PART ONE

Commutative Theory

1

The Origins of Operator-Theoretic Approaches to Function Theory

In this chapter we present some of the highlights of operator theory over the last century and briefly describe their connections with function theory. It is intended to establish notation and provide orientation for the reader, before we go on to a detailed and systematic development in Chapter 2 (which can be understood independently of this scene-setting chapter). The reader should not attempt to master all the contents of this chapter in detail before progressing– some of them are substantial. Our choice, inevitably, is partial; we mention some other major contributions in historical notes at the end of the chapter, but we do not attempt a history of the subject.

1.1 Operators

We adopt the standard Bourbaki symbols $\mathbb{R}$ and $\mathbb{C}$ for the sets of real numbers and complex numbers, respectively, and also the notations

$$\mathbb{T} = \{z \in \mathbb{C}\colon\ |z| = 1\} \ \text{ and } \ \mathbb{D} = \{z \in \mathbb{C}\colon\ |z| < 1\}.$$

For any Banach space $\mathcal{X}$, we shall denote by ball $\mathcal{X}$ the closed unit ball of $\mathcal{X}$, that is,

$$\operatorname{ball}\mathcal{X} = \{u \in \mathcal{X}\colon\ \|u\| \le 1\}.$$

If $\mathcal{X}, \mathcal{Y}$ are Banach spaces, we say that a linear transformation $T\colon \mathcal{X} \to \mathcal{Y}$ is *bounded* if there exists $c \in [0,\infty)$ such that

$$\|Tu\| \le c\|u\| \quad \text{ for all } u \in \mathcal{X}. \tag{1.1}$$

For a bounded linear transformation $T\colon \mathcal{X} \to \mathcal{Y}$ we define $\|T\|$, the *norm of* T, by the formula

$$\|T\| = \sup_{u \in \operatorname{ball}\mathcal{X}} \|Tu\|, \tag{1.2}$$

or equivalently, by the formula

$$\|T\| = \inf\{c\colon\ \|Tu\| \le c\|u\| \quad \text{for all } u \in \mathcal{X}\}.$$

When $\mathcal{X}, \mathcal{Y}$ are Banach spaces, we denote by $\mathcal{B}(\mathcal{X})$ the set of bounded linear transformations from $\mathcal{X}$ to $\mathcal{X}$, and by $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ the set of bounded linear transformations from $\mathcal{X}$ to $\mathcal{Y}$. We refer to the elements of $\mathcal{B}(\mathcal{X})$ or $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ as *operators*. $\mathcal{B}(\mathcal{X})$and $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ are Banach spaces under the norm defined by equation (1.2). Furthermore, $\mathcal{B}(\mathcal{X})$ (with composition as multiplication) is a Banach algebra with identity, as it satisfies the axioms[1] $\|1\| = 1$ and

$$\|T_1 T_2\| \le \|T_1\|\ \|T_2\|$$

for all $T_1, T_2 \in \mathcal{B}(\mathcal{X})$.

Every element of $\mathcal{B}(\mathcal{X})$ has a *spectrum*. For $T \in \mathcal{B}(\mathcal{X})$ we say that T is *invertible* if there exists $S \in \mathcal{B}(\mathcal{X})$ such that $ST = TS = 1$. The *spectrum of* T, denoted by $\sigma(T)$, is defined by

$$\sigma(T) = \{z \in \mathbb{C}\colon\ z - T \text{ is not invertible}\}.$$

The set $\sigma(T)$ is a nonempty compact subset of $\mathbb{C}$ provided that $\mathcal{X} \ne \{0\}$.

1.2 Functional Calculi

Informally, the term *functional calculus* refers to a rule that enables a function f to act on an operator $T \in \mathcal{B}(\mathcal{X})$ to produce an operator $f(T) \in \mathcal{B}(\mathcal{X})$. There are numerous functional calculi that have been defined in various settings. However, in these notes we shall restrict ourselves principally to the case that the functional calculus is *holomorphic*, that is, the function f is assumed to be holomorphic on a neighborhood of $\sigma(T)$. To be an honest functional calculus, the rule should satisfy the following natural conditions.

Properties of a holomorphic functional calculus

Let $\mathcal{X}$ be a Banach space and let $T \in \mathcal{B}(\mathcal{X})$.

(1) If $f(z) = 1$ for all z in a neighborhood of $\sigma(T)$, then $f(T) = 1$, the identity operator on $\mathcal{X}$.
(2) If $f(z) = z$ for all z in a neighborhood of $\sigma(T)$, then $f(T) = T$.
(3) If f, g are holomorphic in a neighborhood of $\sigma(T)$ then

$$(f+g)(T) = f(T) + g(T) \quad \text{and} \quad (fg)(T) = f(T)g(T).$$

(4) If f is holomorphic in a neighborhood of $\sigma(T)$, $\lambda \in \mathbb{C}$ and $g = \lambda f$ then $g(T) = \lambda f(T)$.

[1] Here the first 1 denotes the identity of $\mathcal{B}(\mathcal{X})$, which is the identity operator on $\mathcal{X}$.

(5) The functional calculus has an appropriate continuity property.

For any open set U in $\mathbb{C}$, the algebra of scalar-valued holomorphic functions on U will be denoted by $\mathrm{Hol}(U)$. Observe that, for any $T \in \mathcal{B}(\mathcal{X})$, if U is a neighborhood of $\sigma(T)$ and $\tau\colon \mathrm{Hol}(U) \to \mathcal{B}(\mathcal{X})$ is defined by $\tau(f) = f(T)$, then τ is a homomorphism of algebras that satisfies $\tau(1) = 1$ and $\tau(z) = T$.

The simplest example of a functional calculus is the *polynomial calculus*. Here, for $f \in \mathbb{C}[z]$, the algebra of polynomials with complex coefficients, if

$$f(z) = \sum_{k=1}^{n} a_k z^k,$$

then $f(T)$ is defined by

$$f(T) = \sum_{k=1}^{n} a_k T^k.$$

The polynomial calculus can easily be extended to the *power series calculus*. If

$$f(z) = \sum_{k=1}^{\infty} a_k (z-a)^k$$

defines an analytic function on $\{z \in \mathbb{C}\colon |z-a| < r\}$ for some $r > 0$, then under the assumption that $\sigma(T) \subseteq \{z \in \mathbb{C}\colon |z-a| < r\}$, the operator $f(T)$ is defined by

$$f(T) = \sum_{k=1}^{\infty} a_k (T-a)^k.$$

That this series converges in the operator norm of $\mathcal{B}(\mathcal{X})$ is a simple exercise on the spectral radius formula and Cauchy's formula for the radius of convergence of a power series.

Another simple modification of the polynomial calculus is the *rational calculus*. For K, a compact subset of $\mathbb{C}$, let $\mathrm{Rat}(K)$ denote the algebra of rational functions with poles off K. If $T \in \mathcal{B}(\mathcal{X})$ and $f \in \mathrm{Rat}(\sigma(T))$, then we may write

$$f(z) = p(z)q(z)^{-1},$$

where p and q are polynomials and $q(z) \neq 0$ for all $z \in \sigma(T)$. By the spectral mapping theorem, which asserts that $q(\sigma(T)) = \sigma(q(T))$, the spectrum of $q(T)$ does not contain 0, so that $q(T)$ is invertible. Therefore, $f(T)$ can be defined by the formula

$$f(T) = p(T)q(T)^{-1}. \tag{1.3}$$

A sweet spot in the theory of functional calculi is the *Riesz–Dunford functional calculus*. Treatments of this topic, complete with proofs, are in

[60, VII.4; 140, section 17.2], and many other introductory texts on functional analysis. We shall just state the important results.

Fix an operator $T \in \mathcal{B}(\mathcal{X})$. For any open set U in $\mathbb{C}$ such that $\sigma(T) \subseteq U$, let Γ be a finite collection of closed rectifiable curves in $U \setminus \sigma(T)$ that winds once around each point in the spectrum of $\sigma(T)$ and no times around each point in the complement of U. If $f \in \mathrm{Hol}(U)$ then, by the Cauchy integral formula,

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(w)}{w-z}\, dw,$$

whenever $z \in \sigma(T)$. The Riesz–Dunford functional calculus defines $f(T)$ for $f \in \mathrm{Hol}(U)$ by the substitution $z = T$ in this formula, that is to say,

$$f(T) = \frac{1}{2\pi i} \int_\Gamma f(w)(w-T)^{-1}\, dw. \tag{1.4}$$

Note that this formula makes sense, since the assumption $\Gamma \subseteq \mathbb{C} \setminus \sigma(T)$ implies that $w - T$ is invertible for all $w \in \Gamma$. Moreover, the defining equation (1.4) depends neither on the choice of contour nor on the choice of U, since by Cauchy's theorem, for any vectors u, v in $\mathcal{X}$ and $\mathcal{X}^*$,

$$\langle f(T)u, v\rangle = \frac{1}{2\pi i} \int_\Gamma f(w)\langle (w-T)^{-1}u, v\rangle\, dw$$

is independent of Γ.

Properties of the Riesz–Dunford functional calculus

(1) If $f(z) = 1$, then $f(T) = 1$ and if $f(z) = z$, then $f(T) = T$.

(2) For any neighborhood U of $\sigma(T)$ the map

$$f \mapsto f(T)$$

is a unital algebra-homomorphism from $\mathrm{Hol}(U)$ to $\mathcal{B}(\mathcal{X})$.

(3) The calculus is consistent with the polynomial and rational calculi, that is, if U is a neighborhood of $\sigma(T)$, if r is a rational function with poles off U, and $f \in \mathrm{Hol}(U)$ is defined by $f(\lambda) = r(\lambda)$ for all $\lambda \in U$, then

$$f(T) = r(T).$$

(4) The calculus is consistent with the power series calculus, that is, if U is a disc, $\sigma(T) \subseteq U$, and $g(z)$ is the power series expansion of $f(z)$ valid in U, then $f(T) = g(T)$, where $g(T)$ is defined by the power series calculus.

(5) The calculus is continuous in the following sense. If U is a neighborhood of $\sigma(T)$, if $\{f_n\}$ is a sequence in $\mathrm{Hol}(U)$, $f \in \mathrm{Hol}(U)$, and $f_n \to f$ uniformly on compact subsets of U, then $f_n(T) \to f(T)$ in the operator norm of $\mathcal{B}(\mathcal{X})$.

Another way of thinking of the Riesz–Dunford functional calculus is by way of Runge's theorem, which implies that, for any open set U in $\mathbb{C}$, every function $f \in \text{Hol}(U)$ is the limit, uniformly on compact subsets of U, of a sequence of rational functions r_n with poles off U. As a consequence, property (5) can be used to define $f(T)$ as the norm limit of $r_n(T)$.

The holomorphic functional calculus is one of the most important tools in operator theory and is central to this book. We shall frequently want to apply a holomorphic function of d complex variables to a d-tuple of pairwise commuting operators; for this we need a generalization of the Riesz–Dunford functional calculus. Such a generalization does indeed exist; the elaboration of this theory is a heroic chapter in the history of operator theory. It is more intricate than the one-variable theory and requires a significant input from the theory of several complex variables. It was developed over several decades by many mathematicians and, while "multivariable spectral theory" remains an active research topic, the foundations of the functional calculus now appear to be in a definitive form.

For a holomorphic function f of d variables and a pairwise commuting d-tuple $T = (T^1, \dots, T^d)$ of operators on a Banach space $\mathcal{X}$, to define what we mean by $f(T)$ one strategy is again to require f to be holomorphic on the spectrum of T and to invoke an integral representation formula. To do this one must overcome some technical difficulties, the first of which is to find the appropriate notion of spectrum of T. For a start, the spectrum should contain the *joint eigenvalues* of T. We say that a non-zero vector $x \in \mathcal{X}$ is a *joint eigenvector* of T and that $\lambda = (\lambda^1, \dots, \lambda^d)$ is a corresponding joint eigenvalue of T if $T^j x = \lambda^j x$ for $j = 1, \dots, d$. If $\dim \mathcal{X} < \infty$ (and $\mathcal{X} \neq \{0\}$) then T has at least one joint eigenvalue. For let λ^1 be an eigenvalue of T^1 and let E_1 be the corresponding eigenspace $\{x \in \mathcal{X} \colon T^1 x = \lambda^1 x\}$. Since the T^j commute with T^1, the space E_1 is invariant under each T^j and the operators $T^j | E_1$ commute pairwise. Let λ^2 be an eigenvalue of $T^2 | E_1$ and let E_2 be the corresponding eigenspace $\{x \in E_1 \colon T^2 x = \lambda^2 x\}$. Continuing in this way we arrive at a point $\lambda = (\lambda^1, \dots, \lambda^d) \in \mathbb{C}^d$ and a non-zero subspace E_d of $\mathcal{X}$ such that $T^j x = \lambda^j x$ for $j = 1, \dots, d$ and all $x \in E_d$. This point $\lambda \in \mathbb{C}^d$ is a joint eigenvalue of T. However, as we already know from the case $d = 1$, when $\mathcal{X}$ is infinite-dimensional there need not be any joint eigenvalues of T.

In 1953 Georgii Evgen'evich Shilov [187] constructed a functional calculus for several elements of a commutative Banach algebra. If T is a commuting d-tuple of elements of $\mathcal{B}(\mathcal{X})$, one can choose any commutative Banach subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{X})$ containing the elements of T and then define the spectrum of T with respect to $\mathcal{A}$ by

$$\sigma_{\mathcal{A}}(T) = \mathbb{C}^d \setminus \rho_A(T), \tag{1.5}$$

where

$$\rho_{\mathcal{A}}(T) = \Big\{\lambda \in \mathbb{C}^d\colon \text{ there exist } R^1,\dots,R^d \in \mathcal{A} \text{ such that } (T^1-\lambda^1)R^1+\cdots+(T^d-\lambda^d)R^d = 1\Big\}. \tag{1.6}$$

Any joint eigenvalue of T does belong to $\sigma_{\mathcal{A}}(T)$, for any commutative subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{X})$ containing the T^j (otherwise, simply apply the defining equation (1.6) of $\rho_{\mathcal{A}}(T)$ to the corresponding joint eigenvector x to get a contradiction).

Shilov's theory yields a functional calculus for a commuting tuple T, a calculus that was further developed by numerous authors; see for example [212]. A drawback of this approach is that the spectrum of a tuple is defined relative to the chosen commutative subalgebra of $\mathcal{B}(\mathcal{X})$ and is typically different for different subalgebras. The simplest choice is to take the subalgebra to be the one generated in $\mathcal{B}(\mathcal{X})$ by the operators in question; we shall call the corresponding spectrum the *algebraic spectrum* of T and denote it by $\sigma_{\text{alg}}(T)$. However, this choice often results in an unnecessarily large spectrum.

This imperfection of the Banach algebra approach was remedied in the groundbreaking papers [198, 197] by Joseph Taylor. He introduced a more geometric notion of spectrum, now called the *Taylor spectrum* of T, denoted in this book by $\sigma(T)$.

The Taylor spectrum is a subset, sometimes proper, of $\sigma_{\mathcal{A}}(T)$ for any commutative algebra $\mathcal{A}$ containing T (indeed, Taylor showed that it is a subset of the spectrum with respect to any algebra $\mathcal{A}$ with T in its center [198, lemma 1.1]). This result is important, as a smaller spectrum gives a larger class of holomorphic functions for which $f(T)$ is defined. It is now accepted that the Taylor spectrum is the right notion for the definition of a functional calculus. We define it in Section 9.1. Let us remark that, for commuting normal tuples, there is a simpler description of the Taylor spectrum: see Section 8.3.1. Note also that $\sigma(T)$ contains all the joint eigenvalues of T [67, page 22].

A definitive discussion of the various notions of spectrum for d-tuples and development of the Taylor functional calculus is given in [89], while a more accessible account can be found in [67]. The present book, on the other hand is about the *use* of the functional calculus as a tool for function theory. What the reader will need is a knowledge of the *properties* of the functional calculus, which we shall describe as we need them.

For most of the book (all except Chapter 9) the domains that we study will be polynomially convex.[2] For such a domain U, the Taylor spectrum of T is contained in U if and only if the algebraic spectrum is. Furthermore, every holomorphic function f on U can be approximated uniformly on compact

[2] See Definition 1.43.

subsets by a sequence of polynomials p_n (this result is called the Oka-Weil theorem). The Taylor functional calculus in this case can be defined concretely by $f(T) = \lim p_n(T)$.

1.3 Operators on Hilbert Space

Operator analysis in the sense of this book concerns the special case that $\mathcal{X}$ is a complex Hilbert space $\mathcal{H}$, that is, a complete inner product space over the field of complex numbers. We assume familiarity with the basic properties of Hilbert space (see [214]). In this case, $\mathcal{B}(\mathcal{X})$ has additional structure beyond that of a Banach algebra, a structure that plays a critical role in the development of operator-analytic methods.

Definition 1.7. *For Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ and for $T \in \mathcal{B}(\mathcal{H},\mathcal{K})$, the* adjoint *of T is the operator T^* in $\mathcal{B}(\mathcal{K},\mathcal{H})$ given by*

$$\left\langle T^* v, u \right\rangle_{\mathcal{H}} = \langle v, Tu \rangle_{\mathcal{K}} \qquad \text{for all } u \in \mathcal{H},\ v \in \mathcal{K}.$$

The adjoint satisfies the *C^*-axiom*

$$\|T^*T\| = \|T\|^2. \tag{1.8}$$

In the C^*-axiom, since $T\colon \mathcal{H} \to \mathcal{K}$ and $T^*\colon \mathcal{K} \to \mathcal{H}$, the operator T^*T is well defined and maps $\mathcal{H}$ to $\mathcal{H}$. The techniques that we shall present from operator analysis rely essentially on the meaningfulness of such algebraic expressions in T and T^* (especially on the statement (1.13)). For a general Banach space, if $T\colon \mathcal{X} \to \mathcal{Y}$, then $T^*\colon \mathcal{Y}^* \to \mathcal{X}^*$, so one cannot make sense of T^*T. For this reason, in the remainder of the book, the term *operator* will always refer to a bounded linear operator acting on a complex Hilbert space or from one complex Hilbert space to another.

The techniques of operator analysis depend essentially on the notion of *positivity*. Classically, an $n \times n$ matrix $[a_{ij}]$ is said to be *positive semi-definite* if, for all $c_1, \ldots, c_n \in \mathbb{C}$,

$$\sum_{i,j=1}^{n} a_{ij} c_j \bar{c}_i \geq 0.$$

We extend this notion to operators on Hilbert space in the following way.

Definition 1.9. *An operator $T \in \mathcal{B}(\mathcal{H})$ is said to be* positive*, written $T \geq 0$, if, for all $u \in \mathcal{H}$,*

$$\langle Tu, u \rangle_{\mathcal{H}} \geq 0.$$

There are other revealing characterizations of the positivity of an operator. See [60, theorem VIII.3.6] for a proof of the following statement.

Proposition 1.10. *For $T \in \mathcal{B}(\mathcal{H})$, the following statements are equivalent.*

(i) $T \geq 0$;
(ii) $T^* = T$ *and*

$$\sigma(T) \subseteq \{x \in \mathbb{R} \colon x \geq 0\};$$

(iii) $T = X^*X$ *for some Hilbert space* $\mathcal{K}$ *and some* $X \in \mathcal{B}(\mathcal{H}, \mathcal{K})$;
(iv) $T = X^2$ *for some* $X \in \mathcal{B}(\mathcal{H})$ *such that* $X = X^*$.

There is a close connection between positivity and some other important classes of operators.

Definition 1.11. *Let* $\mathcal{H}, \mathcal{K}$ *be Hilbert spaces. An operator* $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ *is*

(i) *a* contractive operator *if* $\|T\| \leq 1$;
(ii) *an* isometric operator *if* $\|Tx\| = \|x\|$ *for every* $x \in \mathcal{H}$.

A simple but, in this book, all-pervading principle is the following.

Proposition 1.12. **(The fundamental fact of operator analysis)** *If* $\mathcal{H}, \mathcal{K}$ *are Hilbert spaces and* $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, *then*

$$T \text{ is contractive if and only if } 1 - T^*T \geq 0. \tag{1.13}$$

Proof.

$$\begin{aligned} \|T\| \leq 1 \quad &\Leftrightarrow \quad \|x\|^2 \geq \|Tx\|^2 \quad \text{for all } x \in \mathcal{H} \\ &\Leftrightarrow \quad \langle x, x \rangle \geq \langle Tx, Tx \rangle \quad \text{for all } x \in \mathcal{H} \\ &\Leftrightarrow \quad \langle x, x \rangle - \left\langle T^*Tx, x \right\rangle \geq 0 \quad \text{for all } x \in \mathcal{H} \\ &\Leftrightarrow \quad \left\langle (1 - T^*T)x, x \right\rangle \geq 0 \quad \text{for all } x \in \mathcal{H} \\ &\Leftrightarrow \quad 1 - T^*T \geq 0. \end{aligned}$$

□

The importance of the equivalence (1.13) in Proposition 1.12 cannot be overstated. It enables the expression of the analytic concept of size in terms of the algebraic concept of positivity. The norm of T is bounded by 1 if and only if $1 - T^*T$ can be represented as X^*X or X^2, as in Proposition 1.10, a condition that can often be resolved by algebra.

In Proposition 1.12 the extremal case of the inequality $1 - T^*T \geq 0$ is that it hold with equality, that is $1 - T^*T = 0$.

Proposition 1.14. *If* $\mathcal{H}, \mathcal{K}$ *are Hilbert spaces and* $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, *then*

$$T \text{ is isometric if and only if } 1 - T^*T = 0. \tag{1.15}$$