = M_{f_n}$, and moreover, $\|f_n\|_\infty \leq \|T_n\| \leq \|T\|$ for every n. Thus the f_n are uniformly bounded, and we can define a function $f \in L^\infty(X,\mu)$ via $f = f_n$ on X_n, $n = 1, 2, \ldots$. The desired conclusion $T = M_f$ follows. $\square$

Every normal operator N generates a von Neumann algebra $W^*(N)$, namely the closure in the weak operator topology of the $*$-algebra generated by N and $\mathbf{1}$. Since N is normal, $W^*(N)$ is an abelian von Neumann algebra; and in some cases it is a *maximal* abelian von Neumann algebra. These are the normal operators that are "multiplicity-free." A comprehensive treatment of multiplicity theory would be inappropriate here, and we refer the reader to [**2**] for more detail. What we do require is the following *sufficient* condition for a multiplication operator to have this useful property.

THEOREM 4.1.3. *Let X be a compact subset of $\mathbb{C}$, let $f \in C(X)$ be a continuous function that separates points of X in the sense that $f(p) \neq f(q)$ for distinct points $p \neq q \in X$, and let μ be a finite measure on X.*

Consider the multiplication operator $M_f \in \mathcal{B}(L^2(X,\mu))$. Then $W^*(M_f)$ is the multiplication algebra $\mathcal{M}$ of $L^2(X,\mu)$, and every operator A that doubly commutes M_f,

$$AM_f = M_f A, \qquad AM_f^* = M_f^* A \tag{4.4}$$

belongs to $\mathcal{M} = W^*(M_f)$.

PROOF. Since f separates points of X, the Stone–Weierstrass theorem implies that $C(X)$ is generated as a C^*-algebra by f and the constant function 1. Theorem 2.1.3 implies that $\mathcal{A} = \{M_g : g \in C(X)\}$ is a C^*-subalgebra of the multiplication algebra, and in fact, it is the C^*-algebra generated by the two operators M_f and $\mathbf{1}$.

We claim now that the closure of $\mathcal{A}$ in the weak operator topology is the multiplication algebra $\mathcal{M}$. Indeed, for every bounded Borel function $h : X \to \mathbb{C}$ there is a uniformly bounded sequence $g_1, g_2, \dots$ of continuous functions such that $\lim_n g_n(p) = h(p)$ for almost every $p \in X$. Choosing such a sequence g_n, then for every pair of functions $\xi, \eta \in L^2(X, \mu)$ the function $\xi\bar{\eta}$ is integrable, so by the dominated convergence theorem

$$\lim_{n\to\infty} \langle M_{g_n}\xi, \eta\rangle = \lim_{n\to\infty} \int_X g_n(p)\xi(p)\bar{\eta}(p)\, d\mu = \int_X h(p)\xi(p)\bar{\eta}(p)\, d\mu = \langle M_h\xi, \eta\rangle.$$

Hence $M_h \in \overline{\mathcal{A}}^{\text{weak}}$, and we conclude that $\mathcal{M}$ is generated as a von Neumann algebra by M_f and $\mathbf{1}$.

Finally, let A be a bounded operator on $L^2(X, \mu)$ that commutes with M_f and M_f^*. Then A commutes with the weakly closed algebra generated by M_f, M_f^*, and $\mathbf{1}$, which, by the preceding paragraphs, contains the multiplication algebra $\mathcal{M}$. By Theorem 4.1.2, $A \in \mathcal{M}$. □

REMARK 4.1.4. It is significant that the second hypothesis $AM_f^* = M_f^*A$ in (4.4) is redundant. That is a consequence of a theorem of Bent Fuglede ([**19**], Proposition 4.4.12), which asserts that any operator that commutes with a normal operator N must also commute with its adjoint N^*.

We also remark that the finiteness hypothesis on the measure σ can be relaxed to σ-finiteness, in view of the fact that for mutually absolutely equivalent σ-finite measures μ, ν on X, the multiplication algebras of $L^2(X, \mu)$ and $L^2(X, \nu)$ are naturally unitarily equivalent (Exercise (2) of Section 2.6).

Finally, we point out that the hypotheses on f can be replaced with the hypothesis that f is a bounded Borel function that separates points of X; but that generalization requires more information about Borel structures than we have at our disposal (see chapter 3 of [**2**]).

COROLLARY 1. Let X be the standard operator on $L^2[0, 1]$,

$$X\xi(t) = t\xi(t), \qquad 0 \le t \le 1, \quad \xi \in L^2[0, 1].$$

For every operator A that commutes with X, there is a function $f \in L^\infty[0, 1]$ such that $A = M_f$.

COROLLARY 2. Let $\{e_n : n \in \mathbb{Z}\}$ be a bilateral orthonormal basis for a Hilbert space H, and let U be the bilateral shift defined on H by $Ue_n = e_{n+1}$, $n \in \mathbb{Z}$. Then the von Neumann algebra $W^*(U)$ generated by U is maximal abelian, and consists of all operators in $\mathcal{B}(H)$ that commute with U.

PROOF. We have seen that U is unitarily equivalent to the multiplication operator M_ζ acting on $L^2(\mathbb{T})$ by $M_\zeta\xi(z) = \zeta(z)\xi(z)$, ζ being the current variable $\zeta(z) = z$, $z \in \mathbb{T}$. Since M_ζ is unitary, any operator commuting with it must also commute with its adjoint $M_\zeta^* = M_\zeta^{-1}$. On the other hand, since ζ separates points of $\mathbb{T}$, it follows from Theorem 4.1.3 that any operator commuting with $\{M_\zeta, M_\zeta^*\}$ must belong to the multiplication algebra of $L^2(\mathbb{T})$, and that the multiplication algebra coincides with the von Neumann algebra generated by U. $\square$

Exercises.

(1) Show that the unit ball of $\mathcal{B}(H)$ is compact in its weak operator topology. Hint: Show that the unit ball of $\mathcal{B}(H)$ can be embedded as a closed subset of a Cartesian product of copies of the complex unit disk $\Delta = \{z \in \mathbb{C} : |z| \leq 1\}$, and appeal to the Tychonoff theorem.

In the following exercises, H denotes a *separable* Hilbert space.

(2) (a) Let $\xi_1, \xi_2, \dots$ be a sequence of vectors dense in the unit ball of H. Show that

$$d(A, B) = \sum_{m,n=1}^{\infty} 2^{-m-n} \frac{|\langle A\xi_m, \xi_n\rangle - \langle B\xi_m, \xi_n\rangle|}{1 + |\langle A\xi_m, \xi_n\rangle - \langle B\xi_m, \xi_n\rangle|}$$

is a metric on the unit ball of $\mathcal{B}(H)$ that is separately continuous in the weak operator topology.

(b) Show that, with its weak operator topology, the unit ball of $\mathcal{B}(H)$ is homeomorphic to a compact metric space.

(3) Deduce that every von Neumann algebra $\mathcal{M}$ acting on H contains a unital C*-subalgebra $\mathcal{A}$ that is (1) separable (i.e., $\mathcal{A}$ contains a countable norm-dense subset) and (2) weakly dense in $\mathcal{M}$.

In the following exercises, you will show that every maximal abelian von Neumann algebra $\mathcal{M} \subset \mathcal{B}(H)$ is unitarily equivalent to the multiplication algebra of a finite measure space, and deduce the spectral theorem from that result.

(4) If an abelian von Neumann algebra $\mathcal{M} \subset \mathcal{B}(H)$ has a cyclic vector, then there is a compact metric space X and a probability measure μ on X such that $\mathcal{M}$ is unitarily equivalent to the multiplication algebra of $L^2(X, \mu)$. Hint: Use Exercise (3).

(5) Let $\mathcal{M} \subset \mathcal{B}(H)$ be a MASA. Show that there is a sequence of mutually orthogonal cyclic projections in $\mathcal{M}$ that sum to the identity. Hint: The projection onto any $\mathcal{M}$-invariant subspace must belong to $\mathcal{M}$.

(6) Deduce that every MASA has a cyclic vector, and hence is unitarily equivalent to a multiplication algebra as in Exercise (4).

(7) Show that every *commutative* $*$-subalgebra $\mathcal{A} \subseteq \mathcal{B}(H)$ is contained in a maximal abelian von Neumann algebra in $\mathcal{B}(H)$, and deduce the spectral theorem from the result of the preceding exercise.

4.2. Toeplitz Matrices and Toeplitz Operators

Starting with a "symbol" (a function in L^∞), we introduce its associated Toeplitz operator acting on the Hardy space H^2 and develop the basic relations between the symbol and the operator. Then we discuss the more classical notion of Toeplitz matrix, and relate the two. Historically, Toeplitz matrices came first.

We begin by reviewing some notation and terminology that will be used throughout the following sections. L^2 will denote the Hilbert space $L^2(\mathbb{T}, \sigma)$, where σ is the normalized length $d\sigma = d\theta/2\pi$ on the unit circle $\mathbb{T}$ of the complex plane. Let $\zeta \in C(\mathbb{T})$ be the current variable, $\zeta(z) = z$, $z \in \mathbb{T}$. The set $\{\zeta^n : n \in \mathbb{Z}\}$ of powers of ζ is an orthonormal basis for L^2, and H^2 is defined as the closed subspace

$$H^2 = [1, \zeta, \zeta^2, \dots]$$

spanned by the nonnegative powers of ζ. The orthocomplement of H^2 is spanned by the negative powers of ζ,

$$H^{2\perp} = [\zeta^n : n < 0].$$

Elements of H^2 are functions f in L^2 whose Fourier series have the form

$$f(e^{i\theta}) \sim \sum_{n=0}^{\infty} a_n e^{in\theta}. \tag{4.5}$$

Similarly, L^∞ denotes the algebra $L^\infty(\mathbb{T}, \sigma)$. It is a commutative C^*-algebra which, in addition to its norm topology, has a weak* topology defined by its natural pairing with L^1. The corresponding subace of L^∞ is denoted by H^∞,

$$H^\infty = L^\infty \cap H^2.$$

By definition, a bounded measurable function f belongs to H^∞ iff its Fourier series has the form (4.5). Given $f \in L^\infty$, the following observation relates membership in H^∞ to properties of the multiplication operator M_f, and implies that H^∞ is a weak*-closed unital subalgebra of L^∞.

Proposition 4.2.1. $H^\infty = \{\phi \in L^\infty : M_\phi H^2 \subseteq H^2\}$.

Proof. Let $\phi \in L^\infty$. If $\phi \in H^\infty$, then for $n \geq 0$,

$$M_\phi \zeta^n = \zeta^n \cdot \phi \in \zeta^n \cdot H^2 \subseteq H^2;$$

hence M_ϕ leaves $H^2 = [1, \zeta, \zeta^2, \dots]$ invariant.

Conversely, if $M_\phi H^2 \subseteq H^2$, then $\phi = M_\phi 1 \in H^2$. □

For every $\phi \in L^\infty$, let $T_\phi \in \mathcal{B}(H^2)$ be the compression of M_ϕ to H^2,

$$T_\phi = P_+ M_\phi \restriction_{H^2},$$

P_+ denoting the projection of L^2 onto H^2. The operator T_ϕ is called the *Toeplitz operator with symbol* ϕ.

REMARK 4.2.2. The map $\phi \mapsto T_\phi$ is obviously a $*$-preserving bounded linear mapping of the commutative C^*-algebra L^∞ into $\mathcal{B}(H^2)$, which carries the unit of L^∞ to the identity operator and is *positive* in the sense that

$$\phi \geq 0 \implies T_\phi \geq 0.$$

Certainly, it is not a representation, but it has the following restricted multiplicativity property. For $f \in H^\infty$ and $g \in L^\infty$, we have

$$T_{fg} = T_g T_f, \qquad T_{\bar{f}g} = T_{\bar{f}} T_g. \tag{4.6}$$

Indeed, the first formula follows from

$$T_{fg} = P_+ M_{gf} \restriction_{H^2} = P_+ M_g M_f \restriction_{H^2} = P_+ M_g P_+ M_f \restriction_{H^2} = T_g T_f,$$

using $M_f H^2 \subseteq H^2$; the second formula follows from the first by taking adjoints.

A fundamental problem concerning Toeplitz operators is to determine $\sigma(T_\phi)$ in terms of the properties of ϕ. While the answer is known for important classes of symbols (e.g., when ϕ is real-valued, or belongs to H^∞, or is continuous), the general problem remains unsolved. The difficulty stems from the fact that the map $\phi \mapsto T_\phi$ fails to be multiplicative. We now direct our attention to developing tools for calculating $\sigma(T_\phi)$ when $\phi \in C(\mathbb{T})$.

A Toeplitz matrix is a matrix of the form (4.2) whose entries a_{ij} $i, j = 0, 1, \dots,$ depend only on $i - j$. We first show that Toeplitz matrices correspond to Toeplitz operators T_ϕ, and we determine their norm in terms of the symbol ϕ. The unilateral shift is identified in this context as the Toeplitz operator $S = T_\zeta$.

PROPOSITION 4.2.3. *Let A be a bounded operator on H^2. The matrix of A relative to the natural basis $\{\zeta^n : n = 0, 1, 2, \dots\}$ is a Toeplitz matrix iff $S^* A S = A$.*

PROOF. The hypothesis on the matrix entries $a_{ij} = \langle A\zeta^j, \zeta^i \rangle$ of A is equivalent to requiring

$$a_{i+1, j+1} = a_{ij}, \qquad i, j = 0, 1, 2, \dots.$$

Noting that $S\zeta^n = \zeta^{n+1}$ for $n \geq 0$ we find that this is equivalent to the requirement that

$$\langle S^* A S \zeta^j, \zeta^i \rangle = \langle A S \zeta^j, S \zeta^i \rangle = \langle A \zeta^{j+1}, \zeta^{i+1} \rangle = \langle A \zeta^j, \zeta^i \rangle$$

for all $i, j \geq 0$; hence it is equivalent to the formula $S^* A S = A$. $\square$

Thus, in order to determine which Toeplitz matrices (4.2) correspond to bounded operators, we must characterize the bounded operators A on H^2 that have the property $S^*AS = A$. This is accomplished as follows. Notice first that any Toeplitz operator T_ϕ with $\phi \in L^\infty$ satisfies $S^*T_\phi S = T_\phi$, since by (4.6) we have

$$S^*T_\phi S = T_{\bar\zeta}T_\phi T_\zeta = T_{\bar\zeta\phi}T_\zeta = T_{\bar\zeta\zeta\phi} = T_\phi,$$

since $\bar\zeta\zeta = 1$. Conversely:

THEOREM 4.2.4 (Characterization of Toeplitz operators). *Let A be a bounded operator on H^2 satisfying $S^*AS = A$. There is a unique function $\phi \in L^\infty$ such that $A = T_\phi$, and one has $\|A\| = \|\phi\|_\infty$.*

PROOF. For every $n = 0, 1, 2, \dots$ let M_n be the following subspace of L^2:

$$M_n = [\zeta^{-n}, \zeta^{-n+1}, \zeta^{-n+2}, \dots].$$

We have $H^2 = M_0 \subseteq M_1 \subseteq M_2 \subseteq \cdots$ and the union $\cup_n M_n$ is dense in L^2. Let $U = M_\zeta \in \mathcal{B}(L^2)$. U is a unitary operator whose restriction to H^2 is the unilateral shift S, and it maps M_n into M_{n-1} for $n \geq 1$. Thus $U^n M_n \subseteq H^2$, and we can define a sequence of operators $A_n \in \mathcal{B}(M_n)$ as follows:

$$A_n = U^{-n}AU^n \restriction_{M_n}.$$

Each A_n is obviously unitarily equivalent to A; hence $\|A_n\| = \|A\|$. Moreover, we claim:

- The sequence $A_1, A_2, \dots$ is coherent in the sense that

$$\langle A_{n+1}\xi, \eta\rangle = \langle A_n\xi, \eta\rangle, \qquad \xi, \eta \in M_n. \tag{4.7}$$

- For every $n \geq 1$,

$$P_+A_n \restriction_{H^2} = A. \tag{4.8}$$

Indeed, since $U^n\xi$ and $U^n\eta$ belong to H^2, we have

$$\begin{aligned}\langle A_{n+1}\xi, \eta\rangle &= \langle U^{-(n+1)}AU^{n+1}\xi, \eta\rangle = \langle ASU^n\xi, SU^n\eta\rangle \\ &= \langle S^*ASU^n\xi, U^n\eta\rangle.\end{aligned}$$

Since $S^*AS = A$, the right side is $\langle AU^n\xi, U^n\eta\rangle = \langle A_n\xi, \eta\rangle$ as (4.7) asserts. For (4.8), note that for $\xi, \eta \in H^2$ one has

$$\begin{aligned}\langle P_+A_n\xi, \eta\rangle &= \langle U^{-n}AU^n\xi, \eta\rangle = \langle AU^n\xi, U^n\eta\rangle = \langle AS^n\xi, S^n\eta\rangle \\ &= \langle S^{*n}AS^n\xi, \eta\rangle = \langle A\xi, \eta\rangle.\end{aligned}$$

It follows from (4.7) that we can use the Riesz lemma to define a unique operator $\tilde{A} \in \mathcal{B}(L^2)$ as a weak limit

$$\langle \tilde{A}\xi, \eta\rangle = \lim_{n\to\infty}\langle A_n\xi, \eta\rangle, \qquad \xi, \eta \in \cup_{n\geq 1}M_n,$$

and since $\|A_n\| = \|A\|$ for every n, we have $\|\tilde{A}\| \leq \|A\|$. We claim that $\tilde{A}$ is a multiplication operator M_ϕ, $\phi \in L^\infty$. In view of Corollary 2 of

Theorem 4.1.3, this follows from the fact that $\tilde{A}$ commutes with U; indeed, for $\xi, \eta \in \cup_{n\geq 1} M_n$ we have

$$\langle U^{-1}\tilde{A}U\xi, \eta\rangle = \lim_{n\to\infty} \langle U^{-1}A_nU\xi, \eta\rangle = \lim_{n\to\infty} \langle A_{n+1}\xi, \eta\rangle = \langle \tilde{A}\xi, \eta\rangle.$$

We have $\|\phi\|_\infty = \|M_\phi\| = \|\tilde{A}\| \leq \|A\|$. Formula (4.8) implies that the compression of $\tilde{A}$ to H^2 is A; hence $A = T_\phi$. The inequality $\|A\| = \|T_\phi\| \leq \|\phi\|_\infty$ is obvious, and uniqueness of ϕ follows from $\|T_\phi\| = \|\phi\|_\infty$. □

In more concrete terms, Theorem 4.2.4 makes the following assertion: Let (a_{ij}) be a formal Toeplitz matrix

$$(a_{ij}) = \begin{pmatrix} c_0 & c_{-1} & c_{-2} & c_{-3} & \dots \\ c_1 & c_0 & c_{-1} & c_{-2} & \dots \\ c_2 & c_1 & c_0 & c_{-1} & \dots \\ c_3 & c_2 & c_1 & c_0 & \dots \\ & & \dots & & \end{pmatrix},$$

where c_n, $n \in \mathbb{Z}$, is a doubly infinite sequence of complex numbers. Then (a_{ij}) is the matrix of a bounded operator iff there is a function $\phi \in L^\infty$ with Fourier series

$$\phi(e^{i\theta}) \sim \sum_{n=-\infty}^{\infty} c_n e^{in\theta}.$$

When such a function ϕ exists, it is unique, $\|(a_{ij})\| = \|\phi\|_\infty$, and in that case, the operator defined on $\ell^2(\mathbb{Z}_+)$ by the matrix

$$(A\xi)_n = \sum_{k=0}^{\infty} c_{n-k}\xi_k, \qquad \xi \in \ell^2(\mathbb{Z}_+),$$

is unitarily equivalent to the Toeplitz operator $T_\phi \in \mathcal{B}(H^2)$. The function ϕ is called the *symbol* of the Toeplitz matrix (a_{ij}) or of the operator T_ϕ.

Corollary 1. Every Toeplitz operator T_ϕ, $\phi \in L^\infty$, satisfies

$$\inf\{\|T_\phi + K\| : K \in \mathcal{K}\} = \|T_\phi\| = \|\phi\|_\infty.$$

In particular, the only compact Toeplitz operator is 0.

Proof. Let S be the unilateral shift acting on H^2 by $S\zeta^n = \zeta^{n+1}$, $n \geq 0$. It suffices to show that for any operator $A \in \mathcal{B}(H^2)$ satisfying $S^*AS = A$ and for any compact operator K we have

$$\|A + K\| \geq \|A\|.$$

The hypothesis $S^*AS = A$ implies that $S^{*n}AS^n = A$ for every $n = 1, 2, \dots$; noting that $P_n = S^nS^{*n}$ is the projection onto $[\zeta^n, \zeta^{n+1}, \dots]$ we have

$$\|A + K\| \geq \|P_n(A + K)P_n\| = \|S^{*n}(A + K)S^n\| = \|A + S^{*n}KS^n\|.$$

The norm of the compression of K to the subspace $[\zeta^n, \zeta^{n+1}, \dots]$ is given by $\|P_n K P_n\| = \|S^{*n} K S^n\|$, which tends to 0 as $n \to \infty$ because K is compact and $P_n \downarrow 0$. Thus

$$\|A + K\| \geq \lim_{n\to\infty} \|A + S^{*n} K S^n\| = \|A\|,$$

as asserted. □

Exercises. Let Λ be a Banach limit on ℓ^∞ (see the Exercises of Section 3.1). Given a sequence $a = (a_1, a_2, \dots) \in \ell^\infty$ we will write $\Lambda_n a_n$ for the value of Λ on a. Let $S = T_\zeta$ be the natural realization of the unilateral shift on H^2.

(1) Show that for every operator $A \in \mathcal{B}(H^2)$ there is a unique operator $\phi(A) \in \mathcal{B}(H^2)$ satisfying
$$\langle \phi(A)\xi, \eta\rangle = \Lambda_n \langle S^{*n} A S^n \xi, \eta\rangle, \qquad \xi, \eta \in H^2.$$
(2) Show that $\phi(A)$ is a Toeplitz operator (i.e., has the form T_f for some $f \in L^\infty$) for every $A \in \mathcal{B}(H^2)$.
(3) Deduce that ϕ is a projection of norm 1 of the Banach space $\mathcal{B}(H^2)$ onto the subspace $\mathcal{S} = \{T_f : f \in L^\infty\}$ of all Toeplitz operators on H^2, satisfying $\phi(AT_f) = \phi(A)T_f$ for $f \in H^\infty$ and $\phi(K) = 0$ for every compact operator K.

4.3. The Toeplitz C^*-Algebra

Let H be a Hilbert space having an orthonormal basis $e_0, e_1, e_2, \dots$ and let S be the unique operator defined by $Se_n = e_{n+1}$, $n \geq 0$. The operator S is called the *unilateral shift*. The C^*-algebra generated by S is of central importance in modern analysis; it is called the Toeplitz C^*-algebra and is often denoted by $\mathcal{T}$. In this section we give a concrete description of the Fredholm operators in $\mathcal{T}$; and in the next we calculate their index.

This is accomplished by relating $\mathcal{T}$ to Toeplitz operators with continuous symbol. We have seen that S can be realized as the Toeplitz operator $T_\zeta \in \mathcal{B}(H^2)$, ζ being the current variable, and throughout this section we take $S = T_\zeta$. Recall that the map $\phi \in L^\infty \mapsto T_\phi \in \mathcal{B}(H^2)$ is a positive linear map of norm 1, and satisfies $T_1 = \mathbf{1}$.

Proposition 4.3.1. *Let $f, g \in L^\infty$. If one of the functions f, g is continuous, then $T_{fg} - T_f T_g \in \mathcal{K}$.*

Proof. Since $T_{fg}^* = T_{\bar f \bar g}$ and $(T_f T_g)^* = T_{\bar g} T_{\bar f}$, it suffices to prove the following assertion: If $f \in C(\mathbb{T})$ and $g \in L^\infty$, then $T_{fg} - T_f T_g \in \mathcal{K}$. Moreover, since $C(\mathbb{T})$ is the norm-closed linear span of the monomials ζ^n, $n \in \mathbb{Z}$, and $\mathcal{K}$ is a norm-closed linear space, we may reduce to the case $f = \zeta^n$ and $g \in L^\infty$, $n \in \mathbb{Z}$.

If $n \geq 0$, then $\zeta^n \in H^\infty$, so that by (4.3.1) we have $T_{f\zeta^n} = T_f T_{\zeta^n}$. Thus $T_{fg} - T_f T_g = 0$ in this case.

If $n < 0$, say $n = -m$ with $m \geq 1$, then ζ^n is the complex conjugate of the H^∞ function ζ^m, and another application of (4.3.1) gives $T_{f\zeta^n} = T_{\zeta^n}T_f = S^{*m}T_f$. Noting that $S^{*m}T_fS^m = T_f$ (by iterating the basic formula $S^*T_fS = T_f$ valid for any Toeplitz operator) we can write

$$T_fT_{\zeta^n} = T_fS^{*m} = S^{*m}T_fS^mS^{*m} = S^{*m}T_f - S^{*m}T_f(\mathbf{1} - S^mS^{*m}).$$

Hence

$$T_{f\zeta^n} - T_fT_{\zeta^n} = S^{*m}T_f - T_fS^{*m} = -S^{*m}T_f(\mathbf{1} - S^mS^{*m}),$$

which is a finite-rank operator, since $\mathbf{1} - S^mS^{*m}$ is the projection onto $[1, \zeta, \zeta^2, \dots, \zeta^{m-1}]$. □

THEOREM 4.3.2. *The Toeplitz C^*-algebra $\mathcal{T} = C^*(S)$ consists of all operators of the form $T_f + K$, where $f \in C(\mathbb{T})$ and K is compact. Moreover, this decomposition is unique: For $f, g \in C(\mathbb{T})$ and $K, L \in \mathcal{K}$,*

$$T_f + K = T_g + L \implies f = g \quad \text{and} \quad K = L.$$

PROOF. We claim first that the set of operators

$$\mathcal{A} = \{T_f + K : f \in C(\mathbb{T}), \quad K \in \mathcal{K}\}$$

is a C^*-algebra. To see this, consider the map $\rho : C(\mathbb{T}) \to \mathcal{B}(H^2)/\mathcal{K}$ given by

$$\rho(f) = T_f + \mathcal{K}, \qquad f \in C(\mathbb{T}).$$

This defines a self-adjoint linear mapping of $C(\mathbb{T})$ to the Calkin algebra. By Theorem 4.3.1, ρ is actually a homomorphism of C^*-algebras. By Theorem 2.10.4, $\rho(C(\mathbb{T}))$ is a C^*-subalgebra of the Calkin algebra; and the inverse image of this C^*-algebra under the natural projection $T \in \mathcal{B}(H^2) \mapsto \dot{T} \in \mathcal{B}(H^2)/\mathcal{K}$ is exactly $\mathcal{A}$.

Clearly, $\mathcal{A}$ contains $S = T_\zeta$, and hence $\mathcal{A} \supseteq \mathcal{T}$. On the other hand, for $n \geq 0$ we have $T_{\zeta^n} = S^n \in \mathcal{T}$, and for $n < 0$ we have $T_{\zeta^n} = S^{*|n|} \in \mathcal{T}$; thus $T_{\zeta^n} \in \mathcal{T}$ for all $n \in \mathbb{Z}$. Using Exercise (1) below, we see that $\mathcal{T}$ contains all compact operators, and thus

$$T_{\zeta^n} + K \in \mathcal{T}, \qquad n \in \mathbb{Z}, \quad K \in \mathcal{K}.$$

Since $C(\mathbb{T})$ is the norm-closed linear span of the set of functions $\{\zeta^n : n \in \mathbb{Z}\}$, it follows that $\mathcal{T}$ contains all operators $T_f + K$ with $f \in C(\mathbb{T})$, K compact. Hence $\mathcal{A} \subseteq \mathcal{T}$.

Finally, the uniqueness of the representation of operators as compact perturbations of Toeplitz operators is an obvious consequence of Corollary 1 of Theorem 4.2.4. □

REMARK 4.3.3. If we compose the linear map $f \in C(\mathbb{T}) \mapsto T_f \in \mathcal{T}$ with the natural homomorphism of $\mathcal{T}$ to the Calkin algebra, then we obtain an injective $*$-homomorphism $f \mapsto \dot{T}_f$ of $C(\mathbb{T})$ into the Calkin algebra. Using this map to identify $C(\mathbb{T})$ with the quotient $\mathcal{T}/\mathcal{K}$, we obtain a short exact sequence of C^*-algebras and $*$-homomorphisms

$$0 \longrightarrow \mathcal{K} \longrightarrow \mathcal{T} \underset{\pi}{\longrightarrow} C(\mathbb{T}) \longrightarrow 0, \tag{4.9}$$

π being the $*$-homomorphism of $\mathcal{T}$ to $C(\mathbb{T})$ given by $\pi(T_f + K) = f$, $f \in C(\mathbb{T})$, $K \in \mathcal{K}$. The sequence (4.9) is called the *Toeplitz extension* of $\mathcal{K}$ by $C(\mathbb{T})$. The Toeplitz extension is *semisplit* in the sense that there is a natural positive linear map $\phi : C(\mathbb{T}) \to \mathcal{T}$, such that $\phi(1) = \mathbf{1}$, with the property that $\pi \circ \phi$ is the identity map of $C(\mathbb{T})$ (namely, $\phi(f) = T_f$). It is significant that ϕ is not a $*$-homomorphism but rather a positive linear map. Indeed, we will see later that this extension is not split; more explicitly, there does not exist a *$*$-homomorphism* $\theta : C(\mathbb{T}) \to \mathcal{T}$ with the property that $\pi \circ \theta$ is the identity map of $C(\mathbb{T})$. The nonexistence of a splitting homomorphism θ has to do with the Fredholm index (see Exercise (4) below).

We immediately obtain the following description of the Fredholm operators in $\mathcal{T}$:

COROLLARY 1. The Fredholm operators in $\mathcal{T}$ are precisely the operators of the form $T_f + K$ where f is an invertible symbol in $C(\mathbb{T})^{-1}$ and $K \in \mathcal{K}$.

Consider a Fredholm operator in $\mathcal{T}$, say $T_f + K$ where $f \in C(\mathbb{T})$ has no zeros on the circle and K is a compact operator. By the stability results of Chapter 3 we see that T_f is also a Fredholm operator and

$$\operatorname{ind}(T_f + K) = \operatorname{ind}(T_f).$$

We know that for $f = \zeta$, T_f is the shift; hence $\operatorname{ind}(T_f) = -1$. However, we still lack tools for computing the index of more general Toeplitz operators with symbols in $C(\mathbb{T})^{-1}$. This issue will be taken up in the following section.

Exercises. Let $e_0, e_1, \dots$ be an orthonormal basis for a Hilbert space H, and realize the unilateral shift S as the unique operator on H satisfying $Se_n = e_{n+1}$, $n \geq 0$. Let $\mathcal{T} = C^*(S)$ be the C^*-algebra generated by S.

(1) Show that for every $m, n \geq 0$,

$$S^m S^{*n} - S^{m+1} S^{*(n+1)} = S^m(\mathbf{1} - SS^*)S^{*n}$$

is a rank-one operator and describe this operator in terms of its action on $e_0, e_1, \dots$. Deduce that $\mathcal{T}$ contains the C^*-algebra $\mathcal{K}$ of all compact operators on H.

(2) Noting that $\mathcal{K}$ is a closed ideal in $\mathcal{T}$, identify the quotient C^*-algebra by showing that there is a unique $*$-isomorphism $\sigma : \mathcal{T}/\mathcal{K} \to C(\mathbb{T})$ that satisfies $\sigma(S + \mathcal{K}) = \zeta$, where ζ is the current variable in $C(\mathbb{T})$, $\zeta(z) = z$ for all $z \in \mathbb{T}$. Hint: Show that the image of S in the Calkin algebra is a unitary operator whose spectrum is $\mathbb{T}$.

(3) Let K be another Hilbert space, and let W be a unitary operator in $\mathcal{B}(K)$. Deduce that there is a unique representation $\pi : \mathcal{T} \to \mathcal{B}(K)$ such that $\pi(S) = W$.

(4) Let T be a Fredholm operator acting on a Hilbert space.
   (a) Assuming that the index of T is nonzero, show that T cannot be decomposed $T = N + K$ into a compact perturbation of a normal operator N.
   (b) Deduce that the unilateral shift is not a compact perturbation of a unitary operator and that the Toeplitz extension (4.9) is not split.

In the following exercises, V denotes an arbitrary isometry acting on some (separable) Hilbert space K. The subspaces V^nK decrease with n, and V is called a *pure* isometry if $\cap_n V^nK = \{0\}$. A closed subspace $M \subseteq K$ is said to be *reducing* for V if it is invariant under both V and V^*. The (self-adjoint) projections onto reducing subspaces are the projections in $\mathcal{B}(K)$ that commute with V.

(5) Show that for every isometry $V \in \mathcal{B}(K)$ there is a unique decomposition of K into reducing subspaces $K = L \oplus M$, where the restriction of V to L is a pure isometry and the restriction of V to M is a unitary operator. Hint: Let $N = (VK)^\perp$ be the orthogonal complement of the range of V. Show that $V^pN \perp V^qN$ if $p \neq q$ and $N \oplus VN \oplus V^2N \oplus \cdots$ is the orthocomplement of $M = VK \cap V^2K \cap V^3K \cap \cdots$.
(6) Show that the restriction of V to the "pure" subspace L is unitarily equivalent to a (finite or infinite) direct sum $S \oplus S \oplus \cdots$ of copies of the shift S, and that the number of copies is the dimension of N.

The result of Exercises (5) and (6) asserts that every isometry decomposes uniquely into a direct sum of two operators, one of which is a multiple copy of the unilateral shift, the other being a unitary operator. This is called the Wold decomposition of an isometry V, after the statistician who discovered the result in connection with the theory of stationary Gaussian processes. The following result is due to Lewis Coburn (1968), and should be compared with the result of Exercise (2). It implies that the Toeplitz C^*-algebra is universal for all C^*-algebras generated by isometries.

(7) For every isometry V acting on a Hilbert space K, there is a unique representation $\pi : \mathcal{T} \to \mathcal{B}(K)$ such that $\pi(S) = V$. Hint: Use the Wold decomposition.

The result of Exercise (7) is sometimes formulated in purely $*$-algebraic terms as follows. Let A be a C^*-algebra with unit and let v be an element of A. Then the following are equivalent:

- There is a (necessarily unique) $*$-homomorphism $\pi : \mathcal{T} \to A$ such that $\pi(S) = v$.
- $v^*v = \mathbf{1}$.

The difference between Exercise (7) and this more abstract uniqueness result involves the Gelfand–Naimark theorem, which asserts that every abstract C^*-algebra has a nondegenerate isometric representation as a C^*-algebra of operators on some Hilbert space. The Gelfand–Naimark theorem will be established in Section 4.8 below.

4.4. Index Theorem for Continuous Symbols

Consider the multiplicative group $G = C(\mathbb{T})^{-1}$ of all complex-valued continuous functions on the circle that have no zeros. G is a commutative topological group relative to its norm topology. We seek a nontrivial homomorphism of G into the additive group $\mathbb{Z}$. This homomorphism is a generalization of the winding number, about the origin, of piecewise smooth functions in G. We first describe this generalized winding number in some detail. Then we relate this topological invariant of functions $f \in G$ to the index of their Toeplitz operators $T_f \in \mathcal{B}(H^2)$. Throughout, $\mathbb{C}^\times$ denotes the multiplicative group of nonzero complex numbers.

We begin with a result about the general linear group of a related C^*-algebra $C[0,1]$. While one can base that result on the fact that $[0,1]$ is a contractible space, or on the properties of covering maps of spaces, the argument we give uses only elementary methods. The reader should keep in mind that the range of a function $f \in C[0,1]^{-1}$ can be very complicated, perhaps having nontrivial interior.

PROPOSITION 4.4.1. *For every function $F \in C[0,1]$ such that $F(t) \neq 0$ for every $t \in [0,1]$, there is a function $G \in C[0,1]$ such that*

$$F(t) = e^{G(t)}, \qquad 0 \leq t \leq 1.$$

PROOF. On the domain $\{z \in \mathbb{C} : |z-1| < 1\}$, let $\log z$ be the principal branch of the logarithm,

$$\log z = -\sum_{n=1}^{\infty} \frac{(1-z)^n}{n}.$$

The log function is holomorphic, satisfies $\log 1 = 0$, and of course $e^{\log z} = z$ for $|z-1| < 1$. Let

$$M = \sup_{0 \leq t \leq 1} |F(t)|^{-1}.$$

By uniform continuity of F, we can find a finite partition of the interval $[0,1]$, $0 = t_0 < t_1 < \cdots < t_n = 1$, such that

$$\sup_{t_{k-1} \leq t \leq t_k} |F(t) - F(t_{k-1})| < \frac{1}{2M}.$$

It follows that for $k = 1, \dots, n$ and $t \in [t_{k-1}, t_k]$,

$$\left|1 - \frac{F(t)}{F(t_{k-1})}\right| = \frac{|F(t) - F(t_{k-1})|}{|F(t_{k-1})|} \leq \frac{1}{2M|F(t_{k-1})|} \leq \frac{1}{2} < 1. \tag{4.10}$$

Setting

$$G_k(t) = \log(F(t)/F(t_{k-1})), \qquad t_{k-1} \le t \le t_k,$$

we find that G_k is continuous, $G_k(t_{k-1}) = 0$, $G_k(t_k) = \log(F(t_k)/F(t_{k-1}))$, and it satisfies

$$F(t) = F(t_{k-1})e^{G_k(t)}$$

throughout the interval $[t_{k-1}, t_k]$. There is an obvious way to piece the G_k together so as to obtain a continuous function $G : [0,1] \to \mathbb{C}$, namely $G(t) = G_1(t)$ for $t \in [0, t_1]$ and, for $k = 2, \dots, n$,

$$G(t) = G_1(t_1) + \cdots + G_{k-1}(t_{k-1}) + G_k(t), \qquad t \in [t_{k-1}, t_k].$$

It follows that

$$F(t) = F(0)e^{G(t)}, \qquad 0 \le t \le 1.$$

Writing $F(0) \in \mathbb{C}^\times$ as an exponential $F(0) = e^{z_0}$, we obtain a continuous function $\tilde{G}$ satisfying $F = e^{\tilde{G}}$ by way of $\tilde{G}(t) = G(t) + z_0$. $\square$

We can now define the winding number (about the origin) of a function $f \in G = C(\mathbb{T})^{-1}$. The function $F : [0,1] \to \mathbb{C}^\times$ defined by

$$F(t) = f(e^{2\pi i t})$$

is continuous, and hence by Proposition 4.4.1 there is a continuous function $G : [0,1] \to \mathbb{C}$ such that

$$f(e^{2\pi i t}) = e^{2\pi i G(t)}, \qquad 0 \le t \le 1. \tag{4.11}$$

Note that $G(1) - G(0) \in \mathbb{Z}$ because $e^{2\pi i G(1)} = e^{2\pi i G(0)}$. The function G is not uniquely determined, but if $\tilde{G}$ is another such, then $G - \tilde{G}$ is a continuous function with

$$e^{2\pi i (G(t) - \tilde{G}(t))} = f(e^{2\pi i t})/f(e^{2\pi i t}) = 1, \qquad 0 \le t \le 1,$$

and hence $G(t)$ and $\tilde{G}(t)$ differ by a constant. It follows that for any choice of a continuous function G satisfying (4.11),

$$\#(f) = G(1) - G(0) \tag{4.12}$$

is a well-defined integer. This integer is called the *winding number* of f. The properties of this generalized winding number are summarized as follows:

PROPOSITION 4.4.2. *For $f, g \in G = C(\mathbb{T})^{-1}$,*

(1) $\#(fg) = \#(f) + \#(g)$.
(2) $\#(f) = n \in \mathbb{Z}$ *iff there is a function $h \in C(\mathbb{T})$ such that $f = \zeta^n e^h$.*

PROOF. For (1), pick continuous functions $F, G : [0,1] \to \mathbb{C}$ such that

$$f(e^{2\pi i t}) = e^{2\pi i F(t)}, \qquad g(e^{2\pi i t}) = e^{2\pi i G(t)}, \qquad t \in [0,1].$$

Then

$$f(e^{2\pi i t})g(e^{2\pi i t}) = e^{2\pi i (F(t) + G(t))}, \qquad t \in [0,1],$$

and the winding number of fg is given by

$$F(1) + G(1) - (F(0) + G(0)) = \#(f) + \#(g).$$

For (2), consider first the case $n = 0$. If $f = e^h$ is the exponential of a function $h \in C(\mathbb{T})$, then we have

$$f(e^{2\pi it}) = e^{2\pi iF(t)}, \qquad 0 \le t \le 1, \tag{4.13}$$

where $F(t) = (2\pi)^{-1}h(e^{2\pi it})$. Clearly, $F(1) = F(0)$, so that $\#(f) = F(1) - F(0) = 0$. Conversely, if $\#(f) = 0$, then there is a function $F \in C[0,1]$ such that (4.13) is satisfied and $F(1) - F(0) = \#(f) = 0$. Since F is periodic, we have $f = e^h$, where $h \in C(\mathbb{T})$ is the function $h(e^{2\pi it}) = 2\pi iF(t)$, $0 \le t \le 1$.

To deal with the case of arbitrary $n \in \mathbb{Z}$ note first that $\#(\zeta) = 1$. Indeed, this is immediate from the fact that

$$\zeta(e^{2\pi it}) = e^{2\pi it}, \qquad 0 \le t \le 1.$$

From the property (1) it follows that $\#(\zeta^n) = n$ for every $n \in \mathbb{Z}$; hence $\#(\zeta^n e^h) = \#(\zeta^n) + \#(e^h) = n$, as asserted. Conversely, if $f \in C(\mathbb{T})$ satisfies $\#(f) = n$, consider $g = \zeta^{-n}f \in C(\mathbb{T})^{-1}$. Using (1) again we have $\#(g) = 0$, and by the preceding paragraph there is an $h \in C(\mathbb{T})$ such that $g = e^h$. Thus $f = \zeta^n g = \zeta^n e^h$ has the asserted form. □

We now complete the computation of the index of Fredholm operators in the Toeplitz C^*-algebra $\mathcal{T}$:

THEOREM 4.4.3. *For every $f \in G = C(\mathbb{T})^{-1}$,*

$$\operatorname{ind} T_f = -\#(f).$$

PROOF. In view of Proposition 4.4.2, it suffices to show that for $f = \zeta^n e^g$ with $n \in \mathbb{Z}$ and $g \in C(\mathbb{T})$ we have $\operatorname{ind} T_f = -n$.

We claim first that $\operatorname{ind} T_{e^g} = 0$. Indeed,

$$A_\lambda = T_{e^{\lambda g}}, \qquad 0 \le \lambda \le 1,$$

defines a continuous arc of Fredholm operators in $\mathcal{B}(H^2)$ satisfying $\operatorname{ind} A_0 = \operatorname{ind} \mathbf{1} = 0$ and $\operatorname{ind} A_1 = \operatorname{ind} T_{e^g}$. By continuity of the index, we must have $\operatorname{ind} A_1 = \operatorname{ind} A_0 = 0$.

Notice that the map $f \in G = C(\mathbb{T})^{-1} \mapsto \operatorname{ind} T_f \in \mathbb{Z}$ is a homomorphism of abelian groups. For fixed $f, g \in G$ Proposition 4.3.1 implies that $T_{fg} = T_f T_g + K$ for some compact operator K. Hence

$$\operatorname{ind} T_{fg} = \operatorname{ind}(T_{fg} + K) = \operatorname{ind}(T_f T_g) = \operatorname{ind} T_f + \operatorname{ind} T_g$$

by the stability and additivity properties of the index. Finally, since T_ζ is the unilateral shift, its index is -1; hence

$$\operatorname{ind} T_f = \operatorname{ind} T_{\zeta^n e^g} = \operatorname{ind} T_{\zeta^n} + \operatorname{ind} T_{e^g} = n \cdot \operatorname{ind} T_\zeta = -n.$$

□

Exercises. In Exercises (1) through (5), $\{a_n : n \in \mathbb{Z}\}$ denotes a doubly infinite sequence of complex numbers that is summable, $\sum_n |a_n| < \infty$, and ϕ is the continuous function defined on the unit circle by

$$\phi(z) = \sum_{n=-\infty}^{\infty} a_n z^n, \qquad z = e^{i\theta}, \quad 0 \leq \theta \leq 2\pi.$$

As usual, $\mathbb{Z}$ denotes the additive group of integers, and $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ denotes the additive semigroup of nonnegative integers.

(1) Consider the Hilbert space $H = \ell^2(\mathbb{Z})$. Show that the convolution operator A defined by

$$(A\xi)_n = \sum_{k=-\infty}^{\infty} a_{n-k}\xi_k = \sum_{j=-\infty}^{\infty} a_j \xi_{n-j}$$

is bounded, and in fact, $\|A\| \leq \sum_n |a_n|$. Labor-saving hint: Realize A appropriately as $\sum_m a_m T^m$ where T is a translation operator.

(2) Show that A is a normal operator by calculating A^*, AA^*, and A^*A.

(3) Determine the spectrum of A in concrete terms. Hint: A is unitarily equivalent to a multiplication operator on some Hilbert space of functions $L^2(X, \mu)$: What is the multiplication operator?

(4) Assuming that $\{a_n\}$ is not a trivial sequence satisfying $a_n = 0$ for all $n \neq 0$, deduce that A has no point spectrum (i.e., no eigenvalues), determine when it is invertible in terms of ϕ, and calculate $\|A\|$ exactly.

In Exercises (5) and (6), you will consider a related operator B, defined on the subspace $K = \ell^2(\mathbb{Z}^+) \subset H$ by

$$(B\xi)_n = \sum_{k=0}^{\infty} a_{n-k}\xi_k = \sum_{j=-\infty}^{n} a_j \xi_{n-j}, \tag{4.14}$$

for $n = 0, 1, 2, \dots, \quad \xi \in K$.

(5) Show that $B^*B - BB^*$ is compact, and show that the essential spectrum of B is the spectrum of A.

(6) Specialize the operator B in (4.14) as follows: $(B\xi)_n = \xi_{n-1} - \xi_{n-2}$ for $n \geq 2$, $(B\xi)_1 = \xi_0$, $(B\xi)_0 = 0$. Sketch the essential spectrum $\sigma_e(B)$ of B and calculate the Fredholm index of $B - \lambda\mathbf{1}$ for all $\lambda \in \mathbb{C} \setminus \sigma_e(B)$. Give a clear sketch with an indication of the various values of the index; it may help to indicate the points where $\sigma_e(B)$ meets the x-axis and the y-axis. Precise numerical computations are unnecessary, provided that you have a clear picture and good qualitative remarks.

4.5. Some H^2 Function Theory

In this section, we present several results connecting the function theory and the operator theory of the Hardy space H^2. The results are important for many aspects of functional analysis, including but certainly not limited to the computations of operator spectra that we will carry out in the next section.

We begin with a result characterizing the (closed) subspaces of H^2 that are invariant under the unilateral shift. This is a famous result of Arne Beurling [**6**]; it is remarkable because there are very few operators whose invariant subspaces are completely known. Indeed, it is not even known whether an arbitrary operator on a (separable) Hilbert space H must have a closed invariant subspace other than the trivial ones $\{0\}$ and H.

An *inner function* is a function $f \in H^\infty$ satisfying $|f(e^{i\theta})| = 1$ almost everywhere on the unit circle. The term "inner" has classical origins, and refers to the fact that if f is a rational function of a complex variable whose restriction to the unit circle has no poles and defines an inner function as above, then the zeros of f are all contained in the interior of the unit disk $\{z : |z| < 1\}$. Such rational functions are important in linear systems theory (they correspond to "causal" filters), and in the prediction theory of stationary Gaussian random processes.

For every function $f \in H^\infty$, the multiplication operator M_f carries H^2 into itself, $f \cdot H^2 \subseteq H^2$ by Proposition 4.2.1; and if f is an inner function, then $M = f \cdot H^2$ is a closed subspace of H^2 that is invariant under the unilateral shift $T_\zeta = M_\zeta \restriction_{H^2}$.

THEOREM 4.5.1 (Beurling). *For every closed shift-invariant subspace $M \subseteq H^2$ there is an inner function v such that $M = v \cdot H^2$.*

A complete proof of Beurling's theorem is outlined in the exercises at the end of the section. The following consequence is a classical theorem of the brothers Riesz, whose original method was quite different. It has attracted a great deal of attention over the years, and far-reaching generalizations have been discovered that relate to diverse areas, including (a) effective generalizations of H^p theory that can be formulated whenever one has a flow acting on a space [**18**], (b) the theory of one-parameter groups of automorphisms of von Neumann algebras that satisfy a "positive energy" condition [**4**], and (c) the properties of annihilating measures of abstract function algebras [**13**].

THEOREM 4.5.2 (F. and M. Riesz). *The set $Z = \{z \in \mathbb{T} : f(z) = 0\}$ of zeros of any nonzero function $f \in H^2$ is a set of Lebesgue measure* 0.

PROOF. Fix a function $f \neq 0$ in H^2 and consider the closed subspace $M = [f, \zeta f, \zeta^2 f, \dots]$ of H^2. Then $M \neq \{0\}$, it is invariant under the shift T_ζ, and every function in M vanishes almost everywhere on the zero set Z. Beurling's Theorem 4.5.1 implies that M contains an inner function v. Since $|v(e^{i\theta})| = 1$ almost everywhere on $\mathbb{T}$, Z must have measure 0. $\square$

REMARK 4.5.3. *Some remarks on H^1.* We collect some details relating to the function theory of H^1 that will be used in the proof of the following theorem: H^1 is defined as the space of all functions $f \in L^1$ whose Fourier series has the form

$$f(e^{i\theta}) \sim \sum_{n=0}^{\infty} a_n e^{in\theta}. \tag{4.15}$$

If $g \in H^1$ is such that its conjugate $\bar{g}$ also belongs to H^1, then g must be a constant. Indeed, (4.15) implies that all the negative Fourier coefficients of g are zero, while $\bar{g} \in H^1$ implies that the positive coefficients of g are zero. Hence the Fourier series of g is the Fourier series of a constant function, and g must be a constant. Let H^1_0 denote the space of all functions f in H^1 with zero constant term, $\langle f, 1\rangle = 0$. We may conclude from these remarks that

$$H^1 \cap \overline{H^1} = \mathbb{C} \cdot \mathbf{1} \qquad \text{and} \qquad H^1 \cap \overline{H^1_0} = \{0\}.$$

Second, we point out that the product of two functions in H^2 must belong to H^1. Indeed, if $f, g \in H^2$, then $fg \in L^1$, and moreover, $\|fg\|_1 \le \|f\|_2 \|g\|_2$. Thus for a fixed negative integer n the Fourier coefficient $\langle fg, \zeta^n\rangle$ defines a bounded bilinear functional on $H^2 \times H^2$ that vanishes whenever f and g are finite sums of the form $a_0 + a_1\zeta + \cdots + a_p\zeta^p$. It follows that $\langle fg, \zeta^n\rangle = 0$ identically on $H^2 \times H^2$. We conclude that the Fourier series of fg has the required form (4.15).

Finally, let $H^2_0 = \{f \in H^2 : \langle f, 1\rangle = 0\}$. Then $H^2_0 = [\zeta, \zeta^2, \zeta^3, \dots]$; hence the orthocomplement of H^2 in L^2 is related to H^2_0 by

$$H^{2\perp} = \overline{H^2_0} = \{\bar{f} : f \in H^2_0\}.$$

The following result of Lewis Coburn [**7**] implies that when a Toeplitz operator is a *Fredholm* operator of index zero, it must be invertible:

THEOREM 4.5.4 (Coburn). *Let ϕ be any nonzero symbol in L^∞. Then either* $\ker T_\phi = \{0\}$ *or* $\ker T^*_\phi = \{0\}$.

PROOF. We show that if both kernels are nontrivial, then $\phi = 0$. For that, choose nonzero functions $f, g \in H^2$ such that $T_\phi f = T^*_\phi g = 0$. With $P_+ \in \mathcal{B}(H^2)$ denoting the projection onto H^2, we have $P_+\phi f = 0$, i.e., $\phi f \in H^{2\perp} = \overline{H^2_0}$. Therefore,

$$\bar{\phi}\bar{f} \in H^2_0. \tag{4.16}$$

Similarly, $T^*_\phi g = 0$ implies that $P_+\bar{\phi} g = 0$, that is, $\bar{\phi} g \in \overline{H^2_0}$. Therefore,

$$\phi\bar{g} \in H^2_0. \tag{4.17}$$

Multiplying the term of (4.16) by g, we obtain

$$\bar{\phi}\bar{f}g \in H^2_0 \cdot H^2 \subseteq H^1_0,$$

by Remark 4.5.3. On the other hand, multiplying the term of (4.17) by f gives

$$\overline{\bar{\phi}\bar{f}g} = \phi\bar{g}f \in H^2_0 \cdot H^2 \subseteq H^1_0.$$

Thus $\bar{\phi}\bar{f}g \in H^1_0 \cap \overline{H^1_0} = \{0\}$. Since neither f nor g is the zero function, the F. and M. Riesz theorem implies that the product $\bar{f}(z)g(z)$ is nonzero for almost every $z \in \mathbb{T}$. Thus $\phi\bar{g}f = 0$ implies that $\bar{\phi}(z)$ vanishes almost everywhere. $\square$

Exercises. In these exercises, you will deduce Beurling's theorem from the following more general result, which characterizes certain subspaces of L^2 that are invariant under the unitary multiplication operator $U = M_\zeta \in \mathcal{B}(L^2)$. Notice that for any such subspace M, the sequence of subspaces $U^n M$ decreases with n.

Theorem A. Let $M \subseteq L^2$ be a nonzero closed U-invariant subspace of L^2 that is pure in the sense that $\cap_{n\geq 0} U^n M = \{0\}$. There is a function $v \in L^\infty$ such that $|v(e^{i\theta})| = 1$ almost everywhere on the unit circle and $M = v \cdot H^2$.

For the following exercises, let $M \subseteq L^2$ be a nonzero closed subspace satisfying the hypotheses of Theorem A.

(1) Let $N = M \ominus UM$ be the orthocomplement of UM in M. Show that $N \neq \{0\}$ and that it is a *wandering* subspace in the sense that for $m, n \in \mathbb{Z}$ with $m \neq n$ we have $U^m N \perp U^n N$.
(2) For every operator $A \in \mathcal{B}(N)$ define $\bar{A} \in \mathcal{B}(L^2)$ by

$$\bar{A} = \sum_{n=-\infty}^{\infty} U^n A P_N U^{-n},$$

$AP_N \in \mathcal{B}(L^2)$ denoting the composition of A with the projection onto N. Show that $\bar{A}$ belongs to the multiplication algebra $\mathcal{M} = \{M_f : f \in L^\infty\}$.
(3) Deduce that $\mathcal{B}(N)$ is abelian, hence N must be one-dimensional.

Choose an element $v \in N$ with $\|v\|_{L^2} = 1$.

(4) Show that for every $m, n \in \mathbb{Z}$ with $m \neq n$ one has $\langle v \cdot z^m, v \cdot z^n \rangle = 0$, and deduce that $|v(e^{i\theta})| = 1$ almost everywhere on the unit circle.
(5) Show that M is spanned by $N, UN, U^2N, \ldots$ and deduce that $M = v \cdot H^2$.

That completes the proof of Theorem A.

(6) Deduce Beurling's theorem from Theorem A.

4.6. Spectra of Toeplitz Operators with Continuous Symbol

Given a continuous symbol $f \in C(\mathbb{T})$, we are now in position to give a description of $\sigma(T_f)$. Let us first consider the essential spectrum $\sigma_e(T_f) \subseteq$

$\sigma(T_f)$. By the exact sequence (4.9) the essential spectrum of T_f is the spectrum of f as an element of the commutative C^*-algebra $C(\mathbb{T})$, namely,

$$\sigma_e(T_f) = f(\mathbb{T}). \tag{4.18}$$

What remains is to determine the other points of the spectrum. Let us decompose $\mathbb{C} \setminus f(\mathbb{T})$ into its connected components, obtaining an unbounded component Ω_∞ together with a finite, infinite, or possibly empty set of holes $\Omega_1, \Omega_2, \ldots$,

$$\mathbb{C} \setminus f(\mathbb{T}) = \Omega_\infty \sqcup \Omega_1 \sqcup \Omega_2 \sqcup \cdots .$$

Choose $\lambda \in \mathbb{C} \setminus f(\mathbb{T})$. Then $f - \lambda \in C(\mathbb{T})^{-1}$, and hence $T_f - \lambda = T_{f-\lambda}$ is a Fredholm operator. Consider the behavior of $\operatorname{ind}(T_f - \lambda)$ as λ varies over one of the components Ω_k of $\mathbb{C} \setminus f(\mathbb{T})$. Since $\lambda \mapsto T_f - \lambda$ is a continuous function from Ω_k to the set of Fredholm operators on H^2 and since the index is continuous, it follows that $\operatorname{ind}(T_f - \lambda)$ is constant over Ω_k. Let $n_k \in \mathbb{Z}$ be this integer, $k = \infty, 1, 2, \ldots$.

Obviously, $n_\infty = 0$ because $T_f - \lambda$ is invertible for sufficiently large λ (for example, when $|\lambda| > \|T_f\|$). When holes exist, n_k can take on any integral value for $k = 1, 2, \ldots$. In such cases Theorem 4.4.3 allows us to evaluate n_k

$$n_k = \operatorname{ind}(T_f - \lambda) = \operatorname{ind}(T_{(f-\lambda)}) = -\#(f - \lambda),$$

in terms of the generalized winding number of the symbol f about λ. Thus we have calculated $\operatorname{ind}(T_f - \lambda)$ throughout the complement of $f(\mathbb{T})$.

If k is such that $n_k \neq 0$, then $T_f - \lambda$ is a Fredholm operator of nonzero index for all $\lambda \in \Omega_k$. Obviously, such operators cannot be invertible; hence $\Omega_k \subseteq \sigma(T_f)$. On the other hand, if $n_k = 0$, then $T_{f-\lambda}$ is a Fredholm operator of index zero for all $\lambda \in \Omega_k$. By Theorem 4.5.4 such operators must be invertible; hence Ω_k is disjoint from $\sigma(T_f)$. We assemble these remarks about Toeplitz operators with continuous symbol into the following description of their spectra.

THEOREM 4.6.1. *Let $f \in C(\mathbb{T})$, and let $\mathbb{C} \setminus f(\mathbb{T}) = \Omega_\infty \sqcup \Omega_1 \sqcup \Omega_2 \sqcup \cdots$ be the decomposition of the complement of $f(\mathbb{T})$ into its unbounded component Ω_∞ and holes Ω_k, $k \geq 1$. For each finite k and $\lambda \in \Omega_k$, the winding number $w_k = \#(f - \lambda)$ is a constant independent of λ.*

The spectrum of T_f is the union of $f(\mathbb{T})$ and the holes Ω_k for which $w_k \neq 0$.

In particular, the spectrum of a Toeplitz operator with continuous symbol contains no isolated points, and is in fact a connected set. The problem of giving a similarly detailed description of the spectra of Toeplitz operators with symbol in L^∞ remains open in general. However, a theorem of Harold Widom asserts that $\sigma(T_f)$ is connected for every $f \in L^\infty$ (see [**11**]). The case of self-adjoint Toeplitz operators is treated in the Exercises below.

Exercises.

(1) Let $\phi \in L^\infty$, and consider its associated multiplication operator $M_\phi \in \mathcal{B}(L^2)$ and Toeplitz operator $T_\phi \in \mathcal{B}(H^2)$.

(a) Given $\epsilon > 0$ such that $\|T_\phi f\| \geq \epsilon\|f\|$ for all $f \in H^2$, show that $\|M_\phi g\| \geq \epsilon\|g\|$ for all $g \in L^2$. Hint: The union of the spaces $\zeta^n H^2$, $n \leq 0$, is dense in L^2.
(b) Prove: If T_ϕ is invertible, then M_ϕ is invertible.
(c) Deduce the spectral inclusion theorem of Hartman and Wintner: For $\phi \in L^\infty$, $\sigma(T_\phi)$ contains the essential range of ϕ.

Let ϕ be a *real-valued* function in L^∞ and let $m \leq M$ be the essential infimum and essential supremum of ϕ,

$$\begin{aligned} m &= \inf\{t \in \mathbb{R} : \sigma\{z \in \mathbb{T} : \phi(z) < t\} > 0\}, \\ M &= \sup\{t \in \mathbb{R} : \sigma\{z \in \mathbb{T} : \phi(z) > t\} > 0\}, \end{aligned}$$

σ denoting normalized Lebeggue measure on $\mathbb{T}$. Thus, $[m, M]$ is the smallest closed interval $I \subseteq \mathbb{R}$ with the property that $\phi(z) \in I$ almost everywhere $d\sigma(z)$. Equivalently, it is the smallest interval containing the essential range of ϕ. In the remaining exercises you will obtain information about the spectrum of the self-adjoint Toeplitz operator T_ϕ.

(2) Let λ be a real number such that $T_\phi - \lambda$ is invertible. Show that there is a nonzero function $f \in H^2$ such that $T_\phi f - \lambda f = 1$, 1 denoting the constant function in H^2.
(3) Show that $(\phi - \lambda)|f|^2 = (\phi - \lambda)\bar{f} \cdot f$ belongs to H^1 and deduce that there is a real number c such that $(\phi(z) - \lambda)|f(z)|^2 = c$ for σ-almost every $z \in \mathbb{T}$.
(4) Deduce that $\phi(z) - \lambda$ is either positive almost everywhere on $\mathbb{T}$ or negative almost everywhere on $\mathbb{T}$. Hint: Use the F. and M. Riesz theorem.
(5) Deduce the following theorem of Hartman and Wintner (1954): For every real-valued symbol $\phi \in L^\infty$,

$$\sigma(T_\phi) = [m, M],$$

m and M being the essential inf and essential sup of ϕ.

4.7. States and the GNS Construction

Throughout this section, A will denote a Banach $*$-algebra with normalized unit $\mathbf{1}$. A linear functional $\rho : A \to \mathbb{C}$ is said to be *positive* if $\rho(x^*x) \geq 0$ for every $x \in A$. A *state* is a positive linear functional satisfying $\rho(\mathbf{1}) = 1$. This terminology has its origins in the connections between C^*-algebras and quantum physics, an important subject that is not touched on here. Notice that we do not assume that states are bounded, but Proposition 4.7.1 below implies that this is the case. It is a fundamental result that starting with a state ρ of A, one can construct a nontrivial representation $\pi : A \to \mathcal{B}(H)$. This procedure is called the GNS construction after the three mathematicians, I.M. Gelfand, M.A. Naimark, and I.E. Segal, who introduced it. The purpose of this section is to discuss the GNS construction in the general context of unital Banach $*$-algebras. Applications to C^*-algebras will be taken up in Section 4.8.

PROPOSITION 4.7.1. *Every positive linear functional ρ on A satisfies the Schwarz inequality*

$$|\rho(y^*x)|^2 \leq \rho(x^*x)\rho(y^*y) \tag{4.19}$$

and moreover, $\|\rho\| = \rho(\mathbf{1})$. In particular, every state of A has norm 1.

PROOF. Considering A as a complex vector space,

$$x, y \in A \mapsto [x, y] = \rho(y^*x)$$

defines a sesquilinear form which is positive semidefinite in the sense that $[x, x] \geq 0$ for every x. The argument that establishes the Schwarz inequality for complex inner product spaces applies *verbatim* in this context, and we deduce (4.19) from $|[x, y]|^2 \leq [x, x][y, y]$.

Clearly, $\rho(\mathbf{1}) = \rho(\mathbf{1}^*\mathbf{1}) \geq 0$, and we claim that $\|\rho\| \leq \rho(\mathbf{1})$. Indeed, for every $x \in A$ the Schwarz inequality (4.19) implies

$$|\rho(x)|^2 = |\rho(\mathbf{1}^*x)| \leq \rho(x^*x)\rho(\mathbf{1}).$$

If, in addition, $\|x\| \leq 1$, then x^*x is a self-adjoint element in A of norm at most 1; consequently, $\mathbf{1} - x^*x$ must have a self-adjoint square root $y \in A$ (see Exercise (2b) below). It follows that $\rho(\mathbf{1} - x^*x) = \rho(y^2) \geq 0$, i.e., $0 \leq \rho(x^*x) \leq \rho(\mathbf{1})$. Substitution into the previous inequality gives $|\rho(x)|^2 \leq \rho(x^*x)\rho(\mathbf{1}) \leq \rho(\mathbf{1})^2$, and $\|\rho\| \leq \rho(\mathbf{1})$ follows. Since the inequality $\|\rho\| \geq \rho(\mathbf{1})$ is obvious, we conclude that $\|\rho\| = \rho(\mathbf{1})$. □

DEFINITION 4.7.2. Let ρ be a positive linear functional on a Banach $*$-algebra A. By a GNS pair for ρ we mean a pair (π, ξ) consisting of a representation π of A on a Hilbert space H and a vector $\xi \in H$ such that

(1) (Cyclicity) $\overline{\pi(A)}\xi = H$, and
(2) $\rho(x) = \langle \pi(x)\xi, \xi \rangle$, for every $x \in A$.

Two GNS pairs (π, ξ) and (π', ξ') are said to be equivalent if there is a unitary operator $W : H \to H'$ such that $W\xi = \xi'$ and $W\pi(x) = \pi'(x)W$, $x \in A$.

THEOREM 4.7.3. *Every positive linear functional ρ on a unital Banach $*$-algebra A has a GNS pair (π, ξ), and any two GNS pairs for ρ are equivalent.*

PROOF. Consider the set

$$N = \{a \in A : \rho(a^*a) = 0\}.$$

With fixed $a \in A$, the Schwarz inequality (4.19) implies that for every $x \in A$ we have $|\rho(x^*a)|^2 \leq \rho(a^*a)\rho(x^*x)$, from which it follows that $\rho(a^*a) = 0 \iff \rho(x^*a) = 0$ for every $x \in A$. Thus N is a left ideal: a linear subspace of A such that $A \cdot N \subseteq N$.

The sesquilinear form $x, y \in A \mapsto \rho(y^*x)$ promotes naturally to sesquilinear form $\langle \cdot, \cdot \rangle$ on the quotient space A/N via

$$\langle x + N, y + N \rangle = \rho(y^*x), \qquad x, y \in A,$$

and for every x we have

$$\langle x + N, x + N \rangle = \rho(x^*x) = 0 \implies x + N = 0.$$

Hence A/N becomes an inner product space. Its completion is a Hilbert space H, and there is a natural vector $\xi \in H$ defined by

$$\xi = \mathbf{1} + N.$$

It remains to define $\pi \in \operatorname{rep}(A, H)$, and this is done as follows. Since N is a left ideal, for every fixed $a \in A$ there is a linear operator $\pi(a)$ defined on A/N by $\pi(a)(x + N) = ax + N$, $x \in A$. Note first that

$$\langle \pi(a)\eta, \zeta \rangle = \langle \eta, \pi(a^*)\zeta \rangle, \tag{4.20}$$

for every pair of elements $\eta = y + N, \zeta = z + N \in A/N$. Indeed, the left side of (4.20) is $\rho(z^*ay)$, while the right side is $\rho((a^*z)^*y) = \rho(z^*ay)$, as asserted.

We claim next that for every $a \in A$, $\|\pi(a)\| \le \|a\|$, where $\pi(a)$ is viewed as an operator on the inner product space A/N. Indeed, if $\|a\| \le 1$, then for every $x \in A$ we have

$$\begin{aligned} \langle \pi(a)(x + N), \pi(a)(x + N) \rangle &= \langle ax + N, ax + N \rangle = \rho((ax)^*ax) \\ &= \rho(x^*a^*ax). \end{aligned} \tag{4.21}$$

Since a^*a is a self-adjoint element in the unit ball of A, we can find a self-adjoint square root y of $\mathbf{1} - a^*a$ (see Exercise (2b)). It follows that $x^*x - x^*a^*ax = x^*(\mathbf{1} - a^*a)x = x^*y^2x = (yx)^*yx$; hence

$$\rho(x^*x - x^*a^*ax) = \rho((yx)^*yx) \ge 0,$$

from which we conclude that $\rho(x^*a^*ax) \le \rho(x^*x)$. This provides an upper bound for the right side of (4.21), and we obtain

$$\langle \pi(a)(x + N), \pi(a)(x + N) \rangle \le \rho(x^*x) = \langle x + N, x + N \rangle.$$

It follows that $\|\pi(a)\| \le 1$ when $\|a\| \le 1$, and the claim is proved.

Thus, for each $a \in A$ we may extend $\pi(a)$ uniquely to a bounded operator on the completion H by taking the closure of its graph; and we denote the closure $\pi(a) \in \mathcal{B}(H)$ with the same notation. Note that (4.20) implies that $\langle \pi(a)\eta, \zeta \rangle = \langle \eta, \pi(a^*)\zeta \rangle$ for all $\eta, \zeta \in H$, and from this we deduce that $\pi(a^*) = \pi(a^*)$, $a \in A$. It is clear from the definition of π that $\pi(ab) = \pi(a)\pi(b)$ for $a, b \in A$; hence $\pi \in \operatorname{rep}(A, H)$.

Finally, note that (π, ξ) is a GNS pair for ρ. Indeed,

$$\pi(A)\xi = \pi(A)(\mathbf{1} + N) = \{a + N : a \in A\}$$

is obviously dense in H, and

$$\langle \pi(a)\xi, \xi \rangle = \langle a + N, \mathbf{1} + N \rangle = \rho(\mathbf{1}^*a) = \rho(a).$$

For the uniqueness assertion, let (π', ξ') be another GNS pair for ρ, $\pi' \in \mathrm{rep}(A, H')$. Notice that there is a unique linear isometry W_0 from the dense subspace $\pi(A)\xi$ onto $\pi'(A)\xi'$ defined by $W_0 : \pi(a)\xi \mapsto \pi'(a)\xi'$, simply because for all $a \in A$,

$$\langle \pi(a)\xi, \pi(a)\xi \rangle = \langle \pi(a^*a)\xi, \xi \rangle = \rho(a^*a) = \langle \pi'(a)\xi', \pi'(a)\xi' \rangle.$$

The isometry W_0 extends uniquely to a unitary operator $W : H \to H'$, and one verifies readily that $W\xi = \xi'$, and that $W\pi(a) = \pi'(a)W$ on the dense set of vectors $\pi(A)\xi \subseteq H$. It follows that (π, ξ) and (π', ξ') are equivalent. □

REMARK 4.7.4. Many important Banach $*$-algebras do not have units. For example, the group algebras $L^1(G)$ of locally compact groups fail to have units except when G is discrete. C^*-algebras such as $\mathcal{K}$ do not have units. But the most important examples of Banach $*$-algebras have "approximate units," and it is significant that there is an appropriate generalization of the GNS construction (Theorem 4.7.3) that applies to Banach $*$-algebras containing an approximate unit [**10**], [**2**].

Exercises.

(1) (a) Fix α in the interval $0 < \alpha < 1$. Show that the binomial series of $(1 - z)^\alpha$ has the form

$$(1 - z)^\alpha = 1 - \sum_{n=1}^{\infty} c_n z^n,$$

where $c_n > 0$ for $n = 1, 2, \ldots$.

(b) Deduce that

$$\sum_{n=1}^{\infty} c_n = 1.$$

(2) (a) Let A be a Banach algebra with normalized unit, and let $c_1, c_2, \ldots$ be the binomial coefficients of the preceding exercise for the parameter value $\alpha = \frac{1}{2}$. Show that for every element $x \in A$ satisfying $\|x\| \leq 1$, the series

$$\mathbf{1} - \sum_{n=1}^{\infty} c_n x^n$$

converges absolutely to an element $y \in A$ satisfying

$$y^2 = \mathbf{1} - x.$$

(b) Suppose in addition that A is a Banach $*$-algebra. Deduce that for every self-adjoint element x in the unit ball of A, $\mathbf{1} - x$ has a self-adjoint square root in A.

In the remaining exercises, $\Delta = \{z \in \mathbb{C} : |z| \leq 1\}$ denotes the closed unit disk and A denotes the disk algebra, consisting of all functions $f \in C(\Delta)$ that are analytic on the interior of Δ.

(3) (a) Show that the map $f \mapsto f^*$ defined by

$$f^*(z) = \overline{f(\bar{z})}, \qquad z \in \Delta,$$

makes A into a Banach $*$-algebra.

(b) For each $z \in \Delta$, let $\omega_z(f) = f(z)$, $f \in A$. Show that ω_z is a positive linear functional if and only if $z \in [-1, 1]$ is real.

(4) Let ρ be the linear functional defined on A by

$$\rho(f) = \int_0^1 f(x)\, dx.$$

(a) Show that ρ is a state.

(b) Calculate a GNS pair (π, ξ) for ρ in concrete terms as follows. Consider the Hilbert space $L^2[0, 1]$, and let $\xi \in L^2[0, 1]$ be the constant function $\xi(t) = 1$, $t \in [0, 1]$. Exhibit a representation π of A on $L^2[0, 1]$ such that (π, ξ) becomes a GNS pair for ρ.

(c) Show that π is faithful; that is, for $f \in A$ we have

$$\pi(f) = 0 \implies f = 0.$$

(d) Show that the closure of $\pi(A)$ in the weak operator topology is a maximal abelian von Neumann algebra.

4.8. Existence of States: The Gelfand–Naimark Theorem

Turning our attention to C^*-algebras, we now show that every unital C^*-algebra has an abundance of states. The GNS construction implies that every state is associated with a representation; these two principles combine to show that every unital C^*-algebra has an isometric representation as a concrete C^*-algebra of operators on some Hilbert space.

Let A be a unital C^*-algebra, fixed throughout. A *positive* element of A is a self-adjoint element with nonnegative spectrum, $\sigma(x) \subseteq [0, \infty)$. One writes $x \geq 0$. Notice that $x^2 \geq 0$ for every self-adjoint element $x \in A$. Indeed, one can compute $\sigma(x^2)$ relative to any unital C^*-subalgebra containing it, and if one uses the commutative C^*-algebra generated by x and $\mathbf{1}$, the result follows immediately from Theorem 2.2.4 and basic properties of the Gelfand map. Significantly, this argument does *not* imply that z^*z has nonnegative spectrum for nonnormal elements $z \in A$, and in fact, the proof that $z^*z \geq 0$ in general (Theorem 4.8.3) is the cornerstone of the Gelfand–Naimark theorem.

We let A^+ denote the set of all positive elements of A. It is clear that A^+ is closed under multiplication by nonnegative scalars, but it is not obvious that the sum of two positive elements is positive.

LEMMA 4.8.1. *If x, y are two positive elements of A, then $x+y$ is positive.*

PROOF. By replacing x, y with $\lambda x, \lambda y$ for an appropriately small positive number λ, we can assume that $\|x\| \leq 1$ and $\|y\| \leq 1$. This implies that both

x and y have their spectra in the unit interval $[0,1]$. Hence $\mathbf{1}-x$ and $\mathbf{1}-y$ have their spectra in

$$\{1-\lambda : \lambda \in [0,1]\} = [-1,0] \subseteq [-1,+1].$$

Since they are self-adjoint, their norms agree with their spectral radii, and we conclude that $\|\mathbf{1}-x\| \leq 1$ and $\|\mathbf{1}-y\| \leq 1$.

It suffices to show that $z = \frac{1}{2}(x+y)$ is positive. z is obviously self-adjoint and

$$\|\mathbf{1}-z\| = \|\frac{1}{2}(\mathbf{1}-x) + \frac{1}{2}(\mathbf{1}-y)\| \leq \frac{1}{2} + \frac{1}{2} = 1.$$

Hence

$$\sigma(z) \subseteq \{t \in \mathbb{R} : |1-t| \leq 1\} \subseteq [0,\infty).$$

□

LEMMA 4.8.2. *If $a \in A$ satisfies $\sigma(a^*a) \subseteq (-\infty,0]$, then $a = 0$.*

PROOF. If a, b are elements of any Banach algebra with unit, then the nonzero points of $\sigma(ab)$ and $\sigma(ba)$ are the same (see Exercises (3) and (4) of Section 1.2). It follows that $\sigma(aa^*) \subseteq (-\infty,0]$. From the preceding lemma we conclude that $\sigma(a^*a + aa^*) \subseteq (-\infty,0]$.

Let $a = x + iy$ be the Cartesian decomposition of a, with $x = x^*$ and $y = y^*$. Expanding $a^*a = (x-iy)(x+iy)$ and $aa^* = (x+iy)(x-iy)$ and canceling where possible, we obtain

$$a^*a + aa^* = 2x^2 + 2y^2.$$

Hence $-(2x^2+2y^2) \geq 0$. Adding the positive element $2y^2$ we find that $-2x^2 \geq 0$, and thus $-x^2 \geq 0$. Since x^2 is a positive element, the preceding sentence implies that its spectrum is contained in $(-\infty,0] \cap [0,\infty) = \{0\}$; hence $\|x^2\| = r(x^2) = 0$, and $x = 0$ follows. Similarly, $y = 0$. □

The key result on the existence of positive elements is the following:

THEOREM 4.8.3. *In a unital C^*-algebra A, every element of the form a^*a has nonnegative spectrum.*

PROOF. Fix $a \in A$, and consider the following continuous functions $f, g : \mathbb{R} \to \mathbb{R}$:

$$f(t) = \begin{cases} \sqrt{t} & , t \geq 0, \\ 0 & , t < 0, \end{cases}$$

and

$$g(t) = \begin{cases} 0 & , t \geq 0, \\ \sqrt{-t} & , t < 0. \end{cases}$$

We have $f(t)^2 - g(t)^2 = t$ and $f(t)g(t) = 0$, $t \in \mathbb{R}$. The properties of the continuous functional calculus imply that $x = f(a^*a)$ and $y = g(a^*a)$ are self-adjoint elements of A satisfying $xy = yx = 0$ and

$$a^*a = x^2 - y^2.$$

Consider the element $ya^*ay = y(x^2 - y^2)y = -y^4$. The spectrum of ya^*ay is nonpositive, so that Lemma 4.8.2 implies that $ay = 0$. Hence $y^4 = -ya^*ay = 0$, and since y is self-adjoint, this entails $y = 0$. We conclude that $a^*a = x^2$ is the square of a self-adjoint element of A and is therefore positive. □

COROLLARY 1. Let ρ be a linear functional on a unital C^*-algebra A satisfying $\|\rho\| = \rho(\mathbf{1}) = 1$. Then ρ is a state.

PROOF. We have to show that $\rho(a^*a) \geq 0$ for every $a \in A$. By Theorem 4.8.3 it is enough to show that for every self-adjoint element $x \in A$ having nonnegative spectrum, we have $\rho(x) \geq 0$. More generally, we claim that for every *normal* element $z \in A$,

$$\rho(z) \in \overline{\operatorname{conv}}\, \sigma(z).$$

To see this, let B be the commutative C^*-subalgebra generated by z and $\mathbf{1}$. The restriction ρ_0 of ρ to B satisfies the same hypotheses $\|\rho_0\| = \rho_0(\mathbf{1}) = 1$. By Theorem 2.2.4, B is isometrically $*$-isomorphic to $C(X)$, and for $C(X)$ this is the result of Lemma 1.10.3. □

COROLLARY 2. For every element x in a unital C^*-algebra A there is a state ρ such that $\rho(x^*x) = \|x\|^2$.

PROOF. Consider the self-adjoint element $y = x^*x$, and let B be the sub C^*-algebra generated by y and the identity. Again, since $B \cong C(X)$ there is a complex homomorphism $\omega \in \operatorname{sp}(B)$ such that $\omega(y) = \|y\|$. Let ρ be any extension of ω to a linear functional on A with $\|\rho\| = \|\omega\| = 1$. We also have $\rho(\mathbf{1}) = \omega(\mathbf{1}) = 1$. Thus $\|\rho\| = \rho(\mathbf{1}) = 1$, and the preceding corollary implies that ρ is a state. □

Let us examine the implications of Corollary 2. Fixing an element $x \in A$, choose a state ρ satisfying $\rho(x^*x) = \|x\|^2$. Applying the GNS construction to ρ we obtain a Hilbert space H, a vector $\xi \in H$, and a representation $\pi \in \operatorname{rep}(A, H)$ with the property

$$\rho(a) = \langle \pi(a)\xi, \xi \rangle, \qquad a \in A.$$

Taking $a = \mathbf{1}$ we have $\|\xi\|^2 = \rho(\mathbf{1}) = 1$; hence ξ is a unit vector. Taking $a = x$ we find that $\|\pi(x)\xi\|^2 = \rho(x^*x) = \|x\|^2$; hence $\|\pi(x)\| = \|x\|$. We conclude that for every element $x \in A$ there is a representation π_x of A on some Hilbert space H_x such that $\|\pi_x(x)\| = \|x\|$. Considering the direct sum of Hilbert spaces

$$H = \oplus_{x \in A} H_x$$

and the representation $\pi \in \operatorname{rep}(A, H)$ defined by

$$\pi = \oplus_{x \in A} \pi_x,$$

we see that π is an *isometric* representation of A on H. Thus we have proved the following result:

THEOREM 4.8.4 (Gelfand–Naimark). *Every unital C^*-algebra can be represented isometrically and $*$-isomorphically as a C^*-algebra of operators on some Hilbert space.*

Of course, the Hilbert space $\oplus_{x\in A} H_x$ is never separable, and a natural question is whether A can be represented faithfully on a *separable* Hilbert space. There is no satisfactory answer in general, but for the important class of C^*-algebras that are generated by a countable set of elements the answer is yes (see Exercise (4) below).

REMARK 4.8.5. *Pure states: Irreducible representations.* Let A be a unital C^*-algebra. The set $S(A)$ of all states is a convex set in the unit ball of the dual of A, and it is closed and therefore compact in its relative weak*-topology. By the Krein–Milman theorem, $S(A)$ is the closed convex hull of its set of extreme points.

An extreme point of $S(A)$ is called a *pure* state. The result of Exercise (6) below implies that Corollary 2 can be strengthened so that $\rho(x^*x) = \|x\|^2$ is achieved with a pure state ρ. It is significant that pure states correspond to irreducible representations in the sense that a state ρ is pure if, and only if, its GNS pair (π, ξ) has the property that π is an irreducible representation. Thus one may infer that *for every element $x \in A$ there is an irreducible representation $\pi \in \operatorname{rep}(A, H)$ such that $\|\pi(x)\| = \|x\|$*. The reader is referred to [**2**] and [**10**] for more detail and further applications.

Exercises.

(1) Show that the Gelfand–Naimark theorem remains true *verbatim* for C^*-algebras without a unit.

(2) Show that in the disk algebra A, considered as a Banach $*$-algebra with involution $f^*(z) = \overline{f(\bar{z})}$, $z \in \Delta$, there are elements a for which the spectrum of a^*a is the closed unit disk.

A C^*-algebra is *separable* if it contains a countable norm-dense set.

(3) Let A be a C^*-algebra that is generated as a C^*-algebra by a finite or countable set of its elements. Show that A is a separable C^*-algebra.

(4) Show that every separable C^*-algebra can be represented (isometrically and $*$-isomorphically) on a separable Hilbert space.

(5) Let X be a compact Hausdorff space. Show that for every $p \in X$ the point evaluation $f \in C(X) \mapsto f(p)$ is a pure state of $C(X)$.

(6) Let A be a unital C^*-algebra and let x be an element of A. Show that there is a *pure* state ρ of A such that $\rho(x^*x) = \|x\|^2$. Hint: Apply Exercise (5) to the unital C^*-subalgebra $A_0 \subseteq A$ generated by x^*x, and show that a pure state of A_0 can be extended to a pure state of A.

This page intentionally left blank

Bibliography

[1] L. Ahlfors, *Complex analysis*, McGraw Hill, New York (1953).

[2] W. Arveson, *An Invitation to C^*-algebras*, Graduate texts in mathematics, vol. 39, Springer-Verlag, New York (reprinted 1998).

[3] W. Arveson, *Notes on Measure and Integration in Locally Compact Spaces*, unpublished lecture notes, available from `www.math.berkeley.edu/~arveson`.

[4] W. Arveson, *On groups of automorphisms of operator algebras*, Jour. Funct. Anal. vol. 15, no. 3 (1974), 217–243.

[5] S. Axler, *Linear Algebra Done Right*, Undergradutate texts in mathematics, Springer-Verlag, New York (1996).

[6] A. Beurling, *On two problems concerning linear transformations in Hilbert space*, Acta Math. vol. 81 (1949), 239–255.

[7] L. Coburn, *Weyl's theorem for nonnormal operators*, Mich. Math. J. vol. 13 (1966), 285–286.

[8] A. Connes, *Noncommutative Geometry*, Academic Press, San Diego (1994).

[9] K. Davidson, *C^*-algebras by Example*, Fields Institute Monographs, Amer. Math. Soc. (1996).

[10] J. Dixmier, *Les C^*-algèbres et leurs Représentations*, Gauthier-Villars, Paris (1964).

[11] R. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press, New York (1972).

[12] N. Dunford and J. Schwartz, *Linear Operators*, volume I, Interscience, New York (1958).

[13] I. Glicksberg, *The abstract F. and M. Riesz theorem*, Jour. Funct. Anal. vol. 1 (1967), 109–122.

[14] P. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, New York (1967).

[15] H. Helson, *Harmonic Analysis*, Wadsworth & Brooks/Cole, Pacific Grove, CA (1983).

[16] E. Hewitt and K. Ross, *Abstract Harmonic Analysis*, vol. 2, Springer-Verlag, Berlin (1970).

[17] L. Loomis, *An Introduction to Abstract Harmonic Analysis*, Van Nostrand, New York, (1953).

[18] P. Muhly, *Function algebras and flows I*, Acta Sci. Math. (Szeged) vol. 35 (1973), 111–121.

[19] G.K. Pedersen, *Analysis NOW*, Graduate texts in mathematics, vol. 118, Springer-Verlag, New York (1989).

[20] C. Rickart, *General Theory of Banach Algebras*, Van Nostrand, Princeton (1960).

[21] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Frederick Ungar, New York (1955).

This page intentionally left blank

Index