2

Linear Operators on Normed Spaces

Many of the basic problems of applied mathematics share the property of *linearity*, and linear spaces and linear operators provide a general and useful framework for the analysis of such problems. More complicated applications often involve nonlinear operators, and a study of linear operators also offers some useful tools for the analysis of nonlinear operators. In this chapter we review some basic results on linear operators, and we give some illustrative applications to obtain results in numerical analysis. Some of the results are quoted without proof; and usually the reader can find detailed proofs of the results in a standard textbook on functional analysis, e.g. see Conway [58], Kantorovich and Akilov [135], and Zeidler [249], [250].

Linear operators are used in expressing mathematical problems, often leading to equations to be solved or to functions to be optimized. To examine the theoretical solvability of a mathematical problem and to develop numerical methods for its solution, we must know additional properties about the operators involved in our problem. The most important such properties in applied mathematics involve one of the following concepts or some mix of them.

- Closeness to a problem whose solvability theory is known. The *Geometric Series Theorem* given in Section 2.3 is the basis of most results for linear operator equations in this category.

- Closeness to a finite dimensional problem. One variant of this leads to the theory of *completely continuous* or *compact* linear operators, which is taken up in Section 2.8.

K. Atkinson and W. Han, *Theoretical Numerical Analysis: A Functional Analysis Framework*, Texts in Applied Mathematics 39, DOI: 10.1007/978-1-4419-0458-4_2,

- Arguments based on finding the minimum of a function, with the point at which the minimum is attained being the solution to the problem under study. The function being minimized is sometimes called an objective function in optimization theory or an energy function in mechanics applications. This is taken up in later chapters, but some of its framework is provided in the material of this chapter.

There are other important means of examining the solvability of mathematical problems in applied mathematics, based on Fourier analysis, complex analysis, positivity of an operator within the context of partially order linear spaces, and other techniques. However, we make only minimal use of such tools in this text.

2.1 Operators

Given two sets V and W, an *operator* T from V to W is a rule which assigns to each element in a subset of V a unique element in W. The *domain* $\mathcal{D}(T)$ of T is the subset of V where T is defined:

$$\mathcal{D}(T) = \{v \in V \mid T(v) \text{ is defined}\},$$

and the *range* $\mathcal{R}(T)$ of T is the set of the elements in W generated by T:

$$\mathcal{R}(T) = \{w \in W \mid w = T(v) \text{ for some } v \in \mathcal{D}(T)\}.$$

It is also useful to define the *null set*, the set of the zeros of the operator:

$$\mathcal{N}(T) = \{v \in V \mid T(v) = 0\}.$$

An operator is sometimes also called a mapping, a transformation, or a function. Usually the domain $\mathcal{D}(T)$ is understood to be the whole set V, unless it is stated explicitly otherwise. Also, from now on, we will assume both V and W are linear spaces, as this suffices in the rest of the book.

Addition and scalar multiplication of operators are defined similarly to that of ordinary functions. Let S and T be operators mapping from V to W. Then $S+T$ is an operator from V to W with the domain $\mathcal{D}(S)\cap\mathcal{D}(T)$ and the rule

$$(S+T)(v) = S(v) + T(v) \quad \forall\, v \in \mathcal{D}(S) \cap \mathcal{D}(T).$$

Let $\alpha \in \mathbb{K}$. Then αT is an operator from V to W with the domain $\mathcal{D}(T)$ and the rule

$$(\alpha T)(v) = \alpha\, T(v) \quad \forall\, v \in \mathcal{D}(T).$$

Definition 2.1.1 *An operator $T : V \to W$ is said to be* one-to-one *or* injective *if*

$$v_1 \neq v_2 \quad \Longrightarrow \quad T(v_1) \neq T(v_2). \tag{2.1.1}$$

The operator is said to map V onto *W or is called* surjective *if $\mathcal{R}(T) = W$. If T is both injective and surjective, it is called a* bijection *from V to W.*

Evidently, when $T : V \to W$ is bijective, we can define its inverse $T^{-1} : W \to V$ by the rule

$$v = T^{-1}(w) \quad \Longleftrightarrow \quad w = T(v).$$

More generally, if $T : V \to W$ is one-to-one, we can define its inverse from $\mathcal{R}(T) \subset W$ to V by using the above rule.

Example 2.1.2 Let V be a linear space. The *identity operator* $I : V \to V$ is defined by

$$I(v) = v \quad \forall\, v \in V.$$

It is a bijection from V to V; and moreover, its inverse is also the identity operator. □

Example 2.1.3 Let $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, and $L(\boldsymbol{v}) = A\boldsymbol{v}$, $\boldsymbol{v} \in \mathbb{R}^n$, where $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ is a real matrix and $A\boldsymbol{v}$ denotes matrix-vector multiplication. From results in linear algebra, the operator L is injective if and only if $\text{rank}(A) = n$; and L is surjective if and only if $\text{rank}(A) = m$. Recall that the rank of a matrix is the maximal number of independent column vectors, that is also the maximal number of independent row vectors.

The same conclusion holds for complex spaces $V = \mathbb{C}^n$, $W = \mathbb{C}^m$, and complex matrix $A \in \mathbb{C}^{m\times n}$. □

Example 2.1.4 We consider the differentiation operator d/dx from $V = C[0, 1]$ to $W = C[0, 1]$ defined by

$$\frac{d}{dx} : v \mapsto v' \quad \text{for } v \in C^1[0, 1].$$

We take the domain of the operator, $\mathcal{D}(d/dx)$, to be $C^1[0, 1]$ which is a proper subspace of $C[0, 1]$. It can be verified that the differentiation operator is a surjection, $\mathcal{R}(d/dx) = C[0, 1]$. The differentiation operator is not injective, and its null set is the set of constant functions. □

Example 2.1.5 Although the differentiation operator d/dx is not injective from $C^1[0, 1]$ to $C[0, 1]$, the following operator

$$D : v(x) \mapsto \begin{pmatrix} v'(x) \\ v(0) \end{pmatrix}$$

is a bijection between $V = C^1[0, 1]$ and $W = C[0, 1] \times \mathbb{R}$. □

If both V and W are normed spaces, we can talk about the continuity and boundedness of the operators.

Definition 2.1.6 *Let V and W be two normed spaces. An operator $T : V \to W$ is* continuous *at $v \in \mathcal{D}(T)$ if*

$$\{v_n\} \subset \mathcal{D}(T) \text{ and } v_n \to v \text{ in } V \quad \Longrightarrow \quad T(v_n) \to T(v) \text{ in } W.$$

T is said to be continuous *if it is continuous over its domain $\mathcal{D}(T)$. The operator is* bounded *if for any $r > 0$, there is an $R > 0$ such that*

$$v \in \mathcal{D}(T) \text{ and } \|v\| \le r \quad \Longrightarrow \quad \|T(v)\| \le R.$$

We observe that an alternative definition of the boundedness is that for any set $B \subset \mathcal{D}(T)$,

$$\sup_{v \in B} \|v\|_V < \infty \quad \Longrightarrow \quad \sup_{v \in B} \|T(v)\|_W < \infty.$$

Example 2.1.7 Let us consider the differentiation operator again. The spaces $C[0,1]$ and $C^1[0,1]$ are associated with their standard norms

$$\|v\|_{C[0,1]} = \max_{0 \le x \le 1} |v(x)|$$

and

$$\|v\|_{C^1[0,1]} = \|v\|_{C[0,1]} + \|v'\|_{C[0,1]}. \tag{2.1.2}$$

Then the operator

$$T_1 = \frac{d}{dx} : C^1[0,1] \subset C[0,1] \to C[0,1]$$

is not continuous using the infinity norm of $C[0,1]$ for $C^1[0,1]$, whereas the operator

$$T_2 = \frac{d}{dx} : C^1[0,1] \to C[0,1]$$

is continuous using the norm of (2.1.2) for $C^1[0,1]$. □

Exercise 2.1.1 Consider Example 2.1.7. Show that T_1 is unbounded and T_2 is bounded, as asserted in the example.

Exercise 2.1.2 Let $T_1 : C[a,b] \to C[a,b]$ be an operator defined by the formula

$$T_1 v(x) = \int_a^b (x - t)\, v(t)\, dt, \quad a \le x \le b,\ v \in C[a,b].$$

Determine the range of T_1. Is T_1 injective?

Exercise 2.1.3 Extending Exercise 2.1.2, for an arbitrary positive integer n, let $T_n : C[a,b] \to C[a,b]$ be defined by

$$T_n v(x) = \int_a^b (x-t)^n v(t)\,dt, \quad a \le x \le b,\ v \in C[a,b].$$

Determine the range of T_n and decide if T_n is injective.

Exercise 2.1.4 Over the space $C[a,b]$, define the operator T by

$$Tv(x) = \int_a^x v(t)\,dt, \quad a \le x \le b,\ v \in C[a,b].$$

Find the range of T. Is T a bijection between $C[a,b]$ and its range?

2.2 Continuous linear operators

This chapter is focused on the analysis of a particular type of operators called linear operators. From now on, when we write $T : V \to W$, we implicitly assume $\mathcal{D}(T) = V$, unless stated otherwise. As in Chapter 1, $\mathbb{K}$ denotes the set of scalars associated with the vector space under consideration.

Definition 2.2.1 *Let V and W be two linear spaces. An operator $L : V \to W$ is said to be* linear *if*

$$L(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 L(v_1) + \alpha_2 L(v_2) \quad \forall\, v_1, v_2 \in V,\ \forall\, \alpha_1, \alpha_2 \in \mathbb{K}\,,$$

or equivalently,

$$L(v_1 + v_2) = L(v_1) + L(v_2) \quad \forall\, v_1, v_2 \in V,$$
$$L(\alpha\, v) = \alpha\, L(v) \quad \forall\, v \in V,\ \forall\, \alpha \in \mathbb{K}\,.$$

For a linear operator L, we usually write $L(v)$ as Lv.

An important property of a linear operator is that continuity and boundedness are equivalent. We state and prove this result in the form of a theorem after two preparatory propositions which are themselves important.

Proposition 2.2.2 *Let V and W be normed spaces, $L : V \to W$ a linear operator. Then continuity of L over the whole space is equivalent to its continuity at any one point, say at $v = 0$.*

Proof. Continuity over the whole space certainly implies continuity at any point. Now assume L is continuous at $v = 0$:

$$v_n \to 0 \text{ in } V \implies Lv_n \to 0 \text{ in } W. \tag{2.2.1}$$

Let $v \in V$ be arbitrarily given and $\{v_n\} \subset V$ a sequence converging to v. Then $v_n - v \to 0$, and by (2.2.1), $L(v_n - v) = Lv_n - Lv \to 0$, i.e., $Lv_n \to Lv$. Hence L is continuous at v. □

Proposition 2.2.3 *Let V and W be normed spaces, $L : V \to W$ a linear operator. Then L is bounded if and only if there exists a constant $\gamma \geq 0$ such that*

$$\|Lv\|_W \leq \gamma \|v\|_V \quad \forall\, v \in V. \tag{2.2.2}$$

Proof. Obviously (2.2.2) implies the boundedness. Conversely, suppose L is bounded, then

$$\gamma \equiv \sup_{v \in B_1} \|Lv\|_W < \infty,$$

where $B_1 = \{v \in V \mid \|v\|_V \leq 1\}$ is the unit ball centered at 0. Now for any $v \neq 0$, $v/\|v\|_V \in B_1$ and by the linearity of L,

$$\|Lv\|_W = \|v\|_V \|L(v/\|v\|_V)\|_W \leq \gamma\, \|v\|_V,$$

i.e., (2.2.2) holds. □

Theorem 2.2.4 *Let V and W be normed spaces, $L : V \to W$ a linear operator. Then L is continuous on V if and only if it is bounded on V.*

Proof. Firstly we assume L is not bounded and prove that it is not continuous at 0. Since L is unbounded, we can find a bounded sequence $\{v_n\} \subset V$ such that $\|Lv_n\| \to \infty$. Without loss of generality, we may assume $Lv_n \neq 0$ for all n. Then we define a new sequence

$$\tilde{v}_n = \frac{v_n}{\|Lv_n\|_W}\,.$$

This sequence has the property that $\tilde{v}_n \to 0$ and $\|L\tilde{v}_n\|_W = 1$. Thus L is not continuous.

Secondly we assume L is bounded and show that it must be continuous. Indeed from (2.2.2) we have the Lipschitz inequality

$$\|Lv_1 - Lv_2\|_W \leq \gamma\, \|v_1 - v_2\|_V \quad \forall\, v_1, v_2 \in V, \tag{2.2.3}$$

which implies the continuity in an obvious fashion. □

From (2.2.3), we see that for a linear operator, continuity and Lipschitz continuity are equivalent.

We use the notation $\mathcal{L}(V, W)$ for the set of all the continuous linear operators from a normed space V to another normed space W. In the special case $W = V$, we use $\mathcal{L}(V)$ to replace $\mathcal{L}(V, V)$. We see that for a linear operator, boundedness (2.2.2) is equivalent to continuity. Thus if $L \in \mathcal{L}(V, W)$, it is meaningful to define

$$\|L\|_{V,W} = \sup_{0 \neq v \in V} \frac{\|Lv\|_W}{\|v\|_V}. \tag{2.2.4}$$

Using the linearity of L, we have the following relations

$$\begin{aligned} \|L\|_{V,W} &= \sup_{v \in B_1} \|Lv\|_W = \sup_{v:\|v\|_V = 1} \|Lv\|_W \\ &= \frac{1}{r} \sup_{v:\|v\|_V = r} \|Lv\|_W = \frac{1}{r} \sup_{v:\|v\|_V \leq r} \|Lv\|_W \end{aligned}$$

for any $r > 0$. The norm $\|L\|_{V,W}$ is the maximum size in W of the image under L of the unit ball B_1 in V.

Theorem 2.2.5 *The set $\mathcal{L}(V, W)$ is a linear space, and* (2.2.4) *defines a norm over the space.*

We leave the proof of the theorem to the reader. The norm (2.2.4) is usually called the *operator norm* of L, which enjoys the following compatibility property

$$\|Lv\|_W \leq \|L\|_{V,W} \|v\|_V \quad \forall\, v \in V. \tag{2.2.5}$$

If it is not stated explicitly, we always understand the norm of an operator as an operator norm defined by (2.2.4). Another useful inequality involving operator norms is given in the following result.

Theorem 2.2.6 *Let U, V and W be normed spaces, $L_1 : U \to V$ and $L_2 : V \to W$ be continuous linear operators. Then the composite operator $L_2 L_1 : U \to W$ defined by*

$$L_2 L_1(v) = L_2(L_1(v)) \quad \forall\, v \in U$$

is a continuous linear mapping. Moreover,

$$\|L_2 L_1\|_{U,W} \leq \|L_1\|_{U,V} \|L_2\|_{V,W}. \tag{2.2.6}$$

Proof. The composite operator $L_2 L_1$ is obviously linear. We now prove (2.2.6), that also implies the continuity of $L_2 L_1$. By (2.2.5), for any $v \in U$,

$$\begin{aligned} \|L_2 L_1(v)\|_W &= \|L_2(L_1(v))\|_W \\ &\leq \|L_2\|_{V,W} \|L_1 v\|_V \\ &\leq \|L_2\|_{V,W} \|L_1\|_{U,V} \|v\|_U. \end{aligned}$$

Hence, (2.2.6) is valid. □

As an important special case, if V is a normed space and if $L \in \mathcal{L}(V)$, then for any non-negative integer n,

$$\|L^n\| \le \|L\|^n.$$

The operator L^n is defined recursively: $L^n = L(L^{n-1})$, $L^n v = L(L^{n-1}v)$ $\forall\, v \in V$, and $L^0 = I$ is defined to be the identity operator. Both (2.2.5) and (2.2.6) are very useful relations for error analysis of some numerical methods.

For a linear operator, the null set $\mathcal{N}(L)$ becomes a subspace of V, and we have the statement

$$L \text{ is one-to-one} \quad \Longleftrightarrow \quad \mathcal{N}(L) = \{0\}.$$

Example 2.2.7 Let V be a linear space. Then the identity operator $I : V \to V$ belongs to $\mathcal{L}(V)$, and $\|I\| = 1$. □

Example 2.2.8 Recall Example 2.1.3. Let $V = \mathbb{C}^n$, $W = \mathbb{C}^m$, and $L(\boldsymbol{v}) = A\boldsymbol{v}$, $\boldsymbol{v} \in \mathbb{C}^n$, where $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ is a complex matrix. If the norms on V and W are $\|\cdot\|_\infty$, then the operator norm is the matrix ∞-norm,

$$\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|.$$

If the norms on V and W are $\|\cdot\|_1$, then the operator norm is the matrix 1-norm,

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|.$$

If the norms on V and W are $\|\cdot\|_2$, then the operator norm is the spectral norm

$$\|A\|_2 = \sqrt{r_\sigma(A^*A)} = \sqrt{r_\sigma(AA^*)},$$

where A^* denotes the conjugate transpose of A. For a square matrix B, $r_\sigma(B)$ denotes the spectral radius of the matrix B,

$$r_\sigma(B) = \max_{\lambda\in\sigma(B)} |\lambda|$$

and $\sigma(B)$ denotes the spectrum of B, the set of all the eigenvalues of B. Proofs of these results are given in [15, Section 7.3]. □

Example 2.2.9 Let $V = W = C[a,b]$ with the norm $\|\cdot\|_\infty$. Let $k \in C([a,b]^2)$, and define $K : C[a,b] \to C[a,b]$ by

$$(Kv)(x) = \int_a^b k(x,y)\, v(y)\, dy. \tag{2.2.7}$$

The mapping K in (2.2.7) is an example of a *linear integral operator*, and the function $k(\cdot,\cdot)$ is called the *kernel function* of the integral operator. Under the continuity assumption on $k(\cdot,\cdot)$, the integral operator is continuous from $C[a,b]$ to $C[a,b]$. Furthermore,

$$\|K\| = \max_{a\le x\le b} \int_a^b |k(x,y)|\,dy. \tag{2.2.8}$$

The linear integral operator (2.2.7) is later used extensively. □

2.2.1 $\mathcal{L}(V,W)$ as a Banach space

In approximating integral and differential equations, the integral or differential operator is often approximated by a sequence of operators of a simpler form. In such cases, it is important to consider the limits of convergent sequences of bounded operators, and this makes it important to have $\mathcal{L}(V,W)$ be a complete space.

Theorem 2.2.10 *Let V be a normed space, and W be a Banach space. Then $\mathcal{L}(V,W)$ is a Banach space.*

Proof. Let $\{L_n\}$ be a Cauchy sequence in $\mathcal{L}(V,W)$. This means

$$\epsilon_n \equiv \sup_{p\ge 1} \|L_{n+p} - L_n\| \to 0 \quad \text{as } n\to\infty.$$

We must define a limit for $\{L_n\}$ and show that it belongs to $\mathcal{L}(V,W)$.

For each $v\in V$,

$$\|L_{n+p}v - L_n v\|_W \le \epsilon_n \|v\|_V \to 0 \quad \text{as } n\to\infty. \tag{2.2.9}$$

Thus $\{L_n v\}$ is a Cauchy sequence in W. Since W is complete, the sequence has a limit, denoted by $L(v)$. This defines an operator $L: V\to W$. Let us prove that L is linear, bounded, and $\|L_n - L\|_{V,W}\to 0$ as $n\to\infty$.

For any $v_1, v_2\in V$ and $\alpha_1,\alpha_2\in\mathbb{K}$,

$$\begin{aligned} L(\alpha_1 v_1 + \alpha_2 v_2) &= \lim_{n\to\infty} L_n(\alpha_1 v_1 + \alpha_2 v_2)\\ &= \lim_{n\to\infty} (\alpha_1 L_n v_1 + \alpha_2 L_n v_2)\\ &= \alpha_1 \lim_{n\to\infty} L_n v_1 + \alpha_2 \lim_{n\to\infty} L_n v_2\\ &= \alpha_1 L(v_1) + \alpha_2 L(v_2). \end{aligned}$$

Thus L is linear.

Now for any $v\in V$, we take the limit $p\to\infty$ in (2.2.9) to obtain

$$\|Lv - L_n v\|_W \le \epsilon_n \|v\|_V.$$

Thus

$$\|L - L_n\|_{V,W} = \sup_{\|v\|_V \le 1} \|Lv - L_n v\|_W \le \epsilon_n \to 0 \quad \text{as } n \to \infty.$$

Hence $L \in \mathcal{L}(V, W)$ and $L_n \to L$ as $n \to \infty$. □

Exercise 2.2.1 For a linear operator $L : V \to W$, show that $L(0) = 0$.

Exercise 2.2.2 Prove that if a linear operator is discontinuous at one point, then it is discontinuous everywhere.

Exercise 2.2.3 Prove that $\mathcal{L}(V, W)$ is a linear space, and (2.2.4) defines a norm on the space.

Exercise 2.2.4 For $L \in \mathcal{L}(V, W)$, show that its norm is given by

$$\|L\| = \inf\{\alpha \mid \|Lv\|_W \le \alpha \|v\|_V \ \forall\, v \in V\}.$$

Exercise 2.2.5 Assume $k \in C([a, b]^2)$. Show that the integral operator K defined by (2.2.7) is continuous, and its operator norm is given by the formula (2.2.8).

Exercise 2.2.6 Let $\Omega \subset \mathbb{R}^d$ be a domain, $1 \le p \le \infty$. Given $m \in C(\overline{\Omega})$, define an operator $M : L^p(\Omega) \to L^p(\Omega)$ by the formula

$$Mv(\boldsymbol{x}) = m(\boldsymbol{x})\, v(\boldsymbol{x}), \quad v \in L^p(\Omega).$$

Show that M is linear, bounded, and $\|M\| = \|m\|_{C(\overline{\Omega})}$.

Exercise 2.2.7 A linear operator L is called nonsingular if $\mathcal{N}(L) = \{0\}$. Otherwise, it is called singular. Show that if L is nonsingular, then a solution of the equation $Lu = f$ is unique.

Exercise 2.2.8 Let a linear operator $L : V \to W$ be nonsingular and map V onto W. Show that for each $f \in W$, the equation $Lu = f$ has a unique solution $u \in V$.

2.3 The geometric series theorem and its variants

The following result is used commonly in numerical analysis and applied mathematics. It is also the means by which we can analyze the solvability of problems that are “close” to another problem known to be uniquely solvable.

Theorem 2.3.1 (Geometric series theorem) *Let V be a Banach space, $L \in \mathcal{L}(V)$. Assume*

$$\|L\| < 1. \tag{2.3.1}$$

Then $I - L$ is a bijection on V, its inverse is a bounded linear operator,

$$(I-L)^{-1} = \sum_{n=0}^{\infty} L^n,$$

and

$$\|(I-L)^{-1}\| \le \frac{1}{1-\|L\|}. \tag{2.3.2}$$

Proof. Define a sequence in $\mathcal{L}(V)$: $M_n = \sum_{i=0}^{n} L^i$, $n \ge 0$. For $p \ge 1$,

$$\|M_{n+p} - M_n\| = \left\| \sum_{i=n+1}^{n+p} L^i \right\| \le \sum_{i=n+1}^{n+p} \|L^i\| \le \sum_{i=n+1}^{n+p} \|L\|^i.$$

Using the assumption (2.3.1), we have

$$\|M_{n+p} - M_n\| \le \frac{\|L\|^{n+1}}{1-\|L\|}. \tag{2.3.3}$$

Hence,

$$\sup_{p\ge 1} \|M_{n+p} - M_n\| \to 0 \quad \text{as } n \to \infty,$$

and $\{M_n\}$ is a Cauchy sequence in $\mathcal{L}(V)$. Since $\mathcal{L}(V)$ is complete, there is an $M \in \mathcal{L}(V)$ with

$$\|M_n - M\| \to 0 \quad \text{as } n \to \infty.$$

Using the definition of M_n and simple algebraic manipulation,

$$(I-L)\,M_n = M_n(I-L) = I - L^{n+1}.$$

Let $n \to \infty$ to get

$$(I-L)\,M = M\,(I-L) = I.$$

This relation implies $(I-L)$ is a bijection, and

$$M = (I-L)^{-1} = \lim_{n\to\infty} \sum_{i=0}^{n} L^i = \sum_{n=0}^{\infty} L^n.$$

To prove the bound (2.3.2), note that

$$\|M_n\| \le \sum_{i=0}^{n} \|L\|^i \le \frac{1}{1-\|L\|}.$$

Taking the limit $n \to \infty$, we obtain (2.3.2). □

The theorem says that under the stated assumptions, for any $f \in V$, the equation

$$(I - L)\, u = f \tag{2.3.4}$$

has a unique solution $u = (I - L)^{-1} f \in V$. Moreover, the solution depends continuously on the right hand side f: Letting $(I - L)\, u_1 = f_1$ and $(I - L)\, u_2 = f_2$, it follows that

$$u_1 - u_2 = (I - L)^{-1}(f_1 - f_2),$$

and so

$$\|u_1 - u_2\| \le c\, \|f_1 - f_2\|$$

with $c = 1/(1 - \|L\|)$.

The theorem also provides an approach to approximate the solution of the equation (2.3.4). Under the stated assumptions of the theorem, we have

$$u = \lim_{n\to\infty} u_n$$

where

$$u_n = \sum_{j=0}^{n} L^j f. \tag{2.3.5}$$

Example 2.3.2 Consider the linear integral equation of the second kind

$$\lambda\, u(x) - \int_a^b k(x,y)\, u(y)\, dy = f(x), \quad a \le x \le b \tag{2.3.6}$$

with $\lambda \ne 0$, $k(x,y)$ continuous for $x, y \in [a,b]$, and $f \in C[a,b]$. Let $V = C[a,b]$ with the norm $\|\cdot\|_\infty$. Symbolically, we write the equation (2.3.6) as

$$(\lambda I - K)\, u = f, \tag{2.3.7}$$

where K is the linear integral operator generated by the kernel function $k(\cdot,\cdot)$. We also will often write this as $(\lambda - K)\, u = f$, understanding it to mean the same as in (2.3.7).

This equation (2.3.7) can be converted into the form needed in the geometric series theorem:

$$(I - L)\, u = \frac{1}{\lambda} f, \qquad L = \frac{1}{\lambda} K.$$

Applying the geometric series theorem we assert that if

$$\|L\| = \frac{1}{|\lambda|}\|K\| < 1,$$

then $(I-L)^{-1}$ exists and

$$\|(I-L)^{-1}\| \le \frac{1}{1-\|L\|}.$$

Equivalently, if

$$\|K\| = \max_{a\le x\le b} \int_a^b |K(x,y)|\,dy < |\lambda|, \tag{2.3.8}$$

then $(\lambda I-K)^{-1}$ exists and

$$\|(\lambda I-K)^{-1}\| \le \frac{1}{|\lambda|-\|K\|}.$$

Hence under the assumption (2.3.8), for any $f\in C[a,b]$, the integral equation (2.3.6) has a unique solution $u\in C[a,b]$ and

$$\|u\|_\infty \le \|(\lambda I-K)^{-1}\|\,\|f\|_\infty \le \frac{\|f\|_\infty}{|\lambda|-\|K\|}. \qquad \square$$

We observe that the geometric series theorem is a straightforward generalization to linear continuous operators on a Banach space of the power series

$$(1-x)^{-1} = \sum_{n=0}^\infty x^n, \quad x\in\mathbb{R},\ |x|<1$$

or its complex version

$$(1-z)^{-1} = \sum_{n=0}^\infty z^n, \quad z\in\mathbb{C},\ |z|<1.$$

From the proof of the theorem we see that for a linear operator $L\in\mathcal{L}(V)$ over a Banach space V, if $\|L\|<1$, then the series $\sum_{n=0}^\infty L^n$ converges in $\mathcal{L}(V)$ and the value of the series is the operator $(I-L)^{-1}$. More generally, we can similarly define an operator-valued function $f(L)$ of an operator variable L from a real function $f(x)$ of a real variable x (or a complex-valued function $f(z)$ of a complex variable z), as long as $f(x)$ is analytic at $x=0$, i.e., $f(x)$ has a convergent power series expansion:

$$f(x) = \sum_{n=0}^\infty a_n x^n, \quad |x|<\gamma$$

for some constant $\gamma>0$, where $a_n = f^{(n)}(0)/n!$, $n\ge 0$. Now if V is a Banach space and $L\in\mathcal{L}(V)$ satisfies $\|L\|<\gamma$, then we define

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

The series on the right-hand side is a well-defined operator in $\mathcal{L}(V)$, thanks to the assumption $\|L\| < \gamma$. We now give some examples of operator-valued functions obtained by this approach, with $L \in \mathcal{L}(V)$ and V a Banach space:

$$
\begin{aligned}
e^L &= \sum_{n=0}^{\infty} \frac{1}{n!}\, L^n, \\
\sin(L) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\, L^{2n+1}, \\
\arctan(L) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}\, L^{2n+1}, \quad \|L\| < 1.
\end{aligned}
$$

2.3.1 A generalization

To motivate a generalization of Theorem 2.3.1, consider the *Volterra integral equation of the second kind*

$$u(x) - \int_0^x \ell(x,y)u(y)\,dy = f(x), \quad x \in [0,B]. \tag{2.3.9}$$

Here, $B > 0$ and we assume the kernel function $\ell(x,y)$ is continuous for $0 \le y \le x \le B$, and $f \in C[0,B]$. If we apply Theorem 2.3.1, then we will need to assume a condition such as

$$\max_{x\in[0,B]} \int_0^x |\ell(x,y)|\,dy < 1$$

in order to conclude the unique solvability. However, we can use a variant of the geometric series theorem to show that the equation (2.3.9) is uniquely solvable, irregardless of the size of the kernel function $\ell(x,y)$. Symbolically, we write this integral equation as $(I-L)u = f$.

Corollary 2.3.3 *Let V be a Banach space, $L \in \mathcal{L}(V)$. Assume for some integer $m \ge 1$ that*

$$\|L^m\| < 1. \tag{2.3.10}$$

Then $I - L$ is a bijection on V, its inverse is a bounded linear operator, and

$$\|(I-L)^{-1}\| \le \frac{1}{1 - \|L^m\|} \sum_{i=0}^{m-1} \|L^i\|. \tag{2.3.11}$$

Proof. From Theorem 2.3.1, we know that $(I - L^m)^{-1}$ exists as a bounded bijective operator on V to V,

$$(I - L^m)^{-1} = \sum_{j=0}^{\infty} L^{mj} \quad \text{in } \mathcal{L}(V)$$

and

$$\|(I - L^m)^{-1}\| \le \frac{1}{1 - \|L^m\|}.$$

From the identities

$$(I - L)\left(\sum_{i=0}^{m-1} L^i\right) = \left(\sum_{i=0}^{m-1} L^i\right)(I - L) = I - L^m,$$

we then conclude that $(I - L)$ is a bijection,

$$(I - L)^{-1} = \left(\sum_{i=0}^{m-1} L^i\right)(I - L^m)^{-1}$$

and the bound (2.3.11) holds. □

Example 2.3.4 Returning to (2.3.9), define

$$Lv(x) = \int_0^x \ell(x, y)v(y)\,dy, \quad 0 \le x \le B,\ v \in C[0, B].$$

Easily, L is a bounded linear operator on $C[0, B]$ to itself. The iterated operators L^k take the form

$$L^k v(x) = \int_0^x \ell_k(x, y)v(y)\,dy$$

for $k = 2, 3, \ldots$, and $\ell_1(x, y) \equiv \ell(x, y)$. It is straightforward to show

$$\ell_{k+1}(x, y) = \int_y^x \ell_k(x, z)\ell(z, y)\,dz, \quad k = 1, 2, \ldots.$$

Let

$$M = \max_{0 \le y \le x \le B} |\ell(x, y)|.$$

Then

$$|\ell_1(x, y)| \le M, \quad 0 \le y \le x \le B.$$

Assume for an integer $k \ge 1$,

$$|\ell_k(x, y)| \le M^k \frac{(x - y)^{k-1}}{(k - 1)!}, \quad 0 \le y \le x \le B.$$

Then for $0 \le y \le x \le B$,

$$\begin{aligned}
|\ell_{k+1}(x, y)| &\le \int_y^x |\ell_k(x, z)|\,|\ell(z, y)|\,dz \\
&\le M^{k+1} \int_y^x \frac{(x - z)^{k-1}}{(k - 1)!}\,dz \\
&= M^{k+1} \frac{(x - y)^k}{k!}.
\end{aligned}$$

Hence, for any integer $k \geq 1$ and any $x \in [0, B]$,

$$
\begin{aligned}
|L^k v(x)| &\leq \int_0^x |\ell_k(x,y)|\,|v(y)|\,dy \\
&\leq \int_0^x M^k \frac{(x-y)^{k-1}}{(k-1)!}\,dy\,\|v\|_\infty \\
&= \frac{M^k x^k}{k!}\,\|v\|_\infty,
\end{aligned}
$$

and then

$$
\|L^k\| \leq \frac{M^k B^k}{k!}, \quad k = 1, 2, \ldots.
$$

It is clear that the right side converges to zero as $k \to \infty$, and thus (2.3.10) is satisfied for m large enough. We can also use this result to construct bounds for the solutions of (2.3.9), which we leave to Exercise 2.3.11. □

2.3.2 A perturbation result

An important technique in applied mathematics is to study an equation by relating it to a "nearby" equation for which there is a known solvability result. One of the more popular tools is the following perturbation theorem.

Theorem 2.3.5 *Let V and W be normed spaces with at least one of them being complete. Assume $L \in \mathcal{L}(V, W)$ has a bounded inverse $L^{-1} : W \to V$. Assume $M \in \mathcal{L}(V, W)$ satisfies*

$$
\|M - L\| < \frac{1}{\|L^{-1}\|}. \tag{2.3.12}
$$

Then $M : V \to W$ is a bijection, $M^{-1} \in \mathcal{L}(W, V)$ and

$$
\|M^{-1}\| \leq \frac{\|L^{-1}\|}{1 - \|L^{-1}\|\,\|L - M\|}. \tag{2.3.13}
$$

Moreover,

$$
\|L^{-1} - M^{-1}\| \leq \frac{\|L^{-1}\|^2 \|L - M\|}{1 - \|L^{-1}\|\,\|L - M\|}. \tag{2.3.14}
$$

For solutions of the equations $Lv_1 = w$ and $Mv_2 = w$, we have the bound

$$
\|v_1 - v_2\| \leq \|M^{-1}\|\,\|(L - M)\,v_1\|. \tag{2.3.15}
$$

Proof. We write M as a perturbation of L. If W is complete, we write

$$
M = [I - (L - M)\,L^{-1}]\,L;
$$

whereas if V is complete, we write

$$M = L\,[I - L^{-1}(L - M)].$$

Let us prove the result for the case W is complete.

The operator $(L - M)\,L^{-1} \in \mathcal{L}(W)$ satisfies

$$\|(L - M)\,L^{-1}\| \le \|L - M\|\,\|L^{-1}\| < 1.$$

Thus by the geometric series theorem, $[I - (L - M)\,L^{-1}]^{-1}$ exists and

$$\|[I - (L - M)\,L^{-1}]^{-1}\| \le \frac{1}{1 - \|(L - M)\,L^{-1}\|} \le \frac{1}{1 - \|L^{-1}\|\,\|L - M\|}.$$

So M^{-1} exists with

$$M^{-1} = L^{-1}[I - (L - M)\,L^{-1}]^{-1}$$

and

$$\|M^{-1}\| \le \|L^{-1}\|\,\|[I - (L - M)\,L^{-1}]^{-1}\| \le \frac{\|L^{-1}\|}{1 - \|L^{-1}\|\,\|L - M\|}.$$

To prove (2.3.14), we write

$$L^{-1} - M^{-1} = M^{-1}(M - L)\,L^{-1},$$

take norms and use (2.3.13). For (2.3.15), write

$$v_1 - v_2 = (L^{-1} - M^{-1})\,w = M^{-1}(M - L)\,L^{-1}w = M^{-1}(M - L)\,v_1$$

and take norms and bounds. □

The above theorem can be paraphrased as follows: *An operator that is close to an operator with a bounded inverse will itself have a bounded inverse.* This is the framework for innumerable solvability results for linear differential and integral equations, and variations of it are also used with nonlinear operator equations.

The inequality (2.3.14) can be termed the local Lipschitz continuity of the operator inverse. The bound (2.3.15) can be used both as an *a priori* and an *a posteriori* error estimate, depending on the way we use it. First, let us view the equation $Lv = w$ as the exact problem, and we take a sequence of approximation problems $L_n v_n = w$, $n = 1, 2, \ldots$. Assuming the sequence $\{L_n\}$ converges to L, we can apply the perturbation theorem to conclude that at least for sufficiently large n, the equation $L_n v_n = w$ has a unique solution v_n, and we have the error estimate

$$\|v - v_n\| \le \|L_n^{-1}\|\,\|(L - L_n)\,v\|. \tag{2.3.16}$$

The *consistency* of the approximation is defined by the condition

$$\|(L - L_n)\, v\| \to 0,$$

whereas the *stability* is defined by the condition that $\{\|L_n^{-1}\|\}_{n \text{ large}}$ is uniformly bounded. We see that consistency plus stability implies *convergence*:

$$\|v - v_n\| \to 0.$$

The error estimate (2.3.16) provides sufficient conditions for convergence (and order error estimate under regularity assumptions on the solution v) before we actually solve the approximation problem $L_n v_n = w$. Such an estimate is called an *a priori* error estimate. We notice that usually an a priori error estimate does not tell us quantitatively how small is the error.

Another way to use (2.3.15) is to view $Mv = w$ as the exact problem, and $L = M_n$ an approximation of M, $n = 1, 2, \ldots$. Denote by v_n the solution of the approximation equation $M_n v_n = w$; the equation is uniquely solvable at least for sufficiently large n. Then we have the error estimate

$$\|v - v_n\| \le \|M^{-1}\|\, \|(M - M_n)\, v_n\|.$$

Suppose we can estimate the term $\|M^{-1}\|$. Then after the approximate solution v_n is found, the above estimate offers a numerical upper bound for the error. Such an estimate is called an *a posteriori* error estimate.

Example 2.3.6 We examine the solvability of the integral equation

$$\lambda\, u(x) - \int_0^1 \sin(xy)\, u(y)\, dy = f(x), \quad 0 \le x \le 1 \tag{2.3.17}$$

with $\lambda \ne 0$. From the discussion of the Example 2.3.2, if

$$|\lambda| > \|K\| = \int_0^1 \sin(y)\, dy = 1 - \cos(1) \approx 0.4597, \tag{2.3.18}$$

then for every $f \in C[0, 1]$, (2.3.17) admits a unique solution $u \in C[0, 1]$.

To extend the values of λ for which (2.3.17) has a unique solution, we apply the perturbation theorem. Since $\sin(xy) \approx xy$ for small values of $|xy|$, we compare (2.3.17) with

$$\lambda\, v(x) - \int_0^1 x\, y\, v(y)\, dy = f(x), \quad 0 \le x \le 1. \tag{2.3.19}$$

In the notation of the perturbation theorem, equation (2.3.17) is $Mu = f$ and (2.3.19) is $Lv = f$. The normed space is $V = C[0, 1]$ with the norm $\|\cdot\|_\infty$, and $L, M \in \mathcal{L}(V)$.

The integral equation (2.3.19) can be solved explicitly. From (2.3.19), assuming $\lambda \neq 0$, we have that every solution v takes the form

$$v(x) = \frac{1}{\lambda}\,[f(x) + cx]$$

for some constant c. Substituting this back into (2.3.19) leads to a formula for c, and then

$$v(x) = \frac{1}{\lambda}\left[f(x) + \frac{1}{\lambda - 1/3}\int_0^1 x\,y\,f(y)\,dy\right] \quad \text{if } \lambda \neq 0, \frac{1}{3}. \tag{2.3.20}$$

The relation (2.3.20) defines $L^{-1}f$ for all $f \in C[0,1]$.

To use the perturbation theorem, we need to measure several quantities. It can be computed that

$$\|L^{-1}\| \le \frac{1}{|\lambda|}\left(1 + \frac{1}{2\,|\lambda - 1/3|}\right)$$

and

$$\|L - M\| = \int_0^1 (y - \sin y)\,dy = \cos(1) - \frac{1}{2} \approx 0.0403.$$

The condition (2.3.12) is implied by

$$\frac{1}{|\lambda|}\left(1 + \frac{1}{2\,|\lambda - 1/3|}\right) < \frac{1}{\cos(1) - 1/2}. \tag{2.3.21}$$

A graph of the left side of this inequality is given in Figure 2.1. If λ is assumed to be real, then there are three cases to be considered: $\lambda > 1/3$, $0 < \lambda < 1/3$, and $\lambda < 0$. For the case $\lambda < 0$, (2.3.21) is true if and only if $\lambda < \lambda_0 \approx -0.0881$, the negative root of the equation

$$\lambda^2 - \left(\frac{5}{6} - \cos 1\right)\lambda - \frac{5}{6}\left(\cos 1 - \frac{1}{2}\right) = 0.$$

As a consequence of the perturbation theorem, we have that if $\lambda < \lambda_0$, then (2.3.17) is uniquely solvable for all $f \in C[0,1]$. This is a significant improvement over the negative portion of the condition (2.3.18). Bounds can also be given on the solution u, but these are left to the reader, as are the remaining two cases for λ. □

Exercise 2.3.1 Consider the integral equation

$$\lambda u(x) - \int_0^1 \frac{u(y)}{1 + x^2 y^2}\,dy = f(x), \quad 0 \le x \le 1$$

for a given $f \in C[0,1]$. Show this equation has a unique continuous solution u if $|\lambda|$ is chosen sufficiently large. For such values of λ, bound the solution u in terms of $\|f\|_\infty$.

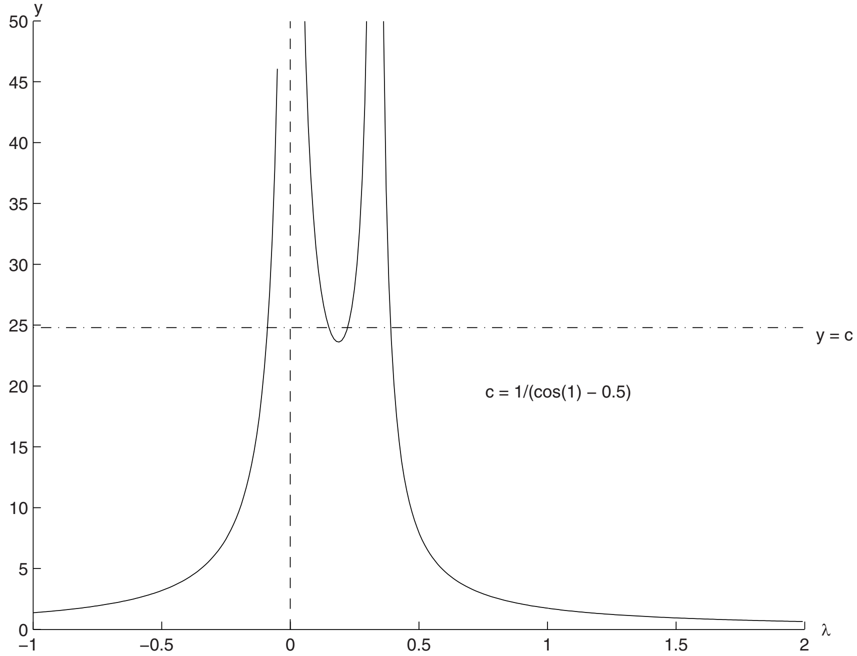

FIGURE 2.1. Graph of the left-hand side of inequality (2.3.21)

Exercise 2.3.2 Show that the integral equation

$$u(x) - \int_0^1 \sin \pi(x-t)\, u(t)\, dt = f(x)$$

has a unique solution $u \in C[0,1]$ for any given $f \in C[0,1]$. As an approximation of the solution u, use the formula (2.3.5) to compute u_2.

Exercise 2.3.3 Let V and W be Banach spaces. Assume $L \in \mathcal{L}(V,W)$ has a bounded inverse $L^{-1} : W \to V$ and $\{M_\lambda\}_{\lambda\in\Lambda} \subset \mathcal{L}(V,W)$ is a family of linear continuous operators such that $\|M_\lambda\| \le \varepsilon$, $\forall\, \lambda \in \Lambda$. Define $L_\lambda = L + M_\lambda$. Find a condition on ε so that given $f \in W$, for any $\lambda \in \Lambda$, the equation $L_\lambda u_\lambda = f$ has a unique solution $u_\lambda \in V$, and the iteration method: $u_0 \in V$ chosen, and for $n = 0, 1, \dots,$

$$u_{n+1} = L^{-1}(f - M_\lambda u_n),$$

converges to the solution u_λ.

Exercise 2.3.4 A linear continuous operator L is said to be nilpotent if for some integer $m \ge 1$, $L^m = 0$. Show that if L is nilpotent, then $(I-L)^{-1}$ exists. Find a formula for $(I-L)^{-1}$.

Exercise 2.3.5 Complete the solvability analysis for Example 2.3.6.

Exercise 2.3.6 Repeat the solvability analysis of Example 2.3.6 for the integral equation

$$\lambda\, u(x) - \int_0^1 u(y) \arctan(xy)\, dy = f(x), \quad 0 \le x \le 1.$$

Use the approximation based on the Taylor approximation

$$\arctan(s) \approx s$$

for small values of s.

Exercise 2.3.7 Let V be a Banach space, $L \in \mathcal{L}(V)$. Assume $\|I - L\| < 1$. Show that L has a bounded inverse, and

$$L^{-1} = \sum_{i=0}^{\infty} (I - L)^i.$$

Exercise 2.3.8 Assume the conditions of the geometric series theorem are satisfied. Then for any $f \in V$, the equation $(I - L)\, u = f$ has a unique solution $u \in V$. Show that this solution can be approximated by a sequence $\{u_n\}$ defined by: $u_0 \in V$, $u_n = f + L u_{n-1}$, $n = 1, 2, \dots$. Derive an error bound for $\|u - u_n\|$.

Exercise 2.3.9 Let $f \in C[0,1]$. Show that the continuous solution of the boundary value problem

$$-\,u''(x) = f(x), \quad 0 < x < 1,$$
$$u(0) = u(1) = 0$$

is

$$u(x) = \int_0^1 k(x,y)\, f(y)\, dy,$$

where the kernel function $k(x,y) = \min(x,y)\,(1 - \max(x,y))$. Let $a \in C[0,1]$. Apply the geometric series theorem to show that the boundary value problem

$$-\,u''(x) + a(x)\, u(x) = f(x), \quad 0 < x < 1,$$
$$u(0) = u(1) = 0$$

has a unique continuous solution u if $\max_{0 \le x \le 1} |a(x)| \le a_0$ is sufficiently small. Give an estimate of the value a_0.

Exercise 2.3.10 Let V and W be Banach spaces. Assume $L \in \mathcal{L}(V, W)$ has a bounded inverse $L^{-1} : W \to V$ and $M \in \mathcal{L}(V, W)$ satisfies

$$\|M - L\| < \frac{1}{2\,\|L^{-1}\|}.$$

For any $f \in W$, the equation $Lu = f$ has a unique solution $u \in V$ which is approximated by the following iteration method: choose an initial guess $u_0 \in V$ and define

$$u_{n+1} = u_n + M^{-1}(f - L u_n), \quad n \ge 0.$$

Prove the convergence $u_n \to u$ as $n \to \infty$.

This iteration method is useful where it is much easier to compute $M^{-1} g$ than $L^{-1} g$ for $g \in W$.

Exercise 2.3.11 Recall the Volterra equation (2.3.9). Bound the solution u using Corollary 2.3.3. Separately, obtain a bound for u by examining directly the convergence of the series

$$u = \sum_{k=0}^{\infty} L^k f$$

and relating it to the Taylor series for $\exp(M\,B)$.

2.4 Some more results on linear operators

In this section, we collect together several independent results which are important in working with linear operators.

2.4.1 An extension theorem

Bounded operators are often defined on a subspace of a larger space, and it is desirable to extend the domain of the original operator to the larger space, while retaining the boundedness of the operator.

Theorem 2.4.1 (Extension theorem) *Let V be a normed space, and let $\widehat{V}$ denote its completion. Let W be a Banach space. Assume $L \in \mathcal{L}(V, W)$. Then there is a unique operator $\widehat{L} \in \mathcal{L}(\widehat{V}, W)$ with*

$$\widehat{L}v = L\,v \quad \forall\, v \in V$$

and

$$\|\widehat{L}\|_{\widehat{V},W} = \|L\|_{V,W}.$$

The operator $\widehat{L}$ is called an extension *of L.*

Proof. Given $v \in \widehat{V}$, let $\{v_n\} \subset V$ with $v_n \to v$ in $\widehat{V}$. The sequence $\{Lv_n\}$ is a Cauchy sequence in W by the following inequality

$$\|Lv_{n+p} - Lv_n\| \le \|L\|\,\|v_{n+p} - v_n\|.$$

Since W is complete, there is a limit $\widehat{L}(v) \in W$. We must show that $\widehat{L}$ is well-defined (i.e., $\widehat{L}(v)$ does not depend on the choice of the sequence $\{v_n\}$), linear and bounded.

To show $\widehat{L}$ is well-defined, let $v_n \to v$ and $\tilde{v}_n \to v$ with $\{v_n\}, \{\tilde{v}_n\} \subset V$. Then as $n \to \infty$,

$$\|Lv_n - L\tilde{v}_n\| \le \|L\|\,\|v_n - \tilde{v}_n\| \le \|L\|\,(\|v_n - v\| + \|\tilde{v}_n - v\|) \to 0.$$

Thus $\{Lv_n\}$ and $\{L\tilde{v}_n\}$ must have the same limit.

To show the linearity, let $u_n \to u$ and $v_n \to v$, and let $\alpha, \beta \in \mathbb{K}$. Then

$$\widehat{L}(\alpha\, u + \beta\, v) = \lim_{n\to\infty} L(\alpha\, u_n + \beta\, v_n) = \lim_{n\to\infty} (\alpha\, L u_n + \beta\, L v_n) = \alpha\, \widehat{L} u + \beta\, \widehat{L} v.$$

To show the boundedness, let $v_n \to v$ and $\{v_n\} \subset V$. Then taking the limit $n \to \infty$ in

$$\|L v_n\|_W \le \|L\|\, \|v_n\|_V = \|L\|\, \|v_n\|_{\widehat{V}},$$

we obtain

$$\|\widehat{L} v\|_W \le \|L\|\, \|v\|_{\widehat{V}}.$$

So $\widehat{L}$ is bounded and

$$\|\widehat{L}\| = \sup_{0 \ne v \in \widehat{V}} \frac{\|\widehat{L} v\|_W}{\|v\|_{\widehat{V}}} \le \|L\|.$$

To see that $\|\widehat{L}\|_{\widehat{V},W} = \|L\|_{V,W}$, we note

$$\|L\|_{V,W} = \sup_{0 \ne v \in V} \frac{\|L v\|_W}{\|v\|_V} = \sup_{0 \ne v \in V} \frac{\|\widehat{L} v\|_W}{\|v\|_{\widehat{V}}} \le \sup_{0 \ne v \in \widehat{V}} \frac{\|\widehat{L} v\|_W}{\|v\|_{\widehat{V}}} = \|\widehat{L}\|_{\widehat{V},W}.$$

To show that $\widehat{L}$ is unique, let $\widetilde{L}$ be another extension of L to $\widehat{V}$. Let $v \in \widehat{V}$ and let $v_n \to v$, $\{v_n\} \subset V$. Then

$$\|\widetilde{L} v - L v_n\|_W = \|\widetilde{L} v - \widetilde{L} v_n\|_W \le \|\widetilde{L}\|\, \|v - v_n\| \to 0 \quad \text{as } n \to \infty.$$

This shows $L v_n \to \widetilde{L} v$ as $n \to \infty$. On the other hand, $L v_n \to \widehat{L} v$. So we must have $\widetilde{L} v = \widehat{L} v$, for any $v \in \widehat{V}$. Therefore, $\widetilde{L} = \widehat{L}$. □

There are a number of ways in which this theorem can be used. Often we wish to work with linear operators which are defined and bounded on some normed space, but the space is not complete with the given norm. Since most function space arguments require complete spaces, the above theorem allows us to proceed with our arguments on a larger complete space, with an operator which agrees with our original one on the original space.

Example 2.4.2 Let $V = C^1[0,1]$ with the inner product norm

$$\|v\|_{1,2} = (\|v\|_2^2 + \|v'\|_2^2)^{1/2}.$$

The completion of $C^1[0,1]$ with respect to $\|\cdot\|_{1,2}$ is the Sobolev space $H^1(0,1)$, which was introduced earlier in Example 1.3.7. (Details of Sobolev spaces are given later in Chapter 7.) Let $W = L^2(0,1)$ with the standard norm $\|\cdot\|_2$.

Define the differentiation operator $D : C^1[0,1] \to L^2(0,1)$ by

$$(Dv)(x) = v'(x), \quad 0 \le x \le 1,\ v \in C^1[0,1].$$

We have

$$\|Dv\|_2 = \|v'\|_2 \le \|v\|_{1,2},$$

and thus

$$\|D\|_{V,W} \le 1.$$

By the extension theorem, we can extend D to $\widehat{D} \in \mathcal{L}(H^1(0,1), L^2(0,1))$, the differentiation operator on $H^1(0,1)$. A more concrete realization of $\widehat{D}$ can be obtained using the notion of weak derivatives. This is also discussed in Chapter 7. □

2.4.2 Open mapping theorem

This theorem is widely used in obtaining boundedness of inverse operators. When considered in the context of solving an equation $Lv = w$, the theorem says that *existence* and *uniqueness* of solutions for all $w \in W$ implies the *stability* of the solution v, i.e., "small changes" in the given data w causes only "small changes" in the solution v. For a proof of this theorem, see [58, p. 91] or [250, p. 179].

Theorem 2.4.3 *Let V and W be Banach spaces. If $L \in \mathcal{L}(V,W)$ is a bijection, then $L^{-1} \in \mathcal{L}(W,V)$.*

To be more precise concerning the stability of the problem being solved, let $Lv = w$ and $L\hat{v} = \hat{w}$. We then have

$$v - \hat{v} = L^{-1}(w - \hat{w}),$$

and then

$$\|v - \hat{v}\| \le \|L^{-1}\| \, \|w - \hat{w}\|.$$

As $w - \hat{w}$ becomes small, so must $v - \hat{v}$. The term $\|L^{-1}\|$ gives a relationship between the size of the error in the data w and that of the error in the solution v. A more important way is to consider the relative changes in the two errors:

$$\frac{\|v - \hat{v}\|}{\|v\|} \le \frac{\|L^{-1}\| \, \|w - \hat{w}\|}{\|v\|} = \|L^{-1}\| \|L\| \frac{\|w - \hat{w}\|}{\|L\| \, \|v\|}.$$

Applying $\|w\| \le \|L\| \, \|v\|$, we obtain

$$\frac{\|v - \hat{v}\|}{\|v\|} \le \|L^{-1}\| \, \|L\| \frac{\|w - \hat{w}\|}{\|w\|}. \tag{2.4.1}$$

The quantity $\operatorname{cond}(L) \equiv \|L^{-1}\| \, \|L\|$ is called the *condition number* of the equation, and it relates the relative errors in the data w and in the solution v. Note that we always have $\operatorname{cond}(L) \ge 1$ as

$$\|L^{-1}\| \, \|L\| \ge \|L^{-1}L\| = \|I\| = 1.$$

Problems with a small condition number are called *well-conditioned*, whereas those with a large condition number *ill-conditioned*.

In a related vein, consider a problem $Lv = w$, $L : V \to W$, in which L is bounded and injective, but not surjective. The inverse operator L^{-1} exists on the range $\mathcal{R}(L) \subset W$. If L^{-1} is unbounded on $\mathcal{R}(L)$ to V, we say the original problem $Lv = w$ is *ill-posed* or *unstable*. Such problems are not considered in this text, but there are a number of important applications (e.g. many indirect sensing devices) which fall into this category. Problems in which L^{-1} is bounded (along with L) are called *well-posed* or *stable*; they can still be ill-conditioned, as was discussed in the preceding paragraph.

2.4.3 Principle of uniform boundedness

Another widely used set of results refer to the collective boundedness of a set of linear operators.

Theorem 2.4.4 *Let $\{L_n\}$ be a sequence of bounded linear operators from a Banach space V to a normed space W. Assume for every $v \in V$, the sequence $\{L_n v\}$ is bounded. Then*

$$\sup_n \|L_n\| < \infty.$$

This theorem is often called the *principle of uniform boundedness*; see [58, p. 95] or [250, p. 172] for a proof and a more extended development. We also have the following useful variant of this principle.

Theorem 2.4.5 (Banach-Steinhaus theorem) *Let V and W be normed spaces with V being complete, and let $L, L_n \in \mathcal{L}(V, W)$. Let V_0 be a dense subspace of V. Then in order for $L_n v \to Lv \ \forall\, v \in V$, it is necessary and sufficient that*
(a) $L_n v \to Lv \ \forall\, v \in V_0$; *and*
(b) $\sup_n \|L_n\| < \infty$.

Proof. ($\Rightarrow$) Assume $L_n v \to Lv$ for all $v \in V$. Then (a) follows trivially; and (b) follows from the principle of uniform boundedness.
($\Leftarrow$) Assume (a) and (b). Denote $B = \sup_n \|L_n\|$. Let $v \in V$ and $\epsilon > 0$. By the denseness of V_0 in V, there is an element $v_\epsilon \in V_0$ such that

$$\|v - v_\epsilon\| \le \frac{\epsilon}{3\max\{\|L\|, B\}}.$$

Then

$$\begin{aligned}\|Lv - L_n v\| &\le \|Lv - Lv_\epsilon\| + \|Lv_\epsilon - L_n v_\epsilon\| + \|L_n v_\epsilon - L_n v\| \\ &\le \|L\|\,\|v - v_\epsilon\| + \|Lv_\epsilon - L_n v_\epsilon\| + \|L_n\|\,\|v_\epsilon - v\| \\ &\le \frac{2\,\epsilon}{3} + \|Lv_\epsilon - L_n v_\epsilon\|.\end{aligned}$$

Using (a), we can find a natural number n_ϵ such that

$$\|Lv_\epsilon - L_n v_\epsilon\| \le \frac{\epsilon}{3}, \quad n \ge n_\epsilon.$$

Combining these results,

$$\|Lv - L_n v\| \le \epsilon, \quad n \ge n_\epsilon.$$

Therefore, $L_n v \to Lv$ as $n \to \infty$. □

Next, we apply Banach-Steinhaus theorem to discuss the convergence of numerical quadratures (i.e., numerical integration formulas).

2.4.4 Convergence of numerical quadratures

As an example, let us consider the convergence of numerical quadratures for the computation of the integral

$$Lv = \int_0^1 w(x)\, v(x)\, dx,$$

where w is a weighted function, $w(x) \ge 0$, $w \in L^1(0,1)$. There are several approaches to constructing numerical quadratures. One popular approach is to replace the function v by some interpolant of it, denoted here by Πv, and then define the corresponding numerical quadrature by the formula

$$\int_0^1 w(x)\, \Pi v(x)\, dx.$$

The topic of function interpolation is discussed briefly in Section 3.2. If Πv is taken to be the Lagrange polynomial interpolant of v on a uniform partition of the integration interval and the weight function $w(x) = 1$, the resulting quadratures are called *Newton-Cotes integration formulas*. It is well-known that high degree polynomial interpolation on a uniform partition leads to strong oscillations near the boundary of the interval and hence divergence of the interpolation in many cases. Correspondingly, one cannot expect the convergence of the Newton-Cotes integration formulas. To guarantee the convergence, one may use the Lagrange polynomial interpolant Πv of v on a properly chosen partition with more nodes placed near the boundary, or one may use a piecewise polynomial interpolant on a uniform partition or a partition suitably refined in areas where the integrand $w\,v$ changes rapidly. With the use of piecewise polynomial interpolants of v, we get the celebrated *trapezoidal rule* (using piecewise linear interpolation) and *Simpson's rule* (using piecewise quadratic interpolation).

A second popular approach to constructing numerical quadratures is by the *method of undetermined parameters*. We approximate the integral by a sequence of finite sums, each of them being a linear combination of some function values (and more generally, derivative values can used in the sums as well). In other words, we let

$$Lv \approx L_n v = \sum_{i=0}^{n} w_i^{(n)} v(x_i^{(n)}) \tag{2.4.2}$$

and choose the weights $\{w_i^{(n)}\}_{i=0}^n$ and the nodes $\{x_i^{(n)}\}_{i=0}^n \subset [0,1]$ by some specific requirements. Some of the weights and nodes may be prescribed *a priori* according to the context of the applications, and the remaining ones are usually determined by requiring the quadrature be exact for polynomials of degree as high as possible. If none of the weights and nodes is prescribed, then we may choose these $2n+2$ quantities so that the quadrature is exact for any polynomial of degree less than or equal to $2n+1$. The resulting numerical quadratures are called *Gaussian quadratures*.

Detailed discussions of numerical quadratures (for the case when the weight function $w \equiv 1$) can be found in [15, Section 5.3]. Here we study the convergence of numerical quadratures in an abstract framework.

Let there be given a sequence of quadratures

$$L_n v = \sum_{i=0}^{n} w_i^{(n)} v(x_i^{(n)}), \tag{2.4.3}$$

where $0 \le x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} \le 1$ is a partition of $[0,1]$. We regard L_n as a linear functional (i.e. a continuous linear operator with scalar values, see next section) defined on $C[0,1]$ with the standard uniform norm. It is straightforward to show that

$$\|L_n\| = \sum_{i=0}^{n} |w_i^{(n)}|, \tag{2.4.4}$$

and this is left as an exercise for the reader.

As an important special case, assume the quadrature scheme L_n has a degree of precision $d(n)$, i.e. the scheme is exact for polynomials of degree less than or equal to $d(n)$,

$$L_n v = Lv \quad \forall\, v \in \mathbb{P}_{d(n)}.$$

Here $\mathbb{P}_{d(n)}$ is the space of all the polynomials of degree less than or equal to $d(n)$, and we assume $d(n) \to \infty$ as $n \to \infty$. Note that the subspace V_0 of all the polynomials is dense in $V = C[0,1]$. Then an application of the Banach-Steinhaus theorem shows that $L_n v \to Lv$ for any $v \in C[0,1]$ if and only if

$$\sup_n \sum_{i=0}^{n} |w_i^{(n)}| < \infty.$$

Continuing the discussion on the convergence of numerical quadratures, we assume all the conditions stated in the previous paragraph are valid. Additionally, we assume the weights $w_i^{(n)} \ge 0$. Then it follows that $L_n v \to Lv$ for any $v \in C[0,1]$ (Exercise 2.4.3).

From the point of view of numerical computations, it is important to have non-negative quadrature weights to avoid round-off error accumulations. It can be shown that for the Gaussian quadratures, all the quadrature weights are non-negative; and if the weight function w is positive on $(0,1)$, then the quadrature weights are positive. See [15, Section 5.3] for an extended discussion of Gaussian quadrature.

In the above discussion, we consider the situation where the degree of precision of the integration formula is increasingly large. Another situation we can consider is where the degree of precision of the integration formula does not increase and the integration formulas integrate exactly piecewise low degree polynomials corresponding to more and more integrating nodes; see Exercise 2.4.5 for such an example.

Exercise 2.4.1 Assume V is a Banach space with either of two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Suppose for some constant $c > 0$,

$$\|v\|_2 \le c\,\|v\|_1 \quad \forall\, v \in V.$$

Show that the two norms are equivalent.

Exercise 2.4.2 Prove the formula (2.4.4).

Exercise 2.4.3 Consider the quadrature formula (2.4.3). Assume all the weights $w_i^{(n)}$ are non-negative and the quadrature formula is exact for polynomials of degree less than or equal to $d(n)$ with $d(n) \to \infty$ as $n \to \infty$. Prove the convergence of the quadrature: $L_n v \to Lv$, for all $v \in C[0,1]$.

Exercise 2.4.4 Let V and W be two Banach spaces. Assume $\{L_n\} \subset \mathcal{L}(V,W)$ is such that for any $v \in V$,

$$\lim_{n\to\infty} L_n(v) = L(v)$$

exists. Show that $L \in \mathcal{L}(V,W)$ and

$$\|L\| \le \liminf_{n\to\infty} \|L_n\|.$$

Hint: Apply the principle of uniform boundedness (Theorem 2.4.4) on the operator sequence $\{L_n\}$.

Exercise 2.4.5 A popular family of numerical quadratures is constructed by approximating the integrand by its piecewise polynomial interpolants. We take the composite trapezoidal rule as an example. The integral to be computed is

$$Lv = \int_0^1 v(x)\,dx.$$

We divide the interval $[0,1]$ into n equal parts, and denote $x_i = i/n$, $0 \le i \le n$, as the nodes. Then we approximate v by its piecewise linear interpolant $\Pi_n v$ defined by

$$\Pi_n v(x) = n\,(x_i - x)\,v(x_{i-1}) + n\,(x - x_{i-1})\,v(x_i)$$

for $x_{i-1} \le x \le x_i$, $1 \le i \le n$. Then the composite trapezoidal rule is

$$L_n v = \int_0^1 \Pi_n v(x)\,dx = \frac{1}{n}\left[\frac{1}{2}v(x_0) + \sum_{i=1}^{n-1} v(x_i) + \frac{1}{2}v(x_n)\right].$$

Show that $L_n v \to Lv$ for any $v \in C[0,1]$.

Using piecewise polynomials of higher degrees based on non-uniform partitions of the integration interval, we can develop other useful numerical quadratures.

Exercise 2.4.6 In the formula (2.4.1), show that the inequality can be made as close as desired to equality for suitable choices of v and $\widetilde{v}$. More precisely, show that

$$\sup_{v,\widetilde{v}}\left(\frac{\|v-\widetilde{v}\|}{\|v\|} \div \frac{\|w-\widetilde{w}\|}{\|w\|}\right) = \|L\|\,\|L^{-1}\|.$$

Exercise 2.4.7 Let the numerical integration formula $L_n v$ of (2.4.2) be the Newton-Cotes formula that is described earlier in Subsection 2.4.4. It is known that there are continuous functions $v \in C[0,1]$ for which $L_n v \not\to Lv$ as $n \to \infty$. Accepting this, show that

$$\sup_n \sum_{i=0}^{n} \left|w_i^{(n)}\right| = \infty.$$

Moreover, show

$$\sum_{i=0}^{n} w_i^{(n)} = 1.$$

These results imply that the quadrature weights must be of varying sign and that they must be increasing in size as $n \to \infty$.

2.5 Linear functionals

An important special case of linear operators is when they take on scalar values. Let V be a normed space, and $W = \mathbb{K}$, the set of scalars associated with V. The elements in $\mathcal{L}(V,\mathbb{K})$ are called *linear functionals*. Since $\mathbb{K}$ is complete, $\mathcal{L}(V,\mathbb{K})$ is a Banach space. This space is usually denoted as V' and it is called the *dual space* of V. Usually we use lower case letters, such as ℓ, to denote a linear functional.

In some references, the term *linear functional* is used for the linear operators from a normed space to $\mathbb{K}$, without the functionals being necessarily bounded. In this work, since we use exclusively linear functionals which are bounded, we use the term "linear functionals" to refer to only bounded linear functionals.

Example 2.5.1 Let $\Omega \subset \mathbb{R}^d$ be a bounded open set. It is a well-known result that for $1 \le p < \infty$, the dual space of $L^p(\Omega)$ can be identified with $L^{p'}(\Omega)$. Here p' is the conjugate exponent of p, defined by the relation

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

By convention, $p' = \infty$ when $p = 1$. In other words, given an $\ell \in (L^p(\Omega))'$, there is a function $u \in L^{p'}(\Omega)$, uniquely determined a.e., such that

$$\ell(v) = \int_\Omega u(\boldsymbol{x})v(\boldsymbol{x})\,dx \quad \forall\, v \in L^p(\Omega). \tag{2.5.1}$$

Conversely, for any $u \in L^{p'}(\Omega)$, the rule

$$v \longmapsto \int_\Omega u(\boldsymbol{x})v(\boldsymbol{x})\,dx, \quad v \in L^p(\Omega)$$

defines a bounded linear functional on $L^p(\Omega)$. It is convenient to identify $\ell \in (L^p(\Omega))'$ and $u \in L^{p'}(\Omega)$, related as in (2.5.1). Then we write

$$(L^p(\Omega))' = L^{p'}(\Omega), \quad 1 \le p < \infty.$$

When $p = 2$, we have $p' = 2$. This special case is examined later in Example 2.5.9 from another perspective. For $p = \infty$, $p' = 1$. The dual space of $L^\infty(\Omega)$, however, is larger than the space $L^1(\Omega)$. □

All the results discussed in the previous sections for general linear operators certainly apply to linear functionals. In addition, there are useful results particular to linear functionals only.

2.5.1 An extension theorem for linear functionals

We have seen that a bounded linear operator can be extended to the closure of its domain. It is also possible to extend linear functionals defined on an arbitrary subspace to the whole space.

Theorem 2.5.2 (Hahn-Banach theorem) *Let V_0 be a subspace of a normed space V, and $\ell : V_0 \to \mathbb{K}$ be linear and bounded. Then there exists an extension $\hat{\ell} \in V'$ of ℓ with $\hat{\ell}(v) = \ell(v)\ \forall\, v \in V_0$, and $\|\hat{\ell}\| = \|\ell\|$.*

A proof can be found in [58, p. 79] or [250, p. 4]. This theorem can be proved by applying the Generalized Hahn-Banach Theorem, Theorem 2.5.5 (Exercise 2.5.1). Note that if V_0 is not dense in V, then the extension need not be unique.

Example 2.5.3 This example is important in the analysis of some numerical methods for solving integral equations. Let $V = L^\infty(0,1)$. This is the space of all cosets (or equivalence classes)

$$\boldsymbol{v} = [v] = \{w \text{ Lebesgue measurable on } [0,1] \mid w = v \text{ a.e. in } [0,1]\}$$

for which

$$\|\boldsymbol{v}\|_\infty \equiv \|v\|_\infty = \operatorname*{ess\,sup}_{0\le x\le 1}|v(x)| < \infty. \tag{2.5.2}$$

With this norm, $L^\infty(0,1)$ is a Banach space.

Let V_0 be the set of all cosets $\boldsymbol{v} = [v]$, where $v \in C[0,1]$. It is a proper subspace of $L^\infty(0,1)$. When restricted to V_0, the norm (2.5.2) is equivalent to the usual norm $\|\cdot\|_\infty$ on $C[0,1]$. It is common to write $V_0 = C[0,1]$; but this is an abuse of notation, and it is important to keep in mind the distinction between V_0 and $C[0,1]$.

Let $c \in [0,1]$, and define

$$\ell_c([v]) = v(c) \quad \forall\, v \in C[0,1]. \tag{2.5.3}$$

The linear functional $\ell_c([v])$ is well-defined on V_0. From

$$|\ell_c([v])| = |v(c)| \le \|v\|_\infty = \|\boldsymbol{v}\|_\infty, \quad \boldsymbol{v} = [v],$$

we see that $\|\ell_c\| \le 1$. By choosing $v \in C[0,1]$ with $v(c) = \|v\|_\infty$, we then obtain

$$\|\ell_c\| = 1.$$

Using Hahn-Banach Theorem, we can extend ℓ_c to $\hat{\ell}_c : L^\infty(0,1) \to \mathbb{K}$ with

$$\|\hat{\ell}_c\| = \|\ell_c\| = 1.$$

The functional $\hat{\ell}_c$ extends to $L^\infty(0,1)$ the concept of point evaluation of a function, to functions which are only Lebesgue measurable and which are not precisely defined because of being members of a coset.

Somewhat surprisingly, many desirable properties of ℓ_c are carried over to $\hat{\ell}_c$. These include the following.

- Assume $[v] \in L^\infty(0,1)$ satisfies $m \le v(x) \le M$ for almost all x in some open interval about c. Then $m \le \hat{\ell}_c([v]) \le M$.
- Assume c is a point of continuity of v. Then

$$\hat{\ell}_c([v]) = v(c),$$
$$\lim_{a\to c} \hat{\ell}_a([v]) = v(c).$$

These ideas and properties carry over to $L^\infty(D)$, with D a closed, bounded set in $\mathbb{R}^d$, $d \ge 1$, and $\boldsymbol{c} \in D$. For additional detail and the application of this extension to numerical integral equations, see [22]. □

In some applications, a stronger form of the Hahn-Banach Theorem is needed. We begin by introducing another useful concept for functionals.

Definition 2.5.4 *A functional p on a real vector space V is said to be* sublinear *if*

$$p(u+v) \le p(u) + p(v) \quad \forall\, u, v \in V,$$
$$p(\alpha v) = \alpha\, p(v) \quad \forall\, v \in V,\ \forall\, \alpha \ge 0.$$

We note that a semi-norm is a sublinear functional. A proof of the following result can be found in [58, p. 78] or [250, p. 2].

Theorem 2.5.5 (Generalized Hahn-Banach Theorem) *Let V be a linear space, $V_0 \subset V$ a subspace. Suppose $p : V \to \mathbb{R}$ is a sublinear functional and $\ell : V_0 \to \mathbb{R}$ a linear functional such that $\ell(v) \le p(v)$ for all $v \in V_0$. Then ℓ can be extended to V such that $\ell(v) \le p(v)$ for all $v \in V$.*

Note that $p(v) = c\,\|v\|_V$, c a positive constant, is a sublinear functional on V. With this choice of p, we obtain the original Hahn-Banach Theorem. Another useful consequence of the generalized Hahn-Banach Theorem is the following result, its proof being left as Exercise 2.5.2.

Corollary 2.5.6 *Let V be a normed space. For any $0 \ne v \in V$, there exists $\ell_v \in V'$ such that $\|\ell_v\| = 1$ and $\ell_v(v) = \|v\|$.*

We can use Corollary 2.5.6 to characterize the norm in a normed linear space.

Corollary 2.5.7 *Let V be a normed space. Then for any $v \in V$,*

$$\|v\| = \sup\{|\ell(v)| \mid \ell \in V',\ \|\ell\| = 1\}. \tag{2.5.4}$$

Proof. The result is obvious for $v = 0$. Assume $v \ne 0$. For any $\ell \in V'$ with $\|\ell\| = 1$, we have

$$|\ell(v)| \le \|\ell\|\, \|v\| = \|v\|.$$

By Corollary 2.5.6, we have an $\ell_v \in V'$ with $\|\ell_v\| = 1$ such that $\ell_v(v) = \|v\|$. Hence, the equality (2.5.4) holds. □

2.5.2 The Riesz representation theorem

On Hilbert spaces, linear functionals are limited in the forms they can take. The following theorem makes this more precise; and the result is one used in developing the solvability theory for some important partial differential equations and boundary integral equations. The theorem also provides a tool for introducing the concept of the adjoint of a linear operator in the next section.

Theorem 2.5.8 (Riesz representation theorem) *Let V be a real or complex Hilbert space, $\ell \in V'$. Then there is a unique $u \in V$ for which*

$$\ell(v) = (v, u) \quad \forall\, v \in V. \tag{2.5.5}$$

In addition,

$$\|\ell\| = \|u\|. \tag{2.5.6}$$

Proof. Assuming the existence of u, we first prove its uniqueness. Suppose $\tilde{u} \in V$ satisfies

$$\ell(v) = (v, u) = (v, \tilde{u}) \quad \forall\, v \in V.$$

Then

$$(v, u - \tilde{u}) = 0 \quad \forall\, v \in V.$$

Take $v = u - \tilde{u}$. Then $\|u - \tilde{u}\| = 0$, which implies $u = \tilde{u}$.

We give two derivations of the existence of u, both for the case of a real Hilbert space.

Standard proof of existence. Denote

$$N = \mathcal{N}(\ell) = \{v \in V \mid \ell(v) = 0\},$$

which is a subspace of V. If $N = V$, then $\|\ell\| = 0$, and we may take $u = 0$.

Now suppose $N \neq V$. Then there exists at least one $v_0 \in V$ such that $\ell(v_0) \neq 0$. It is possible to decompose V as the direct sum of N and $N^\perp$ (see Section 3.6). From this, we have the decomposition $v_0 = v_1 + v_2$ with $v_1 \in N$ and $v_2 \in N^\perp$. Then $\ell(v_2) = \ell(v_0) \neq 0$.

For any $v \in V$, we have the property

$$\ell\left(v - \frac{\ell(v)}{\ell(v_2)} v_2\right) = 0.$$

Thus

$$v - \frac{\ell(v)}{\ell(v_2)} v_2 \in N,$$

and in particular, it is orthogonal to v_2:

$$\left(v - \frac{\ell(v)}{\ell(v_2)} v_2, v_2\right) = 0,$$

i.e.,

$$\ell(v) = \left(v, \frac{\ell(v_2)}{\|v_2\|^2} v_2\right).$$

In other words, we may take u to be $[\ell(v_2)/\|v_2\|^2]v_2$.

Proof using a minimization principle. From Theorem 3.3.12 in Chapter 3, we know the problem

$$\inf_{v \in V} \left[\frac{1}{2}\|v\|^2 - \ell(v)\right]$$

has a unique solution $u \in V$. The solution u is characterized by the relation (2.5.5).

We complete the proof of the theorem by showing (2.5.6). From (2.5.5) and the Schwarz inequality,

$$|\ell(v)| \le \|u\| \, \|v\| \quad \forall\, v \in V.$$

Hence

$$\|\ell\| \le \|u\|.$$

Let $v = u$ in (2.5.5). Then

$$\ell(u) = \|u\|^2$$

and

$$\|\ell\| = \sup_{v \ne 0} \frac{|\ell(v)|}{\|v\|} \ge \frac{|\ell(u)|}{\|u\|} \ge \|u\|.$$

Therefore, (2.5.6) holds. □

This theorem seems very straightforward, and its proof seems fairly simple. Nonetheless, this is a fundamental tool in the solvability theory for elliptic partial differential equations, as we see later in Chapter 8.

Example 2.5.9 Let $\Omega \subset \mathbb{R}^d$ be open bounded. $V = L^2(\Omega)$ is a Hilbert space. By the Riesz representation theorem, there is a one-to-one correspondence between V' and V by the relation (2.5.5). We can identify $\ell \in V'$ with $u \in V$ related by (2.5.5). In this sense, $(L^2(\Omega))' = L^2(\Omega)$. □

For the space $L^2(\Omega)$, the element u in (2.5.5) is almost always immediately apparent. But for spaces such as the Sobolev space $H^1(a, b)$ introduced in Example 1.3.7 of Chapter 1, the determination of u of (2.5.5) is often not as obvious. As an example, define $\ell \in (H^1(a, b))'$ by

$$\ell(v) = v(c), \quad v \in H^1(a, b) \tag{2.5.7}$$

for some $c \in [a, b]$. This linear functional can be shown to be well-defined (Exercise 2.5.5). From the Riesz representation theorem, there is a unique $u \in H^1(a, b)$ such that

$$\int_a^b [u'(x)\, v'(x) + u(x)\, v(x)]\, dx = v(c) \quad \forall\, v \in H^1(a, b). \tag{2.5.8}$$

The element u is the generalized solution of the boundary value problem

$$\begin{aligned} -u'' + u &= \delta(x - c) \ \text{ in } (a, b), \\ u'(a) &= u'(b) = 0, \end{aligned}$$

where $\delta(x-c)$ is the Dirac δ-function at c. This boundary value problem can be written equivalently as

$$\begin{aligned} -u'' + u &= 0 \quad \text{in } (a,c)\cup(c,b), \\ u'(a) &= u'(b) = 0, \\ u(c-) &= u(c+), \\ u'(c-) - u'(c+) &= 1. \end{aligned}$$

Exercise 2.5.1 Use Theorem 2.5.5 to prove Theorem 2.5.2.
Hint: Take $p(v) = \|\ell\|\,\|v\|$.

Exercise 2.5.2 Prove Corollary 2.5.6.

Exercise 2.5.3 Let us discuss the Riesz representation theorem for a finite-dimensional inner product space. Suppose ℓ is a linear functional on $\mathbb{R}^d$. Show directly the formula

$$\ell(\boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{l}) \quad \forall\, \boldsymbol{x} \in \mathbb{R}^d.$$

Determine the vector $\boldsymbol{l} \in \mathbb{R}^d$.

Exercise 2.5.4 Suppose $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis of a real Hilbert space V. Show that the element $u \in V$ satisfying (2.5.5) is given by the formula

$$u = \sum_{j=1}^{\infty} \ell(e_j)\, e_j.$$

Exercise 2.5.5 Show that the functional defined in (2.5.7) is linear and bounded on $H^1(a,b)$.
Hint: Use the following results. For any $f \in H^1(a,b)$, f is continuous on $[a,b]$, and therefore,

$$\int_a^b f(x)\,dx = f(\zeta)$$

for some $\zeta \in (a,b)$. In addition,

$$f(c) = f(\zeta) + \int_\zeta^c f'(x)\,dx.$$

Exercise 2.5.6 Find the function u of (2.5.8) through the boundary value problem it satisfies.

2.6 Adjoint operators

The notion of an adjoint operator is a generalization of the matrix transpose to infinite dimensional spaces. First let us derive a defining property for the matrix transpose. Let $A \in \mathbb{R}^{m\times n}$, which is viewed as a linear continuous operator from $\mathbb{R}^n$ to $\mathbb{R}^m$. We use the conventional Euclidean inner products for the spaces $\mathbb{R}^n$ and $\mathbb{R}^m$. Then

$$\boldsymbol{y}^T A\boldsymbol{x} = (A\boldsymbol{x}, \boldsymbol{y})_{\mathbb{R}^m}, \quad \boldsymbol{x}^T A^T \boldsymbol{y} = (\boldsymbol{x}, A^T\boldsymbol{y})_{\mathbb{R}^n} \quad \forall\, \boldsymbol{x} \in \mathbb{R}^n,\ \boldsymbol{y} \in \mathbb{R}^m.$$

Since $\boldsymbol{y}^T A\boldsymbol{x}$ is a real number, $\boldsymbol{y}^T A\boldsymbol{x} = (\boldsymbol{y}^T A\boldsymbol{x})^T = \boldsymbol{x}^T A^T \boldsymbol{y}$. We observe that the transpose (or adjoint) A^T is uniquely defined by the property

$$(A\boldsymbol{x}, \boldsymbol{y})_{\mathbb{R}^m} = (\boldsymbol{x}, A^T\boldsymbol{y})_{\mathbb{R}^n} \quad \forall\, \boldsymbol{x} \in \mathbb{R}^n,\ \boldsymbol{y} \in \mathbb{R}^m.$$

Turning now to the general situation, assume V and W are Hilbert spaces, $L \in \mathcal{L}(V, W)$. Let us use the Riesz representation theorem to define a new operator $L^* : W \to V$, called the *adjoint* of L. For simplicity, we assume in this section that $\mathbb{K} = \mathbb{R}$ for the set of scalars associated with W and V. Given $w \in W$, define a linear functional $\ell_w \in V'$ by

$$\ell_w(v) = (Lv, w)_W \quad \forall\, v \in V.$$

This linear functional is bounded because

$$|\ell_w(v)| \le \|Lv\|\,\|w\| \le \|L\|\,\|v\|\,\|w\|$$

and so

$$\|\ell_w\| \le \|L\|\,\|w\|.$$

By the Riesz representation theorem, there is a uniquely determined element, denoted by $L^*(w) \in V$ such that

$$\ell_w(v) = (v, L^*(w)) \quad \forall\, v \in V.$$

We write

$$(Lv, w)_W = (v, L^*(w))_V \quad \forall\, v \in V,\, w \in W.$$

We first show that L^* is linear. Let $w_1, w_2 \in W$, and consider the linear functionals

$$\ell_1(v) = (Lv, w_1)_W = (v, L^*(w_1))_V,$$
$$\ell_2(v) = (Lv, w_2)_W = (v, L^*(w_2))_V$$

for any $v \in V$. Add these relations,

$$(Lv, w_1 + w_2)_W = (v, L^*(w_1) + L^*(w_2))_V \quad \forall\, v \in V.$$

By definition,

$$(Lv, w_1 + w_2)_W = (v, L^*(w_1 + w_2))_V;$$

so

$$(v, L^*(w_1 + w_2))_V = (v, L^*(w_1) + L^*(w_2))_V \quad \forall\, v \in V.$$

This implies

$$L^*(w_1 + w_2) = L^*(w_1) + L^*(w_2).$$

By a similar argument, for any $\alpha \in \mathbb{R}$, any $w \in W$,

$$L^*(\alpha\, w) = \alpha\, L^*(w).$$

Hence L^* is linear and we write $L^*(w) = L^*w$, and the defining relation is

$$(Lv, w)_W = (v, L^*w)_V \quad \forall\, v \in V,\, w \in W. \tag{2.6.1}$$

Then we show the boundedness of L^*. We have

$$\|L^*w\| = \|\ell_w\| \le \|L\|\, \|w\| \quad \forall\, w \in W.$$

Thus

$$\|L^*\| \le \|L\| \tag{2.6.2}$$

and L^* is bounded. Let us show that actually the inequality in (2.6.2) can be replaced by an equality. For this, we consider the adjoint of L^*, defined by the relation

$$(L^*w, v)_V = (w, (L^*)^*v)_W \quad \forall\, v \in V,\, w \in W.$$

Thus

$$(w, (L^*)^*v)_W = (w, Lv)_W \quad \forall\, v \in V,\, w \in W.$$

By writing this as $(w, (L^*)^*v - Lv)_W = 0$ and letting $w = (L^*)^*v - Lv$, we obtain

$$(L^*)^*v = Lv \quad \forall\, v \in V.$$

Hence

$$(L^*)^* = L. \tag{2.6.3}$$

We then apply (2.6.1) to L^* to obtain

$$\|L\| = \|(L^*)^*\| \le \|L^*\|.$$

Combining this with (2.6.2), we have

$$\|L^*\| = \|L\|. \tag{2.6.4}$$

From the above derivation, we see that for a continuous linear operator between Hilbert spaces, the adjoint of its adjoint is the operator itself.

In the special situation $V = W$ and $L = L^*$, we say L is a *self-adjoint* operator. When L is a self-adjoint operator from $\mathbb{R}^n$ to $\mathbb{R}^n$, it is represented by a symmetric matrix in $\mathbb{R}^{n\times n}$. Equations of the form $Lv = w$ with L self-adjoint occur in many important physical settings, and the study of them forms a large and important area within functional analysis.

Example 2.6.1 Let $V = W = L^2(a, b)$ with the real numbers as scalars and the standard norm $\|\cdot\|_2$. Consider the linear integral operator

$$Kv(x) = \int_a^b k(x, y)\, v(y)\, dy, \quad a \le x \le b,$$

where the kernel function satisfies the condition

$$B \equiv \left[\int_a^b \int_a^b |k(x, y)|^2 dx\, dy\right]^{1/2} < \infty.$$

For any $v \in L^2(a, b)$,

$$\begin{aligned} \|Kv\|_2^2 &= \int_a^b \left|\int_a^b k(x, y)\, v(y)\, dy\right|^2 dx \\ &\le \int_a^b \left[\int_a^b |k(x, y)|^2 dy\right] \left[\int_a^b |v(y)|^2 dy\right] dx \\ &= B^2 \|v\|_2^2. \end{aligned}$$

Thus,

$$\|Kv\|_2 \le B\, \|v\|_2 \quad \forall\, v \in L^2(a, b),$$

and then

$$\|K\| \le B.$$

Hence we see that K is a continuous linear operator on $L^2(a, b)$.

Now let us find the adjoint of K. By the defining relation (2.6.1),

$$\begin{aligned} (v, K^* w) &= (Kv, w) \\ &= \int_a^b w(x) \left[\int_a^b k(x, y)\, v(y)\, dy\right] dx \\ &= \int_a^b \left[\int_a^b k(x, y)\, w(x)\, dx\right] v(y)\, dy \end{aligned}$$

for any $v, w \in L^2(a, b)$. This implies

$$K^* v(y) = \int_a^b k(x, y)\, v(x)\, dx \quad \forall\, v \in L^2(a, b).$$

The integral operator K is self-adjoint if and only if $k(x, y) = k(y, x)$. □

Given a Hilbert space V, the set of self-adjoint operators on V form a subspace of $\mathcal{L}(V)$. Indeed the following result is easy to verify.

Proposition 2.6.2 *If $L_1, L_2 \in \mathcal{L}(V)$ are self-adjoint, then for any real scalars α_1 and α_2, the operator $\alpha_1 L_1 + \alpha_2 L_2$ is self-adjoint.*

Proof. From Exercise 2.6.1, we have

$$(\alpha_1 L_1 + \alpha_2 L_2)^* = \alpha_1 L_1^* + \alpha_2 L_2^*.$$

Since L_1 and L_2 are self-adjoint,

$$(\alpha_1 L_1 + \alpha_2 L_2)^* = \alpha_1 L_1 + \alpha_2 L_2.$$

Hence $\alpha_1 L_1 + \alpha_2 L_2$ is self-adjoint. □

Proposition 2.6.3 *Assume $L_1, L_2 \in \mathcal{L}(V)$ are self-adjoint. Then $L_1 L_2$ is self-adjoint if and only if $L_1 L_2 = L_2 L_1$.*

Proof. Since L_1 and L_2 are self-adjoint, we have

$$(L_1 L_2 u, v) = (L_2 u, L_1 v) = (u, L_2 L_1 v) \quad \forall\, u, v \in V.$$

Thus

$$(L_1 L_2)^* = L_2 L_1.$$

It follows that $L_1 L_2$ is self-adjoint if and only if $L_1 L_2 = L_2 L_1$ is valid. □

Corollary 2.6.4 *Suppose $L \in \mathcal{L}(V)$ is self-adjoint. Then for any non-negative integer n, L^n is self-adjoint (by convention, $L^0 = I$, the identity operator). Consequently, for any polynomial $p(x)$ with real coefficients, the operator $p(L)$ is self-adjoint.*

We have a useful characterization of the norm of a self-adjoint operator.

Theorem 2.6.5 *Let $L \in \mathcal{L}(V)$ be self-adjoint. Then*

$$\|L\| = \sup_{\|v\|=1} |(Lv, v)|. \tag{2.6.5}$$

Proof. Denote $M = \sup_{\|v\|=1} |(Lv, v)|$. First for any $v \in V$, $\|v\| = 1$, we have

$$|(Lv, v)| \le \|Lv\|\, \|v\| \le \|L\|.$$

So

$$M \le \|L\|. \tag{2.6.6}$$

Now for any $u, v \in V$, we have the identity

$$(Lu, v) = \frac{1}{4}\left[(L(u+v), u+v) - (L(u-v), u-v)\right].$$

Thus

$$|(Lu, v)| \le \frac{M}{4}\left(\|u+v\|^2 + \|u-v\|^2\right) = \frac{M}{2}\,(\|u\|^2 + \|v\|^2).$$

For $u \in V$ with $Lu \ne 0$, we take $v = (\|u\|/\|Lu\|)\,Lu$ in the above inequality to obtain

$$\|u\|\,\|Lu\| \le M\,\|u\|^2,$$

i.e.,

$$\|Lu\| \le M\,\|u\|.$$

Obviously, this inequality also holds if $Lu = 0$. Hence,

$$\|Lu\| \le M\,\|u\| \quad \forall\, u \in V,$$

and we see that $\|L\| \le M$. This inequality and (2.6.6) imply (2.6.5). □

Exercise 2.6.1 Prove the following properties for adjoint operators.

$$\begin{aligned}(\alpha_1 L_1 + \alpha_2 L_2)^* &= \alpha_1 L_1^* + \alpha_2 L_2^*, \quad \alpha_1, \alpha_2 \text{ real},\\ (L_1 L_2)^* &= L_2^* L_1^*,\\ (L^*)^* &= L.\end{aligned}$$

Exercise 2.6.2 Define $K : L^2(0,1) \to L^2(0,1)$ by

$$Kf(x) = \int_0^x k(x,y) f(y)\,dy, \quad 0 \le x \le 1,\ f \in L^2(0,1),$$

with $k(x,y)$ continuous for $0 \le y \le x \le 1$. Show K is a bounded operator. What is K^*? To what extent can the assumption of continuity of $k(x,y)$ be made less restrictive?

Exercise 2.6.3 The right-hand side of (2.6.5) defines a quantity $|||L|||$ for a general linear continuous operator $L \in \mathcal{L}(V)$. Prove the inequality

$$|(Lu, v)| \le |||L|||\,\|u\|\,\|v\|.$$

Hint: First consider the case $\|u\| = \|v\| = 1$.

2.7 Weak convergence and weak compactness

We recall from Definition 1.2.8 that in a normed space V, a sequence $\{u_n\}_n$ is said to converge to an element $u \in V$ if

$$\lim_{n\to\infty} \|u_n - u\| = 0.$$

Such convergence is also called strong convergence or convergence in norm. We write $u_n \to u$ in V as $n \to \infty$ to express this convergence property.

In this section, we introduce convergence in a weak sense.

Definition 2.7.1 *Let V be a normed space, V' its dual space. A sequence* $\{u_n\} \subset V$ converges weakly *to $u \in V$ if*

$$\ell(u_n) \to \ell(u) \quad \text{as } n \to \infty, \ \forall\, \ell \in V'.$$

In this case, we write $u_n \rightharpoonup u$ as $n \to \infty$.

From the definition, it is easy to see that strong convergence implies weak convergence.

Proposition 2.7.2 *Let V be a normed space, and assume $\{u_n\}$ converges weakly to u in V. Then*

$$\sup_n \|u_n\| < \infty.$$

Proof. From the sequence $\{u_n\} \subset V$, we define a sequence $\{u_n^*\} \subset (V')'$ by

$$u_n^*(\ell) = \ell(u_n) \quad \forall\, \ell \in V'.$$

Then $\{u_n^*\}$ is a sequence of bounded linear operators defined on the Banach space V', and for every $\ell \in V'$, the sequence $\{u_n^*(\ell)\}$ is bounded since $u_n^*(\ell) = \ell(u_n)$ converges. By the principle of uniform boundedness (Theorem 2.4.4), we know that

$$\sup_n \|u_n^*\| < \infty.$$

Apply Corollary 2.5.7, for any n,

$$\begin{aligned} \|u_n\| &= \sup\{|\ell(u_n)| \mid \ell \in V',\ \|\ell\| = 1\} \\ &= \sup\{|u_n^*(\ell)| \mid \ell \in V',\ \|\ell\| = 1\} \\ &\le \|u_n^*\|. \end{aligned}$$

Hence, $\sup_n \|u_n\| \le \sup_n \|u_n^*\| < \infty$. □

Thus, a weakly convergent sequence must be bounded.

Example 2.7.3 Let $f \in L^2(0, 2\pi)$. Then we know from Parseval's equality (1.3.10) that the Fourier series of f converges in $L^2(0, 2\pi)$. Therefore the Fourier coefficients converge to zero, and in particular,

$$\int_0^{2\pi} f(x)\, \sin(nx)\, dx \to 0 \quad \forall\, f \in L^2(0, 2\,\pi).$$

This result is known as the *Riemann-Lebesgue Lemma.* Thus the sequence $\{\sin(nx)\}_{n\ge 1}$ converges weakly to 0 in $L^2(0, 2\,\pi)$. But certainly the sequence $\{\sin(nx)\}_{n\ge 1}$ does not converge strongly to 0 in $L^2(0, 2\,\pi)$. □

Strong convergence implies weak convergence, but not vice versa as Example 2.7.3 shows, unless the space V is finite-dimensional. In a finite-dimensional space, it is well-known that a bounded sequence has a convergent subsequence (see Theorem 1.6.2). In an infinite-dimensional space, we expect only a weaker property; but even the weaker property is still useful in proving many existence results.

In the proof of Proposition 2.7.2, we have used the bidual $V'' = (V')'$. The bidual V'' of a normed space V is the dual of its dual V' with the corresponding dual norm. The mapping $J : V \to V''$ through the relation

$$\langle Jv, \ell\rangle_{V''\times V'} = \langle \ell, v\rangle_{V'\times V} \quad \forall\, v \in V,\ \forall\, \ell \in V'$$

defines an isometric isomorphism between V and $J(V) \subset V''$. The isometry refers to the equality

$$\|Jv\|_{V''} = \|v\|_V \quad \forall\, v \in V.$$

This equality is proved as follows: Easily, $\|Jv\|_{V''} \le \|v\|_V$, and by an application of Corollary 2.5.6, $\|Jv\|_{V''} \ge \|v\|_V$. We identify $J(V)$ with V.

Definition 2.7.4 *A normed space V is said to be* reflexive *if $(V')' = V$.*

An immediate consequence of this definition is that a reflexive normed space must be complete (i.e. a Banach space). By the Riesz representation theorem, it is relatively straightforward to show that any Hilbert space is reflexive.

The most important property of a reflexive Banach space is its weak compactness, given in the next theorem. It is fundamental in the development of an existence theory for abstract optimization problems (see Section 3.3). A proof is given in [58, p. 132] and [250, p. 64].

Theorem 2.7.5 *A Banach space V is reflexive if and only if any bounded sequence in V has a subsequence weakly converging to an element in V.*

Let $\Omega \subset \mathbb{R}^d$ be open and bounded. Recall from Example 2.5.1 that for $p \in (1,\infty)$, the dual space of $L^p(\Omega)$ is $L^{p'}(\Omega)$, where p' is the conjugate of p defined by the relation $1/p' + 1/p = 1$. Therefore,

$$(L^p(\Omega)')' = \left(L^{p'}(\Omega)\right)' = L^p(\Omega),$$

i.e., if $p \in (1,\infty)$, then the space $L^p(\Omega)$ is reflexive. Consequently, the above theorem implies the following: If $\{u_n\}$ is a bounded sequence in $L^p(\Omega)$, i.e. $\sup_n \|u_n\|_{L^p(\Omega)} < \infty$, then we can find a subsequence $\{u_{n'}\} \subset \{u_n\}$ and a function $u \in L^p(\Omega)$ such that

$$\lim_{n'\to\infty} \int_\Omega u_{n'}(\boldsymbol{x})\, v(\boldsymbol{x})\, dx = \int_\Omega u(\boldsymbol{x})\, v(\boldsymbol{x})\, dx \quad \forall\, v \in L^{p'}(\Omega).$$

The space $L^1(\Omega)$ is not reflexive since $(L^1(\Omega))' = L^\infty(\Omega)$, but $(L^\infty(\Omega))'$ strictly contains $L^1(\Omega)$ (Example 2.5.1). To state a result on weak compactness of the space $L^1(\Omega)$, we need the concept of uniform integrability. We say a set $U \subset L^1(\Omega)$ is *uniformly integrable* if U is bounded and for any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) > 0$ such that for any measurable subset $\Omega_1 \subset \Omega$ with measure $|\Omega_1| < \delta$, we have

$$\sup_{v\in U} \int_{\Omega_1} |v(\boldsymbol{x})|\, dx < \varepsilon.$$

Then from Dunford-Pettis theorem ([26, Subsection 2.4.5]), U is uniformly integrable if and only if each sequence $\{u_n\}_n \subset U$ has a subsequence $\{u_{n'}\}_{n'}$ converging weakly to some element $u \in L^1(\Omega)$:

$$\lim_{n'\to\infty} \int_\Omega u_{n'}(\boldsymbol{x})\, v(\boldsymbol{x})\, dx = \int_\Omega u(\boldsymbol{x})\, v(\boldsymbol{x})\, dx \quad \forall\, v \in L^\infty(\Omega).$$

A necessary and sufficient condition for $U \subset L^1(\Omega)$ to be uniformly integrable is that there exists a function $f : [0,\infty) \to [0,\infty)$ such that

$$\frac{f(s)}{s} \to \infty \text{ as } s \to \infty, \quad \text{and} \quad \sup_{v\in U} \inf_{\Omega} f(|v(\boldsymbol{x})|)\, dx < \infty.$$

Finally, we introduce the concepts of strong convergence and *weak-** convergence of a sequence of linear operators.

Definition 2.7.6 *Let V and W be normed spaces. A sequence of linear operators $\{L_n\}$ from V to W is said to* converge strongly *to a linear operator $L : V \to W$ if*

$$\lim_{n\to\infty} \|L - L_n\| = 0.$$

In this case, we write $L_n \to L$ as $n \to \infty$. We say $\{L_n\}$ converges weak-* *to L and write $L_n \rightharpoonup^* L$ if*

$$\lim_{n\to\infty} L_n v = Lv \quad \forall\, v \in V.$$

We also say $\{L_n\}$ converges pointwise to L on V.

For a reflexive Banach space V, a sequence $\{L_n\} \subset V'$ converges weak-* to L if and only if the sequence converges weakly to L in V'.

We end this section with the following equivalent definitions of weak convergence in $L^p(\Omega)$ for $1 \le p < \infty$ and weak-* convergence in $L^\infty(\Omega)$: a sequence $\{u_n\}_n \subset L^p(\Omega)$ weakly converges to $u \in L^p(\Omega)$ if

$$\lim_{n\to\infty} \int_\Omega u_n(\boldsymbol{x})\, v(\boldsymbol{x})\, dx = \int_\Omega u(\boldsymbol{x})\, v(\boldsymbol{x})\, dx \quad \forall\, v \in L^{p'}(\Omega);$$

a sequence $\{u_n\}_n \subset L^\infty(\Omega)$ weakly-* converges to $u \in L^\infty(\Omega)$ if

$$\lim_{n\to\infty} \int_\Omega u_n(\boldsymbol{x})\, v(\boldsymbol{x})\, dx = \int_\Omega u(\boldsymbol{x})\, v(\boldsymbol{x})\, dx \quad \forall\, v \in L^1(\Omega).$$

Exercise 2.7.1 Consider the linear operators from $C[a,b]$ to $\mathbb{R}$ defined by

$$Lv = \int_a^b v(x)\, dx$$

and

$$L_n v = \frac{b-a}{n} \sum_{i=1}^n v\left(a + \frac{b-a}{n}\, i\right), \quad n = 1, 2, \ldots.$$

We recognize that $\{L_n v\}$ is a sequence of Riemann sums for the integral Lv. Show that $L_n \rightharpoonup^* L$ but $L_n \not\to L$.

Exercise 2.7.2 If $u_n \rightharpoonup u$, then

$$\|u\| \le \liminf_{n\to\infty} \|u_n\|.$$

Show by an example that in general,

$$\|u\| \ne \liminf_{n\to\infty} \|u_n\|.$$

Exercise 2.7.3 Show that in an inner product space,

$$u_n \to u \quad \Longleftrightarrow \quad u_n \rightharpoonup u \text{ and } \|u_n\| \to \|u\|.$$

Exercise 2.7.4 The equivalence stated in Exercise 2.7.3 can be extended to a class of Banach spaces known as uniformly convex Banach spaces. A Banach space V is said to be uniformly convex if for any sequence $\{u_n\} \subset V$ and $\{v_n\} \subset V$ such that $\|u_n\| = \|v_n\| = 1$ for any n, the following implication holds as $n \to \infty$:

$$\|u_n + v_n\| \to 2 \quad \Longrightarrow \quad \|u_n - v_n\| \to 0.$$

(a) Show that a Hilbert space is uniformly convex.
(b) Apply the Clarkson inequalities (1.5.5) and (1.5.6) to show that $L^p(\Omega)$, $1 < p < \infty$, is uniformly convex.
(c) Show that in a uniformly convex Banach space,

$$u_n \to u \quad \Longleftrightarrow \quad u_n \rightharpoonup u \text{ and } \|u_n\| \to \|u\|.$$

Hint: For the "$\Leftarrow$" part, consider $u_n/\|u_n\|$ and $u/\|u\|$, and note that $u_n/\|u_n\| \rightharpoonup u/\|u\|$.

Exercise 2.7.5 Recall the following Riemann-Lebesgue Theorem ([61, Theorem 2.1.5]): Let $1 \le p \le \infty$ and $\Omega = (a_1, b_1) \times \cdots \times (a_d, b_d)$ for real numbers $a_i < b_i$, $1 \le i \le d$. For any $u \in L^p(\Omega)$, extend it by periodicity from Ω to $\mathbb{R}^d$ and define a sequence of functions $\{u_n\}_{n\ge1}$ by $u_n(\boldsymbol{x}) = u(n\boldsymbol{x})$. Denote

$\overline{u} = \int_\Omega u(\boldsymbol{x})\,dx/\text{meas}(\Omega)$ for the mean value of u over Ω. Then $u_n \rightharpoonup \overline{u}$ in $L^p(\Omega)$ if $1 \le p < \infty$, and $u_n \rightharpoonup^* \overline{u}$ in $L^\infty(\Omega)$ if $p = \infty$.

Apply this result to determine the weak limits of the following sequences: $\{\sin nx\}_{n\ge1} \subset L^p(0,\pi)$; $\{\sin nx\}_{n\ge1} \subset L^p(0,\pi/2)$; $\{\sin^2 nx\}_{n\ge1} \subset L^p(0,2\pi)$; $\{u_n(x)\}_{n\ge1} \subset L^p(0,1)$ with $u_n(x) = u(nx)$ and u is a piecewise constant function: $u(x) = c_i$ for $x \in ((i-1)/m, i/m)$, $1 \le i \le m$.

2.8 Compact linear operators

When V is a finite dimensional linear space and $A : V \to V$ is linear, the equation $Au = w$ has a well-developed solvability theory. To extend these results to infinite dimensional spaces, we introduce the concept of a *compact operator* K and then we present a theory for operator equations $Au = w$ in which $A = I - K$. Equations of the form

$$u - Ku = f \tag{2.8.1}$$

are called "equations of the second kind", and generally K is assumed to have special properties. The main idea is that compact operators are in some sense closely related to finite-dimensional operators, i.e. operators with a finite-dimensional range. If K is truly finite-dimensional, in a sense we define below, then (2.8.1) can be reformulated as a finite system of linear equations and solved exactly. If K is compact, then it is close to being finite-dimensional; and the solvability theory of (2.8.1) is similar to that for the finite-dimensional case.

In the following, recall the discussion in Section 1.6.

Definition 2.8.1 *Let V and W be normed spaces, and let $K : V \to W$ be linear. Then K is* compact *if the set $\{Kv \mid \|v\|_V \le 1\}$ has compact closure in W. This is equivalent to saying that for every bounded sequence $\{v_n\} \subset V$, the sequence $\{Kv_n\}$ has a subsequence that is convergent to some point in W. Compact operators are also called* completely continuous operators.

There are other definitions for a compact operator, but the above is the one used most commonly. In the definition, the spaces V and W need not be complete; but in virtually all applications, they are complete. With completeness, some of the proofs of the properties of compact operators become simpler, and we will always assume V and W are complete (i.e. Banach spaces) when dealing with compact operators.

2.8.1 Compact integral operators on $C(D)$

Let D be a closed bounded set in $\mathbb{R}^d$. The space $C(D)$ is to have the norm $\|\cdot\|_\infty$. Given a function $k : D \times D \to \mathbb{R}$, we define

$$Kv(\boldsymbol{x}) = \int_D k(\boldsymbol{x}, \boldsymbol{y})v(\boldsymbol{y})\,dy, \quad \boldsymbol{x} \in D,\ v \in C(D). \tag{2.8.2}$$

We want to formulate conditions under which $K : C(D) \to C(D)$ is both bounded and compact. We assume $k(\boldsymbol{x}, \boldsymbol{y})$ is integrable as a function of $\boldsymbol{y}$, for all $\boldsymbol{x} \in D$, and further we assume the following.
(A_1) $\lim_{h\to 0} \omega(h) = 0$, with

$$\omega(h) \equiv \sup_{\substack{\boldsymbol{x},\boldsymbol{z}\in D\\ \|\boldsymbol{x}-\boldsymbol{z}\|\le h}} \int_D |k(\boldsymbol{x}, \boldsymbol{y}) - k(\boldsymbol{z}, \boldsymbol{y})|\,dy. \tag{2.8.3}$$

Here, $\|\boldsymbol{x} - \boldsymbol{z}\|$ denotes the Euclidean length of $\boldsymbol{x} - \boldsymbol{z}$.

(A_2)

$$\sup_{\boldsymbol{x}\in D} \int_D |k(\boldsymbol{x}, \boldsymbol{y})|\,dy < \infty. \tag{2.8.4}$$

By (A_1), if $v(\boldsymbol{y})$ is bounded and integrable, then $Kv(\boldsymbol{x})$ is continuous and

$$|Kv(\boldsymbol{x}) - Kv(\boldsymbol{y})| \le \omega(\|\boldsymbol{x} - \boldsymbol{y}\|)\,\|v\|_\infty. \tag{2.8.5}$$

Using (A_2), we have boundedness of K, with its norm

$$\|K\| = \max_{\boldsymbol{x}\in D} \int_D |k(\boldsymbol{x}, \boldsymbol{y})|\,dy. \tag{2.8.6}$$

To discuss compactness of K, we first need to identify the compact sets in $C(D)$. To do this, we apply Arzela-Ascoli theorem (Theorem 1.6.3). Consider the set $S = \{Kv \mid v \in C(D),\ \|v\|_\infty \le 1\}$. This set is uniformly bounded, since $\|Kv\|_\infty \le \|K\|\,\|v\|_\infty \le \|K\|$ for any $v \in S$. In addition, S is equicontinuous from (2.8.5). Thus S has compact closure in $C(D)$, and K is a compact operator on $C(D)$ to $C(D)$.

What kernel functions k satisfy (A_1) and (A_2)? Easily, these assumptions are satisfied if $k(\boldsymbol{x}, \boldsymbol{y})$ is a continuous function of $(\boldsymbol{x}, \boldsymbol{y}) \in D$. In addition, let $D = [a, b]$ and consider

$$Kv(x) = \int_a^b \log|x - y|\,v(y)\,dy \tag{2.8.7}$$

and

$$Kv(x) = \int_a^b \frac{1}{|x - y|^\beta} v(y)\,dy \tag{2.8.8}$$

with $\beta < 1$. These operators K can be shown to satisfy (A_1)–(A_2), although we omit the proof. Later we show by other means that these are compact operators. An important and related example is

$$Kv(\boldsymbol{x}) = \int_D \frac{1}{\|\boldsymbol{x} - \boldsymbol{y}\|^\beta} v(\boldsymbol{y})\, dy, \quad \boldsymbol{x} \in D,\ v \in C(D).$$

The set $D \subset \mathbb{R}^d$ is assumed to be closed, bounded, and have a non-empty interior. This operator satisfies (A_1)–(A_2) provided $\beta < d$, and therefore K is a compact operator from $C(D) \to C(D)$.

Still for the case $D = [a, b]$, another way to show that $k(x, y)$ satisfies (A_1) and (A_2) is to rewrite k in the form

$$k(x, y) = \sum_{i=0}^{n} h_i(x, y) l_i(x, y) \tag{2.8.9}$$

for some $n > 0$, with each $l_i(x, y)$ continuous for $a \le x, y \le b$ and each $h_i(x, y)$ satisfying (A_1)–(A_2). It is left to the reader to show that in this case, k also satisfies (A_1)–(A_2). The utility of this approach is that it is sometimes difficult to show directly that k satisfies (A_1)–(A_2), whereas showing (2.8.9) with h_i, l_i satisfying the specified conditions may be easier.

Example 2.8.2 Let $[a, b] = [0, \pi]$ and $k(x, y) = \log|\cos x - \cos y|$. Rewrite the kernel function as

$$k(x, y) = \underbrace{|x - y|^{-1/2}}_{h(x,y)} \underbrace{|x - y|^{1/2} \log|\cos x - \cos y|}_{l(x,y)}. \tag{2.8.10}$$

Easily, l is continuous. From the discussion following (2.8.8), h satisfies (A_1)–(A_2). Thus k is the kernel of a compact integral operator on $C[0, \pi]$ to $C[0, \pi]$. □

2.8.2 Properties of compact operators

Another way of obtaining compact operators is to look at limits of simpler "finite-dimensional operators" in $\mathcal{L}(V, W)$, the Banach space of bounded linear operators from V to W. This gives another perspective on compact operators, one that leads to improved intuition by emphasizing their close relationship to operators on finite dimensional spaces.

Definition 2.8.3 *Let V and W be linear spaces. The linear operator $K : V \to W$ is of* finite rank *if $\mathcal{R}(K)$, the range of K, is finite dimensional.*

Proposition 2.8.4 *Let V and W be normed spaces, and let $K : V \to W$ be a bounded finite rank operator. Then K is a compact operator.*

Proof. The range $\mathcal{R}(K)$ is a finite-dimensional normed space, and therefore it is complete. Consider the set

$$S = \{Kv \mid \|v\|_V \le 1\}.$$

It is bounded, each of its elements being bounded by $\|K\|$. Notice that $S \subset \mathcal{R}(K)$. Then S has a compact closure, since all bounded closed sets in a finite dimensional space are compact. This shows K is compact. □

Example 2.8.5 Let $V = W = C[a, b]$ with the norm $\|\cdot\|_\infty$. Consider the kernel function

$$k(x, y) = \sum_{i=1}^{n} \beta_i(x)\, \gamma_i(y) \tag{2.8.11}$$

with each β_i continuous on $[a, b]$ and each γ_i absolutely integrable on $[a, b]$. Then the associated integral operator K is a bounded, finite rank operator on $C[a, b]$ to $C[a, b]$:

$$Kv(x) = \sum_{i=1}^{n} \beta_i(x) \int_a^b \gamma_i(y)\, v(y)\, dy, \quad v \in C[a, b]. \tag{2.8.12}$$

Indeed, we have

$$\|K\| \le \sum_{i=1}^{n} \|\beta_i\|_\infty \int_a^b |\gamma_i(y)|\, dy.$$

From (2.8.12), $Kv \in C[a, b]$ and $\mathcal{R}(K) \subset \operatorname{span}\{\beta_1, \dots, \beta_n\}$, a finite dimensional space. □

Kernel functions of the form (2.8.11) are called *degenerate*. Below we see that the associated integral equation $(\lambda I - K)v = f$, $\lambda \ne 0$, is essentially a finite dimensional equation.

Proposition 2.8.6 *Let $K \in \mathcal{L}(U, V)$ and $L \in \mathcal{L}(V, W)$ with at least one of them being compact. Then LK is a compact operator from U to W.*

The proof is left as Exercise 2.8.1 for the reader.

The following result gives the framework for using finite rank operators to obtain similar, but more general compact operators.

Proposition 2.8.7 *Let V and W be normed spaces, with W complete. Assume $\{K_n\} \subset \mathcal{L}(V, W)$ is a sequence of compact operators such that $K_n \to K$ in $\mathcal{L}(V, W)$. Then K is compact.*

This is a standard result found in most books on functional analysis; e.g. see [58, p. 174] or [71, p. 486].

For almost all function spaces V that occur in applied mathematics, the compact operators can be characterized as being the limit of a sequence of bounded finite-rank operators. This gives a further justification for the presentation of Proposition 2.8.7.

Example 2.8.8 Let D be a closed and bounded set in $\mathbb{R}^d$. For example, D could be a region with nonempty interior, a piecewise smooth surface, or a piecewise smooth curve. Let $k(\boldsymbol{x}, \boldsymbol{y})$ be a continuous function of $\boldsymbol{x}, \boldsymbol{y} \in D$. Suppose we can define a sequence of continuous degenerate kernel functions $k_n(\boldsymbol{x}, \boldsymbol{y})$ for which

$$\max_{\boldsymbol{x}\in D} \int_D |k(\boldsymbol{x}, \boldsymbol{y}) - k_n(\boldsymbol{x}, \boldsymbol{y})|\, dy \to 0 \quad \text{as } n \to \infty. \tag{2.8.13}$$

Then for the associated integral operators, it easily follows that $K_n \to K$; and by Proposition 2.8.7, K is compact. The condition (2.8.13) is true for general continuous functions $k(x, y)$, and we leave to the exercises the proof for various choices of D. Of course, we already knew that K was compact in this case, from the discussion following (2.8.8). But the present approach shows the close relationship of compact operators and finite dimensional operators. $\square$

Example 2.8.9 Let $V = W = C[a, b]$ with the norm $\|\cdot\|_\infty$. Consider the kernel function

$$k(x, y) = \frac{1}{|x-y|^\gamma} \tag{2.8.14}$$

for some $0 < \gamma < 1$. Define a sequence of continuous kernel functions to approximate it:

$$k_n(x, y) = \begin{cases} \dfrac{1}{|x-y|^\gamma}, & |x-y| \ge \dfrac{1}{n}, \\ n^\gamma, & |x-y| \le \dfrac{1}{n}. \end{cases} \tag{2.8.15}$$

This merely limits the height of the graph of $k_n(x, y)$ to that of $k(x, y)$ when $|x - y| = 1/n$. Easily, $k_n(x, y)$ is a continuous function for $a \le x, y \le b$, and thus the associated integral operator K_n is compact on $C[a, b]$. For the associated integral operators,

$$\|K - K_n\| = \frac{2\gamma}{1-\gamma} \cdot \frac{1}{n^{1-\gamma}}$$

which converges to zero as $n \to \infty$. By Proposition 2.8.7, K is a compact operator on $C[a, b]$. $\square$

2.8.3 Integral operators on $L^2(a, b)$

Let $V = W = L^2(a, b)$, and let K be the integral operator associated with a kernel function $k(x, y)$. We first show that under suitable assumptions on

k, the operator K maps $L^2(a,b)$ to $L^2(a,b)$ and is bounded. Assume

$$M \equiv \left[\int_a^b \int_a^b |k(x,y)|^2 dy\, dx\right]^{1/2} < \infty. \tag{2.8.16}$$

For $v \in L^2(a,b)$, we use the Cauchy-Schwarz inequality to obtain

$$\begin{aligned} \|Kv\|_2^2 &= \int_a^b \left| \int_a^b k(x,y)v(y)dy \right|^2 dx \\ &\le \int_a^b \left[\int_a^b |K(x,y)|^2 dy\right] \left[\int_a^b |v(y)|^2 dy\right] dx \\ &= M^2 \|v\|_2^2. \end{aligned}$$

Thus, $Kv \in L^2(a,b)$ and

$$\|K\| \le M. \tag{2.8.17}$$

This bound is comparable to the use of the Frobenius matrix norm to bound the operator norm of a matrix A: $\mathbb{R}^n \to \mathbb{R}^m$, when the vector norm $\|\cdot\|_2$ is being used. Recall that the Frobenius norm of a matrix A is given by

$$\|A\|_F = \Big(\sum_{i,j} |A_{i,j}|^2\Big)^{1/2}.$$

Kernel functions k for which $M < \infty$ are called *Hilbert-Schmidt kernel functions*, and the quantity M in (2.8.16) is called the *Hilbert-Schmidt norm* of K.

For integral operators K with a degenerate kernel function as in (2.8.11), the operator K is bounded if all β_i, $\gamma_i \in L^2(a,b)$. This is a straightforward result which we leave as a problem for the reader. From Proposition 2.8.4, the integral operator is then compact.

To examine the compactness of K for more general kernel functions, we assume there is a sequence of kernel functions $k_n(x,y)$ for which (i) $K_n : L^2(a,b) \to L^2(a,b)$ is compact, and (ii)

$$M_n \equiv \left[\int_a^b \int_a^b |k(x,y) - k_n(x,y)|^2 dy\, dx\right]^{1/2} \to 0 \quad \text{as } n \to \infty. \tag{2.8.18}$$

For example, if K is continuous, then this follows from (2.8.13). The operator $K - K_n$ is an integral operator, and we apply (2.8.16)–(2.8.17) to it to obtain

$$\|K - K_n\| \le M_n \to 0 \quad \text{as } n \to \infty.$$

From Proposition 2.8.7, this shows K is compact. For any Hilbert-Schmidt kernel function, (2.8.18) can be shown to hold for a suitable choice of degenerate kernel functions k_n.

We leave it for the reader to show that $\log|x-y|$ and $|x-y|^{-\gamma}$, $\gamma < \frac{1}{2}$, are Hilbert-Schmidt kernel functions (Exercise 2.8.4). For $\frac{1}{2} \le \gamma < 1$, the kernel function $|x-y|^{-\gamma}$ still defines a compact integral operator K on $L^2(a,b)$, but the above theory for Hilbert-Schmidt kernel functions does not apply. For a proof of the compactness of K in this case, see Mikhlin [172, p. 160].

2.8.4 The Fredholm alternative theorem

Integral equations were studied in the 19th century as one means of investigating boundary value problems for the Laplace equation, for example,

$$\begin{aligned} \Delta u(\boldsymbol{x}) &= 0, \qquad && \boldsymbol{x} \in \Omega, \\ u(\boldsymbol{x}) &= f(\boldsymbol{x}), && \boldsymbol{x} \in \partial\Omega, \end{aligned} \tag{2.8.19}$$

and other elliptic partial differential equations. In the early 1900's, Ivar Fredholm found necessary and sufficient conditions for the solvability of a large class of Fredholm integral equations of the second kind. With these results, he then was able to give much more general existence theorems for the solution of boundary value problems such as (2.8.19). In this subsection, we state and prove the most important results of Fredholm; and in the following subsection, we give additional results without proof.

The theory of integral equations has been due to many people, with David Hilbert being among the most important popularizers of the area. The subject of integral equations continues as an important area of study in applied mathematics; and for an introduction that includes a review of much recent literature, see Kress [149]. For an interesting historical account of the development of functional analysis as it was affected by the development of the theory of integral equations, see Bernkopf [36]. From hereon, to simplify notation, for a scalar λ and an operator $K : V \to V$, we use $\lambda - K$ for the operator $\lambda I - K$, where $I : V \to V$ is the identity operator.

Theorem 2.8.10 (Fredholm alternative) *Let V be a Banach space, and let $K : V \to V$ be compact. Then the equation $(\lambda - K)u = f$, $\lambda \ne 0$, has a unique solution $u \in V$ for any $f \in V$ if and only if the homogeneous equation $(\lambda - K)v = 0$ has only the trivial solution $v = 0$. In such a case, the operator $\lambda - K : V \underset{onto}{\overset{1-1}{\to}} V$ has a bounded inverse $(\lambda - K)^{-1}$.*

Proof. The theorem is true for any compact operator K, but here we give a proof only for those compact operators which are the limit of a sequence of bounded finite-rank operators. For a more general proof, see Kress [149, Chap. 3] or Conway [58, p. 217]. We remark that the theorem is a generalization of the following standard result for finite dimensional vector spaces V. For A a matrix of order n, with $V = \mathbb{R}^n$ or $\mathbb{C}^n$ (with A having real entries for the former case), the linear system $Au = w$ has a unique solution $u \in V$ for any $w \in V$ if and only if the homogeneous linear system $Az = 0$ has only the zero solution $z = 0$.

(a) We begin with the case where K is finite-rank and bounded. Let $\{\varphi_1, \ldots, \varphi_n\}$ be a basis for $\mathcal{R}(K)$, the range of K. Rewrite the equation $(\lambda - K)u = f$ as

$$u = \frac{1}{\lambda}\,(f + Ku). \tag{2.8.20}$$

If this equation has a unique solution $u \in V$, then

$$u = \frac{1}{\lambda}\,(f + c_1\varphi_1 + \cdots + c_n\varphi_n) \tag{2.8.21}$$

for some uniquely determined set of constants $c_1, \ldots, c_n$.

By substituting (2.8.21) into the equation $(\lambda - K)\,u = f$, we have

$$\lambda\left(\frac{1}{\lambda}f + \frac{1}{\lambda}\sum_{i=1}^{n} c_i\varphi_i\right) - \frac{1}{\lambda}Kf - \frac{1}{\lambda}\sum_{j=1}^{n} c_j K\varphi_j = f.$$

Multiply by λ, and then simplify to obtain

$$\lambda\sum_{i=1}^{n} c_i\varphi_i - \sum_{j=1}^{n} c_j K\varphi_j = Kf. \tag{2.8.22}$$

Using the basis $\{\varphi_i\}$ for $\mathcal{R}(K)$, write

$$Kf = \sum_{i=1}^{n} \gamma_i\varphi_i$$

and

$$K\varphi_j = \sum_{i=1}^{n} a_{ij}\varphi_i, \quad 1 \le j \le n.$$

The coefficients $\{\gamma_i\}$ and $\{a_{ij}\}$ are uniquely determined. Substitute these expressions into (2.8.22) and rearrange the terms,

$$\sum_{i=1}^{n}\left(\lambda c_i - \sum_{j=1}^{n} a_{ij}c_j\right)\varphi_i = \sum_{i=1}^{n}\gamma_i\varphi_i.$$

By the independence of the basis elements φ_i, we obtain the linear system

$$\lambda c_i - \sum_{j=1}^{n} a_{ij}c_j = \gamma_i, \quad 1 \le i \le n. \tag{2.8.23}$$

Claim: This linear system and the equation $(\lambda - K)u = f$ are completely equivalent in their solvability, with (2.8.21) furnishing a one-to-one correspondence between the solutions of the two of them.

We have shown above that if u is a solution of $(\lambda - K)u = f$, then $(c_1, \ldots, c_n)^T$ is a solution of (2.8.23). In addition, suppose u_1 and u_2 are distinct solutions of $(\lambda - K)u = f$. Then

$$Ku_1 = \lambda u_1 - f \quad \text{and} \quad Ku_2 = \lambda u_2 - f, \quad \lambda \neq 0,$$

are also distinct vectors in $\mathcal{R}(K)$, and thus the associated vectors of coordinates $(c_1^{(1)}, \ldots, c_n^{(1)})^T$ and $(c_1^{(2)}, \ldots, c_n^{(2)})^T$,

$$K\varphi_i = \sum_{k=1}^{n} c_k^{(i)} \varphi_k, \quad i = 1, 2$$

must also be distinct.

For the converse statement, suppose $(c_1, \ldots, c_n)^T$ is a solution of (2.8.23). Define a vector $u \in V$ by using (2.8.21), and then check whether this u satisfies the integral equation:

$$\begin{aligned}
(\lambda - K)u &= \lambda \left(\frac{1}{\lambda} f + \frac{1}{\lambda} \sum_{i=1}^{n} c_i \varphi_i \right) - \frac{1}{\lambda} K f - \frac{1}{\lambda} \sum_{j=1}^{n} c_j K \varphi_j \\
&= f + \frac{1}{\lambda} \left(\lambda \sum_{i=1}^{n} c_i \varphi_i - K f - \sum_{j=1}^{n} c_j K \varphi_j \right) \\
&= f + \frac{1}{\lambda} \left(\sum_{i=1}^{n} \lambda c_i \varphi_i - \sum_{i=1}^{n} \gamma_i \varphi_i - \sum_{j=1}^{n} c_j \sum_{i=1}^{n} a_{ij} \varphi_i \right) \\
&= f + \frac{1}{\lambda} \sum_{i=1}^{n} \underbrace{\left(\lambda c_i - \gamma_i - \sum_{j=1}^{n} a_{ij} c_j \right)}_{=0,\ i=1,\ldots,n} \varphi_i \\
&= f.
\end{aligned}$$

Also, distinct coordinate vectors $(c_1, \ldots, c_n)^T$ lead to distinct solutions u given by (2.8.21), because of the linear independence of the basis vectors $\{\varphi_1, \ldots, \varphi_n\}$. This completes the proof of the claim given above.

Now consider the Fredholm alternative theorem for $(\lambda - K)u = f$ with this finite rank operator K. Suppose

$$\lambda - K : V \underset{onto}{\overset{1-1}{\to}} V.$$

Then trivially, the null space $\mathcal{N}(\lambda - K) = \{0\}$. For the converse, assume $(\lambda - K)\, v = 0$ has only the solution $v = 0$. Note that we want to show that $(\lambda - K)\, u = f$ has a unique solution for every $f \in V$.

Consider the associated linear system (2.8.23). It can be shown to have a unique solution for any right hand side $(\gamma_1, \ldots, \gamma_n)^T$ by showing that the homogeneous linear system has only the zero solution. The latter is done by means of the equivalence of the homogeneous linear system to the homogeneous equation $(\lambda - K)v = 0$, which implies $v = 0$. But since (2.8.23) has a unique solution, so must $(\lambda - K)u = f$, and it is given by (2.8.21).

We must also show that $(\lambda - K)^{-1}$ is bounded. This can be done directly by a further examination of the consequences of K being a bounded and finite rank operator; but it is simpler to just cite the open mapping theorem (see Theorem 2.4.3).

(b) Assume now that $\|K - K_n\| \to 0$, with K_n finite rank and bounded. Rewrite $(\lambda - K)\,u = f$ as

$$[\lambda - (K - K_n)]\,u = f + K_n u, \quad n \geq 1. \tag{2.8.24}$$

Pick an index $m > 0$ for which

$$\|K - K_m\| < |\lambda| \tag{2.8.25}$$

and fix it. By the geometric series theorem (Theorem 2.3.1),

$$Q_m \equiv [\lambda - (K - K_m)]^{-1}$$

exists and is bounded, with

$$\|Q_m\| \leq \frac{1}{|\lambda| - \|K - K_m\|}.$$

The equation (2.8.24) with $n = m$ can now be written in the equivalent form

$$u - Q_m K_m u = Q_m f. \tag{2.8.26}$$

The operator $Q_m K_m$ is bounded and finite rank. The boundedness follows from that of Q_m and K_m. To show it is finite rank, let $\mathcal{R}(K_m) = \text{span}\{\varphi_1, \dots, u_m\}$. Then

$$\mathcal{R}(Q_m K_m) = \text{span}\{Q_m\varphi_1, \dots, Q_m u_m\}$$

is a finite-dimensional space.

The equation (2.8.26) is one to which we can apply part (a) of this proof. Assume $(\lambda - K)v = 0$ implies $v = 0$. By the above equivalence, this yields

$$(I - Q_m K_m)\,v = 0 \quad \Longrightarrow \quad v = 0.$$

But from part (a), this says $(I - Q_m K_m)\,u = w$ has a unique solution u for every $w \in V$, and in particular, for $w = Q_m f$ as in (2.8.26). By the equivalence of (2.8.26) and $(\lambda - K)\,u = f$, we have that the latter is uniquely solvable for every $f \in V$. The boundedness of $(\lambda - K)^{-1}$ follows from part (a) and the boundedness of Q_m. Alternatively, the open mapping theorem can be cited, as earlier in part (a). □

For many practical problems in which K is not compact, it is important to note what makes this proof work. It is *not* necessary to have a sequence of bounded and finite rank operators $\{K_n\}$ for which $\|K - K_n\| \to 0$. Rather, it is necessary to satisfy the inequality (2.8.25) for one finite rank operator K_m; and in applying the proof to other operators K, it is necessary only that K_m be compact. In such a case, the proof following (2.8.25) remains valid, and the Fredholm Alternative still applies to such an equation

$$(\lambda - K)\, u = f.$$

2.8.5 Additional results on Fredholm integral equations

In this subsection, we give additional results on the solvability of compact equations of the second kind, $(\lambda - K)\, u = f$, with $\lambda \neq 0$. No proofs are given for these results, and the reader is referred to a standard text on integral equations; e.g. see Kress [149] or Mikhlin [171].

Definition 2.8.11 *Let $K : V \to V$. If there is a scalar λ and an associated vector $u \neq 0$ for which $Ku = \lambda\, u$, then λ is called an* eigenvalue *and u an associated* eigenvector *of the operator K.*

When dealing with compact operators K, we generally are interested in only the nonzero eigenvalues of K.

In the following, recall that $\mathcal{N}(A)$ denotes the null space of A.

Theorem 2.8.12 *Let $K : V \to V$ be compact, and let V be a Banach space. Then:*

(1) *The eigenvalues of K form a discrete set in the complex plane $\mathbb{C}$, with 0 as the only possible limit point.*

(2) *For each nonzero eigenvalue λ of K, there are only a finite number of linearly independent eigenvectors.*

(3) *Each nonzero eigenvalue λ of K has finite index $\nu(\lambda) \geq 1$. This means*

$$\mathcal{N}(\lambda - K) \underset{\neq}{\subset} \mathcal{N}((\lambda - K)^2) \underset{\neq}{\subset} \cdots \underset{\neq}{\subset} \mathcal{N}((\lambda - K)^{\nu(\lambda)}) = \mathcal{N}((\lambda - K)^{\nu(\lambda)+1}). \qquad (2.8.27)$$

In addition, $\mathcal{N}((\lambda - K)^{\nu(\lambda)})$ is finite dimensional. The elements of the subspace $\mathcal{N}((\lambda - K)^{\nu(\lambda)}) \backslash \mathcal{N}(\lambda - K)$ are called generalized eigenvectors of K.

(4) *For any* $\lambda \neq 0$, $\mathcal{R}(\lambda - K)$ *is closed in* V.

(5) *For each nonzero eigenvalue* λ *of* K,

$$V = \mathcal{N}((\lambda - K)^{\nu(\lambda)}) \oplus \mathcal{R}((\lambda - K)^{\nu(\lambda)}) \tag{2.8.28}$$

is a decomposition of V *into invariant subspaces. This implies that every* $u \in V$ *can be written as* $u = u_1 + u_2$ *with unique choices*

$$u_1 \in \mathcal{N}((\lambda - K)^{\nu(\lambda)}) \quad \text{and} \quad u_2 \in \mathcal{R}((\lambda - K)^{\nu(\lambda)}).$$

Being invariant means that

$$K : \mathcal{N}((\lambda - K)^{\nu(\lambda)}) \to \mathcal{N}((\lambda - K)^{\nu(\lambda)}),$$
$$K : \mathcal{R}((\lambda - K)^{\nu(\lambda)}) \to \mathcal{R}((\lambda - K)^{\nu(\lambda)}).$$

(6) *The Fredholm alternative theorem and the above results* (1)–(5) *remain true if* K^m *is compact for some integer* $m > 1$.

For results on the speed with which the eigenvalues $\{\lambda_n\}$ of compact integral operators K converge to zero, see Hille and Tamarkin [125] and Fenyö and Stolle [80, Section 8.9]. Generally, as the differentiability of the kernel function $k(x, y)$ increases, the speed of convergence to zero of the eigenvalues also increases.

For the following results, recall from Section 2.6 the concept of an adjoint operator.

Lemma 2.8.13 *Let* V *be a Hilbert space with scalars the complex numbers* $\mathbb{C}$, *and let* $K : V \to V$ *be a compact operator. Then* $K^* : V \to V$ *is also a compact operator.*

This implies that the operator K^* also shares the properties stated above for the compact operator K. There is, however, a closer relationship between the operators K and K^*, which is given in the following theorem.

Theorem 2.8.14 *Let* V *be a Hilbert space with scalars the complex numbers* $\mathbb{C}$, *let* $K : V \to V$ *be a compact operator, and let* λ *be a nonzero eigenvalue of* K. *Then:*

(1) $\bar{\lambda}$ *is an eigenvalue of the adjoint operator* K^*. *In addition,* $\mathcal{N}(\lambda - K)$ *and* $\mathcal{N}(\bar{\lambda} - K^*)$ *have the same dimension.*

(2) *The equation* $(\lambda - K)u = f$ *is solvable if and only if*

$$(f, v) = 0 \quad \forall\, v \in \mathcal{N}(\bar{\lambda} - K^*). \tag{2.8.29}$$

An equivalent way of writing this is

$$\mathcal{R}(\lambda - K) = \mathcal{N}(\bar{\lambda} - K^*)^{\perp},$$

the subspace orthogonal to $\mathcal{N}(\bar{\lambda} - K^*)$. *With this, we can write the decomposition*

$$V = \mathcal{N}(\bar{\lambda} - K^*) \oplus \mathcal{R}(\lambda - K). \tag{2.8.30}$$

Theorem 2.8.15 *Let V be a Hilbert space with scalars the complex numbers $\mathbb{C}$, and let $K : V \to V$ be a self-adjoint compact operator. Then all eigenvalues of K are real and of index $\nu(\lambda) = 1$. In addition, the corresponding eigenvectors can be chosen to form an orthonormal set. Order the nonzero eigenvalues as follows:*

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n| \geq \cdots > 0 \tag{2.8.31}$$

with each eigenvalue repeated according to its multiplicity (i.e. the dimension of $\mathcal{N}(\lambda - K)$). Then we write

$$Ku_i = \lambda_i u_i, \quad i \geq 1 \tag{2.8.32}$$

with

$$(u_i, u_j) = \delta_{ij}.$$

Also, the eigenvectors $\{u_i\}$ form an orthonormal basis for $\overline{\mathcal{R}(\lambda - K)}$.

Much of the theory of self-adjoint boundary value problems for ordinary and partial differential equations is based on Theorems 2.8.14 and 2.8.15. Moreover, the completeness in $L^2(D)$ of many families of functions is proven by showing they are the eigenfunctions to a self-adjoint differential equation or integral equation problem.

Example 2.8.16 Let $D = \{\boldsymbol{x} \in \mathbb{R}^3 \mid \|\boldsymbol{x}\| = 1\}$, the unit sphere in $\mathbb{R}^3$, and let $V = L^2(D)$. Here $\|\boldsymbol{x}\|$ denotes the Euclidean length of $\boldsymbol{x}$. Define

$$Kv(\boldsymbol{x}) = \int_D \frac{v(\boldsymbol{y})}{\|\boldsymbol{x} - \boldsymbol{y}\|} dS_{\boldsymbol{y}}, \quad \boldsymbol{x} \in D. \tag{2.8.33}$$

This is a compact operator, a proof of which is given in Mikhlin [172, p. 160]. The eigenfunctions of K are called *spherical harmonics*, a much-studied set of functions; e.g. see [86], [161]. For each integer $k \geq 0$, there are $2k + 1$ independent spherical harmonics of *degree* k; and for each such spherical harmonic φ_k, we have

$$K\varphi_k = \frac{4\pi}{2k+1}\varphi_k, \quad k = 0, 1, \ldots. \tag{2.8.34}$$

Letting $\mu_k = 4\pi/(2k+1)$, we have $\mathcal{N}(\mu_k - K)$ has dimension $2k+1$, $k \geq 0$. It is well-known that the set of all spherical harmonics forms a basis for $L^2(D)$, in agreement with Theorem 2.8.15. □

Exercise 2.8.1 Prove Proposition 2.8.6.

Exercise 2.8.2 Suppose k is a degenerate kernel function given by (2.8.11) with all $\beta_i, \gamma_i \in L^2(a,b)$. Show that the integral operator K, defined by

$$Kv(x) = \int_a^b k(x,y)\, v(y)\, dy$$

is bounded from $L^2(a,b)$ to $L^2(a,b)$.

Exercise 2.8.3 Consider the integral operator (2.8.2). Assume the kernel function k has the form (2.8.9) with each $l_i(x,y)$ continuous for $a \le x, y \le b$ and each $h_i(x,y)$ satisfying (A_1)–(A_2). Prove that k also satisfies (A_1)–(A_2).

Exercise 2.8.4 Show that $\log|x-y|$ and $|x-y|^{-\gamma}$, $\gamma < \frac{1}{2}$, are Hilbert-Schmidt kernel functions.

Exercise 2.8.5 Consider the integral equation

$$\lambda f(x) - \int_0^1 e^{x-y} f(y)\, dy = g(x), \quad 0 \le x \le 1$$

with $g \in C[0,1]$. Denote the integral operator in the equation by K. Consider K as a mapping on $C[0,1]$ into itself, and use the uniform norm $\|\cdot\|_\infty$. Find a bound for the condition number

$$\text{cond}(\lambda - K) \equiv \|\lambda - K\|\, \|(\lambda - K)^{-1}\|$$

within the framework of the space $C[0,1]$. Do this for all values of λ for which $(\lambda - K)^{-1}$ exists as a bounded operator on $C[0,1]$. Comment on how the condition number varies with λ.

Exercise 2.8.6 Similar to Example 2.3.6 of Section 2.3, use the approximation

$$e^{xy} \approx 1 + xy$$

to examine the solvability of the integral equation

$$\lambda u(x) - \int_0^1 e^{xy} u(y)\, dy = f(x), \quad 0 \le x \le 1.$$

To solve the integral equation associated with the kernel $1 + xy$, use the method developed in the proof of Theorem 2.8.10.

Exercise 2.8.7 For any $f \in C[0,1]$, define

$$\mathcal{A}f(x) = \begin{cases} \displaystyle\int_0^x \frac{f(y)}{\sqrt{x^2 - y^2}}\, dy, & 0 < x \le 1, \\ \displaystyle\frac{\pi}{2} f(0), & x = 0. \end{cases}$$

This is called an *Abel integral operator*. Show that $f(x) = x^\alpha$ is an eigenfunction of $\mathcal{A}$ for every $\alpha \ge 0$. What is the corresponding eigenvalue? Can $\mathcal{A}$ be a compact operator?

2.9 The resolvent operator

Let V be a complex Banach space, e.g. $V = C(D)$ the set of continuous complex-valued functions on a closed set D with the uniform norm $\|\cdot\|_\infty$; and let $L : V \to V$ be a bounded linear operator. From the geometric series theorem, Theorem 2.3.1, we know that if $|\lambda| > \|L\|$, then $(\lambda - L)^{-1}$ exists as a bounded linear operator from V to V. It is useful to consider the set of all complex numbers λ for which such an inverse operator $(\lambda - L)^{-1}$ exists on V to V.

Definition 2.9.1 (a) *Let V be a complex Banach space, and let $L : V \to V$ be a bounded linear operator. We say $\lambda \in \mathbb{C}$ belongs to the* resolvent set *of L if $(\lambda - L)^{-1}$ exists as a bounded linear operator from V to V. The resolvent set of L is denoted by $\rho(L)$. The operator $(\lambda - L)^{-1}$ is called the* resolvent operator.
(b) *The set $\sigma(L) = \mathbb{C}\backslash\rho(L)$ is called the* spectrum *of L.*

From the remarks preceding the definition,

$$\{\lambda \in \mathbb{C} \mid |\lambda| > \|L\|\} \subset \rho(L).$$

In addition, we have the following.

Lemma 2.9.2 *The set $\rho(L)$ is open in $\mathbb{C}$; and consequently, $\sigma(L)$ is a closed set.*

Proof. Let $\lambda_0 \in \rho(L)$. We use the perturbation result in Theorem 2.3.5 to show that all points λ in a sufficiently small neighborhood of λ_0 are also in $\rho(L)$; this is sufficient for showing $\rho(L)$ is open. Since $(\lambda_0 - L)^{-1}$ is a bounded linear operator on V to V, consider all $\lambda \in \mathbb{C}$ for which

$$|\lambda - \lambda_0| < \frac{1}{\|(\lambda_0 - L)^{-1}\|}. \tag{2.9.1}$$

Using Theorem 2.3.5, we have that $(\lambda - L)^{-1}$ exists as a bounded operator from V to V, and moreover,

$$\left\|(\lambda - L)^{-1} - (\lambda_0 - L)^{-1}\right\| \le \frac{|\lambda - \lambda_0| \left\|(\lambda_0 - L)^{-1}\right\|^2}{1 - |\lambda - \lambda_0| \left\|(\lambda_0 - L)^{-1}\right\|}. \tag{2.9.2}$$

This shows

$$\{\lambda \in \mathbb{C} \mid |\lambda - \lambda_0| < \varepsilon\} \subset \rho(L)$$

provided ε is chosen sufficiently small. Hence, $\rho(L)$ is an open set.

The inequality (2.9.2) shows that $R(\lambda) \equiv (\lambda - L)^{-1}$ is a continuous function of λ from $\mathbb{C}$ to $\mathcal{L}(V)$. □

A complex number λ can belong to $\sigma(L)$ for several different reasons. Following is a standard classification scheme.

1. *Point spectrum.* $\lambda \in \sigma_P(L)$ means that λ is an eigenvalue of L. Thus there is a nonzero eigenvector $u \in V$ for which $Lu = \lambda u$. Such cases were explored in Section 2.8 with L a compact operator. In this case, the nonzero portion of $\sigma(L)$ consists entirely of eigenvalues, and moreover, 0 is the only possible point in $\mathbb{C}$ to which sequences of eigenvalues can converge.

2. *Continuous spectrum.* $\lambda \in \sigma_C(L)$ means that $(\lambda - L)$ is one-to-one, $\mathcal{R}(\lambda - L) \neq V$, and $\overline{\mathcal{R}(\lambda - L)} = V$. Note that if $\lambda \neq 0$, then L cannot be compact. (Why?) This type of situation, $\lambda \in \sigma_C(L)$, occurs in solving equations $(\lambda - L)u = f$ that are ill-posed. In the case $\lambda = 0$, such equations can often be written as an integral equation of the first kind
$$\int_a^b \ell(x,y)u(y)\,dy = f(x), \quad a \leq x \leq b,$$
with $\ell(x,y)$ continuous and smooth.

3. *Residual spectrum.* $\lambda \in \sigma_R(L)$ means $\lambda \in \sigma(L)$ and that it is in neither the point spectrum nor continuous spectrum. This case can be further subdivided, into cases with $\mathcal{R}(\lambda - L)$ closed and not closed. The latter case consists of ill-posed problems, much as with the case of continuous spectrum. For the former case, the equation $(\lambda - L)u = f$ is usually a well-posed problem; but some change in it is often needed when developing practical methods of solution.

If L is a compact operator on V to V, and if V is infinite dimensional, then it can be shown that $0 \in \sigma(L)$. In addition in this case, if 0 is not an eigenvalue of L, then L^{-1} can be shown to be unbounded on $\mathcal{R}(L)$. Equations $Lu = f$ with L compact make up a significant proportion of ill-posed problems.

2.9.1 $R(\lambda)$ as a holomorphic function

Let $\lambda_0 \in \rho(L)$. Returning to the proof of Lemma 2.9.2, we can write $R(\lambda) \equiv (\lambda - L)^{-1}$ as
$$R(\lambda) = \sum_{k=0}^{\infty} (-1)^k (\lambda - \lambda_0)^k R(\lambda_0)^{k+1} \tag{2.9.3}$$
for all λ satisfying (2.9.1). Thus we have a power series expansion of $R(\lambda)$ about the point λ_0. This can be used to introduce the notion that R is an analytic (or holomorphic) function from $\rho(L) \subset \mathbb{C}$ to the vector space $\mathcal{L}(V)$. Many of the definitions, ideas, and results of *complex analysis* can be extended to analytic operator-valued functions. See [71, p. 566] for an introduction to these ideas.

In particular, we can introduce line integrals. We are especially interested in line integrals of the form

$$g_\Gamma(L) = \frac{1}{2\pi i} \int_\Gamma (\mu - L)^{-1} g(\mu)\, d\mu. \tag{2.9.4}$$

Note that whereas $g : \rho(L) \to \mathbb{C}$, the quantity $g_\Gamma(L) \in \mathcal{L}(V)$. In this integral, Γ is a piecewise smooth curve of finite length in $\rho(L)$; and Γ can consist of several finite disjoint curves. In complex analysis, such integrals occur in connection with studying *Cauchy's theorem*.

Let $\mathcal{F}(L)$ denote the set of all functions g which are analytic on some open set U containing $\sigma(L)$, with the set U dependent on the function g (U need not be connected). For functions in $\mathcal{F}(L)$, a number of important results can be shown for the operators $g(L)$ of (2.9.4) with $g \in \mathcal{F}(L)$. For a proof of the following, see [71, p. 568].

Theorem 2.9.3 *Let $f, g \in \mathcal{F}(L)$, and let $f_\Gamma(L)$, $g_\Gamma(L)$ be defined using* (2.9.4)*, assuming Γ is located within the domain of analyticity of both f and g. Then*
(a) *$f \cdot g \in \mathcal{F}(L)$, and $f_\Gamma(L) \cdot g_\Gamma(L) = (f \cdot g)_\Gamma(L)$;*
(b) *if f has a power series expansion*

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

that is valid in some open disk about $\sigma(L)$, then

$$f_\Gamma(L) = \sum_{n=0}^{\infty} a_n L^n.$$

In numerical analysis, such integrals (2.9.4) become a means for studying the convergence of algorithms for approximating the eigenvalues of L.

Theorem 2.9.4 *Let L be a compact operator from V to V, and let λ_0 be a nonzero eigenvalue of L. Introduce*

$$E(\lambda_0, L) = \frac{1}{2\pi i} \int_{|\lambda - \lambda_0| = \varepsilon} (\lambda - L)^{-1}\, d\lambda \tag{2.9.5}$$

with ε less than the distance from λ_0 to the remaining portion of $\sigma(L)$. Then:
(a) *$E(\lambda_0, L)$ is a projection operator on V to V.*
(b) *$E(\lambda_0, L)V$ is the set of all ordinary and generalized eigenvectors associated with λ_0, i.e.*

$$E(\lambda_0, L)V = \mathcal{N}((\lambda - K)^{\nu(\lambda_0)})$$

with the latter taken from (2.8.27) *and $\nu(\lambda_0)$ the index of λ_0.*

For a proof of these results, see Dunford and Schwartz [71, pp. 566–580].

When L is approximated by a sequence of operators $\{L_n\}$, we can examine the convergence of the eigenspaces of L_n to those of L by means of tools fashioned from (2.9.5). Examples of such analyses can be found in [11], [13], and Chatelin [48].

Exercise 2.9.1 Let $\lambda \in \rho(L)$. Define $d(\lambda)$ to be the distance from λ to $\sigma(L)$,

$$d(\lambda) = \min_{\kappa \in \sigma(L)} |\lambda - \kappa| .$$

Show that

$$\left\|(\lambda - L)^{-1}\right\| \geq \frac{1}{d(\lambda)}.$$

This implies $\left\|(\lambda - L)^{-1}\right\| \to \infty$ as $\lambda \to \sigma(L)$.

Exercise 2.9.2 Let $V = C[0,1]$, and let L be the Volterra integral operator

$$Lv(x) = \int_0^x k(x,y)\, v(y)\, dy, \quad 0 \leq x \leq 1,\ v \in C[0,1]$$

with $k(x,y)$ continuous for $0 \leq y \leq x \leq 1$. What is $\sigma(L)$?

Exercise 2.9.3 Derive the formula (2.9.3).

Exercise 2.9.4 Let $F \subset \rho(L)$ be closed and bounded. Show $(\lambda - L)^{-1}$ is a continuous function of $\lambda \in F$, with

$$\max_{\lambda \in F} \left\|(\lambda - L)^{-1}\right\| < \infty.$$

Exercise 2.9.5 Let L be a bounded linear operator on a Banach space V to V; and let $\lambda_0 \in \sigma(L)$ be an isolated nonzero eigenvalue of L. Let $\{L_n\}$ be a sequence of bounded linear operators on V to V with $\|L - L_n\| \to 0$ as $n \to \infty$. Assume $F \subset \rho(L)$ is closed and bounded. Prove that there exists N such that

$$n \geq N \quad \Longrightarrow \quad F \subset \rho(L_n).$$

This shows that approximating sequences $\{L_n\}$ cannot produce extraneous convergent sequences of approximating eigenvalues.
Hint: Use the result of Exercise 2.9.4.

Exercise 2.9.6 Assume L is a compact operator on V to V, a complex Banach space, and let $\{L_n\}$ be a sequence of approximating bounded linear compact operators with $\|L - L_n\| \to 0$ as $n \to \infty$. Referring to the curve $\Gamma = \{\lambda : |\lambda - \lambda_0| = \varepsilon\}$ of (2.9.5), we have from Exercise 2.9.5 that we can define

$$E(\sigma_n, L_n) = \frac{1}{2\pi i} \int_{|\lambda - \lambda_0| = \varepsilon} (\lambda - L_n)^{-1}\, d\lambda, \quad n \geq N$$

with σ_n denoting the portion of $\sigma(L_n)$ located within Γ. Prove

$$\|E(\sigma_n, L_n) - E(\lambda_0, L)\| \to 0 \quad \text{as } n \to \infty.$$

It can be shown that $\mathcal{R}(E(\sigma_n, L_n))$ consists of combinations of the ordinary and generalized eigenvectors of L_n corresponding to the eigenvalues of L_n within σ_n. In addition, prove that for every $u \in \mathcal{N}((\lambda - K)^{\nu(\lambda_0)})$,

$$E(\sigma_n, L_n)u \to u \quad \text{as } n \to \infty.$$

This shows convergence of approximating simple and generalized eigenfunctions of L_n to those of L.

Suggestion for Further Reading.

See "Suggestion for Further Readings" in Chapter 1.

Springer

http://www.springer.com/978-1-4419-0457-7

Theoretical Numerical Analysis
A Functional Analysis Framework
Atkinson, K.; Han, W.
2009, XVI, 625 p., Hardcover
ISBN: 978-1-4419-0457-7