Introduction

Preliminaries in Operator Theory

In this chapter, we give a summary of the ingredients from operator theory which will be needed in discussions in subsequent chapters of this book. This is only a brief overview of the relevant definitions, basic properties, and main results. Some of them are needed only in certain specific chapters and sections, where more detailed information, when needed, will be provided in the main text or verified by the reader in the "Problems" section after each chapter. We will also fix the notations and the terminology used throughout the book. The general references are the two books [290] and [130] by Halmos and Conway, respectively.

Let $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ consist of positive integers and all integers, respectively. We use $\mathbb{R}$ and $\mathbb{C}$ to denote the fields of real and complex numbers, respectively, and $\mathbb{D}$ the open unit disc $\{z \in \mathbb{C} : |z| < 1\}$ of $\mathbb{C}$. For a nonzero z in $\mathbb{C}$, its *principal argument*, denoted by $\operatorname{Arg} z$, is the unique θ in $[0, 2\pi)$ such that $z = |z|e^{i\theta}$, and its *argument* $\arg z$ assumes the multiple values $\operatorname{Arg} z + 2n\pi$, $n = 0, \pm 1, \pm 2, \ldots$. For a nonempty subset $\triangle$ of $\mathbb{R}$, we use $\inf \triangle$ (resp., $\sup \triangle$) to denote the infimum (resp., supremum) of elements of $\triangle$. If the extremum is attained at some point of $\triangle$, then we use $\min \triangle$ (resp., $\max \triangle$) instead of $\inf \triangle$ (resp., $\sup \triangle$). For a point λ and a set $\triangle$ in the plane, we use $\operatorname{dist}(\lambda, \triangle)$ to denote the distance $\inf\{|\lambda - z| : z \in \triangle\}$ from λ to $\triangle$. For any real t, $\lfloor t \rfloor$ (resp., $\lceil t \rceil$) denotes the *floor* (resp., *ceiling*) of t, that is, the largest (resp., smallest) integer which is less (resp., greater) than or equal to t. For any set $\triangle$, let $\#\triangle$ denote its *cardinal number* or *cardinality*. The cardinalities $\#\mathbb{N} = \#\mathbb{Z}$ and $\#\mathbb{R} = \#\mathbb{C}$ are $\aleph_0$ and $\aleph_1$, respectively. If $\triangle_1$ and $\triangle_2$ are nonempty subsets of a vector space X, we use $\triangle_1 + \triangle_2$ to denote the set $\{x + y : x \in \triangle_1, y \in \triangle_2\}$. When $\triangle_1$ is a singleton $\{x\}$, $\triangle_1 + \triangle_2$ is abbreviated to $x + \triangle_2$.

I.1 Basic Properties

In the discussions throughout, we consider only Hilbert spaces H over $\mathbb{C}$ with inner product $\langle \cdot, \cdot \rangle$ and its associated norm $\| \cdot \|$. The *dimension* of H, denoted by $\dim H$, is the cardinal number of any maximal orthonormal subset of vectors in H. If $\dim H = n < \infty$, then H can be identified with $\mathbb{C}^n$ and H is separable if and only if $\dim H$

is at most countable. A net of vectors $\{x_\alpha\}$ in H is said to *converge weakly* to x if $\langle x_\alpha, y\rangle \to \langle x, y\rangle$ for any vector y in H.

A *bounded linear operator* A on H is a linear transformation on H for which the *operator norm* defined by $\|A\| = \sup\{\|Ax\|/\|x\| : x \in H, x \neq 0\}$ is finite. The collection of all operators on H is denoted by $\mathcal{B}(H)$. It is a complete metric space under the metric induced by the operator norm. Other commonly used topologies in $\mathcal{B}(H)$ are the *strong operator topology* (SOT) and the *weak operator topology* (WOT). A net of operators $\{A_\alpha\}$ converging to A under SOT (resp., WOT) means that $\|(A_\alpha - A)x\| \to 0$ (resp., $\langle (A_\alpha - A)x, y\rangle \to 0$) for any vector x (resp., vectors x and y). It is known that the WOT is properly weaker than SOT and SOT properly weaker than the norm topology.

A prominent property of WOT convergence is the following version of *Alaoglu's theorem* [130, Proposition IX.5.5].

Theorem 1.1 *Any net of operators in a bounded subset of $\mathcal{B}(H)$ has a WOT-convergent subnet.*

For $1 \le n \le \infty$, let H_j, $1 \le j \le n$, be a sequence of Hilbert spaces. Their *direct sum* $\sum_{j=1}^n \oplus H_j$ is the space $\{\sum_j \oplus x_j : x_j \in H_j \text{ for all } j, \sum_j \|x_j\|^2 < \infty\}$ equipped with the inner product $\langle \sum_j \oplus x_j, \sum_j \oplus y_j\rangle = \sum_j \langle x_j, y_j\rangle$. If A_j, $1 \le j \le n$, is an operator on H_j, then $\sum_{j=1}^n \oplus A_j$ is the operator $(\sum_j \oplus A_j)(\sum_j \oplus x_j) = \sum_j \oplus (A_j x_j)$. An operator of the form $\sum_{j=1}^n \oplus A$ on $\sum_{j=1}^n \oplus H$, $1 \le n \le \infty$, is call an *inflation* of A and will be denoted by $A^{(n)}$ on $H^{(n)}$.

Some concrete Hilbert spaces encountered in the book are $\ell^2 = \{(x_0, x_1, x_2, \ldots) : \sum_{n=0}^\infty |x_n|^2 < \infty\}$, $\ell^2(\mathbb{Z}) = \{(\ldots, x_{-1}, x_0, x_1, \ldots) : \sum_{n=-\infty}^\infty |x_n|^2 < \infty\}$, $L^2(\mu) = \{f : X \to \mathbb{C} : f \text{ measurable}, \int_X |f|^2 d\mu < \infty\}$, where μ is a positive measure on a σ-algebra of subsets of set X, $L^2[0,1]$ (resp., $L^2(\partial\mathbb{D})$) when the preceding μ is the Lebesgue measure m on $[0,1]$ (resp., $\partial\mathbb{D}$), and the *Hardy space* $H^2 = \{f : \mathbb{D} \to \mathbb{C} : f$ analytic on $\mathbb{D}$, $\sup_{0<r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta < \infty\}$ or $\{f \in L^2(\partial\mathbb{D}) : f$ has Fourier expansion $f(e^{i\theta}) = \sum_{n=0}^\infty a_n e^{in\theta}$ a.e. $[m]\}$ with the obvious inner products.

If $\dim H = n < \infty$, then an operator A on H can be identified with an n-by-n matrix $[a_{ij}]_{i,j=1}^n$. Similarly, if $\dim H$ is countably infinite, then H can be identified with ℓ^2 or $\ell^2(\mathbb{Z})$ and A on H with $[a_{ij}]_{i,j=1}^\infty$ or $[a_{ij}]_{i,j=-\infty}^\infty$.

Associated with any operator A on H is its *adjoint* A^* defined by $\langle A^*x, y\rangle = \langle x, Ay\rangle$ for all x and y in H. It is known, among other things, that $\|A^*\| = \|A\|$, $A^{**} = A$ and $(AB)^* = B^*A^*$ for operators A and B.

A subspace of a Hilbert space H is always assumed to be closed. If it is not, we refer to it as a *linear submanifold* of H. For any nonempty subset M of H, $\bigvee M$ denotes the subspace spanned by M. If K is a subspace of H, we use $K^\perp$ to denote its *orthogonal complement* $\{x \in H : \langle x, y\rangle = 0 \text{ for all } y \text{ in } K\}$ in H. For two subspaces K_1 and K_2 of H, $K_1 \ominus K_2$ denotes the *orthogonal difference* $K_1 \cap K_2^\perp$ of K_1 and K_2. The kernel and range of an operator A will be denoted by $\ker A$ and $\operatorname{ran} A$, respectively. The identity operator on H is denoted by I_H or, simply, I.

I.2 Spectral Theory

The *spectrum* $\sigma(A)$ of an operator A on H is the set $\{\lambda \in \mathbb{C} : A - \lambda I \text{ not invertible}\}$. It is a nonempty compact subset of the complex plane. In particular, A is invertible if and only if 0 is not in $\sigma(A)$. The *resolvent* of A is the analytic function $f_A(z) = (zI - A)^{-1}$ from $\mathbb{C} \setminus \sigma(A)$ to $\mathcal{B}(H)$.

The *spectral radius* $\rho(A)$ of A is $\max\{|\lambda| : \lambda \in \sigma(A)\}$. It is always less than or equal to $\|A\|$ and can be recaptured from the norms of powers of A: $\rho(A) = \lim_n \|A^n\|^{1/n} = \sup_n \|A^n\|^{1/n}$ (Gelfand's formula).

The points in $\sigma(A)$ can be classified more finely. The *point spectrum* $\sigma_p(A)$ of A consists of the eigenvalues of A, namely, the points λ for which $Ax = \lambda x$ for some nonzero vector x. If $Ax = \lambda x$ and $A^*x = \overline{\lambda}x$ for a nonzero x, then λ is a *reducing eigenvalue* of A, in which case A can be expressed as the direct sum $[\lambda] \oplus B$ for some operator B. The *approximate point spectrum* $\sigma_{ap}(A)$ of A is the set $\{\lambda \in \mathbb{C} : \|(A - \lambda I)x_n\| \to 0$ for some unit vectors x_n in $H\}$. It is equal to the *left spectrum* $\sigma_l(A)$ of A, defined by $\{\lambda \in \mathbb{C} : A - \lambda I \text{ not left invertible}\}$. The *right spectrum* $\sigma_r(A)$ of A can similarly be defined, which equals $\sigma_{ap}(A^*)^*$ of A^*, where, for any subset $\triangle$ of $\mathbb{C}$, $\triangle^*$ denotes its conjugate $\{\overline{\lambda} : \lambda \in \triangle\}$. The three spectra are related by $\partial\sigma(A) \subseteq \sigma_l(A) \cap \sigma_r(A)$.

One useful result relating the spectrum of a function of an operator $f(A)$ to that of A is the following *spectral mapping theorem*.

Theorem 2.1 *If A is an operator on H and f is a function analytic on a neighborhood of $\sigma(A)$, then $\sigma\big(f(A)\big) = f\big(\sigma(A)\big)$, where $f(A)$ is defined following the Riesz functional calculus.*

This can be found in [130, Chapter VII, Section 4].

I.3 Special Types of Operators

A *quadratic operator* A is one which satisfies $p(A) = 0$ for some quadratic polynomial p. If $p(z) = (z - a)(z - b)$, then $\sigma(A)$ consists of the two eigenvalues a and b of A. More generally, if p is allowed to be any polynomial, then A is an *algebraic operator*. Special cases of quadratic operators are *square-zero operators* ($A^2 = 0$) and *idempotent operators* ($A^2 = A$). A canonical form for quadratic operators is given in [555, Theorem 1.1].

A natural generalization of square-zero operators is the nilpotent ones. An operator A is *nilpotent* if $A^n = 0$ for some $n \geq 1$. The smallest integer n for which $A^n = 0$ is called the *nilpotency* of A. Nilpotent operators are always *quasinilpotent*, which means that its spectrum consists of 0 only.

A *Hermitian operator* A is one with $A = A^*$. In this case, its spectrum is always contained in the real line $\mathbb{R}$. Every general operator A can be written as $\operatorname{Re} A + i\operatorname{Im} A$, called the *Cartesian decomposition* of A, where $\operatorname{Re} A = (A + A^*)/2$ and

$\operatorname{Im} A = (A - A^*)/(2i)$ are the *real* and *imaginary parts* of A, respectively. Two notable subclasses of Hermitian operators are the positive semidefinite and positive definite ones. An operator A is *positive semidefinite* (resp., *positive definite*), denoted by $A \geq 0$ (resp., $A > 0$), if $\langle Ax, x\rangle \geq 0$ (resp., $\langle Ax, x\rangle > 0$) for all vectors x (resp., nonzero vectors x). Their spectra are then subsets of the positive real line $\{t \in \mathbb{R} : t \geq 0\}$. A prominent property of positive operators A is the existence of their positive square roots $A^{1/2}$. A partial order among the Hermitian operators can be defined via the positive ones, namely, if A and B are Hermitian, then $A \leq B$ means that $B - A \geq 0$. *Negative semidefinite* (resp., *negative definite*) *operators*, denoted by $A \leq 0$ (resp., $A < 0$), can be defined similarly. Every Hermitian operator A can be decomposed as $A_1 \oplus A_2$ with $A_1 \geq 0$ and $A_2 < 0$.

An idempotent Hermitian operator P on H ($P^2 = P = P^*$) is called an (*orthogonal*) *projection*, in which case its range K is closed. To emphasize this, we also denote the operator as P_K. Depending on the context, we may consider P_K as an operator on H or an operator from H to K.

If K is a subspace of H, then any operator A on H can be expressed as $\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$ on $H = K \oplus K^\perp$, where $A_1 = P_K A|K$, $A_2 = P_K A|K^\perp$, $A_3 = P_{K^\perp} A|K$, and $A_4 = P_{K^\perp} A|K^\perp$.

A subspace K of H is said to be *invariant* under A if $AK \subseteq K$. In this case, A can be expressed as $\begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix}$ on $H = K \oplus K^\perp$, where $A_1 = A|K$. If K and $K^\perp$ are both invariant under A, then we say that K is a *reducing subspace* of A or it reduces A, in which case we may express A as $A_1 \oplus A_4$ with $A_1 = A|K$ and $A_4 = A|K^\perp$.

An operator A on H is (*unitarily*) *reducible* if it has a reducing subspace other than the trivial ones $\{0\}$ and H; otherwise, it is (*unitarily*) *irreducible*. The usual way to check the irreducibility of A is to show that the only (orthogonal) projections commuting with A are 0 and I.

An operator U is *unitary* if $U^*U = UU^* = I$. Two operators A and B are *unitarily similar* (resp., *similar*) if $UA = BU$ (resp., $XA = BX$) for some unitary U (resp., invertible X). More general than the unitary operators are the isometries. An operator V is an *isometry* if $\|Vx\| = \|x\|$ for all x or $V^*V = I$. The norm of V is always equal to 1. Besides the unitaries, the *unilateral shifts* are also isometries. These are operators S_K defined by $S_K(\sum_{n=0}^\infty \oplus x_n) = 0 \oplus (\sum_{n=1}^\infty \oplus x_{n-1})$ on $\sum_{n=0}^\infty \oplus K$. If $\dim K = 1$, then S_K, abbreviated as S, is the *simple unilateral shift* and can be considered as acting on ℓ^2 by $S(x_0, x_1, x_2, \ldots) = (0, x_0, x_1, \ldots)$. More generally, if K is a separable space of dimension k, $1 \leq k \leq \infty$, then S_K is unitarily similar to $\sum_{j=1}^k \oplus S$. In contrast, the *bilateral shift* W_K (resp., *simple bilateral shift* W), which is unitary, is the operator $W_K(\sum_{n=-\infty}^\infty \oplus x_n) = \sum_{n=-\infty}^\infty \oplus x_{n-1}$ on $\sum_{n=-\infty}^\infty \oplus K$ (resp., $W(\ldots, x_{-1}, \underline{x_0}, x_1, \ldots) = (\ldots, x_{-2}, \underline{x_{-1}}, x_0, x_1, \ldots)$ on $\ell^2(\mathbb{Z})$ with the 0th component underlined). The *Wold decomposition* says that any isometry can be decomposed as $U \oplus S_K$ for some unitary U and unilateral shift S_K.

A *unilateral* (resp., *bilateral*) *weighted shift* A with bounded weights w_n, $n \geq 0$ (resp., $-\infty < n < \infty$), is the operator $A(x_0, x_1, x_2, \ldots) = (0, w_0x_0, w_1x_1, \ldots)$ on

ℓ^2 (resp., $A(\ldots, x_{-1}, \underline{x_0}, x_1, \ldots) = (\ldots, w_{-2}x_{-2}, \underline{w_{-1}x_{-1}}, w_0x_0, \ldots)$ on $\ell^2(\mathbb{Z})$). In matrix forms, these are

$$\begin{bmatrix} 0 & & & \\ w_0 & 0 & & \\ & w_1 & 0 & \\ & & \ddots & \ddots \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \ddots & & & & & \\ \ddots & 0 & & & & \\ & w_{-1} & \underline{0} & & & \\ & & w_0 & 0 & & \\ & & & w_1 & 0 & \\ & & & & \ddots & \ddots \end{bmatrix},$$

respectively.

A *partial isometry* V is one which is isometric on $(\ker V)^\perp$. Such operators play a prominent role in the *polar decomposition* of operators, which says that every operator A can be factored as $V(A^*A)^{1/2}$ for some partial isometry V. Here V is in general not unique; it is if we further require that $(\ker V)^\perp = \overline{\operatorname{ran} A^*}$ or, equivalently, $\ker V = \ker A$. Another frequently used form of polar decomposition, when $\ker A$ and $\ker A^*$ have equal dimensions, is $A = U(A^*A)^{1/2}$ with unitary U.

An operator A is *normal* if $A^*A = AA^*$. Examples of normal operators are *diagonal operators* $\operatorname{diag}(a_0, a_1, a_2, \ldots)$ on ℓ^2 with a_n's as diagonals and, more generally, *multiplication operators* M_ϕ on $L^2(\mu)$. The spectral theorem for normal operators says that every normal operator is unitarily similar to a multiplication operator (cf. [130, Theorem IX.4.6 and Proposition IX.4.7]).

Theorem 3.1 *If A is a normal operator on H, then there is a positive measure μ on a σ-algebra of subsets of a set and a function ϕ in $L^\infty(\mu)$ such that A is unitarily similar to the multiplication operator M_ϕ on $L^2(\mu)$ defined by $M_\phi f = \phi f$ for f in $L^2(\mu)$. Moreover, if H is separable, then μ is σ-finite.*

Operator A is *quasinormal* if it commutes with A^*A. A *subnormal* A on H is the restriction of a normal operator N on K to one of its invariant subspaces H. The normal N is unique up to unitary similarity, called the *minimal normal extension* of A, if we further require that N be minimal in the sense that any space in between H and K which reduces N must be equal to K. As an example, the (simple) bilateral shift is the minimal normal extension of the (simple) unilateral shift.

A *hyponormal operator* A is defined by the property $A^*A \geq AA^*$. For such an operator, we have $\|A^n\| = \|A\|^n$ for all $n \geq 1$ and therefore $\|A\| = \rho(A)$ by Gelfand's formula. It is known that hyponormal operators properly contain subnormal operators and subnormal operators properly contain normal operators.

Finally, we come to compact operators. An operator A on H is *compact* if the closure of the image of $\{x \in H : \|x\| \leq 1\}$ under A is compact in H. A is a *finite-rank operator* if the closure of $\operatorname{ran} A$ is finite dimensional. Finite-rank operators are easily seen to be compact. Moreover, any compact operator can be approximated in the operator norm by a sequence of finite-rank operators (even on a nonseparable space). Besides the finite-rank ones, other notable examples of compact operators are diagonal operators diag $(a_0, a_1, a_2, \ldots)$ on ℓ^2 with $a_n \to 0$ and the *Volterra operator* $(Vf)(x) = \int_0^x f(t)dt$ for f in $L^2[0,1]$. Their spectra are $\{a_0, a_1, a_2, \ldots, 0\}$ and $\{0\}$, respectively. According to a result of F. Riesz (cf. [130, Theorem VII.7.1]), the spectra of these three types of operators are representative of those of general compact operators on an infinite-dimensional space. For $p \geq 1$, the *Schatten p-class* $\mathcal{C}_p$ consists of compact operators A for which the eigenvalues s_n, $n \geq 1$, of $(A^*A)^{1/2}$ are such that $\sum_n s_n^p < \infty$. The *trace-class* and *Hilbert–Schmidt operators* are the ones in $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively.

I.4 Matrix Theory

The class of n-by-n complex matrices is denoted by $M_n(\mathbb{C})$ or simply by M_n. The n-by-n zero and identity matrices are 0_n and I_n, respectively. We also use $0_{m,n}$ to denote the m-by-n zero matrix. For the complex matrix $A = [a_{ij}]_{i,j=1}^n$, its *adjoint* and *transpose* are $A^* = [\overline{a}_{ji}]_{i,j=1}^n$ and $A^T = [a_{ji}]_{i,j=1}^n$, respectively. A is *Hermitian* (resp., *symmetric*) if $A = A^*$ (resp., $A = A^T$). We use $\det A$ and $\operatorname{tr} A$ for its determinant and trace. The *characteristic polynomial* p_A and *minimal polynomial* m_A of A are defined as $p_A(z) = \det(zI_n - A)$ and the monic polynomial with smallest degree for which $m_A(A) = 0_n$. It is known that m_A always divides p_A and their zeros both constitute the eigenvalues of A. The *algebraic* (resp., *geometric*) *multiplicity* of an eigenvalue λ of A is the multiplicity of λ as a zero of p_A (resp., the dimension of $\ker(\lambda I_n - A)$). The algebraic multiplicity of an eigenvalue is always bigger than or equal to its geometric multiplicity. Under similarity, every matrix can be transformed to the *Jordan form* or the *rational form*. These canonical forms are built up by taking direct sums of matrices of a special type, namely,

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix},$$

respectively. The former with $\lambda = 0$, denoted by J_n, is call a *Jordan block* and the latter *companion matrix*. On the other hand, via the unitary similarity, every matrix can be transformed to an upper-triangular one.

A matrix $A = [a_{ij}]_{i,j=1}^n$ is *nonnegative* (resp., *positive*) if $a_{ij} \geq 0$ (resp., $a_{ij} > 0$) for all i and j, which we denote by $A \succcurlyeq 0_n$ (resp., $A \succ 0_n$). A *permutation matrix* is such that each of its rows and columns has exactly one component equal to 1 and all others equal to 0. Two matrices A and B of the same size are *permutationally similar* if there is a permutation matrix P such that $P^TAP = B$. An n-by-n matrix A is (*permutationally*) *reducible* (not to be confused with (unitary) reducibility) if $n \geq 2$ and it is permutationally similar to a matrix of the form $\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$ with square matrices B and D; otherwise, A is (*permutationally*) *irreducible*. Thus a positive matrix is always irreducible. The classical *Perron–Frobenius theorem* gives rather precise information on the extremum eigenvalues of irreducible nonnegative matrices. In particular, if A is such a matrix, then $\rho(A) > 0$ is an eigenvalue of A with algebraic multiplicity equal to 1 and having an associated eigenvector with positive components (cf. [315, Theorem 8.4.4]).

I.5 C^*-Algebra Theory

A *normed space* X over $\mathbb{C}$ is a complex vector space with a *norm* $\|\cdot\|$ satisfying (a) $\|x\| = 0$ if and only if $x = 0$, (b) $\|ax\| = |a|\|x\|$, and (c) $\|x + y\| \leq \|x\| + \|y\|$ for all a in $\mathbb{C}$ and x and y in X. It is a *Banach space* if the metric induced by its norm is complete. The celebrated *Hahn–Banach theorem* is one of the fundamental results in functional analysis.

Theorem 5.1 *Let X be a normed space and Y a linear submanifold of X. If $f : Y \to \mathbb{C}$ is a bounded linear functional, then f can be extended to a bounded linear functional $F : X \to \mathbb{C}$ such that $\|F\| = \|f\|$.*

This can be found, together with many of its ramifications, in [130, Section III.6].

For a normed space X, its *dual* X^* is the Banach space consisting of all bounded linear functionals $f : X \to \mathbb{C}$ on X. In X^*, a net $\{f_\alpha\}$ is said to converge to f in the *weak-star topology* if $f_\alpha(x) \to f(x)$ for all x in X. The next theorem is the most important property of this topology: *Alaoglu's theorem*. It is in [130, Theorem V.3.1].

Theorem 5.2 *For a normed space X, the set $\{f \in X^* : \|f\| \leq 1\}$ is weak-star compact.*

A *Banach algebra* $\mathcal{A}$ over $\mathbb{C}$ is an algebra with a norm $\|\cdot\|$ which makes it a Banach space and satisfies $\|ab\| \leq \|a\|\|b\|$ for all a and b in $\mathcal{A}$. A *C^*-algebra* $\mathcal{A}$ is a Banach algebra with an involution a^* for any a in $\mathcal{A}$ satisfying $a^{**} = a$, $(\alpha a + \beta b)^* = \overline{\alpha} a^* + \overline{\beta} b^*$, $(ab)^* = b^* a^*$, and $\|a^* a\| = \|a\|^2$ for all a and b in $\mathcal{A}$ and α and β in $\mathbb{C}$. Examples of C^*-algebras are $\mathcal{B}(H)$ and the Calkin algebra $\mathcal{B}(H)/\mathcal{K}(H)$ for Hilbert spaces H (cf. Section I.6). Basic properties of C^*-algebras can be found in [130, Chapter VIII].

In Section 1.2, we would need the *Berberian representation* for operators on H. This is a useful tool in proving an asymptotic result by reducing it to its exact version.

Theorem 5.3 *For any Hilbert space H, there is another space K which contains H and a unital $*$-isomorphism α from $\mathcal{B}(H)$ to $\mathcal{B}(K)$ such that the following hold for all A in $\mathcal{B}(H)$:*

(a) $A = \alpha(A)|H$,

(b) $\|A\| = \|\alpha(A)\|$,

(c) $\alpha(A)$ *attains its norm, that is, there is a unit vector* x *in* K *such that* $\|\alpha(A)x\| = \|\alpha(A)\|$,
(d) $\sigma(A) = \sigma\big(\alpha(A)\big)$, *and*
(e) $\sigma_{ap}(A) = \sigma_{ap}\big(\alpha(A)\big) = \sigma_p\big(\alpha(A)\big)$.

This was proved in [45], where the constructions of K and α made use of the Banach limit, which we sketch as follows. An analogous construction was considered by Calkin [94] even earlier.

Let ℓ^∞ be the Banach algebra $\{(x_0, x_1, x_2, \ldots) : x_n \in \mathbb{C}$ for all n, $\sup_n |x_n| < \infty\}$ equipped with the supremum norm $\|\cdot\|_\infty$, and let c consist of $(x_0, x_1, \ldots)$ in ℓ^∞ such that $\lim_n x_n$ exists.

Theorem 5.4 *There is a linear functional* $L : \ell^\infty \to \mathbb{C}$ *such that, for* $x = (x_0, x_1, \ldots)$ *in* ℓ^∞,
(a) $\|L\| \equiv \sup\{|L(x)| : x \in \ell^\infty, \|x\|_\infty = 1\}$ *equals* 1,
(b) $L(x) = \lim_n x_n$ *if* x *is in* c,
(c) $L(x) \geq 0$ *if* $x_n \geq 0$ *for all* n, *and*
(d) $L(x) = L(x')$, *where* $x' = (x_1, x_2, \ldots)$.

This is proved in [130, Theorem III.7.1] using the Hahn–Banach theorem.

Here is a brief sketch of the proof of Theorem 5.3. Let X be the vector space consisting of sequences $\widetilde{x} = (x_0, x_1, \ldots)$ of bounded vectors x_n in H with the usual componentwise addition and scalar multiplication. Let $L : \ell^\infty \to \mathbb{C}$ be any Banach limit. For $\widetilde{x} = (x_0, x_1, \ldots)$ and $\widetilde{y} = (y_0, y_1, \ldots)$ in X, the sequence $(\langle x_0, y_0\rangle, \langle x_1, y_1\rangle, \ldots)$ is in ℓ^∞. Hence we may define $\phi : X \times X \to \mathbb{C}$ by $\phi(\widetilde{x}, \widetilde{y}) = L(\langle x_0, y_0\rangle, \langle x_1, y_1\rangle, \ldots)$. It is easily checked that $\phi(\cdot,\cdot)$ is a semi-inner product on X. Let $X_0 = \{\tilde{x} \in X : \phi(\tilde{x}, \tilde{x}) = 0\}$. Using the Cauchy–Schwarz inequality, we obtain that $\phi(\widetilde{x}, \tilde{x}) = 0$ if and only if $\phi(\widetilde{x}, \tilde{y}) = 0$ for all $\tilde{y}$ in X. Hence X_0 is a vector subspace of X. Let Y denote the quotient space X/X_0 with the inner product $\langle \widetilde{x} + X_0, \tilde{y} + X_0\rangle = \phi(\widetilde{x}, \tilde{y})$, and let K be the completion of Y with respect to the induced norm. Then K is a Hilbert space. Define $\alpha : \mathcal{B}(H) \to \mathcal{B}(K)$ by letting, for any A in $\mathcal{B}(H)$, the operator $\alpha(A)$ on K be given by $\alpha(A)((x_0, x_1, \ldots) + X_0) = (Ax_0, Ax_1, \ldots) + X_0$ for any $(x_0, x_1, \ldots)$ in X. Then α is a unital $*$-isomorphism. Let $i : H \to K$ be the linear isometry sending an element x in H to the corresponding $(x, x, \ldots) + X_0$ in K. Then we may identify H with its image $i(H)$ in K and, under this identification, we obtain (a). Parts (b) and (d) follow from the general C^*-algebra theory. To prove (c), let x_n, $n \geq 0$, be a sequence of unit vectors in H such that $\|Ax_n\| \to \|A\|$ as $n \to \infty$, and let $\widetilde{x} = (x_0, x_1, \ldots)$ in X. Then $\|\widetilde{x} + X_0\| = 1$ and $\|\alpha(A)(\widetilde{x} + X_0)\|^2 = L(\|Ax_0\|^2, \|Ax_1\|^2, \ldots) = \|A\|^2 = \|\alpha(A)\|^2$. This proves (c). For (e), $\sigma_{ap}(A) = \sigma_{ap}\big(\alpha(A)\big)$ is an easy consequence of the fact that z is not in $\sigma_{ap}(A)$ if and only if $(A - zI)^*(A - zI) \geq \varepsilon I$ for some $\varepsilon > 0$, and $\sigma_{ap}(A) \subseteq \sigma_p\big(\alpha(A)\big)$ can be proved in a similar fashion as for (c).

I.6 Fredholm Theory

For an infinite-dimensional space H, let $\mathcal{K}(H)$ be the set of all compact operators on H. It forms a self-adjoint ideal in $\mathcal{B}(H)$. Thus we can consider the quotient C^*-algebra $\mathcal{C}(H) \equiv \mathcal{B}(H)/\mathcal{K}(H)$, called the *Calkin algebra* of H. If $\pi : \mathcal{B}(H) \to \mathcal{C}(H)$ is the quotient map, then the *essential spectrum* $\sigma_e(A)$ of an operator A on H is simply the spectrum of the image $\pi(A)$ of A in the Calkin algebra. An operator A is said to be *Fredholm* if there are operators B_1 and B_2 and compact operators K_1 and K_2 such that $B_1 A = I + K_1$ and $AB_2 = I + K_2$. In terms of the Calkin algebra, this means that $\pi(A)$ is invertible in $\mathcal{C}(H)$. Thus $\sigma_e(A)$ can also be phrased as the subset $\{\lambda \in \mathbb{C} : A - \lambda I \text{ not Fredholm}\}$ of $\mathbb{C}$. The *left* (resp., *right*) *essential spectrum* $\sigma_{le}(A)$ (resp., $\sigma_{re}(A)$) of A and the *left-Fredholm* (resp., *right-Fredholm*) operators can be defined similarly. Obviously, $\sigma_e(A) \subseteq \sigma(A)$, $\sigma_{le}(A) \subseteq \sigma_l(A)$, $\sigma_{re}(A) \subseteq \sigma_r(A)$, and A is compact if and only if $\sigma_e(A) = \{0\}$. Just as the containment $\partial\sigma(A) \subseteq \sigma_l(A) \cap \sigma_r(A)$, which relates the three spectra of A, there is a corresponding essential one, namely, $\partial\sigma_e(A) \subseteq \sigma_{le}(A) \cap \sigma_{re}(A)$. In Section 1.2, we will encounter an improvement of these due to Putnam (cf. [130, Theorem XI.6.8 and Proposition XI.6.9]).

Theorem 6.1 *Let A be an operator on the infinite-dimensional H. If λ is in $\partial\sigma(A)$, then either λ is an isolated point of $\sigma(A)$ which is a reducing eigenvalue of A with finite multiplicity or λ is in $\sigma_{le}(A) \cap \sigma_{re}(A)$.*

It is known that A is left Fredholm (resp., right Fredholm) if and only if ran A is closed and $\dim \ker A < \infty$ (resp., ran A is closed and $\dim \ker A^* < \infty$). If A is either left Fredholm or right Fredholm, then its *Fredholm index*, ind A, is defined by $\dim \ker A - \dim \ker A^*$ with values in $\mathbb{Z} \cup \{\pm\infty\}$. The main result of the Fredholm index is in the following theorem [130, Theorem XI.3.7].

Theorem 6.2 *If A and B are left Fredholm* (*resp., right Fredholm*), *then so is AB and* $\text{ind}\, AB = \text{ind}\, A + \text{ind}\, B$.

Other properties of the Fredholm index can be found in [130, Section XI.3].

I.7 Compression and Dilation

If A is an operator on H and K a subspace of H, then the *compression* of A to K is the operator $B = P_K A|K$ on K, where P_K denotes the (orthogonal) projection from H onto K. In this case, we also say that A is a *dilation* of B. In matrix form, this simply means that $A = \begin{bmatrix} B & * \\ * & * \end{bmatrix}$ on $H = K \oplus K^\perp$. Compression and dilation can also be expressed in a slightly more general way by taking into account the unitary similarity, namely, A on H is a dilation of B on K (equivalently, B is a compression of A) if $B = V^*AV$ for some operator V from H to K satisfying $V^*V = I_K$.

A *Toeplitz operator* T_ϕ, where ϕ is an essentially bounded function (with respect to the Lebesgue measure) on $\partial\mathbb{D}$, is the compression of the multiplication operator M_ϕ on $L^2(\partial\mathbb{D})$ to H^2. Relative to the standard basis $\{e_n(e^{i\theta}) = e^{in\theta} : \theta \text{ real }, n \geq 0\}$ of H^2, T_ϕ can be represented as a *Toeplitz matrix*

$$\begin{bmatrix} a_0 & a_{-1} & a_{-2} & \\ a_1 & a_0 & a_{-1} & \ddots \\ a_2 & a_1 & a_0 & \ddots \\ & \ddots & \ddots & \ddots \end{bmatrix},$$

where the a_n's are the Fourier coefficients of ϕ. In case when $\phi(e^{i\theta}) = e^{i\theta}$, the corresponding Toeplitz operator is just the simple unilateral shift. Properties of such operators are in [290, Chapter 25].

An operator A is a *contraction* (resp., *strict contraction*) if $\|A\| \leq 1$ (resp., $\|A\| < 1$). For contractions on a Hilbert space, there is a rich dilation theory. It all started with the *Halmos dilation*, which says that every contraction can be dilated to a unitary operator (cf. [290, Problem 227]). The *Sz.-Nagy dilation theorem* gives a much stronger and more useful dilation result.

Theorem 7.1 *If A is a contraction on H, then there is a unitary operator U on a space K containing H such that $A^n = P_H U^n|H$ for all $n \geq 1$. Moreover, U is unique up to unitary similarity if we further require that $K = \bigvee_{n=-\infty}^{\infty} U^n H$.*

A contraction is *completely nonunitary* if it has no unitary summand. Every contraction can be decomposed as the direct sum of a unitary operator and a completely nonunitary one. If A is a completely nonunitary contraction on H and f is a function in the Hardy space H^∞ of bounded analytic functions on $\mathbb{D}$, then $f(A)$ can be defined, which makes up a functional calculus $f \mapsto f(A)$ from H^∞ to $\mathcal{B}(H)$ (cf. [539, Section III.2]). Centered around the Sz.-Nagy dilation, a whole branch of operator theory was developed by Sz.-Nagy and Foiaş in the 1960s and 1970s, which resulted in a functional model for the class of completely nonunitary contractions (cf. [539]). In Chapter 7 of this book, we will be mainly concerned with the class of S_n-matrices, which is the n-dimensional version of the compressions of the unilateral shift. An n-by-n matrix A is said to be of *class* S_n if $\|A\| \leq 1$, $\sigma(A) \subseteq \mathbb{D}$ and rank $(I_n - A^*A) = 1$, that is, A is a completely nonunitary contraction on an n-dimensional space with its defect index equal to one. Such matrices, or rather the general compressions of the shift, are the building blocks for the Jordan model of C_0 contractions from the Sz.-Nagy–Foiaş dilation theory (cf. [539, Section III.4] and [49]).