Appendix A
Complex Numbers

The purpose of this short appendix is to review the basics of complex numbers and complex arithmetic, which are used throughout much of the text.

A *complex number* is an expression of the form $z = x + \mathrm{i}\, y$, where $x, y \in \mathbb{R}$ are real and $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. The set of all complex numbers is denoted by $\mathbb{C}$. We call $x = \operatorname{Re} z$ the *real part* of z and $y = \operatorname{Im} z$ the *imaginary part* of $z = x + \mathrm{i}\, y$. (*Note*: The imaginary part is the real number y, *not* $\mathrm{i}\, y$.) A real number x is merely a complex number with zero imaginary part, $\operatorname{Im} z = 0$, and so we may regard $\mathbb{R} \subset \mathbb{C}$. Complex addition and multiplication are based on simple adaptations of the rules of real arithmetic to include the identity $\mathrm{i}^2 = -1$, and so

$$\begin{aligned} (x + \mathrm{i}\, y) + (u + \mathrm{i}\, v) &= (x + u) + \mathrm{i}\,(y + v), \\ (x + \mathrm{i}\, y)\,(u + \mathrm{i}\, v) &= (x\,u - y\,v) + \mathrm{i}\,(x\,v + y\,u). \end{aligned} \tag{A.1}$$

Complex numbers enjoy all the usual laws of real addition and multiplication, *including commutativity*: $z\,w = w\,z$.

We can identify a complex number $x + \mathrm{i}\, y$ with a vector $(x, y) \in \mathbb{R}^2$ in the real, two-dimensional plane. For this reason, $\mathbb{C}$ is sometimes referred to as the *complex plane*. (Although keep in mind that, as a complex vector space, $\mathbb{C}$ is only one-dimensional.) Based on this identification, we shall employ the standard terminology of planar vector calculus — domain, curve, etc. — without alteration. Complex addition (A.1) corresponds to vector addition, but the vector interpretation of complex multiplication is more obscure.

The *complex conjugate* of $z = x + \mathrm{i}\, y$ is $\overline{z} = x - \mathrm{i}\, y$. Note that $\operatorname{Re} \overline{z} = \operatorname{Re} z$, while $\operatorname{Im} \overline{z} = -\operatorname{Im} z$. Geometrically, the complex conjugate of z is obtained by reflecting the corresponding vector through the real axis, as illustrated in Figure A.1. In particular, $\overline{z} = z$ if and only if z is real. In general,

$$\operatorname{Re} z = \frac{z + \overline{z}}{2}, \qquad \operatorname{Im} z = \frac{z - \overline{z}}{2\,\mathrm{i}}. \tag{A.2}$$

Complex conjugation is compatible with complex arithmetic:

$$\overline{z + w} = \overline{z} + \overline{w}, \qquad \overline{z\,w} = \overline{z}\;\overline{w}.$$

In particular, the product of a complex number and its conjugate,

$$z\,\overline{z} = (x + \mathrm{i}\, y)\,(x - \mathrm{i}\, y) = x^2 + y^2, \tag{A.3}$$

is real and nonnegative. Its square root is known as the *modulus* or *norm* of the complex number $z = x + \mathrm{i}\, y$, and written

$$|\, z \,| = \sqrt{x^2 + y^2}\,. \tag{A.4}$$

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0, © Springer International Publishing Switzerland 2014

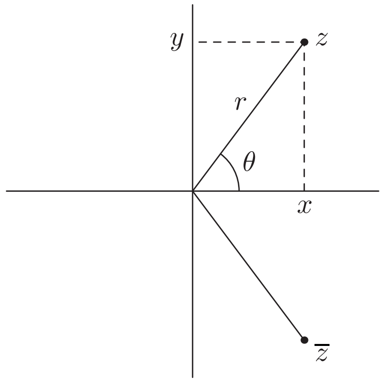

Figure A.1. Complex numbers.

Note that $|z| \geq 0$, with $|z| = 0$ if and only if $z = 0$. The modulus $|z|$ generalizes the absolute value of a real number and coincides with the standard Euclidean norm in the (x, y)–plane. This implies the validity of the triangle inequality

$$|z + w| \leq |z| + |w|. \tag{A.5}$$

Equation (A.3) can be rewritten in terms of the modulus as

$$z\,\overline{z} = |z|^2. \tag{A.6}$$

Rearranging the factors, we deduce the formula for the reciprocal of a nonzero complex number:

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2}, \qquad z \neq 0, \qquad \text{or, equivalently,} \qquad \frac{1}{x + \mathrm{i}\,y} = \frac{x - \mathrm{i}\,y}{x^2 + y^2}. \tag{A.7}$$

The general formula for complex division,

$$\frac{w}{z} = \frac{w\,\overline{z}}{|z|^2} \qquad \text{or} \qquad \frac{u + \mathrm{i}\,v}{x + \mathrm{i}\,y} = \frac{(x\,u + y\,v) + \mathrm{i}\,(x\,v - y\,u)}{x^2 + y^2}, \tag{A.8}$$

is an immediate consequence.

The modulus of a complex number,

$$r = |z| = \sqrt{x^2 + y^2}\,,$$

is one component of its polar coordinate representation

$$x = r\cos\theta, \qquad y = r\sin\theta \qquad \text{or} \qquad z = r(\cos\theta + \mathrm{i}\,\sin\theta). \tag{A.9}$$

The polar angle θ, which measures the angle that the line connecting z to the origin makes with the horizontal axis, is known as the *phase*, and written

$$\theta = \mathrm{ph}\, z. \tag{A.10}$$

As such, the phase is defined only up to an integer multiple of 2π. The unique *principal value* of the phase is restricted to $-\pi < \mathrm{ph}\, z \leq \pi$. A more common term for the polar angle is the *argument* of z, written $\arg z = \operatorname{ph} z$. However, in conformity with [**85**, **86**], we prefer to use "phase" here, in part to avoid confusion with the argument z of a function $f(z)$.

Euler's celebrated formula for the complex exponential,

$$e^{\,\mathrm{i}\,\theta} = \cos\theta + \mathrm{i}\,\sin\theta, \tag{A.11}$$

can be used to compactly rewrite the polar form (A.9) of a complex number as

$$z = r\,e^{\,\mathrm{i}\,\theta}, \qquad \text{where} \qquad r = |\,z\,|, \qquad \theta = \operatorname{ph} z. \tag{A.12}$$

Consequently, the complex logarithm has the form

$$\log z = \log(r\,e^{\,\mathrm{i}\,\theta}) = \log r + \log e^{\,\mathrm{i}\,\theta} = \log r + \mathrm{i}\,\theta = \log|\,z\,| + \mathrm{i}\,\operatorname{ph} z. \tag{A.13}$$

More generally, the complex exponential is given by

$$e^{z} = e^{x}\cos y + \mathrm{i}\,e^{x}\sin y, \qquad \text{for} \qquad z = x + \mathrm{i}\,y. \tag{A.14}$$

We note that the modulus and phase of a product of complex numbers can be readily computed:

$$|\,z\,w\,| = |\,z\,|\,|\,w\,|, \qquad \operatorname{ph}(z\,w) = \operatorname{ph} z + \operatorname{ph} w, \tag{A.15}$$

the latter formula requiring that we allow multiply valued phases; the formula does *not* hold as stated for all z, w when the principal value of the phase is used. Similarly, the modulus and phase of the reciprocal of a nonzero complex number are

$$\left|\,\frac{1}{z}\,\right| = \frac{1}{|\,z\,|}\,, \qquad \operatorname{ph}\left(\frac{1}{z}\right) = -\operatorname{ph} z. \tag{A.16}$$

On the other hand, complex conjugation preserves the modulus, but negates the phase:

$$|\,\overline{z}\,| = |\,z\,|, \qquad \operatorname{ph}\overline{z} = -\operatorname{ph} z. \tag{A.17}$$

The latter formula is not valid for the principal value of the phase when z lies on the negative real axis.

Appendix B
Linear Algebra

In this appendix, we collect basic results and definitions from linear algebra that are used in our study of partial differential equations. The reader is referred to [**89**] for the proofs and further details.

B.1 Vector Spaces and Subspaces

Vector spaces and their ancillary structures provide the common language of linear algebra. The basic definition is modeled on the prototypical finite-dimensional example: the *Euclidean space* $\mathbb{R}^n$, which is the set of all real (column) vectors with n entries, equipped with the operations of vector addition and scalar multiplication. More generally:

Definition B.1. A (real) *vector space* is a set V equipped with two operations:

(i) *Addition*: adding any pair of elements $\mathbf{v}, \mathbf{w} \in V$ produces another vector $\mathbf{v} + \mathbf{w} \in V$.

(ii) *Scalar Multiplication*: multiplying an element $\mathbf{v} \in V$ by a scalar $c \in \mathbb{R}$ produces a vector $c\,\mathbf{v} \in V$.

These are subject to the following axioms: for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all scalars $c, d \in \mathbb{R}$,

(a) *Commutativity of Addition*: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.

(b) *Associativity of Addition*: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.

(c) *Additive Identity*: There is a zero element $\mathbf{0} \in V$ satisfying $\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$.

(d) *Additive Inverse*: For each $\mathbf{v} \in V$ there is an element $-\mathbf{v} \in V$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$.

(e) *Distributivity*: $(c + d)\,\mathbf{v} = (c\,\mathbf{v}) + (d\,\mathbf{v})$, and $c\,(\mathbf{v} + \mathbf{w}) = (c\,\mathbf{v}) + (c\,\mathbf{w})$.

(f) *Associativity of Scalar Multiplication*: $c\,(d\,\mathbf{v}) = (c\,d)\,\mathbf{v}$.

(g) *Unit for Scalar Multiplication*: the scalar $1 \in \mathbb{R}$ satisfies $1\,\mathbf{v} = \mathbf{v}$.

Complex vector spaces are defined in an identical manner, the only difference being that the scalars are allowed to be complex numbers. In this case, the prototype is the space $\mathbb{C}^n$ consisting of column vectors with n complex entries.

While finite-dimensional vector spaces play a significant role in the study of partial differential equations, particularly in the design of numerical solution schemes, for us the more important examples are infinite-dimensional vector spaces whose elements ("vectors") are functions. The main example is the following:

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0,

Example B.2. Let $I \subset \mathbb{R}$ be an interval. The *function space* $\mathcal{F} = \mathcal{F}(I)$, whose elements are all real-valued functions $f(x)$ defined for $x \in I$, has the structure of a vector space. Addition of functions in $\mathcal{F}$ is defined in the usual manner: $(f+g)(x) = f(x)+g(x)$ for all $x \in I$. Multiplication by scalars $c \in \mathbb{R}$ is the same as multiplication by constants, $(c\,f)(x) = c\,f(x)$. The zero element is the constant function that is identically 0 for all $x \in I$. With these operations, all the vector space axioms listed in Definition B.1 are valid, and hence $\mathcal{F}(I)$ is a real vector space.

More generally, if $\Omega \subset \mathbb{R}^n$ is any subset of n-dimensional Euclidean space, the function space $\mathcal{F}(\Omega)$ is defined as the set of all real-valued functions $f(x_1, \ldots, x_n)$ defined for all $x = (x_1, \ldots, x_n) \in \Omega$. Addition and scalar (constant) multiplication of functions are defined in the same manner.

A *subspace* of a vector space V is a subset $W \subset V$ that is a vector space in its own right. In particular, a subspace W *must* contain the zero element of V.

Proposition B.3. *A nonempty subset $W \subset V$ of a vector space is a subspace if and only if*

(*a*) *for every $\mathbf{v}, \mathbf{w} \in W$, the sum $\mathbf{v} + \mathbf{w} \in W$, and*

(*b*) *for every $\mathbf{v} \in W$ and every $c \in \mathbb{R}$, the scalar product $c\mathbf{v} \in W$.*

For example, a complete list of subspaces of $V = \mathbb{R}^3$ is (*i*) the origin $\{\mathbf{0}\}$; (*ii*) every line through the origin; (*iii*) every plane through the origin; (*iv*) all of $\mathbb{R}^3$.

Example B.4. Here are some examples of subspaces of the function space $\mathcal{F}(I)$.

(*a*) The space $\mathcal{P}^{(n)}$ of polynomials of degree $\leq n$.

(*b*) The space $\mathrm{C}^0(I)$ of all continuous functions on the interval I.

(*c*) The space $\mathrm{C}^n(I)$ consisting of all functions $f(x)$ that have n continuous derivatives $f'(x), f''(x), \ldots, f^{(n)}(x)$ on† I.

(*d*) The space $\mathrm{C}^\infty(I) = \bigcap_{n \geq 0} \mathrm{C}^n(I)$ of infinitely differentiable, or *smooth*, functions is also a subspace.

(*e*) The space $\mathcal{A}(I)$ of analytic functions. Recall that a function $f(x)$ is called *analytic* at a point a if it is smooth, and, moreover, its Taylor series

$$f(a) + f'(a)\,(x-a) + \tfrac{1}{2}\,f''(a)\,(x-a)^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}\,(x-a)^n \tag{B.1}$$

converges to $f(x)$ for all x sufficiently close to a. (The series is not required to converge on the entire interval I.) Not every smooth function is analytic, and so $\mathcal{A}(I) \subsetneq \mathrm{C}^\infty(I)$; see Exercise 11.3.21 for an explicit example.

B.2 Bases and Dimension

Definition B.5. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ belong to a vector space V. A sum of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \sum_{i=1}^{k} c_i\mathbf{v}_i, \tag{B.2}$$

† We use one-sided derivatives at any endpoint that belongs to the interval.

where the coefficients $c_1, c_2, \ldots, c_k$ are any scalars, is known as a *linear combination* of the elements $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Their *span* is the subspace $W = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ consisting of all possible linear combinations.

Definition B.6. The elements $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ are called *linearly dependent* if there exist scalars $c_1, \ldots, c_k$, *not all zero*, such that

$$c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = \mathbf{0}. \tag{B.3}$$

Elements that are not linearly dependent are called *linearly independent*.

In particular, a collection of functions $f_1(x), \ldots, f_n(x)$ is linearly dependent if and only if there exist constants $c_1, \ldots, c_n$, *not all zero*, such that the linear combination

$$c_1 f_1(x) + \cdots + c_n f_n(x) \equiv 0 \tag{B.4}$$

is identically zero. Conversely, if the only choice of constants for which (B.4) holds is $c_1 = \cdots = c_n = 0$, then the functions are linearly independent.

Definition B.7. A *basis* of a vector space V is a finite collection of elements $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ that (*a*) spans V, and (*b*) is linearly independent.

The simplest example is the *standard basis* of $\mathbb{R}^n$, consisting of the n vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \ldots, \qquad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \tag{B.5}$$

so that $\mathbf{e}_i$ is the vector with 1 in the i^{th} slot and 0's elsewhere. However, there are many other bases of $\mathbb{R}^n$; indeed, any n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ form a basis.

Lemma B.8. *The elements $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis of V if and only if every $\mathbf{v} \in V$ can be written* uniquely *as a linear combination of the basis elements*:

$$\mathbf{v} = c_1 \mathbf{v}_1 + \cdots + c_n \mathbf{v}_n = \sum_{i=1}^{n} c_i \mathbf{v}_i. \tag{B.6}$$

The coefficients $(c_1, \ldots, c_n)$ are called the coordinates *of the vector $\mathbf{v}$ with respect to the given basis.*

Theorem B.9. *Suppose the vector space V has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then every other basis of V has the same number of elements in it. This number is called the* dimension *of V, and written* $\dim V = n$.

On the other hand, if the vector space contains infinitely many linearly independent elements, then it does not have a basis in the sense of Definition B.7, and is thus *infinite-dimensional*. All of the function spaces and subspaces listed above are infinite-dimensional vector spaces. An example of a finite-dimensional function space is the space $\mathcal{P}^{(n)} \subset \mathcal{F}(\mathbb{R})$ consisting of all polynomials $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ of degree $\leq n$. The monomials $1, x, x^2, \ldots, x^n$ form a basis, and hence $\mathcal{P}^{(n)}$ has dimension $n+1$. (On the other hand, the vector space containing *all* polynomials is infinite-dimensional.)

B.3 Inner Products and Norms

The dot product on Euclidean space $\mathbb{R}^n$ plays an essential role in geometry, analysis, and mechanics. Its basic properties inspire the general definition of an inner product on a vector space.

Definition B.10. An *inner product* on the real vector space V is a pairing that takes two elements $\mathbf{v}, \mathbf{w} \in V$ and produces a real number $\langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{R}$, subject to the following three axioms for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and scalars $c, d \in \mathbb{R}$.

(*i*) *Bilinearity*:

$$\begin{aligned} \langle c\,\mathbf{u} + d\,\mathbf{v}, \mathbf{w} \rangle &= c\,\langle \mathbf{u}, \mathbf{w} \rangle + d\,\langle \mathbf{v}, \mathbf{w} \rangle, \\ \langle \mathbf{u}, c\,\mathbf{v} + d\,\mathbf{w} \rangle &= c\,\langle \mathbf{u}, \mathbf{v} \rangle + d\,\langle \mathbf{u}, \mathbf{w} \rangle. \end{aligned} \tag{B.7}$$

(*ii*) *Symmetry*:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle. \tag{B.8}$$

(*iii*) *Positivity*:

$$\langle \mathbf{v}, \mathbf{v} \rangle > 0 \qquad \text{whenever} \qquad \mathbf{v} \neq \mathbf{0}, \qquad \text{while} \qquad \langle \mathbf{0}, \mathbf{0} \rangle = 0. \tag{B.9}$$

Given an inner product, the associated *norm* of an element $\mathbf{v} \in V$ is defined as the positive square root of its inner product with itself:

$$\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}\,. \tag{B.10}$$

Bilinearity of the inner product implies that

$$\| c\,\mathbf{v} \| = | c | \, \| v \|$$

for any scalar c. The positivity axiom implies that $\| \mathbf{v} \| \geq 0$ is real and nonnegative, and equals 0 if and only if $\mathbf{v} = \mathbf{0}$ is the zero element. A vector space norm induces a notion of *distance* between elements $\mathbf{v}, \mathbf{w} \in V$, with $\operatorname{dist}(\mathbf{v}, \mathbf{w}) = \| \mathbf{v} - \mathbf{w} \|$. In particular, $\operatorname{dist}(\mathbf{v}, \mathbf{w}) = 0$ if and only if $\mathbf{v} = \mathbf{w}$.

Example B.11. The most familiar example of an inner product is the *dot product*[†]

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w} = v_1\,w_1 + v_2\,w_2 + \ \cdots \ + v_n\,w_n \tag{B.11}$$

on the Euclidean space $\mathbb{R}^n$. The associated *Euclidean norm*

$$\| \mathbf{v} \| \ = \ \sqrt{\mathbf{v} \cdot \mathbf{v}} \ = \ \sqrt{v_1^2 + v_2^2 + \ \cdots \ + v_n^2} \tag{B.12}$$

conforms to our usual notion of distance between points in Euclidean space.

To find the most general inner product on $\mathbb{R}^n$, we need to introduce the important class of positive definite matrices.

Definition B.12. An $n \times n$ matrix C is called *positive definite* if it satisfies the positivity condition

$$\mathbf{v}^T C\,\mathbf{v} > 0 \qquad \text{for all} \qquad \mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n. \tag{B.13}$$

We will sometimes write $C > 0$ to mean that C is a positive definite matrix.

---

† The elements $\mathbf{v} \in \mathbb{R}^n$ are to be regarded as column vectors, while the transpose, written $\mathbf{v}^T$, is the corresponding row vector.

Warning: The condition $C > 0$ does *not* mean that all the entries of C are positive. For example, $\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$ is positive definite, whereas $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is not.

Many authors, including [**89**], require that a positive definite matrix also be symmetric. We will *not* impose this condition here a priori. However, most of the positive definite matrices we will encounter in applications will be symmetric (or, more generally, self-adjoint — as in Example 9.15). For a symmetric matrix, the most useful test for positive definiteness is to perform Gaussian Elimination on the matrix C, which is positive definite if and only if no row interchanges are needed, and all the pivots are positive, [**89**].

Proposition B.13. *Every inner product on* $\mathbb{R}^n$ *is given by*

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T C\, \mathbf{w} \qquad \textit{for} \qquad \mathbf{v}, \mathbf{w} \in \mathbb{R}^n, \tag{B.14}$$

where $C > 0$ *is a symmetric positive definite matrix.*

The next example is of particular significance in Fourier analysis and partial differential equations.

Example B.14. Let $[a, b] \subset \mathbb{R}$ be a bounded closed interval. The integral

$$\langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx \tag{B.15}$$

defines an inner product on the space $\mathrm{C}^0[a, b]$ of continuous functions. The associated norm

$$\| f \| = \sqrt{\int_a^b f(x)^2\, dx} \tag{B.16}$$

is known as the L^2 *norm* of the function f over the interval $[a, b]$. The positivity of the norm: $\| f \| > 0$ for $f \neq 0$, follows from the fact that the only continuous nonnegative function $g(x) \geq 0$ that satisfies $\displaystyle\int_a^b g(x)\, dx = 0$ is the zero function $g(x) \equiv 0$. Extending this construction to spaces containing discontinuous functions is trickier, since there are discontinuous functions that are not identically zero, but nevertheless have zero norm integral. An example is a function that is zero except at a single point. Further discussion can be found in Section 3.5.

The two most important inequalities in mathematical analysis apply to any inner product space.

Theorem B.15. *Every inner product satisfies the* Cauchy–Schwarz *and* triangle inequalities

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \;\leq\; \| \mathbf{v} \|\, \| \mathbf{w} \|, \qquad \| \mathbf{v} + \mathbf{w} \| \;\leq\; \| \mathbf{v} \| + \| \mathbf{w} \|, \qquad \textit{for all} \qquad \mathbf{v}, \mathbf{w} \in V. \tag{B.17}$$

Equality holds if and only if $\mathbf{v}$ *and* $\mathbf{w}$ *are parallel, i.e., scalar multiples of each other.*

Proof: We begin with the Cauchy–Schwarz inequality: $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \| \mathbf{v} \|\, \| \mathbf{w} \|$. The case $\mathbf{w} = \mathbf{0}$ is trivial, and so we assume $\mathbf{w} \neq \mathbf{0}$. Let $t \in \mathbb{R}$ be an arbitrary scalar. Using the three inner product axioms, we have

$$0 \leq \| \mathbf{v} + t\, \mathbf{w} \|^2 = \langle \mathbf{v} + t\, \mathbf{w}, \mathbf{v} + t\, \mathbf{w} \rangle = \| \mathbf{v} \|^2 + 2\, t \, \langle \mathbf{v}, \mathbf{w} \rangle + t^2\, \| \mathbf{w} \|^2, \tag{B.18}$$

with equality holding if and only if $\mathbf{v} = -t\,\mathbf{w}$, which requires $\mathbf{v}$ and $\mathbf{w}$ to be parallel vectors. We fix $\mathbf{v}$ and $\mathbf{w}$, and consider the right-hand side of (B.18) as a quadratic function of t. Its minimum value occurs when $t = \|\,\mathbf{w}\,\|^{-2}\,\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle$. Substituting this value into (B.18), we obtain

$$0 \;\le\; \|\,\mathbf{v}\,\|^2 - 2\,\frac{\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle^2}{\|\,\mathbf{w}\,\|^2} + \frac{\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle^2}{\|\,\mathbf{w}\,\|^2} \;=\; \|\,\mathbf{v}\,\|^2 - \frac{\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle^2}{\|\,\mathbf{w}\,\|^2}\,,$$

and hence $\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle^2 \;\le\; \|\,\mathbf{v}\,\|^2\,\|\,\mathbf{w}\,\|^2$, which, upon taking the square root, establishes the Cauchy–Schwarz inequality. Again, as noted above, equality holds if and only if $\mathbf{v}$ and $\mathbf{w}$ are parallel.

To establish the triangle inequality, we compute

$$\begin{aligned}\|\,\mathbf{v}+\mathbf{w}\,\|^2 = \langle\,\mathbf{v}+\mathbf{w}\,,\mathbf{v}+\mathbf{w}\,\rangle &= \|\,\mathbf{v}\,\|^2 + 2\,\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle + \|\,\mathbf{w}\,\|^2 \\ &\le \|\,\mathbf{v}\,\|^2 + 2\,\|\,\mathbf{v}\,\|\,\|\,\mathbf{w}\,\| + \|\,\mathbf{w}\,\|^2 = \big(\,\|\,\mathbf{v}\,\| + \|\,\mathbf{w}\,\|\,\big)^2,\end{aligned}$$

where the middle inequality follows from the Cauchy–Schwarz inequality (which clearly also holds if the absolute value is removed.) Taking square roots of both sides completes the proof. *Q.E.D.*

We will also have occasion to use inner products on complex vector spaces. To ensure that the associated norm remains positive, the real definition must be modified. The complex conjugate of a complex scalar $c = a + \mathrm{i}\,b$, with $a, b \in \mathbb{R}$, will be indicated by an overbar: $\overline{c} = a - \mathrm{i}\,b$. When dealing with a complex inner product space, one must pay careful attention to complex conjugation.

Definition B.16. An *inner product* on the complex vector space V is a pairing that takes two vectors $\mathbf{v}, \mathbf{w} \in V$ and produces a complex number $\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle \in \mathbb{C}$, subject to the following requirements, for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and $c, d \in \mathbb{C}$:

(*i*) *Sesquilinearity*:

$$\begin{aligned}\langle\,c\,\mathbf{u} + d\,\mathbf{v}\,,\mathbf{w}\,\rangle &= c\,\langle\,\mathbf{u}\,,\mathbf{w}\,\rangle + d\,\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle,\\ \langle\,\mathbf{u}\,,c\,\mathbf{v} + d\,\mathbf{w}\,\rangle &= \overline{c}\,\langle\,\mathbf{u}\,,\mathbf{v}\,\rangle + \overline{d}\,\langle\,\mathbf{u}\,,\mathbf{w}\,\rangle.\end{aligned} \tag{B.19}$$

(*ii*) *Conjugate Symmetry*:

$$\langle\,\mathbf{v}\,,\mathbf{w}\,\rangle = \overline{\langle\,\mathbf{w}\,,\mathbf{v}\,\rangle}. \tag{B.20}$$

(*iii*) *Positivity*:

$$\|\,\mathbf{v}\,\|^2 = \langle\,\mathbf{v}\,,\mathbf{v}\,\rangle \ge 0, \qquad \text{and} \qquad \langle\,\mathbf{v}\,,\mathbf{v}\,\rangle = 0 \quad \text{if and only if} \quad \mathbf{v} = \mathbf{0}. \tag{B.21}$$

Example B.17. The simplest example is the *Hermitian dot product*

$$\mathbf{z}\cdot\mathbf{w} = \mathbf{z}^T\,\overline{\mathbf{w}} = z_1\,\overline{w}_1 + z_2\,\overline{w}_2 + \;\cdots\; + z_n\,\overline{w}_n, \quad \text{for} \quad \mathbf{z} = \begin{pmatrix} z_1\\ z_2\\ \vdots\\ z_n\end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} w_1\\ w_2\\ \vdots\\ w_n\end{pmatrix}, \tag{B.22}$$

between complex vectors $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$.

Example B.18. Let $\mathrm{C}^0[\,-\pi,\pi\,]$ denote the complex vector space consisting of all complex-valued continuous functions $f(x) = u(x) + \mathrm{i}\,v(x)$ depending on the *real* variable $-\pi \le x \le \pi$. The L^2 *Hermitian inner product* on $\mathrm{C}^0[\,-\pi,\pi\,]$ is defined as

$$\langle\,f\,,g\,\rangle = \int_{-\pi}^{\pi} f(x)\,\overline{g(x)}\,dx\,, \tag{B.23}$$

i.e., the integral of f times the complex conjugate of g, with corresponding norm

$$\| f \| = \sqrt{\int_{-\pi}^{\pi} | f(x) |^2 \, dx} = \sqrt{\int_{-\pi}^{\pi} \left[u(x)^2 + v(x)^2 \right] dx} \,. \tag{B.24}$$

Inner products on complex vector spaces also satisfy the Cauchy–Schwarz and triangle inequalities (B.17). The proof is left as an exercise for the reader; see [**89**; Exercise 3.6.46].

B.4 Orthogonality

Definition B.19. Two elements $\mathbf{v}, \mathbf{w} \in V$ of an inner product space V are called *orthogonal* if their inner product vanishes: $\langle \mathbf{v} , \mathbf{w} \rangle = 0$.

For ordinary Euclidean space equipped with the dot product, two vectors are orthogonal if and only if they are perpendicular, i.e., meet at a right angle.

Definition B.20. A basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of an inner product space V is called *orthogonal* if $\langle \mathbf{u}_i , \mathbf{u}_j \rangle = 0$ for all $i \neq j$. The basis is called *orthonormal* if, in addition, each vector has unit length: $\| \mathbf{u}_i \| = 1$, for all $i = 1, \ldots, n$.

For example, the standard basis vectors (B.5) form an orthonormal basis of $\mathbb{R}^n$ with respect to the dot product, but they are not orthonormal for any other inner product thereon.

Theorem B.21. *If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form an orthogonal basis, then the corresponding coordinates of a vector*

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n \qquad \text{are given by} \qquad a_i = \frac{\langle \mathbf{v} , \mathbf{v}_i \rangle}{\| \mathbf{v}_i \|^2} \,. \tag{B.25}$$

Moreover, the vector's norm can be computed using the formula

$$\| \mathbf{v} \|^2 = \sum_{i=1}^{n} a_i^2 \, \| \mathbf{v}_i \|^2 = \sum_{i=1}^{n} \left(\frac{\langle \mathbf{v} , \mathbf{v}_i \rangle}{\| \mathbf{v}_i \|} \right)^2 \,. \tag{B.26}$$

Proof: We compute the inner product of (B.25) with one of the basis vectors. By orthogonality,

$$\langle \mathbf{v} , \mathbf{v}_i \rangle = \left\langle \sum_{j=1}^{n} a_j \mathbf{v}_j \, , \, \mathbf{v}_i \right\rangle = \sum_{j=1}^{n} a_j \langle \mathbf{u}_j , \mathbf{u}_i \rangle = a_i \, \| \mathbf{v}_i \|^2.$$

To prove formula (B.26), we similarly expand

$$\| \mathbf{v} \|^2 = \langle \mathbf{v} , \mathbf{v} \rangle = \sum_{i,j=1}^{n} a_i \, a_j \langle \mathbf{v}_i , \mathbf{v}_j \rangle = \sum_{i=1}^{n} a_i^2 \, \| \mathbf{v}_i \|^2. \qquad \textit{Q.E.D.}$$

In the case of an orthonormal basis, the formulas (B.25–26) simplify to

$$\mathbf{v} = c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n, \quad \text{where} \quad c_i = \langle \mathbf{v} , \mathbf{u}_i \rangle, \qquad \| \mathbf{v} \| = c_1^2 + \cdots + c_n^2. \tag{B.27}$$

Example B.22. A particularly important orthogonal basis is provided by the following vectors lying in $\mathbb{C}^n$:

$$\begin{aligned}\boldsymbol{\omega}_k &= \left(1, \zeta^k, \zeta^{2k}, \zeta^{3k}, \ldots, \zeta^{(n-1)k}\right)^T \\ &= \left(1, e^{2k\pi\,\mathrm{i}/n}, e^{4k\pi\,\mathrm{i}/n}, \ldots, e^{2(n-1)k\pi\,\mathrm{i}/n}\right)^T,\end{aligned} \qquad k = 0, \ldots, n-1, \tag{B.28}$$

where

$$\zeta = e^{2\pi\,\mathrm{i}/n}. \tag{B.29}$$

Orthogonality relies on the fact that its powers, $\zeta^k = e^{2k\pi\,\mathrm{i}/n}$, $k = 0, \ldots, n-1$, are the complex roots of the elementary polynomial

$$z^n - 1 = (z-1)(1 + z + z^2 + \cdots + z^{n-1}), \tag{B.30}$$

while

$$\overline{\zeta} = e^{-2\pi\,\mathrm{i}/n} = \zeta^{-1}.$$

Since when $0 < k \leq n-1$, the complex number $\zeta^k \neq 1$ is a root of the polynomial (B.30), it must also be a root of the second factor. This implies that

$$1 + \zeta^k + \zeta^{2k} + \zeta^{3k} + \cdots + \zeta^{(n-1)k} = \begin{cases} n, & k \equiv 0 \mod n, \\ 0, & k \not\equiv 0 \mod n, \end{cases}$$

where the former case $k \equiv 0 \mod n$ follows by direct substitution of $\zeta^k = 1$. Thus, the Hermitian inner products of the vectors (B.28) equal

$$\langle \boldsymbol{\omega}_k, \boldsymbol{\omega}_l \rangle = \sum_{j=0}^{n-1} \zeta^{jk}\,\overline{\zeta}^{jl} = \sum_{j=0}^{n-1} \zeta^{j(k-l)} = \begin{cases} n, & k = l, \\ 0, & k \neq l, \end{cases} \tag{B.31}$$

provided $0 \leq k, l \leq n-1$, thereby establishing orthogonality. These vectors are the discrete analogues of the orthogonal complex exponential functions that are used to construct complex Fourier series. They are the basis of the discrete Fourier transform, [**89**; §5.7], and their orthogonality is the key to modern signal processing.

B.5 Eigenvalues and Eigenvectors

The eigenvalues and eigenvectors of a matrix first appear when solving linear systems of ordinary differential equations. But their essential importance extends across all of mathematics and its manifold applications. Extensions of the eigenvalue method to linear operators on function spaces are critical to the analysis of partial differential equations.

Definition B.23. Let A be an $n \times n$ matrix. A scalar λ is called an *eigenvalue* of A if there is a *nonzero* vector $\mathbf{v} \neq \mathbf{0}$, called an associated *eigenvector*, such that

$$A\mathbf{v} = \lambda\mathbf{v}. \tag{B.32}$$

In particular, a matrix has $\lambda = 0$ as an eigenvalue if and only if it has a *null eigenvector* $\mathbf{v} \neq \mathbf{0}$, satisfying $A\mathbf{v} = \mathbf{0}$, and hence is a singular (non-invertible) matrix, with vanishing determinant: $\det A = 0$. An eigenvalue is called *simple* if it admits only one linearly independent eigenvalue; more generally, the *multiplicity* of an eigenvalue is defined as the dimension of the *eigenspace* consisting of all solutions to the eigenequation (B.32), including $\mathbf{0}$. Thus, a simple eigenvalue has multiplicity 1.

Even if A is a real matrix, we must allow the possibility of complex eigenvectors. Matrices with a "complete" set of eigenvectors are the most common, and also the easiest to deal with.

Definition B.24. An $n \times n$ real or complex matrix A is called *complete* if there exists a basis of $\mathbb{C}^n$ consisting of its (complex) eigenvectors.

It is not hard to show that eigenvectors corresponding to different eigenvalues are necessarily linearly independent. This means that matrices with all distinct (and hence simple) eigenvalues are necessarily complete:

Proposition B.25. *Any $n \times n$ matrix with n distinct eigenvalues is complete.*

Unfortunately, not all matrices with repeated eigenvalues are complete. For instance, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is complete, since, for instance, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ form an eigenvector basis of $\mathbb{C}^2$, whereas $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not, since it has only one independent eigenvector, namely $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Incomplete matrices are much more challenging to deal with, both theoretically and numerically. Fortunately, we can safely ignore the incomplete cases in this text.

The most common way for orthogonal bases to arise is as eigenvector bases of symmetric matrices. (Orthogonality is with respect to the standard dot product on $\mathbb{R}^n$.) The extension of this result to "self-adjoint" operators on function space forms the foundation of Fourier analysis and its generalizations.

Theorem B.26. *Let $A = A^T$ be a real symmetric $n \times n$ matrix. Then*

(a) *All the eigenvalues of A are real.*

(b) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*

(c) *There is an orthonormal basis of $\mathbb{R}^n$ consisting of n eigenvectors of A.*

Let us demonstrate orthogonality, leaving the remaining steps in the proof to [**89**; Theorem 8.20]. If

$$A\mathbf{v} = \lambda\mathbf{v}, \qquad A\mathbf{w} = \mu\,\mathbf{w},$$

where $\lambda \neq \mu$ are distinct real eigenvalues, then, by symmetry of A,

$$\lambda\,\mathbf{v}\cdot\mathbf{w} = (A\mathbf{v})\cdot\mathbf{w} = (A\mathbf{v})^T\mathbf{w} = \mathbf{v}^T A\mathbf{w} = \mathbf{v}\cdot(A\mathbf{w}) = \mathbf{v}\cdot(\mu\,\mathbf{w}) = \mu\,\mathbf{v}\cdot\mathbf{w},$$

and hence

$$(\lambda - \mu)\,\mathbf{v}\cdot\mathbf{w} = 0.$$

Since $\lambda \neq \mu$, this implies that the eigenvectors $\mathbf{v}, \mathbf{w}$ are necessarily orthogonal.

B.6 Linear Iteration

For numerical applications, we will require some basic results on iteration of linear systems. Consider first a *homogeneous linear iterative system* of the form

$$\mathbf{u}^{(k+1)} = A\,\mathbf{u}^{(k)}, \qquad \mathbf{u}^{(0)} = \mathbf{u}_0, \tag{B.33}$$

in which A is an $n \times n$ matrix and $\mathbf{u}_0 \in \mathbb{R}^n$ or $\mathbb{C}^n$. The solution to such a system is evidently obtained by repeatedly multiplying the initial vector $\mathbf{u}_0$ by the matrix A, and so

$$\mathbf{u}^{(k)} = A^k \mathbf{u}_0. \tag{B.34}$$

Definition B.27. A matrix A is called *convergent* if every solution to the homogeneous linear iterative system (B.33) tends to zero in the limit: $\mathbf{u}^{(k)} \to \mathbf{0}$ as $k \to \infty$. Equivalently, A is convergent if and only if its powers converge to the zero matrix: $A^k \to \mathrm{O}$ as $k \to \infty$.

The solution formula (B.34), while elementary, is not particularly enlightening. An alternative approach is to recognize that if λ_j is an eigenvalue of A and $\mathbf{v}_j$ a corresponding eigenvector, then

$$\mathbf{u}_j^{(k)} = \lambda_j^k \mathbf{v}_j \tag{B.35}$$

is a solution, since

$$A\mathbf{u}_j^{(k)} = \lambda_j^k A\mathbf{v}_j = \lambda_j^{k+1} \mathbf{v}_j = \mathbf{u}_j^{(k+1)}.$$

Moreover, linear combinations of such *eigensolutions* are also solutions. In particular, if A is complete, then we can write down the general solution to (B.33) as a linear combination of the independent eigensolutions:

$$\mathbf{u}^{(k)} = c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + \cdots + c_n \lambda_n^k \mathbf{v}_n, \tag{B.36}$$

where $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is the eigenvector basis. The coefficients $c_1, \ldots, c_n$ are uniquely determined by the initial conditions,

$$\mathbf{u}^{(0)} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n = \mathbf{u}_0,$$

which relies on the fact that the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis. Now, A is convergent if and only if all solutions $\mathbf{u}^{(k)} \to \mathbf{0}$. The individual eigensolution (B.35) goes to zero if and only if its associated eigenvalue is strictly less than 1 in modulus: $|\lambda_j| < 1$. This proves the following result for complete matrices. The proof in the incomplete case relies on the Jordan canonical form, [**89**; Chapter 10].

Theorem B.28. *The matrix A is convergent if and only if all its eigenvalues satisfy $|\lambda| < 1$.*

Definition B.29. The *spectral radius* of a matrix A is defined as the maximal modulus of all of its real and complex eigenvalues: $\rho(A) = \max\{|\lambda_1|, \ldots, |\lambda_k|\}$.

Corollary B.30. *The matrix A is convergent if and only if $\rho(A) < 1$.*

Indeed, the spectral radius essentially governs the rate of convergence of the iterative system — the closer it is to 0, the faster the convergence rate.

Next, consider the *inhomogeneous linear iterative system*

$$\mathbf{v}^{(k+1)} = A\mathbf{v}^{(k)} + \mathbf{b}, \qquad \mathbf{v}^{(0)} = \mathbf{v}_0, \tag{B.37}$$

where $\mathbf{b}$ a fixed vector. A *fixed point* is a vector $\mathbf{v}^\star$ that satisfies

$$\mathbf{v}^\star = A\mathbf{v}^\star + \mathbf{b}, \qquad \text{or, equivalently,} \qquad (\mathrm{I} - A)\mathbf{v}^\star = \mathbf{b}, \tag{B.38}$$

where I is the *identity matrix* of the same size as A. Thus, if 1 is not an eigenvalue of A (which cannot happen when A is convergent), then $\mathrm{I} - A$ is nonsingular, and so the iterative system has a unique fixed point.

Theorem B.31. *Assume that 1 is not an eigenvalue of A. Then all solutions to* (B.37) *converge to the fixed point,* $\mathbf{v}^{(k)} \to \mathbf{v}^\star$ *as* $k \to \infty$ *if and only if A is a convergent matrix.*

Proof: Let $\mathbf{u}^{(k)} = \mathbf{v}^{(k)} - \mathbf{v}^\star$, so that $\mathbf{v}^{(k)} \to \mathbf{v}^\star$ if and only if $\mathbf{u}^{(k)} \to \mathbf{0}$. Now,

$$\mathbf{u}^{(k+1)} = \mathbf{v}^{(k+1)} - \mathbf{v}^\star = (A\mathbf{v}^{(k)} + \mathbf{b}) - (A\mathbf{v}^\star + \mathbf{b}) = A(\mathbf{v}^{(k)} - \mathbf{v}^\star) = A\mathbf{u}^{(k)},$$

and hence $\mathbf{u}^{(k)}$ solves the homogeneous version (B.33). Thus, the result is an immediate consequence of Definition B.27. *Q.E.D.*

B.7 Linear Functions and Systems

The most basic structural features of linear differential equations, both ordinary and partial, linear boundary value problems, etc., are founded on the concept of a linear function between vector spaces.

Definition B.32. Let U and V be real vector spaces. A function $L: U \to V$ is called *linear* if it obeys two basic rules:

$$L[\mathbf{u} + \mathbf{v}] = L[\mathbf{u}] + L[\mathbf{v}], \qquad\qquad L[c\,\mathbf{u}] = c\,L[\mathbf{u}], \tag{B.39}$$

for all $\mathbf{u}, \mathbf{v} \in U$ and all scalars c.

We will refer to U as the *domain space* of the function L, and V as the *target space*. The latter is to emphasize the fact that the *range* of L, namely

$$\operatorname{rng} L = \{\, \mathbf{v} \in V \mid \mathbf{v} = L[\mathbf{u}] \text{ for some } \mathbf{u} \in U \,\}, \tag{B.40}$$

may very well be a proper subspace of the target space V.

Theorem B.33. *Every linear function $L: \mathbb{R}^n \to \mathbb{R}^m$ is given by matrix multiplication, $L[\mathbf{v}] = A\mathbf{v}$, where A is an $m \times n$ matrix.*

Proving that matrix multiplication satisfies the linearity conditions (B.39) is easy. The converse is established by seeing what the linear function does to the basis vectors of $\mathbb{R}^n$; see [**89**; Theorem 7.5].

Corollary B.34. *Every linear function $L: \mathbb{R}^n \to \mathbb{R}$ is given by taking the dot product with a fixed vector $\mathbf{a} \in \mathbb{R}^n$:*

$$L[\mathbf{v}] = \mathbf{a} \cdot \mathbf{v}. \tag{B.41}$$

When U is a function space, a linear function is also referred to as a *linear operator* in order to avoid confusion with the elements of U. If the target space $V = \mathbb{R}$, then the term *linear functional* is also often used for $L: U \to \mathbb{R}$.

Here are some representative examples that arise in applications.

Example B.35. (*a*) Evaluation of a function at a point, namely $L[f] = f(x_0)$, defines a linear operator $L: \mathrm{C}^0[a,b] \to \mathbb{R}$.

(*b*) Integration,

$$I[f] = \int_a^b f(x)\,dx, \tag{B.42}$$

also defines a linear functional $I: \mathrm{C}^0[a,b] \to \mathbb{R}$.

(*c*) The operation $M_a[f(x)] = a(x)\, f(x)$ of multiplication by a continuous function a defines a linear operator $M_a\colon \mathrm{C}^0[a,b] \to \mathrm{C}^0[a,b]$.

(*d*) Differentiation of functions, $D[f] = f'$, serves to define a linear operator $D\colon \mathrm{C}^1[a,b] \to \mathrm{C}^0[a,b]$.

(*e*) A general *linear ordinary differential operator* of order n,

$$L = a_n(x)\, D^n + a_{n-1}(x)\, D^{n-1} + \cdots + a_1(x)\, D + a_0(x), \tag{B.43}$$

is obtained by summing such operators. If the coefficient functions $a_0(x), \ldots, a_n(x)$ are continuous, then

$$L[u] = a_n(x)\, \frac{d^n u}{dx^n} + a_{n-1}(x)\, \frac{d^{n-1} u}{dx^{n-1}} + \cdots + a_1(x)\, \frac{du}{dx} + a_0(x) u \tag{B.44}$$

defines a linear operator from $\mathrm{C}^n[a,b]$ to $\mathrm{C}^0[a,b]$.

Linear partial differential equations are based on linear partial differential operators, which are discussed in Chapter 1. They are particular examples of the general concept of a linear system.

Definition B.36. A *linear system* is an equation of the form

$$L[\mathbf{u}] = \mathbf{f}, \tag{B.45}$$

in which $L\colon U \to V$ is a linear function, $\mathbf{f} \in V$, while the desired solution $\mathbf{u} \in U$. The system is *homogeneous* if $\mathbf{f} = \mathbf{0}$; otherwise, it is called *inhomogeneous*.

Note that, by the definition (B.40) of the range of L, the linear system (B.45) will have a solution if and only if $\mathbf{f} \in \operatorname{rng} L$. In particular, a homogeneous linear system always has a solution, namely $\mathbf{u} = \mathbf{0}$. However, it may possibly admit other, nonzero, solutions.

Theorem B.37. *If $\mathbf{z}_1, \ldots, \mathbf{z}_k$ are all solutions to the same homogeneous linear system*

$$L[\mathbf{z}] = \mathbf{0}, \tag{B.46}$$

then any linear combination $c_1 \mathbf{z}_1 + \cdots + c_k \mathbf{z}_k$, for any scalars $c_1, \ldots, c_k$, is also a solution.

In other words, the set of solutions to a homogeneous linear system (B.46) forms a subspace of the domain space U, known as the *kernel* of the linear function L:

$$\ker L = \{\, \mathbf{z} \in U \mid L[\mathbf{z}] = \mathbf{0} \,\}. \tag{B.47}$$

Theorem B.38. *If the inhomogeneous linear system $L[\mathbf{u}] = \mathbf{f}$ has a particular solution $\mathbf{u}^\star$, which requires $\mathbf{f} \in \operatorname{rng} L$, then the general solution is $\mathbf{u} = \mathbf{u}^\star + \mathbf{z}$, where $\mathbf{z} \in \ker L$ is any solution to the corresponding homogeneous system $L[\mathbf{z}] = \mathbf{0}$.*

The *Superposition Principle* for inhomogeneous linear systems allows us to combine solutions corresponding to different right-hand sides.

Theorem B.39. *Suppose that for each $i = 1, \ldots, k$, we know a particular solution $\mathbf{u}_i^\star$ to the inhomogeneous linear system $L[\mathbf{u}] = \mathbf{f}_i$ for some $\mathbf{f}_i \in \operatorname{rng} L$. Then, given scalars $c_1, \ldots, c_k$, a particular solution to the combined inhomogeneous system*

$$L[\mathbf{u}] = c_1 \mathbf{f}_1 + \cdots + c_k \mathbf{f}_k \tag{B.48}$$

is the corresponding linear combination

$$\mathbf{u}^\star = c_1\,\mathbf{u}_1^\star + \cdots + c_k\,\mathbf{u}_k^\star \tag{B.49}$$

of particular solutions. The general solution to the inhomogeneous system (B.48) *is*

$$\mathbf{u} = \mathbf{u}^\star + \mathbf{z} = c_1\,\mathbf{u}_1^\star + \ \cdots\ + c_k\,\mathbf{u}_k^\star + \mathbf{z}, \tag{B.50}$$

where $\mathbf{z} \in \ker L$ *is an arbitrary solution to the associated homogeneous system* $L[\,\mathbf{z}\,] = \mathbf{0}$.

References

[**1**] Abdulloev, K. O., Bogolubsky, I. L., and Makhankov, V. G., One more example of inelastic soliton interaction, *Phys. Lett. A* **56** (1976), 427–428.

[**2**] Ablowitz, M. J., and Clarkson, P. A., *Solitons, Nonlinear Evolution Equations and the Inverse Scattering Transform*, L.M.S. Lecture Notes in Math., vol. 149, Cambridge University Press, Cambridge, 1991.

[**3**] Abraham, R., Marsden, J. E., and Ratiu, T., *Manifolds, Tensor Analysis, and Applications*, Springer–Verlag, New York, 1988.

[**4**] Airy, G. B., On the intensity of light in the neighborhood of a caustic, *Trans. Cambridge Phil. Soc.* **6** (1838), 379–402.

[**5**] Aki, K., and Richards, P. G., *Quantitative Seismology*, W. H. Freeman, San Francisco, 1980.

[**6**] Ames, W. F., *Numerical Methods for Partial Differential Equations*, 3rd ed., Academic Press, New York, 1992.

[**7**] Antman, S. S., *Nonlinear Problems of Elasticity*, Appl. Math. Sci., vol. 107, Springer–Verlag, New York, 1995.

[**8**] Apostol, T. M., *Calculus*, Blaisdell Publishing Co., Waltham, Mass., 1967–1969.

[**9**] Apostol, T. M., *Introduction to Analytic Number Theory*, Springer–Verlag, New York, 1976.

[**10**] Atkinson, K., and Han, W., *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*, Lecture Notes in Math., vol. 2044, Springer, Berlin, 2012.

[**11**] Bank, S. B., and Kaufman, R. P., A note on Hölder's theorem concerning the gamma function, *Math. Ann.* **232** (1978), 115–120.

[**12**] Batchelor, G. K., *An Introduction to Fluid Dynamics*, Cambridge University Press, Cambridge, 1967.

[**13**] Bateman, H., Some recent researches on the motion of fluids, *Monthly Weather Rev.* **43** (1915), 63–170.

[**14**] Benjamin, T. B., Bona, J. L., and Mahony, J. J., Model equations for long waves in nonlinear dispersive systems, *Phil. Trans. Roy. Soc. London A* **272** (1972), 47–78.

[**15**] Berest, Y., and Winternitz, P., Huygens' principle and separation of variables, *Rev. Math. Phys.* **12** (2000), 159–180.

[**16**] Berry, M. V., Marzoli, I., and Schleich, W., Quantum carpets, carpets of light, *Physics World* **14**(6) (2001), 39–44.

[**17**] Birkhoff, G., *Hydrodynamics — A Study in Logic, Fact and Similitude*, 2nd ed., Princeton University Press, Princeton, 1960.

[**18**] Birkhoff, G., and Rota, G.–C., *Ordinary Differential Equations*, Blaisdell Publ. Co., Waltham, Mass., 1962.

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0,

[**19**] Black, F., and Scholes, M., The pricing of options and corporate liabilities, *J. Political Economy* **81** (1973), 637–654.

[**20**] Blanchard, P., Devaney, R. L., and Hall, G. R., *Differential Equations*, Brooks–Cole Publ. Co., Pacific Grove, Calif., 1998.

[**21**] Boussinesq, J., Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond, *J. Math. Pures Appl.* **17** (2) (1872), 55–108.

[**22**] Boussinesq, J., Essai sur la théorie des eaux courants, *Mém. Acad. Sci. Inst. Nat. France* **23** (1) (1877), 1–680.

[**23**] Boyce, W. E., and DiPrima, R. C., *Elementary Differential Equations and Boundary Value Problems*, 7th ed., John Wiley & Sons, Inc., New York, 2001.

[**24**] Bradie, B., *A Friendly Introduction to Numerical Analysis*, Prentice–Hall, Inc., Upper Saddle River, N.J., 2006.

[**25**] Bronstein, M., *Symbolic integration I: Transcendental Functions*, Springer–Verlag, New York, 1997.

[**26**] Burgers, J. M., A mathematical model illustrating the theory of turbulence, *Adv. Appl. Mech.* **1** (1948), 171–199.

[**27**] Cantwell, B. J., *Introduction to Symmetry Analysis*, Cambridge University Press, Cambridge, 2003.

[**28**] Carleson, L., On the convergence and growth of partial sums of Fourier series, *Acta Math.* **116** (1966), 135–157.

[**29**] Carmichael, R., *The Theory of Numbers*, Dover Publ., New York, 1959.

[**30**] Chen, G., and Olver, P. J., Dispersion of discontinuous periodic waves, *Proc. Roy. Soc. London* **469** (2012), 20120407.

[**31**] Coddington, E. A., and Levinson, N., *Theory of Ordinary Differential Equations*, McGraw–Hill, New York, 1955.

[**32**] Cole, J. D., On a quasilinear parabolic equation occurring in aerodynamics, *Q. Appl. Math.* **9** (1951), 225–236.

[**33**] Courant, R., Friedrichs, K. O., and Lewy, H., Über die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.* **100** (1928), 32–74.

[**34**] Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, vol. I, Interscience Publ., New York, 1953.

[**35**] Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, vol. II, Interscience Publ., New York, 1953.

[**36**] Drazin, P. G., and Johnson, R. S., *Solitons: An Introduction*, Cambridge University Press, Cambridge, 1989.

[**37**] Dym, H., and McKean, H. P., *Fourier Series and Integrals*, Academic Press, New York, 1972.

[**38**] Evans, L. C., *Partial Differential Equations*, Grad. Studies Math. vol. 19, Amer. Math. Soc., Providence, R.I., 1998.

[**39**] Feller, W., *An Introduction to Probability Theory and Its Applications*, 3rd ed., J. Wiley & Sons, New York, 1968.

[**40**] Fermi, E., Pasta, J., and Ulam, S., Studies of nonlinear problems. I., preprint, Los Alamos Report LA 1940, 1955; in: *Nonlinear Wave Motion*, A. C. Newell, ed., Lectures in Applied Math., vol. 15, American Math. Soc., Providence, R.I., 1974, pp. 143–156.

[**41**] Forsyth, A. R., *The Theory of Differential Equations*, Cambridge University Press, Cambridge, 1890, 1900, 1902, 1906.

[**42**] Fourier, J., *The Analytical Theory of Heat*, Dover Publ., New York, 1955.

[**43**] Gander, M. J., and Kwok, F., Chladni figures and the Tacoma bridge: motivating PDE eigenvalue problems via vibrating plates, *SIAM Review* **54** (2012), 573–596.

[**44**] Garabedian, P., *Partial Differential Equations*, 2nd ed., Chelsea Publ. Co., New York, 1986.

[**45**] Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M., Method for solving the Korteweg–deVries equation, *Phys. Rev. Lett.* **19** (1967), 1095–1097.

[**46**] Gonzalez, R. C., and Woods, R. E., *Digital Image Processing*, 2nd ed., Prentice–Hall, Inc., Upper Saddle River, N.J., 2002.

[**47**] Gordon, C., Webb, D. L., and Wolpert, S., One cannot hear the shape of a drum, *Bull. Amer. Math. Soc.* **27** (1992), 134–138.

[**48**] Gradshteyn, I. S., and Ryzhik, I. W., *Table of Integrals, Series and Products*, Academic Press, New York, 1965.

[**49**] Gurtin, M. E., *An Introduction to Continuum Mechanics*, Academic Press, New York, 1981.

[**50**] Haberman, R., *Elementary Applied Partial Differential Equations*, 3rd ed., Prentice–Hall, Inc., Upper Saddle River, NJ, 1998.

[**51**] Hairer, E., Lubich, C., and Wanner, G., *Geometric Numerical Integration*, Springer–Verlag, New York, 2002.

[**52**] Hale, J. K., *Ordinary Differential Equations*, 2nd ed., R.E. Krieger Pub. Co., Huntington, N.Y., 1980.

[**53**] Henrici, P., *Applied and Computational Complex Analysis*, vol. 1, J. Wiley & Sons, New York, 1974.

[**54**] Hille, E., *Ordinary Differential Equations in the Complex Domain*, John Wiley & Sons, New York, 1976.

[**55**] Hobson, E. W., *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Dover Publ., New York, 1957.

[**56**] Hopf, E., The partial differential equation $u_t + u u_x = \mu u$, *Commun. Pure Appl. Math.* **3** (1950), 201–230.

[**57**] Howison, S., *Practical Applied Mathematics*: *Modelling, Analysis, Approximation*, Cambridge University Press, Cambridge, 2005.

[**58**] Hydon, P. E., *Symmetry Methods for Differential Equations*, Cambridge Texts in Appl. Math., Cambridge University Press, Cambridge, 2000.

[**59**] Ince, E. L., *Ordinary Differential Equations*, Dover Publ., New York, 1956.

[**60**] Iserles, A., *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge, 1996.

[**61**] Jost, J., *Partial Differential Equations*, Graduate Texts in Mathematics, vol. 214, Springer–Verlag, New York, 2007.

[**62**] Kamke, E., *Differentialgleichungen Lösungsmethoden und Lösungen*, vol. 1, Chelsea, New York, 1971.

[**63**] Keller, H. B., *Numerical Methods for Two-Point Boundary-Value Problems*, Blaisdell, Waltham, MA, 1968.

[**64**] Knobel, R., *An Introduction to the Mathematical Theory of Waves*, American Mathematical Society, Providence, RI, 2000.

[**65**] Korteweg, D. J., and de Vries, G., On the change of form of long waves advancing in a rectangular channel, and on a new type of long stationary waves, *Phil. Mag.* (5) **39** (1895), 422–443.

[**66**] Landau, L. D., and Lifshitz, E. M., *Quantum Mechanics* (*Non-relativistic Theory*), Course of Theoretical Physics, vol. 3, Pergamon Press, New York, 1977.

[**67**] Levine, I. N., *Quantum Chemistry*, 5th ed., Prentice–Hall, Inc., Upper Saddle River, N.J., 2000.

[**68**] Lighthill, M. J., *Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press, Cambridge, 1970.

[**69**] Lin, C. C., and Segel, L. A., *Mathematics Applied to Deterministic Problems in the Natural Sciences*, SIAM, Philadelphia, 1988.

[**70**] McOwen, R. C., *Partial Differential Equations*: *Methods and Applications*, Prentice–Hall, Inc., Upper Saddle River, N.J., 2002.

[**71**] Merton, R. C., Theory of rational option pricing, *Bell J. Econ. Management Sci.* **4** (1973), 141–183.

[**72**] Messiah, A., *Quantum Mechanics*, John Wiley & Sons, New York, 1976.

[**73**] Miller, W., Jr., *Symmetry and Separation of Variables*, Encyclopedia of Mathematics and Its Applications, vol. 4, Addison–Wesley Publ. Co., Reading, Mass., 1977.

[**74**] Milne–Thompson, L. M., *The Calculus of Finite Differences*, Macmillan and Co., Ltd., London, 1951.

[**75**] Misner, C. W., Thorne, K. S., and Wheeler, J. A., *Gravitation*, W. H. Freeman, San Francisco, 1973.

[**76**] Miura, R. M., Gardner, C. S., and Kruskal, M. D., Korteweg–deVries equation and generalizations. II. Existence of conservation laws and constants of the motion, *J. Math. Phys.* **9** (1968), 1204–1209.

[**77**] Moon, F. C., *Chaotic Vibrations*, John Wiley & Sons, New York, 1987.

[**78**] Moon, P., and Spencer, D. E., *Field Theory Handbook*, Springer-Verlag, New York, 1971.

[**79**] Morse, P. M., and Feshbach, H., *Methods of Theoretical Physics*, McGraw–Hill, New York, 1953.

[**80**] Morton, K. W., and Mayers, D. F., *Numerical Solution of Partial Differential Equations*, 2nd ed., Cambridge University Press, Cambridge, 2005.

[**81**] Murray, J. D., *Mathematical Biology*, 3rd ed., Springer-Verlag, New York, 2002–2003.

[**82**] Oberhettinger, F., *Tables of Fourier Transforms and Fourier Transforms of Distributions*, Springer-Verlag, New York, 1990.

[**83**] Øksendal, B., *Stochastic Differential Equations*: *An Introduction with Applications*, Springer–Verlag, New York, 1985.

[**84**] Okubo, A., *Diffusion and Ecological Problems*: *Mathematical Models*, Springer-Verlag, New York, 1980.

[**85**] Olver, F. W. J., *Asymptotics and Special Functions*, Academic Press, New York, 1974.

[**86**] Olver, F. W. J., Lozier, D. W., Boisvert, R. F., and Clark, C. W., eds., *NIST Handbook of Mathematical Functions*, Cambridge University Press, Cambridge, 2010.

[**87**] Olver, P. J., *Applications of Lie Groups to Differential Equations*, 2nd ed., Graduate Texts in Mathematics, vol. 107, Springer–Verlag, New York, 1993.

[**88**] Olver, P. J., Dispersive quantization, *Amer. Math. Monthly* **117** (2010), 599–610.

[**89**] Olver, P. J., and Shakiban, C., *Applied Linear Algebra*, Prentice–Hall, Inc., Upper Saddle River, N.J., 2005.

[**90**] Oskolkov, K. I., A class of I. M. Vinogradov's series and its applications in harmonic analysis, *in*: *Progress in Approximation Theory* , Springer Ser. Comput. Math., 19, Springer, New York, 1992, pp. 353–402.

[**91**] Pinchover, Y., and Rubinstein, J., *An Introduction to Partial Differential Equations*, Cambridge University Press, Cambridge, 2005.

[**92**] Pinsky, M. A., *Partial Differential Equations and Boundary–Value Problems with Applications*, 3rd ed., McGraw–Hill, New York, 1998.

[**93**] Polyanin, A. D., and Zaitsev, V. F., *Handbook of Exact Solutions for Ordinary Differential Equations*, 2nd ed., Chapman & Hall/CRC, Boca Raton, Fl., 2003.

[**94**] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes*: *The Art of Scientific Computing*, 3rd ed., Cambridge University Press, Cambridge, 2007.

[**95**] Reed, M., and Simon, B., *Methods of Modern Mathematical Physics*, Academic Press, New York, 1972.

[**96**] Royden, H. L., and Fitzpatrick, P. M., *Real Analysis*, 4th ed., Pearson Education Inc., Boston, MA, 2010.

[**97**] Rudin, W., *Principles of Mathematical Analysis*, 3rd ed., McGraw–Hill, New York, 1976.

[**98**] Rudin, W., *Real and Complex Analysis*, 3rd ed., McGraw–Hill, New York, 1987.

[**99**] Salsa, S., *Partial Differential Equations in Action*: *From Modelling to Theory*, Springer–Verlag, New York, 2008.

[**100**] Sapiro, G., *Geometric Partial Differential Equations and Image Analysis*, Cambridge University Press, Cambridge, 2001.

[**101**] Schrödinger, E., *Collected Papers on Wave Mechanics*, Chelsea Publ. Co., New York, 1982.

[**102**] Schumaker, L. L., *Spline Functions*: *Basic Theory*, John Wiley & Sons, New York, 1981.

[**103**] Schwartz, L., *Théorie des distributions*, Hermann, Paris, 1957.

[**104**] Scott Russell, J., On waves, *in*: *Report of the 14^{th} Meeting*, British Assoc. Adv. Sci., 1845, pp. 311–390.

[**105**] Sethares, W. A., *Tuning, Timbre, Spectrum, Scale*, Springer–Verlag, New York, 1999.

[**106**] Siegel, C. L., Über einige Anwendungen diophantischer Approximationen, *in*: *Gesammelte Abhandlungen*, vol. 1, Springer–Verlag, New York, 1966, pp. 209–266.

[**107**] Smoller, J., *Shock Waves and Reaction–Diffusion Equations*, 2nd ed., Springer-Verlag, New York, 1994.

[**108**] Stewart, J., *Calculus*: *Early Transcendentals*, vols. 1 & 2, 7th ed., Cengage Learning, Mason, OH, 2012.

[**109**] Stokes, G. G., On a difficulty in the theory of sound, *Phil. Mag.* **33**(3) (1848), 349–356.

[**110**] Stokes, G. G., *Mathematical and Physical Papers*, Cambridge University Press, Cambridge, 1880–1905.

[**111**] Stokes, G. G., *Mathematical and Physical Papers*, 2nd ed., Johnson Reprint Corp., New York, 1966.

[**112**] Strang, G., *Introduction to Applied Mathematics*, Wellesley Cambridge Press, Wellesley, Mass., 1986.

[**113**] Strang, G., and Fix, G. J., *An Analysis of the Finite Element Method*, Prentice–Hall, Inc., Englewood Cliffs, N.J., 1973.

[**114**] Strauss, W. A., *Partial Differential Equations*: *An Introduction*, John Wiley & Sons, New York, 1992.

[**115**] Thaller, B., *Visual Quantum Mechanics*, Springer–Verlag, New York, 2000.

[**116**] Thijssen, J., *Computational Physics*, Cambridge University Press, Cambridge, 1999.

[**117**] Titchmarsh, E. C., *Theory of Functions*, Oxford University Press, London, 1968.

[**118**] Varga, R. S., *Matrix Iterative Analysis*, 2nd ed., Springer–Verlag, New York, 2000.

[**119**] Watson, G. N., *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1952.

[**120**] Weinberger, H. F., *A First Course in Partial Differential Equations*, Dover Publ., New York, 1995.

[**121**] Wiener, N., *I Am a Mathematician*, Doubleday, Garden City, N.Y., 1956.

[**122**] Whitham, G. B., *Linear and Nonlinear Waves*, John Wiley & Sons, New York, 1974.

[**123**] Wilmott, P., Howison, S., and Dewynne, J., *The Mathematics of Financial Derivatives*, Cambridge University Press, Cambridge, 1995.

[**124**] Yong, D., Strings, chains, and ropes, *SIAM Review* **48** (2006), 771–781.

[**125**] Zabusky, N. J., and Kruskal, M. D., Interaction of "solitons" in a collisionless plasma and the recurrence of initial states, *Phys. Rev. Lett.* **15** (1965), 240–243.

[**126**] Zienkiewicz, O. C., and Taylor, R. L., *The Finite Element Method*, 4th ed., McGraw–Hill, New York, 1989.

[**127**] Zwillinger, D., *Handbook of Differential Equations*, Academic Press, Boston, 1992.

[**128**] Zygmund, A., *Trigonometric Series*, 3rd ed., Cambridge University Press, Cambridge, 2002.

Symbol Index

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0, © Springer International Publishing Switzerland 2014

Author Index

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0, © Springer International Publishing Switzerland 2014

Subject Index

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0, © Springer International Publishing Switzerland 2014