

4

Spectral Theory of Linear Operators

In this chapter we examine the spectral properties of operators commonly encountered in electromagnetics. We emphasize the role of eigenvalues and eigenfunctions in spectral theory since these quantities play an important role in many applications, as both mathematical and physical entities. The primary goal of this chapter is to present elements from the spectral theory of operators on infinite-dimensional spaces. Operators on finite-dimensional spaces and their associated matrix representations are also briefly covered, as this material forms an appropriate starting point for discussion.

After considering the finite-dimensional case, we then show that for certain classes of operators acting on infinite-dimensional spaces, simple results may be obtained. In particular, when the operator is compact and self-adjoint, the spectral theory becomes relatively simple. Similar results are obtained if the operator is unbounded, self-adjoint, and positive with compact inverse. Fortunately, these two cases occur quite frequently in physical applications. For example, the first case frequently arises in the study of integral equations pertaining to shielded electrodynamic problems and in the integral equation formulation of electrostatic problems. The second class encompasses various differential operators that commonly occur in electromagnetics applications.

The chapter begins with the presentation of the basic concepts behind spectral elements, including the point, continuous, and residual spectrum. Next, spectral properties of various operators are presented, including discussions of various eigenvalue problems. Spectral representations and expansion theorems are then presented for different classes of operators. After a section on functions of operators, the chapter concludes with a discussion of the solution of operator equations using spectral methods.

4.1 General Concepts

Consider a linear operator $A : H \rightarrow H$ mapping a Hilbert space H to itself. The *spectral problem* consists, in part, of examining solutions of the equation

$$Ax = \lambda x, \quad (4.1)$$

where $\lambda \in \mathbf{C}$ is known as an *eigenvalue* (or *characteristic value*) and $x \in D_A$, $x \neq 0$, is called an *eigenvector* (or often *eigenfunction* if H is a function space). We refer to this as a *standard eigenvalue problem*. It is convenient for later work to write the above as

$$(A - \lambda I)x = 0, \quad (4.2)$$

where I is the identity operator on H .

Another, more general formulation is possible, involving two operators $A, B : H \rightarrow H$, given as

$$Ax = \lambda Bx \quad (4.3)$$

or

$$(A - \lambda B)x = 0 \quad (4.4)$$

and called a *generalized eigenvalue problem*. Of course, (4.3) reduces to (4.1) when operator B is the identity operator on H .

Finally, a third formulation commonly occurs in electromagnetic applications, known as a *nonstandard eigenvalue problem*. It is formulated as

$$A(\gamma)x = 0, \quad (4.5)$$

where γ is called a nonstandard eigenvalue and $x \neq 0$ is a nonstandard eigenvector. The operator $A(\gamma)$, dependent on the variable γ , is called an *operator-valued function* (of γ). The standard and generalized eigenvalue problems can also be viewed as special cases of (4.5), where, for instance, we may consider the nonstandard eigenvalue as a parameter γ such that

$$(A(\gamma) - \lambda(\gamma)I)x = 0.$$

Often we seek the value of the nonstandard eigenvalue γ such that $\lambda(\gamma) = 0$.

Example 4.1.

Consider the microstrip transmission line depicted in Figure 4.1. The strip is infinitely thin and perfectly conducting, and exists over $z \in (-\infty, \infty)$. We seek guided modes of the transmission line in the form

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= \mathbf{E}(x, y) e^{i(\omega t - \beta z)}, \\ \mathbf{H}(\mathbf{r}, t) &= \mathbf{H}(x, y) e^{i(\omega t - \beta z)}, \end{aligned}$$

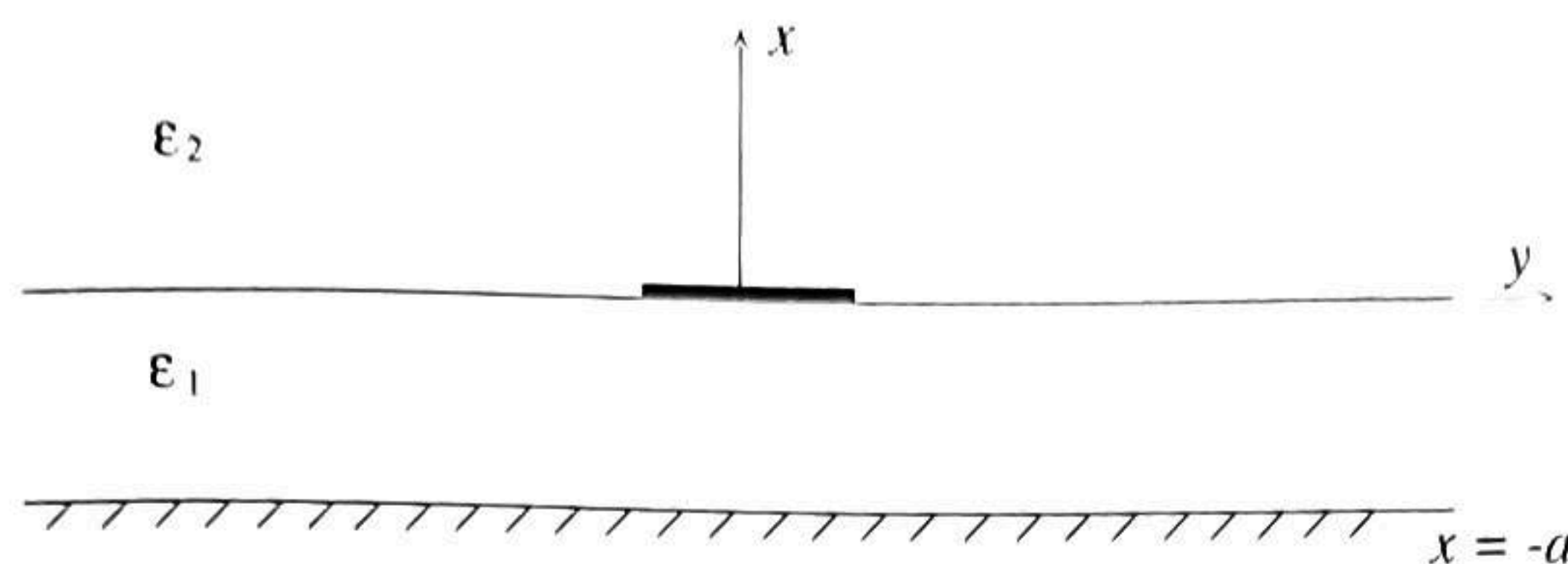


Figure 4.1: Geometry of a microstrip transmission line as considered in Example 4.1. The determination of the guided modes of the microstrip line can be associated with the spectral properties of a self-adjoint positive operator.

where ω is the radian frequency and β is the propagation constant, with $\omega, \beta \in \mathbf{R}$. Letting $\nabla_\beta \times \mathbf{x}$, $\nabla_\beta \cdot \mathbf{x}$, and $\nabla_\beta x$ be the differential operators obtained from curl, divergence, and gradient, respectively, upon the substitution $\partial/\partial z \rightarrow -i\beta$, the problem of determining guided microstrip modes can be formed as the spectral problem [1]

$$A(\beta) \mathbf{E} = \omega^2(\beta) \mathbf{E},$$

where $A : \mathbf{L}^2(\Omega)^3 \rightarrow \mathbf{L}^2(\Omega)^3$ is defined as

$$A(\beta) \mathbf{E} \equiv \frac{1}{\varepsilon} \{ \nabla_\beta \times \nabla_\beta \times \mathbf{E} - \varepsilon \nabla_\beta \nabla_\beta \cdot \varepsilon \mathbf{E} \},$$

$$D_A \equiv \left\{ \begin{array}{l} \mathbf{E} : \mathbf{E}, \nabla_\beta \times \mathbf{E}, \nabla_\beta \times \nabla_\beta \times \mathbf{E} \in \mathbf{L}^2(\Omega)^3, \\ \nabla_\beta \cdot \varepsilon \mathbf{E} \in \mathbf{H}_{0,1}, \mathbf{n} \times \mathbf{E}|_\Gamma = \mathbf{0} \end{array} \right\},$$

with Γ being the contour of both the strip and the ground plane. It can be shown that $A(\beta)$ is self-adjoint and positive, so that eigenvalues $\lambda(\beta) = \omega^2(\beta)$ are real-valued and positive¹ (see Section 4.2.6).

Furthermore, A in (4.5) may depend on a large number of independent variables (in electromagnetics these may be frequency, transform-domain wavenumbers, geometrical values, constitutive quantities, etc.), and usually A is nonlinear in these variables. In particular, (4.5) can be written as $A(\tau, \gamma)x = 0$, where one seeks a solution $\gamma(\tau)$. This often occurs in resonance problems, where γ represents a natural frequency and τ is a vector of geometrical and constitutive quantities. The nonstandard formulation also commonly occurs in complicated waveguiding problems, where γ would represent a propagation constant. In this chapter we concentrate on the standard eigenvalue problem, although both the generalized and nonstandard problems are also discussed.

¹In [1] it is also established that the essential spectrum (for our purposes continuous spectrum; see Section 4.1.2) of A is identical to the essential spectrum of the grounded dielectric without the strip and that guided microstrip modes exist.

4.1.1 Operators on Finite-Dimensional Spaces

Beginning with the finite-dimensional case, consider (4.1) where $A \in \mathbf{C}^{n \times n}$ is an $n \times n$ complex matrix (representing a finite-dimensional operator with respect to some basis, or the projection of an infinite-dimensional operator onto a finite-dimensional space) and $x \in \mathbf{C}^n = D_A$. The eigenvalues of A are the solutions $\lambda \in \mathbf{C}$ of (4.1), which is equivalent to the set of values λ for which $A - \lambda I$ is not invertible. For every eigenvalue λ , (4.2) has nontrivial solutions x , which are the eigenvectors corresponding to that particular eigenvalue. The set of all eigenvalues of A is called the *spectrum* of A , denoted as $\sigma(A)$, i.e.,

$$\sigma(A) \equiv \left\{ \lambda : (A - \lambda I)^{-1} \text{ does not exist} \right\}.$$

Because we are interested in nontrivial solutions of (4.2), we get as a necessary condition

$$\det(A - \lambda I) = 0, \quad (4.6)$$

which leads to an n th-order polynomial known as the *characteristic polynomial*. By the fundamental theorem of algebra, (4.6) has n roots, which form the spectrum² of A .

Once a particular eigenvalue has been determined, its corresponding eigenvector(s) can be found from (4.2). Those eigenvectors, together with the zero vector, form what is known as the *eigenspace* of A for that particular eigenvalue (i.e., the null space of the operator $A - \lambda I$ for a given λ). The case of the union of all eigenspaces spanning H is very important, as is discussed later in the spectral theorems.

The *algebraic multiplicity* of an eigenvalue is the number of times that eigenvalue appears as a root of (4.6) (i.e., its multiplicity as a zero of the characteristic polynomial), and its *geometric multiplicity* is the dimension of the corresponding eigenspace. It can be shown that the geometric multiplicity of an eigenvalue cannot exceed its algebraic multiplicity.

It is convenient to recall that the determinant of a matrix is equal to the product of its eigenvalues, counting algebraic multiplicity, i.e.,

$$\det(A) = \prod_{i=1}^n \lambda_i$$

such that if any eigenvalue has the value 0, then $\det(A) = 0$ and A is not invertible.

Regarding the location of the eigenvalues we have the following theorem.

²It is important to work within complex Hilbert spaces even when dealing with real-valued matrices, because if we make the restriction $\lambda \in \mathbf{R}$ eigenvalues may not exist.

Theorem 4.1. (*Gerschgorin circle theorem*) Every eigenvalue of the operator $A : \mathbf{C}^n \rightarrow \mathbf{C}^n$ lies in one of the complex domains

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}} |a_{ij}|$$

for $i = 1, 2, \dots, n$.

Proof. Let λ be an eigenvalue that does not lie in one of the specified domains. Then, $|\lambda - a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}} |a_{ij}|$ and so $(A - \lambda I)$ is diagonally dominant. By Theorem 3.26, $(A - \lambda I)$ has an inverse, which contradicts the assumption that λ is an eigenvalue. ■

Furthermore, if the elements of the matrix A are continuous functions of a parameter γ , such that the eigenvalue problem is

$$(A(\gamma) - \lambda(\gamma)I)x = 0,$$

then $\det(A(\gamma))$ is a continuous function of γ and the eigenvalues $\lambda(\gamma)$ vary continuously with γ .

For all other values of λ , $A - \lambda I$ is invertible such that (4.2) has only the trivial solution $x = 0$. These values of λ make up the *resolvent set*, denoted as $\rho(A)$, which is the complement of the spectrum, i.e.,

$$\rho(A) \equiv \mathbf{C} - \sigma(A).$$

4.1.2 Operators on Infinite-Dimensional Spaces

It turns out that the spectrum of a linear operator $A : H \rightarrow H$, where H is an infinite-dimensional Hilbert space, is much more complicated than for an operator acting on a finite-dimensional space. Considering the operator $A : H \rightarrow H$ and motivated by (4.2), it is useful to examine properties of the linear operator $A_\lambda \equiv A - \lambda I$, where I is the identity operator on H . We denote the domain of A_λ as $D_{A-\lambda I} = D_A$, which will be assumed dense in H . The range of A_λ will depend on λ and may or may not be dense³ in H . As discussed in the previous chapter, the operator inverse to A_λ is called the *resolvent operator* (or simply the *resolvent*), denoted as

$$R_\lambda(A) \equiv (A - \lambda I)^{-1}.$$

For a given operator A , $R_\lambda(A)$ is an operator-valued function of λ , analytic on $\rho(A)$, with $R_{\lambda=0}(A)$ being simply the inverse operator A^{-1} . If $R_\lambda(A)$ exists for a particular λ , then that λ cannot be an eigenvalue since (4.2) would only have trivial solutions. Therefore, it is not surprising that spectral properties of an operator A can be determined by considering properties of $R_\lambda(A)$ for various values⁴ of λ :

³Recall from Theorem 2.7 that $R_{A-\lambda I}$ is dense in H if and only if $(R_{A-\lambda I})^\perp = \{0\}$. The dimension of $(R_{A-\lambda I})^\perp$ is called the *deficiency* of λ for obvious reasons.

⁴This classification scheme is not unique. Here we follow [2, pp. 125–127] and [37, pp. 194–195], among others.

1. **Values of λ for which $R_\lambda(A)$ does not exist** (i.e., $A - \lambda I$ is not invertible and therefore (4.2) has a nontrivial solution).

The set of these values of λ forms the *point spectrum* of A , denoted as $\sigma_p(A)$. Such values of λ are called eigenvalues of A , and for a given eigenvalue λ the corresponding nontrivial solutions $x \in H$ are eigenvectors corresponding to that eigenvalue. Eigenvalues of A can be considered to be poles of the resolvent operator $R_\lambda(A)$.

2. **Values of λ for which $R_\lambda(A)$ exists** (i.e., $A - \lambda I$ is invertible and (4.2) has only the trivial solution $x = 0$):

- (a) **Closure of the range of $A - \lambda I$ is H** (i.e., the range of $A - \lambda I$ is dense in H):

- i. if the above condition holds and $R_\lambda(A)$ is bounded, then such a value λ is called a *regular value* of A . Regular values are not part of the spectrum of A , and therefore regular values make up the resolvent set $\rho(A)$.
- ii. if the above condition holds and $R_\lambda(A)$ is unbounded (not continuous), then such a value λ is part of the *continuous spectrum* of A , denoted as $\sigma_c(A)$.

- (b) **Closure of the range of $A - \lambda I$ is a proper subset⁵ of H .**

The set of such values of λ forms the *residual spectrum* of A , denoted as $\sigma_r(A)$.

The total spectrum is then

$$\sigma(A) = \mathbf{C} - \rho(A) = \sigma_r(A) \cup \sigma_p(A) \cup \sigma_c(A).$$

Note that we associate with operators on complex finite-dimensional spaces a point spectrum and a resolvent set, while the additional notions of continuous and residual spectrum only arise in the infinite-dimensional case. The residual spectrum does not usually occur in electromagnetic applications,⁶ while the point and continuous spectrum are extremely important in the analysis to follow.

Because the above criteria are fairly easy to apply, except possibly the denseness condition, and since we will not be concerned with the residual spectrum here, then for our purposes it will be sufficient to check the existence of $R_A(\lambda)$ —or the lack thereof. This will divide the continuous spectrum and resolvent set from the point spectrum. If $R_A(\lambda)$ exists, then its boundedness can be used to separate the continuous spectrum from the resolvent set.

⁵A set X can be a subset of itself. To exclude this case, if X is a subset of Y and $X \neq Y$, we say X is a *proper* subset of Y .

⁶For an example relating to the residual spectrum, see the analysis of the shift operator in [2, pp. 126–127].

In all examples presented here the residual spectrum will be empty (the range of $A - \lambda I$ will be dense in H); in a few later theorems the lack of a residual spectrum will be shown to be true for several classes of operators important in electromagnetic applications (see Theorems 4.18 and 4.26). It is worthwhile to note that for a linear operator on a complex Hilbert space, the spectrum is never empty.

Finally, since we are primarily interested in the point and continuous spectrum, and since we will later identify elements of the continuous spectrum as having some properties similar to those of eigenvalues, we introduce the concept of the *approximate spectrum* [2, p. 127].

Definition 4.1. Consider a linear operator $A : H \rightarrow H$. A value λ is in the approximate spectrum of A , denoted by $\sigma_a(A)$, if there exists a sequence of elements $x_n \in D_A$ with $\|x_n\| = 1$ such that $\|(A - \lambda I)x_n\| < 1/n$. The elements x_n are called *approximate eigenvectors* of A , and the corresponding values $\lambda \in \sigma_a(A)$ are said to be *approximate eigenvalues* of A .

Theorem 4.2. The approximate spectrum of a linear operator $A : H \rightarrow H$ contains the point and continuous spectrum of A , but not any elements of the resolvent set.

Proof. Let $\lambda_n \in \sigma_p(A)$, with corresponding eigenfunction x_n , $\|x_n\| = 1$. Because $(A - \lambda_n I)x_n = 0$, the desired inequality is seen to hold for all finite n . Now, let $\lambda \in \sigma_c(A)$. Because by definition $(A - \lambda I)^{-1}$ is unbounded, there exists a sequence of functions y_n such that $\|(A - \lambda I)^{-1}y_n\| > n\|y_n\|$. With $x_n = (A - \lambda I)^{-1}y_n$ we have $\|x_n\| > n\|(A - \lambda I)x_n\|$, or $\|(A - \lambda I)x_n\| < (1/n)\|x_n\|$ and so with $\|x_n\| = 1$, λ clearly belongs to the approximate spectrum. If $\lambda \in \rho(A)$, then $(A - \lambda I)^{-1}$ is bounded and so λ cannot be in the approximate spectrum. ■

Elements of the residual spectrum may or may not be in the approximate spectrum.

Examples of Point Spectrum, Continuous Spectrum, and Resolvent Set

1. For the operator $A : H \rightarrow H$ defined by $Ax \equiv \alpha x$, where $\alpha \in \mathbf{C}$ (i.e., multiplication by a complex number, $A = \alpha I$), all vectors $x \in H$, $x \neq 0$ are eigenvectors with eigenvalues $\lambda = \alpha$. The only point in the spectrum is $\lambda = \alpha$, which has infinite multiplicity assuming H is infinite dimensional. Therefore, the resolvent set is $\rho(A) = \mathbf{C} - \{\alpha\}$. In particular, for $\alpha = 1$ we have the identity operator, and for $\alpha = 0$ we have the zero operator.

2. For the matrix operator $\begin{bmatrix} -5 & 0 \\ 1 & 2 \end{bmatrix}$, eigenvalues and corresponding

eigenvectors are determined from (4.6) and (4.2) as

$$\begin{aligned}\lambda = 2, \quad x &= \begin{bmatrix} 0 \\ \alpha \end{bmatrix}, \\ \lambda = -5, \quad x &= \begin{bmatrix} -7\beta \\ \beta \end{bmatrix},\end{aligned}$$

where $\alpha, \beta \neq 0$, such that $\rho(A) = \mathbf{C} - \{2, -5\}$.

3. The multiplication operator $A : \mathbf{L}^2(a, b) \rightarrow \mathbf{L}^2(a, b)$ defined by

$$(Ax)(t) \equiv f(t)x(t)$$

is a linear bounded operator assuming $\max_t |f(t)| < \infty$.

First, consider the case of $f(t) = c$, where $c \in \mathbf{C}$ is a constant. Then

$$(A - \lambda I)^{-1} x(t) = (c - \lambda)^{-1} x(t)$$

such that the resolvent fails to exist for $\lambda = c$, which is an eigenvalue of A (any nonzero $x \in \mathbf{L}^2(a, b)$ is a corresponding eigenvector). For all other values of λ the resolvent exists and is bounded, and therefore the continuous spectrum is empty and the resolvent set is $\rho(A) = \mathbf{C} - \{c\}$.

Next, consider the case where $f(t) = t$. Then, by considering

$$(A - \lambda I)^{-1} x(t) = (t - \lambda)^{-1} x(t),$$

it is clear that the resolvent generally exists but is unbounded if $a \leq \lambda \leq b$. Therefore, every $\lambda \in [a, b]$ is in the continuous spectrum. For all other values of λ the resolvent exists and is bounded, and so the point spectrum is empty and every value $\lambda \notin [a, b]$ is in the resolvent set (note that although $(t - \lambda)\delta(t - \lambda) = 0$, $\delta(t - \lambda) \notin \mathbf{L}^2$ and so $\delta(t - \lambda)$ is not an eigenvector).

4. The differential operator $A : \mathbf{L}^2(0, \infty) \rightarrow \mathbf{L}^2(0, \infty)$ defined by⁷

$$\begin{aligned}(Ax)(t) &\equiv \frac{d}{dt}x(t), \\ D_A &\equiv \left\{ x : x, \frac{d}{dt}x(t) \in \mathbf{L}^2(0, \infty) \right\}\end{aligned}$$

is a linear unbounded operator. Consider the equation

$$(A - \lambda I)x = (x' - \lambda x) = -y.$$

For $y = 0$ and $\operatorname{Re}(\lambda) < 0$, $(A - \lambda I)e^{\lambda t} = 0$ so that $e^{\lambda t}$ is an eigenfunction of A with eigenvalue λ (with $\operatorname{Re}(\lambda) < 0$; for $\operatorname{Re}(\lambda) >$

⁷Strictly speaking, x must be absolutely continuous on $(0, \infty)$.

0, $e^{\lambda t} \notin \mathbf{L}^2(0, \infty)$ and so in this case $e^{\lambda t}$ is not an eigenfunction). Therefore, $\sigma_p(A)$ is that part of the complex λ -plane with $\operatorname{Re}(\lambda) < 0$ (the left half-plane).⁸

Next, consider the solution of $(A - \lambda I)x = (x' - \lambda x) = -y$ for $y \neq 0$. The formal solution is⁹

$$x(t) = \int_t^\infty e^{\lambda(t-\tau)} y(\tau) d\tau$$

and so the resolvent operator is formally identified by

$$R_\lambda(A)y = \int_t^\infty e^{\lambda(t-\tau)} y(\tau) d\tau.$$

This operator exists and is continuous for all λ such that $\operatorname{Re}(\lambda) > 0$, and so this part of the complex λ -plane (the right half-plane) comprises the resolvent set.

Finally, considering the case $\lambda = 0$ ($x(t) = \int_t^\infty y(\tau) d\tau$) it can be seen that the resolvent exists but is not continuous (i.e., not bounded; compare with Example 12, Section 3.1.2), and so $\lambda = 0$ is in the continuous spectrum. This also remains true for $\operatorname{Re}(\lambda) = 0$, so $\sigma_c(A)$ is the imaginary axis of the complex λ -plane. Because $\sigma(A) = \mathbf{C} = \rho(A) \cup \sigma_r(A) \cup \sigma_p(A) \cup \sigma_c(A)$, we see that the residual spectrum is empty, i.e., $\sigma_r(A) = \emptyset$. Figure 4.2 depicts the sets $\sigma(A)$ and $\rho(A)$ in the complex λ -plane.

5. Consider the differential operator $A : \mathbf{L}^2(0, a) \rightarrow \mathbf{L}^2(0, a)$ defined by

$$(Ax)(t) \equiv -\frac{d^2}{dt^2}x(t),$$

$$D_A \equiv \left\{ x : x, \frac{d^2}{dt^2}x \in \mathbf{L}^2(0, a), x(0) = x(a) = 0 \right\}.$$

Eigenfunctions are found to be $\sin \sqrt{\lambda_n}t$, with corresponding eigenvalues $\lambda_n = (n\pi/a)^2$, $n = 1, 2, \dots$ (defining a branch of the square root $\sqrt{\lambda_n}$ leads to, for example, only positive integers n). The point spectrum, which in this case is discrete, is $\sigma_p(A) = \{\lambda_n\}$, the residual and continuous spectrum are empty, and the resolvent set is $\rho(A) = \mathbf{C} \setminus \{\lambda_n\}$. This is depicted in Figure 4.3.

⁸Note that although the point spectrum is often discrete, this example shows that this need not be the case.

⁹Indeed, differentiating x we obtain

$$\frac{d}{dt}x(t) = \int_t^\infty \frac{d}{dt}e^{\lambda(t-\tau)}y(\tau) d\tau - y(t) = \lambda \int_t^\infty e^{\lambda(t-\tau)}y(\tau) d\tau - y(t) = \lambda x(t) - y(t).$$

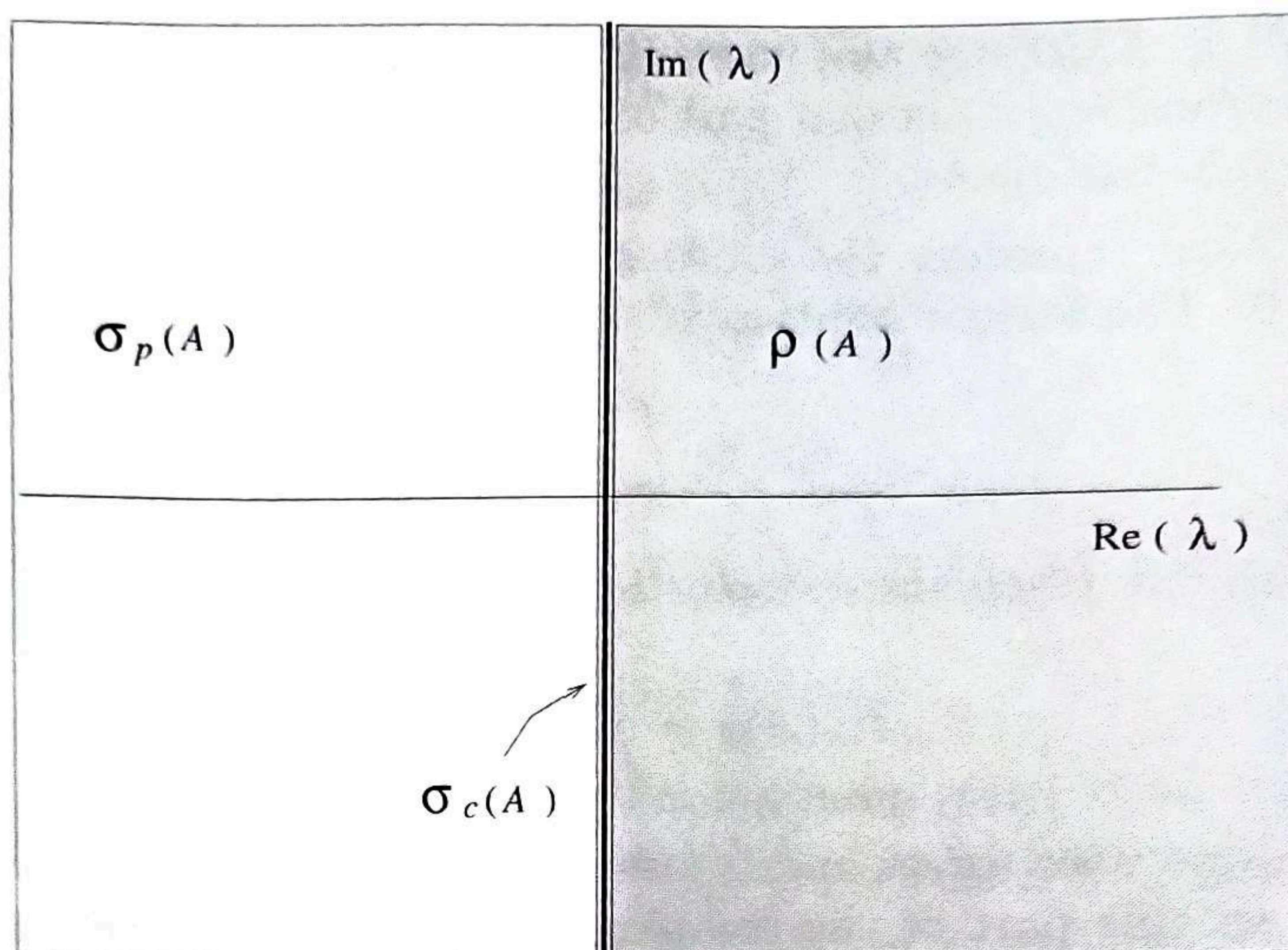


Figure 4.2: Complex λ -plane depicting the sets $\sigma(A)$ and $\rho(A)$ for the differential operator in Example 4. The point spectrum $\sigma_p(A)$ is the left half-plane, the resolvent set $\rho(A)$ is the right half-plane, and the continuous spectrum $\sigma_c(A)$ is the imaginary axis.

It should be noted that eigenfunctions of differential operators never satisfy nonhomogeneous boundary conditions, since this generally wouldn't allow for the determination of λ ; they satisfy either homogeneous conditions, periodic conditions, or some sort of fitness conditions as described in Chapter 5.

4.2 Spectral Properties of Operators

In this section we present some spectral properties important in both theory and applications. We are particularly interested in linear independence, orthogonality, and especially completeness of the eigenvectors of a given operator class in a particular space. Even for general linear operators some strong statements can be made, although the best situations arise when the operators belong to some appropriately restricted class.

4.2.1 General Properties

It is very important to know when eigenvectors of an operator form a linearly independent set. The next theorem addresses this issue.

Theorem 4.3. *For a linear operator $A : H \rightarrow H$, eigenvectors $x_1, x_2, \dots, x_n$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ form a linearly independent set in H .*

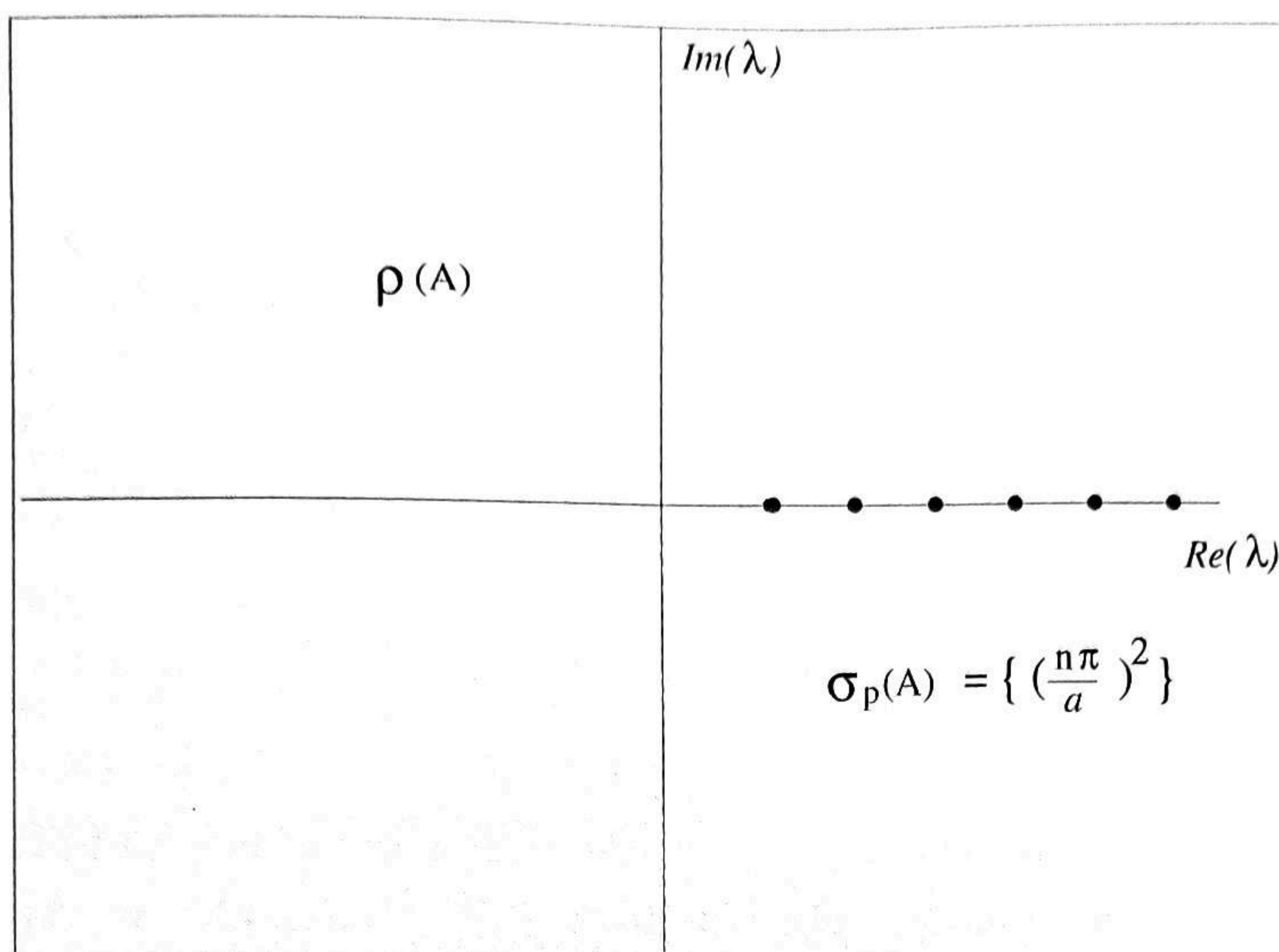


Figure 4.3: Complex λ -plane depicting the sets $\sigma(A)$ and $\rho(A)$ for the differential operator in Example 5. The point spectrum $\sigma_p(A) = \sigma(A)$ is the discrete set of points $\lambda_n = (n\pi/a)^2$, $n = 1, 2, \dots$, and the resolvent set $\rho(A)$ is everything else.

Proof. Assume that $\{x_1, x_2, \dots, x_n\}$ forms a linearly dependent set and that x_m is the first element that can be written as a linear combination of the preceding (and thus linearly independent) elements $\{x_1, x_2, \dots, x_{m-1}\}$ so that $x_m = \sum_{i=1}^{m-1} \alpha_i x_i$. Applying the operator $(A - \lambda_m I)$ to this sum we get

$$(A - \lambda_m I) x_m = 0 = (A - \lambda_m I) \sum_{i=1}^{m-1} \alpha_i x_i = \sum_{i=1}^{m-1} \alpha_i (\lambda_i - \lambda_m) x_i.$$

Because the elements on the right form a linearly independent set that sums to zero with coefficients $\alpha_i (\lambda_i - \lambda_m)$, then these coefficients must all vanish. We then have $\alpha_i (\lambda_i - \lambda_m) = 0$, but $\lambda_i - \lambda_m \neq 0$ by assumption and so $\alpha_i = 0$, $i = 1, \dots, m-1$. But then $x_m = 0$, which contradicts the fact that x_m is an eigenvector. Therefore, we have shown that $\{x_1, x_2, \dots, x_n\}$ forms a linearly independent set. ■

A linearly independent set can be converted to an orthonormal set by the *Gram-Schmidt procedure*, as detailed in Appendix C.

Theorem 4.4. If a linear operator $A : H \rightarrow H$ has eigenvalues λ_n and eigenvectors x_n , i.e., $Ax_n = \lambda_n x_n$, then $\lambda_n - \alpha$ and x_n are eigenvalues and eigenvectors, respectively, of the operator $B : H \rightarrow H$ defined as $B \equiv A - \alpha I$.

Proof. Let the eigenvalues and eigenvectors of A be λ_n and x_n , respectively, and let the eigenvalues and eigenvectors of $(A - \alpha I)$ be γ_n and y_n , respectively. Then $(A - \alpha I)y_n = \gamma_n y_n$ and $Ay_n - \alpha y_n = \gamma_n y_n$, such that $Ay_n = (\gamma_n + \alpha)y_n$, which shows that y_n is an eigenvector of A with eigenvalue $\gamma_n + \alpha$. Therefore, $y_n = x_n$ and $\gamma_n + \alpha = \lambda_n$, or $\gamma_n = \lambda_n - \alpha$. ■

The next theorem shows that parts of the spectrum of an operator are often related to (perhaps different) parts of the spectrum of the adjoint operator.

Theorem 4.5. *If λ is in the residual spectrum of a linear operator $A : H \rightarrow H$ with deficiency m , then $\bar{\lambda}$ is an eigenvalue of A^* of multiplicity m .*

That is, if $\lambda \in \sigma_r(A)$, then $\bar{\lambda} \in \sigma_p(A^*)$.

Proof. For $\lambda \in \sigma_r(A)$, $R_{A-\lambda I}$ is not dense in H (by definition), and so $(R_{A-\lambda I})^\perp$ is not empty. By Theorem 3.7, $(R_{A-\lambda I})^\perp = N_{(A-\lambda I)^*} = N_{A^*-\bar{\lambda}I}$, and so $N_{A^*-\bar{\lambda}I}$ is not empty; hence $\bar{\lambda} \in \sigma_p(A^*)$. The multiplicity argument is fairly obvious. ■

More generally, one can show the following [3, p. 228].

Theorem 4.6. *For any bounded linear operator $A : H \rightarrow H$,*

$$\sigma(A^*) = \bar{\sigma}(A). \quad (4.7)$$

Proof. If $\lambda \in \rho(A)$, then $R_A(\lambda)$ is bounded, with the range of the invertible operator $(A - \lambda I)$ dense in H . The operator $(A - \lambda I)$ then satisfies the conditions of Theorem 3.34, such that $(A - \lambda I)^* = (A^* - \bar{\lambda}I)$ has range dense in H and is invertible with bounded inverse. Therefore, $\bar{\lambda} \in \rho(A^*)$, and we have $\sigma(A^*) \subseteq \bar{\sigma}(A)$. Reversing the process by starting with $\lambda \in \rho(A^*)$ leads to $\sigma(A) \subseteq \bar{\sigma}(A^*)$ and hence $\sigma(A^*) = \bar{\sigma}(A)$. ■

It is important to note that this result applies to the entire spectrum; spectral elements σ_p , σ_r , σ_c do not necessarily individually satisfy (4.7). In fact, the presence of a nonempty residual spectrum complicates relating eigenvalues of A to those of A^* . It turns out, however, that the point spectrum of an operator is often related to the point spectrum of its adjoint, as the following theorem states (see also Theorem 4.24 and [2, p. 90]).

Theorem 4.7. *Let $A, A^* : H \rightarrow H$ have empty residual spectra, and let A have eigenvalues λ_n and corresponding eigenvectors x_n (i.e., $Ax_n = \lambda_n x_n$). Then $\bar{\lambda}_n$ are eigenvalues of the adjoint operator $A^* : H \rightarrow H$ corresponding to adjoint eigenvectors y_n (i.e., $A^*y_n = \bar{\lambda}_n y_n$), and*

$$(\lambda_n - \lambda_m) \langle x_n, y_m \rangle = 0. \quad (4.8)$$

Proof. Assume that λ_n, x_n are eigenvalues and eigenvectors, respectively, of $A : H \rightarrow H$ and y an arbitrary function in the domain of A^* . Then $(A - \lambda_n I)x_n = 0$ and $\langle y, (A - \lambda_n I)x_n \rangle = 0$ and, upon noting

$\langle y, Ax \rangle = \langle A^*y, x \rangle$ we have $\langle (A - \lambda_n I)^* y, x_n \rangle = \langle (A^* - \bar{\lambda}_n I) y, x_n \rangle = 0$. Thus, either $(A^* - \bar{\lambda}_n I) y = 0$, indicating y is an eigenvector of A^* with eigenvalue $\bar{\lambda}_n$, or $(A^* - \bar{\lambda}_n I) y = f \neq 0$. If the latter is true and y is arbitrary, then $\langle f, x_n \rangle = 0$ for f arbitrary, implying $x_n = 0$. Because this cannot occur, the former statement must be true.¹⁰

To prove the relationship among the eigenvectors of A and A^* , start with $Ax_n = \lambda_n x_n$ and $A^*y_m = \bar{\lambda}_m y_m$ (under suitable ordering of the eigenvalues). Taking inner products we have $\langle Ax_n, y_m \rangle = \langle \lambda_n x_n, y_m \rangle = \lambda_n \langle x_n, y_m \rangle$ and $\langle A^*y_m, x_n \rangle = \langle \bar{\lambda}_m y_m, x_n \rangle = \bar{\lambda}_m \langle y_m, x_n \rangle$, where the latter becomes $\langle x_n, A^*y_m \rangle = \lambda_m \langle x_n, y_m \rangle$. Forming $0 = \langle Ax_n, y_m \rangle - \langle x_n, A^*y_m \rangle = \lambda_n \langle x_n, y_m \rangle - \lambda_m \langle x_n, y_m \rangle$ leads to $(\lambda_n - \lambda_m) \langle x_n, y_m \rangle = 0$, or $\langle x_n, y_m \rangle = 0$ for $\lambda_n \neq \lambda_m$. ■

Note that (4.8) can be interpreted as

$$N_{A-\lambda_n I} \perp N_{A-\lambda_m I}$$

for $\lambda_n \neq \lambda_m$. In particular, for a self-adjoint operator the set of eigenvectors (if any even exist) corresponding to distinct eigenvalues forms an orthogonal set. For nonself-adjoint operators, the set of eigenvectors (if any even exist) and adjoint eigenvectors corresponding to distinct eigenvalues forms a bi-orthogonal set.

Rayleigh Quotient

Assuming $x_1, x_2, \dots, x_n$ are orthogonal eigenvectors corresponding to different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of $A : H \rightarrow H$, one can obtain the eigenvalues using the *Rayleigh quotient* representation,

$$\lambda_n = \frac{\langle Ax_n, x_n \rangle}{\langle x_n, x_n \rangle}. \quad (4.9)$$

Of course, for (4.9) to be directly useful for calculation of the eigenvalue λ_n one must know the eigenvector x_n . However, if we consider the functional

$$F(x) \equiv \frac{\langle Ax, x \rangle}{\langle x, x \rangle}, \quad (4.10)$$

then it can be shown that for many types of operators (not necessarily bounded) the functional F is stationary at x_n , in the sense that first-order errors in approximating the element x_n lead to second-order errors in calculating the eigenvalue λ_n . That is, if $x = x_n + O(\varepsilon)$, then $F(x) = \lambda_n + O(\varepsilon^2)$.

¹⁰Lack of a residual spectrum is important for the following reason: By definition, $f \in R_{A^* - \bar{\lambda}_n I}$. By excluding the possibility of a residual spectrum the range of $A^* - \bar{\lambda}_n I$ is dense in H , assuming $\bar{\lambda}_n$ is not an eigenvalue of A^* . The conclusion that $\langle f, x_n \rangle = 0$ for f arbitrary implies $x_n = 0$ is only valid if f is an arbitrary element of H , or of a space dense in H .

If $F(x)$ is stationary at all x_n , then all eigenvalues λ_n can be approximated (to first order) using appropriate trial functions for x_n . This is a common method in applications. Furthermore, if A is self-adjoint with a discrete spectrum, then various eigenvalues can be obtained from (4.10) by minimizing $F(x)$. For instance, the smallest eigenvalue of A is equal to $\min_{x \in D_A} F(x)$, with the minimum attained when $x = x_n$ [2, p. 209]. Also, if A is bounded, the largest eigenvalue can be obtained as $\max_{x \in D_A} F(x)$. Often one can go further and develop so-called minimax theorems for obtaining all eigenvalues λ_n as minimizations of $F(x)$ over specially chosen sets $\{x\}$ within D_A . See, e.g., [4, pp. 419–421] for a detailed discussion.

Example 4.2.

Consider a source-free cavity occupying a region Ω with perfectly conducting boundary surface Γ and filled with a homogeneous medium characterized by μ, ε . The electromagnetic field satisfies

$$\begin{aligned}\nabla \times \nabla \times \mathbf{E} &= k^2 \mathbf{E}, \\ \nabla \cdot \mathbf{E} &= 0, \\ \mathbf{n} \times \mathbf{E}|_{\Gamma} &= \mathbf{0},\end{aligned}$$

where $k^2 = k_n^2 = \omega_n^2 \mu \varepsilon > 0$ represents real-valued eigenvalues of the operator $\nabla \times \nabla \times$ corresponding to eigenfunctions $\mathbf{E} = \mathbf{E}_n$. Then, the functional

$$\omega_n^2(\mathbf{E}) = \frac{1}{\mu \varepsilon} \frac{\langle \nabla \times \nabla \times \mathbf{E}, \mathbf{E} \rangle}{\langle \mathbf{E}, \mathbf{E} \rangle} = \frac{1}{\mu \varepsilon} \frac{\int_{\Omega} \nabla \times \nabla \times \mathbf{E} \cdot \mathbf{E} d\Omega}{\int_{\Omega} \mathbf{E} \cdot \mathbf{E} d\Omega} \quad (4.11)$$

is identically satisfied for $\mathbf{E} = \mathbf{E}_n$ and stationary [5, p. 332] for all “trial” eigenfunctions $\mathbf{E}_n^t = \mathbf{E}_n + \delta \mathbf{E}_n$ such that $\mathbf{n} \times \mathbf{E}_n^t|_{\Gamma} = \mathbf{0}$.

4.2.2 Bounded Operators

If a linear operator is bounded, then it turns out that its spectrum is bounded as well.¹¹

Theorem 4.8. *Let $A : H \rightarrow H$ be a bounded linear operator. The spectrum $\sigma(A)$ is a compact (therefore, necessarily closed and bounded) subset of the complex plane lying within the closed region $|z| \leq \|A\|$.*

It can be easily seen that if $A : H \rightarrow H$ is a bounded linear operator and $|z| > \|A\|$, then z is in the resolvent set of A . Therefore, for a bounded linear operator the resolvent set is never empty.

A more restrictive theorem relating to the eigenvalues themselves is easier to prove.

¹¹See Theorem 4.1 for an explicit formula pertaining to operators on finite-dimensional spaces.

Theorem 4.9. *If λ is an eigenvalue of a linear bounded operator $A : H \rightarrow H$, then $|\lambda| \leq \|A\|$.*

Proof. Starting with $Ax = \lambda x$, we obtain $\|Ax\| = \|\lambda x\| = |\lambda| \|x\|$. However, $\|Ax\| \leq \|A\| \|x\|$ and so $|\lambda| \|x\| \leq \|A\| \|x\|$, implying $|\lambda| \leq \|A\|$. ■

The concept of a *spectral radius* is motivated by the preceding.

Definition 4.2. *The spectral radius of an operator, denoted as $r_\sigma(A)$, is the radius of the smallest closed disk, centered at the origin of the complex plane, containing $\sigma(A)$. The spectral radius is given by*

$$r_\sigma(A) = \sup_{\lambda \in \sigma(A)} |\lambda|.$$

The next theorem shows the connection between bounded operators and their eigenvalues.

Theorem 4.10. *A linear operator $A : H \rightarrow H$ such that $\sigma_p(A) = \sigma(A)$ is continuous (bounded) if and only if the set $\{|\lambda_1|, |\lambda_2|, \dots\}$ is bounded. Furthermore, $\|A\| \geq r_\sigma(A)$.*

Proof. Assume $A : H \rightarrow H$ is bounded. Then $\{|\lambda_n|\}$ is bounded by Theorem 4.9. Now assume $\{|\lambda_n|\}$ is a bounded set. Because $\|Ax\| = \|\lambda x\| = |\lambda| \|x\|$ and $|\lambda|$ is bounded, then by Definition 3.6, A is bounded. ■

4.2.3 Invertible Operators

An important observation can be made about the spectral properties of invertible operators.

Theorem 4.11. *Let $A : H \rightarrow H$ be an invertible linear operator with eigenvalues λ and corresponding eigenvectors x (recall that if A is invertible, $N_A = \{0\}$ from Theorem 3.24, and so $\lambda \neq 0$). Then, $A^{-1} : H \rightarrow H$ has eigenvalues $1/\lambda$ and corresponding eigenvectors x .*

Proof. It is given that we have $Ax = \lambda x$ with $\lambda, x \neq 0$. Because A^{-1} exists we have $A^{-1}Ax = A^{-1}\lambda x$ such that $x = \lambda A^{-1}x$, or $A^{-1}x = (1/\lambda)x$. ■

Furthermore, if A and A^{-1} are bounded, then

$$\sigma(A^{-1}) = \sigma^{-1}(A).$$

The following theorem shows a relationship pertaining to similarity transformations, which will be discussed in Section 4.3.1 in conjunction with the theory of matrix representations of operators.

Theorem 4.12. *Let $A : H \rightarrow H$ be a linear operator and let the linear operator $T : H \rightarrow H$ be invertible. The operators A and TAT^{-1} (which are called similar) have the same eigenvalues.*

Proof. Let λ be an eigenvalue of A with eigenvector x . Then we have $TAT^{-1}(Tx) = TAx = T\lambda x = \lambda(Tx)$, showing that λ is an eigenvalue of TAT^{-1} . If λ is an eigenvalue of TAT^{-1} with eigenvector y , then $TAT^{-1}y = \lambda y$, and so $AT^{-1}y = \lambda T^{-1}y$, showing that λ is an eigenvalue of A . ■

4.2.4 Compact Operators

Compact operators need not have any eigenvalues (as a common example the integral operator $A : L^2(0, 1) \rightarrow L^2(0, 1)$ defined by $(Af)(t) \equiv \int_0^t f(\tau) d\tau$ is Hilbert–Schmidt, and hence compact, but does not possess any eigenvalues¹²). However, if eigenvalues do exist, then the following theorems are useful.

Theorem 4.13. *Every spectral value $\lambda \neq 0$ of a compact linear operator $A : H \rightarrow H$ is an eigenvalue.*

Proof. If $R_\lambda(A) = (A - \lambda I)^{-1}$ does not exist, then $\lambda \in \sigma_p(A)$ by definition. Now, let $\lambda \neq 0$ and assume that $R_\lambda(A)$ exists. Then $(A - \lambda I)x = 0$ implies $x = 0$, and so the Fredholm alternative 3.44 shows that $(A - \lambda I)x = y$ is uniquely solvable for all $y \in H$. Therefore, $R_\lambda(A)$ is defined on all $y \in H$ and is bounded by the open mapping theorem (Theorem 3.27), and $\lambda \in \rho(A)$. ■

Therefore, the continuous spectrum and residual spectrum of a compact operator either are empty or are the zero element.

Theorem 4.14. *Let $A : H \rightarrow H$ be a compact linear operator. Then, the null space of $A - \lambda I$, $\lambda \neq 0$, is finite dimensional.*

Theorem 4.15. *Let $A : H \rightarrow H$ be a compact linear operator. Then, for $\alpha > 0$ the number of eigenvalues λ such that $|\lambda| \geq \alpha$ is finite.*

4.2.5 Self-Adjoint Operators

Self-adjoint operators may not have any eigenvalues (recall the multiplication operator example, p. 226), but if they do possess eigenvalues, then the eigenvalues and corresponding eigenvectors have some very desirable properties.

¹²To see this consider $\lambda = 0$ as a possible eigenvalue, leading to $Ax = 0x = 0$. For the operator in question this indicates $x = 0$. For some $\lambda \neq 0$ as a possible eigenvalue, the equation $Ax = \lambda x$ implies $x(0) = 0$ and leads to $x = \lambda x'$, the only solution of which is $x = 0$.

Theorem 4.16. *Let $A : H \rightarrow H$ be a self-adjoint linear operator. Then*

- a. the eigenvalues of A are real, and*
- b. the eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof. (a) Let x be an eigenvector corresponding to eigenvalue λ , satisfying $Ax = \lambda x$. Then $\langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle$. Also, since A is self-adjoint, $\langle Ax, x \rangle = \langle x, Ax \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle$. Therefore, since $\langle x, x \rangle > 0$, $\lambda = \bar{\lambda}$, proving¹³ $\lambda \in \mathbf{R}$. (b) Let x_1, x_2 be eigenvectors corresponding to distinct eigenvalues λ_1, λ_2 of the self-adjoint operator A , so that we have $Ax_1 = \lambda_1 x_1$ and $Ax_2 = \lambda_2 x_2$. Then $\lambda_1 \langle x_1, x_2 \rangle = \langle Ax_1, x_2 \rangle = \langle x_1, Ax_2 \rangle = \langle x_1, \lambda_2 x_2 \rangle = \bar{\lambda}_2 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle$, where the last equality comes about since eigenvalues are real. We have $(\lambda_1 - \lambda_2) \langle x_1, x_2 \rangle = 0$, and, since $\lambda_1 \neq \lambda_2$, then $\langle x_1, x_2 \rangle = 0$ as desired. ■

Unfortunately, all eigenvectors of a self-adjoint operator are not necessarily mutually orthogonal, only those corresponding to distinct eigenvalues (although within each eigenspace we can orthogonalize the set). If an operator is real and self-adjoint, then, if eigenvectors exist, they either are real-valued or can be made so,¹⁴ although in the latter case orthogonality will generally be lost.

Furthermore, beyond considerations of the eigenvalues themselves, we have the following.

Theorem 4.17. *The spectrum $\sigma(A)$ of a bounded self-adjoint linear operator $A : H \rightarrow H$ is real and is a subset of the interval $[-\|A\|, \|A\|]$. Furthermore, for $\lambda \in \rho(A)$,*

$$\|R_\lambda(A)\| \leq \frac{1}{|\operatorname{Im} \lambda|}.$$

The proof is fairly simple for the eigenvalues themselves, but more difficult for the entire spectrum.

The above indicates that for a self-adjoint operator, every λ with $\operatorname{Im} \lambda \neq 0$ is in the resolvent set. We also have the following theorem regarding the residual spectrum.

Theorem 4.18. *The residual spectrum $\sigma_r(A)$ of a self-adjoint linear operator $A : H \rightarrow H$ is empty.*

¹³Obviously this is also true if A is merely symmetric. Note also that the converse is also true, i.e., that if an eigenvalue λ is real-valued, then A is at least symmetric. To see this assume $\lambda \in \mathbf{R}$, and note that $\langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle = \bar{\lambda} \langle x, x \rangle = \langle x, \lambda x \rangle = \langle x, Ax \rangle$.

¹⁴Let $A : H \rightarrow H$ be real and self-adjoint, with x being an eigenvector of A with corresponding (real) eigenvalue λ . Then $(A - \lambda I)(x - \bar{x}) = (A - \lambda I)\bar{x} = \overline{(A - \lambda I)x} = \overline{(A - \lambda I)x} = 0$, and therefore $(x - \bar{x})$ is a (real-valued) eigenvector of A .

The proof follows as a special case of Theorem 4.26 presented later. Note that, at most, the spectrum of a self-adjoint linear operator will contain a discrete and a continuous part, and both will be bounded on the real axis if the operator is bounded.

For unbounded linear operators, where the distinction between self-adjointness and symmetry is important, we have similar results for operators that are merely symmetric, namely that the point spectrum and continuous spectrum are real (the residual spectrum, if it exists, need not be real) and that eigenvectors corresponding to distinct eigenvalues are orthogonal.

4.2.6 Nonnegative and Positive Operators

The eigenvalues of nonnegative and positive operators (recall that these are special cases of self-adjoint, or at least symmetric, operators) have the expected spectral properties detailed in the next theorem.

Theorem 4.19. *All eigenvalues of a nonnegative linear operator are nonnegative. All eigenvalues of a positive linear operator are positive.*

Proof. Consider $Ax = \lambda x$, with A nonnegative and $x \neq 0$. Then $0 \leq \langle Ax, x \rangle = \lambda \langle x, x \rangle = \lambda \|x\|^2$, which shows that $\lambda \geq 0$ as required. Repeating for $\langle Ax, x \rangle > 0$ leads to $\lambda > 0$. ■

Similar statements can be made about nonpositive and negative operators and about the entire spectrum $\sigma(A)$. For a bounded dissipative operator ($\operatorname{Im} \langle Ax, x \rangle \geq 0$) we have the following.

Theorem 4.20. *The eigenvalues of a bounded dissipative linear operator lie in the half-plane $\operatorname{Im} \lambda \geq 0$.*

Proof. Consider $Ax = \lambda x$, with A dissipative and $x \neq 0$. Then $0 \leq \operatorname{Im} \langle Ax, x \rangle = \operatorname{Im} (\lambda \langle x, x \rangle) = \operatorname{Im} (\lambda) \|x\|^2$, which shows that $\operatorname{Im} \lambda \geq 0$ as required. ■

Furthermore, it can be shown that $\sigma(A)$ lies in the half-plane $\operatorname{Im} \lambda \geq 0$.

4.2.7 Compact Self-Adjoint Operators

It was shown previously that compact operators and self-adjoint operators may not possess any eigenvalues; however, if eigenvalues do exist, then they have some convenient mathematical properties. When one considers the intersection of these two classes of operators, namely compact self-adjoint operators, the situation changes quite a bit as the following theorems demonstrate. This will also be seen in later sections on spectral expansions.

The first thing to state about compact self-adjoint operators is that they do possess an eigenvalue, at least one.

Theorem 4.21. *Let $A : H \rightarrow H$ be a compact self-adjoint linear operator on a nontrivial Hilbert space. Then A has an eigenvalue λ with $|\lambda| = \|A\|$.*

Therefore, compact self-adjoint operators have at least one eigenvalue. This fact is also true for compact normal operators.

Theorem 4.22. *Let $\{\lambda_n\}$ be the set of distinct nonzero eigenvalues of a compact, self-adjoint linear operator. Then, either $\{\lambda_n\}$ is a finite set, or $\lim_{n \rightarrow \infty} \lambda_n = 0$.*

Proof. Assume A is a compact self-adjoint operator with infinitely many distinct eigenvalues λ_n , and corresponding eigenvectors x_n , where the set $\{x_n\}$ is orthogonal simply by the distinctness of the eigenvalues since A is self-adjoint. Assume further that the eigenvectors are normalized to become an orthonormal set. By the discussion in Section 2.5.6, orthonormal sequences are weakly convergent to 0, and from Section 3.6.1 we have $0 = \lim_{n \rightarrow \infty} \|Ax_n\|^2 = \lim_{n \rightarrow \infty} \langle Ax_n, Ax_n \rangle = \lim_{n \rightarrow \infty} \langle \lambda_n x_n, \lambda_n x_n \rangle = \lim_{n \rightarrow \infty} \lambda_n^2 \|x_n\|^2 = \lim_{n \rightarrow \infty} \lambda_n^2$. ■

Note that the proof relies on compactness to convert the weakly convergent sequence into a strongly convergent sequence. Theorem 4.22 also follows from Theorems 4.14 and 4.15.

The last theorem indicates a potential problem with compact self-adjoint operators, namely that their eigenvalues λ_n tend toward the origin in the complex plane as $n \rightarrow \infty$. In later sections, when expansions of operators and operator inverses are discussed, we will see that in this case the fact that the inverse of a compact operator on an infinite-dimensional space is unbounded (as discussed in the previous chapter) is analogous to division by an eigenvalue tending toward the origin.

Similar to Theorem 4.13 we have the following theorem [6, p. 24].

Theorem 4.23. *Every spectral value $\lambda \neq 0$ of a compact self-adjoint linear operator $A : H \rightarrow H$ is an eigenvalue of finite multiplicity that can only accumulate at $\lambda = 0$. Conversely, a self-adjoint operator having these properties is compact.*

4.2.8 Normal and Unitary Operators

Concerning the spectral properties of normal and unitary operators, we have the following.¹⁵

The first theorem provides a relationship between eigenvalues and eigenvectors of a normal operator and those of its adjoint.

¹⁵Recall that because self-adjoint operators are always normal, the following theorems apply to self-adjoint operators as well.

Theorem 4.24. *Let $A : H \rightarrow H$ be a bounded normal linear operator. If λ is an eigenvalue of A with $x \in H$ its corresponding eigenvector, then $\bar{\lambda}$ is an eigenvalue of A^* and the corresponding eigenvector of A^* is x . Furthermore, $N_{A-\lambda I} = N_{A^*-\bar{\lambda}I}$.*

Proof. If A is normal, then $A - \lambda I$ is also normal. From Theorem 3.13 we see that $\|(A - \lambda I)x\| = 0$ if and only if $\|(A^* - \bar{\lambda}I)x\| = 0$. From Definition 2.31 this indicates that $(A^* - \bar{\lambda}I)x = 0$. ■

The next theorem extends the orthogonality of eigenvectors corresponding to distinct eigenvalues in the self-adjoint case (Theorem 4.16) to the case of normal operators.

Theorem 4.25. *Let $A : H \rightarrow H$ be a normal linear operator. Then $N_{A-\lambda_n I} \perp N_{A-\lambda_m I}$ for $\lambda_n \neq \lambda_m$.*

Proof. Let $x_n \in N_{A-\lambda_n I}$ and $x_m \in N_{A-\lambda_m I}$, $x_n, x_m \neq 0$. From $\langle Ax_n, x_m \rangle = \langle x_n, A^*x_m \rangle$, we obtain $\langle \lambda_n x_n, x_m \rangle = \langle x_n, \bar{\lambda}_m x_m \rangle$. Then $(\lambda_n - \bar{\lambda}_m) \langle x_n, x_m \rangle = 0$, which implies $\langle x_n, x_m \rangle = 0$ since $\lambda_n - \bar{\lambda}_m \neq 0$. ■

The next theorem is worthwhile to note since it makes any consideration of the residual spectrum for normal operators unnecessary.

Theorem 4.26. *The residual spectrum of a normal linear operator is empty.*

Proof. From Section 4.1.2 (see p. 223) it is enough to show that if $A - \lambda I$ is invertible, then $R_{A-\lambda I}$ is dense in H . To show this, assume $A - \lambda I$ is invertible (and so one-to-one) and let y be orthogonal to $R_{A-\lambda I}$. We then have $0 = \langle (A - \lambda I)x, y \rangle = \langle x, (A^* - \bar{\lambda}I)y \rangle$ for all $x \in H$. It then follows from Definition 2.41 that $(A^* - \bar{\lambda}I)y = 0$, and so $y \in N_{A^*-\bar{\lambda}I}$. From Theorem 4.24 it follows that $y \in N_{A-\lambda I}$. Because $A - \lambda I$ is one-to-one, then $N_{A-\lambda I} = \{0\}$, and therefore $y = 0$. We then have $R_{A-\lambda I}^\perp = \{0\}$, and by Theorem 2.7(c) $R_{A-\lambda I}$ is dense in H . ■

Furthermore, for typical differential operators the residual spectrum is empty, but this is not a general result.

The next theorem provides a nice geometrical picture of the eigenvalues of unitary operators.

Theorem 4.27. *All eigenvalues of a unitary linear operator $A : H \rightarrow H$ are complex numbers of modulus 1. Furthermore, $\sigma(A)$ lies on the unit circle $|\lambda| = 1$. For $|\lambda| \neq 1$,*

$$\|R_\lambda(A)\| \leq \frac{1}{|1 - |\lambda||}.$$

Proof. We prove only the first part of the theorem. Let A be unitary and let λ be an eigenvalue of A with corresponding eigenvector $x \in H$. Then $\langle Ax, Ax \rangle = \langle \lambda x, \lambda x \rangle = |\lambda|^2 \|x\|^2$. Also, $\langle Ax, Ax \rangle = \langle x, A^*Ax \rangle = \langle x, x \rangle = \|x\|^2$, leading to $|\lambda| = 1$. ■

4.2.9 Generalized Eigenvalue Problems

While attention is usually focused on the standard eigenvalue problem, generalized and nonstandard eigenvalue problems, and eigenvalue problems utilizing a pseudo inner product, are also of interest in electromagnetics. In this and the next two sections we briefly present some results concerning more general eigenvalue problems. Much of the material follows that presented in [7, Ch. 3] and [2].

Recall that the generalized eigenvalue problem is given as

$$Ax = \lambda Bx \quad (4.12)$$

or, equivalently,

$$(A - \lambda BI)x = 0. \quad (4.13)$$

In this case it is not enough for the operators A and B to be self-adjoint for the spectrum to be real.

Theorem 4.28. *Let $A, B : H \rightarrow H$ be self-adjoint (or symmetric) linear operators such that $\langle Ax, x \rangle \neq 0$ or $\langle Bx, x \rangle \neq 0$. Then eigenvalues corresponding to $Ax = \lambda Bx$ are real-valued.*

Proof. First assume that $\langle Bx, x \rangle \neq 0$. Then, given $Ax = \lambda Bx$ we have $\langle Ax, x \rangle = \lambda \langle Bx, x \rangle$ and $\langle x, Ax \rangle = \bar{\lambda} \langle x, Bx \rangle$. Since A and B are symmetric, then $\lambda \langle Bx, x \rangle = \bar{\lambda} \langle Bx, x \rangle$, or $(\lambda - \bar{\lambda}) \langle Bx, x \rangle = 0$. Therefore, $(\lambda - \bar{\lambda}) = 0$, which proves eigenvalues λ are real if $\langle Bx, x \rangle \neq 0$. The value of $\langle Ax, x \rangle$ is immaterial. Now assume that $\langle Ax, x \rangle \neq 0$. Then, given $Ax = \lambda Bx$ with $\lambda \neq 0$ we have $(1/\lambda) \langle Ax, x \rangle = \langle Bx, x \rangle$ and $(1/\bar{\lambda}) \langle x, Ax \rangle = \langle x, Bx \rangle$. Since A and B are symmetric, then $(1/\lambda) \langle Ax, x \rangle = (1/\bar{\lambda}) \langle Ax, x \rangle$, or $((1/\lambda) - (1/\bar{\lambda})) \langle Ax, x \rangle = 0$. Therefore, $((1/\lambda) - (1/\bar{\lambda})) = 0$, which proves that eigenvalues λ are real if $\langle Ax, x \rangle \neq 0$. The value of $\langle Bx, x \rangle$ is immaterial. ■

Therefore, if A and B are self-adjoint, or at least symmetric, and either A or B is definite (positive, positive definite, negative, or negative definite as defined in Section 3.5), then the eigenvalue problem (4.12) admits only real eigenvalues. For A and B self-adjoint or symmetric, complex eigenvalues can occur only if both A and B are indefinite.

As with the ordinary eigenvalue problem, one can investigate orthogonality of eigenvectors for the generalized eigenvalue problem.

Theorem 4.29. *Let $A, B : H \rightarrow H$ be self-adjoint (or symmetric) operators. Then eigenvectors corresponding to $Ax = \lambda Bx$ satisfy the orthogonality relationships*

$$\begin{aligned} (\lambda_n - \bar{\lambda}_m) \langle Bx_n, x_m \rangle &= 0, \\ \left(\frac{1}{\lambda_n} - \frac{1}{\bar{\lambda}_m} \right) \langle Ax_n, x_m \rangle &= 0 \end{aligned}$$

($\lambda_n, \lambda_m \neq 0$ in the last expression).

Proof. We have $\langle Ax_n, x_m \rangle = \langle \lambda_n Bx_n, x_m \rangle = \lambda_n \langle Bx_n, x_m \rangle = \langle x_n, Bx_m \rangle$ and $\langle Ax_m, x_n \rangle = \langle \lambda_m Bx_m, x_n \rangle = \lambda_m \langle Bx_m, x_n \rangle$. Using both expressions we obtain $\langle Ax_m, x_n \rangle = \lambda_m \langle Bx_m, x_n \rangle = \bar{\lambda}_m \langle x_n, Bx_m \rangle = \lambda_n \langle x_n, Bx_m \rangle$; therefore, $(\lambda_n - \bar{\lambda}_m) \langle x_n, Bx_m \rangle = 0$ and using self-adjointness of B we have the first result. A similar procedure leads to the second relationship. ■

If one of the operators A or B is definite, then $\lambda_n \in \mathbf{R}$ and the orthogonality relation holds for distinct eigenvalues $\lambda_n \neq \lambda_m$.

Example 4.3.

Consider Maxwell's equations applied to the interior of a source-free cavity Ω bounded by a perfectly conducting, closed surface Γ ,

$$\begin{aligned}\nabla \times \mathbf{E}(\mathbf{r}) &= -i\omega\mu\mathbf{H}(\mathbf{r}), \\ \nabla \times \mathbf{H}(\mathbf{r}) &= i\omega\varepsilon\mathbf{E}(\mathbf{r}), \\ \mathbf{n} \times \mathbf{E}|_{\Gamma} &= \mathbf{0}.\end{aligned}\tag{4.14}$$

This set of equations can be described in operator form as

$$-i \begin{pmatrix} 0 & \nabla \times \\ -\nabla \times & 0 \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \omega \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}$$

or

$$A \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \lambda B \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}$$

with the operators $A, B : \mathbf{L}^2(\Omega)^6 \rightarrow \mathbf{L}^2(\Omega)^6$ defined by

$$\begin{aligned}A &\equiv -i \begin{pmatrix} 0 & \nabla \times \\ -\nabla \times & 0 \end{pmatrix}, & B &\equiv \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix}, \\ D_A &\equiv \left\{ \mathbf{x} : \mathbf{x}, \nabla \times \mathbf{x} \in \mathbf{L}^2(\Omega)^6, \mathbf{n} \times \mathbf{x}|_{\Gamma} = \mathbf{0} \right\}, \\ D_B &\equiv \left\{ \mathbf{x} : \mathbf{x} \in \mathbf{L}^2(\Omega)^6 \right\}.\end{aligned}$$

The operators A and B can be seen to be self-adjoint, and since B is clearly positive definite for $\varepsilon, \mu > 0$, by Theorem 4.28 we see that the resonance frequencies $\lambda_n = \omega_n$ are real. Eigenfunctions $\mathbf{x}_n = \begin{pmatrix} \mathbf{E}_n \\ \mathbf{H}_n \end{pmatrix}$, corresponding to distinct eigenvalues ω_n , are orthogonal in the sense

$$(\omega_n - \omega_m) \langle B\mathbf{x}_n, \mathbf{x}_m \rangle = (\omega_n - \omega_m) \int_{\Omega} (\varepsilon \mathbf{E}_n \cdot \bar{\mathbf{E}}_m + \mu \mathbf{H}_n \cdot \bar{\mathbf{H}}_m) d\Omega = 0.$$

If A and/or B is not symmetric, then we have the following.

Theorem 4.30. *If x_n are eigenvectors of $Ax = \lambda Bx$, with corresponding eigenvalues λ_n , then $\bar{\lambda}_n$ are eigenvalues of the adjoint equation $A^*y = \lambda^* B^*y$ corresponding to adjoint eigenvectors y_n (i.e., $\lambda_n^* = \bar{\lambda}_n$), and*

$$\begin{aligned}(\lambda_n - \lambda_m) \langle Bx_n, y_m \rangle &= 0, \\ \left(\frac{1}{\lambda_n} - \frac{1}{\lambda_m} \right) \langle Ax_n, y_m \rangle &= 0\end{aligned}$$

($\lambda_n, \lambda_m \neq 0$ in the last expression).

As with Theorem 4.7, we assume that the residual spectrum is empty. The proof follows along the same lines as in the previous cases. Moreover, the result for eigenvalues holds for the entire spectrum,

$$\sigma(A, B) = \overline{\sigma}(A^*, B^*).$$

4.2.10 Eigenvalue Problems under a Pseudo Inner Product

We now restate three theorems presented in previous sections, this time using a pseudo inner product. Since a pseudo inner product does not generate a Hilbert space, here we consider linear operators $A, B : X \rightarrow X$, where X is a pseudo inner product space (see Section 2.5.2).

For the ordinary eigenvalue problem $Ax = \lambda x$, Theorem 4.7 becomes the following.

Theorem 4.31. *If x_n are eigenvectors of the linear operator $A : X \rightarrow X$, with corresponding eigenvalues λ_n (i.e., $Ax_n = \lambda_n x_n$), then λ_n are also eigenvalues of the pseudo adjoint operator $A^{p*}(X)$ corresponding to adjoint eigenvectors y_n (i.e., $A^{p*}y_n = \lambda_n y_n$), and*

$$(\lambda_n - \lambda_m) \langle x_n, y_m \rangle_p = 0.$$

The proof follows along the same lines as the proof of Theorem 4.7, and, in fact, more generally

$$\sigma(A) = \sigma(A^{p*}).$$

For a pseudo self-adjoint problem ($A = A^{p*}$ and, upon suitable ordering, $y_n = x_n$), we have the orthogonality relationship $(\lambda_n - \lambda_m) \langle x_n, x_m \rangle_p = 0$.

For a generalized eigenvalue problem $Ax = \lambda Bx$, Theorem 4.29 becomes the following.

Theorem 4.32. *Let $A, B : X \rightarrow X$ be pseudo self-adjoint (or pseudo symmetric) operators. Then eigenvectors corresponding to $Ax = \lambda Bx$ satisfy*

the orthogonality relationships

$$\begin{aligned}(\lambda_n - \lambda_m) \langle Bx_n, x_m \rangle_p &= 0, \\ \left(\frac{1}{\lambda_n} - \frac{1}{\lambda_m} \right) \langle Ax_n, x_m \rangle_p &= 0\end{aligned}$$

($\lambda_n, \lambda_m \neq 0$ in the last expression).

For the case of a generalized nonpseudo self-adjoint problem, Theorem 4.30 becomes the next result.

Theorem 4.33. *If x_n are eigenvectors of $Ax = \lambda Bx$, with corresponding eigenvalues λ_n (i.e., $Ax_n = \lambda_n Bx_n$), then λ_n are eigenvalues of the pseudo adjoint equation $A^{p*}y = \lambda^{p*}B^{p*}y$ (i.e., $\lambda_n^{p*} = \lambda_n$) corresponding to pseudo adjoint eigenvectors y_n , and*

$$\begin{aligned}(\lambda_n - \lambda_m) \langle Bx_n, y_m \rangle_p &= 0, \\ \left(\frac{1}{\lambda_n} - \frac{1}{\lambda_m} \right) \langle Ax_n, y_m \rangle_p &= 0\end{aligned}$$

($\lambda_n, \lambda_m \neq 0$ in the last expression).

Again, the more general relationship

$$\sigma(A, B) = \sigma(A^{p*}, B^{p*})$$

holds, and we assume that the residual spectrum is empty in all cases.

4.2.11 Pseudo Self-Adjoint, Nonstandard Eigenvalue Problems

Recall that the nonstandard eigenvalue problem (4.5) is given as

$$A(\gamma)x = 0, \tag{4.15}$$

where γ is called a nonstandard eigenvalue and x is a nonstandard eigenvector. Usually A is a nonlinear operator-valued function of γ .

In the general case, not much can be said concerning the spectral properties of the nonstandard eigenvalues γ and eigenfunctions x . Furthermore, recalling that pseudo self-adjointness is a far weaker property (for complex spaces) than self-adjointness, it is interesting and very useful to note that if a nonstandard eigenvalue problem is such that $A(\gamma)$ is pseudo self-adjoint, then it is easy to obtain a stationary functional for the nonstandard eigenvalue γ . In fact, that stationary functional is related to a Galerkin solution of the operator equation. The analysis presented here follows [8], where a more general problem is treated.

We assume that the operator $A(\gamma)$ in the nonstandard problem is pseudo self-adjoint, $\langle Ax, y \rangle_p = \langle y, Ax \rangle_p$, which is physically associated with reciprocity. A variational expression for a nonstandard eigenvalue γ is obtained from

$$F(\gamma, x) = \langle x, A(\gamma)x \rangle_p = 0.$$

We form the first variation of F by replacing γ and x with $\gamma + \delta\gamma$ and $x + \delta x$, respectively. Therefore,

$$F(\gamma + \delta\gamma, x + \delta x) = \langle x + \delta x, A(\gamma + \delta\gamma)(x + \delta x) \rangle_p = F(\gamma, x) + \delta F. \quad (4.16)$$

Defining $A' \equiv \partial A(\gamma) / \partial \gamma = [A(\gamma + \delta\gamma) - A(\gamma)] / \delta\gamma$ [9, Ch. 4], which approaches the derivative of A as $\delta\gamma \rightarrow 0$, and neglecting second-order terms result in

$$\delta F = \delta\gamma \langle x, A'x \rangle_p + \langle x, A\delta x \rangle_p \quad (4.17)$$

where we have used (4.15). Exploiting the pseudo self-adjointness of A , $\langle x, A\delta x \rangle_p = \langle Ax, \delta x \rangle_p$, the second term in (4.17) vanishes, resulting in

$$\delta F = \delta\gamma \langle x, A'x \rangle_p. \quad (4.18)$$

Therefore, $\delta\gamma = 0$ is a stationary point of the functional F (unless it happens that $\langle x, A'(\gamma)x \rangle_p = 0$), such that at $\delta\gamma = 0$ the first variation of F vanishes. Thus, first-order errors in x result in second-order errors in γ .

In a method-of-moments solution of $Ax = 0$,

$$\langle x, A(\gamma)x \rangle_p = \langle A(\gamma)x, x \rangle_p = 0$$

is simply Galerkin's method, which is widely used to solve a variety of integral equations occurring in applied electromagnetics. It is easy to see that if x is approximated by a series of basis functions, $x \sim \tilde{x} = \sum_{i=1}^N a_i x_i$, then $\langle A(\gamma)x, x \rangle_p = 0$ becomes a stationary (nonlinear) functional equation

$$H(\gamma) = \det[Z(\gamma)] = 0,$$

where Z is the matrix with elements $Z_{ij} = \langle A(\gamma)x_j, x_i \rangle_p$. Since for electromagnetic applications in reciprocal media the resulting operators are generally pseudo self-adjoint (see the footnote on p. 162), natural resonance and spectral problems typically exhibit this stationary characteristic. An example is provided in Section 8.3.

We assume that the equation $H(\gamma) = 0$ leads to the first-order root $\gamma = \gamma_0$. A second-order root is obtained by imposing the additional condition $\partial H(\gamma) / \partial \gamma = H'(\gamma) = 0$. The system of equations

$$H(\gamma) = H'(\gamma) = 0$$

(with some additional constraints) has been applied for the numerical determination of various critical points in waveguiding problems in [10] and references therein. The above procedure also holds if A is self-adjoint, $\langle Ax, y \rangle = \langle y, Ax \rangle$, by replacing $\langle \cdot, \cdot \rangle_p$ with $\langle \cdot, \cdot \rangle$.

4.2.12 Steinberg's Theorems for Compact Operators

Several theorems relating to the spectral properties of identity-plus-compact Fredholm operators are presented by Steinberg in [11] (see also [12]). In the following theorems A is a compact operator on a Banach space B .

Theorem 4.34. (*Analytic Fredholm theorem*) *If $A(\gamma)$ is an analytic family of compact operators¹⁶ for $\gamma \in \Omega$, where Ω is an open, connected subset of the complex plane, then either*

- $(I - A(\gamma))$ is nowhere invertible in Ω , or else
- $(I - A(\gamma))^{-1}$ is meromorphic in Ω .

If the second result holds, then there exists a discrete set of values $\{\gamma_n\}$ that are the poles of the inverse operator. In this case $(I - A(\gamma))^{-1}$ exists and is analytic at all points $\gamma \neq \gamma_n$, and for $\gamma = \gamma_n$ the equation

$$(I - A(\gamma))x = 0$$

has a nontrivial solution in B .

Note that the set $\{\gamma_n\}$ may be empty. Therefore, Theorem 4.34 is not an existence theorem for poles. Rather, it guarantees that if the second condition holds, then if singularities are present in Ω , they will be pole singularities.

Example 4.4.

Consider scattering from a perfectly conducting, finite-sized object having a sufficiently smooth surface S located in free space. The magnetic field integral equation (1.118)

$$\int_S [\underline{\mathbf{G}}_{me}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}_e^s(\mathbf{r}')] \times \mathbf{n} dS' + \frac{1}{2} \mathbf{J}_e^s(\mathbf{r}) = \mathbf{n} \times \mathbf{H}^i(\mathbf{r}), \quad \mathbf{r} \in S, \quad (4.19)$$

can be used to determine the induced current density $\mathbf{J}_e^s$ caused by an excitation $\mathbf{H}^i$, which can, in turn, be used to determine the scattered field. From (1.79),

$$\begin{aligned} \underline{\mathbf{G}}_{me}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}_e^s(\mathbf{r}') &= (\nabla g(\mathbf{r}, \mathbf{r}') \times \underline{\mathbf{I}}) \cdot \mathbf{J}_e^s(\mathbf{r}') \\ &= \nabla g(\mathbf{r}, \mathbf{r}') \times \mathbf{J}_e^s(\mathbf{r}'), \end{aligned}$$

where $g = (e^{-jkR}) / (4\pi R)$ is the usual free-space scalar Green's function and $k = \omega\sqrt{\mu_0\epsilon_0}$, we obtain the integral equation

$$\mathbf{J}_e^s(\mathbf{r}) - 2 \int_S \mathbf{n} \times (\nabla g(\mathbf{r}, \mathbf{r}') \times \mathbf{J}_e^s(\mathbf{r}')) dS' = 2\mathbf{n} \times \mathbf{H}^i(\mathbf{r}), \quad \mathbf{r} \in S. \quad (4.20)$$

¹⁶This means that $A(\gamma)$ is analytic and compact for each $\gamma \in \Omega$.

This can be written as

$$(I - A(k)) \mathbf{J}_e^s(\mathbf{r}) = 2\mathbf{n} \times \mathbf{H}^i(\mathbf{r})$$

for $\mathbf{r} \in S$, where the operator $A(k) : \mathbf{L}^2(\Omega)^3 \rightarrow \mathbf{L}^2(\Omega)^3$, defined as

$$(A\mathbf{J})(\mathbf{r}) \equiv 2 \int_S \mathbf{n} \times \left[\nabla \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \times \mathbf{J}(\mathbf{r}') \right] dS'$$

for suitable currents $\mathbf{J}$, is compact as discussed in Section 3.6.2 and clearly analytic in the finite k -plane. $(I - A)$ is invertible for some k and therefore, by the above theorem, $(I - A(k))^{-1}$ is a meromorphic operator-valued function of k . This establishes the existence of a discrete set of nonstandard eigenvalues $\{k_n\} \in \mathbf{C}$, at which points the inverse does not exist. It is well known that this set is not empty.

Viewed as a spectral problem,

$$(I - A(k)) \mathbf{x}_n = \lambda_n(k) \mathbf{x}_n,$$

the values of k_n are such that $\lambda_n(k_n) = 0$. These points lead to complex-valued resonance frequencies $\omega_n = k_n/\sqrt{\mu_0\epsilon_0}$, which are called the *exterior resonances* of the object and which may be determined from

$$\det(I - A(k_n)) = 0.$$

In many cases it is assumed that the poles are simple, which can be proven for S corresponding to a spherical surface. By the Fredholm alternative (Theorem 3.44), because $(I - A(k))^{-1}$ does not exist at $k = k_n$, nontrivial solutions $\mathbf{x}_n$ and $\mathbf{y}_n$ exist such that

$$\begin{aligned} (I - A(k)) \mathbf{x}_n &= 0, \\ (I - A^*(k)) \mathbf{y}_n &= 0. \end{aligned}$$

The eigenfunctions and adjoint eigenfunctions $\mathbf{x}_n$ and $\mathbf{y}_n$, respectively, form a bi-orthogonal set (although not necessarily a basis) in $\mathbf{L}^2(\Omega)^3$.

For $k \neq k_n$, the solution of the integral equation is

$$\mathbf{J}_e^s(\mathbf{r}) = 2(I - A(k))^{-1} \mathbf{n} \times \mathbf{H}^i(\mathbf{r}),$$

and since poles are the only singularities in the finite k -plane the Mittag-Leffler theorem [13] can be used to expand the inverse in terms of the associated spectral parameters (see [14], and, in particular, [15]). This important result, originally obtained from the magnetic field integral equation by a somewhat different method in [16], was used as a theoretical basis for the *singularity expansion method* developed in [17]. By the uniqueness theorem the exterior resonances are determined by the geometry of the object, and not the mathematical formulation.

For some additional applications in electromagnetic scattering, see [18].

The next theorem extends the above result to include the dependence of the inverse on a nonspectral parameter $\tau \in \mathbf{R}$.

Theorem 4.35. *Let $A(\gamma, \tau)$ be a family of compact operators analytic in $\gamma \in \Omega$, where Ω is an open, connected subset of the complex plane, and let $A(\gamma, \tau)$ be jointly continuous in (γ, τ) for each $(\gamma, \tau) \in \Omega \times \mathbf{R}$. Then,*

- *if $(I - A(\gamma, \tau))$ is somewhere invertible for each τ , $(I - A(\gamma, \tau))^{-1}$ is meromorphic in γ for each τ ,*
- *if γ_0 is not a pole of $(I - A(\gamma, \tau_0))^{-1}$, $(I - A(\gamma, \tau))^{-1}$ is jointly continuous in (γ, τ) at (γ_0, τ_0) , and*
- *the poles $\gamma(\tau)$ of $(I - A(\gamma, \tau))^{-1}$ depend continuously on τ and can appear and disappear only at the boundary of Ω (which may extend to infinity).*

In typical electromagnetics applications, τ represents a geometrical parameter, and γ represents frequency in a three-dimensional problem or propagation constant in a two-dimensional problem. The above theorem can be applied to establish (a) the existence of a discrete set (possibly empty) of nonstandard eigenvalues γ_n for a given τ , (b) the continuity of the inverse operator with respect to spectral and structural parameters, and (c) the continuity of the spectral values $\gamma_n(\tau)$ as a function of τ . For example, this theorem can be used to show that the exterior resonances k_n described in Example 4.4 smoothly vary as the shape of the scatterer S smoothly varies.

In electromagnetics problems involving lossy media or infinite media, the domain of analyticity Ω often is constrained by the presence of branch point singularities, as the next example demonstrates.

Example 4.5.

As in Example 4.4, consider scattering from a perfectly conducting, finite-sized object having a sufficiently smooth surface S , but this time let the object reside within an infinite parallel-plate region with plate separation a . As in the previous example, the magnetic field integral equation (1.118)

$$\int_S [\underline{\mathbf{G}}_{me}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}_e^s(\mathbf{r}')] \times \mathbf{n} dS' + \frac{1}{2} \mathbf{J}_e^s(\mathbf{r}) = \mathbf{n} \times \mathbf{H}^i(\mathbf{r}), \quad \mathbf{r} \in S, \quad (4.21)$$

can be used to determine the induced current density $\mathbf{J}_e^s$ caused by an excitation $\mathbf{H}^i$. In this case, the Green's dyadic $\underline{\mathbf{G}}_{me}$ accounts for the infinite parallel plates (see Sections 8.1.2 and 8.3.2) and is given by (8.197) with (8.203) and (8.208),

$$\begin{aligned} \underline{\mathbf{G}}_{me}(\mathbf{r}, \mathbf{r}') &= \nabla \times \underline{\mathbf{G}}_{\pi}(\mathbf{r}, \mathbf{r}') \\ &= \nabla \times (\hat{\mathbf{x}}\hat{\mathbf{x}} G_{\pi}^E(\mathbf{r}, \mathbf{r}') + (\hat{\mathbf{y}}\hat{\mathbf{y}} + \hat{\mathbf{z}}\hat{\mathbf{z}}) G_{\pi}^H(\mathbf{r}, \mathbf{r}')), \end{aligned}$$

where

$$\begin{aligned} G_{\pi}^H(\mathbf{r}, \mathbf{r}') &= \frac{1}{2ia} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{n\pi}{a}x'\right) H_0^{(2)}(\psi_n \rho), \\ G_{\pi}^E(\mathbf{r}, \mathbf{r}') &= \frac{1}{4ia} \sum_{n=0}^{\infty} \epsilon_n \cos\left(\frac{n\pi}{a}x\right) \cos\left(\frac{n\pi}{a}x'\right) H_0^{(2)}(\psi_n \rho), \end{aligned} \quad (4.22)$$

with

$$\psi_n = \sqrt{k^2 - \left(\frac{n\pi}{a}\right)^2}.$$

We obtain the operator equation

$$(I - A(k)) \mathbf{J}_e^s(\mathbf{r}) = 2\mathbf{n} \times \mathbf{H}^i(\mathbf{r})$$

for $\mathbf{r} \in S$. The operator $A(k) : \mathbf{L}^2(\Omega)^3 \rightarrow \mathbf{L}^2(\Omega)^3$, defined as

$$(A\mathbf{J})(\mathbf{r}) \equiv -2 \int_S [\underline{\mathbf{G}}_{me}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')] \times \mathbf{n} dS'$$

for suitable currents $\mathbf{J}$, is compact but clearly not analytic in the finite k -plane, since the Green's dyadic is a function of the parallel-plate waveguide dispersion function ψ_n .

Concentric regions Ω_n may be defined in the complex k -plane wherein $A(k)$ is analytic,

$$\begin{aligned} \Omega_0 &\equiv \left\{ k : |k| < \frac{\pi}{a} \right\}, \\ \Omega_1 &\equiv \left\{ k : \frac{\pi}{a} < |k| < \frac{2\pi}{a} \right\}, \\ &\vdots \\ \Omega_n &\equiv \left\{ k : \frac{n\pi}{a} < |k| < \frac{(n+1)\pi}{a} \right\}, \end{aligned}$$

such that $(I - A(k))^{-1}$ will be meromorphic in each Ω_n . The poles of the resolvent, k_n , may appear or disappear at the values $k = n\pi/a$, which defines the cutoff condition for parallel-plate waveguide modes (at cutoff forward and backward propagating modes intersect, and below cutoff they become evanescent).

4.3 Expansions and Representations, Spectral Theorems

In this section we are concerned with simple representations of linear operators. We first discuss the matrix representation of linear operators on

finite-dimensional spaces, where it is often of interest to obtain as simple a matrix representation as possible. Conditions under which an operator has a diagonal matrix representation are stated. If a diagonal matrix representation is not possible, then other fairly simple representations are discussed.

We will see that the prime consideration for this topic is the choice of an appropriate basis set. For example, in Section 3.1.2 we found that a bounded linear operator $A : H \rightarrow H$ has the matrix representation

$$\begin{bmatrix} \langle Ax_1, x_1 \rangle & \langle Ax_2, x_1 \rangle & \cdots \\ \langle Ax_1, x_2 \rangle & \langle Ax_2, x_2 \rangle & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

with respect to the basis $\mathbf{x} = \{x_1, x_2, \dots\}$ of H (the matrix has finite (infinite) dimension if H is finite- (infinite-) dimensional). It is clear that if the basis $\mathbf{x}$ is such that $\langle Ax_i, x_j \rangle = 0$ for $i \neq j$, the matrix representation of A will be diagonal. Conditions for this to occur will be discussed, but as a preview if x_i is an eigenvector of A corresponding to the eigenvalue λ_i , then $\langle Ax_i, x_j \rangle = \langle \lambda_i x_i, x_j \rangle = \lambda_i \langle x_i, x_j \rangle$. If $\langle x_i, x_j \rangle = 0$ for $i \neq j$, the off-diagonal entries of the matrix will be zero. However, there also needs to be enough eigenvectors, so that they form a basis for the space H .

This section begins by considering the matrix representation of operators on finite-dimensional spaces with respect to different bases. Operators that have a diagonal representation with respect to some (eigenvector) basis are identified, and similarity transformations from an arbitrary basis to an eigenvector basis are discussed. Operators on infinite-dimensional spaces are considered next. Although one may view the infinite-dimensional case in the same way as the finite-dimensional case—that is, as a question of finding a diagonal matrix representation of the operator—it is more useful to focus on the (essentially equivalent) eigenfunction expansion problem. Operators on infinite-dimensional spaces that possess a complete eigenfunction basis are identified, and associated expansions are discussed.

4.3.1 Operators on Finite-Dimensional Spaces

Because the spectral theory of operators on finite-dimensional spaces is much simpler than that for operators on infinite-dimensional spaces (recall that in the finite-dimensional case σ_c and σ_r do not exist, and σ_p is purely discrete), and since electromagnetic problems are often ultimately cast as numerical matrix problems, we begin this section with some elements from the spectral theory of matrices. For generality, we work within finite-dimensional Hilbert spaces H_1^n, H_2^m , with the superscript indicating the dimension of the space.

Equivalent and Similar Matrices

Recall from the examples in Section 3.1.2 that the bounded operator $A : H_1^n \rightarrow H_2^m$ is uniquely represented, for $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ a basis of H_1^n and $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ a basis of H_2^m , by the $m \times n$ matrix $[a_{ij}]_{\mathbf{x}, \mathbf{y}} \in \mathbf{C}^{m \times n}$, where $a_{ij} = \langle A\mathbf{x}_j, \mathbf{y}_i \rangle$ and the subscript indicates the basis pairs.¹⁷ If $\mathbf{x}, \mathbf{y}$ are the standard bases in $H_1^n = \mathbf{C}^n, H_2^m = \mathbf{C}^m$, respectively, then $[a_{ij}]_{\mathbf{x}, \mathbf{y}}$ is called the *standard matrix representation* of the linear operator A . The matrix representation of a given operator will, of course, be different for different bases, and it is often desirable to achieve a matrix representation that is as simple as possible, perhaps triangular or even diagonal in form.

It can be shown that for a given linear operator $A : H_1^n \rightarrow H_2^m$, represented as $[a_{ij}]_{\mathbf{x}, \mathbf{y}}$ in the pair of bases $\{\mathbf{x}_n\}$ and $\{\mathbf{y}_n\}$, and as $[a_{ij}]_{\xi, \zeta}$ in the pair of bases $\{\xi_n\}$ and $\{\zeta_n\}$, the two matrix operator representations are related as

$$[a_{ij}]_{\mathbf{x}, \mathbf{y}} = [P]_{\zeta \rightarrow \mathbf{y}} [a_{ij}]_{\xi, \zeta} [Q]_{\xi \rightarrow \mathbf{x}}^{-1}. \quad (4.23)$$

In the above, the $m \times m$ matrix $[P]_{\zeta \rightarrow \mathbf{y}}$ and the $n \times n$ matrix $[Q]_{\xi \rightarrow \mathbf{x}}$ are nonsingular *transition matrices* from the bases ζ to $\mathbf{y}$ and ξ to $\mathbf{x}$, respectively [19, p. 127].¹⁸ Note also that if $[Q]_{\xi \rightarrow \mathbf{x}}$ is a nonsingular transition matrix from the basis ξ to the basis $\mathbf{x}$, then inversion of the matrix reverses the operation, i.e., $[Q]_{\xi \rightarrow \mathbf{x}}^{-1}$ is the transition matrix from the basis $\mathbf{x}$ to the basis ξ . If the bases are orthonormal, then $[P]$ and $[Q]$ are unitary (therefore, $[P][P]^* = [I]$); by Example 2 in Section 3.3.2 if the bases are orthonormal, then the inverse of $[P]$ or $[Q]$ is easily computed as the conjugate-transpose matrix.

The two matrices $[a_{ij}]_{\mathbf{x}, \mathbf{y}}$ and $[a_{ij}]_{\xi, \zeta}$ in (4.23) are said to be *equivalent*, in the sense that the set of all matrix representations of A with respect to different pairs of bases defines an equivalence class of matrices that all represent the “parent” operator $A : H_1^n \rightarrow H_2^m$ in $\mathbf{C}^{m \times n}$.

Now, for simplicity, consider a finite-dimensional operator $A : H^n \rightarrow H^n$, represented as $[a_{ij}]_{\mathbf{x}}$ in the basis $\mathbf{x}$ ($a_{ij} = \langle A\mathbf{x}_j, \mathbf{x}_i \rangle$) and as $[a_{ij}]_{\xi}$ in the basis ξ ($a_{ij} = \langle A\xi_j, \xi_i \rangle$). The two operator representations (matrices) are related as [19, p. 130]

$$[a_{ij}]_{\mathbf{x}} = [P] [a_{ij}]_{\xi} [P]^{-1}, \quad (4.24)$$

where the $n \times n$ matrix $[P] = [P]_{\xi \rightarrow \mathbf{x}}$ is a nonsingular transition matrix from the basis ξ to the basis $\mathbf{x}$. When two matrices are related as in (4.24), they are said to be *similar*. Similar matrices have the same rank, determinant, and eigenvalues, and related eigenvectors.

¹⁷The space of bounded operators $H_1^n \rightarrow H_2^m$ and the space $\mathbf{C}^{m \times n}$ are isomorphic, and so one can work with the operator, or a matrix representation of the operator, in an equivalent manner.

¹⁸Explicitly, the matrix $[P]_{\zeta \rightarrow \mathbf{y}}$ is such that the two bases are related as $\zeta_j = \sum_{i=1}^m P_{ij} \mathbf{y}_i$, $j = 1, 2, \dots, m$, and similarly for $[Q]$.

Spectral Representations (Matrix Form)

The primary reason for introducing the concept of similar matrices is to facilitate an understanding of when and how one obtains a “simple” matrix representation of a given operator on a finite-dimensional space. The simplest possible matrix has a diagonal form, and the next theorem provides conditions leading to a diagonal matrix representation of an operator $A : H^n \rightarrow H^n$ (alternately, one may say that the matrix representation of the operator in some basis is *diagonalizable*, or simply that a given matrix is diagonalizable). A clear way to envision this problem is to recognize that the set of all matrix representations of A with respect to different bases defines an equivalence class of similar matrices. The problem is then to find the simplest matrix in this class, which will undoubtedly most clearly show the salient features of the “parent” operator.

Theorem 4.36. *Let $A : H^n \rightarrow H^n$ be a linear operator on a finite-dimensional complex Hilbert space. If A has n distinct eigenvalues $\{\lambda_n\}$, then A has a diagonal matrix representation with respect to a basis of eigenvectors¹⁹ of A .*

Actually, the last theorem provides a sufficient but unnecessary condition. More generally, we do not actually require n distinct eigenvalues; instead we require that the eigenvectors of A form a basis for the space H^n .

Theorem 4.37. *Let $A : H^n \rightarrow H^n$ be a linear operator on a finite-dimensional, complex Hilbert space. Then if A has n linearly independent eigenvectors $\{\mathbf{x}_n\}$ (i.e., an eigenbasis of H^n), corresponding to the (not necessarily distinct) eigenvalues $\{\lambda_n\}$, A has a diagonal matrix representation with respect to its eigenbasis.*

Such an operator is said to have *simple structure* (or, equivalently, to be represented by a simple matrix). The elements on the main diagonal of the resulting matrix representation are exactly the eigenvalues of the operator.

The next theorem states precisely when an operator on a finite-dimensional space will possess an eigenbasis.

Theorem 4.38. *(Complex spectral theorem for finite-dimensional spaces) Let $A : H^n \rightarrow H^n$ be a linear operator on a finite-dimensional, complex Hilbert space. Then H^n has an orthonormal basis of eigenvectors $\{\mathbf{x}_n\}$ if and only if A is normal.*

Therefore, if $A : H^n \rightarrow H^n$ is normal (or self-adjoint), it has a diagonal matrix representation with respect to the eigenbasis $\mathbf{e} = \{\mathbf{x}_n\}$,

$$[A]_{\mathbf{e}} = \text{diag} [\lambda_1, \lambda_2, \dots, \lambda_n].$$

¹⁹A basis of a space consisting of eigenvectors of an operator is called an *eigenbasis*.

It also turns out that two self-adjoint operators A and B can be simultaneously diagonalized (by a single basis) if they commute, i.e., if $[A, B] = AB - BA = 0$. The above theorem holds for H^n being a real space if the condition that A be normal is replaced with the condition that A be self-adjoint (symmetric). For a more complete discussion, see [20, p. 133].

It is again useful to recall that all self-adjoint operators are normal, but normal operators may not be self-adjoint. If A is self-adjoint, then the resulting diagonal matrix representation has real entries on the main diagonal.

It can also be shown that if $A : H^n \rightarrow H^n$ admits a diagonal matrix representation, then any other nondiagonal matrix representation of the same operator (that is, a matrix representation with respect to a basis ξ that is not an eigenbasis) may be diagonalized by (4.24).

Example 4.6.

To demonstrate these concepts, consider the self-adjoint operator $A : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ defined by

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \equiv \begin{bmatrix} \alpha_1 + \alpha_2 \\ \alpha_1 - \alpha_2 \end{bmatrix}$$

and the two orthonormal bases $(\mathbf{y}_1, \mathbf{y}_2)$ and (ξ_1, ξ_2) , where

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

is the standard basis and

$$\xi_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \xi_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then the matrix representations with respect to the two bases are

$$[A]_{\mathbf{y}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad [A]_{\xi} = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}.$$

The transition matrix from the basis (ξ_1, ξ_2) to the basis $(\mathbf{y}_1, \mathbf{y}_2)$ is found to be

$$[P]_{\xi \rightarrow \mathbf{y}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

such that (4.24) is easily seen to be satisfied, i.e., $[A]_{\mathbf{y}} = [P]_{\xi \rightarrow \mathbf{y}} [A]_{\xi} [P]_{\xi \rightarrow \mathbf{y}}^{-1}$.

Neither basis yields a diagonal matrix representation. Because the operator is self-adjoint, and so necessarily normal, we can find an eigenbasis that will yield a diagonal matrix representation. Alternatively, we can define a similarity transformation from, say, the basis $(\mathbf{y}_1, \mathbf{y}_2)$ to the eigenbasis that will result in a diagonal matrix representation.

The eigenvalues and eigenvectors of the operator A are found to be

$$\lambda_{1,2} = \pm\sqrt{2}, \quad \mathbf{e}_{1,2} = \begin{bmatrix} 1 \\ \pm\sqrt{2} - 1 \end{bmatrix},$$

which, when normalized to form an orthonormal eigenbasis, become

$$\mathbf{x}_1 = \frac{1}{2\sqrt{1-2^{-1/2}}} \begin{bmatrix} 1 \\ \sqrt{2}-1 \end{bmatrix}, \quad \mathbf{x}_2 = \frac{1}{2\sqrt{1+2^{-1/2}}} \begin{bmatrix} 1 \\ -\sqrt{2}-1 \end{bmatrix}.$$

Noting that the transition matrix from the eigenbasis $\mathbf{e}$ to the basis $\mathbf{y}$ is

$$[P]_{\mathbf{e} \rightarrow \mathbf{y}} = \begin{bmatrix} \frac{1}{2\sqrt{1-2^{-1/2}}} & \frac{1}{2\sqrt{1+2^{-1/2}}} \\ \frac{\sqrt{2}-1}{2\sqrt{1-2^{-1/2}}} & \frac{-\sqrt{2}-1}{2\sqrt{1+2^{-1/2}}} \end{bmatrix} = [P]_{\mathbf{y} \rightarrow \mathbf{e}},$$

then, using (4.24), we obtain

$$[A]_{\mathbf{e}} = [P][A]_{\mathbf{y}}[P]^{-1} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix},$$

the desired diagonal matrix representation of the operator A .

In summary, every finite-dimensional normal operator is similar to a diagonal matrix operator. As such, every finite-dimensional normal operator $A : H^n \rightarrow H^n$ can be represented as a multiplication operator, $(PAP^{-1}\mathbf{x})_i = \lambda_i x_i$. This fact also holds for normal operators on infinite-dimensional spaces.

Jordan Canonical Form

It is seen that Theorem 4.38 characterizes explicitly when an operator on a finite-dimensional, complex Hilbert space will have a diagonal matrix representation. If an operator is not normal, a diagonal matrix representation will not be possible. In this case it is of interest to investigate what other fairly simple forms the matrix representation of such a linear operator may assume, which brings up the concept of *generalized eigenvectors* and the *Jordan canonical form*. For completeness this topic is included here, although the treatment is very brief. This material does, however, relate in an analogous way to the topic of generalized eigenvectors in infinite-dimensional space problems.

Consider a linear operator on a finite-dimensional space $A : H \rightarrow H$ with an eigenvalue λ_n . Recall that $\mathbf{0} \neq \mathbf{x}_n \in H$ is an eigenvector of A corresponding to an eigenvalue λ_n if $(A - \lambda_n I)\mathbf{x}_n = \mathbf{0}$. To generalize this concept, an element $\mathbf{x}_{n,m-1} \in H$ is called a *generalized eigenvector* (or *root vector*) of rank m of the operator A corresponding to the eigenvalue λ_n if

$$\begin{aligned} (A - \lambda_n I)^m \mathbf{x}_{n,m-1} &= \mathbf{0}, \\ (A - \lambda_n I)^{m-1} \mathbf{x}_{n,m-1} &\neq \mathbf{0}, \end{aligned}$$

where m is a positive integer.²⁰ The rank m cannot be higher than the algebraic multiplicity of the corresponding eigenvalue λ_n . Note that every (ordinary) eigenvector of A is a generalized eigenvector of rank 1 ($\mathbf{x}_{n,0} \equiv \mathbf{x}_n$). We will call generalized eigenvectors having rank $m > 1$ *associated eigenvectors*, in order to distinguish them from ordinary eigenvectors.²¹ The system of generalized eigenvectors is called the *root system* of A and is composed of the union of the ordinary and associated eigenvectors.

The next theorem highlights the importance of generalized eigenvectors.

Theorem 4.39. *Consider the linear operator $A : H^n \rightarrow H^n$, where H^n is a complex, finite-dimensional Hilbert space. Then the maximal set of all linearly independent generalized eigenvectors of A forms a basis (generalized eigenbasis) of H^n .*

Therefore, in H^n this basis, also called the *root basis*, exists for any operator $A : H^n \rightarrow H^n$. For infinite-dimensional space this is not necessarily the case.

Furthermore, let us call a basis of H^n a *Jordan basis* for A if, with respect to this basis, the matrix representation of A has a block-diagonal form

$$\begin{bmatrix} M_1 & & 0 \\ & \ddots & \\ 0 & & M_n \end{bmatrix}, \quad (4.25)$$

where the matrices M_i are lower-triangular, e.g.,

$$M_i = \begin{bmatrix} \lambda_i & 0 & 0 & 0 \\ 1 & \lambda_i & 0 & 0 \\ 0 & 1 & \lambda_i & 0 \\ 0 & 0 & 1 & \lambda_i \end{bmatrix}.$$

The dimension of the matrix M_i is equal to the algebraic multiplicity of λ_i , and, in the case of simple eigenvalues, the matrix representation of A is diagonal.

For any complex-valued matrix operator $A : H^n \rightarrow H^n$ there is a basis of generalized eigenvectors that is a Jordan basis. Therefore, any $A : H^n \rightarrow H^n$ with H^n complex admits a matrix representation of the form (4.25), which is fairly simple. Matrix representations of A in some other basis can be brought into the Jordan form by a similarity transformation, representing a change of basis to the Jordan basis. Unfortunately, while the

²⁰Note that we define an operator raised to a power as $A^n x = \underbrace{A A A \cdots A}_{n \text{ times}} x$, where for this to make sense $R_A \subseteq D_A$ (see Section 4.4).

²¹It can be shown that self-adjoint operators only possess generalized eigenvectors of rank 1, i.e., that they only possess ordinary eigenvectors (see p. 259). However, this does not imply that the dimension of the eigenspace associated with an eigenvalue is necessarily unity.

Jordan basis is made up of generalized eigenvectors, its actual construction is somewhat complicated and will not be discussed here (see, for instance, [19, Ch. 6]).

Spectral Representations

In preparation for examining eigenfunction expansions in the next section, we consider the *spectral representation* of an operator acting on a finite-dimensional space. Assume $A : H^n \rightarrow H^n$ admits a diagonal representation $[A]_{\mathbf{e}}$ with respect to an eigenbasis $\mathbf{e} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, and write

$$[A]_{\mathbf{e}} = \sum_{i=1}^n \lambda_i [\tilde{P}_i],$$

where $[\tilde{P}_i]$ is the matrix for which all entries are zero except the i th element on the main diagonal ($\sum_{i=1}^n [\tilde{P}_i] = I_{n \times n}$). Then, the representation of A in an arbitrary basis $\mathbf{x}$ is

$$[A]_{\mathbf{x}} = [P] \left(\sum_{i=1}^n \lambda_i [\tilde{P}_i] \right) [P]^{-1} = \sum_{i=1}^n \lambda_i [P_i],$$

where $[P_i] \equiv [P] [\tilde{P}_i] [P]^{-1}$, with $[P] = [P]_{\mathbf{e} \rightarrow \mathbf{x}}$ being the nonsingular transition matrix from the basis $\mathbf{e}$ to the basis $\mathbf{x}$. In operator form we have

$$A = \sum_{i=1}^n \lambda_i P_i. \quad (4.26)$$

Note the simple geometric meaning presented by the representation (4.26): The space H^n decomposes into n orthogonal subspaces $H_1, H_2, \dots, H_n$, each H_i being associated²² with the i th eigenvector $\mathbf{x}_i$. For any $\mathbf{x} \in H^n$,

$$P_i \mathbf{x} = \alpha_i \mathbf{x}_i \in H_i.$$

Therefore, the $P_i : H^n \rightarrow H_i$ are seen to be projection operators from the space H^n to the subspace H_i (note that²³ $\sum_{i=1}^n P_i = I_{n \times n}$). This representation for A is called a *weighted sum of projections* and is depicted in Figures 4.4 and 4.5.

The above decomposition for normal or self-adjoint operators on a finite-dimensional space H^n can be extended to normal and self-adjoint operators on infinite-dimensional spaces, most transparently using Riemann–Stieltjes integrals. The procedure is, nevertheless, quite complicated and so will be omitted here, although if we add the condition that A is compact we are back to a simple situation, as shown in the next section.

²²The null space for each λ_i may not be one-dimensional, but the space spanned by $\mathbf{x}_i$, i.e., H_i , is one-dimensional. Therefore, more than one H_i may be associated with a single eigenvalue.

²³ $\sum_{i=1}^n P_i = I_{n \times n}$ is called a *resolution of the identity*.

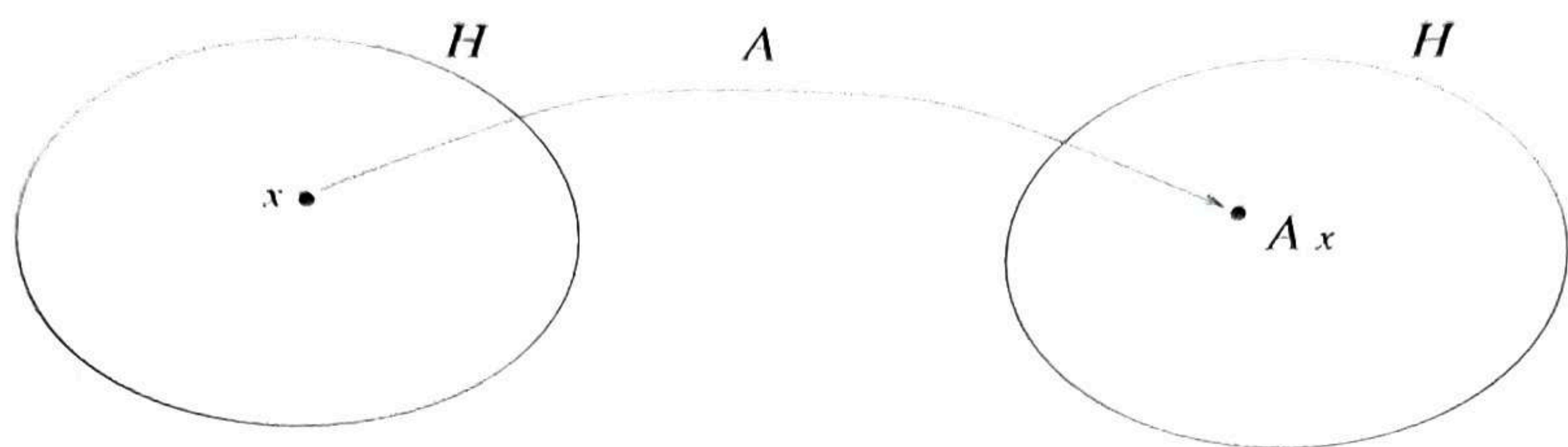


Figure 4.4: Depiction of the mapping $A : H \rightarrow H$.

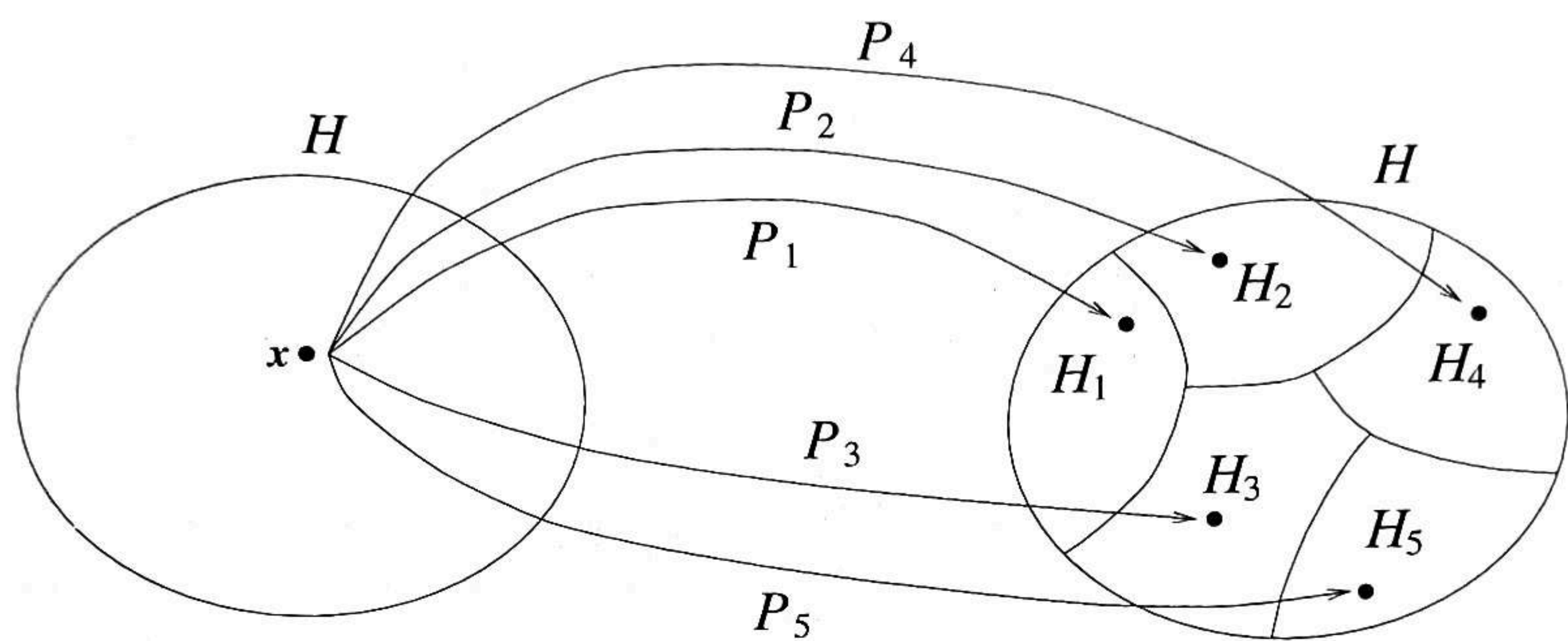


Figure 4.5: Decomposition of a simple operator $A : H \rightarrow H$ as $A = \sum_{i=1}^n \lambda_i P_i$, where H is finite-dimensional. The space H is decomposed into n orthogonal subspaces, $H = H_1 \oplus H_2 \oplus \cdots \oplus H_n$ ($n = 5$ in this figure), where $H_i = N_{A-\lambda_i I}$ and $P_i \mathbf{x} = \langle \mathbf{x}, \mathbf{x}_i \rangle \mathbf{x}_i \in H_i$.

Example 4.7.

As an illustration of a spectral representation, consider the operator in the example on p. 251, where

$$[A]_{\mathbf{y}} = \sum_{i=1}^2 \lambda_i [P_i] = \sqrt{2} [P_1] - \sqrt{2} [P_2]$$

and

$$\begin{aligned} [P_1] &= [P]^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} [P], \\ [P_2] &= [P]^{-1} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} [P]. \end{aligned}$$

This form is particularly useful for constructing functions of operators and matrices, as shown in Section 4.4.

A form related to (4.26) is easily developed in the finite-dimensional case and also extends to the infinite-dimensional case for compact normal

operators. Assume that $A : H^n \rightarrow H^n$ is a simple operator, i.e., that it possesses an eigenbasis. Then, any $\mathbf{x} \in H^n$ can be written as

$$\mathbf{x} = \sum_{i=1}^n \langle \mathbf{x}, \mathbf{x}_i \rangle \mathbf{x}_i,$$

where $\mathbf{x}_i$ is the eigenvector corresponding to the eigenvalue λ_i , i.e., $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$. With $A\mathbf{x} = \sum_{i=1}^n \langle \mathbf{x}, \mathbf{x}_i \rangle A\mathbf{x}_i = \sum_{i=1}^n \lambda_i \langle \mathbf{x}, \mathbf{x}_i \rangle \mathbf{x}_i$, one may write the operator A as

$$A = \sum_{i=1}^n \lambda_i \langle \cdot, \mathbf{x}_i \rangle \mathbf{x}_i. \quad (4.27)$$

If we identify $P_i : H^n \rightarrow H_i$ as the (projection) operator that takes $\mathbf{x} \in H^n$ into the subspace spanned by $\mathbf{x}_i$, i.e., $P_i \mathbf{x} = \langle \mathbf{x}, \mathbf{x}_i \rangle \mathbf{x}_i$, then the spectral representations (4.26) and (4.27) are seen to be equivalent. However, (4.26) is usually preferred for theoretical constructions. The identity operator is given by

$$I_{n \times n} = \sum_{i=1}^n \langle \cdot, \mathbf{x}_i \rangle \mathbf{x}_i,$$

where it is sometimes convenient to write

$$I_{n \times n} = \sum_{i=1}^n \mathbf{x}_i \bar{\mathbf{x}}_i$$

such that $I_{n \times n} \mathbf{x} = \mathbf{x} = \sum_{i=1}^n \mathbf{x}_i \bar{\mathbf{x}}_i \cdot \mathbf{x}$.

Finally, it is important to make a connection between the eigenvalues and eigenvectors of a linear operator on a finite-dimensional complex space $A : H^n \rightarrow H^n$, and the eigenvalues and eigenvectors of its matrix representation $[a_{ij}]_{\mathbf{y}}$ with respect to some basis $\mathbf{y}$. It can be shown [19, pp. 152–153] that

- the spectrum of an operator and the spectrum of its matrix representation coincide, and
- an element $\mathbf{x} \in H^n$ is an eigenvector of A if and only if the representation of $\mathbf{x}$ with respect to a basis of H^n is an eigenvector of the matrix representation of A with respect to the same basis.

However, for operators acting on infinite-dimensional spaces projected onto a finite-dimensional subspace, eigenvalues of the actual operator may or may not be well approximated by eigenvalues of the resulting matrix, depending on the chosen finite subspace. For example, an operator on an infinite-dimensional space may have no eigenvalues, yet an $n \times n$ matrix, which in some approximate way is supposed to represent the operator, always has n eigenvalues.

4.3.2 Operators on Infinite-Dimensional Spaces

Spectral Expansions

We have already seen that the spectrum of an operator acting on an infinite-dimensional space H may be much more complicated than that of an operator acting on a finite-dimensional space. For the finite-dimensional case, one usually is interested in a simple matrix representation of a given operator. To determine the simplest matrix representation possible one needs to know the eigenvalues and eigenvectors of the operator. Although for the infinite-dimensional case one may investigate matrix representations, it is more often desirable to consider the possibility of representing an element of H in terms of eigenvectors of A . The following theorems provide sufficient conditions for when such representations are possible.

Since operators that are merely compact, or merely self-adjoint, need not possess any eigenvalues, in the general case one needs to concentrate on compact self-adjoint operators to obtain sufficiently strong results. In the case of typical boundary value problems, self-adjointness alone is often sufficient to get similarly strong results because the inverse operator is often compact and self-adjoint.

The first theorem concerns the representation of an element $x \in H$ in terms of the eigenvectors corresponding to nonzero eigenvalues of an operator $A : H \rightarrow H$ [21, pp. 188–191].

Theorem 4.40. (*Hilbert–Schmidt theorem*) Let $A : H \rightarrow H$ be a compact, self-adjoint linear operator acting on an infinite-dimensional Hilbert space H . Then there exists an orthonormal system of eigenvectors $\{u_n\}$ corresponding to nonzero eigenvalues $\{\lambda_n\}$ such that every $x \in H$ can be uniquely represented as

$$x = x_0 + \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n,$$

where x_0 satisfies $Ax_0 = 0$. Furthermore, if $\{\lambda_n\}$ is an infinite set of distinct eigenvalues, then $\lim_{n \rightarrow \infty} \lambda_n = 0$.

Note that the expansion coefficients are the generalized Fourier coefficients $\alpha_n = \langle x, u_n \rangle$. If we denote by S the space spanned by $\{u_n\}$ (then a closed linear subspace of H), by Theorem 2.9 (projection) we see that the described representation is simply a decomposition of x into an element $\sum_{n=1}^{\infty} \alpha_n u_n \in S$ and an element $x_0 \in S^\perp$.

Although the set $\{u_n\}$ does not form a basis for H unless $N_A = \{0\}$, it is clear that it does form a basis for the range of A . For $x \in H$,

$$Ax = \sum_n \lambda_n \langle x, u_n \rangle u_n.$$

²⁴Note that $\{u_n\}$ will also be a basis for H if R_A is dense in H .

Such a set $\{u_n\}$ is called a *basic system* of eigenvectors.

Next we consider the complete set of eigenvectors (including those corresponding to eigenvalues $\lambda_n = 0$), which are found to form a basis for H .

Theorem 4.41. (*Spectral theorem for compact self-adjoint operators*) Let $A : H \rightarrow H$ be a compact, self-adjoint (or more generally, normal) linear operator acting on an infinite-dimensional Hilbert space H . Then there exists an orthonormal basis for H of eigenvectors $\{x_n\}$ with corresponding eigenvalues $\{\lambda_n\}$. For every $x \in H$,

$$x = \sum_n \langle x, x_n \rangle x_n$$

and

$$Ax = \sum_n \lambda_n \langle x, x_n \rangle x_n.$$

To gain insight into the above, let $\{\nu_n\}$ be a basis for the Hilbert space N_A consisting of eigenvectors associated with eigenvalues $\lambda_n = 0$, and note that $\{x_n\} = \{u_n\} \cup \{\nu_n\}$ is an orthonormal basis for H . Then $x_0 = \sum_n \langle x, \nu_n \rangle \nu_n$ where $Ax_0 = 0$ [43, p. 101]. From Theorem 3.17, $N_{A-\lambda_n I}$ is finite-dimensional for $\lambda_n \neq 0$, although N_A will be infinite-dimensional. It can then be observed that the space H is decomposed as $H = N_A \oplus \bigoplus_n N_{A-\lambda_n I}$.

Although Theorem 4.41 is fundamentally important, it is often convenient to use the notation of Theorem 4.40 and explicitly separate out the term x_0 , such that

$$\begin{aligned} x &= x_0 + \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n, \\ Ax &= \sum_{n=1}^{\infty} \lambda_n \langle x, u_n \rangle u_n, \end{aligned} \tag{4.28}$$

where $Ax_0 = 0$. In this form we are concerned only with the basic system of eigenvectors u_n , corresponding to nonzero eigenvalues. As usual, all equalities are understood in the norm sense.

It is worthwhile to note that the converse of the spectral theorem is also true [23, p. 117].

Theorem 4.42. Let $\{x_n\}$ be an orthonormal set in H and $\{\lambda_n\}$ a sequence of real numbers that is either finite or that converges to zero. The linear operator defined by $Ax \equiv \sum_n \lambda_n \langle x, x_n \rangle x_n$ is compact and self-adjoint.

Because compact operators $A : H \rightarrow H$, where H is infinite-dimensional, are necessarily bounded, they can always be represented by an infinite matrix. From the properties of compact operators, and those of self-adjoint operators, we see that the matrix representation of a compact,

self-adjoint operator $A : H \rightarrow H$ with respect to the eigenbasis $\{x_n\}$ is a diagonal matrix with real-valued eigenvalues on the main diagonal, tending towards zero as $n \rightarrow \infty$. Furthermore, it can be shown that two commuting compact, self-adjoint operators $A, B : H \rightarrow H$ possess a common eigenbasis (see Theorem 3.12).

Finally, if an operator $A : H \rightarrow H$ is compact but not self-adjoint, we have the following expansion theorem.

Theorem 4.43. *Let $A : H \rightarrow H$ be a compact operator. Then, there exist orthonormal sets $\{x_n\}$ and $\{y_n\}$ and nonnegative real numbers $\{s_n\}$ with $s_n \rightarrow 0$ such that*

$$Ax = \sum_n s_n \langle x, x_n \rangle y_n.$$

It turns out that $\{x_n\}$ are eigenvectors of A^*A , $\{y_n\}$ are eigenvectors of AA^* , and $\{s_n^2\}$ are eigenvalues of A^*A (and also of AA^*). The s_n are called *singular values* of the operator A , and the series Ax is known as a *singular-value decomposition* of A . The foundation of this representation is that while A is compact and not necessarily normal, A^*A and AA^* are compact and self-adjoint by Theorem 3.8 and the discussion in Section 3.6.1.

Jordan Chain

Unfortunately, while for operators acting on finite-dimensional spaces one can always obtain a basis of generalized eigenvectors (Jordan basis) by Theorem 4.39, the same is not true for the infinite-dimensional case. However, if generalized eigenvectors exist, then they are defined as in the finite-dimensional case, i.e., an element $0 \neq x_{n,m-1} \in H$ is a generalized eigenvector (root vector) of rank m of the operator $A : H \rightarrow H$ corresponding to an eigenvalue λ_n if

$$\begin{aligned} (A - \lambda_n I)^m x_{n,m-1} &= 0, \\ (A - \lambda_n I)^{m-1} x_{n,m-1} &\neq 0, \end{aligned} \tag{4.29}$$

where m is a positive integer. Every (ordinary) eigenvector of A is a generalized eigenvector of rank 1 ($x_{n,0} \equiv x_n$), and we again call the generalized eigenvectors having rank $m > 1$ associated eigenvectors to distinguish them from ordinary eigenvectors. The root system of A is defined as the union of the ordinary and associated eigenvectors. This system may or may not form a basis for H , depending on the operator²⁵.

In practice, to determine the associated eigenvectors, one starts with an eigenvector x_n satisfying $(A - \lambda_n I)x_n = 0$. If the equation

$$(A - \lambda_n I)x_{n,1} = x_n$$

²⁵ An example of a differential operator that leads to a root basis is provided in Section 5.3.2.

has a solution $x_{n,1}$, then $x_{n,1}$ is a generalized eigenvector of rank 2; more specifically, an associated eigenvector associated with the eigenvalue λ_n and eigenvector x_n . Continuing, if

$$(A - \lambda_n I) x_{n,2} = x_{n,1}$$

is solvable, then $x_{n,2}$ is another associated eigenvector (rank 3), associated with the eigenvalue λ_n and eigenvector x_n . In general, we consider

$$(A - \lambda_n I) x_{n,k} = x_{n,k-1}$$

such that the chain $\{x_n, x_{n,1}, x_{n,2}, \dots, x_{n,j}\}$ consisting of the ordinary and associated eigenvectors is called a Jordan chain of length $j + 1$.

It is worthwhile to note that if generalized eigenvectors of a self-adjoint operator exist, then they have rank 1 (i.e., no associated eigenvectors exist).²⁶

4.3.3 Spectral Expansions Associated with Boundary Value Problems

Theorems in the previous section detailed sufficient conditions for eigenfunctions of an operator to form a basis in H . Fortunately, many operators that do not satisfy the previous theorems also lead to an eigenbasis. Important examples are unbounded operators corresponding to self-adjoint boundary value problems. We make the following observation.

The eigenfunctions of a self-adjoint boundary value problem on a Hilbert space H form an orthonormal basis of H .

The term “boundary value problem” in common usage typically means a differential operator together with some specified boundary conditions. Since the boundary conditions define (in part) the domain of the operator, they are part of the definition of the operator, such that by “self-adjoint boundary value problem” we actually refer to a self-adjoint differential operator, which we will denote by L . The above classification of self-adjoint boundary value problems relies on the fact that such operators have a compact self-adjoint inverse (integral) operator L^{-1} on H , the Green’s operator, with an associated orthonormal eigenbasis by Theorem 4.41. By Theorem 4.11 the eigenfunctions of the inverse operator L^{-1} are also eigenfunctions of L (zero eigenvalues are not an issue), and therefore L possesses an eigenbasis for H .

²⁶ Assume $A : H \rightarrow H$ is self-adjoint, with x an associated eigenvector of rank 2 corresponding to eigenvalue λ . Then, $0 = \langle x, (A - \lambda I)^2 x \rangle = \langle (A - \lambda I) x, (A - \lambda I) x \rangle = \|(A - \lambda I) x\|^2$, which implies, from Definition 2.31, that $(A - \lambda I) x = 0$ which is a contradiction, and so no generalized eigenvectors of rank 2 exist. Therefore, no generalized eigenvectors of rank higher than 1 exist.

The above conclusion holds for problems on a bounded region of space, $\Omega \subset \mathbf{R}^n$. If Ω is unbounded the inverse will not, in general, be compact, and Theorem 4.41 is not applicable. However, if the differential operator is self-adjoint and certain fitness conditions are satisfied at infinity, then the eigenfunctions (proper and improper) form an eigenbasis (possibly discrete and continuous) for H . This is discussed in more detail in later sections (see, e.g., Sections 5.2 and 8.1).

In particular, we have the following [4, pp. 396–397].

Theorem 4.44. *The p th-order differential operator $L: \mathbf{L}_w^2(a, b) \rightarrow \mathbf{L}_w^2(a, b)$ defined by*

$$-L \equiv a_p \frac{d^p}{dt^p} + \cdots + a_1 \frac{d}{dt} + a_0,$$

$$D_L \equiv \left\{ x: \frac{d^j x}{dt^j} \in \mathbf{L}_w^2(a, b) \text{ for } j = 1, 2, \dots, p, B_a(x) = B_b(x) = 0 \right\},$$

where a, b are finite, p is even, $a_n \in \mathbf{C}^p(a, b)$ are real-valued coefficients with $a_p \neq 0$, and where $B_{a,b}$ are given as

$$B_a(x) = \alpha_1 x(a) + \alpha_2 x'(a) + \cdots + \alpha_p x^{(p-1)}(a) = 0,$$

$$B_b(x) = \beta_1 x(b) + \beta_2 x'(b) + \cdots + \beta_p x^{(p-1)}(b) = 0,$$

with $\alpha = (\alpha_1, \dots, \alpha_p)$ and $\beta = (\beta_1, \dots, \beta_p)$ independent vectors in $\mathbf{R}^p$, is self-adjoint (and so forms a self-adjoint boundary value problem) on $H = \mathbf{L}_w^2(a, b)$.

Therefore, the eigenvalue problem

$$Lx_n = \lambda_n x_n$$

with L defined as in Theorem 4.44 leads to an orthonormal eigenbasis of H and to the expansion $x = \sum_n \langle x, x_n \rangle x_n$ for any $x(t) \in H = \mathbf{L}_w^2(a, b)$. In the above, H is a weighted $\mathbf{L}^2$ space, where the weight w (possibly $w=1$) depends on the coefficients and is chosen so that the operator is self-adjoint. The inner product for the space is

$$\langle f, g \rangle = \int_a^b f(\xi)g(\xi)w(\xi) d\xi$$

Second-order operators of this type are extensively studied, where self-adjointness is shown and a sketch of the compactness is presented.

For scalar partial differential operators we have the following [4, pp. 542–543] (see Example 8 in Section 3.4.2).

Theorem 4.45. *The negative Laplacian operator $L : \mathbf{L}^2(\Omega) \rightarrow \mathbf{L}^2(\Omega)$ defined by*

$$(Lx)(\mathbf{t}) \equiv -\nabla^2 x(\mathbf{t}),$$

$$D_L \equiv \{x : x(\mathbf{t}), -\nabla^2 x(\mathbf{t}) \in \mathbf{L}^2(\Omega), x(\mathbf{t})|_{\Gamma} = 0\},$$

where Γ is the sufficiently smooth boundary of the region $\Omega \subset \mathbf{R}^3$, forms a self-adjoint boundary value problem on $\mathbf{L}^2(\Omega)$.

As such, the orthonormal eigenbasis defined by

$$-\nabla^2 x_n = \lambda_n x_n$$

leads to the expansion $x = \sum_n \langle x, x_n \rangle x_n$ for any $x(\mathbf{t}) \in H = \mathbf{L}^2(\Omega)$. This topic is discussed further in Section 6.3 for electrostatic problems, where self-adjointness is shown and the completeness proof outlined. Additionally, these eigenfunctions are important in separation-of-variables solutions for many partial differential equations, as shown in Section 5.4.

Furthermore, for the vector differential operator $\nabla \times \nabla \times$ we have a similar theorem (see Example 9, Section 3.4.2).

Theorem 4.46. *The vector differential operator $L : \mathbf{L}^2(\Omega)^3 \rightarrow \mathbf{L}^2(\Omega)^3$ defined by*

$$(L\mathbf{x})(\mathbf{t}) \equiv \nabla \times \nabla \times \mathbf{x}(\mathbf{t}),$$

$$D_L \equiv \left\{ \mathbf{x} : \mathbf{x}(\mathbf{t}), \nabla \times \nabla \times \mathbf{x}(\mathbf{t}) \in \mathbf{L}^2(\Omega)^3, \mathbf{n} \times \mathbf{x}(\mathbf{t})|_{\Gamma} = \mathbf{0} \right\},$$

where Γ is the sufficiently smooth boundary of the region $\Omega \subset \mathbf{R}^3$, forms a self-adjoint boundary value problem on $\mathbf{L}^2(\Omega)^3$.

Therefore, the orthonormal eigenbasis defined by

$$\nabla \times \nabla \times \mathbf{x}_n = \lambda_n \mathbf{x}_n$$

with appropriate vector boundary conditions leads to the expansion $\mathbf{x} = \sum_n \langle \mathbf{x}, \mathbf{x}_n \rangle \mathbf{x}_n$ for any $\mathbf{x}(\mathbf{t}) \in H = \mathbf{L}^2(\Omega)^3$. The same result holds for the vector eigenvalue problem

$$\nabla^2 \mathbf{x}_n = \lambda_n \mathbf{x}_n.$$

The situation is quite different if Ω is an unbounded region of space.

Because we deal with vector eigenfunctions, boundary conditions generally need to be specified for both tangential and normal field components to uniquely identify the desired eigenfunctions. For example, vector Dirichlet conditions are $\mathbf{x}_n|_{\Gamma} = \mathbf{0}$, although the boundary conditions $\mathbf{n} \times \mathbf{x}_n|_{\Gamma} = \mathbf{0}$ and $\nabla \cdot \mathbf{x}_n|_{\Gamma} = 0$ are generally more useful in electromagnetics applications

(see Chapters 9 and 10). Boundary conditions do not need to be imposed on all three scalar components of $\mathbf{x}_n$ if the eigenfunction is required to satisfy homogeneous Maxwell's equations. In such cases it is only necessary to impose tangential conditions.

The above theorems concerning self-adjoint differential operators are extremely important because they justify many eigenfunction expansion methods used in electromagnetic applications.

4.3.4 Spectral Expansions Associated with Integral Operators

We now consider some spectral expansions arising from integral operators involving the free-space scalar Green's function. For all of the resulting eigenfunctions, completeness in certain spaces is well known and can be established using nonspectral techniques. Therefore, as with the boundary value problems just considered, identification of the (integral) operator as being compact and self-adjoint is not required, although some of the operators fall into this category. Many of the same expansions are obtained from a differential operator (Sturm–Liouville) approach in Section 5.5.

Chebyshev Polynomials of the First Kind

Consider the weighted Hilbert space $\mathbf{L}_w^2(-a, a)$ with weight $w = (a^2 - x^2)^{-\frac{1}{2}}$ and inner product

$$\langle f, g \rangle = \int_{-a}^a f(x)g(x) \frac{dx}{\sqrt{a^2 - x^2}}.$$

The compact, self-adjoint integral operator $A : \mathbf{L}_w^2(-a, a) \rightarrow \mathbf{L}_w^2(-a, a)$ defined by²⁷

$$(Af)(x) \equiv \int_{-a}^a f(x') \frac{\ln(|x - x'|)}{\sqrt{a^2 - x'^2}} dx'$$

arises, for example, in the study of static and quasi-static electromagnetic diffraction by a strip or by a slot in a perfectly conducting infinite screen. For two-dimensional dynamic problems these types of logarithmic kernels represent the source-point singularity associated with the two-dimensional principal Green's function (Hankel function). The eigenvalue problem

$$Af = \lambda f$$

leads to *first-kind Chebyshev polynomials* $T_n(x/a)$ as eigenfunctions, with corresponding eigenvalues $\lambda_0 = -\pi \ln(2/a)$ for $n = 0$, and $\lambda_n = -\pi/n$ for

²⁷ $\ln(|x - \xi|)$ is the natural restriction of the two-dimensional, static free-space Green's function $\ln(|\mathbf{r} - \xi'|)$ to a planar surface.

$n \neq 0$ (note $N_A = \{0\}$ if $a \neq 2$), i.e.,

$$\int_{-a}^a \frac{\ln|x-x'|}{\sqrt{a^2-x'^2}} T_n(x'/a) dx' = \begin{cases} -\pi \ln(2/a) T_0(x/a), & n=0, \\ -\frac{\pi}{n} T_n(x/a), & n>0. \end{cases} \quad (4.30)$$

The eigenfunctions are normalized as

$$\frac{\varepsilon_n}{\pi} \langle T_n, T_m \rangle = \frac{\varepsilon_n}{\pi} \int_{-a}^a T_n(x/a) T_m(x/a) \frac{1}{\sqrt{a^2-x^2}} dx = \delta_{nm}, \quad (4.31)$$

where $\varepsilon_0 = 1$ and $\varepsilon_n = 2$ for $n \neq 0$, and the set of Chebyshev polynomials $\left\{ \sqrt{\varepsilon_n/\pi} T_n(x/a) \right\}$, $n = 0, 1, 2, \dots$, forms a basis in $\mathbf{L}_w^2(-a, a)$. Therefore, any function $f(x) \in \mathbf{L}_w^2(-a, a)$ can be expanded as

$$f = \frac{1}{\pi} \sum_{n=0}^{\infty} \varepsilon_n \langle f, T_n \rangle T_n$$

and

$$Af = \frac{1}{\pi} \sum_{n=0}^{\infty} \varepsilon_n \lambda_n \langle f, T_n \rangle T_n.$$

For example, in $\mathbf{L}_w^2(-a, a)$ the equation

$$Ag = f$$

can be solved by writing $f = (1/\pi) \sum_{n=0}^{\infty} \varepsilon_n \langle f, T_n \rangle T_n$, leading to

$$g = \frac{1}{\pi} \sum_{n=0}^{\infty} \frac{\varepsilon_n}{\lambda_n} \langle f, T_n \rangle T_n. \quad (4.32)$$

Convergence in the form $\lim_{N \rightarrow \infty} \|g - g_N\| = 0$, where

$$g_N \equiv (1/\pi) \sum_{n=0}^N (\varepsilon_n/\lambda_n) \langle f, T_n \rangle T_n,$$

depends on the condition $\sum_n (1/\lambda_n^2) |\langle f, T_n \rangle|^2 < \infty$ by the Riesz–Fischer theorem (Theorem 2.11). Since $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$, which is a general property of compact operators, convergence problems may occur depending on the properties of f . This type of problem is typical of first-kind operator equations involving compact operators.

Chebyshev Polynomials of the Second Kind

Consider the weighted Hilbert space $\mathbf{L}_w^2(-a, a)$ with weight $w = (a^2 - x^2)^{\frac{1}{2}}$ and inner product

$$\langle f, g \rangle = \int_{-a}^a f(x)g(x) \sqrt{a^2 - x^2} dx.$$

As with the previous case, the operator $A : \mathbf{L}_w^2(-a, a) \rightarrow \mathbf{L}_w^2(-a, a)$ defined by

$$(Af)(x) \equiv \frac{d^2}{dx^2} \int_{-a}^a f(x') \ln(|x - x'|) \sqrt{a^2 - x'^2} dx'$$

arises in the study of strip and slot diffraction. The eigenvalue problem

$$Af = \lambda f$$

leads to *second-kind Chebyshev polynomials* $U_n(x/a)$ as eigenfunctions, with corresponding eigenvalues $\lambda_n = \pi(n+1)$, i.e.,

$$\frac{d^2}{dx^2} \int_{-a}^a U_n(x'/a) \ln|x - x'| \sqrt{a^2 - x'^2} dx' = \pi(n+1) U_n(x/a).$$

The eigenfunctions are normalized as

$$\frac{2}{\pi a^2} \langle U_n, U_m \rangle = \frac{2}{\pi a^2} \int_{-a}^a U_n(x/a) U_m(x/a) \sqrt{a^2 - x^2} dx = \delta_{nm},$$

and the set of Chebyshev polynomials $\left\{ \sqrt{2/(\pi a^2)} U_n(x/a) \right\}$, $n = 0, 1, 2, \dots$, forms a basis in $\mathbf{L}_w^2(-a, a)$. Therefore, any function $f(x) \in \mathbf{L}_w^2(-a, a)$ can be expanded as

$$f = \frac{2}{\pi a^2} \sum_{n=0}^{\infty} \langle f, U_n \rangle U_n$$

and

$$Af = \frac{2}{\pi a^2} \sum_{n=0}^{\infty} \lambda_n \langle f, U_n \rangle U_n.$$

Laguerre Polynomials

The weighted Hilbert space $\mathbf{L}_w^2(0, \infty)$ with weight $e^{-x}/\sqrt{x}$ and inner product

$$\langle f, g \rangle = \int_0^{\infty} f(x) g(x) \frac{1}{\sqrt{x}} e^{-x} dx$$

along with the integral operator $A : \mathbf{L}_w^2(0, \infty) \rightarrow \mathbf{L}_w^2(0, \infty)$ defined by

$$(Af)(x) \equiv \int_0^{\infty} f(x') K_0(|x - x'|) \frac{e^{-x'}}{\sqrt{x'}} dx'$$

is encountered in the study of E-polarized wave diffraction by [24]. The kernel

$$K_0(x) = -\frac{i\pi}{2} H_0^{(2)}(-ix) = \frac{i\pi}{2} H_0^{(1)}(ix)$$

is the modified Bessel function [25], sometimes called the *MacDonald function*. In this case we have the generalized eigenvalue problem

$$Af = \lambda e^{-x} f.$$

which leads to $L_n^{-1/2}(2x)$ as eigenfunctions with $\lambda_n = (\pi/\sqrt{2})\gamma_n^E$ as the corresponding eigenvalues, i.e.,

$$\int_0^\infty L_n^{-1/2}(2x') K_0(|x - x'|) \frac{e^{-x'}}{\sqrt{x'}} dx' = \frac{\pi}{\sqrt{2}} \gamma_n^E e^{-x} L_n^{-1/2}(2x).$$

The functions L_n^v are *associated Laguerre polynomials* ($v > -1$, and for $v = 0$, $L_n^0 = L_n$ are called *Laguerre polynomials*), and

$$\gamma_n^E = \Gamma(n + 1/2) / \Gamma(n + 1),$$

where Γ is the gamma-function. The eigenfunctions are normalized as

$$\frac{1}{\gamma_n^E} \langle L_n^{-1/2}, L_m^{-1/2} \rangle = \frac{1}{\gamma_n^E} \int_0^\infty L_n^{-1/2}(x) L_m^{-1/2}(x) \frac{e^{-x}}{\sqrt{x}} dx = \delta_{nm}.$$

and the set $\{\sqrt{1/\gamma_n^E} L_n^{-1/2}(x)\}$, $n = 0, 1, 2, \dots$, forms a basis in $L_w^2(0, \infty)$. Therefore, any function $f(x) \in L_w^2(0, \infty)$ can be expanded as

$$f = \frac{1}{\gamma_n^E} \sum_{n=0}^\infty \langle f, L_n^{-1/2} \rangle L_n^{-1/2}.$$

In a similar manner, the operator $A : L_w^2(0, \infty) \rightarrow L_w^2(0, \infty)$ defined by

$$(Af)(x) \equiv \left(\frac{d^2}{dx^2} - 1 \right) \int_0^\infty f(x') K_0(|x - x'|) \sqrt{x'} e^{-x'} dx'$$

arises in the study of H-polarized wave diffraction by a half-space [24], where the weight is $\sqrt{x}e^{-x}$ with the inner product

$$(f, g) = \int_0^\infty f(x)g(x)\sqrt{x}e^{-x}dx.$$

The generalized eigenvalue problem

$$Af = \lambda e^{-x} f$$

leads to $L_n^{1/2}(2x)$ as eigenfunctions with $\lambda_n = -\pi\sqrt{2}\gamma_n^H$ as the corresponding eigenvalues, i.e.,

$$\left(\frac{d^2}{dx^2} - 1 \right) \int_0^\infty L_n^{1/2}(2x') K_0(|x - x'|) \sqrt{x'} e^{-x'} dx' = -\pi\sqrt{2}\gamma_n^H e^{-x} L_n^{1/2}(2x)$$

where $\gamma_n^H = \Gamma(n + 3/2) / \Gamma(n + 1)$. The eigenfunctions are normalized as

$$\frac{1}{\gamma_n^H} \langle L_n^{1/2}, L_m^{1/2} \rangle = \frac{1}{\gamma_n^H} \int_0^\infty L_n^{1/2}(x) L_m^{1/2}(x) \sqrt{x} e^{-x} dx = \delta_{nm}$$

and the set $\left\{ \sqrt{1/\gamma_n^H} L_n^{1/2}(x) \right\}$, $n = 0, 1, 2, \dots$, forms a basis in $\mathbf{L}_w^2(0, \infty)$. Therefore, any function $f(x) \in \mathbf{L}_w^2(0, \infty)$ can be expanded as

$$f = \frac{1}{\gamma_n^H} \sum_{n=0}^{\infty} \langle f, L_n^{1/2} \rangle L_n^{1/2}.$$

Expansions involving Chebyshev and Laguerre polynomials are further considered in Section 5.5.

Circular Exponential Functions

Consider the Hilbert space $\mathbf{L}^2(0, 2\pi)$ with inner product

$$\langle f, g \rangle = \int_0^{2\pi} f(\phi) \overline{g(\phi)} d\phi.$$

The integral operator $A : \mathbf{L}^2(0, 2\pi) \rightarrow \mathbf{L}^2(0, 2\pi)$ defined by²⁸

$$(Af)(\phi) \equiv \int_0^{2\pi} f(\phi') H_0^{(1,2)} \left(2ka \sin \frac{|\phi - \phi'|}{2} \right) d\phi', \quad (4.33)$$

where $H_0^{(1,2)}$ is the zeroeth-order Hankel function (see Section 5.4.1), arises in the study of scattering by circular cross-section cylinders. The eigenvalue problem

$$Af = \lambda f$$

leads to complex exponentials $e^{in\phi}$ as eigenfunctions, $n = 0, \pm 1, \pm 2, \dots$, with corresponding eigenvalues $\lambda_n = 2\pi J_n(ka) H_n^{(1,2)}(ka)$, i.e.,

$$\int_0^{2\pi} e^{in\phi'} H_0^{(1,2)} \left(2ka \sin \frac{|\phi - \phi'|}{2} \right) d\phi' = 2\pi J_n(ka) H_n^{(1,2)}(ka) e^{in\phi}.$$

The eigenfunctions are normalized as

$$\frac{1}{2\pi} \langle e^{in\phi}, e^{im\phi} \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{in\phi} e^{-im\phi} d\phi = \delta_{nm},$$

²⁸ $H_0^{(1,2)}(2ka \sin(|\phi - \xi|/2))$ is the natural restriction of the two-dimensional, dynamic Green's function $H_0^{(1,2)}(k|\mathbf{r} - \xi|)$ to the surface of a circular cylinder having radius a . We assume $k \in \mathbf{R}$.

and form the well known basis $\{(1/\sqrt{2\pi})e^{in\phi}\}$, $n = 0, \pm 1, \pm 2, \dots$, in $\mathbf{L}^2(0, 2\pi)$. Any function $f(\phi) \in \mathbf{L}^2(0, 2\pi)$ can be expanded as

$$f = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \langle f, e^{in\phi} \rangle e^{in\phi}$$

and

$$Af = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \lambda_n \langle f, e^{in\phi} \rangle e^{in\phi}.$$

Spherical Harmonics

Consider the Hilbert space $\mathbf{L}^2(\Gamma)$, where Γ is a spherical surface having radius a , with inner product

$$\langle f, g \rangle = \int_0^\pi \int_0^{2\pi} f(\theta, \phi) \overline{g(\theta, \phi)} a^2 \sin \theta d\theta d\phi.$$

The integral operator $A : \mathbf{L}^2(\Gamma) \rightarrow \mathbf{L}^2(\Gamma)$ defined by

$$(Af)(\theta, \phi) \equiv \int_0^\pi \int_0^{2\pi} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} f(\theta', \phi') a^2 \sin \theta' d\theta' d\phi' \quad (4.34)$$

for $\mathbf{r}|_\Gamma$ arises in the study of scattering by spheres, where we assume $k \in \mathbf{R}$. The eigenvalue problem

$$Af = \lambda f$$

leads to spherical harmonics $Y_{n,m}(\theta, \phi)$ (see Section 5.4.2 for more details on spherical harmonics) as eigenfunctions,²⁹ with corresponding eigenvalues

²⁹From the expansion (5.172) we have

$$\begin{aligned} & \int_0^\pi \int_0^{2\pi} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} Y_{p,q}(\theta', \phi') a^2 \sin \theta' d\theta' d\phi' \\ &= \int_0^\pi \int_0^{2\pi} (-ik) \sum_{n=0}^{\infty} j_n(ka) h_n^{(2)}(ka) \sum_{m=-n}^n Y_{n,m}(\theta, \phi) \overline{Y_{n,m}(\theta', \phi')} \\ & \quad \cdot Y_{p,q}(\theta', \phi') a^2 \sin \theta' d\theta' d\phi' \\ &= (-ika^2) \sum_{n=0}^{\infty} j_n(ka) h_n^{(2)}(ka) \sum_{m=-n}^n Y_{n,m}(\theta, \phi) \int_0^\pi \int_0^{2\pi} \overline{Y_{n,m}(\theta', \phi')} \\ & \quad \cdot Y_{p,q}(\theta', \phi') \sin \theta' d\theta' d\phi' \\ &= (-ika^2) \sum_{n=0}^{\infty} j_n(ka) h_n^{(2)}(ka) \sum_{m=-n}^n Y_{n,m}(\theta, \phi) \delta_{np} \delta_{mq} \\ &= (-ika^2) j_p(ka) h_p^{(2)}(ka) Y_{p,q}(\theta, \phi), \end{aligned}$$

where the orthonormality relation (5.164) has been used.

$\lambda_n = -ik j_n(ka) h_n^{(2)}(ka)$, i.e.,

$$\int_0^\pi \int_0^{2\pi} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \bigg|_{\mathbf{r} \in \Gamma} Y_{n,m}(\theta', \phi') \sin \theta' d\theta' d\phi' \\ = (-ik) j_n(ka) h_n^{(2)}(ka) Y_{n,m}(\theta, \phi),$$

where $m = \{\dots, -2, -1, 0, 1, 2, \dots\}$ and $n \geq m$. The eigenfunctions are normalized as

$$\langle Y_{n,m}, Y_{n',m'} \rangle = \int_0^{2\pi} \int_0^\pi Y_{n,m}(\theta, \phi) \bar{Y}_{n',m'}(\theta, \phi) a^2 \sin \theta d\theta d\phi \\ = a^2 \delta_{nn'} \delta_{mm'}, \quad (4.35)$$

forming the well known basis $\{Y_{n,m}(\theta, \phi)\}$ in $\mathbf{L}^2(\Gamma)$.

Any function $f(\theta, \phi)$ square-integrable on a sphere of radius a can be expanded as

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \langle f, Y_{n,m} \rangle Y_{n,m}(\theta, \phi)$$

and

$$Af = \sum_{n=0}^{\infty} \sum_{m=-n}^n \lambda_n \langle f, Y_{n,m} \rangle Y_{n,m}(\theta, \phi).$$

4.3.5 Generalized Eigenvectors and the Root System

Considering operators on infinite-dimensional Hilbert spaces, Theorem 4.41 provides sufficient conditions under which operators admit simple eigenfunction representations (i.e., the existence of an eigenbasis is assured). In Sections 4.3.3 and 4.3.4, differential and integral operators are considered that also lead to an eigenbasis, even though they do not necessarily satisfy the conditions of Theorem 4.41.

If the eigenfunctions do not form a basis for the space in question, it may happen that the set of generalized eigenfunctions (i.e., the root system) forms a basis in H . Further, if the root system is a basis, then one can find a set orthogonal to the root system which is also a basis, and a generalized expansion in terms of the resulting bi-orthogonal set will be valid. This is a rather difficult subject (see, e.g., [26] and [27]), and here we mention only a few results, with particular attention to the compact operator $A : \mathbf{L}^2(\Gamma) \rightarrow \mathbf{L}^2(\Gamma)$ defined by

$$(Af)(\mathbf{r}) \equiv \int_{\Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} f(\mathbf{r}') d\Gamma', \quad (4.36)$$

where Γ is a smooth closed surface in $\mathbf{R}^3$ and $k > 0$. This topic is also treated briefly for the case of a second-order differential operator in Section 5.3.2.

The following theorem is proved in [26] (see also [28]).

Theorem 4.47. *Let $Q : H \rightarrow H$ be a nonnegative compact operator, and let $D : H \rightarrow H$ be a dissipative nuclear³⁰ operator. Then the system of root vectors of the operator $B = Q + D$ is complete in H .*

In particular, this theorem applies to the operator adjoint to (4.36), where

$$(Qf)(\mathbf{r}) = \int_{\Gamma} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} f(\mathbf{r}') d\Gamma',$$

$$(Df)(\mathbf{r}) = \int_{\Gamma} \frac{(e^{+ik|\mathbf{r}-\mathbf{r}'|} - 1)}{4\pi |\mathbf{r} - \mathbf{r}'|} f(\mathbf{r}') d\Gamma',$$

so that the root vectors of the operator adjoint to (4.36) form a basis in $L^2(\Gamma)$.

Furthermore, when a compact operator $A : H \rightarrow H$ is normal, the root system coincides with the system of ordinary eigenfunctions [28] (see also [29]), that is, all generalized eigenfunctions are of rank 1 (this result is obviously in harmony with Theorem 4.41). While the operator in (4.36) is clearly not self-adjoint (unless $k = 0$ or $ik \in \mathbf{R}$), it is normal under certain circumstances. Indeed, because for A to be normal we must have $AA^* - A^*A = 0$, this condition applied to (4.36) yields

$$\begin{aligned} (AA^*f)(\mathbf{r}) - (A^*Af)(\mathbf{r}) &= \int_{\Gamma} \int_{\Gamma} \frac{(e^{-ik(|\mathbf{r}-\mathbf{r}''|-|\mathbf{r}''-\mathbf{r}'|)} - e^{+ik(|\mathbf{r}-\mathbf{r}''|-|\mathbf{r}''-\mathbf{r}'|)})}{(4\pi)^2 |\mathbf{r} - \mathbf{r}''| |\mathbf{r}'' - \mathbf{r}'|} f(\mathbf{r}') d\Gamma' d\Gamma'' \\ &= \int_{\Gamma} \int_{\Gamma} \frac{-\sin k(|\mathbf{r} - \mathbf{r}''| - |\mathbf{r}'' - \mathbf{r}'|)}{2i(4\pi)^2 |\mathbf{r} - \mathbf{r}''| |\mathbf{r}'' - \mathbf{r}'|} f(\mathbf{r}') d\Gamma' d\Gamma'' = 0. \end{aligned}$$

Therefore, we obtain

$$\int_{\Gamma} \frac{-\sin k(|\mathbf{r} - \mathbf{r}''| - |\mathbf{r}' - \mathbf{r}'|)}{2i(4\pi)^2 |\mathbf{r} - \mathbf{r}''| |\mathbf{r}'' - \mathbf{r}'|} f(\mathbf{r}') d\Gamma' = 0 \quad (4.37)$$

for $\mathbf{r}, \mathbf{r}'' \in \Gamma$, which is a condition on the surface Γ . This condition is satisfied if Γ is a spherical surface, but not, for instance, if Γ is an ellipsoid.

We also have the following theorem [30], which establishes conditions under which all generalized eigenfunctions of a compact operator A corresponding to an eigenvalue λ have rank 1.

Theorem 4.48. *The eigenspace and the root space of a compact operator A , corresponding to an eigenvalue λ , coincide if and only if*

$$a. \lambda \text{ is a simple pole of the resolvent } (A - \lambda I)^{-1}, \text{ or}$$

³⁰An operator D is called *nuclear* if $\sum s_n < \infty$, where s_n are the eigenvalues of $(D^*D)^{1/2}$.

b. $(A - \lambda I)^2 f = 0$ implies $(A - \lambda I) f = 0$, or

c. the operator $(A - \lambda I)$ does not have zeros in the subspace $R_{A - \lambda I}$.

It should be noted that it is also possible to show that eigenfunctions of an operator form an eigenbasis using other methods. For instance, it is shown in [31, p. 774] that sufficient conditions for an operator to possess a complete set of eigenfunctions in an $\mathbf{L}^2$ space are that the operator is positive definite, and hence self-adjoint, and that the eigenvalue equation $Lu = \lambda u$ corresponds to a variational principle. Determining completeness of the eigenfunctions for nonself-adjoint operators is, in general, much more difficult, although some further results can be stated for dissipative operators [26, Ch. 6]. For our purposes we will call an operator *simple* if it possesses an orthonormal eigenbasis of ordinary eigenfunctions.

4.3.6 Spectral Representations

For a compact, self-adjoint operator $A : H \rightarrow H$, or, in general, whenever the ordinary eigenfunctions of $A : H \rightarrow H$ form an orthonormal basis for H , one may develop a spectral expansion of the operator in terms of eigenfunctions. From Theorem 4.41 with $x = \sum_n \langle x, x_n \rangle x_n$, we identify the form analogous to (4.27),

$$A = \sum_{n=1}^{\infty} \lambda_n \langle \cdot, x_n \rangle x_n. \quad (4.38)$$

Note that $I = \sum_{n=1}^{\infty} \langle \cdot, x_n \rangle x_n$ and that on a function space where $x = x(t)$ one may write, in a distributional sense,³¹

$$\delta(t - t') = \sum_{n=1}^{\infty} x_n(t) \bar{x}_n(t'). \quad (4.39)$$

31

$$\begin{aligned} x(t) &= \int \delta(t - t') x(t') dt' = \int \sum_{n=1}^{\infty} x_n(t) \bar{x}_n(t') x(t') dt' \\ &= \sum_{n=1}^{\infty} x_n(t) \int \bar{x}_n(t') x(t') dt' = \sum_{n=1}^{\infty} x_n \langle x, x_n \rangle. \end{aligned}$$

In the event of a function space of vectors, $\mathbf{I} \delta(\mathbf{t} - \mathbf{t}') = \sum_{n=1}^{\infty} \mathbf{x}_n(\mathbf{t}) \bar{\mathbf{x}}_n(\mathbf{t}')$, and

$$\begin{aligned} \mathbf{x}(\mathbf{t}) &= \int \mathbf{I} \delta(\mathbf{t} - \mathbf{t}') \cdot \mathbf{x}(\mathbf{t}') d\mathbf{t}' = \int \sum_{n=1}^{\infty} \mathbf{x}_n(\mathbf{t}) \bar{\mathbf{x}}_n(\mathbf{t}') \cdot \mathbf{x}(\mathbf{t}') d\mathbf{t}' \\ &= \sum_{n=1}^{\infty} \mathbf{x}_n(\mathbf{t}) \int \bar{\mathbf{x}}_n(\mathbf{t}') \cdot \mathbf{x}(\mathbf{t}') d\mathbf{t}' = \sum_{n=1}^{\infty} \mathbf{x}_n \langle \mathbf{x}, \mathbf{x}_n \rangle. \end{aligned}$$

If $P_n : H \rightarrow H_n$ is the projection operator that takes $x \in H$ into the one-dimensional subspace spanned by the n th eigenfunction x_n (i.e., $P_n x = \langle x, x_n \rangle x_n$, and $P_n P_m = \delta_{nm}$), then the spectral representation is seen to have the same weighted sum of projections form as (4.26),

$$A = \sum_{n=1}^{\infty} \lambda_n P_n \quad (4.40)$$

with

$$x = \sum_{n=1}^{\infty} P_n x, \quad Ax = \sum_{n=1}^{\infty} \lambda_n P_n x, \quad I = \sum_{n=1}^{\infty} P_n.$$

If we let the indices n_j be those values of n that correspond to the eigenfunctions $\nu_n = \nu_{n_j}$ (corresponding to eigenvalues $\lambda_{n_j} = 0$), then $\sum_{n_j=1}^{\infty} P_{n_j} x = P_0 x = x_0$, and we have

$$x = P_0 x + \sum_{\substack{n=1 \\ n \neq n_j}}^{\infty} P_n x,$$

in agreement with (4.28).

It can further be shown that if A is bounded, or unbounded but densely defined,³²

$$A^* = \sum_{n=1}^{\infty} \bar{\lambda}_n P_n$$

or, equivalently,

$$A^* = \sum_{n=1}^{\infty} \bar{\lambda}_n \langle \cdot, x_n \rangle x_n.$$

This representation for the adjoint assumes A possesses an orthonormal basis of eigenvectors (e.g., the formula is useful for compact normal operators) and does not hold for arbitrary operators. Functions of operators in terms of projections, including inversion formulas, are discussed in Section 4.4.

To summarize, compact, self-adjoint operators $A : H \rightarrow H$ possess an orthonormal eigenbasis for H . The same can be said for compact normal operators, self-adjoint boundary value problems, as well as some noncompact integral operators. Moreover, the stated conditions are sufficient but not necessary.

³²To see this, let $x = \sum_n \langle x, x_n \rangle x_n$ and $y = \sum_n \langle y, x_n \rangle x_n$, such that $\langle x, y \rangle = \sum_n \langle x, x_n \rangle \overline{\langle y, x_n \rangle}$. With $\langle Ax, y \rangle = \sum_n \lambda_n \langle x, x_n \rangle \overline{\langle y, x_n \rangle}$, we look for $z \in H$ such that $\langle Ax, y \rangle = \langle x, z \rangle$. If such a $z = \sum_n \langle z, x_n \rangle x_n$ exists, then $\overline{\langle z, x_n \rangle} = \lambda_n \overline{\langle y, x_n \rangle}$, or $\langle z, x_n \rangle = \bar{\lambda}_n \langle y, x_n \rangle$. Therefore, if $z = A^* y$, then $A^* y = \sum_n \bar{\lambda}_n \langle y, x_n \rangle x_n$, i.e., $A^* = \sum_{n=1}^{\infty} \bar{\lambda}_n P_n$.

4.4 Functions of Operators

It is often necessary to consider functions of operators, such as determining the exponential of an operator A . In this section we consider functions of an operator via three methods: spectral representations, series expansions, and the Dunford integral representation.

4.4.1 Functions of Operators via Spectral Representations

We assume that the operator $A : H \rightarrow H$ is bounded and simple,³³ with H generally infinite-dimensional. Therefore, for every $x \in H$ we have $x = \sum_n \langle x, x_n \rangle x_n$, $Ax = \sum_n \lambda_n \langle x, x_n \rangle x_n$, and $Ix = \sum_n \langle x, x_n \rangle x_n$.³⁴ In what follows we also use the form (4.40), such that $x = \sum_n P_n x$, $A = \sum_n \lambda_n P_n$, and $I = \sum_n P_n$.

Beginning with operators raised to an integer power, we have $A^2 x \equiv AAx = \sum_n \lambda_n^2 \langle x, x_n \rangle x_n$, and one can see that, generally,

$$A^m x = \sum_n \lambda_n^m \langle x, x_n \rangle x_n \quad (4.41)$$

for any power $m < \infty$. Therefore, for any polynomial $p(t) = \sum_{m=0}^M a_m t^m$ (we assume $x = x(t) \in H$),

$$p(A)x = \sum_n p(\lambda_n) \langle x, x_n \rangle x_n \quad (4.42)$$

or

$$p(A) = \sum_n p(\lambda_n) P_n.$$

Generalizing³⁵ to bounded continuous functions $f(t)$, we define the operator-valued function $f(A)$ by

$$f(A)x = \sum_n f(\lambda_n) \langle x, x_n \rangle x_n \quad (4.43)$$

or

$$f(A) = \sum_n f(\lambda_n) P_n. \quad (4.44)$$

³³If the operator is not simple, its set of eigenfunctions will not necessarily be orthogonal or complete in H . In light of Theorem 4.7, it can be seen that in this case an expansion of the form $x = \sum_n \langle x, y_n \rangle x_n$ may be useful, where y_n are eigenfunctions of the adjoint operator.

³⁴As usual, when index limits are not specified, we assume a finite set $n = 1, \dots, N$ for operators on finite-dimensional spaces, and an infinite set $n = 1, 2, \dots$ for operators on infinite-dimensional spaces.

³⁵We can achieve this generalization since any real-valued continuous function can be written as the limit of a sequence of polynomials $\{p_n(t)\}$ via the Weierstrass theorem (see Example 2, Section 2.2.2).

As special cases, if $f(t) = 1$ then $f(A) = I$, and if $f(t) = t$ then $f(A) = A$.

Note also that because we have $A = \sum_n \lambda_n P_n$, then (4.44) implies that if A is a diagonal matrix, $A = \text{diag}[a_{11}, a_{22}, \dots]$, then $f(A)$ is the diagonal matrix³⁶

$$f(A) = \text{diag}[f(a_{11}), f(a_{22}), \dots]. \quad (4.45)$$

The spectral representations (4.43) and (4.44) allow one to define such quantities as $\sin(A)$, $\log(A)$, etc., assuming $f(\lambda_n)$ is meaningful. For example, by Theorem 4.19 we see that eigenvalues of a nonnegative operator are nonnegative. Therefore, for a compact, self-adjoint, nonnegative operator A , the square root of A ,

$$A^{1/2} = \sum_n \lambda_n^{1/2} P_n,$$

is well defined. In general, ambiguities can be avoided if f is analytic on $\sigma(A)$.

As another example, the exponential of the operator $A : H \rightarrow H$, applied to $x \in H$, is

$$e^{iA}x = \sum_n e^{i\lambda_n} \langle x, x_n \rangle x_n,$$

where λ_n, x_n are the eigenvalues and eigenfunctions, respectively, of the operator A . Alternatively,

$$e^{iA} = \sum_n e^{i\lambda_n} P_n. \quad (4.46)$$

Example 4.8.

Consider a system of n first-order, constant-coefficient differential equations

$$\mathbf{x}'(t) = A\mathbf{x}(t),$$

where $A \in \mathbf{C}^{n \times n}$ is a matrix of constants, $\mathbf{x}$ is an n -dimensional vector of functions, and $\mathbf{x}'$ is the vector of derivatives taken individually on elements of $\mathbf{x}$. Initial conditions are given as the vector of constants $\mathbf{x}(0) = \mathbf{x}_0 = [x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(n)}]^\top \in \mathbf{C}^n$. For the case of one equation ($n = 1$), A is a 1×1 matrix (a scalar, call it α) and the general solution is obviously

$$x(t) = e^{At}x_0 = e^{\alpha t}x_0,$$

³⁶Moreover, for any pair of operators A, B related by a similarity transformation, i.e., $A = PBP^{-1}$, then $A^m = \underbrace{(PBP^{-1})(PBP^{-1}) \cdots (PBP^{-1})}_{n \text{ times}} = (PB^mP^{-1})$ and

we see (directly for polynomials and by extension (Weierstrass) to at least continuous functions) that for f defined on $\sigma(A) = \sigma(B)$,

$$f(A) = Pf(B)P^{-1}.$$

where $x_0 = x_0^{(1)}$ is a scalar. In the case of n equations, the general solution is

$$\mathbf{x}(t) = e^{At} \mathbf{x}_0 = \sum_{m=1}^n e^{\lambda_m t} \langle \mathbf{x}_0, \mathbf{x}_m \rangle \mathbf{x}_m,$$

where λ_m and $\mathbf{x}_m$ are eigenvalues and eigenfunctions of the matrix A . If A is diagonal (i.e., if the equations are not coupled), the solution is particularly simple,³⁷

$$[x_1(t), x_2(t), \dots, x_n(t)]^\top = [e^{\lambda_1 t} x_0^{(1)}, e^{\lambda_2 t} x_0^{(2)}, \dots, e^{\lambda_n t} x_0^{(n)}]^\top.$$

If A is not diagonal but is at least diagonalizable, then it is easiest to diagonalize A as $D = PAP^{-1}$ such that $\mathbf{z}'(t) = D\mathbf{z}(t)$, solve this new system of equations, and find $\mathbf{x}$ as $\mathbf{x}(t) = P^{-1}\mathbf{z}(t)$.

A particularly important function is inversion, leading to the inverse operator. Letting

$$f(t) = (t - \lambda)^{-1},$$

then

$$f(A) = R_\lambda(A) = \sum_n (\lambda_n - \lambda)^{-1} P_n \quad (4.47)$$

for the resolvent, where obviously if $\lambda = \lambda_n$ we must have $P_n = 0$ for $f(A)$ to possibly be meaningful. For $\lambda = 0$ (note $\lambda_n \neq 0$), $f(t) = t^{-1}$, leading to the inverse operator $f(A) = A^{-1}$,

$$A^{-1} = \sum_n \lambda_n^{-1} P_n \quad (4.48)$$

or

$$A^{-1}x = \sum_n \lambda_n^{-1} \langle x, x_n \rangle x_n. \quad (4.49)$$

It can be shown that for a bounded operator A , the spectrum of a function of the operator is equal to the function of the spectrum,

$$\sigma(f(A)) = f(\sigma(A)),$$

which is known as the *spectral mapping theorem*. For example, if $\sigma(A) = \sigma_p(A) = \{\lambda_1, \lambda_2, \dots\}$, then $\sigma(f(A)) = \{f(\lambda_1), f(\lambda_2), \dots\}$.

Also, from Theorem 4.10 note that for A bounded, $\|A\| \geq r_\sigma(A) \equiv \sup_{\lambda \in \sigma(A)} |\lambda|$. For A normal it can be shown that equality is achieved, i.e., $\|A\| = r_\sigma(A)$, and further,

$$\|f(A)\| = r_\sigma(f(A)) = \sup_{\lambda \in \sigma(A)} |f(\lambda)|.$$

³⁷Note that for a diagonal matrix the eigenvalues are the main diagonal entries and the eigenvectors are the standard vectors, i.e., $\mathbf{x}_1 = [1, 0, \dots]^\top$, $\mathbf{x}_2 = [0, 1, 0, \dots]^\top$, ...

The above representations for functions of an operator $f(A)$ are valid assuming the operator can be written as a weighted discrete sum of projections, i.e., that the operator is simple. For the operators considered here, each projection is associated with the span of an eigenfunction. For example, compact normal operators belong to this class. It can be shown that normal operators (not necessarily compact) can be represented as a weighted continuous sum of projections, and formulas similar to a continuous version of (4.44) (i.e., in terms of integrals) can be developed for both bounded and unbounded operators. Furthermore, rather than considering continuous functions, one may develop a theory for generally measurable functions [12, Chs. VII and VIII]. For operators that are not normal, weighted projection methods do not generally apply, and some other method of constructing functions of operators must be used. One such method is to use series expansions.

4.4.2 Functions of Operators via Series Expansions

A nonspectral method of constructing functions of operators is via series expansions. For instance, consider a function given by a power series,

$$f(t) = \sum_{n=0}^{\infty} a_n t^n$$

with radius of convergence R . Then, if A is bounded with $\|A\| < R$, the series $\sum_{n=0}^{\infty} a_n A^n$ converges³⁸ and we set

$$f(A) = \sum_{n=0}^{\infty} a_n A^n, \quad (4.50)$$

where $A^0 = I$. For example,

$$e^{i\omega t} = \sum_{n=0}^{\infty} (i\omega)^n t^n / n!,$$

and so³⁹

$$e^{i\omega A} = \sum_{n=0}^{\infty} (i\omega)^n A^n / n!.$$

It is interesting to note that $(e^{i\omega A})^n = e^{i\omega n A}$ and, in particular, that $e^{i\omega A}$ is invertible with inverse $e^{-i\omega A}$ even if A itself is not invertible (this

³⁸Since $\|A^n\| \leq \|A\|^n$ and $\|A\| < R$, then $\sum_{n=0}^{\infty} |a_n| \|A^n\|$ is convergent because we are within the radius of convergence of the original power series. When $\sum_{n=0}^{\infty} |a_n| \|A^n\|$ is convergent, $\sum_{n=0}^{\infty} a_n A^n$ is called absolutely convergent, which, in a Hilbert (more generally Banach) space implies convergence.

³⁹Recalling that power series for trigonometric functions have infinite radii of convergence, these series of operators will converge for any bounded operator.

also shows that $e^{i\omega A}$ is unitary). Exponentials of operators often arise in electromagnetics via propagator matrix approaches used in problems associated with a multilayered medium [32, Ch. 2], and in multiconductor transmission-line problems (see Section 7.4).

4.4.3 Functions of Operators via the Dunford Integral Representation

Finally, another method of constructing functions of operators is via a generalization of the Cauchy integral formula. First, consider the case of scalar functions. Let $f(\lambda)$ be analytic in a simply connected region Ω of the complex λ -plane, let Γ be an arbitrary closed piecewise smooth curve within Ω , and let z be a point interior to Γ . Then,

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda) (\lambda - z)^{-1} d\lambda \quad (4.51)$$

is known as *Cauchy's integral representation* of the function f [33, p. 37], where the integral is taken in the usual counterclockwise direction.

Now let $A : H \rightarrow H$ be bounded and $f(z)$ analytic within a region Ω of the complex plane, with $\sigma(A) \subset \Omega$. Then

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda) (\lambda I - A)^{-1} d\lambda \quad (4.52)$$

is known as the *Dunford integral representation* [34] of $f(A)$, which is seen to be a generalization of the Cauchy integral representation for ordinary functions to operator-valued functions. In particular,

$$I = \frac{1}{2\pi i} \int_{\Gamma} (\lambda I - A)^{-1} d\lambda$$

and

$$A = \frac{1}{2\pi i} \int_{\Gamma} \lambda (\lambda I - A)^{-1} d\lambda.$$

Like the Cauchy integral representation for scalar functions, the Dunford integral representation is primarily useful for theoretical developments rather than computation.

The various definitions of $f(A)$, i.e., (4.44), (4.50), and (4.52), may have different domains of validity. For instance, in order for (4.44) to hold, A must admit a representation as a weighted sum of projections, while for (4.50) to hold, A must be appropriately bounded, such that the series of operators converges. The various definitions will agree whenever they are mutually applicable for a certain operator [22, p. 10]. For instance, let f , analytic on $\sigma(A)$ where A is a bounded operator, be given by the power series

$$f(\lambda) = \sum_{n=0}^{\infty} a_n \lambda^n$$

with radius of convergence R . Let Γ be a circular (radius r), positively oriented contour in the λ -plane such that $\sigma(A)$ is contained within Γ (i.e., $r_\sigma(A) < r$). Then

$$\begin{aligned} f(A) &= \frac{1}{2\pi i} \int_{\Gamma} f(\lambda) (\lambda I - A)^{-1} d\lambda \\ &= \frac{1}{2\pi i} \int_{\Gamma} \left(\sum_{n=0}^{\infty} a_n \lambda^n \right) (\lambda I - A)^{-1} d\lambda \end{aligned}$$

is meaningful. Using

$$(\lambda I - A)^{-1} = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n,$$

which converges by Theorem 3.38 assuming $|\lambda| > \|A\|$, we see that if r satisfies $r_\sigma(A) \leq \|A\| < r < R$, then

$$\begin{aligned} f(A) &= \frac{1}{2\pi i} \int_{\Gamma} \left(\sum_{n=0}^{\infty} a_n \lambda^n \right) \left(\sum_{k=0}^{\infty} \frac{1}{\lambda^{k+1}} A^k \right) d\lambda \\ &= \sum_{n=0}^{\infty} a_n \sum_{k=0}^{\infty} A^k \frac{1}{2\pi i} \int_{\Gamma} \lambda^{n-k-1} d\lambda, \end{aligned}$$

where the interchange of operators is permissible because of the convergence properties of the series. From the residue theorem [33, Sec. 22] we see that if $k \leq n-1$ or $k > n$, the integral vanishes, whereas for $k = n$ the integral is $2\pi i$. Therefore,

$$f(A) = \sum_{n=0}^{\infty} a_n A^n,$$

and in this case $f(A)$ defined by the Dunford integral representation (4.52) is in agreement with the series expansion form (4.50).

4.5 Spectral Methods in the Solution of Operator Equations

Spectral representations are especially useful for solving inhomogeneous operator equations. In the following analysis the operator $A : H \rightarrow H$ is assumed simple (e.g., compact and self-adjoint, in which case $A - \lambda I$ is a Fredholm operator), with an orthonormal eigenbasis $\{x_n\}$ for H . The set $\{u_n\} \subseteq \{x_n\}$ consists of eigenfunctions of A corresponding to nonzero eigenvalues, forming a basis of R_A , and the set $\{\nu_n\} \subset \{x_n\}$ consists of eigenfunctions corresponding to zero eigenvalues ($A\nu_n = 0$), forming a basis of N_A , and $\{x_n\} = \{u_n\} \cup \{\nu_n\}$.

4.5.1 First- and Second-Kind Operator Equations

First-Kind Operator Equation, $N_A = \{0\}$

Consider a first-kind operator equation of the form

$$Ax = y$$

with $N_A = \{0\}$ (i.e., $\lambda = 0$ is not an eigenvalue of A , and so the set $\{u_n\}$ is an orthonormal basis for H) and expand $x, y \in H$ as

$$\begin{aligned} x &= \sum_n \langle x, u_n \rangle u_n, \\ y &= \sum_n \langle y, u_n \rangle u_n, \end{aligned}$$

both series being convergent in norm because $\{u_n\}$ is a basis for H . Since $Ax = \sum_n \lambda_n \langle x, u_n \rangle u_n = y = \sum_n \langle y, u_n \rangle u_n$, we solve for the coefficients of x as $\langle x, u_n \rangle = (1/\lambda_n) \langle y, u_n \rangle$, leading to the unique solution

$$x = \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle u_n,$$

which can also be obtained by application of the inverse operator (4.49). By the Riesz–Fischer theorem (Theorem 2.11) $\sum_n (1/\lambda_n) \langle y, u_n \rangle u_n$ will converge if and only if $\sum_n |(1/\lambda_n) \langle y, u_n \rangle|^2 < \infty$, i.e., if $(1/\lambda_n) \langle y, u_n \rangle \in l^2$.

The solution will not necessarily be continuously dependent on y (i.e., A^{-1} may be unbounded). To see this, change y by a small amount, say εu_k , where u_k is some eigenfunction of A and ε is a small parameter (the form εu_k is chosen for convenience). Then, with $\tilde{y} = y + \varepsilon u_k$, the solution of $A\tilde{x} = \tilde{y}$ is

$$\tilde{x} = \sum_n \frac{1}{\lambda_n} \langle \tilde{y}, u_n \rangle u_n = \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle u_n + \frac{\varepsilon}{\lambda_k}.$$

Therefore, $\|x - \tilde{x}\| = |\varepsilon/\lambda_k|$. If λ_k is very small, the difference $\|x - \tilde{x}\|$ can become very large, even for a small perturbation ε [4, p. 406]. Such is the case when A is compact (and so A^{-1} is unbounded), because $\lambda_k \rightarrow 0$ as $k \rightarrow \infty$.

For the compact case, or generally where $\lambda_n \rightarrow 0$, it is interesting to note that in the numerical solution of first-kind operator equations by discretization methods (e.g., the method of moments), often a reasonably accurate solution can be obtained by a somewhat coarse discretization. As the discretization becomes finer, rather than converging, the solution typically diverges, reflecting the more accurate modeling of the unbounded inverse. Furthermore, often eigenvalues associated with more singular kernels tend toward the origin less quickly compared with eigenvalues associated with

smoother kernels. Therefore, the solution of first-kind equations having more singular kernels is sometimes easier from an approximate numerical standpoint, compared with solving first-kind integral equations having smoother kernels.

First-Kind Operator Equation, $N_A \neq \{0\}$

If $\lambda = 0$ is an eigenvalue of A , then it is convenient to write

$$x = x_0 + \sum_n \langle x, u_n \rangle u_n,$$

where $Ax_0 = 0$ and $x_0 = \sum_n \langle x_0, \nu_n \rangle \nu_n$. Since $Ax = \sum_n \lambda_n \langle x, u_n \rangle u_n = y$ and $\{u_n\} \perp \{\nu_n\}$, clearly for y to be in the range of A it must be orthogonal to $\{\nu_n\}$ (i.e., $y \perp N_A$; note that if $N_A = \{0\}$, as in the case considered above, then $y \perp N_A$ trivially). Assuming this is the case, we write $y = \sum_n \langle y, u_n \rangle u_n$. Exploiting orthonormality of the set $\{u_n\}$ leads to $\langle x, u_n \rangle = (1/\lambda_n) \langle y, u_n \rangle$, such that the solution is given by

$$x = x_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle u_n,$$

which is, however, nonunique since x_0 is an arbitrary element of N_A . As before, $\sum_n (1/\lambda_n) \langle y, u_n \rangle u_n$ will converge if and only if $\sum_n |(1/\lambda_n) \langle y, u_n \rangle|^2 < \infty$.

To summarize, the operator equation $Ax = y$, where A is a simple operator, is solvable if $y \perp N_A$. The solution,

$$x = x_0 + \sum_n \frac{1}{\lambda_n} \langle y, u_n \rangle u_n,$$

where $Ax_0 = 0$, converges if $\sum_n |(1/\lambda_n) \langle y, u_n \rangle|^2 < \infty$ and is unique if $N_A = \{0\}$, that is, if $Ax_0 = 0$ implies $x_0 = 0$.

If the operator is not simple, the set of eigenfunctions will not necessarily be mutually orthogonal or complete. In this case, one may expand $x \in H$ in the bi-orthogonal set $\{x_n, y_n\}$ as $x = \sum_n \langle x, y_n \rangle x_n$, with the sets $\{x_n\}$ and $\{y_n\}$ being the eigenfunctions of A and the adjoint operator A^* , respectively (of course, completeness of this set and validity of the expansion need to be established). The solution is obtained as

$$x = x_0 + \sum_n \frac{1}{\lambda_n} \langle y, y_n \rangle x_n,$$

and we require $y \perp N_{A^*}$.

Second-Kind Operator Equation, λ Is Not an Eigenvalue of A

Consider the second-kind operator equation

$$(A - \lambda I)x = y.$$

The case $\lambda = 0$ was considered above, so we will assume $\lambda \neq 0$.

We first consider the case where λ is not an eigenvalue of A (i.e., $\lambda \notin \sigma_p(A)$). Therefore, $(A - \lambda I)x = 0$ has only the trivial solution, and $N_{A-\lambda I} = \{0\}$.⁴⁰ In this case we can make the expansions

$$\begin{aligned} Ax &= \sum_n \lambda_n \langle x, u_n \rangle u_n, \\ x &= x_0 + \sum_n \langle x, u_n \rangle u_n, \\ y &= y_0 + \sum_n \langle y, u_n \rangle u_n, \end{aligned} \tag{4.53}$$

where $x_0, y_0 \in N_A$. Exploiting orthonormality of the sets $\{u_n\}$ and $\{\nu_n\}$ leads to $(\lambda_n - \lambda) \langle x, u_n \rangle = \langle y, u_n \rangle$ and $\lambda x_0 = -y_0$, respectively. The solution is therefore

$$x = -\frac{y_0}{\lambda} + \sum_n \frac{1}{(\lambda_n - \lambda)} \langle y, u_n \rangle u_n. \tag{4.54}$$

If $N_A = \{0\}$, then the y_0 term is omitted and the solution is unique. Otherwise y_0 is retained and the solution is not unique. By the Riesz-Fischer theorem (Theorem 2.11) we require $\sum_n |(1/(\lambda_n - \lambda)) \langle y, u_n \rangle|^2 < \infty$ for the series to converge. Note that trivially $y \perp N_{A-\lambda I}$.

Second-Kind Operator Equation, λ Is an Eigenvalue of A

Now consider the case when λ is an eigenvalue of A , and let $\lambda = \lambda_m$. For simplicity, assume unit multiplicity of λ_m . Therefore, $(A - \lambda I)x = (A - \lambda_m I)x = 0$ has nontrivial solutions in $N_{A-\lambda_m I}$. In this case we again make the expansions (4.53). Exploiting orthonormality of the sets $\{u_n\}$ and $\{\nu_n\}$ leads to $(\lambda_m - \lambda_n) \langle x, u_n \rangle = \langle y, u_n \rangle$ and $\lambda x_0 = -y_0$, respectively. Then,

- if $\langle y, u_m \rangle \neq 0$, we cannot solve for $\langle x, u_m \rangle$, and the equation has no solution,
- if $\langle y, u_m \rangle = 0$, then $\langle x, u_m \rangle$ is arbitrary; call it α . The (nonunique) solution is

$$x = -\frac{y_0}{\lambda_m} + \sum_{n \neq m} \frac{1}{(\lambda_n - \lambda_m)} \langle y, u_n \rangle u_n + \alpha u_m.$$

⁴⁰Note that this does not imply $N_A = \{0\}$.

By the Riesz-Fischer theorem we need $\sum_{n \neq m} |(1/(\lambda_n - \lambda_m)) \langle y, u_n \rangle|^2 < \infty$ for the series to converge. Note that we require $y \perp N_{A - \lambda I}$.

In both of the above cases we require $y \perp N_{A - \lambda I}$ for y to be in the range of the operator, such that a solution will exist. This is in agreement with the Fredholm alternative 3.44 (generally we require $y \perp N_{(A - \lambda I)^*}$, but the adjoint is not needed since we are assuming that the operator is simple⁴¹).

4.5.2 Spectral Methods and Green's Functions

In Section 3.9.3 we showed that a Green's function is the kernel of an integral operator, the Green's operator, which is inverse to a differential operator. In this section we continue the discussion of Green's functions, emphasizing spectral theory. The manipulations are merely formal and intended to highlight salient points of the theory. Explicit formulations are considered in Chapters 1 and 5 and in Part II.

Referring to (3.34), consider the differential operator equation

$$(\mathbf{L} - \lambda I) \mathbf{x} = \mathbf{y}, \quad (4.55)$$

where $\mathbf{L}$ is a vector differential operator on a Hilbert space H (for example, $\mathbf{L} = \nabla \times \nabla \times$ or ∇^2), where $\mathbf{x} = \mathbf{x}(\mathbf{r}) \in D_{\mathbf{L}}$ and $\mathbf{y} = \mathbf{y}(\mathbf{r}) \in R_{\mathbf{L} - \lambda I}$. We assume that $\mathbf{x}$ as a vector will generally not be oriented in the same direction as the source vector $\mathbf{y}$, and so we anticipate a Green's operator $\mathbf{G}_\lambda$ (inverse operator to $(\mathbf{L} - \lambda I)$) with a dyadic Green's function kernel $\underline{\mathbf{g}}$, satisfying

$$(\mathbf{L} - \lambda I) \underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) = \mathbf{I} \delta(\mathbf{r} - \mathbf{r}') \quad (4.56)$$

such that

$$\mathbf{x}(\mathbf{r}) = (\mathbf{G}_\lambda \mathbf{y})(\mathbf{r}) = \int_{\Omega} \underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) \cdot \mathbf{y}(\mathbf{r}') d\Omega'. \quad (4.57)$$

In this case, where we assume that $(\mathbf{L} - \lambda I)$ is invertible, the Green's operator is exactly the resolvent of $\mathbf{L}$, i.e., $\mathbf{G}_\lambda = (\mathbf{L} - \lambda I)^{-1} = R_\lambda(\mathbf{L})$.

To determine a spectral representation, assume that (4.55) forms a self-adjoint boundary value problem, such that the vector eigenfunctions defined by

$$\mathbf{L} \mathbf{x}_n = \lambda_n \mathbf{x}_n$$

form an orthonormal eigenbasis in H subject to some appropriate boundary conditions. The method leading to (4.54) results in the solution of (4.55) as

$$\mathbf{x} = \sum_n \frac{\langle \mathbf{y}, \mathbf{x}_n \rangle}{(\lambda_n - \lambda)} \mathbf{x}_n = \mathbf{G}_\lambda \mathbf{y}, \quad (4.58)$$

⁴¹See Example 5.14 for an example when the operator is not simple, and associated eigenfunctions are implicated. In such cases the expansion (4.53) needs to be modified to include the associated eigenfunctions.

where for simplicity we assume that $N_{\mathbf{L}} = \{0\}$ and that $\lambda = 0$ is not an eigenvalue of $\mathbf{L}$. The summation is generally multidimensional. A similar form for the Green's function is also available. Indeed, comparing (4.57) and (4.58) we identify the Green's function as

$$\underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) = \sum_n \frac{\mathbf{x}_n(\mathbf{r}) \bar{\mathbf{x}}_n(\mathbf{r}')}{\lambda_n - \lambda}, \quad (4.59)$$

which is known as the *bilinear series* form for $\underline{\mathbf{g}}$.

Using the notation developed in Section 4.4, we can express the Green's operator as a weighted sum of projections

$$\mathbf{G}_\lambda = \sum_n (\lambda_n - \lambda)^{-1} \mathbf{P}_n, \quad (4.60)$$

where $\mathbf{P}_n \mathbf{x} = \langle \mathbf{x}, \mathbf{x}_n \rangle \mathbf{x}_n$ projects $\mathbf{x}$ onto the one-dimensional space spanned by the eigenfunction $\mathbf{x}_n$. This also naturally leads to the idea that the Green's operator is simply one element (the inverse function) in a family of possible functions of the operator $(\mathbf{L} - \lambda I)$.

In higher dimensions the eigenfunction expansion of the Green's function is generally not convergent when the source point $\mathbf{r}'$ coincides with the field point $\mathbf{r}$. In these situations the Green's function is to be considered as an operator within the theory of distributions, such that when substituted as the kernel of an integral expression (like (1.49) or (1.73)) the resulting terms can be integrated term by term, leading to a convergent series.

The representation (4.59) provides the Green's function $\underline{\mathbf{g}}$ given the eigenfunctions and eigenvalues $\mathbf{x}_n$ and λ_n . Alternatively, given the Green's function one may determine the eigenfunctions and eigenvalues. Assume the series (4.59) represents a *meromorphic function*⁴² of λ such that the Green's function has simple poles at $\lambda = \lambda_n$ in the complex λ -plane. Since $\mathbf{L}$ is self-adjoint, these will be on the real λ -axis. Taking the integral of (4.59) over a closed, counter-clockwise oriented contour C in the complex λ -plane that encloses all of the poles, conveniently chosen as a circle having

⁴²A meromorphic function $f(\lambda)$ is one in which the only singularities of f in the finite λ -plane are poles. Pole singularities may be absent as well, such that a meromorphic function is a function whose singularities in the finite λ -plane are "no worse" than poles (i.e., no branch points or essential singularities may exist in the finite λ -plane).

It is useful to note that the derivative of a meromorphic function f is a meromorphic function f' and that f and f' share the same poles.

radius tending toward infinity in the λ -plane, leads to⁴³

$$\begin{aligned} \oint_C \underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) d\lambda &= \sum_n \mathbf{x}_n(\mathbf{r}) \bar{\mathbf{x}}_n(\mathbf{r}') \oint_C \frac{d\lambda}{\lambda_n - \lambda} \\ &= -2\pi i \sum_n \mathbf{x}_n(\mathbf{r}) \bar{\mathbf{x}}_n(\mathbf{r}'), \end{aligned} \quad (4.61)$$

where the series (4.59) is viewed as a distribution and integrated term by term, and the last equality is via the residue theorem [33, Sec. 22.1]. The expression (4.61) allows one to obtain the eigenfunctions by considering the poles of the Green's function and its associated residues.

To obtain a further useful expression, consider the expansion of the delta-function $\underline{\mathbf{I}} \delta(\mathbf{r} - \mathbf{r}')$ in eigenfunctions of $\mathbf{L}$ (see the footnote on p. 271),

$$\underline{\mathbf{I}} \delta(\mathbf{r} - \mathbf{r}') = \sum_n \mathbf{x}_n(\mathbf{r}) \bar{\mathbf{x}}_n(\mathbf{r}'), \quad (4.62)$$

which is called the *spectral representation of the delta-function*. Relations of type (4.62) are also known as *completeness relations*, since taking the (posterior) dot product of (4.62) with $\mathbf{x}(\mathbf{r}')$ and integrating over space lead to the expansion $\mathbf{x}(\mathbf{r}) = \sum_n \langle \mathbf{x}, \mathbf{x}_n \rangle \mathbf{x}_n(\mathbf{r})$.

Combining (4.61) and (4.62) we have⁴⁴

$$\frac{1}{2\pi i} \oint_C \underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) d\lambda = -\underline{\mathbf{I}} \delta(\mathbf{r} - \mathbf{r}'), \quad (4.63)$$

an expression of principal interest as evidenced by its usefulness in later chapters. Throughout the text the integral in relations of the form (4.63) will be taken in a counterclockwise direction.

For unbounded region problems $\underline{\mathbf{g}}$ may not be meromorphic, and often $\underline{\mathbf{g}}$ admits branch-point singularities in the λ -plane. In this case the discrete set of eigenfunctions is not complete⁴⁵ in H , and this set must be augmented by a continuum (representing the continuous spectrum of $\mathbf{L}$),⁴⁶ leading to

$$\underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) = \sum_n \frac{\mathbf{x}_n(\mathbf{r}) \bar{\mathbf{x}}_n(\mathbf{r}')}{\lambda_n - \lambda} + \int_v \frac{\mathbf{x}(\mathbf{r}, v) \bar{\mathbf{x}}(\mathbf{r}', v)}{\lambda(v) - \lambda} dv, \quad (4.64)$$

⁴³The fact that the poles are simple is important, because the residue theorem is utilized in obtaining the second equality in (4.61). For more general operators, poles having multiplicity greater than 1 may be present. In this case, the corresponding residue contributions need to be included and lead to associated eigenfunctions. This is discussed in more detail in Section 5.3.2.

⁴⁴This is a very general expression, such that the equality holds whenever the root system of the operator forms a basis for the space in question, assuming that the Green's function has the proper behavior at infinity. See also Section 5.3.2.

⁴⁵Note that this does not imply that H lacks a discrete orthonormal basis. In fact, every separable H possesses such a basis.

⁴⁶See Section 5.2 for a detailed discussion of the continuous spectrum for a scalar one-dimensional problem.

where $\int_v (\cdot) dv$ represents a multidimensional integration, and again we assume that the poles λ_n are simple. Elements of the continuum $\mathbf{x}(\mathbf{r}, v)$ are called improper eigenfunctions (since they do not belong to H), with improper eigenvalues $\lambda(v)$, and satisfy

$$\mathbf{L}\mathbf{x}(\mathbf{r}, v) = \lambda(v) \mathbf{x}(\mathbf{r}, v).$$

In this case the analog of (4.61) and (4.63) is

$$\begin{aligned} \frac{1}{2\pi i} \oint_C \underline{\mathbf{g}}(\mathbf{r}, \mathbf{r}', \lambda) d\lambda &= -\underline{\mathbf{I}} \delta(\mathbf{r} - \mathbf{r}') \\ &= -\sum_n \mathbf{x}_n(\mathbf{r}) \bar{\mathbf{x}}_n(\mathbf{r}') - \int_v \mathbf{x}(\mathbf{r}, v) \bar{\mathbf{x}}(\mathbf{r}', v) dv. \end{aligned} \quad (4.65)$$

The generalization to the nonself-adjoint case is considered in Section 5.3.2, which includes some comments on the validity of the two equalities in the above equation.

4.5.3 Convergence of Nonstandard Eigenvalues in Projection Techniques

In addition to the usefulness of the stationary form presented in Section 4.2.11, it is also possible under certain conditions to ascertain the convergence of spectral parameters obtained by the projection of infinite-dimensional operators onto finite-dimensional spaces. Consider the spectral problem

$$A(\gamma) x = 0, \quad (4.66)$$

where $A : B \rightarrow B$ is a bounded operator with B an infinite-dimensional Banach space. We assume that A is approximated by the operator $A_n : B_n \rightarrow B_n$, such that $A_n \rightarrow A$ as $n \rightarrow \infty$, where B_n is a finite-dimensional Banach space. In practice, we need to solve

$$A_n(\gamma_n) x_n = 0, \quad (4.67)$$

which is the discretized version of (4.66), and we would like to know whether $\gamma_n \rightarrow \gamma$ as $n \rightarrow \infty$. Indeed, under the conditions that A and A_n are Fredholm operator-valued functions, analytic on an open connected region Ω of the complex γ -plane, then [35]

- for every $\gamma \in \sigma(A)$, there exists $\gamma_n \in \sigma(A_n)$ such that $\gamma_n \rightarrow \gamma$ as $n \rightarrow \infty$.

More importantly,

- for every $\gamma_n \in \sigma(A_n)$ such that $\gamma_n \rightarrow \gamma \in \Omega$, $\gamma \in \sigma(A)$.

Therefore, under the stated conditions the (approximate) nonstandard eigenvalue γ_n , determined by solving a finite-dimensional version of (4.66), converges to an actual nonstandard eigenvalue γ of the original operator. For an application of this result to a dielectric waveguide problem, see [36].