1

The fundamental ideas

1.1 Unbounded linear operators

One of the key notions in any introductory course on functional analysis is that of a bounded linear operator. If A is such an operator on the Banach space $\mathscr{B}$ then there is a closed bounded subset Spec(A) of the complex plane called its spectrum. The proof that the spectrum is always non-empty is rather indirect, and this is related to the fact that the explicit determination of the spectrum of particular operators is often very difficult.

In this chapter we describe the appropriate context in which one can define and analyse the spectrum of unbounded linear operators, particularly those which are closed or self-adjoint. The description of the spectrum of particular operators will be the main focus of attention throughout the book.

Before one can start to study a differential operator one has to choose the Banach or Hilbert space in which it acts; we mention here that all Banach and Hilbert spaces in the book are assumed to be complex. It turns out that the spectrum of an operator can change depending upon the Banach space in which it acts. There is, however, another problem, namely that differential operators are unbounded when considered as acting on any of the usual Banach or Hilbert spaces. Because of this we cannot even start to study them until we have given a more general definition of a linear operator.

The key to this new definition is to drop the requirement that the domain of the operator is the whole of the Banach space in which the operator acts, and allow it to be a dense linear subspace. The precise specification of that subspace is very important since it turns out that the choice of different subspaces corresponds to the application of different boundary conditions to the same formal operator, which frequently leads to totally different spectra. For these reasons when using the term 'differential operator' we shall understand that we have already chosen the boundary conditions if we are thinking in more applied terms, or that we have already chosen the precise domain of definition of the operator if we are thinking in abstract terms.

Definition We define a linear operator on a Banach space $\mathcal{B}$ to be a pair consisting of a dense linear subspace L of $\mathcal{B}$ together with a linear map $A : L \to \mathcal{B}$. We call L the domain of the operator A and write $\mathrm{Dom}(A) := L$. If $\tilde{L}$ is a linear subspace of $\mathcal{B}$ which contains L and $\tilde{A}f = Af$ for all $f \in L$ then we say that $\tilde{A}$ is an extension of A.

A complex number λ is said to be an eigenvalue of such an operator A if there exists a non-zero $f \in \mathrm{Dom}(A)$ such that $Af = \lambda f$. Since the Banach spaces we are interested in are all spaces of functions, we call f an eigenfunction of the operator A. As in the more elementary theory of bounded linear operators, the set of eigenvalues is not to be confused with the spectrum (defined below), which is often a much larger set.

As an elementary example we choose $\mathcal{B}$ to be the space of all continuous functions on the interval $[a,b]$ and put $Af = -f''$ where $\mathrm{Dom}(A)$ is the set of all smooth (i.e. infinitely differentiable) functions on $[a,b]$. Every complex number is an eigenvalue of this operator, whose spectrum is therefore equal to $\mathbf{C}$. If, however, we take $\mathcal{B}$ to be the space of all continuous periodic functions on the interval $[a,b]$ and the domain to be the set of all periodic smooth functions on $[a,b]$, then the same formula defines a different operator with countable spectrum. The following example is more typical of those which we shall study later.

Example 1.1.1 We consider the operator H given formally by

$$Hf := -f'' \tag{1.1.1}$$

on the following alternative domains in the Hilbert space $L^2(a,b)$. To treat Dirichlet boundary conditions we take the domain L_D consisting of all twice continuously differentiable functions f on $[a,b]$ for which $f(a) = f(b) = 0$. To treat Neumann boundary conditions, however, we take the domain L_N of all twice continuously differentiable functions f on $[a,b]$ for which $f'(a) = f'(b) = 0$. Because we have two different domains the equation (1.1.1) determines two different operators which we shall call H_D and H_N. It is straightforward to determine the eigenvalues of these two operators and to see that 0 is an eigenvalue of H_N but not of H_D. In higher dimensions the difference between the spectrum of the Laplacian under Dirichlet and Neumann boundary conditions is much greater than in this example. □

In this chapter we shall illustrate our ideas by means of very simple operators such as H_D and H_N. The only reason for this is that we do not wish the reader to have to cope with the abstract theory and its applications at the same time. We ask the reader at this point to accept our reassurance that the conditions of the abstract theorems which we shall prove are actually verifiable in a wide range of more interesting applications.

The continuity of bounded linear operators is so useful that we need to have a replacement for it in our more general situation. This is provided by the notion of closedness. We will henceforth use the expression $\lim_{n\to\infty} f_n = f$ without further comment to mean that $\|f_n - f\|$ converges to zero as $n \to \infty$.

Definition Let A be an operator on $\mathcal{B}$ with domain L. We say that A is closed if whenever f_n is a sequence in L with limit $f \in \mathcal{B}$ and there exists $g \in \mathcal{B}$ such that $\lim_{n\to\infty} Af_n = g$, it follows that $f \in L$ and that $Af = g$.

There is an alternative formulation of this idea. The product $\mathcal{B}_1 \times \mathcal{B}_2$ of two Banach spaces $\mathcal{B}_1$ and $\mathcal{B}_2$ becomes a Banach space if we provide it with the norm

$$\|(f, g)\| := \{\|f\|^2 + \|g\|^2\}^{1/2}.$$

Other equivalent choices of the norm can be made, but one advantage of the above definition is that if $\mathcal{B}_i$ are both Hilbert spaces then $\mathcal{B}_1 \times \mathcal{B}_2$ is also a Hilbert space for the above norm. If we define the graph of A to be the set

$$\{(f, g) : f \in L,\ g \in \mathcal{B} \quad \text{and} \quad Af = g\},$$

then the operator is closed if and only if its graph is a closed subspace of $\mathcal{B} \times \mathcal{B}$.

The closed graph theorem states that if a closed operator has domain equal to $\mathcal{B}$, then it has finite norm. While conceptually extremely valuable, this result has the weakness of not giving any information about the size of the norm. We shall see below that the size of the norm of resolvent operators is important in locating the spectrum of A.

Definition If A is a linear operator on $\mathscr{B}$ with domain L then its spectrum Spec(A) is defined as follows. We say that a complex number z does not lie in Spec(A) if the operator $(z - A)$ maps L one-one onto $\mathscr{B}$, and the inverse (or resolvent) operator, which we shall denote by $R(z, A)$ or $(z - A)^{-1}$, is bounded.

The following lemma explains why the notion of closedness is so important. The statement and proof involve using analytic function theory for operator-valued functions of a complex variable. The definitions and proofs of the relevant results are simple adaptations of the corresponding results for complex-valued functions, and we leave readers to write out the details for themselves.

Lemma 1.1.2 *If the operator A does not have spectrum equal to the whole of the complex plane* $\mathbf{C}$ *then A must be closed. The spectrum* Spec(A) *of a linear operator A is always closed. More specifically let $z \notin$* Spec(A) *and let $c = \|R(z, A)\|$. Then the spectrum does not intersect the ball*

$$\{w \in \mathbf{C} : |z - w| < c^{-1}\}.$$

The resolvent operator is a norm analytic function of z and satisfies the resolvent equations

$$R(z, A) - R(w, A) = -(z - w)R(z, A)R(w, A), \tag{1.1.2}$$

$$R(z, A)R(w, A) = R(w, A)R(z, A), \tag{1.1.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}z}R(z, A) = -R(z, A)^2, \tag{1.1.4}$$

for all $z, w \notin$ Spec(A).

Proof Suppose that $z \notin$ Spec(A) and let $B = (z - A)^{-1}$ be the inverse operator, which is bounded by hypothesis. Let $f_n \in$ Dom(A), $\lim_{n\to\infty} f_n = f$, $\lim_{n\to\infty} Af_n = g$ and $h_n := (z - A)f_n$. Then

$$\lim_{n\to\infty} h_n = \lim_{n\to\infty}\{zf_n - Af_n\} = zf - g,$$

so

$$B(zf - g) = \lim_{n\to\infty}\{Bh_n\} = \lim_{n\to\infty}\{f_n\} = f.$$

This implies that $f \in$ Dom(A) and $(z - A)f = zf - g$, or $Af = g$. Hence A is closed.

The remainder of the proof is very similar to the case when A is bounded. Consider the bounded operator C defined by

$$C := \sum_{n=0}^{\infty}(-u)^n B^{n+1}, \qquad (1.1.5)$$

where the series is norm convergent if $|u| < \|B\|^{-1}$. The operator C satisfies the identities

$$C = B - uBC \quad , \quad C = B - uCB.$$

The first identity implies that the kernel $\mathrm{Ker}(C)$ and range $\mathrm{Ran}(C)$ of C satisfy

$$\mathrm{Ker}(C) \subseteq \mathrm{Ker}(B) \quad , \quad \mathrm{Ran}(C) \subseteq \mathrm{Ran}(B),$$

while the second implies

$$\mathrm{Ker}(B) \subseteq \mathrm{Ker}(C) \quad , \quad \mathrm{Ran}(B) \subseteq \mathrm{Ran}(C).$$

Since B has kernel $\{0\}$ and range $\mathrm{Dom}(A)$ we deduce that C is a bounded linear operator mapping $\mathscr{B}$ one-one onto $\mathrm{Dom}(A)$. If $f \in \mathrm{Dom}(A)$ and $g = (z-A)f$ then $f = Bg$, so $Cg = f - uCf$. Hence $C(z+u-A)f = f$. Since $(z+u-A)Cf = f$ for all $f \in \mathscr{B}$ by a similar calculation, we conclude that $C = (z+u-A)^{-1}$. This establishes both that $z+u \notin \mathrm{Spec}(A)$ if $|u| < \|(z-A)^{-1}\|^{-1}$ and that

$$(z+u-A)^{-1} = \sum_{n=0}^{\infty}(-u)^n (z-A)^{-(n+1)}. \qquad (1.1.6)$$

The norm convergence of this series is more than enough justification for saying that the resolvent is a norm analytic function of z. The resolvent equation (1.1.4) is proved by differentiating the series (1.1.5) term by term, the justification for this being the same as for complex-valued power series. The resolvent equation (1.1.2) is proved by multiplying both sides by $(z-A)$, in which case it becomes an elementary identity. Upon interchanging w and z we see that (1.1.2) implies (1.1.3). □

It is possible to construct closed operators whose spectrum is either empty or equal to the whole complex plane. However, we shall mainly be interested in studying self-adjoint operators, whose spectrum is always a non-empty subset of the real line.

Although all of the above suggest that we should only study closed operators, there is a practical problem with this, namely that differential operators are usually defined initially on simple domains where they are not closed. This problem is overcome by yet more definitions.

Lemma 1.1.3 *An operator A on $\mathcal{B}$ with domain L is said to be closable if it has a closed extension $\tilde{A}$. In this case there is a closed extension $\overline{A}$, which we call its closure, whose domain is smallest among all closed extensions.*

Proof We define $\mathcal{D}$ to be the set of $f \in \mathcal{B}$ for which there exist $f_n \in \mathrm{Dom}(A)$ and $g \in \mathcal{B}$ such that $\lim_{n\to\infty} f_n = f$ and $\lim_{n\to\infty} Af_n = g$. Since $\tilde{A}$ is a closed extension of A it follows that $f \in \mathrm{Dom}(\tilde{A})$ and $\tilde{A}f = g$. Hence g is uniquely determined by f in the above situation. We define $\overline{A}f = g$ with $\mathrm{Dom}(\overline{A}) = \mathcal{D}$. Clearly $\overline{A}$ is an extension of A and every closed extension of A is also an extension of $\overline{A}$. The graph of $\overline{A}$ is the closure of the graph of A in the Banach space $\mathcal{B} \times \mathcal{B}$. Hence $\overline{A}$ is a closed operator. □

Many of the differential operators which we will study have the key property of self-adjointness. An intermediate but much more elementary property is that of being symmetric.

Definition We say that an operator H with dense domain L in a Hilbert space $\mathcal{H}$ is symmetric if for all $f, g \in L$ we have

$$\langle Hf, g\rangle = \langle f, Hg\rangle.$$

Considering again the operator $Hf = -f''$ on $L^2(a,b)$, choose either of the domains of Example 1.1.1. By use of the identity

$$\int_a^b (f''\overline{g} - f\overline{g''})\mathrm{d}x = [f'\overline{g} - f\overline{g'}]_a^b$$

we see that both H_D and H_N are symmetric.

Lemma 1.1.4 *Every symmetric operator H is closable and its closure is also symmetric.*

Proof Let $\mathcal{D}$ be the set of $f \in \mathcal{H}$ for which there exist $f_n \in \mathrm{Dom}(H)$ and $g \in \mathcal{H}$ such that $\lim_{n\to\infty} f_n = f$ and $\lim_{n\to\infty} Hf_n = g$. It is easy to see that $\mathcal{D}$ is a linear subspace of $\mathcal{H}$ containing $\mathrm{Dom}(H)$. If $h \in \mathrm{Dom}(H)$ then

$$\langle g, h\rangle = \lim_{n\to\infty}\langle Hf_n, h\rangle = \lim_{n\to\infty}\langle f_n, Hh\rangle = \langle f, Hh\rangle.$$

Now g is uniquely determined by the functional $h \to \langle g, h\rangle$ on $\mathrm{Dom}(H)$ because $\mathrm{Dom}(H)$ is dense in $\mathcal{H}$ by hypothesis. Therefore g is uniquely determined by f. If we define $\overline{H}f = g$ then it follows that $\overline{H}$ is linear on its domain $\mathcal{D}$. Moreover, the graph of $\overline{H}$ is the closure of the graph of H. If $h_n \in \mathrm{Dom}(H)$, $\lim_{n\to\infty} h_n = h \in \mathscr{D}$ and $\lim_{n\to\infty} Hh_n = k$ then we have already shown that

$$\langle \overline{H}f, h_n\rangle = \langle f, Hh_n\rangle.$$

In the limit as $n \to \infty$ we get

$$\langle \overline{H}f, h\rangle = \langle f, k\rangle = \langle f, \overline{H}h\rangle$$

which establishes that $\overline{H}$ is symmetric. ☐

The above lemma enables us to concentrate henceforth on closed operators. The above definitions and Lemma 1.1.4 allow us to be a little careless in distinguishing between a closable operator and its closure and we shall often take advantage of this.

1.2 Self-adjointness

There is a difference between symmetry and self-adjointness for an unbounded operator A on a Hilbert space $\mathscr{H}$, which does not correspond to anything in the theory of bounded linear operators. At first this seems to be an annoying technicality, but in fact it is of profound importance. The condition of self-adjointness is much more demanding and difficult to verify, but unless it is met one cannot apply the very powerful machinery of spectral theory. In the context of differential operators, the issue is whether one has fully specified the boundary conditions appropriate to the particular differential operator one is studying.

Definition If A is a linear operator on a Hilbert space $\mathscr{H}$ then the adjoint operator A^* is determined by the condition that

$$\langle Af, g\rangle = \langle f, A^*g\rangle$$

for all $f \in \mathrm{Dom}(A)$ and $g \in \mathrm{Dom}(A^*)$. The domain of A^* is defined to be the set $\mathscr{D}$ of all $g \in \mathscr{H}$ for which there exists $k \in \mathscr{H}$ such that

$$\langle Af, g\rangle = \langle f, k\rangle$$

for all $f \in \mathrm{Dom}(A)$. After showing that k is unique and we will put $A^*g = k$.

Lemma 1.2.1 *If A is a closed linear operator with dense domain then the adjoint A^* is also a closed linear operator with dense domain.*

Proof If $g \in \mathscr{D}$ and k, k' are two elements of $\mathscr{H}$ such that

$$\langle Af, g \rangle = \langle f, k \rangle = \langle f, k' \rangle$$

for all $f \in \mathrm{Dom}(A)$ then $\langle f, k - k' \rangle = 0$ for all such f. The density of $\mathrm{Dom}(A)$ implies that $k = k'$. It is thus permissible to define A^* on $\mathscr{D}$ by $A^* g = k$. If L is the graph of A^* then $(g, k) \in L$ if and only if (g, k) is orthogonal to

$$M := \{(Af, -f) \in \mathscr{H} \times \mathscr{H} : f \in \mathrm{Dom}(A)\}.$$

The orthogonal complement $M^\perp$ of the linear subspace M must be a closed linear subspace, so A^* is a closed linear operator.

It only remains to prove that A^* is densely defined. If $h \in \mathscr{H}$ satisfies $\langle h, g \rangle = 0$ for all $g \in \mathscr{D}$, then $(h, 0)$ is orthogonal to L. But $L^\perp = M^{\perp\perp} = \overline{M}$. Hence there exists a sequence $\{f_n\}_{n=1}^\infty \in \mathrm{Dom}(A)$ such that $\lim_{n\to\infty} f_n = 0$ and $\lim_{n\to\infty} Af_n = h$. But A is assumed to be closed, so $h = A0 = 0$. Thus $\mathscr{D}^\perp = 0$, and $\mathscr{D}$ is a dense linear subspace of $\mathscr{H}$. ☐

If H is a symmetric operator then it is easy to see that the adjoint H^* is an extension of H. We say that H is self-adjoint if H is symmetric and $\mathrm{Dom}(H) = \mathrm{Dom}(H^*)$. This is equivalent to requiring that $H = H^*$, and implies that H is closed. We say that H is essentially self-adjoint if it is symmetric and its closure is self-adjoint. Our next lemma gives a method of proving essential self-adjointness, but it is only useful for simple operators whose eigenvectors can be determined explicitly. In the proof of the lemma we assume that $\mathscr{H}$ is separable, or equivalently that it has a countable complete orthonormal set. This is valid in all applications to differential operators, and the interested reader can no doubt provide the necessary modification to non-separable Hilbert spaces. We shall make the same assumption at many other places in the book without comment.

Lemma 1.2.2 *Let H be a symmetric operator on $\mathscr{H}$ with domain L, and let $\{f_n\}_{n=1}^\infty$ be a complete orthonormal set in $\mathscr{H}$. If each f_n lies in L and there exist $\lambda_n \in \mathbf{R}$ such that $Hf_n = \lambda_n f_n$ for every n, then H is essentially self-adjoint. Moreover, the spectrum of $\overline{H}$ is the closure in $\mathbf{R}$ of the set of all λ_n.*

Proof If $f = \sum_{n=1}^\infty \alpha_n f_n$ lies in L and

$$g := Hf = \sum_{n=1}^\infty \beta_n f_n,$$

then

$$\beta_m = \langle g, f_m \rangle = \langle Hf, f_m \rangle = \langle f, Hf_m \rangle = \lambda_m \langle f, f_m \rangle = \lambda_m \alpha_m.$$

The requirements that $f, g \in \mathscr{H}$ force

$$\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty \quad , \quad \sum_{n=1}^{\infty} |\beta_n|^2 < \infty$$

and hence

$$\sum_{n=1}^{\infty} (1 + \lambda_n^2)|\alpha_n|^2 < \infty.$$

We now define an operator $\tilde{H}$ as follows. Let $\tilde{L}$ be the set of all $f \in \mathscr{H}$ of the form $f = \sum_{n=1}^{\infty} \alpha_n f_n$ where

$$\sum_{n=1}^{\infty} (1 + \lambda_n^2)|\alpha_n|^2 < \infty,$$

and for such f define

$$\tilde{H}f := \sum_{n=1}^{\infty} \alpha_n \lambda_n f_n.$$

It is clear that $\tilde{H}$ is an extension of H. We first determine its spectrum.

Let S be the closure of the set $\{\lambda_n : 1 \le n < \infty\}$. Each λ_n is an eigenvalue of $\tilde{H}$, and $\mathrm{Spec}(\tilde{H})$ is closed, so $S \subseteq \mathrm{Spec}(\tilde{H})$. If $z \notin S$ then the operator A defined on $\mathscr{H}$ by

$$A\left(\sum_{n=1}^{\infty} \alpha_n f_n\right) := \sum_{n=1}^{\infty} \alpha_n (z - \lambda_n)^{-1} f_n$$

is bounded and one-one. By expanding everything in terms of the complete orthonormal set $\{f_n\}_{n=1}^{\infty}$ one can check that its range is precisely $\tilde{L}$ and that $(z - \tilde{H})Af = f$ for all $f \in \mathscr{H}$. Thus $z \notin \mathrm{Spec}(\tilde{H})$ and $A = (z - \tilde{H})^{-1}$. The above together imply that $S = \mathrm{Spec}(\tilde{H})$.

We claim that $\tilde{H}$ is equal to the closure of H. Since $\mathrm{Spec}(\tilde{H})$ is not equal to $\mathbf{C}$, Lemma 1.1.2 implies that $\tilde{H}$ is a closed operator. If $g := \sum_{n=1}^{\infty} \alpha_n f_n \in \mathrm{Dom}(\tilde{H})$ and we put $g_m := \sum_{n=1}^{m} \alpha_n f_n$ then $\lim_{m\to\infty} g_m = g$ and

$$\lim_{m\to\infty} \{Hg_m\} = \lim_{m\to\infty} \left\{\sum_{n=1}^{m} \alpha_n \lambda_n f_n\right\} = \sum_{n=1}^{\infty} \alpha_n \lambda_n f_n = \tilde{H}g.$$

This establishes the stated claim.

We finally prove that $\tilde{H}$ is self-adjoint. If $f \in \mathrm{Dom}(\tilde{H}^*)$ and $\tilde{H}^* f = k$, then

$$\begin{aligned}\langle k, f_n \rangle &= \langle \tilde{H}^* f, f_n \rangle \\ &= \langle f, \tilde{H} f_n \rangle \\ &= \lambda_n \langle f, f_n \rangle.\end{aligned}$$

If $f = \sum_{n=1}^{\infty} \alpha_n f_n$ the above implies that $k = \sum_{n=1}^{\infty} \lambda_n \alpha_n f_n$. It follows that $f \in \mathrm{Dom}(\tilde{H})$, and hence that $\tilde{H}^* = \tilde{H}$. □

Example 1.2.3 The operator H_D defined in Example 1.1.1 is essentially self-adjoint on its domain L_D. To see this we first define the orthonormal sequence of functions

$$f_n(x) := \left(\frac{2}{b-a} \right)^{1/2} \sin \left\{ \frac{n\pi(x-a)}{(b-a)} \right\},$$

where $n \in \mathbf{N}$. The fact that these are eigenfunctions with eigenvalues $n^2\pi^2/(b-a)^2$ is easy to verify. The harder fact needed to apply Lemma 1.2.2 is that $\{f_n\}_{n=1}^{\infty}$ is a complete orthonormal set in $L^2(a,b)$; this is a standard result of Fourier analysis, but was proved long after the classic paper of Fourier (1822). A similar argument applies to H_N. □

Example 1.2.4 Another example of a similar type arises in connection with Legendre's equation

$$-\frac{\mathrm{d}}{\mathrm{d}x} \left\{ (1-x^2) \frac{\mathrm{d}f}{\mathrm{d}x} \right\} = \lambda f.$$

We show that the symmetric operator H on $L^2(-1,1)$ defined by

$$Hf := -\frac{\mathrm{d}}{\mathrm{d}x} \left\{ (1-x^2) \frac{\mathrm{d}f}{\mathrm{d}x} \right\}$$

is essentially self-adjoint on the domain of all twice continuously differentiable functions on $[-1,1]$. It is elementary but tedious to check that the Legendre polynomials $P_n(x)$ defined for $n \geq 0$ by

$$P_n(x) := \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \{(x^2-1)^n\}$$

satisfy Legendre's equation with eigenvalues $\lambda_n = n(n+1)$. A further direct computation, using integration by parts repeatedly, establishes that the functions

$$f_n(x) := \left(\frac{2n+1}{2} \right)^{1/2} P_n(x)$$