Frontiers in Mathematics

Mohammad Sal Moslehian
Hiroyuki Osaka

Advanced Techniques with Block Matrices of Operators

Birkhäuser

Birkhäuser

Frontiers in Mathematics

This series is designed to be a repository for up-to-date research results which have been prepared for a wider audience. Graduates and postgraduates as well as scientists will benefit from the latest developments at the research frontiers in mathematics and at the "frontiers" between mathematics and other fields like computer science, physics, biology, economics, finance, etc. All volumes are online available at SpringerLink.

Mohammad Sal Moslehian · Hiroyuki Osaka

Advanced Techniques with Block Matrices of Operators

Birkhäuser

Mohammad Sal Moslehian
Department of Pure Mathematics
Ferdowsi University of Mashhad
Mashhad, Iran

Hiroyuki Osaka
Department of Mathematical Sciences
Ritsumeikan University
Kyoto, Japan

ISSN 1660-8046 ISSN 1660-8054 (electronic)
Frontiers in Mathematics
ISBN 978-3-031-64545-7 ISBN 978-3-031-64546-4 (eBook)
https://doi.org/10.1007/978-3-031-64546-4

Mathematics Subject Classification: 15A60, 47B10, 47B15, 15B48, 46L05, 46L10, 46L30, 47A05, 47A08, 47A30, 47A50, 47A63, 47A64, 47C15

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to our dear math teachers, who introduced us to the mysterious world of mathematics:

Ali Abaei, Mohammad Ashraf Hesari, Mahmoud Elhami, Ebrahim Rezazadeh, and Hassan Taghizadeh (in high school)

Michi Nishizuka (in junior high school)
with respect and affection

Preface

The main goal of this modern book is to introduce several powerful techniques and fundamental ideas involving block matrices of operators (matrices) as well as matrices with entries in a C^*-algebra $\mathscr{A}$. These techniques and ideas can be used to solve problems that are difficult to address in $\mathscr{A}$ itself. In particular, 2×2 operator matrices provide important mathematical inequalities in many areas of operator theory and matrix analysis. We employ these matrices to simplify problems. For example, the classical proof of the Putnam–Fuglede theorem is based on 2×2 matrix techniques. Moreover, such methods are applied to investigate n-positive maps, completely positive maps, operator means, nonlinear positive maps, and various operator and norm inequalities. In recent decades, operator matrices in quantum information and computing theories have received significant attention.

This book is suitable as a textbook for an advanced undergraduate or graduate course or as a supplement for researchers and students in mathematics and physics who have a basic knowledge of linear algebra, functional analysis, and operator theory. The book provides detailed arguments and relevant technical material for most results. Some portions are drawn from various sources and presented in a self-contained, unified, and logically consistent manner.

By addressing existing literature, we ensure that readers can effectively understand our aim and derive essential techniques for working with block matrices in a clear, coherent, and integrated manner. This approach allows us to enrich the book with a deep contextual background and established methods.

This book is divided into five chapters.

Chapter 1 introduces the reader to basic concepts and theorems from functional analysis, operator theory, and matrix analysis. These concepts and results serve as essential tools for the subsequent chapters.

Chapter 2 is the heart of the book. It introduces the concept of block matrices of operators through the isomorphism $\mathbb{M}_n(\mathbb{B}(\mathscr{H})) \simeq \mathbb{B}(\mathscr{H}^{\oplus n})$. This chapter provides a comprehensive exposition of dilation theory, presenting numerous characterizations of the positivity of 2×2 operator matrices of the form $\begin{bmatrix} A & X \\ X^* & B \end{bmatrix}$. It also investigates the properties of 2×2 matrices with entries in a C^*-algebra. In addition, operator matrices are used to derive several inequalities related to the eigenvalues and unitarily invariant norms of matrices.

Chapter 3 is devoted to the study of operator monotone and operator convex functions, along with a thorough investigation of their foundational characteristics. In this chapter, we establish connections between operator monotone functions and operator means using the Kubo–Ando theory. We also address the verification of positive, n-positive, weakly n-positive, and completely positive maps.

Chapter 4 extends the concepts of variance and covariance beyond classical probability theory to a noncommutative framework. In this context, we provide upper bounds for unitarily invariant norms of the covariance of bounded linear operators and matrices.

Chapter 5 is concerned with the topic of nonlinear positive maps. We examine Lieb maps and their essential properties, explore the concept of 3-positivity in nonlinear maps, and investigate the continuity of 3-positive maps. Throughout this chapter, we extensively utilize block techniques to facilitate our analysis.

At the end of each chapter, readers can expect a variety of exercises and problems with references to the relevant literature. Some of these problems involve open questions, while others are challenging and provide suggestions for future research. The book also includes an extensive bibliography with about 230 references. It is worth noting that several results are due to esteemed mathematicians in the field such as Tsuyoshi Ando, Rajendra Bhatia, Jean-Christophe Bourin, Man-Duen Choi, Fuad Kittaneh, Vern I. Paulsen, and others. In addition, the book includes our favorite strategies involving block matrices. However, readers may find techniques in the literature other than those covered in this book.

The authors would like to express their gratitude to Jean-Christophe Bourin, Ali Dadkhah, Masatoshi Fujii, Fumio Hiai, Minghua Lin, Pei Yuan Wu, and Qingxiang Xu for their valuable comments and suggestions.

Mashhad, Iran Mohammad Sal Moslehian
Kyoto, Japan Hiroyuki Osaka
Spring 2024

Contents

1 Matrices and Hilbert Space Operators

We introduce basic concepts and key results from functional analysis, operator theory, and matrix analysis. These concepts and results will play a central role as indispensable tools in the upcoming chapters.

1.1 Fundamental Information

Throughout the book, all vector spaces are considered to be complex vector spaces, and all operators are assumed to be linear and bounded unless explicitly stated otherwise. A capital letter stands for an operator, a matrix, or an element of a C^*-algebra. We use the following notations:

A *normed space* is a vector space $\mathcal{X}$ equipped with a so-called *norm* $\|\cdot\| : \mathcal{X} \to [0, \infty)$ satisfying the following properties: (i) $\|x\| = 0$ if and only if $x = 0$; (ii) $\|\lambda x\| = |\lambda|\,\|x\|$; (iii) $\|x + y\| \le \|x\| + \|y\|$ for all $x, y \in \mathcal{X}$ and all scalars λ in the field $\mathbb{C}$ of complex numbers. If (i) replaced by $\|x\| = 0$ if $x = 0$, then $\|\cdot\|$ is called a *seminorm*.

If a normed space $(\mathcal{X}, \|\cdot\|)$, endowed with the metric $d(x, y) := \|x - y\|$, is a complete metric space, in the sense that every Cauchy sequence in $\mathcal{X}$ converges to some vector in $\mathcal{X}$, then it is called a *Banach space*. The *dual space* $\mathcal{X}'$ of a normed space $\mathcal{X}$ is defined as the Banach space consisting of all continuous linear functionals $f : \mathcal{X} \to \mathbb{C}$.

Every finite-dimensional normed space $\mathcal{X}$ is a Banach space and any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on it are equivalent. This means that there exist positive numbers α and β such that $\alpha\|x\|_2 \le \|x\|_1 \le \beta\|x\|_2$ for all $x \in \mathcal{X}$. The space $\mathbb{M}_n$ of $n \times n$ matrices with complex entries together with any one of the norms $\|[a_{ij}]\|_\sigma = \sum_{i,j=1}^n |a_{ij}|$ and $\|[a_{ij}]\|_{\max} = \max_{1\le i,j\le n} |a_{ij}|$ is a Banach space. Hilbert spaces are significant examples of Banach spaces as defined below.

M. S. Moslehian and H. Osaka, *Advanced Techniques with Block Matrices of Operators*, Frontiers in Mathematics, https://doi.org/10.1007/978-3-031-64546-4_1

Suppose that $\mathscr{H}$ is a vector space. A function $\langle \cdot, \cdot \rangle : \mathscr{H} \times \mathscr{H} \to \mathbb{C}$ is called a *semi-inner product* if

(i) $\langle x + \lambda y, z \rangle = \langle x, z \rangle + \lambda \langle y, z \rangle$,
(ii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$,
(iii) $\langle x, x \rangle \geq 0$

for all $x, y, z \in \mathscr{H}$ and all scalars $\lambda \in \mathbb{C}$. Then, the pair $(\mathscr{H}, \langle \cdot, \cdot \rangle)$ is called a *semi-inner product space* and $\|x\| = \sqrt{< x, x >}$ gives a seminorm. In such a space, it holds that

$$\langle x, y \rangle = \frac{1}{4} \sum_{k=0}^{3} i^k \|x + i^k y\|^2 \qquad (x, y \in \mathscr{H}) \tag{1.1.1}$$

which is called the *polarization identity*.

If $x = 0$ whenever $\langle x, x \rangle = 0$, then $\mathscr{H}$ is called an *inner product space* and $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ can be shown to be a norm on $\mathscr{H}$ by using the following *Cauchy–Schwarz inequality*:

Theorem 1.1.1 (Cauchy–Schwarz inequality) *In a semi-inner product space $\mathscr{H}$, it holds that*

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \langle y, y \rangle^{1/2} \qquad (x, y \in \mathscr{H}).$$

If the norm $\| \cdot \|$ is complete, then $\mathscr{H}$ is called a *Hilbert space*. It is known that every Hilbert space is of the form $L_2(\Omega, \mu)$, which is the *space of square integrable functions* (by identifying functions that are equal almost everywhere) in a measure space (Ω, μ). In particular, the space ℓ_2 consists of square summable sequences (x_n) equipped with the inner product

$$\langle (x_n), (y_n) \rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}.$$

Also, the closed subspace $\mathbb{C}^n$ of ℓ^2 is known as *n-dimensional Euclidean space*.

A vector $x \in \mathscr{H}$ is called *orthogonal* to $y \in \mathscr{H}$ if $\langle x, y \rangle = 0$. The set $\mathcal{M}^{\perp}$ consists of all $x \in \mathscr{H}$ that are orthogonal to every element in a subset $\mathcal{M}$ of $\mathscr{H}$, and it is called the *orthogonal complement* of $\mathcal{M}$. Two vectors are called *orthonormal* if they are unit vectors (that is, of norm 1) and orthogonal to each other. An orthonormal basis for $\mathscr{H}$ is an orthonormal set $(e_j)_{j \in J}$ of vectors such that

$$x = \sum_j \langle x, e_j \rangle e_j$$

for every $x \in \mathscr{H}$. Also,

$$\|x\|^2 = \sum_j |\langle x, e_j \rangle|^2$$

is called the *Parseval identity*. The Gram–Schmidt process is a method used to create orthogonal vectors from a finite number of linearly independent vectors in a Hilbert space. Employing Zorn's lemma along with the Gram–Schmidt process, we can show that every Hilbert space admits an orthonormal basis. The standard basis for ℓ_2 is composed of the vectors $e_n = (\delta_{1n}, \delta_{2n}, \delta_{3n}, \ldots)$ for $n = 1, 2, \ldots$, where δ represents the Kronecker delta defined as

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

A linear map $A : \mathscr{H} \to \mathscr{H}$ is called a *bounded operator* if

$$\sup\{\|Ax\| : x \in \mathscr{H}, \|x\| = 1\} < \infty.$$

In this case, its *operator norm* is defined by

$$\|A\| = \sup\{\|Ax\| : x \in \mathscr{H}, \|x\| = 1\}.$$

The space $\mathbb{B}(\mathscr{H})$ of all bounded linear operators on a complex Hilbert space $\mathscr{H}$ equipped with the operator norm $\|\cdot\|$ is a Banach space. The identity operator in $\mathbb{B}(\mathscr{H})$ is denoted by I. We can turn the space $\mathbb{B}(\mathscr{H})$ into an algebra by defining the multiplication of $A, B \in \mathbb{B}(\mathscr{H})$ as the composition $AB := A \circ B$.

On $\mathbb{B}(\mathscr{H})$, we can define the *weak operator topology* by the convergence of nets as $A_\alpha \xrightarrow{\text{WOT}} A$ (or $\text{w}-\lim A_\alpha = A$) whenever $\langle A_\alpha x, y\rangle \to \langle Ax, y\rangle$ for all $x, y \in \mathscr{H}$. We can also define the *strong operator topology* by $A_\alpha \xrightarrow{\text{SOT}} A$ (or $\text{s}-\lim A_\alpha = A$) whenever $A_\alpha x \to Ax$ for all $x \in \mathscr{H}$.

We denote the range, kernel, and rank of an operator A by $\operatorname{ran} A$, $\ker A$, and $\operatorname{rank}(A) = \dim(\operatorname{ran} A)$. The closure of a set $\mathscr{D}$ in the norm topology is denoted by $\overline{\mathscr{D}}$. The restriction of an operator A to a set $\mathcal{D}$ is presented as $A|_{\mathcal{D}}$. A linear map between Hilbert spaces is called *bounded below* if there is a positive number c such that $\|Ax\| \geq c\|x\|$ for all $x \in \mathscr{H}$. In general, a bounded below linear map may not be bounded. For example, the map $A : \ell_2 \to \ell_2$ defined by $A(x_1, x_2, x_3, \ldots) = (x_1, 2x_2, 3x_3, \ldots)$ is bounded below since

$$\|A(x_1, x_2, x_3, \ldots)\|_2^2 = \sum_{n=1}^{\infty} n^2|x_n|^2 \geq \sum_{n=1}^{\infty} |x_n|^2 = \|(x_1, x_2, x_3, \ldots)\|.$$

However, it is not bounded since $\|A(e_n)\| = n$.

If $\dim \mathscr{H} = n$, we can identify $\mathbb{B}(\mathscr{H})$ with the matrix algebra $\mathbb{M}_n$. We denote the identity of $\mathbb{M}_n$ as I_n, or simply I if there is no ambiguity. Therefore, a matrix can be viewed as an operator. The standard *system of matrix units* for $\mathbb{M}_n$ is the family $(E_{ij})_{1\leq i,j\leq n}$. The (i, j)-entry of E_{ij} is one, while all other entries are zero. In fact, we can express $E_{ij} = e_i^* e_j$, where e_i is the ith vector in the standard basis of $\mathbb{C}^n$. A matrix obtained by deleting certain rows or columns from a given matrix is called a *submatrix* of the original matrix. A *principal* *submatrix* is a specific type of submatrix where the remaining row indices are identical to the remaining column indices.

Let us recall the tensor product of matrices $A = [a_{ij}] \in \mathbb{M}_m$ and $B = [b_{pq}] \in \mathbb{M}_n$. Let $(e_1, e_2, \ldots, e_m)$ and $(f_1, f_2, \ldots, f_n)$ be the standard orthonormal bases for $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively. The *tensor product of matrices* A and B is the matrix $A \otimes B$ represented relative to the basis elements $e_i \otimes f_p$ of the tensor product $\mathbb{C}^m \otimes \mathbb{C}^n$ via

$$\langle (A \otimes B)(e_j \otimes f_q), e_i \otimes f_p \rangle = \langle Ae_j \otimes Bf_q, e_i \otimes f_p \rangle = \langle Ae_j, e_i \rangle \langle Bf_q, f_p \rangle = a_{ij} b_{pq}.$$

It is easy to see that the tensor product of positive semidefinite matrices is positive semidefinite.

The following result is commonly used. To prove that two operators $A, B \in \mathbb{B}(\mathscr{H})$ are equal it is sufficient to use it and show that $\langle Ax, x \rangle = \langle Bx, x \rangle$ for all $x \in \mathscr{H}$.

Proposition 1.1.2 *An operator $A \in \mathbb{B}(\mathscr{H})$ is the zero operator* 0 *if and only if $\langle Ax, x \rangle = 0$ for all $x \in \mathscr{H}$.*

Proof Suppose that $\langle Ax, x \rangle = 0$ for all $x \in \mathscr{H}$. It follows from (1.1.1) that

$$\langle Ax, y \rangle = \frac{1}{4} \sum_{k=0}^{3} i^k \langle A(x + i^k y), (x + i^k y) \rangle$$

that $\langle Ax, y \rangle = 0$ for all $x, y \in \mathscr{H}$. Set $y = Ax$ to get $\|Ax\|^2 = \langle Ax, Ax \rangle = 0$. Hence, $A = 0$. The reverse assertion is clear. □

The following theorem characterizes the dual of a Hilbert space.

Theorem 1.1.3 (Riesz representation theorem) *Let $\mathscr{H}$ be a Hilbert space and $f : \mathscr{H} \to \mathbb{C}$ be a bounded linear functional. Then there exists a unique element $z \in \mathscr{H}$ such that $\|z\| = \|f\|$ and $f(x) = \langle x, z \rangle$ for all $x \in \mathscr{H}$.*

Proof If $f = 0$, then taking $z = 0$ is sufficient. Let's assume that $f \neq 0$. Thus, $\ker f$ is a proper closed subspace. This implies that $\mathscr{H} = \mathbb{C}z' \oplus \ker f$ for some $z' \in \mathscr{H}$ with $z' \perp \ker f$. We can assume that $f(z') = 1$. Then, $x - f(x)z' \in \ker f$ and

$$\langle x, z' \rangle = \langle x - f(x)z', z' \rangle + \langle f(x)z', z' \rangle = f(x)\|z'\|^2.$$

By choosing $z = \frac{z'}{\|z'\|^2}$, we obtain $f(x) = \langle x, z \rangle$. If there exist two vectors z_1 and z_2 such that $f(x) = \langle x, z_1 \rangle = \langle x, z_2 \rangle$, then $\langle x, z_1 - z_2 \rangle = 0$ for all $x \in \mathscr{H}$. Setting $x = z_1 - z_2$, we deduce that $\langle z_1 - z_2, z_1 - z_2 \rangle = 0$, which ensures that $z_1 = z_2$. This confirms the uniqueness assertion. □

For each $A \in \mathbb{B}(\mathscr{H})$ and $y \in \mathscr{H}$, we can use the Riesz representation Theorem 1.1.3 to the bounded linear functional $x \mapsto \langle Ax, y\rangle$ to find a unique vector denoted by A^*y satisfying $\|A^*y\| \le \|A\|\|y\|$ and $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for all $x \in \mathscr{H}$. Therefore, we get a bounded linear operator $y \mapsto A^*y$ on $\mathscr{H}$. Thus, the so-called *adjoint operation* $A \mapsto A^*$ can be defined. The algebra $\mathbb{B}(\mathscr{H})$ endowed with the adjoint operation can be considered as a $*$-*algebra*, which means an algebra endowed with an involution $*$ that satisfies the following properties:

- $(A + \lambda B)^* = A^* + \overline{\lambda} B^*$;
- $(AB)^* = B^*A^*$;
- $(A^*)^* = A$.

For matrices, if $A = [a_{ij}]$, then $A^* = [\overline{a_{ji}}]$. The matrix $A^t = [a_{ji}]$ is said to be the *transpose* of A.

A C^*-*algebra* is a complex $*$-algebra $\mathscr{A}$, which is at the same time a Banach space and satisfies the submultiplicative property $\|AB\| \le \|A\|\,\|B\|$ and the C^*-condition

$$\|A^*A\| = \|A\|^2 \tag{1.1.2}$$

for all $A, B \in \mathscr{A}$. The identity element of a C^*-algebra $\mathscr{A}$ is denoted by I. It is easy to verify that $\mathbb{B}(\mathscr{H})$ is a C^*-algebra. In addition, if Ω is any compact Hausdorff space, then the algebra $C(\Omega)$ of all continuous complex-valued functions on Ω equipped with the sup-norm $\|f\|_\infty = \sup_{x\in\Omega} |f(x)|$ and the $*$-operation $f \mapsto \bar{f}$ is a C^*-algebra. Here, $\bar{f}(x) := \overline{f(x)}$ for all $x \in \Omega$.

The C^*-subalgebra generated by a subset of $\mathscr{A}$ is defined as the intersection of all C^*-subalgebras of $\mathscr{A}$ containing the subset. Especially, the C^*-algebra generated by a self-adjoint element $A \in \mathscr{A}$ and the identity operator I, denoted by $C^*(A, I)$, is the closure of all polynomials in A, and so it is commutative. We use $\mathscr{A}$ and $\mathscr{B}$ to denote C^*-algebras and Φ and Ψ for arbitrary (linear or nonlinear) maps between C^*-algebras.

By a $*$-*homomorphism* $\pi : \mathscr{A} \to \mathscr{B}$, we mean a linear map that satisfies the multiplicative property $\pi(AB) = \pi(A)\pi(B)$ and the $*$-preserving property $\pi(A^*) = \pi(A)^*$. A $*$-homomorphism $\pi : \mathscr{A} \to \mathbb{B}(\mathscr{H})$ is called a *representation*. An *irreducible representation* is a representation $\pi : \mathscr{A} \to \mathbb{B}(\mathscr{H})$, where the algebra $\pi(\mathscr{A})$ has no invariant closed subspace other than 0 and itself. Recall that a closed subspace $\mathscr{K}$ of $\mathscr{H}$ is called an *invariant subspace* for a subset $\mathcal{S}$ of $\mathbb{B}(\mathscr{H})$ if for each operator $A \in \mathcal{S}$ we have $A(\mathscr{K}) \subseteq \mathscr{K}$.

Every C^*-algebra can be regarded as a norm-closed $*$-subalgebra of $\mathbb{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$. This representation is known as the *Gelfand–Naimark–Segal representation*; refer to [176, Theorem 3.4.1].

A one-to-one surjective $*$-homomorphism is called a $*$-isomorphism. If a C^*-algebra $\mathscr{A}$ is $*$-isometric to a C^*-algebra $\mathscr{B}$, then we write $\mathscr{A} \simeq \mathscr{B}$. It is well-known [176, Theorem 3.1.5] that every $*$-isomorphism $\pi : \mathscr{A} \to \mathscr{B}$ between C^*-algebras is isometric, that is, $\|\pi(A)\| = \|A\|$ for all $A \in \mathscr{A}$.

The Gelfand–Naimark–Segal representation enables us to naturally define various types of operators such as self-adjoint operators, normal operators, contractions, isometries, coisometries partial isometries, unitaries, idempotents, and projections for elements of a C^*-algebra. An operator A in a C^*-algebra $\mathscr{A}$ is said to be

- *self-adjoint* (*Hermitian matrix*) if $A^* = A$;
- *normal* if $A^*A = AA^*$;
- *contraction* if $\|A\| \leq 1$;
- *isometry* if $A^*A = I$;
- *coisometry* if A^* is an isometry;
- *partial isometry* if $AA^*A = A$;
- *unitary* if $A^*A = AA^* = I$;
- *idempotent* if $A^2 = A$;
- *nilpotent* if $A^n = 0$ for some n in the set of natural numbers $\mathbb{N}$;
- *(orthogonal) projection* if it is a self-adjoint idempotent.

For a closed subspace $\mathscr{M}$, it holds that $\mathscr{H} = \mathscr{M} \oplus \mathscr{M}^\perp$, which means that every element $x \in \mathscr{H}$ can be uniquely expressed as $x = y + z$ where $y \in \mathscr{M}$ and $z \in \mathscr{M}^\perp$. The map $x \mapsto y$ gives a projection, denoted by $P_{\mathscr{M}}$, on $\mathscr{H}$ whose range is $\mathscr{M}$. Conversely, for every projection $P \in \mathbb{B}(\mathscr{H})$, the range of P is a closed subspace of $\mathscr{H}$. In fact, there exists a one-to-one correspondence between the set of all closed subspaces of $\mathscr{H}$ and the set of projections in $\mathbb{B}(\mathscr{H})$. Given two projections P and Q, the projection onto $\operatorname{ran}(P) \cap \operatorname{ran}(Q)$ is denoted as $P \wedge Q$.

If A is a partial isometry, then A^*A (AA^*, respectively) is a projection and its range is called the *initial space* (*final space*, respectively) of A.

The Cartesian decomposition of an element $A \in \mathscr{A}$ is $A = B + iC$, where $B = \operatorname{Re}(A) := (A + A^*)/2$ and $C = \operatorname{Im}(A) := (A - A^*)/(2i)$ are self-adjoint and called the *real part* and *imaginary part* of A. The real linear space of all self-adjoint elements of $\mathscr{A}$ is denoted by $\mathscr{A}_{\text{sa}}(\mathscr{H})$. In particular, the set of self-adjoint operators on a Hilbert space $\mathscr{H}$ is denoted by $\mathbb{B}_{\text{sa}}(\mathscr{H})$.

The spectrum of an operator $A \in \mathbb{B}(\mathscr{H})$ refers to the set $\operatorname{sp}(A)$ of all complex numbers λ such that $A - \lambda I$ is not invertible in $\mathbb{B}(\mathscr{H})$. It is a compact and nonempty set. When A is a matrix in $\mathbb{M}_n$, the set of its eigenvalues is exactly $\operatorname{sp}(A)$. The *spectral radius* of an operator A is defined as

$$r(A) = \max\{|\lambda| : \lambda \in \operatorname{sp}(A)\}.$$

The *Gelfand–Beurling formula* states that

$$r(A) = \inf_n \|A^n\|^{1/n} = \lim_n \|A^n\|^{1/n}.$$

In particular, $r(A) \leq \|A\|$. For $A, B \in \mathbb{B}(\mathscr{H})$,

$$\operatorname{sp}(AB) \cup \{0\} = \operatorname{sp}(BA) \cup \{0\}, \tag{1.1.3}$$

and so $r(AB) = r(BA)$. The equality (1.1.3) follows from the relation

$$\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}.$$

Example 1.1.4 Suppose that $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then

(i) $A^2 = 0$, so $r(A) = \inf_n \|A^n\|^{\frac{1}{n}} = 0$, but $\|A\| = 1$. Hence, equality may not hold in $r(A) \le \|A\|$ ([134, Theorem 3.2.3]).
(ii) $r(AB) > r(A)r(B)$ and $r(A+B) > r(A) + r(B)$. Thus, the spectral radius is neither submultiplicative nor subadditive in general; see [134, Proposition 3.2.10].

Example 1.1.5 Let $A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Then $\mathrm{sp}(A+B)$ is not a subset of $\mathrm{sp}(A) + \mathrm{sp}(B)$. It is notable that if A and B commute, then $\mathrm{sp}(A+B) \subseteq \mathrm{sp}(A) + \mathrm{sp}(B)$ [134, Proposition 3.2.10].

Let $A \in \mathbb{B}(\mathscr{H})$. The set $W(A) = \{\langle Ax, x\rangle : \|x\| = 1\}$ is said to be the *numerical range* of A and $w(A) = \sup\{|\lambda| : \lambda \in W(A)\}$ is called the *numerical radius* of A.

It is known that $W(A)$ is convex, according to the Toeplitz-Hausdorff theorem. It is invariant under unitary similarity, which means that $W(A) = W(U^*AU)$ for all A and all unitaries U. In addition, thus the convex hull of $\mathrm{sp}(A)$ is contained in the closure of $W(A)$. However, the numerical range does not have similarity invariance. To illustrate this, consider the matrix

$$A_\lambda = \begin{bmatrix} 0 & \lambda \\ 0 & 0 \end{bmatrix}.$$

Then, $W(A_\lambda) = \lambda W(A_1)$, which is the closed disc centered at the origin with radius $|\lambda|$. Thus, $W(A_\lambda)$'s are distinct sets. However, for $\lambda \neq 0$, the matrix A_λ is similar to A_1. This can be seen by considering $B_\lambda = \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}$, which satisfies $A_\lambda = B_\lambda A_1 B_\lambda^{-1}$; see [207]. Interestingly, $w(\cdot)$ is a norm on $\mathbb{B}(\mathscr{H})$ that is equivalent to the operator norm. More precisely,

$$\frac{1}{2}\|A\| \le w(A) \le \|A\| \tag{1.1.4}$$

holds for each $A \in \mathbb{B}(\mathscr{H})$.

Theorem 1.1.6 *If $A \in \mathbb{B}_{sa}(\mathscr{H})$, then*

(i) $r(A) = \|A\|$;

(ii) sp(A) *is a subset of the set* $\mathbb{R}$ *of real numbers;*
(iii) *at least one of* $\|A\|$ *or* $-\|A\|$ *is in* $\operatorname{sp}(A)$ *and* $\operatorname{sp}(A) \subseteq [-\|A\|, \|A\|]$*;*
(iv)

$$\|A\| = w(A).$$

Proof (i) $\|A^{2^n}\| = \|A\|^{2^n}$ for all n, so

$$r(A) = \lim_n \|A^n\|^{1/n} = \lim_n \|A^{2^n}\|^{1/2^n} = \lim_n \|A\| = \|A\|\,.$$

(ii) Let $\lambda = \alpha + i\beta \in \operatorname{sp}(A)$, where α and β are real numbers. Assume that $\beta \neq 0$. For each n, put $A_n = A - (\alpha - in\beta)I$. Then $i(n+1)\beta \in \operatorname{sp}(A_n)$ and

$$\begin{aligned}(n+1)^2\beta^2 &= \ |i(n+1)\beta|^2 \leq r(A_n)^2 \leq \|A_n\|^2 = \|A_n^* A_n\| \\ &= \ \|(A - (\alpha + in\beta)I)(A - (\alpha - in\beta)I)\| = \|(A - \alpha I)^2 + n^2\beta^2 I\| \\ &\leq \ \|A - \alpha I\|^2 + n^2\beta^2\,,\end{aligned}$$

whence $n \leq \frac{1}{2}\left(\|A - \alpha I\|^2/\beta^2 - 1\right)$ for all n, which is impossible. Hence, $\lambda = \alpha \in \mathbb{R}$.
(iii) It follows from (i) and compactness of $\operatorname{sp}(A)$ that $\|A\| \in \{|\lambda| : \lambda \in \operatorname{sp}(A) \subseteq \mathbb{R}\}$. Thus, $\|A\|$ or $-\|A\|$ is in $\operatorname{sp}(A)$. In addition, $\operatorname{sp}(A) \subseteq [-\|A\|, \|A\|]$, since $r(A) \leq \|A\|$.
(iv) Since $r(A) \leq w(A) \leq \|A\|$, it follows from (i) that

$$\|A\| = w(A).$$ □

There is a valuable formula in which the numerical radius of an operator is expressed in terms of the norm of specific operators as follows.

Corollary 1.1.7 ([233, p. 85]) *Let* $A \in \mathbb{B}(\mathscr{H})$. *Then*

$$w(A) = \sup_{\theta \in \mathbb{R}} \|\operatorname{Re}(e^{i\theta}A)\| = \frac{1}{2} \sup_{\theta \in \mathbb{R}} \|A + e^{-2i\theta}A^*\|. \tag{1.1.5}$$

Proof Since $\operatorname{Re}(e^{i\theta}A)$ is self-adjoint, we have

$$\begin{aligned}\sup_{\theta \in \mathbb{R}} \|\operatorname{Re}(e^{i\theta}A)\| &= \sup_{\theta \in \mathbb{R}} w\left(\operatorname{Re}(e^{i\theta}A)\right) = \sup_{\theta \in \mathbb{R}} \sup_{\|x\|=1} |\langle \operatorname{Re}(e^{i\theta}A)x, x\rangle| \\ &= \sup_{\|x\|=1} \sup_{\theta \in \mathbb{R}} |\operatorname{Re}\left(e^{i\theta}\langle Ax, x\rangle\right)| = \sup_{\|x\|=1} |\langle Ax, x\rangle| = w(A).\end{aligned}$$

The last equality is concluded from the definition of the real part of an operator. □

An element $A \in \mathscr{A}$ is called *positive* if it is self-adjoint and its spectrum is contained in the interval $[0, \infty)$. In this case, we write $A \geq 0$. We say A is strictly positive (positive definite in the setting of matrices) and write $A > 0$ if it is positive and invertible. For self-adjoint elements (Hermitian matrices, respectively) A and B, we say $B \geq A$ ($B > A$, respectively) if $B - A \geq 0$ ($B - A > 0$, respectively). This order is known as the *Löwner order*. The *Schur product theorem* states that the *Hadamard product* or *Schur product* $A \circ B = [a_{ij}b_{ij}]$ of two positive semidefinite matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ is also positive semidefinite; see [116, Theorem 2.18].

The (continuous) functional calculus for a self-adjoint operator provides a powerful tool for establishing connections between continuous functions and bounded linear operators. Let us consider an operator $A \in \mathbb{B}_{sa}(\mathscr{H})$. If we take a function f in the space $C(\mathrm{sp}(A))$ of all continuous functions defined on the spectrum of A, we can find a unique operator $f(A)$ in $\mathbb{B}(\mathscr{H})$. This operator possesses a special property: if we have a sequence of polynomials (p_n) such that $\lim_{n\to\infty} p_n = f$ in sup-norm on $C(\mathrm{sp}(A))$, then it follows that $f(A) = \lim_{n\to\infty} p_n(A)$ in operator norm on $\mathbb{B}(\mathscr{H})$.

This function-to-operator map, denoted as $f \mapsto f(A)$, represents a unique isometric $*$-isomorphism from the space of continuous functions defined on the spectrum of A to the C^*-subalgebra of $\mathbb{B}(\mathscr{H})$ generated by A and the identity operator I. Notably, this map assigns the function $f(t) = t$ to the operator A and the constant function $f(t) = 1$ to the identity operator I. Consequently, if we have two functions f and g in $C(\mathrm{sp}(A))$ such that $f(t) \leq g(t)$ for all $t \in \mathrm{sp}(A)$, then it follows that $f(A) \leq g(A)$. In addition, $f(A)$ is self-adjoint if and only if $f(t) = \overline{f(t)}$ for all $t \in \mathrm{sp}(A)$ or, equivalently, the range of f is a subset of $\mathbb{R}$.

The *spectral theorem* further asserts that $\mathrm{sp}(f(A)) = f(\mathrm{sp}(A))$ for any $A \in \mathscr{A}$.

Since A is self-adjoint, $\mathrm{sp}(A) \subseteq \mathbb{R}$. By the compactness of the spectrum, there are real numbers m, M such that $\mathrm{sp}(A) \subseteq [m, M]$. By functional calculus, this is equivalent to $mI \leq A \leq MI$ or simply $m \leq A \leq M$.

Let $A \in \mathbb{B}_{sa}(\mathscr{H})$. Consider the continuous functions $f_+(t) = \max\{t, 0\}$ and $f_-(t) = \max\{-t, 0\}$. These functions satisfy $f = f_+ - f_-$, $f_+f_- = 0$, $|f_\pm| \leq |f|$. Using the functional calculus for A, we get two positive operators $A_+, A_- \in \mathbb{B}(\mathscr{H})$ such that $A = A_+ - A_-$, $A_+A_- = 0 = A_-A_+$, $\|A_\pm\| \leq \|A\|$. The decomposition $A = A_+ - A_-$ is called the *Jordan decomposition*, and A_+ and A_- are called the *positive part* and the *negative part* of A, respectively.

Let us consider the case where $A \in \mathbb{B}(\mathscr{H})$ is a normal operator and $\mathscr{B}$ is the C^*-algebra generated by A and I. Then, there exists a $*$-isometrically isomorphism between $\mathscr{B}$ and $C(\mathrm{sp}(A))$, which maps A to the inclusion map of $\mathrm{sp}(A)$ in $\mathbb{C}$. Consequently, if $\mathrm{sp}(A) \subseteq \mathbb{R}$, then $z = \bar{z}$ on $\mathrm{sp}(A)$, implying that A is self-adjoint.

We can use the aforementioned fact to show that if P_1 and P_2 are projections such that P_1P_2 is normal, then P_1 and P_2 commute. To see this, note that $\mathrm{sp}(P_1P_2) \cup \{0\} = \mathrm{sp}(P_1P_2P_2) \cup \{0\} = \mathrm{sp}(P_2P_1P_2) \cup \{0\} \subseteq \mathbb{R}$ since $P_2P_1P_2$ is self-adjoint. Hence, P_1P_2 is self-adjoint, and thus $P_1P_2 = (P_1P_2)^* = P_2P_1$.

A *von Neumann algebra* $\mathscr{A}$ acting on a Hilbert space $\mathscr{H}$ is a $*$-subalgebra of the algebra $\mathbb{B}(\mathscr{H})$ such that $\mathscr{A} = (\mathscr{A}^c)^c$. Here, the *commutant of a set* $\mathcal{D} \subseteq \mathbb{B}(\mathscr{H})$ is defined by

$\mathcal{D}^c = \{Y \in \mathbb{B}(\mathscr{H}) : \; XY = YX \;\text{ for all }\; X \in \mathcal{D}\}$. Equivalently, a von Neumann algebra is a C^*-algebra $\mathscr{A}$ that is the dual of a Banach space $\mathscr{A}_*$. The latter space is indeed the space of all normal linear functionals on $\mathscr{A}$. A bounded linear functional is called a *normal linear functional* when for a bounded increasing net of self-adjoint operators (A_α), we have $f(\sup_\alpha A_\alpha) = \sup_\alpha f(A_\alpha)$. Here, $\sup_\alpha A_\alpha$ means the least upper bound of the self-adjoint operators A_α's with respect to the Löwner order.

Projections P and Q are called *Murray–von Neumann equivalent*, denoted as $P \sim Q$, if there is a partial isometry U such that $U^*U = P$ and $UU^* = Q$. A von Neumann algebra is said to be *properly infinite* if there exist projections P_1 and P_2 such that $P_1 \sim I$, $P_2 \sim I$, and $P_1 P_2 = 0$.

The *spectral representation* for a self-adjoint operator $A \in \mathbb{B}_{sa}(\mathscr{H})$ states that if $m = \min\{\lambda : \lambda \in \mathrm{sp}\,(A)\}$ and let $M = \max\{\lambda : \lambda \in \mathrm{sp}(A)\}$, then there exists a certain family of projections $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ known as the *spectral family* of A, which has the following properties

(a) $E_\lambda \le E_{\lambda'}$ for $\lambda \le \lambda'$;
(b) $E_{m-0} = 0$, $E_M = I$ and $E_{\lambda+} = E_\lambda$ for all $\lambda \in \mathbb{R}$;
(c) $A = \int_{m-0}^{M} \lambda d E_\lambda$;
(d) for every continuous complex-valued function f defined on $\mathbb{R}$, the Riemann–Stieltjes operator-valued integral

$$f(A) = \int_{m-0}^{M} f\,(\lambda)\, dE_\lambda$$

holds. This integral means that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\left\| f(A) - \sum_{k=1}^{n} f\left(\lambda'_k\right)\left(E_{\lambda_k} - E_{\lambda_{k-1}}\right) \right\| \le \varepsilon$$

when $\lambda_0 < m = \lambda_1 < \cdots < \lambda_{n-1} < \lambda_n = M, \lambda_k - \lambda_{k-1} \le \delta$ for $1 \le k \le n$, and $\lambda'_k \in (\lambda_{k-1}, \lambda_k)$ for $1 \le k \le n$.

If $\dim \mathscr{H} = n$, then A can be considered as a Hermitian matrix in $\mathbb{M}_n$. The Schur decomposition for matrices states that there exists a unitary matrix $U \in \mathbb{M}_n$ such that $A = U^*DU$, where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix whose diagonal entries are the eigenvalues λ_j $(1 \le j \le n)$ of A. If $E_j = U^*\mathrm{diag}(\underbrace{1, 1, \ldots, 1}_{j \text{ terms}}, 0, \ldots, 0)U$, then we get the spectral representation of A as follows:

$$A = U^*DU = \sum_{j=1}^{n} \lambda_j \Delta E_j,$$

where $\Delta E_j = E_j - E_{j-1}$ and $E_0 = 0$ are pairwise orthogonal projections; see also [116, p. 18]. Also, if f is a real-valued continuous function on an interval containing the eigenvalues of A, then

$$f(A) = \sum_{j=1}^{n} f(\lambda_j)\Delta E_j .$$

Here $f(A)$ is understood as $f(A)$ defined by functional calculus, or simply as $f(A) = U^*D'U$, where $D' = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$.

Now, we present a fundamental theorem.

Theorem 1.1.8 *Let $A \in \mathbb{B}(\mathscr{H})$. The following assertions are equivalent:*
(a) A is positive;
(b) A is of the form B^2 for some positive operator $B \in \mathbb{B}(\mathscr{H})$;
(c) A is of the form B^*B for some $B \in \mathbb{B}(\mathscr{H})$;
(d) $\langle Ax, x\rangle \geq 0$ holds for every $x \in \mathscr{H}$.

Proof $(a) \Longrightarrow (b)$. Since $\operatorname{sp}(A) \subseteq [0, \infty)$, the function $t^{1/2}$ is continuous on $\operatorname{sp}(A)$. Using the functional calculus for A, we get the positive operator $A^{1/2}$ satisfying $A = (A^{1/2})A^{1/2}$. So we reach (b) with $B = A^{1/2}$.
$(b) \Longrightarrow (c)$. It is clear.
$(c) \Longrightarrow (d)$. We have

$$\langle Ax, x\rangle = \langle B^*Bx, x\rangle = \langle Bx, Bx\rangle = \|Bx\|^2 \geq 0 .$$

$(d) \Longrightarrow (a)$. It follows from $\langle Ax, x\rangle \geq 0$ that

$$\langle Ax, x\rangle = \overline{\langle x, Ax\rangle} = \langle x, Ax\rangle = \langle A^*x, x\rangle .$$

Hence, $\langle (A - A^*)x, x\rangle = 0$ for all $x \in \mathscr{H}$. It follows from Proposition 1.1.2 that $A = A^*$. Let $A = A_+ - A_-$ be the Jordan decomposition of A. Since $\langle Ax, x\rangle \geq 0$, we have $\langle A_-x, x\rangle \leq \langle A_+x, x\rangle$. Replacing x with A_-x in the latter inequality, we get $0 \leq \langle A_-^3x, x\rangle = \langle A_-^2x, A_-x\rangle \leq \langle A_+A_-x, A_-x\rangle = 0$, because $A_+A_- = 0$. By Proposition 1.1.2, we have $A_-^3 = 0$. Applying functional calculus, we infer that $A_- = 0$. Therefore, $A = A_+ - A_- = A_+ \geq 0$. □

The operator B in Part (b) of the above theorem is unique and called the *positive square root* of A and is denoted by $A^{1/2}$. In addition, for each operator $A \in \mathbb{B}(\mathscr{H})$, the positive square root of the positive operator A^*A is denoted by $|A|$ and is called *absolute value* of A.

Let $A \in \mathbb{B}_{sa}(\mathscr{H})$. Then A is positive if there is a nonnegative real number c such that $\|A - cI\| \leq c$, since by utilizing the functional calculus for A, this norm inequality is equivalent to $\sup_{t\in \operatorname{sp}(A)} |t - c| \leq c$, which yields $t \geq 0$ and this, in turn, implies that $A \geq 0$. Conversely, in the same way, we observe that if $A \geq 0$, then $\|A - cI\| \leq c$ for all nonnegative real numbers $c \geq \|A\|/2$.

The following properties of the Löwner order and positive operators are frequently used without being referred to.

Theorem 1.1.9 *Let $A, B \in \mathbb{B}(\mathscr{H})$ be self-adjoint. Then,*

(i) if $A \geq 0$ and $t \in [0, \infty)$, then $tA \geq 0$;
(ii) if A and B are positive, then so is $A + B$;
(iii) If $-B \leq A \leq B$, then $\|A\| \leq \|B\|$;
(iv) if $A \leq B$, then $X^*AX \leq X^*BX$ for all $X \in \mathbb{B}(\mathscr{H})$;
(v) $0 < A$ if and only if there is $m > 0$ such that $0 < m \leq A$;
(vi) if $0 < A \leq B$, then B is invertible and $B^{-1} \leq A^{-1}$;
(vii) the set of positive operators is closed in $\mathbb{B}(\mathscr{H})$;
(viii) if $A, B \in \mathbb{B}(\mathscr{H})$ are positive and $AB = BA$, then $AB \geq 0$;
(ix) if $A, B \in \mathbb{B}(\mathscr{H})$ are positive, then $\mathrm{sp}(AB) \subseteq [0, \infty)$.

Proof (i) It is deduced from $\mathrm{sp}(tA) = t\,\mathrm{sp}(A)$.
(ii) By the note preceding this theorem, we have $\|A - \|A\|\| \leq \|A\|$ and $\|B - \|B\|\| \leq \|B\|$. Hence

$$\|(A + B) - (\|A\| + \|B\|)\| \leq \|A - \|A\|\| + \|B - \|B\|\| \leq \|A\| + \|B\|,$$

from which we conclude, from the note preceding this theorem, that $A + B \geq 0$.
(iii) By making use of (ii), we infer that $\leq$ is a partial order on $\mathbb{B}_{sa}(\mathscr{H})$. It follows from $t \leq \sup_{t\in\mathrm{sp}(A)} |t|$ and the functional calculus for C that $C \leq \|C\|I$ for any self-adjoint operator C. Thus, $-\|B\|I \leq A \leq B \leq \|B\|I$. By the functional calculus for B, $-\|B\| \leq t \leq \|B\|$ for all $t \in \mathrm{sp}(A)$. Thus, $\sup_{t\in\mathrm{sp}(A)} |t| \leq \|B\|$, whence $\|A\| \leq \|B\|$.
(iv) Since $0 \leq A \leq B$, we have $B - A \geq 0$, so

$$X^*(B - A)X = X^*(B - A)^{1/2}(B - A)^{1/2}X = ((B - A)^{1/2}X)^*((B - A)^{1/2}X) \geq 0,$$

from which we conclude that $X^*AX \leq X^*BX$.
(v) Let A be a positive invertible operator. Hence, $\mathrm{sp}(A) \subseteq (0, \infty)$. The spectrum of any operator is compact, so there is a real positive number m such that $0 < m \leq t$ for all $t \in \mathrm{sp}(A)$. Applying the functional calculus for A, we get $m \leq A$. The proof of the reverse assertion is similar.
(vi) Since $0 < m \leq A \leq B$ we infer that B is invertible. Using the functional calculus for B, we observe that the operator $B^{-1/2}$ corresponding to the continuous function $t^{-1/2}$ is well-defined. Employing (iv), we deduce from $A \leq B$ that $B^{-1/2}AB^{-1/2} \leq B^{-1/2}BB^{-1/2} = I$. Using the functional calculus for $B^{-1/2}AB^{-1/2}$ we see that $t \leq 1$ for all $t \in \mathrm{sp}(B^{-1/2}AB^{-1/2})$. Hence, $t^{-1} \geq 1$ for all $t \in \mathrm{sp}(B^{-1/2}AB^{-1/2})$. Another use of