Lecture Notes on
NON-SELF ADJOINT OPERATORS AND RELATED TOPICS

Leszek F. Demkowicz

March 25, 2024

Contents

List of Figures

Preface

The reported studies on non-self adjoint operators have been motivated with a model acoustical waveguide problem illustrated in Fig 1. We are looking for pressure p satisfying the Helmholtz equation, hard boundary condition (BC) at $x = 0$, initial condition at $z = 0$, nonlocal Dirichlet-to-Neumann (DtN) BC at $z = L$, and an impedance BC at $x = a$ with d being the impedance constant. The nonlocal DtN BC is formulated in terms of decomposition of the solu-

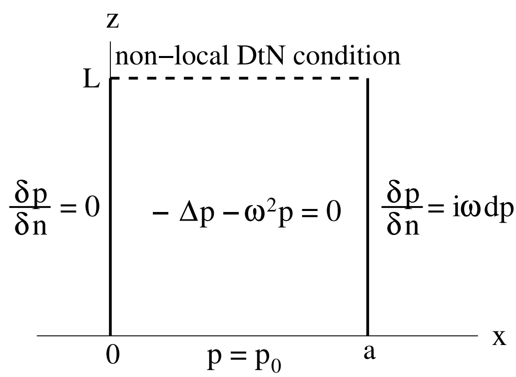

Figure 1: Model acoustical waveguide problem.

tion into the waveguide modes and it implies the separation of variables as a method of choice. Assuming $p = X(x)Z(z)$, we arrive at the eigenvalue problem in x,

$$\left\{\begin{array}{rll} -X'' - \omega^2 X & = \lambda X & x \in (0,a) \\ X' & = 0 & x = 0 \\ X' - i\omega d X & = 0 & x = a\,. \end{array}\right.$$

The corresponding standard variational formulation looks as follows.

$$\left\{\begin{array}{ll} X \in H^1(0,a) & \\ (X',Y') + (X,Y) - i\omega d X(a)\overline{Y(a)} & = (1+\lambda+\omega^2)(X,Y) \qquad Y \in H^1(0,a)\,. \end{array}\right.$$

What makes the problem non-standard is the fact that the operator on the left is *not* self-adjoint[1]. The standard Sturm-Liouville spectral theory does not apply and we even do not know whether the system of eigenvectors is complete in $H^1(0,a)$ (the energy space). However, we can rewrite the variational formulation in the operator form:

$$RX + CX = (1+\lambda+\omega^2)MX$$

[1]The adjoint has a flipped sign in front of the impedance term.

where R is the Riesz operator in $H^1(0,a)$, C is a compact operator representing the boundary term and M represents the compact embedding of $H^1(0,a)$ into $L^2(0,a)$. Upon applying the inverse R^{-1} to both sides, we obtain:

$$X + R^{-1}CX = (1 + \lambda + \omega^2)R^{-1}MX \, .$$

The left-hand side represents a compact perturbation of the identity operator in $H^1(0,a)$, and the operator on the right represents a compact and self-adjoint operator in $H^1(0,a)$. As we will learn, the completeness of the system of eigenvectors (modes) in $H^1(0,a)$ follows from the Second Keldyš Theorem 5.5 concluding these notes.

The notes have been extracted from the book of Gohberg and Krein [2] and the book of Levin [3]. As a starting point, we assume that the reader is familiar with our textbook [4]. Otherwise, the notes are self-contained and simply represent my reading of the two books. Many thanks to Peter Monk for making us aware of these results.

Acknowledgment. The work has been supported with AFOSR Grant FA9550-23-1-0103.

Leszek F. Demkowicz

Austin, March 25, 2024

Chapter 1

Preliminaries

Throughout these notes, X will denote a separable Hilbert space and all considered operators A go from X into itself, and are bounded,

$$A \in \mathcal{L}(X) := \{A \in L(X,X) \; : A \text{ is bounded} \} .$$

A *subspace* $V \subset X$ is *always* assumed to be *closed*.

1.1 ▪ Introduction

Invariant subspaces of an operator. A subspace $V \subset X$ is *an invariant subpace* of $A \in \mathcal{L}(X)$ if

$$v \in V \quad \Rightarrow \quad Av \in V .$$

Lemma 1.1.
Let $P \in \mathcal{L}(X)$ be a projection, i.e., $P^2 = P$, and $Q = I - P$ be the corresponding projection onto $\mathcal{N}(P)$.

(i) PX is invariant with respect to A iff $PAP = AP$.

(ii) Assume that PX is invariant wrt A. Then QX is invariant wrt A iff $PA = AP$.

(iii) Let $A = A^*$, and P be an orthogonal projection. Then

$$PAP = AP \quad \Leftrightarrow \quad PA = AP .$$

In other words, PX is invariant wrt $A = A^*$ iff P and A commute.

Proof.

(i) $(\Rightarrow)$

$$y \in PX \quad \Rightarrow \quad Ay \in PX \quad \Rightarrow \quad PAy = Ay$$

so, $PAPx = APx$, $x \in X$.
$(\Leftarrow)$ Assume

$$PAPx = APx, \quad x \in X .$$

Consequently,

$$y \in PX \quad \Rightarrow PAy = Ay \quad \Rightarrow \quad Ay \in PX .$$

(ii) (⇒) By (i),

$$\begin{aligned} (I-P)A(I-P) &= A(I-P) \quad \text{implies} \\ A - AP - PA + \underbrace{PAP}_{=AP} &= A - AP \quad \text{and, so} \\ -PA &= -AP\,. \end{aligned}$$

(⇐) Reverse the argument.

(iii) (⇒)

$$PAP = AP \quad \Rightarrow \quad (PAP)^* = (AP)^* \quad \Rightarrow \quad PAP = PA \quad \Rightarrow \quad AP = PA\,.$$

(⇐)

$$AP = PA \quad \Rightarrow \quad PAP = PPA = PA\,.$$

QED

Lemma 1.2.
V be an invariant subspace of $A \in \mathcal{L}(X)$. Then $V^\perp$ is an invariant subspace of A^*.

Proof. Let P be the orthogonal projection of X onto V, and $Q := I - P$. We have,

$$\begin{aligned} PAP = AP \quad &\Rightarrow \quad \underbrace{(I-P)}_{=Q} AP = 0 \\ &\Rightarrow \quad \underbrace{P}_{I-Q} A^*Q = 0 \\ &\Rightarrow \quad QA^*Q = A^*Q\,, \end{aligned}$$

i.e., $QX = V^\perp$ is invariant subspace of A^*. QED

Lemma 1.3.
Let V be an invariant subspace of $A \in \mathcal{L}(X)$, $P : X \to V$ the orthogonal projection, and $Q = I - P$.

(a) If two of the operators:

$$A, \quad PAP + Q, \quad P + QAQ$$

are invertible, then so is the third, and

(b)

$$\begin{aligned} (PAP+Q)^{-1} &= PA^{-1}P + Q \\ (QAQ+P)^{-1} &= QA^{-1}Q + P\,. \end{aligned}$$

Proof. $QAP = QPAP = 0$ implies

$$A = (P+Q)A(P+Q) = PAP + PAQ + QAQ\,.$$

Direct computation shows now that

$$A = (QAQ+P)(I+PAQ)(Q+PAP)\,. \tag{1.1.1}$$

Also,

$$(PAQ)^2 = PA\underbrace{QP}_{=0}AQ = 0$$

implies that $(I+PAQ)^{-1}$ exists and equals $I-PAQ$. Consequently, invertibility of two remaining operators in (1.1.1) implies the invertibility of the third. Once all operators are invertible,

$$\begin{aligned} A^{-1} &= (Q+PAP)^{-1}(I-PAQ)(QAQ+P)^{-1'} & \Rightarrow \\ A^{-1}(QAQ+P) &= (Q+PAP)^{-1}(I-PAQ) & \Rightarrow \\ PA^{-1}\underbrace{(QAQ+P)P}_{=0+P} &= P(Q+PAP)^{-1}\underbrace{(I-PAQ)P}_{=P}\,. \end{aligned}$$

and so,

$$PA^{-1}P = P(Q+PAP)^{-1}P\,.$$

Similarly,

$$\begin{aligned} (Q+PAP)A^{-1} &= (I-PAQ)(QAQ+P)^{-1} & \Rightarrow \\ \underbrace{Q(Q+PAP)}_{=Q+0}A^{-1}Q &= \underbrace{Q(I-PAQ)}_{=Q}(QAQ+P)^{-1}Q & \Rightarrow \\ QA^{-1}Q &= Q(QAQ+P)^{-1}Q\,. \end{aligned}$$

Now,

$$\begin{aligned} Q(Q+PAP)^{1} = Q \quad &\text{since} \quad Q = Q(Q+PAP) = Q \quad &\text{, and} \\ (Q+PAP)^{-1}Q = Q \quad &\text{since} \quad Q = (Q+PAP)Q = Q\,. \end{aligned}$$

Therefore,

$$\begin{aligned} P(Q+PAP)^{-1}P &= P(Q+PAP)^{-1} - \underbrace{P\underbrace{(Q+PAP)^{-1}Q}_{=Q}}_{=0} \\ &= (Q+PAP^{-1} - \underbrace{Q(Q+PAP)^{-1}}_{=Q} \\ &= (Q+PAP)^{1} - Q \end{aligned}$$

which implies

$$PA^{-1}P = (Q+PAP)^{-1} - Q$$

and, so,

$$(Q+PAP)^{-1} = PA^{-1}P + Q\,.$$

Proof of the other identity in b) is fully analogous. QED

Resolvents. Let $A \in \mathcal{L}(X)$. The inverse (if it exists and is continuous):

$$R(\lambda) := (A-\lambda I)^{-1} \in \mathcal{L}(X)\,,$$

is called the *resolvent* of operator A at λ. The collection of all λ's for which $R(\lambda)$ exists and is continuous, denote $\rho(A)$ is called the *resolvent set* of operator A. Complement of the resolvent set is called the *spectrum* of operator A. The resolvent set $\rho(A)$ is an open subset of complex plane $\mathbb{C}$. Indeed, let $\lambda_0 \in \rho(A)$. We have:

$$A - \lambda I = A - \lambda_0 I + (\lambda_0 - \lambda)I = (A - \lambda_0 I)(I + (\lambda_0 - \lambda)R(\lambda_0))$$

where, by the Neumann series argument,

$$(I + (\lambda_0 - \lambda)R(\lambda_0))^{-1} = \sum_{k=0}^{\infty}(\lambda - \lambda_0)^k R^k(\lambda_0)\,,$$

for all $|\lambda_0 - \lambda| < \|R(\lambda_0)\|^{-1}$. Consequently, $R(\lambda)$ exists, and

$$R(\lambda) = (I + (\lambda_0 - \lambda)R(\lambda_0))^{-1}R(\lambda_0) = \sum_{k=0}^{\infty}(\lambda - \lambda_0)^k R^{k+1}(\lambda_0)\,,$$

which proves that $R(\lambda)$ is a *holomorphic* (analytic) operator-valued function.

The Riesz integral. Let Γ be a counterclockwise oriented (ccw) contour enclosing a region G_Γ and lying in the interior of the resolvent set of operator $A \in \mathcal{L}(X)$. We define the *Riesz integral* as:

$$P_\Gamma := -\frac{1}{2\pi i}\int_\Gamma R(\lambda)\,d\lambda\,. \tag{1.1.2}$$

Theorem 1.4 (Riesz).
The following properties hold:

(i) P_Γ *is a projection commuting with* A *and, therefore,*

$$X = Y \oplus Z, \qquad Y = \mathcal{R}(P_\Gamma) = P_\Gamma X$$
$$Z = \mathcal{N}(P_\Gamma) = \mathcal{R}(I - P_\Gamma) = (I - P_\Gamma)X\,,$$

and Y, Z *are invariant subspaces of* A*;*

- *spectrum of* $A\big|_Y$ *is contained in* G_Γ*;*
- *spectrum of* $A\big|_Z$ *lies outside of* G_Γ*,*
- *if* $G_{\Gamma_1} \cap G_{\Gamma_2} = \emptyset$ *then* $P_{\Gamma_1}, P_{\Gamma_2}$ *are mutually orthogonal in the sense that:*

$$P_{\Gamma_1}P_{\Gamma_2} = P_{\Gamma_2}P_{\Gamma_1} = 0\,.$$

1.2 ▪ Polar Representation of a Bounded Operator

Lemma 1.5.
Let $A \in \mathcal{L}(X)$ *and* A^* *denote its adjoint. The following orthogonal decompositions hold:*

$$X = \overline{\mathcal{R}(A)} \overset{\perp}{\oplus} \mathcal{N}(A^*) = \overline{\mathcal{R}(A^*)} \overset{\perp}{\oplus} \mathcal{N}(A)\,. \tag{1.2.3}$$

Proof. Let $C := \overline{\mathcal{R}(A)}^{\perp}$.
$C \subset \mathcal{N}(A^*)$. Let $z \in C$. Then

$$(x, A^*z) = (Ax, z) = 0 \quad \forall\, x \in X \qquad \Rightarrow \qquad A^*z = 0\,.$$

$\mathcal{N}(A^*) \subset C$. Let $z \in \mathcal{N}(A^*)$ and $y \in \overline{\mathcal{R}(A)}$, i.e., $y = \lim_{n\to\infty} Ax_n,\ x_n \in X$. Then

$$0 = (x_n, A^*z) = (Ax_n, z) \to (y, z)$$

and, therefore, $z \in C$. The proof of the second decomposition is analogous. QED

Partially isometric operators. Operator $B \in \mathcal{L}(X)$ is *partially isometric* if it maps $\mathcal{N}(B)^\perp = \overline{\mathcal{R}(B^*)}$ *isometrically onto* $\mathcal{R}(B)$. This proves that range $\mathcal{R}(B)$ is also closed. By the Closed Range Theorem, range $\mathcal{R}(B^*)$ must be closed as well (and equal $\mathcal{N}(B)^\perp$), so

$$\mathcal{R}(B^*) \xrightarrow{B} \mathcal{R}(B)$$

is an isometry. One can show, comp. Exercise 1.2.1, that if B is partially isometric then so is adjoint B^*. Moreover, B^*B is the orthogonal projection of X onto range $\mathcal{R}(B^*)$, and BB^* is the orthogonal projection of X onto range $\mathcal{R}(B)$.

Polar decomposition. Let A be a bounded operator. Then composition A^*A is self-adjoint, and we can use[2] the Spectral Theorem for Self-Adjoint Operators (see [4], Theorem 6.11.1) to define the square root of A^*A, i.e., a bounded, self-adjoint operator H such that $H^2 = A^*A$. We have,

$$\|Au\|^2 = (Au, Au) = (A^*Au, u) = (H^2u, u) = (Hu, Hu) = \|Hu\|^2$$

which shows that operator

$$U\ :\ \mathcal{R}(H) \to \mathcal{R}(A), \quad Hx \to Ax\,,$$

is an isometric isomorphism. Extending U to $\overline{\mathcal{R}(H)}$ by continuity and setting $U = 0$ on null space $\mathcal{N}(H^*) = \mathcal{N}(H)$, we obtain a partially isometric operator:

$$U\ :\ X \to \overline{\mathcal{R}(A)}\,.$$

Lemma 1.6.
The following equalities hold:

$$\overline{\mathcal{R}(H)} = \overline{\mathcal{R}(H^2)} = \overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)}\,.$$

Proof. We need to prove only the last equality. The second one is trivial and the first one follows from the last with A replaced by H. One inclusion is immediate,

$$\mathcal{R}(A^*A) \subset \mathcal{R}(A^*) \quad \Rightarrow \quad \overline{\mathcal{R}(A^*A)} \subset \overline{\mathcal{R}(A^*)}\,.$$

To prove the reverse inclusion, pick an $x \in \overline{\mathcal{R}(A^*)}$. By Lemma 1.5,

$$x = \lim_{n\to\infty} A^*y_n \qquad y_n = X = \overline{\mathcal{R}(A)} \overset{\perp}{\oplus} \mathcal{N}(A^*)$$

$$y_n = \lim_{k\to\infty} Az_n^k + y_n^0 \qquad\qquad z_n^k \in X,\ y_n^0 \in \mathcal{N}(A^*)$$

$$A^*y_n = A^*(\lim_{k\to\infty} Az_n^k) = \lim_{k\to\infty} A^*Az_n^k$$

[2] Is there a more elementary argument, we could use ?

and, so,

$$x = \lim_{n\to\infty} \lim_{k\to\infty} A^*Az_n^k \,.$$

The double limit presents no problem. Let $\epsilon = \frac{1}{l}$. It follows from the definition of the limit that there exists y_n such that $\|x - A^*y_n\| < \frac{\epsilon}{2}$. In turn, there exists z_n^k such that $\|A^*y_n - A^*Az_n^k\| < \frac{\epsilon}{2}$. Set $x_l := z_n^k$. Then

$$\|x - A^*Ax_l\| \le \|x - A^*y_n\| + \|A^*y_n - A^*Az_n^k\| < \epsilon = \frac{1}{l}\,.$$

Consequently,

$$x = \lim_{l\to\infty} A^*Ax_l \quad \Rightarrow \quad x \in \overline{\mathcal{R}(A^*A)}\,.$$

QED

We arrive at the following result.

Theorem 1.7.
For every $A \in \mathcal{L}(X)$, there exist unique $U, H \in \mathcal{L}(X)$ such that:

$$A = UH, \quad H = H^*, \text{ and}$$
$$U \text{ maps } \overline{\mathcal{R}(A^*)} = \overline{\mathcal{R}(H)} \text{ isometrically onto } \overline{\mathcal{R}(A)}\,.$$

Corollary 1.8.
Let $A = UH$ be the spectral decomposition of a bounded operator A. The following identities hold:

$$(i) \quad U^*A = H\,,$$
$$(ii) \quad H_1 = UHU^*, \quad H = U^*H_1U \quad \text{where } H_1 := (AA^*)^{\frac{1}{2}}\,,$$
$$(iii) \quad A = H_1U, \quad H_1 = AU^*\,.$$

We leave the proof for Exercise 1.2.2.

Rank (dimension) of a bounded operator. We define the *rank* [3] of a bounded operator A as:

$$r(A) := \dim \overline{\mathcal{R}(A)}\,.$$

One can show that (Exercise 1.2.3):

$$r(A) = r((A^*A)^{\frac{1}{2}}) = r((AA^*)^{\frac{1}{2}}) = r(A^*)\,. \tag{1.2.4}$$

Every finite rank operator A can be represented in the form:

$$Au = \sum_{j=1}^{n} (u, \phi_j)\psi_j$$

where $\psi_j, j = 1, \ldots, n$, is a basis for $\mathcal{R}(A)$, and $\phi_j \in X$. Indeed, let $\beta_j \in \mathcal{R}(A)$ be a cobasis of ψ_j in the range $\mathcal{R}(A)$. Then,

$$Au = \sum_{j=1}^{n} (Au, \beta_j)\psi_j = \sum_{j=1}^{n} (u, \underbrace{A^*\beta_j}_{=:\phi_j})\psi_j\,.$$

[3] Gohberg calls it the dimension of operator.

Exercises

1.2.1. Let $A \in \mathcal{L}(X)$. Show that the following conditions are equivalent to each other.

(i) A is partially isometric.

(ii) A^* is partially isometric.

(iii) A^*A is the orthogonal projection of X onto $\mathcal{R}(A^*)$.

(iv) AA^* is the orthogonal projection of X onto $\mathcal{R}(A)$.

(5 points)

1.2.2. Prove Corollary 1.8.
(5 points)

1.2.3. Prove relation (1.2.4).
(5 points)

1.3 ▪ Regular Eigenvalues of a Bounded Operator

Regular eigenvalue of a bounded operator. An eigenvalue of λ_0 of operator $A \in \mathcal{L}(X)$ is said to be *regular*[4] iff, by definition,

(i) the algebraic multiplicity r of λ_0, i.e., the dimension of its generalized eigenspace X_{λ_0} is finite;

(ii) we have the decomposition:

$$X = X_{\lambda_0} \oplus Y_{\lambda_0}$$

where Y_{λ_0} is invariant subspace of A in which $A - \lambda_0 I$ has a bounded inverse.

Note that decomposition above must be unique, i.e., space Y_{λ_0} is unique. Indeed, let r be an integer for which operator $(A - \lambda_0 I)\big|_{X_{\lambda_0}}$ is nilpotent. Then,

$$(A - \lambda_o I)^r X = \underbrace{(A - \lambda_0 I)^r X_{\lambda_0}}_{=0} + (A - \lambda_0 I)^r Y_{\lambda_0} \,,$$

and invertibility of $A - \lambda_0 I$ on Y_{λ_0} implies that $(A - \lambda_0 I)^r Y_{\lambda_0} = Y_{\lambda_0}$. Consequently,

$$Y_{\lambda_0} = (A - \lambda_0 I)^r X \,.$$

Theorem 1.9.
The following conditions are equivalent to each other.

(i) λ_0 *is a regular eigenvalue of operator A.*

(ii) λ_0 *is an isolated point of spectrum of A, and projector*

$$P_{\lambda_0} \,:\, X \to X_{\lambda_0} \qquad P_{\lambda_0} := -\frac{1}{2\pi i}\int_{|\lambda - \lambda_0| = \delta} R(\lambda)\, d\lambda$$

has finite rank.

[4] Gohberg calls them *normal*.

If λ_0 is regular then the projector P_{λ_0} is surjective, i.e., rank of P_{λ_0} equals the algebraic multiplicity of λ_0.

Proof. (i) $\Rightarrow$ (ii) Define:

$$A_1 := A\big|_{X_{\lambda_0}}, \qquad A_2 := A\big|_{Y_{\lambda_0}} .$$

Let n be the smallest integer for which $(A_1 - \lambda_0 I)^n = 0$. Set $B_1 = A_1 - \lambda_0 I$. We have:

$$\begin{aligned} -(\lambda-\lambda_0)^n I &= \underbrace{B_1^n}_{=0} - (\lambda-\lambda_0)^n I \\ &= (A_1 - \lambda I)\left[(\lambda-\lambda_0)^{n-1} I + (\lambda-\lambda_0)^{n-1} B_1 + \ldots + B_1^{n-1}\right] \end{aligned}$$

Hence,

$$-(A_1 - \lambda I)^{-1} = (\lambda-\lambda_0)^{-1} I + \sum_{j=1}^{n-1} (\lambda-\lambda_0)^{-j-1} B_1^j .$$

On the other hand, as $A_2 - \lambda_0 I$ is invertible in Y_{λ_0}, we know[5] that for all λ such that

$$|\lambda - \lambda_0| < \frac{1}{\|(A_2 - \lambda_0 I)^{-1}\|} ,$$

the inverse $(A_2 - \lambda I)^{-1}$ exists and it can be represented by the convergent series:

$$(A_2 - \lambda I)^{-1} = R_0 + (\lambda-\lambda_0) R_0^2 + \ldots + (\lambda-\lambda_0)^n R_0^{n-1} + \ldots .$$

We thus obtain the following representation for the resolvent of operator A,

$$\begin{aligned} R(\lambda) &= (A - \lambda I)^{-1} \\ &= -\left[(\lambda-\lambda_0)^{-n} B_1^{n-1} + \ldots + (\lambda-\lambda_0)^{-2} B_1 + (\lambda-\lambda_0)^{-1} I\right] P \\ &\quad + \sum_{k=0}^{\infty} (\lambda-\lambda_0)^k R_0^{k-1} (I - P) \end{aligned}$$

where $P : X \to X_{\lambda_0}$ is the linear projection in the direction of Y_{λ_0}. The Riesz integral defines the desired projection:

$$-\frac{1}{2\pi i} \int_{|\lambda-\lambda_0|=\delta} R(\lambda)\, d\lambda =: P_{\lambda_0}$$

where δ is sufficiently small.
(ii) $\Rightarrow$ (i) Define:

$$X_{\lambda_0} := P_{\lambda_0} X, \qquad Y_{\lambda_0} := (I - P_{\lambda_0}) X .$$

By the Riesz Theorem 1.4,

- $X = X_{\lambda_0} \oplus Y_{\lambda_0}$,
- $A\big|_{X_{\lambda_0}} : X_{\lambda_0} \to X_{\lambda_0}$ has a unique eigenvalue $\lambda = \lambda_0$ and its algebraic multiplicity is bounded by $\dim X_{\lambda_0}$,
- $(A - \lambda_0 I)$ is invertible in Y_{λ_0}.

[5]A standard result for resolvents.

Finally, if the generalized subspace corresponding to λ_0 were only a proper subspace of X_{λ_0} then restriction of A to X_{λ_0} would have an eigenvalue different form λ_0, a contradiction. QED

Corollary 1.10. *If λ_0 is a regular eigenvalue of operator A then $\bar{\lambda}_0$ is a regular eigenvalue of adjoint A^* with the same algebraic multiplicity.*

1.4 ▪ Compact operators

By $\mathcal{L}_c(X)$ we denote the space of compact operators forming a *closed* subspace of $\mathcal{L}(X)$. Recall some fundamental properties of compact operators.

- Composition of a bounded and a compact operator (in any order) is compact,

$$K \in \mathcal{L}_c(X), A \in \mathcal{L}(X) \quad \Rightarrow \quad KA, AK \in \mathcal{L}_c(X)\,.$$

- A compact operator has at most a countable set of non-zero eigenvalues. If the number is infinite, they sequence converges to zero, $0 \neq \lambda_n \to 0$ ([4], Theorem 6.10.1).

- Each eigenvalue $\lambda \neq 0$ has a *finite algebraic multiplicity* defined as the dimension of the *generalized eigenspace* X_λ,

$$X_\lambda := \dim \mathcal{N}((A - \lambda I)^r)\,,$$

  for sufficiently large r. Elements of the generalized eigenspace are called *generalized eigenvectors*[6]. The actual eigenspace is a subspace of the generalized eigenspace. By $\nu(A)$ we will denote the sum of the algebraic multiplicities for all non-zero eigenvalues (may be infinite).

Volterra operator. A compact operator A is *a Volterra operator* if it does not have non-zero eigenvalues.

Lemma 1.11.
Let A be a compact operator. Assume that

$$X_A := \overline{\mathrm{span}\{\textit{generalized eigenvectors of } A\}} \neq X\,,$$

and let

$$Q_A \,:\, X \to X_A^\perp$$

be the orthogonal projection. Then $Q_A A Q_A$ is a Volterra operator.

Proof. Let $X_{\bar{\lambda}_j}(A^*)$ denote the generalized eigenspace for adjoint A^* corresponding to an eigenvalue $\bar{\lambda}_j$. Let

$$X = X_{\bar{\lambda}_j}(A^*) \oplus Y_j \quad Y_j = X_{\bar{\lambda}_j}^\perp(A)$$

be the *unique* decomposition of X reducing operator A. Define

$$Y = \bigcap_j Y_j\,.$$

[6] Gohberg calls them *root vectors*.

And so,

$$f \in Y \quad \Leftrightarrow \quad f \perp X_{\lambda_j}(A) \quad \forall\, j\,.$$

In other words,

$$X = Y \overset{\perp}{\oplus} X_A\,.$$

Each subspace Y_j is invariant wrt A^* and, therefore, so is subspace Y. Any eigenvector of $A^*\big|_Y : Y \to Y$ corresponding to a non-zero eigenvalue would also have to be an eigenvector for A^* which is impossible. In other words, $A^*\big|_Y$ is Volterra. But then operator $Q_A A Q_A$ is Volterra as well. Indeed,

$$Q_A A Q_A v = \lambda v \quad \Rightarrow \quad A Q_A v = Q_A A Q_A v = \lambda Q_A v$$

implies that $Q_A v$ is an eigenvector for $A^*\big|_Y$, a contradiction. Finally,

$$Q_A A Q_A \quad (= (Q_A A^* Q_A)^*)$$

is Volterra as well. QED

Chapter 2

Weyl's Results

2.1 ▪ Weyl's Lemmas

The first Weyl lemma is purely algebraic.

Lemma 2.1 (First Weyl Lemma). *Let A be a compact operator, and $s_j, j = 1, \dots$ denote its singular values in the decreasing order. Let $\phi_1, \dots, \phi_n$ be arbitrary elements of X. Then*

$$\det(A\phi_j, A\phi_k) \le s_1^2 \dots s_n^2 \det(\phi_j, \phi_k) \quad 1 \le j, k \le n\,. \tag{2.1.1}$$

Proof. Let e_j, $j = 1, 2, \dots$ be a complete orthonormal system of eigenvectors of A^*A. Expanding ϕ_j into e_i's, we obtain:

$$\begin{aligned} \phi_j \quad &= \sum_{i=1}^{\infty} (\phi_j, e_i) e_i\,, \\ A^*A\phi_j \quad &= \sum_{i=1}^{\infty} s_i^2 (\phi_j, e_i)\, e_i\,. \end{aligned}$$

Consequently,

$$\underbrace{(A\phi_j, A\phi_k)}_{=:A_{jk}} = (A^*A\phi_j, \phi_k) = \sum_{i=1}^{\infty} s_i^2 (\phi_j, e_i)\,(e_i, \phi_k) = \sum_{i=1}^{\infty} s_i^2 (\phi_j, e_i)\,\overline{(\phi_k, e_i)}\,,$$

or,

$$A_{jk} = \sum_{i=1}^{\infty} B_{ji} \overline{B_{ki}} = \sum_{i=1}^{\infty} B_{ji} B_{ik}^*$$

where

$$B_{ji} = s_i \underbrace{(\phi_j, e_i)}_{=:\Phi_{ji}}\,.$$

By the Binet-Cauchy Theorem (comp. Exercise 2.1.1),

$$\det A = \sum_{1 \le r_1 < r_2 < \dots, r_n < \infty} \det \begin{pmatrix} B_{1r_1} & B_{1,r_2} & \dots & B_{1r_n} \\ \vdots & & & \vdots \\ B_{nr_1} & B_{n,r_2} & \dots & B_{nr_n} \end{pmatrix} \det \begin{pmatrix} B^*_{r_1 1} & B^*_{r_2 1} & \dots & B^ast_{r_n,1} \\ \vdots & & & \vdots \\ B^*_{r_1 n} & B^*_{r_2 n} & \dots & B^*_{r_n n} \end{pmatrix}$$

By the multilinearity of determinant and the monotonicity of s_j,

$$\det\begin{pmatrix} B_{1r_1} & B_{1,r_2} & \dots & B_{1r_n} \\ \vdots & & & \vdots \\ B_{nr_1} & B_{n,r_2} & \dots & B_{nr_n} \end{pmatrix} = \det\begin{pmatrix} s_{r_1}\Phi_{1r_1} & s_{r_2}\Phi_{1,r_2} & \dots & s_{r_n}\Phi_{1r_n} \\ \vdots & & & \vdots \\ s_{r_1}\Phi_{nr_1} & s_{r_2}\Phi_{n,r_2} & \dots & s_{r_n}\Phi_{nr_n} \end{pmatrix}$$

$$= s_{r_1}s_{r_2}\dots s_{r_n}\det\begin{pmatrix} \Phi_{1r_1} & \Phi_{1,r_2} & \dots & \Phi_{1r_n} \\ \vdots & & & \vdots \\ \Phi_{nr_1} & \Phi_{n,r_2} & \dots & \Phi_{nr_n} \end{pmatrix}$$

$$\leq s_1 s_2 \dots s_n \det\begin{pmatrix} \Phi_{1r_1} & \Phi_{1,r_2} & \dots & \Phi_{1r_n} \\ \vdots & & & \vdots \\ \Phi_{nr_1} & \Phi_{n,r_2} & \dots & \Phi_{nr_n} \end{pmatrix}.$$

Note that, by semipositive definitness of A_{jk} and the Sylvester criterion, all involved determinants are non-negative. Consequently, by the Binet-Cauchy Theorem again,

$$\det A \leq s_1^2 \dots s_n^2 \sum_{1\leq r_1<r_2<\dots,r_n<\infty} \det\begin{pmatrix} \Phi_{1r_1} & \Phi_{1,r_2} & \dots & \Phi_{1r_n} \\ \vdots & & & \vdots \\ \Phi_{nr_1} & \Phi_{n,r_2} & \dots & \Phi_{nr_n} \end{pmatrix} \det\begin{pmatrix} \Phi^*_{r_1 1} & \Phi^*_{r_2 1} & \dots & \Phi^*_{r_n 1} \\ \vdots & & & \vdots \\ \Phi^*_{r_1 n} & \Phi^*_{r_2 n} & \dots & \Phi^*_{r_n n} \end{pmatrix}$$

$$= s_1^2 \dots s_n^2 \det\left(\sum_{i=1}^{\infty} \Phi_{ji}\Phi^*_{ik}\right).$$

QED.

Recall Spectral Theorem for Compact and Normal Operators ([4], Theorems 6.10.2 and 6.10.3). If λ_n denote the eigenvalues of a compact and normal operator A ($\lambda_n \to 0$ if the operator is of infinite rank), the corresponding finite-dimensional eigenspaces X_n are orthogonal to each other, and the operator can be represented in the form:

$$Au = \sum_{n=1}^{\infty} \lambda_n P_n u \quad \text{(convergence in operator norm)}$$

where $P_n : X \to X_n$ are the orthogonal projections onto the eigenspaces. In other words, one can always select (unit) eigenvectors e_i in such a way that

$$Au = \sum_{i=1}^{\infty} \lambda_i (u, e_i) e_i .$$

If we complement the eigenvectors e_i with an an additional orthonormal basis for null space $\mathcal{N}(A)$, we obtain an orthonormal basis for space X (Resolution of Identity). Recall that orthogonal projections and self-adjoint operators are examples of normal operators. The spectral theorem says that *every compact and normal operator* is in fact a sum of orthogonal projections. The spectral representation for the adjoint A^* shares the same eigenvectors with the corresponding eigenvalues being the complex conjugates of λ_i,

$$A^* u = \sum_{i=1}^{\infty} \overline{\lambda}_i (u, e_i) e_i .$$

It follows immediately that

$$A^* A u = \sum_{i=1}^{\infty} |\lambda_i|^2 (u, e_i) e_i ,$$

i.e., vectors e_i are also the eigenvectors of A^*A, and the singular values of A are

$$s_i(A) = |\lambda_i| = |(Ae_i, e_i)|\,.$$

The next lemma establishes the uniqueness of the relation above. If an orthonormal system is related to the singular values by the relation above then the operator is normal and e_i's must be its eigenvectors.

Lemma 2.2. *Let A be a compact operator, and s_j denote its singular values in the decreasing order. Let $r(A)$ denote the rank of operator A (possibly infinite). Let $\phi_j, j = 1, \dots, r(A)$ be an arbitrary orthonormal system in X such that*

$$|(A\phi_j, \phi_j)| = s_j(A) \quad j = 1, \dots, r(A)\,.$$

Then operator A is normal, and ϕ_j are eigenvectors of A forming a complete system in $\overline{\mathcal{R}(A)}$.

Proof. We first use the min-max properties of eigenvalues of a self-adjoint operator to establish that ϕ_j's are eigenvectors of A^*A corresponding to its eigenvalues s_j^2. We start with the first eigenvalue,

$$s_1^2(A) = |(A\phi_1, \phi_1)|^2 \le \|A\phi_1\|^2 = (A^*A\phi_1, \phi_1) \le \max_{\|\phi\|=1} (A^*A\phi, \phi) = s_1^2(A)\,.$$

Consequently, all inequalities are actually equalities which proves that

$$A^*A\phi_1 = s_1^2(A)\phi_1\,.$$

Similarly,

$$s_2^2(A) = |(A\phi_2, \phi_2)|^2 \le \|A\phi_2\|^2 = (A^*A\phi_2, \phi_2) \le \max_{\|\phi\|=1,\,(\phi,\phi_1)=0} (A^*A\phi, \phi) = s_2^2(A)$$

which shows that

$$A^*A\phi_2 = s_2^2(A)\phi_2\,.$$

By induction,

$$A^*A\phi_j = s_j^2(A)\phi_j \quad j = 1, 2, \dots r(A)\,,$$

and, by the same argument,

$$AA^*\phi_j = s_j^2(A)\phi_j \quad j = 1, 2, \dots r(A)\,.$$

Let now

$$X_0 := \{u \in X \,:\, (u, \phi_j) = 0 \quad j = 1, \dots, r(A)\}$$

denote the subspace of vectors orthogonal to all eigenvectors ϕ_j. Resolution of identity for the compact self-adjoint operator A^*A (see [4], Theorem 6.10.3) implies that $X_0 = \mathcal{N}(A^*A)$ and, therefore, $X_0 = \mathcal{N}(A^*A) = \mathcal{N}(A)$ (Exercise 2.1.2). Repeating the argument for adjoint A^*, we learn that also $X_0 = \mathcal{N}(A^*)$.

We claim that any $f \in \overline{\mathcal{R}(A)}$ can be decomposed into vectors ϕ_j, i.e.,

$$f = \sum_{j=1}^{r(A)} (f, \phi_j)\phi_j\,.$$

Indeed,

$$(f - \sum_{j=1}^{r(A)} (f, \phi_j)\phi_j, \phi_i) = (f, \phi_i) - \sum_{j=1}^{r(A)} (f, \phi_j)\delta_{ji} = 0\,,$$

and, at the same time,

$$(f - \sum_{j=1}^{r(A)} (f, \phi_j)\phi_j, v) = (f, v) = 0 \quad \forall\, v \in X_0 = \mathcal{N}(A^*)\,,$$

since $\overline{\mathcal{R}(A)} = \mathcal{N}(A^*)^\perp$. As $X_0 = \mathcal{N}(A^*A)$, and (resolution of identity)

$$X = \overline{\text{span}\{e_j\}} \oplus \mathcal{N}(A^*A)\,,$$

there must be $f - \sum_{j=1}^{r(A)} (f, \phi_j)\phi_j = 0$, the claim has been proved. In particular, for $f = A\phi_j$,

$$A\phi_j = \sum_{k=1}^{r(A)} (A\phi_j, \phi_k)\phi_k$$

and, consequently,

$$\|A\phi_j\|^2 = (A\phi_j, A\phi_j) = \sum_{k=1}^{r(A)} |(A\phi_j, \phi_k)|^2\,.$$

But,

$$|(A\phi_j, \phi_j)|^2 = s_j^2(A) = (A\phi_j, A\phi_j)$$

and, therefore,

$$(A\phi_j, \phi_k) = 0 \quad \forall\, k \neq j\,.$$

In conclusion,

$$A\phi_j = (A\phi_j, \phi_j)\phi_j \quad j = 1, \ldots, r(A)$$

and, by the same token,

$$A^*\phi_j = (A^*\phi_j, \phi_j)\phi_j \quad j = 1, \ldots, r(A)\,.$$

Consequently,

$$\begin{aligned} Au &= \sum_{j=1}^{r(A)} (Au, \phi_j)\phi_j = \sum_{j=1}^{r(A)} (u, A^*\phi_j)\phi_j \\ &= \sum_{j=1}^{r(A)} (u, (A^*\phi_j, \phi_j)\phi_j)\phi_j = \sum_{j=1}^{r(A)} \overline{(A^*\phi_j, \phi_j)}(u, \phi_j)\phi_j \\ &= \sum_{j=1}^{r(A)} (A\phi_j, \phi_j)(u, \phi_j)\phi_j\,. \end{aligned}$$

The operator A is thus normal, and $(\underbrace{(A\phi_j, \phi_j)}_{=:\lambda_j}, \phi_j)$ are its eigenpairs. QED.

Let A be a compact operator. Recall that $\nu(A)$ denotes the sum of the algebraic multiplicities of eigenvalues λ_j of operator A, i.e.,

$$\nu(A) = \sum_j \dim X_j$$

where X_j denote the generalized eigenspaces of operator A corresponding to eigenvalues λ_j, i.e., $X_j = \mathcal{N}((A-\lambda_j)^{r_j})$, for some (finite) r_j.

Define now[7]:

$$X_A := \overline{\text{span}\{\text{generalized eigenspaces of } A\}}$$

and let $\hat{A}$ denote the reduction of operator A to X_A,

$$\hat{A} := A|_{X_A} \; : \; X_A \to X_A \, .$$

The following very useful result holds.

Lemma 2.3 (Schur's lemma).
Let A be a compact operator. There exists an orthonormal basis ω_j, $j = 1, \ldots, \nu(A)$, for subspace X_A in which the reduced operator $\hat{A}$ has an upper triangular matrix representation,

$$A\omega_j = \alpha_{j1}\omega_1 + \alpha_{j2}\omega_2 + \ldots + a_{jj}\omega_j \quad j = 1, 2, \ldots, \nu(A) \tag{2.1.2}$$

where $\alpha_{jj} = (A\omega_j, \omega_j) = \lambda_j(A)$.

Proof. Let λ be an eigenvalue of A with the corresponding generalized eigenspace X_λ. Choose Jordan chains for a basis for X_λ,

$$\begin{aligned} A\phi_1 \quad &= \lambda\phi_1 \\ A\phi_k \quad &= \lambda\phi_k + \phi_{k-1} \quad k = 2, \ldots \end{aligned}$$

Orthonormalize now the (collective) basis ϕ_j using the Gram-Schmidt orthonormalization to obtain system ω_j. It follows from the Gram-Schmidt procedure that each vector ω_j is a linear combination of vectors $\phi_1, \ldots, \phi_j$ and, conversely, each vector ϕ_j is a linear combination of vectors $\omega_1, \ldots, \omega_j$,

$$\phi_j = \beta_{jj}\omega_j + \sum_{k=1}^{j-1} \beta_{kj}\omega_k \qquad \beta_{jj} \neq 0 \, .$$

The relation,

$$A\phi_j = \lambda_j\phi_j \qquad \underbrace{(+\phi_{j-1})}_{\text{possible extra term}}$$

translates into:

$$\beta_{jj}A\omega_j + \sum_{k=1}^{j-1}\beta_{kj}A\omega_k = \lambda_j\beta_{jj}\omega_j + \lambda_j\sum_{k=1}^{j-1}\beta_{kj}\omega_k \quad (+\sum_{k=1}^{j-1}\beta_{kj-1}\omega_k) \, .$$

As, for each $k < j$, $A\omega_k$ is a linear combination of vectors $\omega_l, l \le k < j$, multiplying both sides with ω_j yields,

$$\beta_{jj}(A\omega_j, \omega_j) = \lambda_j\beta_{jj}\underbrace{(\omega_j, \omega_j)}_{=1}$$

[7] Kohberg calls generalized subspaces *root subspaces*.

from which the equality $(A\omega_j, \omega_j) = \lambda_j$ follows. QED

As we may have multiple Jordan chains for an eigenvalue, and we do not attempt to order them, the Schur orthonormal system may not be unique.

Lemma 2.4 (Second Weyl Lemma).
Let A be a compact operator. Then

$$|\lambda_1(A)\,\lambda_2(A)\ldots\lambda_n(A)| \leq s_1(A)\,s_2(A)\ldots s_n(A) \quad \forall\, n = 1,\ldots,\nu(A)\,. \tag{2.1.3}$$

If $\nu(A) = r(A)(\leq \infty)$ then inequality in (2.1.3) turns into equality for all n, if and only if operator A is normal.

Proof. Let $\omega_j,\ j = 1,\ldots,\nu(A)$ denote an orthornomal Schur system for operator A, i.e.,

$$A\omega_j = \alpha_{j1}\omega_1 + \alpha_{j2}\omega_2 + \ldots + \alpha_{jj}\omega_j \quad j = 1,\ldots,\nu(A)$$

where

$$\alpha_{jj} = (A\omega_j, \omega_j) = \lambda_j(A)\,.$$

By the First Weyl Lemma,

$$\det\left((A\omega_j A\omega_k)_1^n\right) \leq s_1^2(A)s_2^2(A)\ldots s_n^2(A) \quad n = 1,2,\ldots,\nu(A)\,. \tag{2.1.4}$$

The Schur representation and the orthonormality of ω_j imply that

$$(A\omega_j, A\omega_k) = \sum_{l=1}^{\min\{j,k\}} (A\omega_j, \omega_l)\overline{(A\omega_k, \omega_l)}.$$

Consequently, by the Cauchy theorem for determinants,

$$\det\left((A\omega_j, A\omega_k)_1^n\right) = \det\left((A\omega_j, \omega_l)_1^n\right)\,\det\left(\overline{(A\omega_k, \omega_l)}_1^n\right) = |\det\left((A\omega_j, \omega_l)_1^n\right)|^2\,.$$

The upper triangular Schur representation implies now that

$$\det\left((A\omega_j, A\omega_k)_1^n\right) = |\lambda_1(A)|^2\ldots|\lambda_n(A)|^2$$

which, along with inequality (2.1.4), proves the first assertion of the lemma.

To prove the second assertion, it is sufficient to notice that the equalities in (2.1.3) imply that

$$s_j(A) = |\lambda_j(A)| = |(A\omega_j, \omega_j)|$$

and apply Lemma 2.2. QED.

Exercises

2.1.1. Formulate and prove the Binet Cauchy Theorem. (5 points)

2.1.2. Show that $\mathcal{N}(A^*A) = \mathcal{N}(A)$. (1 point)

2.2 ▪ Weyl's Majorant Theorem

Lemma 2.5 (Weyl, Hardy, Littlewood, Polya).
Let

$$\Phi : [-\infty, \infty) \to \mathbb{R}, \quad \Phi(-\infty) = 0$$

be a convex function.

(i) Let $a_j, b_j \in \mathbb{R}$, $j = 1, \ldots, \omega (\le \infty)$ *be two (weakly) decreasing sequences such that*

$$\sum_{j=1}^{k} a_j \le \sum_{j=1}^{k} b_j \quad k = 1, \ldots, \omega\,.$$

Then,

$$\sum_{j=1}^{k} \Phi(a_j) \le \sum_{j=1}^{k} \Phi(b_j) \quad k = 1, \ldots, \omega\,.$$

(ii) If, additionally, Φ *is strictly convex, then*

$$\sum_{j=1}^{\omega} \Phi(a_j) = \sum_{j=1}^{\omega} \Phi(b_j) \quad \Leftrightarrow \quad a_j = b_j \quad j = 1, \ldots, \omega\,.$$

Proof. Let $\Phi'(x)$, $x \in (-\infty, \infty)$ denote the *left* derivative of Φ. Recall its standard properties:

- the derivative $\Phi'(x)$ exists everywhere,
- $\Phi'(x) \ge 0$,
- the derivative is (weakly) increasing in x.

We claim the following relation between $\Phi(x)$ and its derivative:

$$\Phi(x) = \int_{-\infty}^{\infty} (x - \mu)_+ \, d\Phi'(\mu) = \int_{-\infty}^{x} (x - \mu) \, d\Phi'(\mu)\,.$$

Let $N > 0$. Integration by parts,

$$0 \le \int_{-N}^{x} (x - \mu) \, d\Phi'(\mu) = \int_{-N}^{N} \Phi'(\mu) \, d\mu - (x + N)\Phi'(-N)$$

implies

$$(x + N)\Phi'(-N) \le \int_{-N}^{N} \Phi'(\mu) \, d\mu = \Phi'(x) - \Phi(-N) \le \Phi(x)$$

and, in turn,

$$\Phi'(-N) \le \frac{\Phi(x)}{x + N}\,.$$

Consequently,

$$\lim_{N \to \infty} \Phi'(-N) = 0\,.$$

Similarly,

$$N\Phi'(-N) \le \Phi(x) - x\Phi'(-N)$$

implies

$$\limsup_{N\to\infty} N\Phi'(-N) \le \Phi(x) \quad (<\infty)\,.$$

Passing with $x \to -\infty$, we obtain:

$$\limsup_{N\to\infty} \underbrace{N\Phi'(-N)}_{\ge 0} \le 0$$

and, therefore,

$$\lim_{N\to\infty} N\Phi'(-N) = 0\,.$$

We can conclude that

$$\lim_{N\to\infty} (x+N)\Phi'(-N) = 0\,.$$

Finally, passing with $N \to \infty$ in

$$\int_{-N}^{x} (x-\mu)\, d\Phi'(\mu) = \underbrace{\int_{-N}^{N} \Phi'(\mu)\, d\mu}_{=\Phi(x)-\Phi(-N)} - (x+N)\Phi'(-N)$$

we obtain the desired representation result.

The representation for $\Phi(x)$ implies :

$$\begin{aligned} \sum_{j=1}^{k} \Phi(a_j) \quad &= \int_{-\infty}^{\infty} \underbrace{\sum_{j=1}^{k} (a_j - x)_+}_{=:A_k(x)} \, d\Phi'(x)\,, \\ \sum_{j=1}^{k} \Phi(b_j) \quad &= \int_{-\infty}^{\infty} \underbrace{\sum_{j=1}^{k} (b_j - x)_+}_{=:B_k(x)} \, d\Phi'(x)\,. \end{aligned}$$

We claim that

$$A_k(x) \le B_k(x) \quad -\infty < x < \infty, \quad k = 1, 2, \ldots\,.$$

We proceed by considering three cases.
Case: $x \le \min\{a_k, b_k\}$ follows directly from $a_j \le b_j$.
Case: $b_k \le x$ is satisfied trivially, both sides are zero.
Case:

$$a_{q+1} \le x < a_q \quad \text{and} \quad b_{p+1} \le x < b_p \qquad \text{for some } p, q \le k\,.$$

For $p \ge q$ we have:

$$\begin{aligned} A_k(x) \quad &= \sum_{j=1}^{q} (a_j - x) = \sum_{j=1}^{q} a_j - qx \\ &\le \sum_{j=1}^{q} b_j - qx + \underbrace{(b_{q+1} - x)}_{\ge 0} + \ldots + \underbrace{(b_p - x)}_{\ge 0} = B_k(x) \end{aligned}$$

whereas for $p < q$,

$$\begin{aligned} A_k(x) &= \sum_{j=1}^{q} a_j - qx \\ &\le \sum_{j=1}^{q} b_j - qx - \underbrace{(b_{p+1} - x)}_{\le 0} - \ldots - \underbrace{(b_q - x)}_{\le 0} \\ &= \sum_{j=1}^{p} b_j - px = B_k(x)\,. \end{aligned}$$

We have proved the first part of the lemma. We prove the second part for the more difficult case $\omega = \infty$. We have:

$$A_\infty(x) = \sum_{j=1}^{\infty}(a_j - x)_+ \le b_\infty(x) = \sum_{j=1}^{\infty}(b_j - x)_+$$

in the sense that, if the right-hand side is finite then so its the left-hand side, and the inequality holds. Consequently,

$$\sum_{j=1}^{\infty} \Phi(a_j) = \int_{-\infty}^{\infty} A_\infty(x)\, d\Phi'(x) \le \int_{-\infty}^{\infty} B_\infty(x)\, d\Phi'(x) = \sum_{j=1}^{\infty} \Phi(b_j)\,.$$

If the extreme sides are equal then

$$\int_{-\infty}^{\infty} (B_\infty - A_\infty)\ \underbrace{d\Phi'(x)}_{>0} = 0$$

implies $A_\infty = B_\infty$ and, therefore, $a_j = b_j$ for all j. QED

We arrive at the main result of this section.

Theorem 2.6 (Weyl's Majorant Theorem).
Let A be a compact operator, and λ_j, s_j denote its eigen- and singular values, resp.

(i) Let $f(x), x \in [0,\infty), f(0) = 0$ be a real-valued function such that

$$\Phi(t) := f(e^t), \quad t \in (-\infty, \infty)$$

is a convex function. Then

$$\sum_{j=1}^{k} f(|\lambda_j|) \le \sum_{j=1}^{k} f(s_j) \quad k = 1, \dots, \nu(A).$$

(ii) If function $\Phi(t)$ is strictly convex then equality:

$$\sum_{j=1}^{\nu(A)} f(|\lambda_j|) = \sum_{j=1}^{\infty} f(s_j) \quad (< \infty)$$

holds if and only if operator A is normal.

Proof. Second Weyl lemma implies that

$$\sum_{j=1}^{k} \ln|\lambda_j| \le \sum_{j=1}^{k} \ln s_j .$$

Use Lemma 2.5 for $a_j = \ln|\lambda_j|$, $b_j = \ln s_j, j = 1, \dots, \nu(A)$, to conclude that

$$\sum_{j=1}^{k} f(|\lambda_j|) \le \sum_{j=1}^{k} f(s_j) .$$

If the inequality above turns into equality then $|\lambda_j| = s_j$, and the second Weyl lemma implies the result. QED

Corollary 2.7. *Choosing $f(x) = x^p, p > 0$ in Theorem 2.6, we obtain*

$$\sum_{j=1}^{k} |\lambda_j(A)|^p \le \sum_{j=1}^{k} s_j^p(A) \quad k = 1, \dots, \nu(A).$$

Choosing $f(x) = \ln(1 + rx), r > 0$ in Theorem 2.6, we obtain

$$\prod_{j=1}^{k} (1 + r|\lambda_j(A)|) \le \prod_{j=1}^{k} (1 + r s_j(A)) \quad k = 1, \dots, \nu(A).$$

2.3 ▪ Nuclear Operators

Operator A is called *nuclear* if $A \in \mathcal{C}_1$, i.e., A is compact and

$$\sum_j s_j(A) < \infty .$$

We say that operator A has a *finite trace* if the series

$$\sum_{j=1}^{\infty}(A\chi_j,\chi_j)$$

converges to a finite value *for any* orthormal basis χ_j. Since a permutation of an orthonormal basis is also an orthonormal basis, the operator has a finite trace if and only if the series above converges *absolutely* for any orthonormal basis χ_j.

Lemma 2.8. *Let H be a bounded linear and nonnegative operator. Then the sum*

$$\sum_{j=1}^{\infty}(H\chi_j,\chi_j)$$

has the same (finite or infinite) value for any orthonormal basis χ_j of space X.

The operator H belongs to $\mathcal{C}_1$ if and only if the sum above is finite.

Proof. Let ϕ_k be another orthonormal basis. We have,

$$\begin{aligned}\sum_{j=1}^{\infty}(H\chi_j,\chi_j) &= \sum_{j=1}^{\infty}\|H^{\frac{1}{2}}\chi_j\|^2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}|(H^{\frac{1}{2}}\chi_j,\phi_k)|^2 \\ &= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}|(H^{\frac{1}{2}}\phi_k,\chi_j)|^2 = \sum_{k=1}^{\infty}\|H^{\frac{1}{2}}\phi_k\|^2 = \sum_{k=1}^{\infty}(H\phi_k,\phi_k)\,.\end{aligned}$$

Note that the sums above may be finite or infinite.

Assume now that the sum above is finite for an orthonormal basis χ_j. We claim first that the operator must be compact. Indeed, define a series of finite rank operators

$$K_n x := \sum_{j=1}^{n}(x,\chi_j)H^{\frac{1}{2}}\chi_j\,.$$

Then

$$\begin{aligned}\|H^{\frac{1}{2}}x - K_n x\| &= \|\sum_{j=n+1}^{\infty}(x,\chi_j)H^{\frac{1}{2}}\chi_j\| \le \sum_{j=n+1}^{\infty}\|H^{\frac{1}{2}}\chi_j\||(x,\chi_j)| \\ &\le (\sum_{j=n+1}^{\infty}\|H^{\frac{1}{2}}\chi_j\|^2)^{\frac{1}{2}}\,\|x\| = (\sum_{j=n+1}^{\infty}(H\chi_j,\chi_j))^{\frac{1}{2}}\,\|x\|\end{aligned}$$

which proves that K_n converge to $H^{\frac{1}{2}}$ in the operator norm. Thus, $H^{\frac{1}{2}}$ and, therefore H as well, are compact. Choosing for χ_j the complete system of eigenvectors of H,

$$\infty > \sum_{j=1}^{\infty}(H\chi_j,\chi_j) = \sum_{j=1}^{\infty}\lambda_j(H)$$

we learn that $H \in \mathcal{C}_1$. Vice versa, if $H \in \mathcal{C}_1$ then the sum above is finite for the eigensystem. QED.

We generalize now the result to arbitrary bounded operators.

Theorem 2.9.
A bounded linear operator A has a finite matrix trace if and only it is nuclear, i.e., $A \in \mathcal{C}_1$. Then the sum

$$\sum_{j=1}^{\infty}(A\chi_j, \chi_j) \tag{2.3.5}$$

takes the same value for any orthonormal basis χ_j.

Proof. QED.

Matrix trace of an operator. Let $A \in \mathcal{C}_1$. The sum (2.3.5) is called the *matrix trace of operator A*, denoted $\mathrm{sp}A$.

The following two properties follow immediately from the definition.

$$\mathrm{sp}(\alpha A + \beta B) = \alpha\,\mathrm{sp}\,A + \beta\,\mathrm{sp}\,B\,. \tag{2.3.6}$$

$$\mathrm{sp}\,A^* = \overline{\mathrm{sp}\,A}\,. \tag{2.3.7}$$

Hilbert-Schmidt Operators. A bounded linear operator A is *a Hilbert-Schmidt operator* iff

$$\mathrm{sp}(A^*A) < \infty\,.$$

Note that

$$\mathrm{sp}(A^*A) = \sum_{j=1}^{r(A)} \lambda_j(A^*A) = \sum_{j=1}^{\infty} s_j^2(A)$$

and

$$\mathrm{sp}(A^*A) = \sum_{j=1}^{\infty} \|A\chi_j\|^2 = \sum_{j,k=1}^{\infty} |(A\chi_j, \chi_k)|^2\,,$$

for *any* ortonormal basis χ_j. The number

$$\|A\|_2 := (\mathrm{sp}(A^*A))^{\frac{1}{2}}$$

is identified as the *Hilbert-Schmidt norm* of operator A.

Note that

$$\mathrm{sp}((QA)^*QA) = \mathrm{sp}(A^*Q^*QA) = \mathrm{sp}(A^*A)\,,$$

for any unitary operator Q. The composition QA of a Hilbert-Schmidt operator A with a unitary operator Q is a Hilbert-Schmidt operator with equal Hilbert-Schmidt norm.

Chapter 3

Elements of Theory of Entire Functions

In this chapter, we study the Weierstrass infinite product:

$$\Pi(z) := \prod_{n=1}^{\infty}(1 - \frac{z}{a_n})$$

and its relation with the sequence $0 \neq a_n \to \infty$ of its zeros[8]. We will study relations between *growth order* ρ of the product:

$$M_\Pi(r) := \sup_{|z|=r} |f(z)| \qquad \rho := \limsup_{r\to\infty} \frac{\ln \ln M_\Pi(r)}{\ln r} \,,$$

convergence exponent λ of the sequence,

$$\lambda := \inf\{\mu \,:\, \sum_{n=1}^{\infty} \frac{1}{|a_n|^\mu} < \infty\} \,,$$

and the *order* ρ_1 of its *counting function* $n(r)$,

$$n(r) := \#\{n \,:\, |a_n| \leq r\} \qquad \rho_1 := \limsup_{r\to\infty} \frac{\ln n(r)}{\ln r} \,.$$

It turns out that $\lambda = \rho_1$, and ρ equals the two constants in the range $(0, 1]$.

The results in this chapter are reproduced from the book of Levin [3].

3.1 ▪ Jensen's Formula and the Counting Function

We begin by recalling the Poisson formula for harmonic functions (comp. Exercise 3.1.1),

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{|\zeta|^2 - |z|^2}{|\zeta - z|^2} u(\zeta)\, d\psi \tag{3.1.1}$$

where $\zeta = Re^{i\psi}$, u is a function harmonic in ball $B(0, R)$ and continuous in $\bar{B}(0, R)$, and point $z \in B(0, R)$. Direct computation shows that

$$\frac{|\zeta|^2 - |z|^2}{|\zeta - z|^2} = \Re\left(\frac{\zeta + z}{\zeta - z}\right) \,.$$

[8]We will assume once and for ever that $|a_n|$ is (weakly) increasing.

Lemma 3.1 (Schwarz's formula).
Let f be an analytic function in a domain D containing $\bar{B}(0,R)$, and $|z|<R$. Then

$$f(z)=\frac{1}{2\pi}\int_0^{2\pi}\frac{\zeta+z}{\zeta-z}\,u(\zeta)\,d\psi+iv(0) \tag{3.1.2}$$

where $f=u+iv$ and $\zeta=Re^{i\psi}$.

Proof. Both sides of the equality are analytic in z. The real parts are equal by the Poisson formula. If a real part of an analytic function is known then one can integrate the Cauchy-Riemann equations for the imaginary part which is unique up to a constant. It is sufficient thus to notice that the imaginary parts of both sides of the formula coincide at $z=0$.

Poisson-Jensen formula. Let f be analytic in a domain D containing $\bar{B}(0,R)$.
Case : $f(z)\neq 0$ in $\bar{B}(0,R)$.
Applying the Schwarz formula to analytic function $\ln f(z)$ (with a predefined branch of $\ln$), we obtain

$$\ln f(z)=\frac{1}{2\pi}\int_0^{2\pi}\frac{\zeta+z}{\zeta-z}\underbrace{\ln|f(\zeta)|}_{=\Re\ln f(\zeta)}\,d\psi+iC \qquad \zeta=Re^{i\psi}\,,$$

and taking the real part of both sides,

$$\ln|f(z)|=\frac{1}{2\pi}\int_0^{2\pi}\frac{|\zeta|^2-|z|^2}{|\zeta-z|^2}\,\ln|f(\zeta)|\,d\psi \qquad \zeta=Re^{i\psi}\,. \tag{3.1.3}$$

Case : f vanishes at $a_1,\ldots,a_n\in B(0,R)$, $f(z)\neq 0$ for $|z|=R$. Assume

$$|a_1|\leq|a_2|\leq\ldots\leq|a_n|\,.$$

Introduce an auxiliary function

$$\varphi(z)=f(z)\prod_{m=1}^{n}\frac{R^2-\overline{a_m}z}{R(z-a_m)} \qquad \varphi(z)\neq 0 \text{ in } \bar{B}(0,R)\,.$$

Check that $|\varphi(\zeta)|=|f(\zeta)|$ and apply formula (3.1.3) to function φ to obtain:

$$\ln\varphi(z)=\frac{1}{2\pi}\int_0^{2\pi}\frac{|\zeta|^2-|z|^2}{|\zeta-z|^2}\,\ln|f(\zeta)|\,d\psi+iC$$

and the final formula expressed in $f(z)$ alone:

$$\ln f(z)=\frac{1}{2\pi}\int_0^{2\pi}\frac{|\zeta|^2-|z|^2}{|\zeta-z|^2}\,\ln|f(\zeta)|\,d\psi+\sum_{|a_m|<R}\ln\frac{R^2-\overline{a_m}z}{R(z-a_m)}+iC$$

with properly adjusted brunch cut for the ln function to avoid collision with the roots a_m. Taking real part of both sides we obtain the *Poisson-Jensen* formula[9]:

$$\ln|f(z)|=\frac{1}{2\pi}\int_0^{2\pi}\frac{|\zeta|^2-|z|^2}{|\zeta-z|^2}\,\ln|f(\zeta)|\,d\psi+\sum_{|a_m|<R}\ln\left|\frac{R^2-\overline{a_m}z}{R(z-a_m)}\right|\,. \tag{3.1.4}$$

[9]The formula was derived by the Finnish mathematician Rolf Nevanlinna [1895 – 1980] who named it after Poisson and Jensen.

Jensen's formula. *Case:* $f(0) \neq 0$.
Setting $z = 0$ in the Poisson-Jensen formula, we obtain:

$$\ln|f(0)| = \frac{1}{2\pi}\int_0^{2\pi} \ln|f(Re^{i\psi})|\,d\psi + \sum_{|a_m|<R} \ln\frac{|a_m|}{R}\,. \tag{3.1.5}$$

Let $a_n \in \mathbb{C}$ be now a growing in modulus sequence of complex numbers converging to ∞,

$$|a_{n+1}| \geq |a_n|, \quad |a_n| \to \infty\,.$$

We define the *counting function* for sequence a_n by:

$$n(r) := \#\{n\,:\,|a_n| \leq r\}, \quad r \in [0,\infty)\,.$$

It is easy to see that $n(r)$ is integer-valued, piece-wise constant, increasing and right-continuous. The Riemann-Stieljes integral allows us to relate the discrete sum on the right of (3.1.5) to an integral of the counting function for the sequence of roots a_n,

$$\sum_{|a_m|<R} \ln\frac{R}{|a_m|} = \int_0^R \ln\frac{R}{t}\,dn(t) = \underbrace{n(t)\ln\frac{R}{t}|_0^R}_{=0} + \int_0^R \frac{n(t)}{t}\,dt\,.$$

Combining the result above with (3.1.5) we obtain the *Jensen formula:*

$$\int_0^R \frac{n(t)}{t}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \ln|f(Re^{i\psi}|\,d\psi - \ln|f(0)|\,. \tag{3.1.6}$$

Case: $f(0) = 0$ with multiplicity k.
Applying (3.1.6) to $f(z)/z^k$, we obtain a modified version of the Jensen formula,

$$\int_0^R \frac{n(t)-n(0)}{t}\,dt + n(0)\ln R = \frac{1}{2\pi}\int_0^{2\pi} \ln|f(Re^{i\psi}|\,d\psi - \ln\left|\frac{f^{(k)}(0)}{k!}\right|\,. \tag{3.1.7}$$

Recall that the growth of an entire function f is measured with function

$$M_f(r) = \max_{|z|=r}|f(z)| = \max_{|z|\leq r}|f(z)|\,.$$

Corollary 3.2. *Assume additionally* $|f(0)| = 1$. *It follows from the Jensen formula that*

$$\int_0^{er} \frac{n(t)}{t}\,dt \leq \frac{1}{2\pi}\int_0^{2\pi} \ln M_f(er)\,d\psi = \ln M_f(er)\,.$$

But,

$$\int_0^{er} \frac{n(t)}{t}\,dt \geq \int_r^{er} \frac{n(t)}{t}\,dt \geq n(r)\int_0^{er} \frac{1}{t}\,dt = n(r)\ln t|_r^{er} = n(r)\,.$$

We obtain thus a bound for the counting function for roots of $f(z)$ in terms of its growth function,

$$n(r) \leq \ln M_f(er)\,. \tag{3.1.8}$$

Exercises

3.1.1. Prove Poisson formula (3.1.1). (5 points)

3.2 ▪ Convergence Exponent of Sequence of Zeros

Convergence exponent of a sequence. Let $0 \neq a_n \to \infty$ be a (weakly) increasing in modulus sequence converging to infinity. Number

$$\lambda =: \inf\{\mu \,:\, \sum_{n=1}^{\infty} \frac{1}{|a_n|^{\mu}} < \infty\} \tag{3.2.9}$$

is called the *convergence exponent of the sequence* a_n. Record a few simple observations:

- If $\sum_{n=1}^{\infty} \frac{1}{|a_n|^{\mu}} < \infty$ for some $\mu > 0$, then $\sum_{n=1}^{\infty} \frac{1}{|a_n|^{\nu}} < \infty$ for any $\nu > \mu$. Indeed,

  $$\sum_{n=1}^{\infty} \frac{1}{|a_n|^{\nu}} = \sum_{n=1}^{\infty} \frac{1}{|a_n|^{\mu}} \frac{1}{|a_n|^{\nu-\mu}} \leq C \sum_{n=1}^{\infty} \frac{1}{|a_n|^{\mu}} < \infty$$

  where C is a bound for sequence $1/|a_n|^{\nu-\mu}$ converging to zero.

- The convergence exponent may be infinite. Recall that

  $$a_n := \ln n \overset{\text{as}}{<} n^{\epsilon}$$

  for any $\epsilon > 0$. Assume that the series (3.2.9) converges for some $\mu < \infty$. Then

  $$\frac{1}{(\ln n)^{\mu}} > \frac{1}{n^{\epsilon\mu}}$$

  and, for any finite μ, we can find $\epsilon > 0$ such that $\epsilon\mu \leq 1$ for which the series on the right diverges.

- The infimum in (3.2.9) may or may not be attained. For instance, for $a_n = n$, $\lambda = 1$, but for $\mu = 1$ we have the harmonic series which diverges. But for $a_n = n \ln^2 n$, the series does converge for $\mu = 1$ but it does not for any $\mu < 1$ (comp. Exercise 3.2.3).

Let $n(r)$ be the counting function of sequence a_n. We define the *order* ρ_1 of $n(r)$ as:

$$\rho_1 := \limsup_{r\to\infty} \frac{\ln n(r)}{\ln r} = \lim_{N\to\infty} \sup_{r \geq N} \frac{\ln n(r)}{\ln r} .$$

It follows that, for any $\epsilon > 0$, there exists N such that

$$\sup_{r \geq N} \frac{\ln n(r)}{\ln r} \leq \rho_1 + \epsilon$$

and, so,

$$\ln n(r) \leq (\rho_1 + \epsilon) \ln r \qquad r \geq N \qquad \Rightarrow \qquad n(r) \leq r^{\rho_1+\epsilon} \quad r \geq N ,$$

i.e.,

$$n(r) \overset{\text{as}}{\leq} r^{\rho_1+\epsilon} \quad \text{for any } \epsilon > 0 ,$$

At the same time, there exists a sequence $r_n \to \infty$ such that

$$\lim_{n\to\infty} \frac{\ln n(r_n)}{\ln r_n} = \rho_1$$

which implies that, for any $\epsilon > 0$, there exists N such that for $n > N$,

$$\rho_1 - \epsilon < \frac{\ln n(r_n)}{\ln r_n} \qquad \Rightarrow \qquad r_n^{\rho_1 - \epsilon} < n(r_n)\,.$$

We denote this fact as:

$$r^{\rho_1-\epsilon} \overset{\text{n}}{\leq} n(r)\,.$$

Lemma 3.3.
Assume

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^\lambda} < \infty\,.$$

Then

$$\int_0^\infty \frac{n(t)}{t^{\lambda+1}}\,dt < \infty \quad \textit{and} \quad \lim_{t\to\infty} \frac{n(t)}{t^\lambda} = 0\,.$$

Proof. We have:

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^\lambda} = \int_0^\infty \frac{dn(t)}{t^\lambda}\,.$$

At the same time,

$$\int_0^r \frac{dn(t)}{t^\lambda} = \underbrace{\frac{n(t)}{t^\lambda}\Big|_0^r}_{=\frac{n(r)}{r^\lambda}} + \lambda \int_0^r \frac{n(t)}{t^{\lambda+1}}\,dt\,.$$

As the left-nad side converges to a number as $r \to \infty$, both non-negative terms on the right must remain bounded. Since,

$$r \to \int_0^r \frac{n(t)}{t^{\lambda+1}}\,dt$$

is increasing (and bounded), it must converge to a number, i.e.,

$$\int_0^\infty \frac{n(t)}{t^{\lambda+1}}\,dt < \infty\,.$$

Consequently,

$$\frac{n(r)}{r^\lambda} = n(r)\lambda \int_r^\infty \frac{dt}{t^{\lambda+1}} \leq \lambda \int_r^\infty \frac{n(t)}{t^{\lambda+1}}\,dt \to 0 \quad \text{as } r \to \infty\,.$$

QED

Lemma 3.4.
Convergence exponent λ of a sequence a_n is equal to the order ρ_1 of its counting function $n(r)$.

Proof. Let λ be the convergence exponent of a_n. Let $\mu > \lambda$. Then

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^\mu} < \infty\,,$$

and, by Lemma 3.3,

$$\frac{n(r)}{r^\mu} \to 0 \quad \text{as } r \to \infty\,,$$

which implies that the order of counting function $\rho_1 \le \mu$. Passing with $\mu \to \lambda$, we obtain $\rho_1 \le \lambda$. On the other hand, for $\epsilon > 0$,

$$n(t) \overset{\text{as}}{\le} t^{\rho_1 + \frac{\epsilon}{2}}\,.$$

Therefore, for $\mu = \rho_1 + \epsilon$,

$$\int_0^\infty \frac{n(t)}{t^{\mu+1}}\,dt < \infty \quad \text{and} \quad \frac{n(t)}{t^\mu} \to 0 \quad \text{as } t \to \infty\,.$$

It follows from the proof of Lemma 3.3 that

$$\sum_{n=1}^\infty \frac{1}{|a_n|^\mu} < \infty$$

and, therefore, $\lambda < \mu = \rho_1 + \epsilon$. Passing with $\epsilon \to 0$, we obtain $\lambda \le \rho_1$. QED

Theorem 3.5 (Hadamard).
Convergence exponent λ of zeros of an entire function, equal to the order ρ_1 of its counting function, is bounded by the growth order ρ of the function.

Proof. Recall the estimate (3.1.8),

$$n(r) \le \ln M_f(er)\,.$$

We have:

$$\begin{aligned} \rho_1 := \limsup_{r\to\infty} \frac{\ln n(r)}{\ln r} \quad &\le \limsup_{r\to\infty} \frac{\ln\ln M_f(er)}{\ln r} \\ &= \limsup_{r\to\infty} \frac{\ln\ln M_f(er)}{\ln(er)} \lim_{r\to\infty} \frac{\ln(er)}{\ln r} \\ &= \limsup_{r\to\infty} \frac{\ln\ln M_f(r)}{\ln(r)}\,. \end{aligned}$$

Use Lemma 3.4 to finish the proof. QED

Exercises

3.2.1. (Borel lemma.) Let $0 \le a_{n+1} \le a_n$ be a (weakly) decreasing sequence of non-negative real numbers such that the series

$$\sum_{n=1}^\infty a_n$$

converges. Prove that

$$\lim_{n\to\infty} n a_n = 0\,.$$

(5 points)

3.2.2. Prove that

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^{\lambda}} < \infty \quad \Leftrightarrow \quad \int_0^{\infty} \frac{n(t)}{t^{\lambda+1}}\, dt < \infty\,.$$

(2 points)

3.2.3. Show that, for the sequence $a_n = n \ln^2 n$,

$$\sum_{n=1}^{\infty} \frac{1}{a_n} < \infty \qquad \text{but} \qquad \sum_{n=1}^{\infty} \frac{1}{a_n^{\mu}} = \infty \quad \forall, \mu < 1\,.$$

Consequently, the sequence convergence exponent $\lambda = 1$ and the infimum in definition (3.2.9) is attained. (2 points)

3.3 ▪ Weierstrass Products

Let $0 \neq a_n \in \mathbb{C}$ be a sequence of non-zero complex numbers increasing in modulus, such that

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^{p+1}} < \infty\,,$$

for some natural number $p = 1, 2, \ldots$. The *Weierstrass canonical product of genus* p is defined as:

$$\Pi(z) := \prod_{n=1}^{\infty} G(\frac{z}{a_n}, p) \tag{3.3.10}$$

where $G(u, p)$ are the *Weierstrass primary factors*:

$$G(u,p) := \begin{cases} 1-u & p = 0\,, \\ (1-u)\exp[u + \frac{u^2}{2} + \ldots + \frac{u^p}{p}] & p > 0\,. \end{cases} \tag{3.3.11}$$

Expanding $\ln(1-u)$ into its Taylor series at $u = 0$, we learn that

$$\ln G(u,p) = \ln(1-u) + u + \frac{u^2}{2} + \ldots \frac{u^p}{p} = -\sum_{k=p+1}^{\infty} \frac{u^k}{k}\,.$$

This leads to the following estimate for $|u| < \frac{1}{2}$,

$$\begin{aligned} |\ln G(u,p)| &\leq \sum_{k=p+1}^{\infty} \frac{|u|^k}{k} \\ &\leq \tfrac{|u|^{p+1}}{p+1}[1 + \tfrac{p+1}{p+2}\tfrac{1}{2} + \tfrac{p+1}{p+3}\tfrac{1}{2^2} + \ldots] \\ &\leq \tfrac{2}{p+1}|u|^{p+1}\,. \end{aligned}$$

Consequently,

$$\ln|\Pi(z)| \leq \frac{2}{p+1}\sum_{n=1}^{\infty} \left|\frac{z}{a_n}\right|^{p+1} = \frac{2|z|^{p+1}}{p+1}\sum_{n=1}^{\infty} \frac{1}{|a_n|^{p+1}}\,.$$

Note that for $|z| \leq R$, and sufficiently large n, $|\frac{z}{a_n}| < \frac{1}{2}$ and, therefore, the series above converges absolutely and uniformly in any disk $\{|z| \leq R < \infty\}$. Consequently, the same holds for the Weierstrass product which represents an entire function[10].

[10] Uniformly convergent sequence of analytic functions is analytic.

Lemma 3.6 (Borel estimates).
The following estimates hold:

$$\begin{aligned} \ln|G(u,0)| &\le \ln(1+|u|) \\ \ln|G(u,p)| &\le A_p \frac{|u|^{p+1}}{1+|u|} \quad A_p := 3e(2+\ln p)\,. \end{aligned} \tag{3.3.12}$$

Proof. Case $p=0$ is obvious. Let $p>0$.
Case: $|u| < \frac{p}{p+1}$. We have (see above):

$$\ln|G(u,p)| \le \sum_{n=p+1}^{\infty} \frac{|u|^n}{n} \le \frac{|u|^{p+1}}{(p+1)(1-|u|)} \le |u|^{p+1}\,,$$

since

$$|u| < \frac{p}{p+1} \quad \Leftrightarrow \quad \frac{1}{1-|u|} < p+1\,.$$

Case: $|u| > \frac{p}{p+1}$. We have:

$$\begin{aligned} \ln|G(u,p)| &\le \underbrace{|\ln(1-u)|}_{\le|u|} + |u| + \tfrac{|u|^2}{2} + \ldots + \tfrac{|u|^p}{p} \\ &= |u|^p \left(\tfrac{1}{p} + \tfrac{1}{p-1}\tfrac{1}{|u|} + \ldots + \tfrac{1}{2}\tfrac{1}{|u|^{p-2}} + \tfrac{2}{|u|^{p-1}} \right) \\ &\le |u|^p \left(\tfrac{p+1}{p}\right)^{p-1} (2 + \tfrac{1}{2} + \ldots + \tfrac{1}{p})\,, \end{aligned}$$

since

$$\begin{aligned} &|u| > \tfrac{p}{p+1} \quad \Rightarrow \\ &\tfrac{1}{|u|} < \tfrac{p+1}{p}, \quad \tfrac{1}{|u|^{p-1}} < \left(\tfrac{p+1}{p}\right)^{p-1}, \quad \tfrac{1}{|u|^{p-2}} < \left(\tfrac{p+1}{p}\right)^{p-2} < \left(\tfrac{p+1}{p}\right)^{p-1} = (*) \end{aligned}$$

etc. Additionally,

$$\begin{aligned} &\left(\frac{p+1}{p}\right)^{p-1} = \left(1+\frac{1}{p}\right)^p \frac{p}{p+1} \le e\frac{p}{p+1} \le e\,, \\ &\frac{1}{2} + \ldots + \frac{1}{p} < \int_0^p \frac{dx}{x} = \ln p\,. \end{aligned}$$

Continuing the estimate above,

$$\begin{aligned} (*) \quad &\le e(2+\ln p)|u|^p \underbrace{\frac{1+|u|}{1+|u|}}_{=1} \\ &= e(2+\ln p)\left(1+\tfrac{1}{|u|}\right)\tfrac{|u|^{p+1}}{1+|u|} \\ &\le \underbrace{3e(2+\ln p)}_{=:A_p}\tfrac{|u|^{p+1}}{1+|u|} \qquad \text{(since } 1+\tfrac{1}{|u|} < 1+\tfrac{p+1}{p} = \tfrac{2p+1}{p} \le 3)\,. \end{aligned}$$

QED

Theorem 3.7.
Let

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^{p+1}} < \infty\,.$$

Then the Weierstrass product

$$\Pi(z) := \prod_{n=1}^{\infty} G(\frac{z}{a_n}, p)$$

converges uniformly on every compact set, and the following estimate holds:

$$\ln|\Pi(z)| \le K_p r^p \left\{ \int_0^r \frac{n(t)}{t^{p+1}}\,dt + r\int_r^{\infty} \frac{n(t)}{t^{p+2}}\,dt \right\} \tag{3.3.13}$$

with $K_p := (p+1)A_p, r = |z|$ where A_p is the constant from Borel estimates.

Proof. We have already discussed the convergence.
Case: $p = 0$.

$$\begin{aligned}
\ln|\Pi(z)| &\le \sum_{n=1}^{\infty} \ln\left(1 + \frac{r}{|a_n|}\right) \\
&= \int_0^{\infty} \ln(1+\frac{r}{t})\,dn(t) \\
&= \underbrace{\ln(1+\frac{r}{t})n(t)\big|_0^{\infty}}_{=0} + r\int\int_0^{\infty} \frac{n(t)}{t(t+r)}\,dt \quad \left(\tfrac{d}{dt}(\ln(1+\tfrac{r}{t})) = \tfrac{1}{1+\frac{r}{t}}(-\tfrac{r}{t^2}) = \tfrac{1}{t(t+r)}\right) \\
&\le \int_0^{\infty} \frac{n(t)}{t}\,dt + r\int_r^{\infty} \frac{n(t)}{t^2}\,dt\,.
\end{aligned}$$

Case: $p > 0$. Borel estimate implies

$$\begin{aligned}
\ln|\Pi(z)| &\le A_p \sum_{n=1}^{\infty} \frac{r^{p+1}}{|a|^p(r+|a_n|)} \\
&= A_p r^{p+1} \int_0^{\infty} \frac{dn(t)}{t^p(t+r)} \\
&= A_p r^{p+1} \frac{n(t)}{t^p(t+r)}\Big|_0^{\infty} + A_p r^{p+1} \int_0^{\infty} \left[\frac{p}{t^{p+1}(t+r)} + \frac{1}{t^p(t+r)^2}\right] n(t)\,dt \\
&\left(\frac{n(t)}{t^{p+1}} \to 0 \Rightarrow \quad \frac{n(t)}{t^p(t+r)} \to 0 \quad \text{as } t\to\infty \text{ since } r \text{ is fixed}\right) \\
&= a_p r^{p+1}\{\int_0^r + \int_r^{\infty}\}\,[\ldots]\; n(t)dt \\
&\le (p+1)A_p r^p \left\{ \int_0^r \frac{n(t)}{t^{p+1}}\,dt + r\int_r^{\infty} \frac{n(t)}{t^{p+2}}\,dt \right\}
\end{aligned}$$

since

$$r\int_0^r \frac{p}{t^{p+1}(t+r)} n(t)\,dt \le p\int_0^r \frac{n(t)}{t^{p+1}\frac{t+r}{t}}\,dt \le p\int_0^r \frac{n(t)}{t^{p+1}}\,dt$$

and

$$r\int_r^\infty \frac{p}{t^{p+1}(t+r)}n(t)\,dt \le pr\int_r^\infty \frac{n(t)}{t^{p+2}}\,dt \quad (t+r>t \;\Rightarrow\; \frac{1}{t+r}<\frac{1}{t})$$

along with

$$r\int_r^\infty \frac{n(t)}{t^p(t+r)^2}\,dt \le r\int_r^\infty \frac{n(t)}{t^{p+2}}\,dt$$

$$\int_0^r \frac{n(t)}{t^p(t+r)(\frac{t}{r}+1)} \le \int_0^r \frac{n(t)}{t^{p+1}}\,dt\,.$$

QED

Theorem 3.8 (Borel).
Let $p \in \mathbb{N}$ be the smallest natural number such that

$$\sum_{n=1}^{\infty} \frac{1}{|a_n|^{p+1}} < \infty \, .$$

The growth order of the Weierstrass product of genus p for sequence a_n is equal to the convergence exponent λ of a_n.

Proof. It follows from the definition of convergence exponent λ of a_n that

$$p \leq \lambda \leq p + 1 \, .$$

Case: $\lambda < p + 1$. Pick an $\epsilon > 0$ such that $\lambda + \epsilon < p + 1$. According to our opening discussion in Section 3.2 and Lemma 3.4,

$$n(t) \overset{\text{as}}{\leq} t^{\lambda+\epsilon} \, .$$

Estimate (3.3.13) implies then

$$\begin{aligned} \ln M_\Pi(r) \quad &\leq K_p r^p \{ \underbrace{O(1)}_{\text{asymptotics}} + \int_0^r t^{\lambda+\epsilon-p-1} \, dt + r \int_r^\infty t^{\lambda+\epsilon-p-2} \, dt \} \\ &\leq K_p \left\{ O(1) + \frac{r^{\lambda+\epsilon-p}}{\lambda+\epsilon-p} + \frac{r^{\lambda+\epsilon-p}}{p+1-\lambda-\epsilon} \right\} \\ &\overset{\text{as}}{\leq} r^{\lambda+\epsilon} \, . \end{aligned}$$

Case: $\lambda = p + 1$. Lemma 3.3 implies that

$$\int_r^\infty \frac{n(t)}{t^{p+1}} \, dt \to 0 \quad \text{as } r \to \infty \, .$$

In turn, it follows from estimate (3.3.13) that

$$\ln M_\Pi(r) \leq K_p r^p \left\{ O(1) + \int_0^r \frac{n(t)}{t^{p+1}} \, dt + r \int_r^\infty \frac{n(t)}{t^{p+2}} \, dt \right\} \, .$$

However,

$$\frac{1}{r} \int_0^r \frac{n(t)}{t^{p+1}} \, dt = \int_0^r \frac{n(t)}{t^{p+2}} \frac{t}{r} \, dt = \int_0^\infty \frac{n(t)}{t^{p+2}} \underbrace{\frac{t}{r} \chi_{[0,r]}}_{\leq 1, \to 0 \text{ as } r \to \infty} \, dt \leq \int_0^\infty \underbrace{\frac{n(t)}{t^{p+2}}}_{\text{dominating function}} \, dt < \infty$$

so, by the Lebesgue Dominated Convergence Theorem, the term converges to zero as $r \to \infty$. Since

$$\lim_{r \to \infty} \int_r^\infty \frac{n(t)}{t^{p+2}} \, dt = 0 \, ,$$

we arrive at the asymptotic estimate:

$$\ln M_\Pi(r) \overset{\text{as}}{\leq} \epsilon r^{p+1} = \epsilon r^\lambda$$

for any $\epsilon > 0$.
Both discussed cases imply that the growth order of function $M_\Pi(r)$ is bounded by λ. For the reverse inequality, see the Hadamard Theorem 3.5. QED

3.4 ▪ Phragmén and Lindelöf Result

Theorem 3.9 (Phragmén, Lindelöf, 1908).
Let D be an infinite sector of the complex plane,

$$D = \{z \,:\, \alpha < \arg z < \beta\} \qquad \beta - \alpha = \frac{\pi}{\lambda}$$

and let f be an analytic function in D, continuous on its boundary[11] *with the growth:*

$$\ln M_f(r) \overset{\text{as}}{\leq} r^\rho \quad \textit{for } \rho < \lambda\,.$$

Assume function f is bounded on the sides of sector D,

$$|f(z)| \leq M \quad z \in \partial D\,.$$

Then f must be bounded by M in the whole sector,

$$|f(z)| \leq M \qquad \forall\, z \in D\,.$$

Proof. Without loosing generality assume

$$D = \{re^{i\theta} \,:\, |\theta| < \alpha\} \quad \alpha = \frac{\pi}{2\lambda}\,.$$

Pick $\rho_1 \in (\rho, \lambda)$ and define

$$\varphi_\rho(z) := f(z)e^{-\delta z^{\rho_1}}, \quad \delta > 0\,.$$

Then (see Exercise 3.4.1 for details),

$$|\varphi_\delta(z)| \leq |f(z)|\, e^{-\delta|z|^{\rho_1}\cos\rho_1\alpha} \overset{\text{as}}{\leq} e^{|z|^\rho - \delta|z|^{\rho_1}\cos\rho_1\alpha}\,.$$

As

$$\rho < \rho_1, \quad \cos\rho_1\alpha > 0 \quad (\rho_1\alpha < \lambda\alpha = \frac{\pi}{2})\,,$$

we have

$$|\varphi_\delta(Re^{i\theta})| \leq M \quad \text{for sufficiently large } R > R(\delta)\,.$$

Maximum Principle implies then

$$|\varphi_\delta(z)| \leq M \qquad \forall\, z \in D_R := \{re^{i\theta} \,:\, r < R,\, |\theta| < \alpha\}\,.$$

As $f(z) = \varphi_\delta(z)e^{\delta z \rho_1}$, this implies that

$$|f(z)| \leq \underbrace{Me^{\delta|z|^{\rho_1}\cos\theta\rho_1}}_{\text{no dependence upon } R} \leq Me^{\delta|z|^{\rho_1}} \qquad (\cos\theta\rho_1 \leq 1 \Rightarrow e^{\cos\theta\rho_1} \leq e)\,.$$

Consequently,

$$|f(z)| \leq Me^{\delta|z|^{\rho_1}} \quad \text{in } D \quad (\text{not just in } D_R)\,.$$

Pass with $\delta \to 0$ to get $|f(z)| \leq M$. QED

Exercises

3.4.1. Fill in the details of the estimate used in proof of Theorem 3.9 (1 point)

[11] The assumption may be replaced with a weaker assumption that $\limsup_{z\to\zeta}|f(z)| \leq M \quad \forall\,\zeta \in \partial D$.

Chapter 4

Macaev's Results

4.1 ▪ Additional Properties of Singular Values

Schmidt representation of a compact operator[12]**.** Let A be a compact operator, and $A = UH$ be its polar representation. Let $\phi_k, k = 1, \ldots, r(H)$ denote the orthonormal system dense in $\mathcal{R}(H)$, i.e.

$$Hu = \sum_{j=1}^{\infty} s_j (u, \phi_j)\phi_j \qquad \text{(convergence in norm)}\ .$$

Applying the unitary operator U to both sides, we obtain

$$Au = \sum_{j=1}^{\infty} s_j (u, \phi_j) \underbrace{U\phi_j}_{=:\psi_j}$$

where ψ_j form an orthonormal system in $\mathcal{R}(A)$. This is the *Schmidt representation (sum, series)* of operator A. Direct computation shows:

$$A^* u = \sum_{j=1}^{\infty} s_j (u, \psi_j)\phi_j\ .$$

The representations for operators A and A^* imply that

$$A^* A\phi_j = s_j^2 \phi_j \qquad \text{and} \quad AA^*\psi_j = s_j^2 \psi_j$$

from which, in turn, follows that $s_j(A) = s_j(A^*)$.

Theorem 4.1.
Let A be a compact operator. Then,

$$s_{n+1}(A) = \min_{\operatorname{rank} K = n} \|A - K\| = \|A - K_n\| \tag{4.1.1}$$

where the minimizer K_n equals the n-th partial Schmidt sum of operator A,

$$K_n = \sum_{j=1}^{n} s_j(A)(u, \phi_j)\psi_j\ .$$

[12]For a derivation avoiding the use of polar representation, see [4], p.587.

Proof. Recall the min-max variational property for the eigenvalues of a self-adjoint compact operator,

$$s_{n+1} = \min_{V\subset X, \dim V = n+1} \max_{u\in V} \frac{\|Av\|}{\|v\|} .$$

Let K be a compact operator of rank n. The min-max property implies that

$$\begin{aligned} s_{n+1} &\le \min_{V\subset \mathcal{N}(K)\subset X, \dim V = n+1} \max_{u\in V} \frac{\|Av\|}{\|v\|} \\ &= \min_{V\subset \mathcal{N}(K)\subset X, \dim V = n+1} \max_{u\in V} \frac{\|(A-K)v\|}{\|v\|} \\ &\le \|A-K\| . \end{aligned}$$

To claim the equality, it is sufficient to notice that

$$\|Au - \sum_{k=1}^{n} s_k(A)(u,\phi_k)\psi_k\| = \|\sum_{k=n+1}^{r(A)} s_k(A)(u,\phi_k)\psi_k\| = s_{n+1}(A) .$$

QED

Corollary 4.2. *Let A be a compact operator, and let T be an operator of rank r. Then*

$$s_{n+r}(A) \le s_n(A+T) \le s_{n-r}(A) . \tag{4.1.2}$$

Proof. Let K_n be the partial Schmidt sum for operator A. By Theorem 4.1,

$$s_{n+1}(A) = \|(A+T) - (T+K_n)\| \ge s_{n+r+1}(A+T), \qquad n = 0,1,\dots .$$

Trading A for $A+T$, we obtain,

$$s_{n+1}(A+T) \ge s_{n+r+1}(A), \qquad n = 0,1,\dots .$$

The two inequalities imply (4.1.2). QED

Corollary 4.3. *Let A, B be two compact operators. The following inequalities hold:*

$$\begin{aligned} s_{m+n-1}(A+B) &\le s_m(A) + s_n(B) \qquad && m,n = 1,2,\dots \\ s_{m+n-1}(AB) &\le s_m(A)s_n(B) \qquad && m,n = 1,2,\dots \end{aligned}$$

In particular,

$$s_n(A^q) \le s^q_{[\frac{n}{q}+1]}(A) \quad n = 1,2,\dots .$$

Proof. Let K_1, K_2 be the $(m-1)$- and $(n-1)$-dimensional operators such that

$$s_m(A) = \|A - K_1\| \quad \text{and} \quad s_n(B) = \|B - K_2\| .$$

Then

$$s_{m+n-1} \le \|A + B - \underbrace{(K_1 + K_2)}_{\text{of rank} \le m+n-2} \| \le \|A - K_1\| + \|B - K_2\| = s_m(A) + s_n(B) .$$

Similarly, since

$$(A - K_1)(B - K_2) = AB - AK_2 - K_1(B - K_2)$$

and the rank of $AK_2 + K_1(B - K_2)$ is bounded by $m + n - 2$, we obtain

$$s_{m+n-1}(AB) \leq \|AB - AK_2 - K_1(B - K_2)\| \leq \|A - K_1\| \, \|B - K_2\| = s_m(A)s_n(B)\,.$$

By induction,

$$s_{qn-(q-1)}(A^q) \leq s_n^q(A)$$

from which the last inequality follows. QED

4.2 ▪ Determinant of an Operator

Class $\mathcal{C}_\mu$ of compact operators. We say that a compact operator A belongs to class $\mathcal{C}_\mu$, for some $\mu > 0$, if

$$\sum_{n=1}^{\infty} s_n^\mu(A) < \infty$$

where $s_n(A)$ are the singular values of the operator.

Let $A \in \mathcal{C}_1$. We define,

$$\det(I - A) := \prod_{j=1}^{\nu(A)} (1 - \lambda_j(A)) \tag{4.2.3}$$

where the right-hand side converges, see below.

Characteristic determinant of operator A is defined as:

$$D_A(z) = \det(I - zA)\,. \tag{4.2.4}$$

We have,

$$\begin{aligned} |D_A(z)| &\leq \prod_{j=1}^{\nu(A)} (1 + |z| \, |\lambda_j(A)|) \\ &\leq \prod_{j=1}^{\infty} (1 + |z| s_j(A)|) \qquad (\text{ Corollary 2.7}) \\ &\leq \exp(|z| \sum_{j=1}^{\infty} s_j(A))\,. \end{aligned}$$

The characteristic determinant $D_A(z)$ is a Weierstrass canonical product of genus zero. Recall that compact operator A is a *Volterra operator* if it does not have non-zero eigenvalues. Then $\det(I - A) = 1$ and $D_A(z) = 1$.

4.3 ▪ A Resolvent Estimate

Theorem 4.4.
Let $A \in \mathcal{C}_1$. The following estimate holds:

$$\|(I - zA)^{-1}\| \leq \frac{1}{|D_A(z)|} \prod_{j=1}^{\infty} (1 + |z| \, s_j(A))\,. \tag{4.3.5}$$

Proof. Let ϕ, ψ be unit vectors and $\xi > 0$. Consider operator

$$A_1 := A + \xi(\cdot, \psi)\phi .$$

By Corollary 4.1.2,

$$s_{j+1}(A_1) \le s_j(A) .$$

Also,

$$s_1(A_1) = \min_{\|u\|=1} \|A_1 u\| \le \min_{\|u\|=1} (\|Au\| + \|\xi(u,\psi)\phi\|) \le \min_{\|u\|=1} \|Au\| + \xi = s_1(A) + \xi .$$

Consequently,

$$|D_{A_1}(z)| \le (1 + |z|(s_1(A) + \xi)) \prod_{j=1}^{\infty} (1 + |z|\, s_j(A)) . \tag{4.3.6}$$

By a generalization of the Cauchy theorem for determinants,

$$\det((I - zA_1)(I - zA)^{-1}) = \frac{D_{A_1}(z)}{D_A(z)} .$$

At the same time,

$$\begin{aligned} \det((I - zA_1)(I - zA)^{-1}) &= \det((I - zA - z\xi(\cdot,\psi)\phi)(I - zA)^{-1}) \\ &= \det(I - z\xi(\cdot,\psi)\phi(I - zA)^{-1}) \\ &= \det(I - \underbrace{z\xi((I - zA)^{-1}\cdot,\psi)\phi}_{\text{rank 1 operator}} \\ &= 1 - z\xi((I - zA)^{-1}\phi, \psi) . \end{aligned}$$

The last equality follows form the fact that the rank one operator $((I - zA)^{-1}\cdot, \psi)\phi$ has a single eigenvector ϕ with the corresponding eigenvalue equal $((I - zA)^{-1}\phi, \psi)$. Consequently, utilizing estimate (4.3.6), we obtain,

$$\begin{aligned} |z(I - zA)^{-1}\phi, \psi)| &\le \frac{1}{\xi} + \frac{1}{\xi}\frac{|D_{A_1}(z)|}{|D_A(z)|} \\ &\le \frac{1}{\xi} + \frac{1}{|D_A(z)|}\left(\frac{1}{\xi} + |z|\left(\frac{s_1(A)}{\xi} + 1\right)\right)\prod_{j=1}^{\infty}(1 + |z| s_j(A)) . \end{aligned}$$

Letting $\xi \to \infty$, and dividing both sides by $|z|$, we obtain

$$|(I - zA)^{-1}\phi, \psi)| \le \frac{1}{|D_A(z)|}\prod_{j=1}^{\infty}(1 + |z| s_j(A)) .$$

Taking supremum with respect ψ and ϕ, we obtain the final estimate. QED

Corollary 4.5. *For a Volterra operator A, $D_A(z) = 1$, and the estimate reduces to*

$$\|(I - zA)^{-1}\| \le \prod_{j=1}^{\infty}(1 + |z|\, s_j(A)) . \tag{4.3.7}$$

4.4 ▪ Macaev's Result

The following crucial theorem is stated in [2], p.244.

Theorem 4.6 (Macaev).
Let A be a compact Volterra operator and $p > 0$ an arbitrary number. If

$$s_n(A) = O(n^{-\frac{1}{p}}) \qquad \left(or\ o(n^{-\frac{1}{p}}) \right)$$

then,

$$\ln M_A(r) = O(r^{p+[p]})^{13} \qquad \left(or\ o(r^{p+[p]}) \right) .$$

I have managed to reproduce the proof under the stronger assumption $A \in \mathcal{C}_p$, and a technical assumption that the convergence exponent λ of sequence $s_n^{-1}(A)$ is strictly less than p. Note that, by Borel lemma (comp. Exercise 3.2.1),

$$\begin{aligned} \sum_{n=1}^{\infty} s_n^p(A) < \infty \quad &\Rightarrow \quad n s_n^p(A) \to 0 \\ &\Rightarrow \quad n^{\frac{1}{p}} s_n(A) \to 0 \quad \Rightarrow \quad s_n(A) = o(n^{-\frac{1}{p}}) . \end{aligned}$$

Proof. *Case:* $p < 1$. By the Borel Theorem 3.8, the growth order ρ of Weierstrass canonical product of genus 0,

$$\Pi(z) = \prod_{n=1}^{\infty} (1 - \frac{z}{a_n}) = \prod_{n=1}^{\infty} (1 - z s_n(A)) \qquad a_n = s_n^{-1}(A) ,$$

equals the convergence exponent λ of sequence a_n. Note that

$$M_\Pi(r) = \prod_{n=1}^{\infty} (1 + r s_n(A)) .$$

It follows from the definition of growth order ρ that

$$\ln M_\Pi(r) \overset{\text{as}}{\leq} r^{\rho+\epsilon} .$$

Under the additional assumption on $\rho < p$, for $\rho + \epsilon < p$, we obtain:

$$\ln M_\Pi(r) \overset{\text{as}}{\leq} r^{\rho+\epsilon} = r^p \underbrace{r^{\rho+\epsilon-p}}_{\to 0}$$

and, so,

$$\ln M_\Pi(r) = o(r^p) .$$

It remains to apply estimate (4.3.7).
Case: $p \geq 1$. Take integer $q = [p] + 1$, so that $p_1 = \frac{p}{q} < 1$, and consider Volterra operator $B = A^q$. We have:

$$\begin{aligned} 1 - z^q B \quad &= (I - zA)(I + zA + \ldots + z^{q-1} A^{q-1}) \\ (I - zA)^{-1} \quad &= (I + zA + \ldots + z^{q-1} A^{q-1})(I - z^q B)^{-1} \end{aligned}$$

[13] Gohberg has $O(r^{\frac{1}{p}})$, a typo ??

which implies:

$$M_A(r) \le M_B(r^q) \sum_{k=0}^{q-1} r^k \|A^k\| \,. \tag{4.4.8}$$

But (see Corollary 4.3),

$$s_n(B) = s_n(A^q) \le s^q_{[\frac{n}{q}+1]}(A) \quad n = 1, 2, \dots$$

and, therefore,

$$\sum_{n=1}^{\infty} s_n^{p_1}(B) \le \sum_{n=1}^{\infty} s^p_{[\frac{n}{q}+1]}(A) \le q \sum_{n=1}^{\infty} s_n^p(A) < \infty \,.$$

By the first case result,

$$\ln M_B(r) = o(r^{\frac{p}{q}}) \quad \Rightarrow \quad \ln M_B(r^q) = o(r^p) \,.$$

Use estimate (4.4.8) to conclude the final result. QED

Chapter 5

Keldyš' Results

5.1 ▪ Keldyš' Lemma

Lemma 5.1.
Let H be an injective normal compact operator. Assume that almost all characteristic numbers of H lie outside the open sector:

$$F := \{z \in \mathbb{C} \,:\, \theta_1 < \arg z < \theta_2\}\,,$$

see Fig. 5.1 for illustration. Let T be another compact operator. Then, for every $\epsilon > 0$,

$$\lim_{|z|\to\infty} \|T(I + zH)^{-1}\| = 0$$

uniformly in closed sector

$$F_\epsilon := \{z \in \mathbb{C} \,:\, \theta_1 + \epsilon \le \arg z \le \theta_2 - \epsilon\}\,.$$

Proof. Recall that

$$\|(H - \lambda I)^{-1}\| = \frac{1}{d(\lambda, \mathrm{sp}(H))}$$

where $d(\lambda, \mathrm{sp}\,(H))$ is the distance of λ to spectrum of operator H. Therefore,

$$\|(I - zH)^{-1}\| = \frac{|\frac{1}{z}|}{d(\frac{1}{z}, \mathrm{sp}(H))}\,.$$

Notice that, for $|z| = c, z \in F_\epsilon$, the smallest distance between $1/z$ and the exterior of F' (and, therefore, the spectrum of H as well) is attained on the boundary of F^ϵ. This produces a lower bound for $d(\frac{1}{z}, \mathrm{sp}(H))$:

$$|\frac{1}{z}| \sin\epsilon \le d(\frac{1}{z}, \mathrm{sp}(H))\,.$$

Consequently,

$$\frac{|\frac{1}{z}|}{d(\frac{1}{z}, \mathrm{sp}H)} \le \frac{1}{\epsilon} \quad \Rightarrow \quad \|(I - zH)^{-1}\| \le \frac{1}{\sin\epsilon}\,.$$

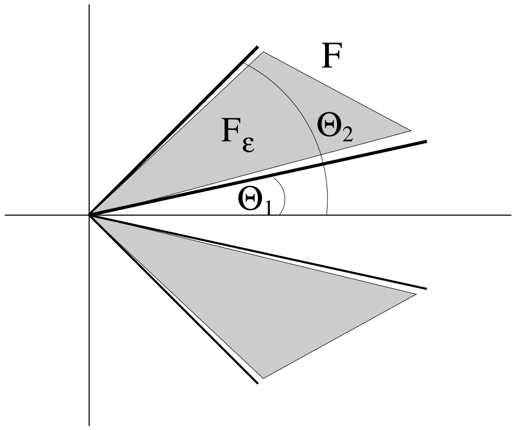

Figure 5.1: Proof of Lemma 5.1: Sector F and its image F' under $z \to \frac{1}{z}$ transformation. The shaded set illustrates subsector F_ϵ for a small ϵ.

By the density of finite rank operators in the subspace of compact operators, for any $\delta > 0$, we can decompose operator into a finite rank operator K and a remainder M such that

$$T = K + M, \qquad \|M\| < \frac{\delta}{2} \sin\epsilon\,.$$

Representing K using its Schmidt's sum,

$$Ku = \sum_{j=1}^{n} (u, \psi_j)\phi_j \quad \|\phi_j\| = 1,\ j = 1, \dots, n$$

and introducing an orthonormal basis e_i formed by eigenvectors of operator H, we have:

$$\begin{aligned} Hu &= \sum_{i=1}^{\infty} \lambda_i (u, e_i)\, e_i \\ (I - zH)u &= \sum_{i=1}^{\infty} (1 - z\lambda_i)\,(u, e_i)\, e_i \\ (I - zH)^{-1}u &= \sum_{i=1}^{\infty} (1 - z\lambda_i)^{-1}\,(u, e_i)\, e_i = \sum_{i=1}^{\infty} \frac{\mu_i}{\mu_i - z}\,(u, e_i)\, e_i \end{aligned}$$

where $\mu_i = 1/\lambda_i$ are the characteristic values of operator H. Select now a sufficiently large N such that

$$\left(\sum_{j=N+1}^{\infty} |(\psi_k, e_j)|^2 \right)^{\frac{1}{2}} < \frac{\delta \sin\epsilon}{4n} \quad \text{for } k = 1, \dots, n$$

and then a corresponding, sufficiently large R such that

$$\left(\sum_{j=1}^{N} \left| \frac{\mu_j}{\mu_j - z} \right| |(\psi_k, e_j)|^2 \right)^{\frac{1}{2}} < \frac{\delta}{4n} \quad \text{for } |z| \geq R,\ k = 1, \dots, N\,.$$

We have,

$$\begin{aligned} K(I - zH)^{-1} f &= \sum_{j=1}^{\infty} \sum_{k=1}^{n} \frac{\mu_j (\psi_k, e_j)(f, e_j)}{\mu_j - z} \phi_k \\ \|K(I - zH)^{-1} f\| &\leq \sum_{k=1}^{\infty} \left(\sum_{j=1}^{\infty} \left| \frac{\mu_j}{\mu_j - z} (\psi_k, e_j) \right|^2 \right)^{\frac{1}{2}} \underbrace{\left(\sum_{j=1}^{\infty} |(f, e_j)|^2 \right)^{\frac{1}{2}}}_{=\|f\|\|} . \end{aligned}$$

But,

$$\left| \frac{\mu_j}{\mu_j - z} \right| = \frac{|z^{-1}|}{|z^{-1} - \mu_j^{-1}|} \leq \frac{1}{\sin \epsilon}$$

so,

$$\|K(I - zH)^{-1} f\| \leq \frac{\delta}{2} \|f\| .$$

QED

5.2 ▪ Keldyš' Theorems

Let A be a compact operator. We will use the notation:

$$p(A) := \inf\{p : \sum_{j=1}^{\infty} |s_j|^p < \infty\} .$$

Theorem 5.2 (First Keldyš Theorem).
Let

$$A = H(I + S)$$

where H is a self-adjoint compact operator with $p(H) < \infty$, and S is a compact operator. We assume that A is injective. Then

(i) The system of generalized eigenvectors of A is complete in X.

(ii) For any $\epsilon > 0$, almost all eigenvalues of A lie in the sectors

$$-\epsilon < \arg z < \epsilon \qquad \pi - \epsilon < \arg z < \pi + \epsilon .$$

If the operator H has only a finite number of negative (positive) eigenvalues, then A has at most a finite number of eigenvalues in the second (first) sector.

Proof. Injectivity of A implies that $I + S$ must be injective as well. Fredholm alternative implies that $(I + S)^{-1}$ exists and it is continuous. Consequently, H is injective as well, and the eigenvectors of H form a complete system for X. Consider compact operator $T := I - (I + S)^{-1}$ (comp. Exercise 5.2.1). Lemma 5.1 implies that $\forall \epsilon > 0 \quad \exists r > 0$ such that

$$z \in F_\epsilon := \{\epsilon \leq |\arg z| \leq \pi - \epsilon, \quad |z| \geq r\} \quad \Rightarrow \quad \|T(I - zH)^{-1}\| < q < 1 .$$

By the same lemma, if H has only a *finite* number of negative eigenvalues, set F_ϵ can be enlarged to:

$$F_\epsilon := \{\epsilon \leq \arg z \leq 2\pi - \epsilon, \quad |z| \geq r\} .$$

From now on, we will consider this case only. The reasoning for the other case(s) is fully analogous. We have:

$$\begin{aligned} I - zA &= (I+S)^{-1}(I+S) - zH(I+S) \\ &= \left((I+S)^{-1} - zH\right)(I+S) \\ &= (I - T - zH)(I+S) \\ &= (I - zH - T)(I+S) \\ &= \left[I - T(I-zH)^{-1}\right](I - zH)(I+S) \end{aligned}$$

Now, operator $\left[I - T(I-zH)^{-1}\right]$ is invertible in set F_ϵ by the Neumann series argument, operator $(I - zH)$ is invertible in set F_ϵ by definition of F_ϵ, and we have already shown that operator $(I+S)$ is invertible as well. Consequently, for $z \in F_\epsilon$, inverse $(I - zA)^{-1}$ exists as well and,

$$(I - zA)^{-1} = (I+S)^{-1}(I-zH)^{-1}\left[\sum_{n=q}^{\infty}(T(I-zH)^{-1})^n\right].$$

We have proved thus already the second assertion of the theorem.

By reasoning identical to the one in proof of Lemma 5.1, we can estimate norm $\|(I-zH)^{-1}\|$ by $\frac{1}{\sin\epsilon}$ which, by the representation above, implies the estimate:

$$\|(I-zA)^{-1}\| \le \frac{\|(I+S)^{-1}\|}{\sin\epsilon}(1-q) \qquad z \in F_\epsilon.$$

Denote,

$$X_A := \overline{\text{span}\{\text{generalized eigenvectors of } A\}}.$$

We need to prove now that $X_A = X$. Suppose, by contrary, that $X_A \neq X$. Let $P : X \to X_A$ be the orthogonal projection onto the closed subspace X_A. Lemma 1.11 implies that operator:

$$A_1 := QAQ, \qquad Q := I - P,$$

is a Volterra operator. Consequently, the operator valued function:

$$z \to (I - zA_1)^{-1}$$

is an entire function. By the same Lemma 1.11,

$$\begin{aligned} (I - zA_1)^{-1} &= (I - zQAQ)^{-1} \\ &= (P = \underbrace{Q}_{=Q^2=QIQ} - zQAQ)^{-1} \\ &= (P + Q(I-zA)Q)^{-1} \\ &= Q(I-zA)^{-1}Q + P. \end{aligned}$$

Consequently, $(I - zA_1)^{-1}$ is bounded in F_ϵ. As $A_1 \in \mathcal{C}_p$ for $p > p(H)$, Theorem 4.6 implies that

$$\ln\|(I - zA_1)^{-1}\| = o(|z|^{p+[p]}).$$

Choose now $\epsilon < \frac{\pi}{p+[p]}$. As $\|(I - zA_1)^{-1}\|$ is bounded on the sides of sector F_ϵ, and we control its growth outside the sector, by the Phragmén-Lindelöf Theorem 3.9, function $\|(I - zA_1)^{-1}\|$

must be bounded on whole complex plane. But *the only bounded entire function* is a constant, and $(I - 0A_1)^{-1} = I^{-1} = I$, so

$$(I - zA_1)^{-1} = I\,.$$

This implies that $I = I - zA_1$ and, therefore, $A_1 = 0$. But (both PX and QX are invariant wrt A),

$$A_1 = QAQ = QQA = QA$$

which implies

$$A^*Q = 0 \quad \Rightarrow \quad A^*(QX) = 0 \quad \Rightarrow \quad QX \subset \mathcal{N}(A^*)$$

which contradicts injectivity of $A^* = (I + S^*)H$. QED

Jordan chains. Let A be a compact operator, μ_0 an eigenvalue of A with the corresponding eigenvector v_0. Recall that vectors $v_0, v_1, \dots v_k$ form a Jordan chain corresponding to eigenvector v_0 if

$$(A - \mu_0 I)v_j = v_{j-1} \quad j = 1, \dots, k\,.$$

As the eigenspace X_μ corresponding to eigenvalue μ may be multidimensional, there may be multiple Jordan chains corresponding to different, linearly independent, eigenvectors v_0 for the same eigenvalue μ_0. Assuming $\mu_0 \neq 0$, we may rewrite the definition of the Jordan chain in terms of the singular value $\lambda_0 = \frac{1}{\mu_0}$,

$$(I - \underbrace{\frac{1}{\mu_0}}_{=\lambda_0} A)v_j = -\frac{1}{\mu_0}v_{j-1} \quad j = 1, \dots, k\,.$$

Thus, at the cost of rescaling vectors v_j, we may redefine the Jordan chain corresponding to a singular value λ_0 and eigenvector v_0 by the relation:

$$(I - \lambda_0 A)v_j = v_{j-1} \quad j = 1, \dots, k\,.$$

It follows from the definition that

$$(I - \lambda_0 A)^{j+1}v_j = 0 \quad j = 0, 1, \dots, k\,.$$

Conversely, given a vector v such that

$$(I - \lambda_0 A)^{k+1}v = 0 \quad \text{and} \quad (I - \lambda_0 A)^k v \neq 0$$

we can reconstruct the corresponding Jordan chain by:

$$v_{k+1} := v, \qquad v_{j-1} := (I - \lambda_0 A)v_j, \quad j = k, \dots, 1\,.$$

The null space $\mathcal{N}((I - \lambda_0 A)^j)$ is identified as the *space of generalized eigenvectors of order* j. The spaces $\mathcal{N}((I - \lambda_0 A)^j)$ form an increasing sequence. It is known that, if A is compact, this sequence eventually stops growing and becomes constant. The corresponding space $\mathcal{N}((I - \lambda_0 A)^j)$ is identified as the *generalized*[14] *eigenspace* corresponding to singular value λ_0. Consequently, length k of Jordan chains is limited by the dimension of the generalized eigenspace.

[14]Gohberg calls it the *root* space.

Keldyš chains. In the next theorem we will consider a bundle

$$L(\lambda) := I - T - \lambda H$$

where T and H are compact, and H is injective.

A number λ_0 is a *characteristic value* of the bundle if there exists a (non-zero) eigenvector x_0 such that

$$(I - T - \lambda_0 H)x_0 = 0 \quad \Leftrightarrow \quad (I - T)x_0 = \lambda_0 H x_0 \,.$$

Vectors $x_1, \ldots, x_k$ form a *Keldyš chain corresponding to eigenvector* x_0 if

$$(I - T - \lambda_0 H)x_j = Hx_{j-1}, \quad j = 1, \ldots, k \,.$$

Case: $T = 0$**.** If H had a range dense in X and the corresponding inverse were bounded[15], we could apply H^{-1} to both sides of the equation above to obtain:

$$(H^{-1} - \lambda_0 I)x_j = x_{j-1}, \quad j = 1, \ldots, k \,.$$

The Keldyš chain would coincide then with the Jordan chain for the inverse H^{-1} corresponding to eigenvalue λ_0 and eigenvector x_0. But the range of operator H may not be dense in X and/or its inverse may not be bounded so this simple interpretation of Keldyš chains, in general, may not be possible.

Definition of the Keldyš chain implies that

$$(I - \lambda_0 H)^{k+1} x_k = H(I - \lambda_0 H)^k x_{k-1} = \ldots = H^k (I - \lambda_0 H)x_0 = 0 \,.$$

Injectivity of H implies thus that vectors $x_0, \ldots, x_k$ are also generalized (root) eigenvectors of operator H. Conversely, let x be a generalized eigenvector of operator H of order k, i.e.,

$$(I - \lambda_0 H)^{k+1} x = 0 \quad \text{and} \quad (I - \lambda_0 H)^k x \neq 0 \,.$$

Setting $x_k = x$, we define:

$$x_{k-1} = \sum_{i=1}^{k+1} \binom{k+1}{i} (-\lambda_0)^i H^{i-1} x_k - \lambda_0 x_k \,.$$

We verify that

$$(I - \lambda_0 H)x_k - Hx_{k-1} = (I - \lambda_0 H)^{k+1} x = 0$$

and, consequently,

$$H(I - \lambda_0 H)^k x_{k-1} = (I - \lambda_0 H)^k H x_{k-1} = (I - \lambda_0 H)^{k+1} x_k = 0 \,.$$

By injectivity of H this implies that $(I - \lambda_0 H)^k x_{k-1} = 0$, i.e., x_{k-1} is a generalized (root) eigenvector of order $k - 1$. Repeating the construction for $x = x_{k-1}$, we obtain a Keldyš chain of vectors $x_0, \ldots, x_k$. We have proved thus the following lemma.

Lemma 5.3.
Let A be a compact and injective operator. Let λ_0 be a characteristic value of the bundle:

$$L(\lambda) := (I - \lambda A) \,.$$

Then the set of all Keldyš chain vectors corresponding to λ_0 and linearly independent eigenvectors x_0 spans the space X_0 of generalized eigenvectors for operator A. In particular, the space is finite-dimensional.

[15] A bounded operator defined on a dense subset admits a unique extension to the whole space.

Remark 5.4. Notice from the reasoning above that Keldyš chains for H do not go into Jordan chains of H.

Theorem 5.5 (Second Keldyš Theorem).
Consider the bundle:

$$L(\lambda) := I - T - \lambda H$$

where T is an arbitrary compact operator, and H is compact, self-adjoint with $p(H) < \infty$. Then the system of Keldyš chain vectors for the bundle is complete in X.

Proof. We will show that the completeness results stated in the two Keldyš theorems are actually equivalent to each other.
Step 1: We first observe that, without loosing any generality, we can assume that operator $I - T$ is injective. Indeed, replacing T with $T + aH$, $a \in \mathbb{C}$, shifts all characteristic values, $\lambda \to \lambda + a$ but does not alter the corresponding generalized eigenspaces. By Lemma 5.1, we can[16] find an a such that

$$\|(I - aH)^{-1}T\| < 1\,.$$

Then

$$I - (T + aH) = (I - aH)\left[I - (I - aH)^{-1}T\right]$$

is invertible. Assuming thus that the inverse $(I - T)^{-1}$ exists and is continuous, we represent it as

$$(I - T)^{-1} = I + S$$

where S is a compact operator, comp. Exercise 5.2.1. Multiplying bundle $L(\lambda) = I - T - \lambda H$ on the left by $(I - T)^{-1}$, we get:

$$(I - T)^{-1}L(\lambda) = \underbrace{I - \lambda\, \underbrace{(I - T)^{-1}H}_{=(I+S)H=:A_1}}_{=:L_1(\lambda)}\,.$$

Note that the new bundle $L_1(\lambda) = I - \lambda A_1$ shares with bundle $L(\lambda)$ all eigenvectors and corresponding Keldyš chain vectors.

Step 2: It is now sufficient to apply Lemma 5.3 to operator A_1 and check that operator A_1 satisfies assumptions of Theorem 5.2. As Keldyš chain vectors for A_1 span the generalized eigenspace X_{λ_0}, the completeness of Keldyš chain vectors for A_1 implies the completeness result stated in Theorem 5.2. QED

Exercises

5.2.1. Let K be a compact operator. Prove that the inverse of the compact perturbation of identity $I + K$ (if it exists), is a compact perturbation of identity as well. (1 point)

[16] Note that $\|A\| = \|A^*\|$.

Chapter 6

Non-orthogonal Bases

6.1 ▪ Introduction to Non-orthogonal Bases

Schauder basis. Let X be a separable Banach space. A sequence $\phi_j \in X, j = 1, 2, \dots$ is a *Schauder basis* for space X iff

$$\forall\, x \in X \quad \exists!\, x_j,\ j = 1, 2, \dots \quad : \quad x = \sum_{j=1}^{\infty} x_j \phi_j \,.$$

Numbers x_j are the components of x wrt the basis. By assumption, they exist and they are unique. We will examine now what these assumptions imply about the basis ϕ_j.

First of all, each component x_j defines a linear functional of x,

$$\psi_j^* \,: X \to \mathbb{C}(\mathbb{R}), \quad x \to x_j \,.$$

Indeed, if

$$x = \sum_{j=1}^{\infty} \underbrace{x_j}_{=\psi_j^*(x)} \phi_j \quad \text{and} \quad y = \sum_{j=1}^{\infty} \underbrace{y_j}_{=\psi_j^*(y)} \phi_j \,,$$

then, for any α, β,

$$\alpha x + \beta y = \sum_{j=1}^{\infty} (\underbrace{\alpha x_j + \beta y_j}_{=\psi_j^*(\alpha x + \beta y)}) \phi_j \,.$$

Define now the partial sum projections:

$$P_n x := \sum_{j=1}^{n} x_j \phi_j = \sum_{j=1}^{n} \psi_j^*(x) \phi_j \,.$$

The assumed convergence of the series $\sum_{j=1}^{\infty} x_j \phi_j$ implies that $P_n x$ are bounded uniformly in n. The *Uniform Boundedness Theorem* ([4], Theorem 5.8.1) implies that the projections are bounded uniformly in the operator norm,

$$\|P_j\| \le C \quad j = 1, 2, \dots \,.$$

This implies that ψ_j^* are not only linear but also bounded. Indeed,

$$x_j \phi_j = P_j x - P_{j-1} x \quad \Rightarrow \quad |x_j|\, \|\phi_j\| \le (\|P_j\| + \|P_{j-1}\|)\|x\| \quad \Rightarrow \quad |\psi_j^*(x)| \le 2C \|\phi_j\|^{-1}\, \|x\| \,.$$

We also have a lower bound:

$$\|\psi_j^*\| = \sup_x \frac{|\psi_j^*(x)|}{\|x\|} \geq \frac{1}{\|\phi_j\|} \quad (x = \phi_j)\,.$$

We will change thus the notation from ψ_j^* to ψ_j'. Moving on from the Banach space to a Hilbert space X, we can introduce now the Riesz representations ψ_j of ψ_j',

$$\psi_j'(x) = (x, \psi_j)\,.$$

Note that by the dual space we mean the space of linear and not anti-linear functionals. We have:

$$\|\phi_j\|^{-1} \leq \|\psi_j\| \leq 2C\|\phi_j\|^{-1}\,. \tag{6.1.1}$$

We can rewrite the condition defining the basis in the form:

$$x = \sum_{j=1}^{\infty}(x, \psi_j)\phi_j\,. \tag{6.1.2}$$

Vectors ψ_j form thus a *biorthogonal sequence* to sequence ϕ_i. The existence of the unique biorthogonal system ψ_j implies two properties of the Schauder basis (Exercise 6.1.1):

(i) "Linear independence" of vectors ϕ_i:

$$\phi_i \notin \overline{\operatorname{span}\{\phi_k \,:\, k \neq i\}}\,,$$

(ii) Sequence ϕ_i is *complete* in X, i.e.,

$$(x, \phi_i) = 0\,, i = 1, 2, \ldots \quad \Rightarrow \quad x = 0\,.$$

Theorem 6.1 (Banach).
Vectors ψ_j form a Schauder basis as well.

Proof. Expansion:

$$f = \sum_{j=1}^{\infty}(f, \psi_j)\phi_j = \lim_{n\to\infty}\sum_{j=1}^{n}(f, \psi_j)\phi_j \qquad f \in X\,,$$

implies that, for any $\chi \in X$,

$$\begin{aligned}(f, \chi) = &\lim_{n\to\infty}(\sum_{j=1}^{n}(f, \psi_j)\phi_j, \chi)\\ = &\lim_{n\to\infty}\sum_{j=1}^{n}(f, \psi_j)(\phi_j, \chi) = \lim_{n\to\infty}(f, \underbrace{\sum_{j=1}^{n}(\chi, \phi_j)\psi_j}_{=:Q_n\chi})\,.\end{aligned}$$

In other words, $Q_n\chi$ converges weakly to χ, $Q_n\chi \rightharpoonup \chi$. But every weakly convergent sequence is bounded ([4], Prop. 5.14.2), i.e. $\|Q_n\chi\| \leq C(\chi)\, n = 1, 2, \ldots$. By the Uniform Boundedness Theorem, operators Q_n must be uniformly bounded,

$$\|Q_n\| \leq C \quad n = 1, 2, \ldots\,.$$

On the other side, representation (6.1.2) implies that vectors ψ_j are complete in X,

$$(f, \psi_j) = 0 \quad j = 1, 2, \ldots \quad \Rightarrow \quad f = 0\,.$$

The completeness condition is equivalent to the density of span of ψ_j in X, i.e.,

$$\forall\, \chi \quad \forall\, \epsilon > 0 \quad \exists\, N_\epsilon\,, c_j^{(\epsilon)},\, j = 1, \ldots, N_\epsilon \;:\; \|\chi - \sum_{j=1}^{N_\epsilon} c_j^{(\epsilon)} \psi_j\| < \epsilon\,.$$

Consequently,

$$\|Q_n(\chi - \sum_{j=1}^{N_\epsilon} c_j^{(\epsilon)} \psi_j)\| = \|Q_n \chi - \sum_{j=1}^{N_\epsilon} c_j^{(\epsilon)} \psi_j\| < C\epsilon \qquad \forall\, n \ge N_\epsilon\,.$$

By triangle inequality,

$$\|Q_n\chi - \chi\| \le \|Q_n\chi - \sum_{j=1}^{N_\epsilon} c_j^{(\epsilon)} \psi_j\| + \|\sum_{j=1}^{N_\epsilon} c_j^{(\epsilon)} \psi_j - \chi\| < (C+1)\epsilon \qquad n \ge N_\epsilon\,,$$

which proves the strong convergence $Q_n\chi \to \chi$. Finally, the biorthogonality condition and completeness of ψ_j imply that the components in the expansions:

$$\chi = \sum_{j=1}^{\infty} c_j \psi_j\,, \qquad c_j = (\chi, \phi_j)\,,$$

are unique (comp. Exercise 6.1.1). QED

Almost normalized sequence. We say that a sequence $\phi_j \in X$ is *almost normalized*, if

$$\inf_j \|\phi_j\| > 0 \quad \text{and} \quad \sup_j \|\phi_j\| < \infty\,.$$

Estimates (6.1.1) imply immediately that, if a basis ϕ_j is almost normalized then the corresponding biorthogonal basis is almost normalized as well.

Exercises

6.1.1. Biorthogonal sequences. Let X be a Hilbert space. Sequences ϕ_i, ψ_j are *biorthogonal* if $(\phi_i, \psi_j) = \delta_{ij}$. Let $\phi_j \in X$ be given. Show that:

(i) A biorthogonal sequence ψ_j exists iff

$$\phi_j \notin \underbrace{\overline{\text{span}\{\phi_k \;:\; k \ne j\}}}_{=:X_j}\,.$$

(ii) If it exists, biorthogonal sequence ψ_j is unique iff ϕ_j is complete in X, i.e.,

$$(x, \phi_j) = 0, \quad j = 1, 2, \ldots \qquad \Rightarrow \quad x = 0\,.$$

(5 points)

6.1.2. Generalize Theorem 6.1 to Banach spaces. (5 points)

6.2 ▪ Riesz Bases

A Riesz basis. Let $A \in \mathcal{L}(X)$ be a linear bounded operator with a bounded inverse, and let χ_j be an orthonormal basis in X. For every $x \in X$, we have the unique expansion:

$$A^{-1}x = \sum_{j=1}^{\infty}(A^{-1}x, \chi_j)\chi_j = \sum_{j=1}^{\infty}(x, (A^*)^{-1}\chi_j)\chi_j .$$

Applying operator A to both sides, we obtain:

$$x = \sum_{j=1}^{\infty}(x, \underbrace{(A^*)^{-1}\chi_j}_{=:\psi_j})\,\underbrace{A\chi_j}_{=:\phi_j}$$

where $\phi_j := A\chi_j$, and $\psi_j := (A^*)^{-1}\chi_j$ are now biorthogonal bases. Any basis ϕ_j that can be obtained from an orthonormal basis χ_j by means of such a transformation, is called a *Riesz basis*. Note that the biorthogonal basis ψ_j being the image of χ_j under operator $(A^*)^{-1}$ is automatically a Riesz basis as well. As

$$\sup_j \|\phi_j\| \leq \|A\| \quad \text{and} \quad \inf_j \|\phi_j\| \geq \frac{1}{\|A^{-1}\|} ,$$

every Riesz basis is also automatically almost normalized. If we normalize a Riesz basis to define:

$$\hat{\phi}_j := \frac{\phi_j}{\|\phi_j\|} \quad j = 1, 2, \dots ,$$

we obtain a new Riesz basis. Indeed, a map B setting the original orthonormal basis χ_j into vectors $\chi_j/\|\phi_j\|$:

$$B : X \to X, \qquad B\chi_j = \frac{\chi_j}{\|\phi_j\|} \quad j = 1, 2, \dots ,$$

obviously represents a bounded invertible operator, and

$$\hat{\phi}_j = AB\chi_j$$

where operator AB is invertible.

Gelfand's Lemma. Let X be a vector space. Recall that a function $p : X \to \mathbb{R}$ is a semi-norm on X, if p is homogeneous and it satisfies the triangle inequality,

$$\begin{aligned} & p(\alpha x) = |\alpha|\, p(x) && \alpha \in \mathbb{C},\ x \in X && \text{(homogeneity)} \\ & p(x + y) \leq p(x) + p(y) && x, y \in X && \text{(triangle inequality)}. \end{aligned}$$

The homogeneity implies that $p(0) = 0$, and the two conditions imply now that p is non-negative, $p(x) \geq 0,\ x \in X$. Indeed, we have for any $x \in X$,

$$0 = p(0) = p(x - x) \leq p(x) + p(-x) = 2p(x) .$$

Every seminorm is also convex,

$$p(\alpha x + (1 - \alpha)y) \leq p(\alpha x) + p((1 - \alpha)y) = \alpha p(x) + (1 - \alpha)p(y) \qquad \alpha \in [0, 1],\ x, y \in X .$$

We will need the following fundamental result of Gelfand.

Lemma 6.2 (Gelfand). *Let X be a Banach space, and $p : X \to [0, \infty)$, a seminorm on X. Assume p is lower semicontinuous on X, i.e., for each $x_0 \in X$ and for each $\epsilon > 0$, there exists $\delta > 0$ such that*

$$p(x) - p(x_0) > -\epsilon \qquad \|x - x_0\| < \delta\,.$$

Then there exists $C > 0$ such that

$$p(x) \le C\|x\| \qquad x \in X\,. \tag{6.2.3}$$

Proof. Condition (6.2.3) is equivalent to the boundedness of p on the unit ball. Indeed, if $p(y) \le C$ for $\|y\| < 1$ then, taking $y = x/\|x\|$, we have,

$$\frac{1}{\|x\|}p(x) = p(\frac{x}{\|x\|}) = p(y) \le C \qquad \Rightarrow \qquad p(x) \le C\|x\|\,.$$

Secondly, we observe that if p were not bounded on the unit ball $B(0,1)$, then p would not be bounded on any ball $B(x_0, \delta)$. Indeed, assume to the contrary that

$$p(y) \le C \qquad \text{for } y \in B(x_0, \delta)\,,$$

Let $\|x\| < 1$. Then $y = x_0 + \delta x \in B(x_0, \delta)$, and

$$p(x) = p(\frac{1}{\delta}(y - x_0)) = \frac{1}{\delta}p(y - x_0) \le \frac{1}{\delta}(p(y) + p(x_0)) \le \frac{2C}{\delta}\,,$$

a contradiction.

Assume now to the contrary that p is not bounded on the unit ball. Consequently, p is not bounded on any ball. Choose a point $x_1 \in B(0,1)$ such that $p(x_1) > 1$. Lower semicontinuity of p at x_1 implies that there exists a sufficiently small ρ_1 such that (choose $\epsilon = p(x_1) - 1$)

$$p(x) - p(x_1) > 1 - p(x_1) \quad x \in B(x_1, \rho_1) \qquad \Rightarrow \qquad p(x) > 1 \quad x \in B(x_1, \rho_1)\,.$$

We can always assume additionally that $\bar{B}(x_1, \rho_1) \subset B(0,1)$ and $\rho_1 < \frac{1}{2}$. But p is not bounded on $B(x_1, \rho_1)$ either, so there exists $x_2 \in B(x_1, \rho_1)$ such that $p(x_2) > 2$. Lower semicontinuity at x_2 implies again that there exists ρ_2 such that $\bar{B}(x_2, \rho_2) \subset B(x_1, \rho_1)$, $\rho_2 < \frac{1}{2}\rho_1$, and

$$p(x) - p(x_2) > 2 - p(x_2) \quad x \in B(x_2, \rho_2) \qquad \Rightarrow \qquad p(x) > 2 \quad x \in B(x_2, \rho_2)\,,$$

and so on. We obtain a sequence of balls

$$B(1,0) \supset\supset B(x_1, \rho_1) \supset\supset B(x_2, \rho_2) \supset\supset \ldots \supset\supset B(x_n, \rho_n) \supset\supset \ldots$$

such that

$$p(x) > n \qquad x \in B(x_n, \rho_n)\,.$$

It follows from the construction that

$$\begin{aligned} &\|x_1\| < 1 \\ &\|x_2 - x_1\| < \rho_1 < \tfrac{1}{2} \\ &\|x_3 - x_2\| < \rho_2 < \tfrac{1}{2}\rho_1 < \tfrac{1}{2^2} \end{aligned}$$

and, by induction,

$$\|x_{n+1} - x_n\| < \frac{1}{2^n}\,.$$

In turn, for $k > 0$,

$$\begin{aligned} \|x_{n+k} - x_n\| &\leq \|x_{n+k} - x_{n+k-1}\| + \ldots + \|x_{n+1} - x_n\| \\ &\leq \tfrac{1}{2^{n+k-1}} + \ldots + \tfrac{1}{2^n} = \tfrac{1}{2^n}\left(\tfrac{1}{2^{k-1}} + \ldots + 1\right) \\ &\leq \tfrac{1}{2^{n-1}} \end{aligned}$$

which implies that x_n is Cauchy and, therefore, $x_n \to x$, for some $x \in X$. It follows from the construction that, for every n, $x \in \bar{B}(x_n, \rho_n)$ and, therefore, $x \in B(x_n, \rho_n)$ as well and, therefore, $p(x) > n$ for every n which is impossible. QED.

Remark 6.3. Note the boundedness of the seminorm is equivalent to its continuity [17]. Indeed, if p is continuous at 0 then there exists $\delta > 0$ such that

$$p(x) = p(x) - p(0) \leq 1 \qquad \|x\| = \|x - 0\| \leq \delta\,.$$

Consequently, for any $x \neq 0$,

$$\frac{\delta}{\|x\|} p(x) = p\left(\frac{\delta x}{\|x\|}\right) \leq 1 \qquad \Rightarrow \qquad p(x) \leq \frac{1}{\delta}\|x\|\,.$$

Conversely, assuming that p is bounded, we have,

$$p(x) - p(x_0) = p(x_0 + x - x_0) - p(x_0) \leq p(x_0) + p(x - x_0) - p(x_0) \leq C\|x - x_0\|$$

and, interchanging x with x_0,

$$p(x_0) - p(x) \leq C\|x_0 - x\| = C\|x - x_0\|\,.$$

Consequently, the Gelfand Lemma may be reformulated by stating that any lower semicontinuous seminorm is automatically continuous.

Corollary 6.4. *Let $p_n(x)$ be a sequence of continuous seminorms defined on a Banach space X. Assume that, for every x, $p_n(x)$ is uniformly bounded in n, i.e., there exists $C_x > 0$ such that*

$$p_n(x) \leq C_x \qquad \forall n\,. \tag{6.2.4}$$

Then the pointwise supremum,

$$p(x) := \sup_n p_n(x)$$

is a bounded seminorm as well.

Proof. Condition (6.2.4) implies that $p(x)$ is well-defined (it is a real number). Passing to the supremum in the homogeneity and triangle inequality conditions, we verify immediately that p is a seminorm. In view of Gelfand's result, it is sufficient to show that p is lower semi-continuous. Let $x_0 \in X$ and $\epsilon > 0$. It follows from the definition of supremum that there exists N such that

$$p(x_0) - p_N(x_0) < \frac{\epsilon}{2}\,.$$

In turn, continuity of p_N at x_0 implies that there exists $\delta > 0$ such that

$$|p_N(x) - p_N(x_0)| \leq \frac{\epsilon}{2} \qquad \|x - x_0\| < \delta\,.$$

[17] The reasoning is identical with that for linear operators.

Consequently, for $\|x - x_0\| < \delta$,

$$p(x) - p(x_0) > \sup_n p_n(x) - p_N(x_0) - \frac{\epsilon}{2} \geq p_N(x) - p_N(x_0) - \frac{\epsilon}{2} \geq -\epsilon\,.$$

QED.

Theorem 6.5 (Bari).
The following five conditions are equivalent to each other.

(i) Sequence ϕ_j is a Riesz basis.

(ii) Sequence ϕ_j represents an orthonormal basis in a new inner product norm equivalent to the original inner product in X.

(iii) Sequence ϕ_j is complete in X, and there exist positive constants α_1, α_2 such that

$$\alpha_1 \sum_{j=1}^{n} |x_j|^2 \leq \| \sum_{j=1}^{n} x_j \phi_j \|^2 \leq \alpha_2 \sum_{j=1}^{n} |x_j|^2 \tag{6.2.5}$$

for any $n > 0$, and any sequence of complex numbers $x_j,\ j = 1, \dots, n$.

(iv) Sequence ϕ_j is complete in X, and its Gram matrix:

$$(\phi_j, \phi_k) \quad j,k = 1, \dots \tag{6.2.6}$$

represents a bounded invertible operator in ℓ^2.

(v) Sequence ϕ_j is complete in X and it has a complete biorthogonal sequence ψ_j, and

$$\sum_{j=1}^{\infty} |(x, \phi_j)|^2 < \infty \quad \textit{and} \quad \sum_{j=1}^{\infty} |(x, \psi_j)|^2 < \infty \qquad \forall\, x \in X\,.$$

Proof. $(i) \Rightarrow (ii)$. Let $\phi_j = A\chi_j$ where χ_j is an orthonormal basis, and A a bounded linear operator with a bounded inverse. Define the new inner product as

$$((x,y)) := (A^{-1}x, A^{-1}y)\,.$$

Then

$$((\phi_i, \phi_j)) = (A^{-1}\phi_i, A^{-1}\phi_j) = (\chi_i, \chi_j) = \delta_{ij}\,.$$

$(ii) \Rightarrow (iii)$. Let $((\cdot,\cdot))$ be the inner product in which basis ϕ_j is orthonormal. By the equivalence of norms, there exist positive constants β_1, β_2 such that

$$\beta_1 \| \sum_{j=1}^{n} x_j \phi_j \|^2 \leq ((\sum_{i=1}^{n} x_i \phi_i, \sum_{j=1}^{n} x_j \phi_j)) = \sum_{j=1}^{n} |x_j|^2 \leq \beta_2 \| \sum_{j=1}^{n} x_j \phi_j \|^2$$

from which the required inequality follows. At the same time, by the equivalence of norms, the density of $\operatorname{span}\{\phi_j,\ j = 1, \dots\}$ in X in norm $((x,x))^{\frac{1}{2}}$ implies the density in the original norm. Consequently, ϕ_j is complete in X.

$(iii) \Rightarrow (i)$. Let χ_j be an *arbitrary* orthonormal basis for X. We define operators A and A_1 defined on the spans of χ_j's, and ϕ_j's by setting χ_j into ϕ_j and vice versa,

$$A(\sum_j a_j\chi_j) := \sum_j a_j\phi_j \qquad \text{and} \qquad A_1(\sum_j a_j\phi_j) := \sum_j a_j\chi_j \,.$$

By (6.2.5),

$$\|A(\sum_j a_j\chi_j)\|^2 = \|\sum_j a_j\phi_j\|^2 \le \alpha_2 \sum_j |a_j|^2 = \|\sum_j a_j\chi_j\|^2$$

and.

$$\|A_1(\sum_j a_j\phi_j)\|^2 = \|\sum_j a_j\chi_j\|^2 = \sum_j |a_j|^2 \le \alpha_1^{-1}\|\sum_j a_j\phi_j\|^2 \,,$$

As both sequences χ_j and ϕ_j are complete, maps A and A_1 admit unique continuous extensions to whole X. Using the continuity argument, we have $AA_1 = A_1A = I$. Consequently, ϕ_j is a Riesz basis.

$(i) \Rightarrow (iv)$. Let A be a bounded invertible operator which carries an orthonormal basis χ_j into ϕ_j. As

$$(A^*A\chi_j, \chi_k) = (A\chi_j, A\chi_k) = (\phi_j, \phi_k) \,,$$

Gram matrix (6.2.6) is the matrix representation of operator A^*A in the orthonormal basis χ_j and, therefore, it represents a bounded invertible operator in the coefficients space ℓ^2.

$(iv) \Rightarrow (iii)$. Let χ_j be an arbitrary orthonormal basis in X. Define an operator H by:

$$H(\sum_j a_j\chi_j) := \sum_j \left(\sum_{k=1}^{\infty}(\phi_k, \phi_j)a_k\right)\chi_j \qquad \sum_j |a_j|^2 < \infty \,.$$

By the isomorphism of bounded invertible operators in X with bounded invertible operators in the coefficient space ℓ^2, the operator H is a bounded, positive and invertible operator in X. We have,

$$\begin{aligned}\|\sum_j a_j\phi_j\|^2 &= (H(\sum_j a_j\chi_j), \sum_j a_j\chi_j) = (H^{\frac{1}{2}}(\sum_j a_j\chi_j), H^{\frac{1}{2}}\sum_j a_j\chi_j) \\ &\le \|H^{\frac{1}{2}}\|^2\|\sum_j a_j\chi_j\|^2 = \|H\|\sum_j |a_j|^2 \,.\end{aligned}$$

By the same token,

$$\sum_j |a_j|^2 \le \|H^{-1}\|\,\|\sum_j a_j\phi_j\|^2 \,.$$

$(i) \Rightarrow (v)$. This was proved in the beginning of this section.

$(v) \Rightarrow (i)$. By Corollary 6.4, the seminorm

$$p(x) := \left(\sum_{j=1}^{\infty}|(x,\phi_j)|^2\right)^{\frac{1}{2}} = \sup_n \underbrace{\left(\sum_{j=1}^{n}|(x,\phi_j)|^2\right)^{\frac{1}{2}}}_{=:p_n(x)}$$

must be bounded, i.e., there exists $C_1 > 0$ such that

$$\left(\sum_{j=1}^{\infty}|(x,\phi_j)|^2\right)^{\frac{1}{2}} \le C_1\|x\| \,,$$

By the same argument, there exists $C_2 > 0$ such that

$$\left(\sum_{j=1}^{\infty} |(x,\psi_j)|^2\right)^{\frac{1}{2}} \leq C_2\|x\|\,,$$

Let χ_j be an arbitrary orthonormal basis in X. Define linear operators A_1 and A_2 by setting vectors ϕ_j and ψ_j into χ_j. Let $x := \sum_j a_j\phi_j$. Then $a_j = (x,\psi_j)$ and the inequality above implies

$$\|A_1(\sum_j a_j\phi_j)\|^2 = \sum_j |a_j|^2 \leq C_2^2\|\sum_j a_j\phi_j\|^2\,.$$

Similarly,

$$\|A_2(\sum_j a_j\psi_j)\|^2 = \sum_j |a_j|^2 \leq C_1^2\|\sum_j a_j\psi_j\|^2\,.$$

By completeness of ϕ_j and ψ_j, the operators A_2 and A_3 can be extended to continuous operators defined on the whole X. We have,

$$(A_1(\sum_j a_j\phi_j), A_2(\sum_k b_k\psi_k)) = \sum_j a_j\overline{b_j} = (\sum_j a_j\phi_j, \sum_k b_k\psi_k)$$

and, by completeness,

$$(A_1x, A_2y) = (x,y) \qquad x,y,\in X\,.$$

Consequently, $A_2^*A_1 = I$ and, in particular, $A_2\chi_j = \phi_j$. By the same argument, $A_1\chi_j = \psi_j$. We thus have,

$$(A_2^*(\sum_j a_j\chi_j), A_1^*(\sum_k b_k\chi_k)) = (\sum_j a_j\phi_j, \sum_k b_k\psi_k) = \sum_j a_j\bar{b}_j = (\sum_j a_j\chi_j, \sum_k b_k\chi_k)$$

and, by completeness again,

$$(A_2^*x, A_1^*y) = (x,y) \qquad x,y,\in X\,.$$

Consequently, $A_1A_2^* = I$ which proves that A_1 is invertible and has a bounded inverse $A_1^{-1} = A_2^*$. QED.

Permutable bases. A Schauder basis ϕ_i of X is *permutable (unconditional)* if every permutation of the basis is a Schauder basis as well. Every orthonormal basis is permutable and every Riesz basis is permutable as well. It turns out that the permutability is unique for the Riesz bases.

Lemma 6.6 (Orlicz). *Let x_n, $n = 1,\ldots$, be a sequence of vectors in a Banach space X. Let $n(k)$, $k = 1,\ldots$ be an arbitrary permutation of the indices. Assume that, for each such a permutations, the corresponding partial sums are uniformly bounded,*

$$\|\sum_{k=1}^{m} x_{n(k)}\| \leq C\,.$$

Then

$$\sup_{n,|\epsilon_j|\leq 1} \|\sum_{j=1}^{n} \epsilon_j x_j\| < \infty\,.$$

Proof. Let $f \in X'$ and $n(k)$ be an arbitrary permutation of the indices. Then

$$|\sum_{k=1}^{m} f(x_{n(k)})| = |f(\sum_{k=1}^{m} x_{n(k)})| \leq \|f\| \, \|\sum_{k=1}^{m} x_{n(k)}\| \leq C\|f\| \, .$$

Consequently, the number series above is absolutely convergent (comp. Exercise 6.2.1). By Corollary 6.4, the convex functional

$$p(f) := \sum_{j=1}^{\infty} |f(x_j)|, \qquad f \in X' \, ,$$

is continuous, i.e., there exists $c > 0$ such that

$$|p(f)| \leq c\|f\| \, .$$

Consequently, for any $|\epsilon_j| \leq 1$,

$$|f(\sum_{j=1}^{n} \epsilon_j x_j)| = |\sum_{j=1}^{n} \epsilon_j f(x_j)| \leq \sum_{j=1}^{n} |\epsilon_j| \, |f(x_j)| \leq \sum_{j=1}^{\infty} |f(x_j)| \leq c\|f\| \, ,$$

as well. By Exercise 6.2.2,

$$\|\sum_{j=1}^{n} \epsilon_j x_j\| = \sup_{f \neq 0} \frac{|f(\sum_{j=1}^{n} \epsilon_j x_j)|}{\|f\|} \leq c \, .$$

QED.

Corollary 6.7 (Orlicz). *Let x_n, $n = 1, \ldots$ be a sequence of vectors in a Hilbert space X, satisfying the assumptions of Lemma 6.6. Then*

$$\sum_{j=1}^{\infty} \|x_j\|^2 < \infty \, .$$

Proof. For any two vectors $x, y \in X$, we can always choose a number ϵ, $|\epsilon| = 1$, such that

$$\|x\|^2 + \|y\|^2 \leq \|x + \epsilon y\|^2 \, .$$

Indeed, expanding,

$$\|x + \epsilon y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re(\bar{\epsilon}(x, y)) = \|x\|^2 + \|y\|^2 + 2|(x, y)| \geq \|x\|^2 + \|y\|^2$$

for $\epsilon = \frac{(x,y)}{|(x,y)|}$. For three vectors x_1, x_2, x_3, we have then:

$$\|x_1\|^2 + \|x_2\|^2 + \|x_3\|^2 \leq \|x_1 + \epsilon_2 x_2\|^2 + \|x_3\|^2 \leq \|x_1 + \epsilon_2 x_2 + \epsilon_3 x_3\|^2$$

and, by induction,

$$\sum_{j=1}^{n} \|x_j\|^2 \leq \|\sum_{j=1}^{n} \epsilon_j x_j\|^2$$

where $\epsilon_1 = 1, |\epsilon_j| = 1,\ j = 1, \ldots$. The final result follows now from Lemma 6.6. QED.

Theorem 6.8 (Lorch).
A Schauder basis $\phi_j \in X,\ j = 1, \ldots$ in a Hilbert space X is a Riesz basis iff it is permutable and almost normalized.

Proof. We need to prove only the sufficiency. For an arbitrary $x \in X$,

$$x = \sum_{j=1}^{\infty} (x, \psi_j)\phi_j$$

where ψ_j is the biorthogonal basis to ϕ_j. By assumption, the series converges for an arbitrary permutation of indices. By Corollary 6.7 then

$$\sum_{j=1}^{\infty} |(x, \psi_j)|^2 \, \|\phi_j\|^2 < \infty$$

and, since ϕ_j is almost normalized,

$$\sum_{j=1}^{\infty} |(x, \psi_j)|^2 < \infty\,.$$

As a basis biorthogonal to a permutable and almost normalized basis is also permutable and almost normalized, we also have

$$\sum_{j=1}^{\infty} |(x, \phi_j)|^2 < \infty\,.$$

The conclusion follows now from Theorem 6.5 (v). QED.

Quadratically close sequences of vectors. Sequences of vectors χ_j and ϕ_j are said to be *quadratically close* if

$$\sum_{j=1}^{\infty} \|\chi_j - \phi_j\|^2 < \infty\,.$$

ω-linear independence. A sequence of vectors ϕ_j is *ω-linearly independent* if the condition

$$\sum_{j=1}^{\infty} x_j \phi_j = 0$$

is not possible for components x_j such that

$$0 < \sum_{j=1}^{\infty} |x_j|^2 \, \|\phi_j\|^2 < \infty\,.$$

This is equivalent to say (comp. Exercise 6.2.3) that

$$\left(\sum_{j=1}^{\infty} x_j \phi_j = 0 \quad \text{and} \quad \sum_{j=1}^{\infty} |x_j|^2 \, \|\phi_j\|^2 < \infty \right) \quad \Rightarrow \quad x_j = 0,\ j = 1, 2, \ldots\,.$$

If sequence ϕ_j is almost normalized then the condition above is equivalent to

$$0 < \sum_{j=1}^{\infty} |x_j|^2 < \infty \,,$$

i.e., the sequence of components x_j is a non-zero element of ℓ^2.

Theorem 6.9 (Bari).
Let $\phi_j \in X$ be a Riesz basis. Let ψ_j be a ω-linearly independent sequence of vectors which is quadratically close to basis ϕ_j. Then ψ_j is a Riesz basis as well.

Proof. Let A be a bounded linear invertible operator mapping an orthonormal basis χ_j onto basis ϕ_j,

$$A\chi_j = \phi_j \quad j = 1, 2, \dots \,.$$

Define an operator T by setting $T\chi_j = \phi_j - \psi_j$. Equivalently,

$$T\left(\sum_{j=1}^{\infty} x_j \chi_j\right) = \sum_{j=1}^{\infty} x_j(\phi_j - \psi_j) \qquad \text{where} \quad \sum_{j=1}^{\infty} |x_j|^2 < \infty \,.$$

Operator T is bounded, and

$$\|T\|^2 \le \sum_{j=1}^{\infty} \|\phi_j - \psi_j\|^2 \,,$$

i.e., $T \in \mathcal{C}_2$ and, in particular, it is compact.

The ω-linear independence of ψ_j implies that equation $(A - T)x = 0$ has only a trivial solution. Indeed, let

$$(A - T)x = \sum_{j=1}^{\infty} x_j \psi_j = 0, \quad x = \sum_{j=1}^{\infty} x_j \chi_j, \quad \sum_{j=1}^{\infty} |x_j|^2 < \infty \,.$$

As

$$\|\psi_j\| \le \|\psi_j - \phi_j\| + \|\phi_j\| \le (\sum_{i=1}^{\infty} \|\psi_i - \phi_i\|^2)^{\frac{1}{2}} + \|A\| \,.$$

$$\sum_{j=1}^{\infty} |x_j|^2 < \infty \quad \Rightarrow \quad \sum_{j=1}^{\infty} |x_j|^2 \|\psi_j\|^2 < \infty$$

and, consequently, $x_j = 0$, $j = 1, \dots$.

By the Fredholm Alternative, operator $A - T$ has thus a bounded inverse, and it sets the orthonormal basis χ_j into basis ψ_j. QED.

Exercises

6.2.1. Unconditional and absolute convergence. Let $x_n \in \mathbb{R}$, $n = 1, 2, \dots$, be an arbitrary sequence.

(i) Let $x_n \geq 0$ and $\sum_{n=1}^{\infty} x_n < \infty$. Prove that the series

$$\sum_{k=1}^{\infty} x_{n(k)}$$

converges for any bijection: $\mathbb{N} \ni k \to n(k) \in \mathbb{N}$ (permutation of indices) to the same number.

(ii) Let x_n be absolutely convergent, i.e., $\sum_{n=1}^{\infty} |x_n| < \infty$. Prove that the permuted series

$$\sum_{k=1}^{\infty} x_{n(k)}$$

converges (to a number), for any permutation $n(k)$.

(iii) Assume now that the series

$$\sum_{k=1}^{\infty} x_{n(k)}$$

converges (to a number), for every permutation of indices $n(k)$. Prove that the series must be absolutely convergent.

(iv) Argue why the equivalence of the unconditional and absolute convergence generalizes to any finite dimensional vector space including $\mathbb{C}$.

(5 points)

6.2.2. Duality pairing. Let X be a Banach space. Consider the duality pairing:

$$X' \times X \ni (f, x) \to \langle f, x \rangle := f(x) \in \mathbb{R}(\mathbb{C})\,.$$

By definition,

$$\|f\|_{X'} := \sup_{x \in X} \frac{|\langle f, x \rangle|}{\|x\|_X}\,.$$

Prove that, in turn,

$$\|x\|_X := \sup_{f \in X'} \frac{|\langle f, x \rangle|}{\|f\|_{X'}}\,.$$

(2 points)

6.2.3. Prove the tautology;

$$((p \wedge q) \Rightarrow \sim r) \quad \Leftrightarrow \quad (r \Rightarrow (\sim p \vee \sim q)) \quad \Leftrightarrow \quad ((r \wedge p) \Rightarrow \sim q)$$

(1 point)

6.3 ▪ Bari Bases

Bari bases. By Theorem 6.9, any ω-linearly independent system of vectors quadratically close to an orthonormal basis is a Riesz basis. Such bases will be called a *Bari bases* after Russian mathematician Nina Bari.

Any permutation of a Bari basis is also a Bari basis. Moreover, if ψ_j, $j = 1, \dots$ is a Bari basis then so is the basis of the normalized vectors $\hat{\psi}_j := \psi_j / \|\psi_j\|$, $j = 1, \dots$. Indeed, let ϕ_j be an orthormal basis quadratically close to ψ_j and let $\epsilon_j = (\hat{\psi}_j, \phi_j)/|(\hat{\psi}_j, \phi_j)|$. Then, $|\epsilon_j| = 1$, and

$$\|\hat{\psi}_j - \epsilon_j \phi_j\|^2 = 2(1 - |(\hat{\psi}_j, \phi_j)|) \leq 2(1 - |(\hat{\psi}_j, \phi_j)|^2) = 2 \min_{z \in \mathbb{C}} \|\phi_j - z\psi_j\|^2 \leq 2\|\phi_j - \psi_j\|^2\,,$$

i.e., the rescaled basis $\hat{\psi}_j$ is quadratically close to rescaled (and still orthonormal) basis $\epsilon_j \phi_j$.

Lemma 6.10. *The following conditions are equivalent to each other.*

(i) ψ_j *is a Riesz basis quadratically close to an orthonormal basis* χ_j*.*

(ii) There exists an operator $T \in \mathcal{C}_2$ *such that*

(a) $T\chi_j = \psi_j - \chi_j$, $j = 1, \ldots$*, and*

(b) $I + T$ *is invertible with a bounded inverse.*

Proof. $(a) \Rightarrow (b)$. Let A be a bounded operator that takes basis χ_j into basis ψ_j. For $T := A - I$, $T\chi_j = \psi_j - \chi_j$, and

$$\sum_j \|T\chi_j\|^2 = \sum_j \|\psi_j - \chi_j\|^2 < \infty ,$$

i.e. $T \in \mathcal{C}_2$.
$(b) \Rightarrow (a)$. It follows that ψ_j is a Riesz basis quadratically close to χ_j. QED.

Corollary 6.11. *If a basis* ψ_j *is quadratically close to an orthonormal basis* χ_j *, then its biorthogonal basis* ϕ_j *is also quadratically close to basis* χ_j*. Consequently the biorthogonal bases* ψ_j, ϕ_j *are also quadratically close to each other.*

Proof. Let

$$(I + T)\chi_j = \psi_j \quad \text{where} \quad T \in \mathcal{C}_2 .$$

Then

$$(\chi_k, (I + T^*)\phi_j) = ((I + T)\chi_k, \phi_j) = (\psi_k, \phi_j) = \delta_{jk} \quad j, k = 1, 2, \ldots$$

implies that $\chi_j = (I + T^*)\phi_j$ which, in turn, implies (see Exercise 6.3.2) that

$$\phi_j = (I + T_1)\chi_j \quad \text{where} \quad T_1 = (I + T^*)^{-1} - I \in \mathcal{C}_2 .$$

Finally,

$$\|\psi_j - \phi_j\|^2 \leq 2\left(\|\psi_j - \chi_j\|^2 + \|\chi_j - \phi_j\|^2\right)$$

implies that ψ_j, ϕ_j are quadratically close as well. QED.

Let ψ_j, $j = 1, \ldots$ be a sequence of linearly independent vectors. It follows from the positive definitness of the inner product and the Sylvester criterion that

$$D(\psi_1, \ldots, \psi_n) := \det\left((\psi_j, \psi_k)_1^n\right) > 0 .$$

If the sequence is normalized then (see Exercise 6.3.3),

$$\frac{D(\psi_1, \ldots, \psi_n)}{D(\psi_1, \ldots, \psi_{n-1})} = d_j^2 \leq \|\psi_j\|^2 = 1, \qquad d_j = \operatorname{dist}(\psi_n, \operatorname{span}\{\psi_1, \ldots, \psi_{n-1}\}) ,$$

so

$$D(\psi_1, \ldots, \psi_n) \leq D(\psi_1, \ldots, \psi_{n-1}) .$$

The sequence is thus (weakly) decreasing and, therefore, convergent and,

$$\Delta := \lim_{n \to \infty} D(\psi_1, \ldots, \psi_n) \geq 0 .$$

The limit Δ can be interpreted as the square of the volume of the (infinite dimensional) parallelpiped spanned by the unit vectors ψ_j.

Theorem 6.12. *Let $\psi_j \in X,\ j = 1, \dots$ be a complete sequence of unit vectors. The following conditions are equivalent to each other,*

(i) The sequence is a basis quadratically close to an orthonormal basis, i.e., it is a Bari basis.

(ii) $\Delta > 0$.

(iii) There exists a sequence ϕ_j biorthogonal to ψ_j, and the two sequences are quadratically close.

Proof. $(i) \Rightarrow (ii)$. Let ψ_j be a normalized basis quadratically close to an orthonormal basis, and let ϕ_j be the biorthogonal basis. We have already learned that

$$\sum_{j=1}^{\infty} \|\psi_j - \phi_j\|^2 < \infty\,.$$

The unit vector $e_j = \phi_j / \|\phi_j\|$ is orthogonal to

$$X_j := \overline{\operatorname{span}\{\psi_j,\, j = 1, \dots, \quad j \neq i\}}$$

and, therefore, the distance δ_j from the unit vector ψ_j to X_j equal to:

$$\delta_j := (\psi_j, e_j) = \|\phi_j\|^{-1} \qquad 0 < \delta_j \leq 1, \quad j = 1, 2, \dots\,.$$

Since $\|\psi_j - \phi_j\|^2 = \|\phi_j\|^2 - 1 = \delta_j^{-2} - 1$, it follows that

$$\sum_{j=1}^{\infty} (1 - \delta_j^2) \leq \sum_{j=1}^{\infty} \frac{1 - \delta_j^2}{\delta_j^2} = \sum_{j=1}^{\infty} (\delta_j^{-2} - 1) < \infty\,.$$

Obviously,

$$\delta_j \leq d_j = \operatorname{dist}(\psi_j, \underbrace{\operatorname{span}\{\psi_1, \dots, \psi_{j-1}\}}_{=:Y_j})\,.$$

But $d_j^2 = D_j / D_{j-1}$ (see Exercise 6.3.3), so

$$\sum_{j=1}^{\infty} \left(1 - \frac{D_j}{D_{j-1}}\right) < \infty\,.$$

But the inequality above is a sufficient and necessary condition for the existence of a positive limit of the product

$$\lim_{n \to \infty} \prod_{i-1}^{n} \frac{D_j}{D_{j-1}} = \lim_{n \to \infty} D_n \qquad (D_0 = 1 \text{ by definition})\,.$$

$(ii) \Rightarrow (i)$. It follows from the lines above that

$$\sum_{j=1}^{\infty} (1 - \delta_j^2) < \infty\,.$$

It follows from

$$(1-\delta_j) = \frac{1-\delta_j^2}{1+\delta_j} \quad \text{and} \quad 0 \le \delta_j \le 1$$

that the series above converges iff the series

$$\sum_{j=1}^{\infty}(1-\delta_j) < \infty\,.$$

Applying the Gram-Schmidt orthonormalization to ψ_j, we obtain an orthormal sequence χ_j,

$$\chi_j = c_{j1}\psi_1 + c_{j2}\psi_2 + \ldots c_{jj}\psi_j \quad c_{jj} > 0, \quad j = 1, 2, \ldots\,.$$

Consequently, $\chi_j \perp Y_{j-1}$ and $d_j = (\chi_j, \psi_j) = c_{jj}$, so

$$\sum_{j=1}^{\infty} \|\chi_j - \psi_j\|^2 = 2\sum_{j=1}^{\infty}(1-(\chi_j,\psi_j)) = 2\sum_{j=1}^{\infty}(1-d_j) < \infty\,.$$

It remains to show that vectors ψ_j are ω-linearly independent. Suppose, to the contrary, that there exists $0 \neq \{c_j\} \in \ell^2$ such that $\sum_{j=1}^{\infty} c_j\psi_j = 0$. We can always rescale the sequence c_j to have $c_1 = 1$. We have then,

$$\epsilon_n := \|\psi_1 + c_2\psi_2 + \ldots + c_n\psi_n\|^2 \to 0 \quad \text{as } n \to \infty\,.$$

But (comp. Exercise 6.3.3),

$$\epsilon_n \ge \min_{\xi} \|\psi_1 + \xi_2\psi_2 + \ldots + \xi_n\psi_n\|^2 = \frac{D(\psi_1,\psi_2,\ldots,\psi_n)}{D(\psi_2,\ldots,\psi_n)} \ge D(\psi_1,\psi_2,\ldots,\psi_n)$$

since, by the Hadamard inequality, $D(\psi_2,\ldots,\psi_n) \le \|\psi_2\|\ldots\|\psi_n\| = 1$. Consequently, $\lim_{n\to\infty} D(\psi_1,\ldots,\psi_n) = 0$, a contradiction.
$(i) \Leftrightarrow (iii)$. The proof is already contained in the reasoning above. QED.

Lemma 6.13.
Let A be a bounded linear operator with a bounded inverse. If $A^*A - I$ belongs to $\mathcal{C}_2$ then so does the operator $(A^*A)^{\frac{1}{2}} - I$ and, for any unitary operator U, the following inequality holds:

$$\|(A^*A)^{\frac{1}{2}} - I\|_2 \le \|A - U\|_2\,.$$

In the relation above, the equality holds iff U is the unitary operator from the polar decomposition of A, i.e., $U = A(A^*A)^{-\frac{1}{2}}$.

Proof. Let $H := (A^*A)^{\frac{1}{2}} - I$. We have,

$$H((A^*A)^{\frac{1}{2}} + I) = A^*A - I\,.$$

Operator $(A^*A)^{\frac{1}{2}}+I)$ is bounded below and self-adjoint and, therefore, invertible with a bounded inverse. Hence, by Exercise 6.3.5, H is a Hilbert-Schmidt operator. Let U_1 be the unitary operator from the polar decomposition of A,

$$A = U_1(A^*A)^{\frac{1}{2}} = U_1(I + H)\,.$$

From

$$A - U = U_1(I + H) - U = U_1(I + H - \underbrace{U_1^{-1}U}_{=:V})$$

follows that

$$\begin{aligned} \|A - U\|_2^2 = \|I + H - V\|_2^2 &= \mathrm{sp}[(H + I - V^*)(H + I - V)] \\ &= \mathrm{sp}[(H + I - V)(H + I - V^*)] = \mathrm{sp}H^2 + \mathrm{sp}C \end{aligned}$$

where

$$C = 2I + 2H - V - V^* - HV^* - VH$$

is a self-adjoint operator from $\mathcal{C}_1$. Let χ_j be a complete eigensystem for H with the corresponding eigenvalues λ_j. Then,

$$(C\chi_j, \chi_j) = 2 + 2\lambda_j - 2\Re(V\chi_j, \chi_j) - 2\lambda_j\Re(V\chi_j, \chi_j) = 2(\lambda_j + 1)(1 - \Re(V\chi_j, \chi_j)) .$$

Since

$$1 - \Re(V\chi_j, \chi_j) \geq 0 \quad j = 1, 2, \ldots$$

we have $\mathrm{sp}(C) \geq 0$ and so the relation

$$\|A - U\|_2^2 \geq \mathrm{sp}\, H^2$$

holds true. The equality holds if $\Re(V\chi_j, \chi_j) = 1$, for all j. In view of the fact that $|(V\chi_j, \chi_j)| = 1$, this implies that actually $(V\chi_j, \chi_j) = 1$. Hence $V = I$ which implies $U_1 = U$. QED.

Theorem 6.14.
Let $\psi_j \in X$, $j = 1, \ldots$, be a sequence of vectors complete in X. The following conditions are equivalent to each other.

(i) The sequence ψ_j is a Bari basis in X.

(ii) The sequence ψ_j is ω-linearly independent, and matrix $(\psi_j, \psi_k) - \delta_{jk}$ is of Hilbert-Schmidt class, i.e.

$$\sum_{j,k=1}^{\infty} |(\psi_j, \psi_k) - \delta_{jk}|^2 < \infty . \tag{6.3.7}$$

Proof. $(i) \Rightarrow (ii)$. If ψ_j is (any Schauder) basis, and $\sum_{j=1}^{\infty} c_j\psi_j = 0$ (for *any* sequence c_j) then, by uniqueness of the components with respect to a basis, all c_j must be equal zero. The ω-linear independence condition is thus satisfied trivially. Let χ_j be an orthonormal basis such that

$$\sum_{j=1}^{\infty} \|\psi_j - \chi_j\|^2 < \infty ,$$

and let A be the operator taking χ_j into ψ_j. By Theorem 6.9, A is a bounded linear operator with a bounded inverse, and $T := A - I \in \mathcal{C}_2$. We have then,

$$(\psi_j, \psi_k) - \delta_{jk} = (A\chi_j, A\chi_k) - (\chi_j, \chi_k) = (\underbrace{(A^*A - I)}_{=:B}\chi_j, \chi_k)$$

and $B = (T+I)^*(T+I) - I = T + T^* + T^*T \in \mathcal{C}_2$. Consequently,

$$\sum_{j,k=1}^{\infty} |(\psi_j, \psi_k) - \delta_{jk}|^2 = \sum_{j=1}^{\infty} |B\chi_j|^2 < \infty \,.$$

$(ii) \Rightarrow (i)$. Let χ_j be an arbitrary othonormal basis in X. Define a linear self-adjoint operator $I + G$ such that

$$((I+G)\chi_j, \chi_k) = (\psi_j, \psi_k) \quad j,k = 1, \dots \,.$$

It follows from condition 6.3.7 that $G \in \mathcal{C}_2$. Next, define an operator A taking χ_j into ψ_j and extend it by linearity to the span of χ_j. Then, for any $x = \sum_{j=1}^{n} x_j \chi_j$,

$$\|Ax\|^2 = \sum_{j,k=1}^{n} x_j \bar{x}_k (\psi_j, \psi_k) = ((I+G)x, x) \le (1 + \|G\|)\|x\|^2 \,,$$

i.e., the operator A is continuous and, by continuity, it can be extended to a continuous operator with the same norm to the whole X. We will denote the extension with the same symbol A. The ω-linear independence implies now that $I + G$ is injective. Indeed, if

$$(I+G)x = 0 \quad \text{where } x = \sum_{j=1}^{\infty} x_j \chi_j, \quad \{x_j\} \in \ell^2$$

then

$$\|Ax\| = 0 \quad \Rightarrow \quad Ax = \sum_{j=1}^{\infty} x_j \psi_j = 0 \quad \Rightarrow \quad x_j = 0, \quad j = 1, \dots \,.$$

Compactness of G implies then that $I + G$ is invertible with a bounded inverse, i.e., there exists $\delta > 0$ such that

$$\delta \|x\|^2 \le ((I+G)x, x) = \|Ax\|^2 \,.$$

The density of span of ψ_j in X (and, therefore, the range of A) implies then that A is invertible on X with a bounded inverse. Since $G = A^*A - I \in \mathcal{C}_2$, by Lemma 6.13, $(A^*A)^{\frac{1}{2}} - I \in \mathcal{C}_2$ as well, and

$$\|A - U\|^2_{\mathcal{C}_2} = \|(A^*A)^{\frac{1}{2}} - I\|^2_{\mathcal{C}_2} < \infty \,,$$

with the unitary operator $U = A(A^*A)^{-\frac{1}{2}}$. Finally, introducing an orthonormal basis $\omega_j = U\chi_j$, we have,

$$\sum_{j=1}^{\infty} \|\psi_j - \omega_j\|^2 + \sum_{j=1}^{\infty} \|(A-U)\chi_j\|^2 < \infty \,.$$

QED.

Exercises

6.3.1. Let $u, v \in X$ be any two elements of a Hilbert space X. Define a function:

$$f(z) = \|v - zu\|^2 = (v - zu, v - zu) \,.$$

Show that

$$\min_{z \in \mathbb{C}} f(z) = f(z_0)$$

where

$$z_0 = \frac{(v,u)}{\|u\|^2}$$

and

$$f(z_0) = \|v\|^2 - \frac{|(v,u)|^2}{\|u\|^2} = \|v\|^2 - |(v,\hat{u})|^2$$

with $\hat{u} = u/\|u\|$. (1 point)

6.3.2. Let $T \in \mathcal{C}_2$, and $T_1 := (I+T)^{-1} - I$. Prove that $T_1 \in \mathcal{C}_2$ as well. (5 points)

6.3.3. Prove that

$$\min_{\xi} \|\xi_1\psi_1 + \xi_2\psi_2 + \ldots + \psi_n\|^2 = \mathrm{dist}^2(\psi_n, \mathrm{span}\{\psi_1, \ldots, \psi_{n-1}\}) = \frac{D(\psi_1, \psi_2, \ldots, \psi_n)}{D(\psi_1, \ldots, \psi_{n-1})}.$$

(5 points)

6.3.4. Let $\psi_j, j = 1, \ldots$ be a sequence of linearly independent vectors in a Hilbert space X. Define

$$D(\psi_1, \ldots, \psi_n) := \det\left((\psi_i, \psi_j)_1^n\right).$$

Prove the *Hadamard inequality:*

$$D(\psi_1, \ldots, \psi_{n+m}) \le D(\psi_1, \ldots, \psi_n)\, D(\psi_{n+1}, \ldots, \psi_{n+m}).$$

Hint: Proceed in the following steps.

Step 1: Let S be a subspace of X and P_S denote the orthogonal projection onto S. Use the induction in n (Exercise 6.3.3 may be useful) to prove that

$$D(P_S\psi_1, \ldots, P_S\psi_n) \le D(\psi_1, \ldots, \psi_n).$$

Step 2: Let $M = \mathrm{span}\{\psi_1, \ldots, \psi_n\}$ and let $M^\perp$ be the orthogonal complement of M in $Y := \mathrm{span}\{\psi_1, \ldots, \psi_n, \psi_{n+1}, \ldots, \psi_{n+m}\}$. Let $P : Y \to M^\perp$ denote the orthogonal projection. Prove that

$$D(\psi_1, \ldots, \psi_n, \psi_{n+1}, \ldots, \psi_{n+m}) = D(\psi_1, \ldots, \psi_n, P\psi_{n+1}, \ldots, P\psi_{n+m}).$$

Step 3: Show that

$$D(\psi_1, \ldots, \psi_n, P\psi_{n+1}, \ldots, P\psi_{n+m}) = D(\psi_1, \ldots, \psi_n)\, D(P\psi_{n+1}, \ldots, P\psi_{n+m})$$

and conclude the proof by using Step 1 result.

(10 points)

6.3.5. Let X be a Hilbert space, and $A : X \to X$ a Hilbert-Schmidt operator. Let $B : X \to X$ be a continuous operator with a bounded inverse. Prove that the composition AB is a Hilbert-Schmidt operator as well. (5 points)

6.3.6. (5 points)

6.3.7. (5 points)

6.4 ▪ Glazman's Criterion for Eigenvectors of a Dissipative Operator to Form a Basis

Dissipative operators. A linear operator

$$X \supset D(A) \ni x \to Ax \in X$$

is called *dissipative* if

$$\Im(Ax, x) \geq 0 \quad x \in D(A) .$$

If A is bounded (and, therefore, defined on the whole X), then

$$\Im(Ax, x) = \frac{1}{2i}[(Ax, x) - \overline{(Ax, x)}] = (\frac{1}{2i}(A - A^*)x, x) ,$$

so the condition is equivalent to the semi-positive definitness of $\frac{1}{2i}(A - A^*)$.

Theorem 6.15 (Glazman).
Let ψ_j, $j = 1, 2, \ldots$, be a system of unit eigenvectors corresponding to distinct eigenvalues λ_j of a dissipative operator such that

$$\sum_{\substack{j,\,k\,=\,1 \\ j\,\neq\,k}}^{\infty} \frac{\Im\lambda_j \, \Im\lambda_k}{|\lambda_j - \overline{\lambda_k}|^2} < \infty .$$

Then the system ψ_j forms a Riesz basis for the closure of its span,

$$\overline{\mathit{span}\{\psi_j,\, j = 1, 2, \ldots\}} .$$

Note that the theorem does not provide results on the completeness of the basis for X, we still need Keldysh's results for this. The criterion is expressed only in terms of the eigenvalues so, in principle, we do not need to check any conditions for the eigenvectors. Finally, the eigenvalues are not assumed to be simple.

In what follows, we will prove a more general result that will cover the Glazman's Theorem as a special case. Let λ_j be the system of *non-real* eigenvalues of a continuous operator A with a compact imaginary component, with corresponding generalized eigenspaces X_j. By Theorem **??**, the numbers λ_j are isolated and spaces X_j are finite-dimensional. Recall the Riesz projectors,

$$P_k = -\frac{1}{2\pi i} \int_{|\lambda - \lambda_k| = r_k} (A - \lambda I)^{-1} \, d\lambda \quad k = 1, 2 \ldots$$

where radius r_k is sufficiently small so the circle does not contain any of the other eigenvalues. We have:

$$P_k X = X_k \quad \text{and} \quad P_k X_j = 0 \quad \text{for } j \neq k .$$

Lemma 6.16. *A sequence ψ_j made up of bases of spaces X_k is ω-linearly independent.*

Proof. Let

$$\sum_j c_j \psi_j = 0, \quad \{c_j\}_1^\infty \in \ell^2 .$$

Then

$$P_k \left(\sum_j c_j \psi_j \right) = \sum_{j = m_k + 1}^{m_{k+1}} c_j \psi_j$$

where $\{\psi_j \,:\, j = m_k + 1, \dots, m_{k+1}\}$ is a basis for X_j. Consequently, $c_j = 0$. QED.

Lemma 6.17. *Let ψ_j, $j = 1, \dots$ be a sequence of ω-linearly independent vectors in a Hilbert space. Then, for any $N > 0$,*

$$\underbrace{\text{span}\{\psi_j\}_1^N}_{=:X_N} \oplus \underbrace{\overline{\text{span}\{\psi_j\}_{N+1}^\infty}}_{=:Y_N} = \underbrace{\overline{\text{span}\{\psi_j\}_1^\infty}}_{=:X} .$$

Proof.

Step 1: $(X_N + \overline{Y_N} \subset \overline{X})$. Indeed,

$$x \in X_N,\, y_n \in Y_N,\, y_n \to y \quad \Rightarrow \quad x + y_n \in X,\, x + y_n \to x + y \in \overline{X} .$$

Step 2: The algebraic sum on the left is indeed a direct sum. Suppose to the contrary that we have

$$0 \neq x = \sum_{j=1}^N x_j \psi_j \in \overline{\text{span}\{\psi_j\}_{N+1}^\infty} .$$

Assume $x_k \neq 0$ for some $1 \leq k \leq N$. By Step 1 result,

$$x_k \psi_k \in \text{span}\{\psi_j, 1 \leq j \leq N,\, j \neq k\} + \overline{\text{span}\{\psi_j\}_{N+1}^\infty} \subset \overline{\text{span}\{\psi_j, j \in \mathbb{N}, k \neq l\}} .$$

But this contradicts the ω-linear independence of ψ_j's.

Step 3: $(X_N \oplus \overline{Y_N} \supset \overline{X})$. Once we have established that the algebraic sum of the two closed subspaces on the left is a direct sum, we can introduce a linear (skewed) projection P projecting the space on the left-hand side onto X_N (in the direction of $\overline{Y_N}$). Let now $z \in \overline{X}$. There exists thus a sequence $z_n \in X$, $z_n \to z$. Trivially, $z_n \in X_N \oplus Y_N \subset X_N \oplus \overline{Y_N}$. From the uniqueness of the direct sum decomposition,

$$z_n = \sum_{j=1}^{m_n} z_n^j \psi_j = \underbrace{\sum_{j=1}^{N} z_n^j \psi_j}_{=:x_n} + \underbrace{\sum_{j=N+1}^{m_n} z_n^j \psi_j}_{=:y_n} ,$$

follows that $x_n = P z_n$. As z_n is Cauchy, so must be x_n (P is continuous) and, therefore, $x_n \to x \in X_N$, for some x. Consequently,

$$x + y_n = \underbrace{x - x_n}_{\to 0} + \underbrace{x_n + y_n}_{= z_n \to z} \to z$$

which proves that $z \in X_N + \overline{Y_N}$.

QED

Theorem 6.18.

(i) Let A be a linear, continuous and dissipative operator with a compact imaginary component. Let $\lambda_j,\ j=1,2\ldots$ be the eigenvalues of the operator with corresponding eigenspaces[18] *$X_j,\ n_j=\dim X_j$. If*

$$\sum_{\substack{j,\,k\,=\,1\\ j\,\neq\,k}}^{\infty} \min\{n_j,n_k\}\,\frac{\Im\lambda_j\Im\lambda_k}{|\lambda_j-\overline{\lambda}_k|^2}<\infty^{19}\,, \tag{6.4.8}$$

then a sequence made up of orthonormal basis for X_j forms a Bari basis for the closure of its linear span.

(ii) If a weaker condition holds:

$$\sum_{\substack{j,\,k\,=\,1\\ j\,\neq\,k}}^{\infty} \frac{\Im\lambda_j\Im\lambda_k}{|\lambda_j-\overline{\lambda}_k|^2}<\infty\,, \tag{6.4.9}$$

then the sequence is a Riesz basis for the closure of its span.

Proof. The strategy in the first case is to show that the Gram matrix is a Hilbert-Schmidt matrix, i.e., it satisfies (6.3.7), and invoke Theorem 6.14. The strategy for the second case is to show that the Gram matrix represents a bounded invertible operator in ℓ^2, and utilize Theorem 6.5.

Part (i). Let $\Im A:=\frac{1}{2i}(A+A^*)$ be the imaginary part of operator A. By the Cauchy-Schwarz inequality,

$$|(\Im A\phi,\psi)|^2\le(\Im A\phi,\phi)\,(\Im A\psi,\psi)\,.$$

Pick unit vectors $\phi\in X_j,\ \psi\in X_k$. We have,

$$(\Im A\phi,\psi)=\frac{1}{2i}|(A\phi,\psi)-(\phi,A\psi)|=\frac{1}{2i}(\lambda_j-\overline{\lambda}_k)(\phi,\psi)$$

and

$$(\Im A\phi,\phi)=\Im\lambda_j,\quad(\Im A\psi,\psi)=\Im\lambda_k\,.$$

Consequently,

$$|(\phi,\psi)|^2\le 4\frac{\Im\lambda_j\Im\lambda_k}{|\lambda_j-\overline{\lambda}_k|^2}=:c_{jk}\,.$$

Let $\phi_r,\ r=1,\ldots,n_j$, and $\psi_q,\ q=1,\ldots,n_k$ be orthonormal bases for X_j and X_k, respectively. Then for vectors $\psi=\sum_q(\phi_r,\psi_q)\psi_q\in X_k$, we have:

$$c_{jk}\ge|(\psi,\phi_r)|^2=\sum_{q=1}^{n_k}|(\psi_q,\phi_r)|^2,\quad r=1,2,\ldots,n_j$$

and, therefore,

$$\sum_{r=1}^{n_j}\sum_{q=1}^{n_k}|(\psi_q,\phi_r)|^2\le n_jc_{jk}\,.$$

[18]*Not* generalized eigenspaces.

[19]We assume that, if $\Im\lambda_j\Im\lambda_k=0$, the corresponding contribution vanishes, even in the case when $\min\{n_j,n_k\}=\infty$.

Interchanging the roles of X_j and X_k, we obtain an analogous inequality with n_j and n_k switched. Consequently,

$$\sum_{r=1}^{n_j}\sum_{q=1}^{n_k} |(\psi_q,\phi_r)|^2 \le \min\{n_j,n_k\}\frac{4\Im\lambda_j\Im\lambda_k}{|\lambda_j-\overline{\lambda}_k|^2}\,.$$

Let $\mathcal{A}_{ij} = (\psi_i,\psi_j)$ be the Gram matrix for the sequence of vectors ψ_j. It follows from the assumption (6.4.8) that matrix $\mathcal{A}-I$ is of Hilbert-Schmidt class. Consequently, by Lemma 6.16 and Theorem 6.14, the first part of the theorem is proved.

Part (ii). Let ϕ_p, $p=1,\ldots,n_j$, and ψ_q, $q=1,\ldots,n_k$ be again orthonormal bases for X_j and X_k, respectively. Consider the orthogonal projection P_{jk} from X_j into X_k. Equivalently,

$$P_{jk}\phi_p = \sum_{q=1}^{n_k}(\phi_p,\psi_q)\psi_q \qquad p=1,\ldots,n_j\,.$$

Thus, the Gram matrix

$$(\phi_p,\psi_q) \quad p=1,\ldots,n_j,\; q=1,\ldots,n_k$$

is the matrix representation of the orthogonal projection map in bases ϕ_p and ψ_q. We have,

$$\begin{aligned}\|P_{jk}\|^2 &= \max_{\|\phi\|=1}\|P_{jk}\phi\|^2\\ &= \max_{\|\phi\|=1}\max_{\|\psi\|=1}|(P_{jk}\phi,\psi)|^2\\ &\le \frac{\Im\lambda_j\Im\lambda_k}{|\lambda_j-\lambda_k|^2}\,.\end{aligned}$$

By Exercise 6.4.1, the Gram matrix $\mathcal{A}_{ij} = (\psi_i,\psi_j)$ corresponding to the union of eigenvectors ψ_i generates a discrete operator $A:\ell^2\to\ell^2$ such that

$$\|A-I\|^2 \le \sum_{\substack{j,\,l\,=\,1\\ j\,\neq\,l}}^{\infty} c_{jl} < \infty\,.$$

In order to apply Theorem 6.5(iv), we need operator A to have a bounded inverse as well. To apply the Neumann series argument, we need the sum above to be strictly bounded by one. This may not be true for the whole series but it is certainly true for its remainder,

$$\sum_{\substack{j,\,l\,=\,N+1\\ j\,\neq\,l}}^{\infty} c_{jl} < 1\,,$$

for sufficiently large N. However, by Lemma 6.17,

$$\overline{\operatorname{span}\{\psi_j, j=1,2,\ldots\}} = \operatorname{span}\{\psi_1,\ldots,\psi_N\}\oplus\overline{\operatorname{span}\{\psi_j, j=N+1,\ldots\}} \tag{6.4.10}$$

and, therefore, it is sufficient to show that $\psi_j, j=N+1,\ldots$ provide a Riesz basis for the the closure of its span, for sufficiently large N. QED.

Exercises

6.4.1. Let $x = \{x_j\}_1^\infty$ where $x_j = \{x_{jl}\}_1^{n_j}$, and $y = \{y_k\}_1^\infty$ where $y_k = \{y_{km}\}_1^{n_k}$. Define discrete ℓ^2 norms:

$$\|x\|^2 = \sum_{j=1}^{\infty} \|x_j\|^2 = \sum_{j=1}^{\infty}\sum_{l=1}^{n_j} |x_{jl}|^2, \qquad \|y\|^2 = \sum_{k=1}^{\infty} \|y_k\|^2 = \sum_{k=1}^{\infty}\sum_{m=1}^{n_k} |y_{km}|^2 .$$

Prove the inequality:

$$\sup_{\|x\|=1} \sup_{\|y\|=1} |\sum_{j=1}^{\infty}\sum_{l=1}^{n_j}\sum_{k=1}^{\infty}\sum_{m=1}^{n_k} A_{kmjl} x_{jl} \bar{y}_{km}| \le \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \|A_{kj}\|^2$$

where $A_{kj} : \mathbb{C}^{n_j} \to \mathbb{C}^{n_k}$ is the map generated by matrix $A_{k \cdot j \cdot}$, i.e.

$$A_{kj} x_j = A_{kj}(\{x_{jl}\}) := \sum_{m=1}^{n_k} (\sum_{l=1}^{n_j} A_{kmjl} x_{jl}) e_m$$

with e_m denoting the canonical basis in $\mathbb{C}^{n_k}$.
(5 points)

6.4.2. (5 points)

Bibliography

[1] L. Demkowicz. Lecture Notes on Mathematical Methods in Science and Engineering. https://users.oden.utexas.edu/ leszek/, 2022. (Not cited)

[2] I.C. Gohberg and M.G. Krein. *Introduction to the Theory of Linear Nonselfadjoint Operators in Hilbert Space*. American Mathematical Society, Providence, 1965. (Cited on pp. vi, 39)

[3] B. Ya. Levin. *Lectures on Entire Functions*. American Mathemtical Society, 1996. (Cited on pp. vi, 23)

[4] J.T. Oden and L.F. Demkowicz. *Applied Functional Analysis for Science and Engineering*. Chapman & Hall/CRC Press, Boca Raton, 2018. Third edition. (Cited on pp. vi, 5, 9, 12, 13, 35, 49, 50)