I

Concepts from Complex Vector Analysis and the Fourier Transform

In this chapter we present the basic properties of complex vector spaces and the Fourier transform. Sections 1.1 and 1.2 prepare the subject through the standard definitions of linear independence, bases, coordinates, inner product, and norm. In Section 1.3 we introduce linear transformations in vector spaces, emphasizing the conceptual difference between passive and active ones: the former refer to changes in reference coordinates, while the latter imply a "physical" process actually transforming the points of the space. Permutations of reference axes and the Fourier transformation are prime examples of coordinate changes (Section 1.4), while the second-difference operator in particular and self-adjoint operators in general (Section 1.5) will be important in applications. We give, in Section 1.6, the elements of invariance group considerations for a finite N-point lattice. Finally, in Section 1.7 we examine the axes of a transformation and develop the properties of self-adjoint and unitary operators.

If the reader so wishes, he can proceed from Section 1.4 directly to Chapter 3 for applications in communication and the fast Fourier transform algorithm. The rest of the sections are needed, however, for the treatment of coupled systems in Chapter 2.

1.1. N-Dimensional Complex Vector Spaces

The elements of real vector analysis are surely familiar to the reader, so the material in this section will serve mainly to fix notation and to enlarge slightly the concepts of this analysis to the field $\mathscr{C}$ of complex numbers.

1.1.1. Axioms

Let $c_1, c_2, \ldots$ be complex numbers, elements of $\mathscr{C}$, and let $\mathbf{f}_1, \mathbf{f}_2, \ldots$ be the elements of a set $\mathscr{V}$ called *vectors* and denoted by boldface letters. We shall allow for two operations within $\mathscr{V}$:

(a) To every pair $\mathbf{f}_1$ and $\mathbf{f}_2$ in $\mathscr{V}$, there is an associated element $\mathbf{f}_3$ in $\mathscr{V}$, called the *sum* of the pair: $\mathbf{f}_3 = \mathbf{f}_1 + \mathbf{f}_2$.
(b) To every $\mathbf{f} \in \mathscr{V}$ ("$\mathbf{f}$ element of $\mathscr{V}$") and every $c \in \mathscr{C}$, there is an associated element $c\mathbf{f}$ in $\mathscr{V}$, referred to as the product of $\mathbf{f}$ by c.

With respect to the sum, $\mathscr{V}$ must satisfy the following:

(a1) *Commutativity:* $\mathbf{f}_1 + \mathbf{f}_2 = \mathbf{f}_2 + \mathbf{f}_1$,
(a2) *Associativity:* $(\mathbf{f}_1 + \mathbf{f}_2) + \mathbf{f}_3 = \mathbf{f}_1 + (\mathbf{f}_2 + \mathbf{f}_3)$,
(a3) $\mathscr{V}$ must contain a *zero* vector $\mathbf{0}$ such that $\mathbf{f} + \mathbf{0} = \mathbf{f}$ for all $\mathbf{f} \in \mathscr{V}$,
(a4) For every $\mathbf{f} \in \mathscr{V}$, there exists a $(-\mathbf{f}) \in \mathscr{V}$ such that $\mathbf{f} + (-\mathbf{f}) = \mathbf{0}$.

With respect to the product it is required that $\mathscr{V}$ satisfy

(b1) $1 \cdot \mathbf{f} = \mathbf{f}$,
(b2) $c_1(c_2\mathbf{f}) = (c_1c_2)\mathbf{f}$.

Finally, the two operations are to intertwine *distributively*, i.e.,

(c1) $c(\mathbf{f}_1 + \mathbf{f}_2) = c\mathbf{f}_1 + c\mathbf{f}_2$,
(c2) $(c_1 + c_2)\mathbf{f} = c_1\mathbf{f} + c_2\mathbf{f}$.

The last requirement relates the sum in $\mathscr{C}$ with the sum in $\mathscr{V}$. We use the same symbol "+" for both. Immediate consequences of these axioms are $0\mathbf{f} = \mathbf{0}$ and $(-1)\mathbf{f} = -\mathbf{f}$.

1.1.2. Linear Independence

Except for allowing the numbers $c_1, c_2, \ldots$ to be complex, the main concepts from ordinary vector analysis remain unchanged: A set of (nonzero) vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_N$ is said to be *linearly independent* when

$$\sum_{n=1}^{N} c_n\mathbf{f}_n = \mathbf{0} \Leftrightarrow c_n = 0, \qquad n = 1, 2, \ldots, N. \tag{1.1}$$

If the implication to the right does not hold, the set of vectors is said to be linearly *dependent*. A complex vector space $\mathscr{V}$ is said to be *N-dimensional* when it is possible to find at most N linearly independent vectors. We affix N to $\mathscr{V}$ as a superscript: $\mathscr{V}^N$. Let $\{\boldsymbol{\varepsilon}_n\}_{n=1}^N := \{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\}$ be a maximal set of linearly independent vectors, called a *basis* for $\mathscr{V}^N$. We can then express any $\mathbf{f} \in \mathscr{V}^N$ as a linear combination of the basis vectors as

$$\mathbf{f} = \sum_{n=1}^{N} f_n \boldsymbol{\varepsilon}_n, \tag{1.2}$$

where $f_n \in \mathscr{C}$ is the *nth coordinate* of $\mathbf{f}$ with respect to the basis $\{\boldsymbol{\varepsilon}_n\}_{n=1}^N$. If $\mathbf{f}$ has coordinates $\{f_n\}_{n=1}^N$ and $\mathbf{g}$ coordinates $\{g_n\}_{n=1}^N$, then the coordinates of a vector $\mathbf{h} = a\mathbf{f} + b\mathbf{g}$ will be $h_n = af_n + bg_n$ for $n = 1, 2, \ldots, N$, as implied by (1.1) and the linear independence of the basis vectors. The vector $\mathbf{0}$ has all its coordinates zero.

1.1.3. Canonical Representation

Any two N-dimensional vector spaces are isomorphic, as we need only establish a one-to-one correspondence between the basis vectors. A most convenient realization of $\{\boldsymbol{\varepsilon}_n\}_{n=1}^N$ is given through the *canonical* column-vector representation

$$\boldsymbol{\varepsilon}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\varepsilon}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \boldsymbol{\varepsilon}_N = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad \text{i.e., } \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_{N-1} \\ f_N \end{pmatrix}. \tag{1.3}$$

Throughout Part I, we shall consider finite-dimensional complex vector spaces.

Exercise 1.1. Map the complex vector space $\mathscr{V}^N$ onto a $2N$-dimensional *real* vector space (i.e., only real numbers allowed). You can number the basis vectors in the latter as $\boldsymbol{\varepsilon}_n{}^R := \boldsymbol{\varepsilon}_n$ and $\boldsymbol{\varepsilon}_{N+n}^{\mathrm{I}} := i\boldsymbol{\varepsilon}_n$, $n = 1, 2, \ldots, N$. (Any other choice?) How do the coordinates of a vector $\mathbf{f} \in \mathscr{V}^N$ relate to the coordinates of the corresponding vector in the real space?

For economy of notation we shall henceforth indicate summations as in (1.2) by $\sum_n$, the range of the index being implied by the context. Double sums will appear as $\sum_{n,m}$, etc. If any ambiguities should arise, we shall revert to the full summation symbol.

1.2. Inner Product and Norm in $\mathscr{V}^N$

In this section we shall generalize the inner (or "scalar") product and norm of ordinary vector analysis to corresponding concepts in complex vector spaces.

1.2.1. Inner Product

To every ordered pair of vectors **f**, **g** in $\mathscr{V}^N$, we associate a complex number (**f**, **g**), their *inner product*. It has the properties of being *linear* in the second argument, i.e.,

$$(\mathbf{f}, c_1\mathbf{g}_1 + c_2\mathbf{g}_2) = c_1(\mathbf{f}, \mathbf{g}_1) + c_2(\mathbf{f}, \mathbf{g}_2), \tag{1.4}$$

and *antilinear* in the first,

$$(c_1\mathbf{f}_1 + c_2\mathbf{f}_2, \mathbf{g}) = c_1^*(\mathbf{f}_1, \mathbf{g}) + c_2^*(\mathbf{f}_2, \mathbf{g}), \tag{1.5}$$

where the asterisk denotes complex conjugation. Such an inner product is thus a *sesquilinear* ("$1\frac{1}{2}$ linear") operation: $\mathscr{V}^N \times \mathscr{V}^N \to \mathscr{C}$. We shall assume that the inner product is *positive*; that is, $(\mathbf{f}, \mathbf{f}) > 0$ for every $\mathbf{f} \neq \mathbf{0}$.

1.2.2. Orthonormal Bases

Two vectors whose inner product is zero are said to be *orthogonal.* A basis such that its vectors satisfy

$$(\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_m) = \delta_{n,m} := \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{if } n \neq m \end{cases} \tag{1.6}$$

is said to be an *orthonormal basis.* It can easily be shown as in real vector analysis, by the Schmidt construction, that one can always find an orthonormal basis for $\mathscr{V}^N$. Conversely, we can define the inner product by demanding (1.6) for a given basis and then extend the definition through (1.4) and (1.5) to the whole space $\mathscr{V}^N$. For two arbitrary vectors **f** and **g** written in terms of the basis, we have

$$\begin{aligned} (\mathbf{f}, \mathbf{g}) &= \Big(\sum_n f_n\boldsymbol{\varepsilon}_n, \sum_m f_m\boldsymbol{\varepsilon}_m\Big) && \text{[from (1.2)]} \\ &= \sum_m g_m\Big(\sum_n f_n\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_m\Big) && \text{[from (1.4)]} \\ &= \sum_{n,m} f_n^* g_m(\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_m) && \text{[from (1.5)]} \\ &= \sum_n f_n^* g_n. && \text{[from (1.6)]} \end{aligned} \tag{1.7}$$

It is now easy to verify that

$$(\mathbf{f}, \mathbf{f}) \geqslant 0, \qquad (\mathbf{f}, \mathbf{f}) = 0 \Leftrightarrow \mathbf{f} = \mathbf{0}, \tag{1.8}$$

$$(\mathbf{f}, \mathbf{g}) = (\mathbf{g}, \mathbf{f})^*. \tag{1.9}$$

[In fact, Eqs. (1.4), (1.8), and (1.9) are sometimes used to *define* the inner product in a vector space: the two sets of axioms are equivalent whenever an orthonormal basis exists. This is the case for finite N-dimensional spaces but not always when N is infinite. In the latter, the definition (1.4)–(1.8)–(1.9) is used.]

1.2.3. Coordinates

The nth coordinate of a vector $\mathbf{f}$ in the orthonormal basis $\{\boldsymbol{\varepsilon}_n\}_{n=1}^{N}$ is easily recovered from $\mathbf{f}$ itself through the inner product: Performing the inner product of a fixed $\boldsymbol{\varepsilon}_m$ with Eq. (1.2), we find

$$(\boldsymbol{\varepsilon}_m, \mathbf{f}) = \left(\boldsymbol{\varepsilon}_m, \sum_n f_n \boldsymbol{\varepsilon}_n\right) = \sum_n f_n(\boldsymbol{\varepsilon}_m, \boldsymbol{\varepsilon}_n) = f_m. \tag{1.10}$$

Hence, we can write

$$\mathbf{f} = \sum_n \boldsymbol{\varepsilon}_n(\boldsymbol{\varepsilon}_n, \mathbf{f}). \tag{1.11}$$

1.2.4. Schwartz Inequality

Two vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ were said to be orthogonal if $(\mathbf{f}_1, \mathbf{f}_2) = 0$. On the other hand, two vectors $\mathbf{g}_1$ and $\mathbf{g}_2$ are *parallel* if $\mathbf{g}_1 = c\mathbf{g}_2$, $c \in \mathscr{C}$, in which case

$$(\mathbf{g}_1, \mathbf{g}_2) = c^*(\mathbf{g}_2, \mathbf{g}_2) = c^{-1}(\mathbf{g}_1, \mathbf{g}_1) = [c^*c^{-1}(\mathbf{g}_1, \mathbf{g}_1)(\mathbf{g}_2, \mathbf{g}_2)]^{1/2}, \tag{1.12}$$

where, note, $|c^*c^{-1}| = 1$. For $|(\mathbf{f}, \mathbf{g})|$, zero is a lower bound, while, in the event $\mathbf{f}$ and $\mathbf{g}$ are parallel, $|(\mathbf{f}, \mathbf{g})| = [(\mathbf{f}, \mathbf{f})(\mathbf{g}, \mathbf{g})]^{1/2}$. These are the extreme values, as stated in the well-known *Schwartz inequality*:

$$|(\mathbf{f}, \mathbf{g})|^2 \leqslant (\mathbf{f}, \mathbf{f})(\mathbf{g}, \mathbf{g}). \tag{1.13}$$

We can prove (1.13) as follows. Consider the vector $\mathbf{f} - c\mathbf{g}$. Then, because of (1.8),

$$0 \leqslant (\mathbf{f} - c\mathbf{g}, \mathbf{f} - c\mathbf{g}) = (\mathbf{f}, \mathbf{f}) - c(\mathbf{f}, \mathbf{g}) - c^*(\mathbf{g}, \mathbf{f}) + |c|^2(\mathbf{g}, \mathbf{g}). \tag{1.14}$$

Now choose (for $\mathbf{g} \neq \mathbf{0}$)

$$c = (\mathbf{f}, \mathbf{g})^*/(\mathbf{g}, \mathbf{g}). \tag{1.15}$$

Replacement in (1.14) and a rearrangement of terms yield (1.13).

1.2.5. Norm

The *norm* (or *length*) of a vector $\mathbf{f} \in \mathscr{V}^N$ is defined as

$$\|\mathbf{f}\| := (\mathbf{f}, \mathbf{f})^{1/2}. \tag{1.16}$$

It is a mapping from $\mathscr{V}^N$ onto $\mathscr{R}^+$ (the nonnegative halfline), having the properties

$$\|\mathbf{f}\| \geqslant 0, \qquad \|\mathbf{f}\| = 0 \Leftrightarrow \mathbf{f} = \mathbf{0}, \tag{1.17}$$

$$\|c\mathbf{f}\| = |c| \, \|\mathbf{f}\|, \tag{1.18}$$

$$\|\mathbf{f} + \mathbf{g}\| \leqslant \|\mathbf{f}\| + \|\mathbf{g}\|. \tag{1.19}$$

Equations (1.17) and (1.18) are easily proven from (1.8) and (1.4)–(1.5), while Eq. (1.19) is the *triangle inequality*, which states, quite geometrically, that the length of the sum of two vectors cannot exceed the sum of the lengths of the vectors. It can be proven from (1.14), setting $c = -1$, that

$$\begin{aligned} 0 \leqslant \|\mathbf{f} + \mathbf{g}\|^2 &= \|\mathbf{f}\|^2 + 2\,\mathrm{Re}(\mathbf{f}, \mathbf{g}) + \|\mathbf{g}\|^2 \\ &\leqslant \|\mathbf{f}\|^2 + 2|(\mathbf{f}, \mathbf{g})| + \|\mathbf{g}\|^2 \qquad (\text{from } \mathrm{Re}\, z \leqslant |z|) \\ &\leqslant \|\mathbf{f}\|^2 + 2\|\mathbf{f}\| \cdot \|\mathbf{g}\| + \|\mathbf{g}\|^2 \qquad [\text{from } (1.13)]. \end{aligned} \tag{1.20}$$

The square root of the second and last terms yields Eq. (1.19).

Exercise 1.2. From (1.14) show that

$$\|\mathbf{f} - \mathbf{g}\| \geqslant \big| \, \|\mathbf{f}\| - \|\mathbf{g}\| \, \big|. \tag{1.21}$$

This is another form of the triangle inequality.

We have obtained the properties of the norm, Eqs. (1.17)–(1.19), as consequences of the definition and properties of the inner product. The abstract definition of a *norm*, however, is that of a mapping from $\mathscr{V}^N$ onto $\mathscr{R}^+$, with properties (1.17)–(1.19). It is a weaker requirement than that of an inner product and quite independent of it. The definition (1.16) only represents a particular kind of norm. Again, in infinite-dimensional spaces one may define a norm but have no inner product.

Exercise 1.3. Prove the *polarization* identity

$$(\mathbf{f}, \mathbf{g}) = \tfrac{1}{4}(\|\mathbf{f} + \mathbf{g}\|^2 - \|\mathbf{f} - \mathbf{g}\|^2) + i\tfrac{1}{4}(\|\mathbf{f} - i\mathbf{g}\|^2 - \|\mathbf{f} + i\mathbf{g}\|^2). \tag{1.22}$$

Note that this identity hinges on the validity of (1.16). It *cannot* be used to define an inner product from a norm.

Exercise 1.4. Define the complex angle between two vectors by

$$\cos\Theta := (\mathbf{f}, \mathbf{g})/\|\mathbf{f}\| \cdot \|\mathbf{g}\|, \qquad \Theta = \theta_R + i\theta_I. \tag{1.23}$$

Show that this restricts Θ to a region $|\sinh\theta_I| \leqslant |\sin\theta_R| \leqslant 1$.

1.3. Passive and Active Transformations

In this section we shall introduce two kinds of transformations on the coordinates of vectors in $\mathscr{V}^N$, those which arise from a change in the basis used for the description of the space, referred to as *passive* transformations, and *active* transformations produced by operators which bodily move the vectors in $\mathscr{V}^N$. Although the resulting expressions for the two kinds of transformations are quite similar, the difference in their interpretation is important.

1.3.1. Transformation of the Basis Vectors

Consider the complex vector space $\mathscr{V}^N$ and the orthonormal basis $\{\boldsymbol{\varepsilon}_n\}_{n=1}^N$ (henceforth called the $\boldsymbol{\varepsilon}$-basis, for short). Out of the $\boldsymbol{\varepsilon}$-basis we can construct the set of vectors

$$\bar{\boldsymbol{\varepsilon}}_m = \sum_n V_{nm}\boldsymbol{\varepsilon}_n, \qquad n = 1, 2, \ldots, N, \tag{1.24}$$

where $V_{nm} \in \mathscr{C}$. The question of the linear independence of the vector set (1.24) can be posed as follows. Let $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_N$ be a set of constants such that

$$\mathbf{0} = \sum_m \bar{c}_m \bar{\boldsymbol{\varepsilon}}_m = \sum_{m,n} \bar{c}_m V_{nm}\boldsymbol{\varepsilon}_n = \sum c_n \boldsymbol{\varepsilon}_n, \tag{1.25}$$

where $c_n = \sum_m \bar{c}_m V_{nm}$. Now, the vectors of the $\boldsymbol{\varepsilon}$-basis are linearly independent, so $c_n = 0$ for $n = 1, 2, \ldots, N$. For this to imply that all the $\bar{c}_m = 0$, $m = 1, 2, \ldots, N$, it is necessary that the matrix $\mathbf{V} = \|V_{nm}\|$ have a nonvanishing determinant. Thus, if $\det \mathbf{V} \neq 0$, the linear independence of the $\boldsymbol{\varepsilon}$-basis implies the linear independence of the N vectors in (1.24). The latter are then a basis as well. Henceforth it will be called the $\bar{\boldsymbol{\varepsilon}}$-basis. The $\bar{\boldsymbol{\varepsilon}}$-basis will not in general consist of mutually orthogonal vectors, but

$$\begin{aligned}(\bar{\boldsymbol{\varepsilon}}_n, \bar{\boldsymbol{\varepsilon}}_m) &= \sum_{j,k} (V_{jn}\boldsymbol{\varepsilon}_j, V_{km}\boldsymbol{\varepsilon}_k) \\ &= \sum_k V_{kn}^* V_{km} = (\mathbf{V}^\dagger\mathbf{V})_{nm},\end{aligned} \tag{1.26}$$

where $\mathbf{V}^\dagger = \mathbf{V}^{T*}$ is the transposed conjugate or *adjoint* of the matrix $\mathbf{V}$ and $(\mathbf{V}^\dagger)_{nm} = V_{mn}^*$.

1.3.2. Passive Transformations

We can regard the matrix $\mathbf{V} = \|V_{nm}\|$ as effecting a *change of basis* for $\mathscr{V}^N$: a *passive transformation* whereby the description of the vectors of $\mathscr{V}^N$ in terms of the $\boldsymbol{\varepsilon}$-basis is replaced by their description in terms of the $\bar{\boldsymbol{\varepsilon}}$-basis.

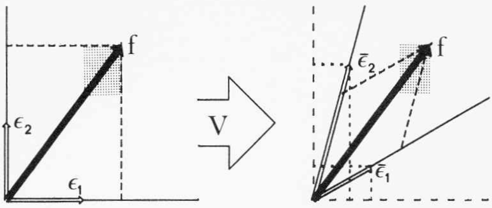


Fig. 1.1. Passive transformation V of a (two-dimensional) vector space. Its description in terms of a basis $\{\boldsymbol{\varepsilon}_i\}$ is replaced by its description in terms of a transformed basis $\{\bar{\boldsymbol{\varepsilon}}_i\}$. The vectors $\mathbf{f}$ in the space are unchanged.

Let $\mathbf{f} \in \mathscr{V}^N$ be a (fixed) vector with coordinates f_n, $n = 1, 2, \ldots, N$, relative to the $\boldsymbol{\varepsilon}$-basis and coordinates $\bar{f}_m$, $m = 1, 2, \ldots, N$, relative to the $\bar{\boldsymbol{\varepsilon}}$-basis. Then (see Fig. 1.1)

$$\sum_n f_n \boldsymbol{\varepsilon}_n = \mathbf{f} = \sum_m \bar{f}_m \bar{\boldsymbol{\varepsilon}}_m = \sum_{n,m} \bar{f}_m V_{nm} \boldsymbol{\varepsilon}_n \qquad \text{(passive).} \tag{1.27}$$

The first and last members of this equation, due to the linear independence of the basis vectors, yield

$$f_n = \sum_m V_{nm} \bar{f}_m, \qquad \bar{f}_m = \sum_n (\mathbf{V}^{-1})_{mn} f_n. \tag{1.28}$$

The matrix $\mathbf{V}^{-1}$ exists as $\mathbf{V}$ is assumed to be nonsingular ($\det \mathbf{V} \neq 0$).

Exercise 1.5. Let the coordinates of $\mathbf{f}$ relative to the $\bar{\boldsymbol{\varepsilon}}$-basis be $\bar{f}_m$ [i.e., second and third members of Eq. (1.27)]. Performing the inner product with $\bar{\boldsymbol{\varepsilon}}_n$ and using (1.26), find $\bar{f}_m$ in terms of $(\bar{\boldsymbol{\varepsilon}}_n, \mathbf{f})$.

Exercise 1.6. Using the result of Exercise 1.5, define the set of vectors $\bar{\boldsymbol{\varepsilon}}_n{}^D$ ($n = 1, 2, \ldots, N$) so that $\bar{f}_n = (\bar{\boldsymbol{\varepsilon}}_n{}^D, \mathbf{f})$. Show that this defines a basis for $\mathscr{V}^N$. It is called the basis *dual* to the $\bar{\boldsymbol{\varepsilon}}$-basis, since (prove!) $(\bar{\varepsilon}_n, \bar{\varepsilon}_m{}^D) = \delta_{n,m}$. If the $\bar{\boldsymbol{\varepsilon}}$-basis is orthonormal, then $\bar{\boldsymbol{\varepsilon}}_n{}^D = \bar{\boldsymbol{\varepsilon}}_n$ ($n = 1, 2, \ldots, N$).

Exercise 1.7. Express $(\mathbf{f}, \mathbf{g})$ in terms of the coordinates of $\mathbf{f}$ and $\mathbf{g}$ in the $\bar{\boldsymbol{\varepsilon}}$-basis.

1.3.3. Active Transformations

Active transformations are produced by operators $\mathbb{A}$ mapping $\mathscr{V}^N$ onto $\mathscr{V}^N$, which transform the vectors of the space as $\mathbf{f} \mapsto \mathbf{f}' = \mathbb{A}\mathbf{f}$. We shall assume these operators to be *linear*, i.e.,

$$\mathbb{A}(a\mathbf{f} + b\mathbf{g}) = a\mathbb{A}\mathbf{f} + b\mathbb{A}\mathbf{g}. \tag{1.29}$$

The linearity requirement allows us to find the transformation undergone by every vector in the space when we know the way the vectors in a given basis (say, the $\boldsymbol{\varepsilon}$-basis) are transformed. Let

$$\boldsymbol{\varepsilon}'_m = \mathbb{A}\boldsymbol{\varepsilon}_m, \qquad m = 1, 2, \ldots, N, \tag{1.30}$$

and define the N^2 constants

$$A_{nm} := (\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}'_m) = (\boldsymbol{\varepsilon}_n, \mathbb{A}\boldsymbol{\varepsilon}_m). \tag{1.31}$$

Using Eq. (1.11) with $\boldsymbol{\varepsilon}'_m$ in place of $\mathbf{f}$, we find

$$\boldsymbol{\varepsilon}'_m = \sum_n A_{nm}\boldsymbol{\varepsilon}_n, \tag{1.32}$$

which is formally identical to (1.24) with A_{nm} in place of V_{nm}. The interpretation of (1.32) as a linear active transformation, however, requires that the vectors $\mathbf{f} \in \mathscr{V}^N$ and the basis $\boldsymbol{\varepsilon}$ undergo the same transformation; that is, the coordinates of $\mathbf{f}'$ in the new basis $\boldsymbol{\varepsilon}'$ *continue to be* f_n, $n = 1, 2, \ldots, N$. Now, denoting by f'_n ($n = 1, 2, \ldots, N$) the coordinates of $\mathbf{f}'$ with respect to the *original* $\boldsymbol{\varepsilon}$-basis, we have

$$\sum_n f'_n\boldsymbol{\varepsilon}_n = \mathbf{f}' = \sum_m f_m\boldsymbol{\varepsilon}'_m = \sum_{m,n} f_m A_{nm}\boldsymbol{\varepsilon}_n \qquad \text{(active)}, \tag{1.33}$$

and this implies

$$f'_n = \sum_m A_{nm} f_m, \tag{1.34}$$

so the coordinates of $\mathbf{f}$ transform as a column vector under the matrix $\mathbf{A} = \|A_{nm}\|$.

1.3.4. Operators and Their Matrix Representatives

As a consequence of the construction (1.31), we see that any linear *operator* $\mathbb{A}$ can be *represented* by a matrix $\mathbf{A}$, acting on the column-vector canonical realization (1.3). The matrix $\mathbf{A}$ was determined uniquely from the linear operator $\mathbb{A}$. Conversely, $\mathbb{A}$ is uniquely determined by $\mathbf{A}$ since the transformation of the basis vectors (1.32) specifies the transformation of any vector in the space. See Fig. 1.2.

We shall now see that this one-to-one correspondence between linear operators and $N \times N$ matrices holds under sum and product of the corresponding quantities. We define the linear combination of two operators

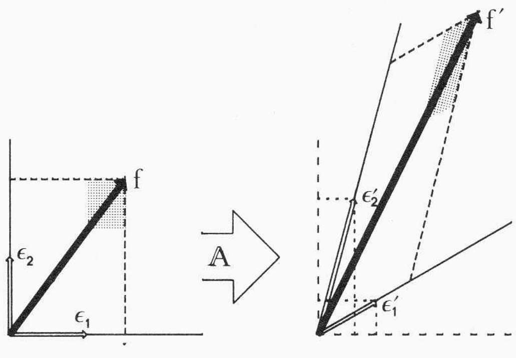


Fig. 1.2. Active transformation $\mathbb{A}$ of a (two-dimensional) vector space. All vectors—basis vectors included—are changed. As the transformation is linear, however, the coordinates of $\mathbf{f}' = \mathbb{A}\mathbf{f}$ in the transformed basis $\{\boldsymbol{\varepsilon}'_i\} = \{\mathbb{A}\boldsymbol{\varepsilon}_i\}$ are the same as those of $\mathbf{f}$ in the original basis.

$\mathbb{C} = a\mathbb{A} + b\mathbb{B}$, quite naturally, as

$$(a\mathbb{A} + b\mathbb{B})\mathbf{f} := a\mathbb{A}\mathbf{f} + b\mathbb{B}\mathbf{f}. \tag{1.35}$$

Now let **A**, **B**, and **C** be the representing matrices. Then, using (1.31),

$$\begin{aligned} C_{nm} &= \big(\boldsymbol{\varepsilon}_n, (a\mathbb{A} + b\mathbb{B})\boldsymbol{\varepsilon}_m\big) = a(\boldsymbol{\varepsilon}_n, \mathbb{A}\boldsymbol{\varepsilon}_m) + b(\boldsymbol{\varepsilon}_n, \mathbb{B}\boldsymbol{\varepsilon}_m) \\ &= aA_{nm} + bB_{nm}, \end{aligned} \tag{1.36}$$

so that $\mathbf{C} = a\mathbf{A} + b\mathbf{B}$. Similarly, for the product $\mathbb{D} = \mathbb{A}\mathbb{B}$,

$$(\mathbb{A}\mathbb{B})\mathbf{f} := \mathbb{A}(\mathbb{B}\mathbf{f}). \tag{1.37}$$

The correspondence with the representing matrices **D**, **A**, and **B** can be established using (1.31), (1.11) for $\mathbb{B}\boldsymbol{\varepsilon}_m$, and the linearity of the operators involved,

$$\begin{aligned} D_{nm} &= (\boldsymbol{\varepsilon}_n, \mathbb{A}\mathbb{B}\boldsymbol{\varepsilon}_m) = \Big(\boldsymbol{\varepsilon}_n, \mathbb{A}\sum_k \boldsymbol{\varepsilon}_k(\boldsymbol{\varepsilon}_k, \mathbb{B}\boldsymbol{\varepsilon}_m)\Big) \\ &= \sum_k (\boldsymbol{\varepsilon}_n, \mathbb{A}\boldsymbol{\varepsilon}_k)(\boldsymbol{\varepsilon}_k, \mathbb{B}\boldsymbol{\varepsilon}_m) = \sum_k A_{nk}B_{km}, \end{aligned} \tag{1.38}$$

so that $\mathbf{D} = \mathbf{AB}$.

1.3.5. Representations in Different Bases

We shall use passive transformations when a given system lends itself to a more convenient description in terms of a new set of coordinates. Active transformations, on the other hand, will describe, for instance, the time evolution of the *state vector* of a system. Note that active transformations of $\mathscr{V}^N$ should not depend on the basis used for the description of the space. Indeed, the representation of $\mathbb{A}$ by a matrix $\mathbf{A} = \|A_{nm}\|$ in (1.31) was made relative to the $\boldsymbol{\varepsilon}$-basis, but under any (passive) change of basis to, say, the $\bar{\boldsymbol{\varepsilon}}$-basis, the same operator $\mathbb{A}$ would be described by a different matrix $\overline{\mathbf{A}} = \|\overline{A}_{nm}\|$ whose elements are

$$\begin{aligned} \overline{A}_{nm} &= (\bar{\boldsymbol{\varepsilon}}_n, \mathbb{A}\bar{\boldsymbol{\varepsilon}}_m) = \sum_{j,k} (V_{jn}\boldsymbol{\varepsilon}_j, \mathbb{A}V_{km}\boldsymbol{\varepsilon}_k) \\ &= \sum_{j,k} V_{jn}^* A_{jk} V_{km} = (\mathbf{V}^\dagger\mathbf{A}\mathbf{V})_{nm}. \end{aligned} \tag{1.39}$$

Exercise 1.8. Show that

$$(\mathbb{A}\mathbf{f}, \mathbb{A}\mathbf{g}) = \sum_{m,n} f_m^*(\mathbf{A}^\dagger\mathbf{A})_{mn} g_n. \tag{1.40}$$

Do the same in terms of coordinates in a nonorthonormal basis.

Exercise 1.9. Define the operator $\mathbb{A}^\dagger$ as that having a matrix representation $\mathbf{A}^\dagger$ in some (orthonormal) basis. We call $\mathbb{A}^\dagger$ the adjoint of $\mathbb{A}$. Show that

$$(\mathbf{f}, \mathbb{A}^\dagger\mathbf{g}) = (\mathbb{A}\mathbf{f}, \mathbf{g}). \tag{1.41}$$

Show that this definition of $\mathbb{A}^\dagger$ does not depend on the matrix representation and that, for any other basis, $\overline{\mathbf{A}^\dagger} = (\bar{\mathbf{A}})^\dagger$.

Exercise 1.10. Let $\mathbb{A}$ and $\mathbb{B}$ be linear operators, $\mathbb{C} = a\mathbb{A} + b\mathbb{B}$ and $\mathbb{D} = \mathbb{A}\mathbb{B}$. Find the representing matrices $\bar{\mathbf{C}}$ and $\bar{\mathbf{D}}$ in the (in general nonorthogonal) $\bar{\boldsymbol{\varepsilon}}$-basis.

Exercise 1.10 should convince the reader that by far the simplest description of vector space operations is in terms of orthonormal bases. In fact, from now on we shall deal exclusively with these bases. This imposes severe restrictions on the allowed **V** in (1.24), which will be examined below. If the reader wants to deepen this necessarily brief account of vector spaces, inner products, and linear transformations, he may find useful the excellent text by Bowen and Wang (1976, Chapters 0–5).

1.4. Unitary Transformations: Permutations and the Fourier Transform

1.4.1. Definition of Unitarity

A transformation **V** which maps an orthonormal basis of the space $\mathscr{V}^N$, $\boldsymbol{\varepsilon}$, to another orthonormal basis $\bar{\boldsymbol{\varepsilon}}$ is called a *unitary* transformation. The necessary and sufficient condition for this to happen can be seen from (1.26) to be

$$\sum_k V_{kn}^* V_{km} = \delta_{n,m}, \quad \text{i.e.,} \quad \mathbf{V}^\dagger\mathbf{V} = \mathbf{1}, \tag{1.42}$$

where **1** is the $N \times N$ unit matrix. Such a matrix **V** is also called *unitary*, and clearly satisfies $\mathbf{V}^{-1} = \mathbf{V}^\dagger$: Its inverse equals its adjoint. As now both the $\boldsymbol{\varepsilon}$- and $\bar{\boldsymbol{\varepsilon}}$-bases are orthonormal, it follows that

$$(\mathbf{f}, \mathbf{g}) = \sum_n f_n^* g_n = \sum_n \bar{f}_n^* \bar{g}_n. \tag{1.43}$$

This is the *Parseval identity* between the coordinates of two vectors **f** and **g** in two bases related by a unitary transformation. (Compare with the result of Exercise 1.7.)

1.4.2. Groups of Unitary Matrices

Geometrically, a unitary transformation can be seen as a *rigid rotation* and/or *reflection* in a complex N-dimensional space: the angle Θ between any two vectors [Eq. (1.23)] is unchanged. Note that, as $\det \mathbf{V}^\dagger = (\det \mathbf{V})^*$, it follows from (1.42) that

$$|\det \mathbf{V}| = 1 \qquad (\mathbf{V} \text{ unitary}). \tag{1.44}$$

One general property of unitary transformations is that they constitute a *group*. This will be defined now. Consider the set $\mathscr{U}$ of unitary matrices. Then, as will be verified below,

(a) $\mathbf{V}_1, \mathbf{V}_2 \in \mathscr{U} \Rightarrow \mathbf{V}_1 \cdot \mathbf{V}_2 \in \mathscr{U}$.
(b) $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3 \in \mathscr{U} \Rightarrow (\mathbf{V}_1 \cdot \mathbf{V}_2) \cdot \mathbf{V}_3 = \mathbf{V}_1 \cdot (\mathbf{V}_2 \cdot \mathbf{V}_3)$.
(c) There exists a *unit* element $\mathbf{E} \in \mathscr{U}$ such that $\mathbf{E} \cdot \mathbf{V} = \mathbf{V}$ for all $\mathbf{V} \in \mathscr{U}$.
(d) For every $\mathbf{V} \in \mathscr{U}$, there exists a $\mathbf{V}^{-1} \in \mathscr{U}$ such that $\mathbf{V}\mathbf{V}^{-1} = \mathbf{E}$.

Abstractly, if the set $\mathscr{U}$ satisfies (a)–(d), it is said to constitute a *group* under the product operation "$\cdot$". In our case, "$\cdot$" is matrix multiplication, and we can verify (a): Let $\mathbf{V}_1^\dagger = \mathbf{V}_1^{-1}$ and $\mathbf{V}_2^\dagger = \mathbf{V}_2^{-1}$; then $(\mathbf{V}_1\mathbf{V}_2)^\dagger = \mathbf{V}_2^\dagger\mathbf{V}_1^\dagger = \mathbf{V}_2^{-1}\mathbf{V}_1^{-1} = (\mathbf{V}_1\mathbf{V}_2)^{-1}$. (b) Complex matrix multiplication is always associative. (c) The unit matrix $\mathbf{1}$ satisfies $\mathbf{1}^\dagger = \mathbf{1} = \mathbf{1}^{-1}$ and thus belongs to $\mathscr{U}$; it has the property $\mathbf{1} \cdot \mathbf{V} = \mathbf{V}$, so we identify $\mathbf{E} = \mathbf{1}$. (d) $(\mathbf{V}^{-1})^\dagger = \mathbf{V}^{\dagger\dagger} = \mathbf{V} = (\mathbf{V}^{-1})^{-1}$.

We conclude that *the set of unitary matrices constitutes a group under multiplication.* Although we shall use the notions of groups sparingly in this text, showing on occasion that sets of operations or objects have the group property under the appropriate product and drawing some immediate consequences, we should emphasize that group theory has been one of the fastest growing branches in applied mathematics. For the reader interested in further study on this field, we can suggest the books by Hamermesh (1962) and Miller (1972).

Exercise 1.11. Show that a vector space has the structure of a group under the "+" operation. The unit element is the zero vector.

We shall now examine two particularly important unitary transformations: permutations and the Fourier transformation.

1.4.3. Permutations

A *permutation* p of the basis vectors $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\}$ (or of any set of numbered objects) is a transformation to a new basis $\{\bar{\boldsymbol{\varepsilon}}_1, \bar{\boldsymbol{\varepsilon}}_2, \ldots, \bar{\boldsymbol{\varepsilon}}_N\} = \{\boldsymbol{\varepsilon}_{p(1)}, \boldsymbol{\varepsilon}_{p(2)}, \ldots, \boldsymbol{\varepsilon}_{p(N)}\}$, where only the order of the elements in the set is changed. The string of numbers $p(1), p(2), \ldots, p(N)$ is a permutation p of $1, 2, \ldots, N$, and $p(m) = p(n) \Leftrightarrow m = n$.

1.4.4. Representing Permutations by Matrices

We can display p by the symbol

$$\begin{pmatrix} 1 & 2 & \cdots & N \\ p(1) & p(2) & \cdots & p(N) \end{pmatrix} =: \begin{pmatrix} m \\ p(m) \end{pmatrix},$$

which on any numbered set replaces the mth element by the $p(m)$th one. The order of the columns in this symbol is irrelevant. Such a permutation is achieved in (1.24) by a matrix $\mathbf{P} = \|P_{nm}\|$ whose mth column has a single nonzero element—of value 1—in the $p(m)$th row, i.e.,

$$\left[\mathbf{P}\begin{pmatrix} p & 2 & \cdots & N \\ p(1) & p(2) & \cdots & p(N)\end{pmatrix}\right]_{nm} = \delta_{n,p(m)}, \tag{1.45}$$

so that $\bar{\boldsymbol{\varepsilon}}_m = \boldsymbol{\varepsilon}_{p(m)}$. The inverse p^{-1} of the permutation p permutes the set $p(1), p(2), \ldots, p(N)$ back to $1, 2, \ldots, N$. This is achieved by

$$\left[\mathbf{P}\begin{pmatrix} 1 & 2 & \cdots & N \\ p(1) & p(2) & \cdots & p(N)\end{pmatrix}^{-1}\right]_{nm} = \left[\mathbf{P}\begin{pmatrix} p(1) & p(2) & \cdots & p(N) \\ 1 & 2 & \cdots & N\end{pmatrix}\right]_{nm}$$

$$= \delta_{m,p(n)} = \left[\mathbf{P}\begin{pmatrix} 1 & 2 & \cdots & N \\ p(1) & p(2) & \cdots & p(N)\end{pmatrix}^{\dagger}\right]_{nm}, \tag{1.46}$$

so that $\bar{\boldsymbol{\varepsilon}}_{p(m)} = \boldsymbol{\varepsilon}_m$. As the elements of $\mathbf{P}$ and $\mathbf{P}^{-1}$ are real, it was concluded in (1.46) that $\mathbf{P}^{-1} = \mathbf{P}^{\dagger}$. It follows that the permutation matrix (1.45) is unitary and that *the permutation of basis vectors is a unitary transformation in* $\mathscr{V}^N$.

The product of two permutations p_1, p_2 is a permutation p_3 since, applied to any numbered set on the right of the symbols,

$$\begin{pmatrix} m \\ p_2(m)\end{pmatrix}\begin{pmatrix} m \\ p_1(m)\end{pmatrix} = \begin{pmatrix} m \\ p_2(p_1(m))\end{pmatrix} = \begin{pmatrix} m \\ p_3(m)\end{pmatrix}. \tag{1.47}$$

Note that the product of two permutations is not commutative in general. The identity permutation $e = \binom{m}{m}$ which leaves every element in a numbered set in its position is obviously a permutation represented in (1.45) by the unit matrix. Finally, as the inverse p^{-1} of a given permutation p is also a permutation represented in (1.46) by $\mathbf{P}^{-1} = \mathbf{P}^{\dagger}$, *the set of all permutations is a group*. We shall denote it by π_N. It has $N!$ elements. As permutations are a subset of the group of unitary transformations, they are said to be a *subgroup* of the latter.

Exercise 1.12. Note that, since the matrices involved are real, $\det \mathbf{P} = \pm 1$. Show that the matrices representing *transpositions* of two elements [i.e., $\left(\begin{smallmatrix} 1 & 2 & \cdots & k & \cdots & l & \cdots & N \\ 1 & 2 & \cdots & l & \cdots & k & \cdots & N\end{smallmatrix}\right)$, where only k and l are exchanged] have a determinant equal to -1.

Exercise 1.13. Show that the product of two transpositions which have one element in common has the form of a *three-cycle* [i.e., $\left(\begin{smallmatrix} 1 & 2 & \cdots & k & \cdots & l & \cdots & m & \cdots & N \\ 1 & 2 & \cdots & l & \cdots & m & \cdots & k & \cdots & N\end{smallmatrix}\right)$, where only k, l, and m are exchanged]. Show that the matrix representing a three-cycle has determinant $+1$.

Exercise 1.14. Consider the real space $\mathscr{V}^3$ for definiteness. Show that permutations with determinant $+1$ are rotations of the coordinate axes, while permutations with determinant -1 involve reflections across planes.

Exercise 1.15. For $N > 3$, one can produce four-cycles, etc. (up to N-cycles), from the product of a three-cycle, etc. [up to $(N - 1)$-cycle], and a transposition with one element in common. Show that n-cycles are represented by matrices with determinant $(-1)^{n+1}$. When this is $+1$ they can be realized as rotations of the coordinate axes in N-space.

Exercise 1.16. Show that all permutations represented by matrices (1.45) with determinant $+1$ form a *subgroup* of the permutation group. Show that those with determinant -1 do not.

1.4.5. The Fourier Transformation

The unitary transformation which is the prime subject of this part is the Fourier transformation, defined in $\mathscr{V}^N$ by the matrix $\mathbf{F} = \|F_{mn}\|$ with elements

$$F_{mn} := N^{-1/2} \exp(-2\pi i m n/N) = F_{nm}. \tag{1.48}$$

We can verify directly that (1.48) is a unitary transformation, i.e.,

$$(\mathbf{F}^{\dagger}\mathbf{F})_{mn} = \sum_k F^*_{mk} F_{nk} = N^{-1} \sum_k \exp[-2\pi i k(n - m)/N] = \delta_{m,n}. \tag{1.49}$$

Using the geometric progression formula with $b + 1$ terms

$$x^a + x^{a+1} + \cdots + x^{a+b} = \begin{cases} (1 - x)^{-1} x^a (1 - x^{b+1}), & x \neq 1, \\ b + 1, & x = 1, \end{cases} \tag{1.50}$$

and letting $x = \exp[-2\pi i(n - m)/N]$, $a = 1$, and $b = N - 1$, we see that for $m \neq n$, $x \neq 1$, and the sum in (1.49) adds to zero, while for $m = n$, $x = 1$, and it adds to N.

1.4.6. Coordinates in the ε- and φ-Bases

From the above it follows that the coordinates of a vector $\mathbf{f}$ in two bases related by the Fourier transformation are given by (1.28), with $\mathbf{F}^{-1} = \mathbf{F}^{\dagger}$ by

$$\tilde{f}_n = N^{-1/2} \sum_m f_m \exp(2\pi i m n/N), \tag{1.51a}$$

$$f_n = N^{-1/2} \sum_m \tilde{f}_m \exp(-2\pi i m n/N). \tag{1.51b}$$

The set $\{\tilde{f}_n\}_{n=1}^N$ is said to be the (finite) *Fourier transform* of the set $\{f_n\}_{n=1}^N$ and the latter, the *inverse* Fourier transform of the former. In our approach, we want to emphasize, they are the coordinates of the same vector **f** in two bases.

The reasons for regarding the Fourier transformation as a particularly important unitary transformation should become clear in the applications of Chapters 2 and 3. Meanwhile, we reserve the tildes in Eqs. (1.51a) and (1.51b) for Fourier transforms, and we shall call the corresponding basis (called the $\bar{\boldsymbol{\varepsilon}}$-basis in Section 1.3) the $\boldsymbol{\varphi}$-basis. Explicitly,

$$\boldsymbol{\varphi}_n = N^{-1/2} \sum_m \boldsymbol{\varepsilon}_m \exp(-2\pi imn/N), \tag{1.52a}$$

$$\boldsymbol{\varepsilon}_n = N^{-1/2} \sum_m \boldsymbol{\varphi}_m \exp(2\pi imn/N), \tag{1.52b}$$

$$\sum_n f_n \boldsymbol{\varepsilon}_n = \mathbf{f} = \sum_n \tilde{f}_n \boldsymbol{\varphi}_n. \tag{1.52c}$$

Clearly, the mth coordinate of $\boldsymbol{\varphi}_n$ in the $\boldsymbol{\varepsilon}$-basis is F_{mn}. In Fig. 1.3 we show the real and imaginary parts of these coordinates for $N = 7$. We have let m take on continuous values and have drawn them as dotted lines in the figure.

1.4.7. Powers of the Fourier Transformation

One of the properties of the Fourier transformation matrix (1.48) is that it is *a fourth root of the unit matrix*. Indeed,

$$(\mathbf{F}^2)_{mn} = \sum_k F_{mk} F_{kn} = N^{-1} \sum_k \exp[-2\pi ik(m + n)/N]. \tag{1.53}$$

Using (1.50), we see that (1.53) equals 1 whenever $m + n = N$ or $2N$, as then all the N summands are N^{-1}. Hence, $\mathbf{F}^2$ is a matrix with 1's above the main antidiagonal and in the N–N position, with zeros elsewhere. In fact, it is a permutation

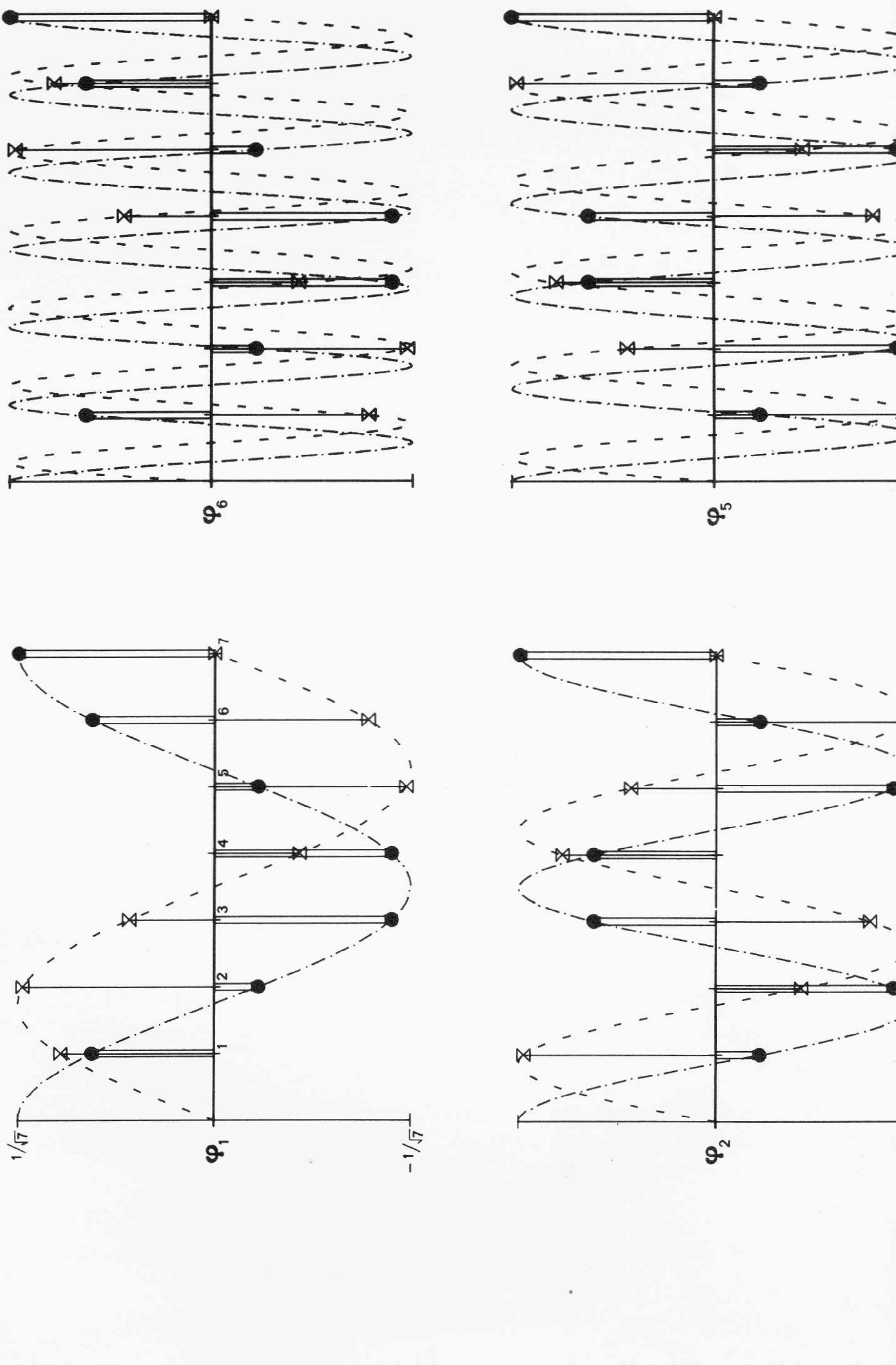
$$\mathbf{F}^2 = \mathbf{P}\begin{pmatrix} 1 & 2 & \cdots & N-2 & N-1 & N \\ N-1 & N-2 & \cdots & 2 & 1 & N \end{pmatrix} =: \mathbf{I}_0. \tag{1.54}$$

We shall call this the *inversion* matrix. Squaring (1.54), we find

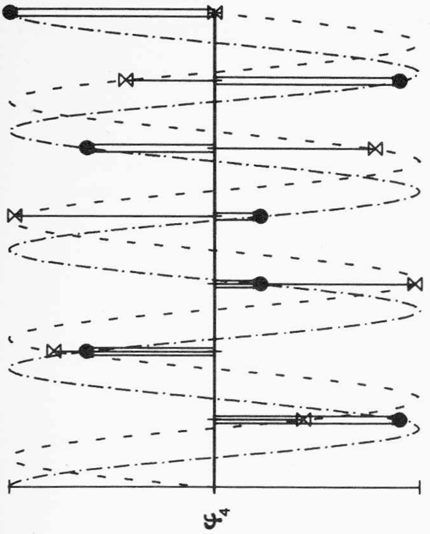
$$\mathbf{F}^4 = \mathbf{1}. \tag{1.55}$$

Exercise 1.17. Decompose the coordinates of a vector **f** in $\mathscr{V}^N$ into their real and imaginary parts as $f_n = f_n^R + if_n^I$ and $\tilde{f}_n = \tilde{f}_n^R + i\tilde{f}_n^I$. Relate these by (1.51a) and (1.51b).

Exercise 1.18. Associate to every vector **f** in $\mathscr{V}^N$ another vector $\mathbf{f}^*$ whose coordinates in the $\boldsymbol{\varepsilon}$-basis are $(\mathbf{f}^*)_n = (f_n)^*$, the complex conjugates of the original



φ_1
φ_2
φ_5
φ_6
$1/\sqrt{7}$
$-1/\sqrt{7}$
1
2
3
4
5
6
7



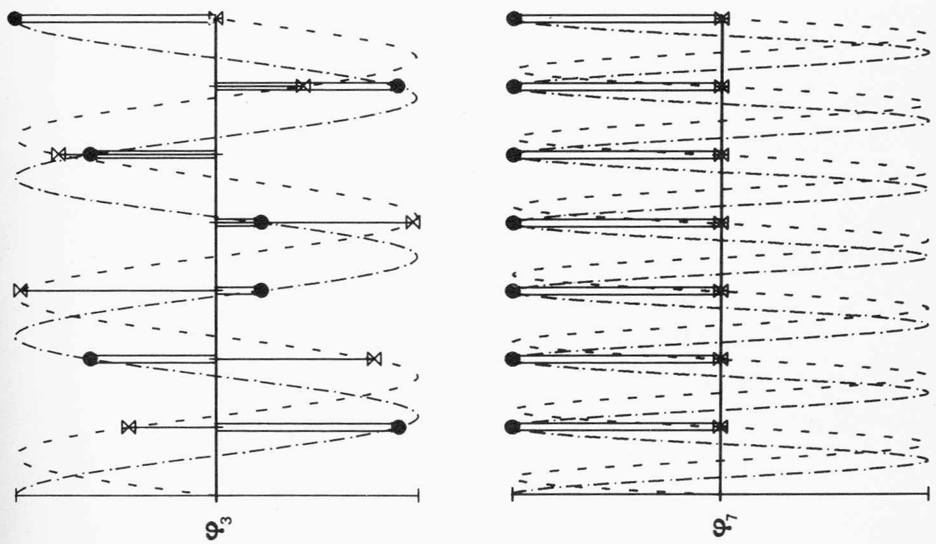


Fig. 1.3. The coordinates of the $\boldsymbol{\varphi}$-basis vectors in the $\boldsymbol{\varepsilon}$-basis reference frame for $N = 7$. The real and imaginary parts of the mth coordinate of $\boldsymbol{\varphi}_n$ are represented, respectively, by full circles ● and hourglasses ⧗. They fall on different broken sinusoidal lines plotted for real m in $(0, 7)$, but their values for integer m can coincide: The coordinates of $\boldsymbol{\varphi}_{N-n}$ are the complex conjugates of the coordinates of $\boldsymbol{\varphi}_n$.

vector. Note that this *cannot* be produced by a linear operator. Show that, in the $\boldsymbol{\varphi}$-basis, the coordinates of $\mathbf{f}^*$ are $\widetilde{(\mathbf{f}^*)}_n = (\tilde{f}_{N-n})^*$.

Exercise 1.19. Prove that the components of a vector $\mathbf{f}$ are *positive* if and only if their Fourier transforms are *positive definite*, i.e.,

$$f_n > 0 \Leftrightarrow \sum_{m,m'} \tilde{f}_{m-m'} \tilde{g}_m^* \tilde{g}_{m'} > 0 \tag{1.56}$$

for an arbitrary vector $\mathbf{g}$ with components $\tilde{g}_m$ in the $\boldsymbol{\varphi}$-basis. This is easy when you show that the second expression in (1.56) is $N^{1/2} \sum_n f_n |g_n|^2$. The coordinates f_n are numbered modulo N ($n \equiv n \bmod N$).

1.5. Self-Adjoint Operators

1.5.1. Definition of Adjunction

We have seen that linear operators $\mathbb{A}$ producing active transformations in $\mathscr{V}^N$ could be represented by matrices $\mathbf{A}$. We define the *adjoint* of $\mathbb{A}$, $\mathbb{A}^\dagger$, as that operator fulfilling

$$(\mathbf{f}, \mathbb{A}^\dagger \mathbf{g}) = (\mathbb{A}\mathbf{f}, \mathbf{g}) \tag{1.57}$$

for every pair of $\mathbf{f}, \mathbf{g} \in \mathscr{V}^N$. Equation (1.57) defines $\mathbb{A}^\dagger$ uniquely if it defines its matrix representative in a unique way. That this is so can be seen letting $\mathbf{f}$ and $\mathbf{g}$ be $\boldsymbol{\varepsilon}_n$ and $\boldsymbol{\varepsilon}_m$ (for $n, m = 1, 2, \ldots, N$) and setting a matrix $\mathbf{A}^\dagger$ to represent $\mathbb{A}^\dagger$. Equation (1.57) then tells us that

$$(\mathbf{A}^\dagger)_{nm} = (\boldsymbol{\varepsilon}_n, \mathbb{A}^\dagger \boldsymbol{\varepsilon}_m) = (\mathbb{A}\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_m) = (\boldsymbol{\varepsilon}_m, \mathbb{A}\boldsymbol{\varepsilon}_n)^* = A_{mn}^*, \tag{1.58}$$

so that $\mathbf{A}^\dagger$ is indeed, as the notation suggested, the adjoint (transposed conjugate) of the matrix $\mathbf{A}$. Now, Eq. (1.57) is independent of the basis used to describe the space, so this property is independent of the matrix realization of the operator (see Exercise 1.9).

1.5.2. Self-Adjointness and Hermiticity

One particularly important class of operators is comprised of those which are equal to their adjoints, i.e.,

$$\mathbb{H}^\dagger = \mathbb{H}. \tag{1.59}$$

Such operators are called *self-adjoint*. They are represented by *hermitian* matrices $\mathbf{H}^\dagger = \mathbf{H}$. (The distinction between hermiticity and self-adjointness may be a matter of semantics for finite-dimensional spaces $\mathscr{V}^N$; it becomes important, though, for infinite-dimensional ones.)

1.5.3. The Second-Difference Operator $\mathbb{\Delta}$

The operator which will occupy us through most of Chapter 2 is the *second-difference* operator $\mathbb{\Delta}$ whose representing matrix in the $\boldsymbol{\varepsilon}$-basis is

$$\boldsymbol{\Delta} := \begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2 & 1 & & & \\ 0 & 0 & 1 & -2 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & & & \ddots & -2 & 1 \\ 1 & 0 & \cdots & & 0 & 1 & -2 \end{pmatrix}. \tag{1.60}$$

Again, as the notation suggests, $\mathbb{\Delta}$ is the finite-dimensional analogue of the Laplacian. As $\boldsymbol{\Delta}$ is manifestly hermitian (as well as real), $\mathbb{\Delta}$ is self-adjoint.

1.5.4. The $\mathbb{\Delta}$ Operator Represented in the $\boldsymbol{\varphi}$-Basis

The matrix $\tilde{\boldsymbol{\Delta}}$ representing $\mathbb{\Delta}$ in the $\boldsymbol{\varphi}$-basis can be found from (1.39), (1.48), and (1.60) as

$$\begin{aligned}\tilde{\Delta}_{mn} &= (\mathbf{F}^\dagger \boldsymbol{\Delta} \mathbf{F})_{mn} = N^{-1} \sum_{j,k} \Delta_{jk} \exp[2\pi i(kn - jm)/N] \\ &= N^{-1} \sum_k (-2\exp[2\pi ik(n-m)/N] + \exp\{2\pi i[kn - (k+1)m]/N\} \\ &\qquad + \exp\{2\pi i[kn - (k-1)m]/N\}).\end{aligned} \tag{1.61a}$$

The last step uses (1.60) explicitly. From the second and third summands we can extract factors $\exp(\mp 2\pi im/N)$, respectively, so that

$$\begin{aligned}\tilde{\Delta}_{mn} &= N^{-1}[-2 + \exp(-2\pi im/N) + \exp(2\pi im/N)] \sum_k \exp[2\pi ik(n-m)/N] \\ &= [-2 + 2\cos(2\pi m/N)]\delta_{m,n}\end{aligned} \tag{1.61b}$$

[see Eq. (1.49)]; hence $\tilde{\boldsymbol{\Delta}}$ is seen to be a *diagonal matrix*:

$$\tilde{\Delta}_{mn} = \delta_{m,n}\lambda_m \tag{1.62a}$$

$$\lambda_m := -4\sin^2(\pi m/N). \tag{1.62b}$$

In fact, the usefulness of the $\boldsymbol{\varphi}$-basis is the property that $\mathbb{\Delta}$ is represented in it by a diagonal matrix. This will be seen time and again in the following sections. At the end of Section 1.7 we indicate how $\mathbf{F}$ is found by asking for the property (1.62a).

Exercise 1.20. If the linear operator $\mathbb{A}$ is represented in the $\boldsymbol{\varepsilon}$- and $\boldsymbol{\varphi}$-bases by matrices $\mathbf{A}$ and $\tilde{\mathbf{A}}$, and similarly for $\mathbb{B}$, show that

$$\tilde{\mathbf{A}} + \tilde{\mathbf{B}} = \widetilde{(\mathbf{A} + \mathbf{B})}, \tag{1.63a}$$

$$\tilde{\mathbf{A}}\tilde{\mathbf{B}} = \widetilde{\mathbf{A}\mathbf{B}}. \tag{1.63b}$$

Exercise 1.21. Show that the product of two hermitian matrices is not necessarily hermitian. The set of these matrices therefore does *not* form a group.

Exercise 1.22. Show that the matrix $\boldsymbol{\Delta}$ in (1.60) has zero determinant and hence $\boldsymbol{\Delta}^{-1}$ does not exist (nor does $\mathbb{\Delta}^{-1}$). This can be done using the fact that, in (1.62b), $\lambda_N = 0$. Which subspace of $\mathscr{V}^N$ is mapped on $\mathbf{0}$?

Exercise 1.23. Show that the matrix representing $\mathbb{\Delta}^2$ in the $\boldsymbol{\varepsilon}$-basis is

$$\boldsymbol{\Delta}^2 = \begin{pmatrix} 6 & -4 & 1 & 0 & \cdots & 0 & 1 & -4 \\ -4 & 6 & -4 & 1 & \cdots & 0 & 0 & 1 \\ 1 & -4 & 6 & -4 & \ddots & & & 0 \\ 0 & 1 & -4 & 6 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 1 & 0 \\ 0 & 0 & & \ddots & \ddots & 6 & -4 & 1 \\ 1 & 0 & & & 1 & -4 & 6 & -4 \\ -4 & 1 & 0 & \cdots & 0 & 1 & -4 & 6 \end{pmatrix}. \tag{1.64}$$

Which matrix represents $\mathbb{\Delta}^2$ in the $\boldsymbol{\varphi}$-basis?

Exercise 1.24. Show that, for $2p + 1 \leqslant N$, $\mathbb{\Delta}^p$ is represented in the $\boldsymbol{\varepsilon}$-basis by the matrix $\boldsymbol{\Delta}^p$ with elements

$$(\boldsymbol{\Delta}^p)_{mn} = (-1)^{m-n+p}\binom{2p}{p + m - n} = (\boldsymbol{\Delta}^p)_{m,N-n} = (\boldsymbol{\Delta}^p)_{N-m,n}, \tag{1.65}$$

where $\binom{r}{s}$ is the binomial coefficient. Verify the cases $p = 0$, $p = 1$ [Eq. (1.60)], and $p = 2$ [Eq. (1.64)]. Show that $\widetilde{\boldsymbol{\Delta}^p} = (\tilde{\boldsymbol{\Delta}})^n$ represents $\mathbb{\Delta}^p$ in the $\boldsymbol{\varphi}$-basis, and find its elements.

Exercise 1.25. Consider the *projection* operator $\mathbb{P}_k$ which maps every vector in $\mathscr{V}^N$ to its projection along $\boldsymbol{\varepsilon}_k$. Show that this is represented in the $\boldsymbol{\varepsilon}$- and $\boldsymbol{\varphi}$-bases by matrices with elements

$$(\mathbf{P}_k)_{mn} = \delta_{mn}\delta_{kn}, \tag{1.66a}$$

$$(\tilde{\mathbf{P}}_k)_{mn} = N^{-1}\exp[2\pi i k(m - n)/N]. \tag{1.66b}$$

These operators are self-adjoint.

Exercise 1.26. From the projection operators of Exercise 1.25, build (for N even) $\mathbb{E}_1 := \sum_{k=1}^{N/2} \mathbb{P}_{2k-1}$ and $\mathbb{E}_2 := \sum_{k=1}^{N/2} \mathbb{P}_{2k}$. Show that these are represented in the $\boldsymbol{\varepsilon}$- and $\boldsymbol{\varphi}$-bases by

$$\mathbf{E}_1 = \begin{pmatrix} 1 & & & & & 0 \\ & 0 & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 1 \\ 0 & & & & & 0 \end{pmatrix}, \qquad \tilde{\mathbf{E}}_1 = \begin{pmatrix} \tfrac{1}{2}\mathbf{1} & -\tfrac{1}{2}\mathbf{1} \\ -\tfrac{1}{2}\mathbf{1} & \tfrac{1}{2}\mathbf{1} \end{pmatrix}, \tag{1.67a}$$

$$\mathbf{E}_2 = \begin{pmatrix} 0 & & & & & 0 \\ & 1 & & & & \\ & & 0 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 0 \\ 0 & & & & & 1 \end{pmatrix}, \qquad \tilde{\mathbf{E}}_2 = \begin{pmatrix} \tfrac{1}{2}\mathbf{1} & \tfrac{1}{2}\mathbf{1} \\ \tfrac{1}{2}\mathbf{1} & \tfrac{1}{2}\mathbf{1} \end{pmatrix}. \tag{1.67b}$$

These operators will be used in Chapter 2.

1.5.5. Functions of Matrices and Operators

The results (1.35) to (1.38) linking linear combination and product of operators with the corresponding operations between the representing matrices show that if $P_n(x)$ is a polynomial of degree n in x, then the operator $\mathbb{B} = P_n(\mathbb{A})$ is well defined and represented by the matrices $\mathbf{B} = P_n(\mathbf{A})$ and $\tilde{\mathbf{B}} = P_n(\tilde{\mathbf{A}}) = \widetilde{P_n(\mathbf{A})}$, which are also well defined. [The last equality is a direct consequence of (1.63).] Now, if we allow n to grow without bound into an *analytic function* $P(x)$ expressible as a Taylor series convergent in some region $|x| < \rho$, what happens with $P(\mathbf{A})$? If the result is a well-defined matrix, this will give us a working definition for the operator $P(\mathbb{A})$. The functions $P(x)$ we shall use later are of the type $\exp(ax)$, $\cosh bx$, $(\sinh x)/x$, while others such as $(1 - x)^{-1}$ are quite commonly used [see, for example, the book by Goertzel and Tralli (1960).] Thus assume that

$$P(x) = \sum_{n=0}^{\infty} p_n x^n \tag{1.68}$$

converges for $|x| < \rho$. Consider now the same series, replace x by the matrix $\mathbf{A}$, and let α be the maximum of the absolute values of its matrix elements: $|A_{nm}| \leqslant \alpha$. Then a bound for the matrix elements of $\mathbf{A}^2$ will be $N\alpha^2$ (the equality will happen when a row of maximal elements α of **A** meets a column of similar elements). Inductively, we see that a bound for the elements of $\mathbf{A}^n$ is $N^{-1}(N\alpha)^n$, and hence *a bound for the elements of* $P(\mathbf{A})$ *is* $N^{-1}P(N\alpha)$ *so that for* $\alpha < \rho/N$, $P(\mathbf{A})$ *exists.* For the exponential and hyperbolic functions, ρ is infinite, so $P(\mathbf{A})$ exists for any **A**. Correspondingly, the operator $P(\mathbb{A})$ is defined. Note that if the constants p_n in (1.68) are real, as $(\mathbb{A}^\dagger)^n = (\mathbb{A}^n)^\dagger$, it follows that $P(\mathbb{A}^\dagger) = P(\mathbb{A})^\dagger$. Hence, if $\mathbb{H}$ is self-adjoint, $P(\mathbb{H})$ will be also.

1.5.6. Multiplication and Commutation

Although we can comfortably work with functions of matrices and operators, we have to be careful about their *composition*, since the rules of ordinary algebra may not apply when several matrices or operators are involved. Consider the well-known relation $\exp(a + b) = \exp a \cdot \exp b$ for numbers a and b. The direct proof proceeds as follows:

$$\exp(a+b) = \sum_{n=0}^{\infty} \frac{1}{n!}(a+b)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

$$= \sum_{m=0}^{\infty} \frac{1}{m!} a^m \sum_{l=0}^{\infty} \frac{1}{l!} b^l = \exp a \cdot \exp b, \tag{1.69}$$

where in the third step we used a double-sum exchange relation. See Appendix C. The *second* step [expanding the binomial $(a + b)^n$] is only true, however, if a and b commute, i.e., $ab = ba$, so that all powers of a can be put to the left and those of b to the right. This is *not* true for matrices **A** and **B**, which do not commute. [For a thorough treatment of such problems, called Baker–Campbell–Hausdorff relations, see the article by Mielnik and Plebañski (1970).] The most we can say here is that Eq. (1.69) is valid for matrices **A** and **B** when these commute, as is the case when they are *both diagonal matrices* or when **B** is a *multiple* of **A**, i.e.,

$$\exp(a\mathbf{A})\cdot\exp(b\mathbf{A}) = \exp[(a+b)\mathbf{A}]. \tag{1.70}$$

This relation will be used often.

1.5.7. Diagonalization and Exponentiation of Various Matrices

The problem of actually exponentiating or obtaining any function of a matrix **A** is another matter. If **A** represents an operator $\mathbb{A}$, it may be that in some basis $\mathbb{A}$ is represented by a diagonal matrix $\overline{\mathbf{A}}$. This is the case for $\mathbb{\Delta}$, represented by $\mathbf{\Delta}$ in the $\boldsymbol{\varepsilon}$-basis [Eq. (1.60)] and by a diagonal matrix $\tilde{\mathbf{\Delta}}$ in the $\boldsymbol{\varphi}$-basis [Eq. (1.62)]. For self-adjoint and unitary operators this is developed in Section 1.7 in some detail. Since any sum or power of diagonal matrices is a diagonal matrix, its elements are given by the sum or powers of the original diagonal elements. If $\mathbf{V}$ is a transformation relating the representing matrices $\mathbf{A}$ and the diagonal $\bar{\mathbf{A}}$ of $\mathbb{A}$ [see Eq. (1.37)], then $P(\mathbf{A})$ can be explicitly calculated by

$$P(\mathbf{A}) = \mathbf{V}\mathbf{V}^{-1}P(\mathbf{A})\mathbf{V}\mathbf{V}^{-1} = \mathbf{V}\left(\sum_{n=0}^{\infty} p_n \mathbf{V}^{-1}\mathbf{A}^n\mathbf{V}\right)\mathbf{V}^{-1}$$

$$= \mathbf{V}\left[\sum_{n=0}^{\infty} p_n(\mathbf{V}^{-1}\mathbf{A}\mathbf{V})^n\right]\mathbf{V}^{-1} = \mathbf{V}P(\bar{\mathbf{A}})\mathbf{V}^{-1}. \tag{1.71}$$

Exercise 1.27. Using Eq. (1.71) as well as (1.60) and (1.62), show that

$$\mathbb{G}^{1,0,1}(\tau) := \exp(\tau\mathbb{\Delta}) \tag{1.72a}$$

is represented in the $\boldsymbol{\varphi}$- and $\boldsymbol{\varepsilon}$-bases by

$$\tilde{G}^{1,0,1}_{m,n}(\tau) = [\exp(\tau\tilde{\boldsymbol{\Delta}})]_{mn} = \delta_{m,n}\exp(\tau\lambda_m), \tag{1.72b}$$

$$G^{1,0,1}_{m,n}(\tau) = [\exp(\tau\boldsymbol{\Delta})]_{mn} = N^{-1}\sum_k \exp(\tau\lambda_k)\exp[2\pi ik(n-m)/N], \tag{1.72c}$$

$$\lambda_m = -4\sin^2(\pi m/N). \tag{1.72d}$$

The operator $\mathbb{G}^{1,0,1}(\tau)$ will appear in Exercise 2.17 as the time-evolution operator for the finite-difference analogue of the heat equation.

Exercise 1.28. From (1.60) it is obvious that $\sum_n \Delta_{nm} = 0$ for $m = 1, 2, \ldots, N$. Prove that $\sum_n (\boldsymbol{\Delta}^p)_{nm} = 0$ and

$$\sum_n [\exp(\tau\boldsymbol{\Delta})]_{mn} = 1, \qquad m = 1, 2, \ldots, N. \tag{1.73}$$

Note that this property holds only for the $\boldsymbol{\varepsilon}$-basis, where $\mathbb{\Delta}$ is represented by $\boldsymbol{\Delta}$. Innocuous as it seems, Eq. (1.73) will lead to the (discrete analogue of) total heat conservation (Exercise 2.18).

Exercise 1.29. Noting that $(\sinh x)/x$ contains only even powers of x in its Taylor expansion, define

$$\mathbb{G}^{0,1,1}(\tau) = \mathbb{\Delta}^{-1/2}\sinh\tau\mathbb{\Delta}^{1/2}. \tag{1.73a}$$

Show that this is represented in the $\boldsymbol{\varphi}$- and $\boldsymbol{\varepsilon}$-bases by

$$\tilde{G}^{0,1,1}_{m,n}(\tau) = \delta_{mn}\frac{1}{\omega_m}\sin\omega_m\tau \tag{1.73b}$$

$$G^{0,1,1}_{m,n}(\tau) = N^{-1}\sum_k \omega_k^{-1}\sin\omega_k\tau\exp[2\pi ik(n-m)/N] \tag{1.73c}$$

$$\omega_m = (-\lambda_m)^{1/2} = 2\sin(\pi m/N). \tag{1.73d}$$

This operator will appear in Section 2.2. It is the time-evolution operator for the finite difference analogue of the wave equation.

Exercise 1.30. Prove that if $\mathbf{H}$ is a hermitian matrix, it *generates* a set of unitary matrices

$$\mathbf{U}(\tau) = \exp(i\tau\mathbf{H}) \tag{1.74}$$

with τ real. Define thus *unitary operators*. They have the property

$$(\mathbb{U}\mathbf{f}, \mathbb{U}\mathbf{g}) = (\mathbf{f}, \mathbf{g}) \tag{1.75}$$

for all $\mathbf{f}, \mathbf{g} \in \mathscr{V}^N$.

Exercise 1.31. Let $\mathbf{A}(\tau)$ be a matrix whose elements are differentiable functions of τ. Define the derivative of $\mathbf{A}(\tau)$ with respect to τ as

$$\frac{d}{d\tau}\mathbf{A}(\tau) = \lim_{\varepsilon\to 0} \varepsilon^{-1}[\mathbf{A}(\tau + \varepsilon) - \mathbf{A}(\tau)] = \mathbf{A}'(\tau). \tag{1.76}$$

From this see that matrix differential calculus is similar to ordinary calculus. In particular, the Leibnitz rule

$$\frac{d}{d\tau}(\mathbf{AB}) = \mathbf{A}'\mathbf{B} + \mathbf{AB}', \qquad \mathbf{A} = \mathbf{A}(\tau), \qquad \mathbf{B} = \mathbf{B}(\tau) \tag{1.77}$$

holds. As commutativity does not hold, we must keep straight the order of the factors.

Exercise 1.32. Let $\mathbf{A}^{-1}(\tau)$ be the matrix inverse to $\mathbf{A}(\tau)$. Show that

$$\frac{d}{d\tau}\mathbf{A}^{-1} = -\mathbf{A}^{-1}\mathbf{A}'\mathbf{A}^{-1}. \tag{1.78}$$

Exercise 1.33. Regarding (1.74), show that

$$\mathbf{H} = -i\frac{d}{d\tau}\mathbf{U}(\tau)|_{\tau=0}. \tag{1.79}$$

If $\mathbf{U}(\tau)$ is a unitary matrix, show that $\mathbf{H}$ is hermitian. You will be using Eq. (1.78).

Exercise 1.34. Show that *every linear operator* $\mathbb{A}$ *can be written as*

$$\mathbb{A} = \mathbb{H}_1 + i\mathbb{H}_2, \tag{1.80}$$

where $\mathbb{H}_1$ *and* $\mathbb{H}_2$ *are self-adjoint.* Equation (1.80) recalls the decomposition of an arbitrary complex number into a real and an imaginary part.

1.6. The Dihedral Group

In Section 1.4 we introduced the permutation group π_N and saw some of its properties. Here we shall study a subset of π_N which constitutes a group by itself, called the dihedral group D_N, which will be seen to mesh in interesting ways with the Fourier transform and the $\mathbb{\Delta}$ operator, and will be used extensively later on.

1.6.1. Rotations and Inversions of a Finite, Closed Lattice

Consider the two permutations $\mathbb{R}$ and $\mathbb{I}_0$, represented by the matrices [Eq. (1.45)]

$$\mathbb{R}:\quad \mathbf{R} := \mathbf{P}\begin{pmatrix} 1 & 2 & \cdots & N-1 & N \\ 2 & 3 & \cdots & N & 1 \end{pmatrix} = \begin{pmatrix} 0 & & & 1 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix}, \tag{1.81}$$

$$\mathbb{I}_0:\quad \mathbf{I}_0 := \mathbf{P}\begin{pmatrix} 1 & 2 & \cdots & N-1 & N \\ N-1 & N-2 & \cdots & 1 & N \end{pmatrix} = \begin{pmatrix} 0 & & 1 & 0 \\ & \cdot^{\cdot^{\cdot}} & & \\ 1 & & & \\ 0 & & & 1 \end{pmatrix}. \tag{1.82}$$

[Note that $\mathbf{I}_0$ has already appeared in (1.54).] The concepts we shall present here can be illustrated as applied to a finite lattice of N masses joined by pairs through springs, as shown in Fig. 1.4(a). Assume the lattice is vibrating. Although the precise description of the time development of the system will be undertaken in Chapter 2, this system will serve to apply the ideas involved.

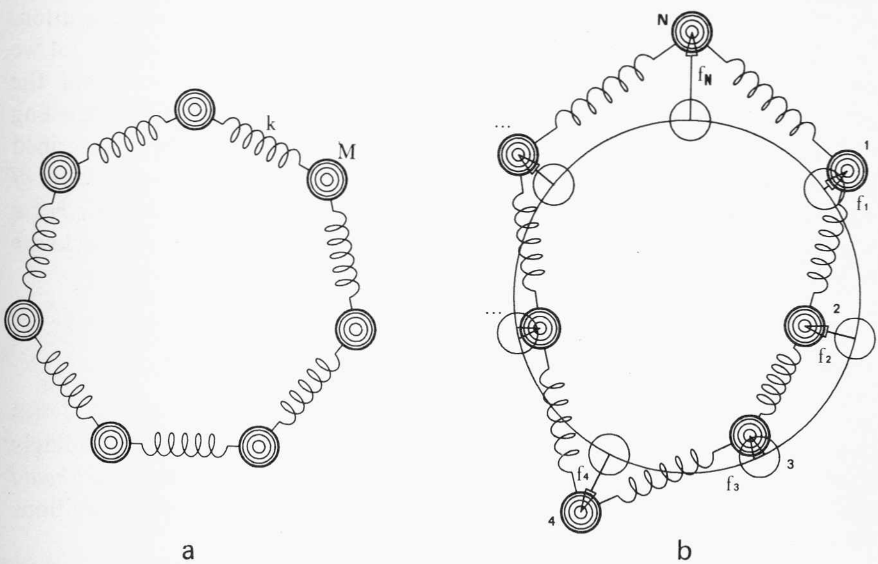


Fig. 1.4. (a) A linear, closed *lattice* constituted by masses M and springs k. (b) The same lattice undergoing vibrational motion. The time-dependent coordinates $f_n(t)$, $n = 1, 2, \ldots, N$, of the masses define the coordinates of the state vector $\mathbf{f}(t)$ of the system.

Let $f_n(t)$ be the elongation of the nth mass at time t, and construct the time-dependent N-dimensional vector $\mathbf{f}(t) = \sum_n f_n(t)\boldsymbol{\varepsilon}_n$, which will be referred to as the *state* vector describing the system. The components of the state vector $\mathbf{f}$ are shown as the arrows indicating the elongations of the vibrating lattice, for some fixed time t, in Fig. 1.4(b). We can express $\mathbf{f}(t)$ in a new basis $\{\bar{\boldsymbol{\varepsilon}}_n\}_{n=1}^N$ by the use of a (passive) transformation which for (1.81) is $\bar{\boldsymbol{\varepsilon}}_n = \boldsymbol{\varepsilon}_{n+1}$. Here and in what follows it will serve us to consider, as Fig. 1.4 suggests, that mass number $N + 1$ is the same as mass number 1, $N + 2$ the same as 2, etc., thus letting the component label n be numbered, as before, modulo N, so that the statement $\bar{\boldsymbol{\varepsilon}}_n = \boldsymbol{\varepsilon}_{n+1}$ implies, in particular, $\bar{\boldsymbol{\varepsilon}}_N = \boldsymbol{\varepsilon}_1$. In the $\bar{\boldsymbol{\varepsilon}}$-basis, $\mathbf{f}(t)$ has its coordinates given by $\bar{f}_n(t) = f_{n-1}(t)$, $n \equiv n \bmod N$. Insofar as Fig. 1.4 is concerned, the same shape $\mathbf{f}(t)$ of the lattice is described by a relabeling of the masses which shifts the old labels clockwise by one unit, while the elongations $\bar{f}_n(t)$ are correspondingly shifted counterclockwise by one unit.

1.6.2. Producing New Solutions from Old Ones

Now consider (1.81) and (1.82) as matrices representing in the $\boldsymbol{\varepsilon}$-basis *active* transformations in $\mathscr{V}^N$, $\mathbb{R}$, and $\mathbb{I}_0$. In this case, under (1.81), again $\boldsymbol{\varepsilon}'_n = \boldsymbol{\varepsilon}_{n+1}$, but the lattice is now bodily moved clockwise by one unit, and the new state vector $\mathbf{f}'(t) = \sum_n f_n(t)\boldsymbol{\varepsilon}'_n = \mathbb{R}\mathbf{f}(t)$ will describe its time evolution. It is quite obvious, however, that in applying $\mathbb{R}$ to the system in Fig. 1.4 we have preserved the neighbor relation between the masses and that the original and the rotated lattices are indistinguishable except for our labeling of the masses. The physical lattices are the same and thus should be described by the same equations of motion. What we have done then is *to produce out of the state vector* $\mathbf{f}(t)$ *a new state vector* $\mathbf{f}'(t) = \mathbb{R}\mathbf{f}(t)$ *which also describes a possible vibration state for Fig. 1.4*, which is the old solution rotated clockwise by one unit.

1.6.3. Invariance of the Equations of Motion

Any transformation which maps the undeformed lattice in Fig. 1.4(a) onto itself (*invariance* transformations of the figure) will correspondingly *produce a new solution vector* $\mathbf{f}'(t)$ *out of any given solution* $\mathbf{f}(t)$ *and should leave the equations of motion invariant*, changing only the initial conditions which determine their subsequent time development.

The *dihedral transformations are the largest set* D_N *of permutations leaving Fig. 1.4(a) invariant* as described and will *constitute a group* since, again quite clearly, the successive application of two invariance transformations is a transformation which also leaves the figure invariant, and the identity transformation $E = P\binom{n}{n}$ is an element in this set. These two observations state that the set of invariance transformations of a system satisfies the group axioms (a) and (c) (Section 1.4.2). Now axiom (b) (associativity) is satisfied since every element of D_N is within π_N, where this property holds. Last, since π_N has only a finite number of elements, for any $\mathbb{T} \in D_N$ we can construct the successive powers $\mathbb{T}, \mathbb{T}^2, \mathbb{T}^3, \ldots \in D_N$ and eventually reach $\mathbb{T}^p = \mathbb{1}$, the identity transformation, so $\mathbb{T}^{p-1} = \mathbb{T}^{-1} \in D_N$ and axiom (d) is also satisfied. Hence $D_N \subset \pi_N$ is a group by itself and a subgroup of π_N.

Exercise 1.35. Show that the successive powers $\mathbb{T}, \mathbb{T}^2, \mathbb{T}^3, \ldots$ cannot enter into a "loop" without involving the identity element.

1.6.4. Multiplying Rotations and Inversions for N Odd

Let N be an odd number and consider the two permutations $\mathbb{R}$ and $\mathbb{I}_0$ represented by (1.81) and (1.82). We saw that $\mathbb{R}$ effects a clockwise rotation by $2\pi/N$. See Fig. 1.5a. Similarly, $\mathbb{I}_0$ reflects the figure across a line which passes through the Nth mass and the midpoint of the spring joining the $[\frac{1}{2}(N - 1)]$th and the $[\frac{1}{2}(N + 1)]$th masses. They are invariance transformations of the figure. Applying $\mathbb{R}$ k times in succession, we see that $\mathbb{R}^k$ performs a rotation by $2\pi k/N$ and that $\mathbb{R}^N = \mathbb{1}$. Under $\mathbb{R}^k$, the mth mass is brought onto the $(m + k)$th mass. We can use the shorthand

$$\mathbb{R}^k[m] = [m + k]. \tag{1.83}$$

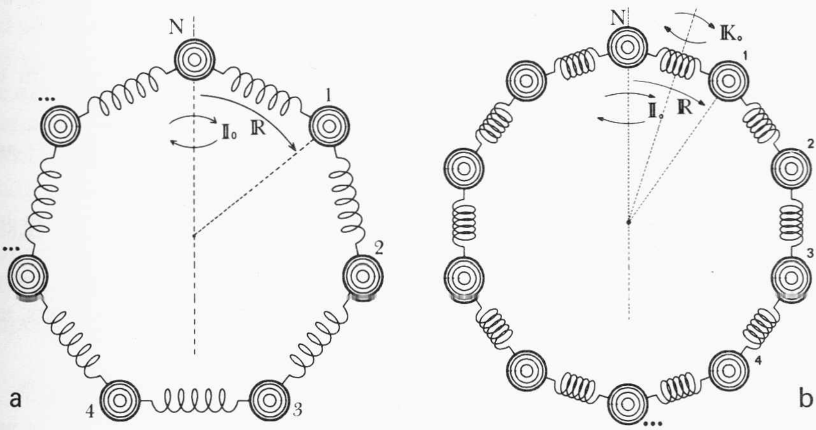


Fig. 1.5. (a) Dihedral group symmetries for an N-mass lattice when N is odd: rotations $\mathbb{R}$ and inversions $\mathbb{I}_0$ through a mass center. (b) Dihedral group symmetries for an N-mass lattice when N is even: In addition to rotations $\mathbb{R}$ and inversions $\mathbb{I}_0$ through a mass center, we have the possibility of inversions $\mathbb{K}_0$ through a spring midpoint.

Next consider $\mathbb{I}_0$, leaving mass N in its place; i.e.,

$$\mathbb{I}_0[m] = [N - m] \tag{1.84}$$

and $[N] \equiv [0] \bmod N$. It is not difficult to conclude that

$$\mathbb{I}_k := \mathbb{R}^k \mathbb{I}_0 \mathbb{R}^{-k}, \tag{1.85}$$

where $\mathbb{R}^{-k} := (\mathbb{R}^{-1})^k$ is a reflection of the figure which leaves mass k invariant, since $\mathbb{R}^{-k}$ maps mass k into the position N left invariant by $\mathbb{I}_0$ and $\mathbb{R}^k$ maps this back to its original place. This can be verified using the shorthand (1.83)–(1.84):

$$\begin{aligned} \mathbb{I}_k[m] &= \mathbb{R}^k \mathbb{I}_0 \mathbb{R}^{-k}[m] = \mathbb{R}^k \mathbb{I}_0[m - k] = \mathbb{R}^k[N - m + k] \\ &= [N + 2k - m]. \end{aligned} \tag{1.86}$$

The products of an $\mathbb{R}$ and an $\mathbb{I}$ or of two $\mathbb{I}$'s can be easily calculated in the same way.

Exercise 1.36. For N odd, prove that the $2N$ elements of D_N satisfy the "multiplication" table:

$$\mathbb{R}^k \mathbb{R}^l = \mathbb{R}^{k+l}, \tag{1.87a}$$

$$\mathbb{R}^k \mathbb{I}_l = \begin{cases} \mathbb{I}_{l+1/2(k)}, & k \text{ even}, \\ \mathbb{I}_{l+1/2(N+k)}, & k \text{ odd}, \end{cases} \tag{1.87b}$$

$$\mathbb{I}_l \mathbb{R}^k = \begin{cases} \mathbb{I}_{l-1/2(k)}, & k \text{ even}, \\ \mathbb{I}_{l-1/2(N+k)}, & k \text{ odd}, \end{cases} \tag{1.87c}$$

$$\mathbb{I}_k \mathbb{I}_l = \mathbb{R}^{2(k-l)}. \tag{1.87d}$$

Exercise 1.37. Show that the set $C_N := \{\mathbb{1}, \mathbb{R}, \mathbb{R}^2, \ldots, \mathbb{R}^{N-1}\}$ is a subgroup of D_N. It is called the *cyclic* group of N elements.

Exercise 1.38. Show that the operators $\mathbb{R}^k$ and $\mathbb{I}_l$ are represented, in the $\boldsymbol{\varepsilon}$-basis, by the matrices

$$\mathbb{R}^k = \begin{pmatrix} \mathbf{0} & \mathbf{1}_k \\ \mathbf{1}_{N-k} & \mathbf{0} \end{pmatrix}, \tag{1.88a}$$

$$\mathbb{I}_l = \begin{pmatrix} \mathbf{1}^A_{N-l-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}^A_{l+1} \end{pmatrix}, \tag{1.88b}$$

where $\mathbf{1}_p$ is the unit $p \times p$ matrix, $\mathbf{1}_q{}^A$ is the unit antidiagonal $q \times q$ matrix, and the $\mathbf{0}$'s are adequate rectangular null matrices. Compare with (1.81), its powers, and (1.82).

Exercise 1.39. Verify that the elements of (1.88) can be written as

$$(\mathbf{R}^k)_{mn} = \delta_{m,n+k}, \tag{1.89a}$$

$$(\mathbf{I}_l)_{mn} = \delta_{m,N+2l-n}, \tag{1.89b}$$

recalling that row (and column) labels are to be considered modulo N. In particular, matrices (1.88)–(1.89) acting on row vectors should transform the entries according to (1.83) and (1.86).

1.6.5. Representative Matrices in the Fourier Basis

Since the elements of D_N are now operators, we can ask for their representing matrices in the $\boldsymbol{\varphi}$-basis. Indeed, using (1.39), the Fourier transform (1.48), and (1.89), we find

$$\begin{aligned}(\tilde{\mathbf{R}})_{mn} &= (\mathbf{F}^{\dagger}\mathbf{R}\mathbf{F})_{mn} \\ &= N^{-1} \sum_{j,l} \delta_{j,l+1} \exp[2\pi i(jm - ln)/N] \\ &= N^{-1} \exp(2\pi i m/N) \sum_{l} \exp[2\pi i l(m - n)/N]. \end{aligned} \tag{1.90}$$

Hence, $\widetilde{\mathbf{R}^k} = \tilde{\mathbf{R}}^k$ is a diagonal matrix,

$$(\tilde{\mathbf{R}}^k)_{mn} = \delta_{m,n} \exp(2\pi i k m/N). \tag{1.91}$$

Similarly, we can show that $\tilde{\mathbf{I}}_l$ is antidiagonal:

$$(\tilde{\mathbf{I}}_l)_{mn} = \delta_{m,N-n} \exp(4\pi i l m/N), \tag{1.92}$$

and in particular

$$\tilde{\mathbf{I}}_0 = \mathbf{I}_0, \tag{1.93}$$

which is obvious from (1.54). All the operators in D_N are unitary since the matrices of π_N-transformations are. In addition, all $\mathbb{I}_l$ are self-adjoint.

In performing the calculations leading to (1.91) and (1.92) we can see a real advantage in treating the row and column indices modulo N, since we can automatically keep track of $n \equiv n + mN \bmod N$ and $-n \equiv N - n \bmod N$. We are dealing with matrices $M_{mn} = f(m, n)$ which are periodic functions of m and n of period N in both variables. This property holds not only for the matrices representing operators in D_N but for the matrix $\mathbf{F}$ representing the Fourier transform (1.48) as well as $\boldsymbol{\Delta}$ and $\tilde{\boldsymbol{\Delta}}$ [Eqs. (1.60) and (1.62)] representing $\mathbb{\Delta}$. Since we are identifying the $(N + k)$th row with the kth one and similarly for columns, we are actually bending any such matrix into a torus. The linear combination and product of any two such matrices have the same property.

Exercise 1.40. Show that if $\mathbf{g} = \mathbb{R}^k\mathbf{f}$, then their coordinates relate as

$$g_n = f_{n-k}, \qquad \tilde{g}_n = \exp(2\pi i k n/N)\tilde{f}_n, \tag{1.94}$$

while if $\mathbf{h} = \mathbb{I}_0\mathbf{f}$,

$$h_n = f_{N-n}, \qquad \tilde{h}_n = \tilde{f}_{N-n}. \tag{1.95}$$

Out of (1.94) and (1.95) you can find the coordinates of $\mathbb{I}_l\mathbf{f}$.

1.6.6. Invariance of $\mathbb{\Delta}$ under the Dihedral Group

One last property we want to point out for the operators in D_N is that they all commute with the operator $\mathbb{\Delta}$ introduced in Section 1.5.3, i.e.,

$$\mathbb{\Delta}\mathbb{R}^k = \mathbb{R}^k\mathbb{\Delta}, \tag{1.96a}$$

$$\mathbb{\Delta}\mathbb{I}_l = \mathbb{I}_l\mathbb{\Delta}. \tag{1.96b}$$

Equation (1.96a) can easily be proven using the representatives of $\mathbb{\Delta}$ and $\mathbb{R}^k$ in the $\boldsymbol{\varphi}$-basis, which are diagonal matrices [Eqs. (1.62) and (1.91)], since all diagonal matrices commute among themselves. Equation (1.96b) can be proven for $l = 0$, noting that for any matrix $\mathbf{A}$, $(\mathbf{I}_0\mathbf{A}\mathbf{I}_0)_{mn} = A_{N-m,N-n}$. For arbitrary l, $\mathbf{I}_l$ can be written in terms of $\mathbf{R}^l$ and $\mathbf{I}_0$ by (1.85) and the equation can be proven thence.

Exercise 1.41. Show that any operator *function* of $\mathbb{\Delta}$ will also commute with operators in the dihedral group.

1.6.7. *N* Even

When the number N of masses in a lattice is even, then, in addition to the $\mathbb{I}_l$-transformations which leave masses l and $\frac{1}{2}N + l$ in their places, we can perform transformations which invert the lattice with respect to the centers of two opposite springs so that no mass is left in its place (see Fig. 1.5b). We thus define the operator $\mathbb{K}_0$ by its action on the lattice masses:

$$\mathbb{K}_0[m] = [N - m + 1]. \tag{1.97}$$

Note that $\mathbb{K}_0[1] = [N]$, $\mathbb{K}_0[\frac{1}{2}N] = [\frac{1}{2}N + 1]$, and $(\mathbb{K}_0)^2 = \mathbb{1}$. In analogy with (1.85) we can define $\mathbb{K}_l = \mathbb{R}^l\mathbb{K}_0\mathbb{R}^{-l}$, which reflects through the midpoints of the springs joining masses $[l]$ and $[l + 1]$ and masses $[\frac{1}{2}N + l]$ and $[\frac{1}{2}N + l + 1]$.

Exercise 1.42. Construct the multiplication table of $\mathbb{K}$'s, $\mathbb{R}$'s, and $\mathbb{I}$'s. Among others, show the relations

$$\mathbb{K}_k\mathbb{K}_l = \mathbb{R}^{2(k-l)}, \qquad \mathbb{K}_k\mathbb{I}_l = \mathbb{R}^{2(k-l)+1}, \qquad \mathbb{I}_l\mathbb{K}_k = \mathbb{R}^{2(l-k)-1}. \tag{1.98}$$

In particular, if N is a multiple of 2 but *not* a multiple of 4, note that $\mathbb{K}_0$ and $\mathbb{I}_{(N+2)/4}$ commute. Why?

Exercise 1.43. Show that the matrices representing $\mathbb{K}_0$ in the $\boldsymbol{\varepsilon}$- and $\boldsymbol{\varphi}$-bases are hermitian and unitary:

$$(\mathbf{K}_0)_{mn} = \delta_{m,N-n+1}, \qquad (\tilde{\mathbf{K}}_0)_{mn} = \delta_{m,N-n}\exp(2\pi i m/N). \tag{1.99}$$

Show that for N even, $\mathbf{K}_0$ is an antidiagonal matrix all of whose diagonal elements are zero, while $\tilde{\mathbf{K}}_0$ has the same shape as $\tilde{\mathbf{I}}_0$ but different elements. Also show that, as in (1.96),

$$\mathbb{\Delta}\mathbb{K}_l = \mathbb{K}_l\mathbb{\Delta}. \tag{1.100}$$

Exercise 1.44. Show that D_N has $2N$ distinct elements: N rotations (including the identity element) and N inversions. Verify this for N odd as well as even.

1.6.8. Polar Decomposition of Operators

We make one last remark about the role of self-adjoint and unitary operators with respect to the set of all operators in $\mathscr{V}^N$: *every non-singular operator* $\mathbb{A}$ (i.e., such that $\det \mathbf{A} \neq 0$) *can be represented in the form*

$$\mathbb{A} = \mathbb{H}\mathbb{U}, \tag{1.101}$$

where $\mathbb{H}$ *is self-adjoint and* $\mathbb{U}$ *unitary*. For the proof we refer to Gel'fand (1961, Section II-15). Recalling Eq. (1.80), we are reminded by (1.101) of the decomposition of an arbitrary complex number into the product of its modulus (a positive real number) and its phase. The phase itself is the imaginary exponential of a real number. Here, see Eq. (1.74).

1.7. The Axes of a Transformation: Eigenvalues and Eigenvectors

When applying an operator $\mathbb{A}$ to the vectors of $\mathscr{V}^N$, a good insight into the nature of $\mathbb{A}$ is given by the directions in $\mathscr{V}^N$ left invariant by the operator. As we are particularly interested in self-adjoint and unitary operators, we shall develop here the results for these cases. In fact, the knowledge of these transformation axes (and the eigenvalues) specify the operator uniquely.

1.7.1. Invariant Directions

Assume a vector $\mathbf{x} \in \mathscr{V}^N$ is mapped by the action of $\mathbb{A}$ into a multiple of itself:

$$\mathbb{A}\mathbf{x} = \mu\mathbf{x}, \qquad \mu \in \mathscr{C}. \tag{1.102}$$

This only means that the *direction* defined by $\mathbf{x}$ is invariant under the action of $\mathbb{A}$. When an equation such as (1.102) holds, $\mathbf{x}$ is said to be an *eigenvector* of $\mathbb{A}$ with *eigenvalue* μ. This is the problem, for instance, of determining which directions in the $\mathscr{V}^2$-plane in Fig. 1.2 are left invariant by the action of the operator. In a basis where $\mathbb{A}$ is represented by a matrix $\mathbf{A}$, Eq. (1.102) can be written as

$$(\mathbf{A} - \mu\mathbf{1})\mathbf{x} = \mathbf{0}. \tag{1.103}$$

1.7.2. Characteristic Equation

If there exists a nonzero vector $\mathbf{x}$ satisfying (1.102), then (1.103), being a set of N homogeneous equations, requires

$$p_N(\mu) := \det(\mathbf{A} - \mu\mathbf{1}) = 0. \tag{1.104}$$

This is called the *characteristic equation* for $\mathbb{A}$, and $p_N(\mu)$ is its *characteristic polynomial.* As this is an Nth-degree polynomial in μ, we are assured by the fundamental theorem of algebra that there exist exactly N roots μ_i, $i = 1, 2, \ldots, N$, of p_N such that (1.104) holds. Of course some of these can be multiple roots of p_N, but not all of them can be zero, as $p_N(\mu) = \mu^N = 0$ would imply that $\mathbf{A} = \mathbf{0}$ and $\mathbb{A}$ would map all of $\mathscr{V}^N$ into $\mathbf{0}$. The set of eigenvalues is said to be the *spectrum* of the operator. This is a property of the operator, not of the particular matrix representation. This is true as long as the defining bases are all nondegenerate, for suppose we subject the basis in which $\mathbb{A}$ is represented by $\mathbf{A}$ to an invertible transformation $\mathbf{V}$ as given by (1.39). Then, from (1.103) it follows that

$$\mathbf{0} = \mathbf{V}^{-1}(\mathbf{A} - \mu_i\mathbf{1})\mathbf{x}_i = (\mathbf{V}^{-1}\mathbf{A}\mathbf{V} - \mu_i\mathbf{1})\mathbf{V}^{-1}\mathbf{x}_i = (\bar{\mathbf{A}} - \mu_i\mathbf{1})\mathbf{V}^{-1}\mathbf{x}_i = \mathbf{0}, \quad (1.105)$$

i.e., the vectors $\mathbf{V}^{-1}\mathbf{x}_i$ are eigenvectors of $\bar{\mathbf{A}}$ (representing $\mathbb{A}$ in the $\bar{\boldsymbol{\varepsilon}}$-basis) with the *same* eigenvalues μ_i.

1.7.3. Spectrum and Eigenbasis of a Self-adjoint Operator

We consider now the case when $\mathbb{A}$ is a self-adjoint operator $\mathbb{H}^\dagger = \mathbb{H}$. When this happens, we shall prove that (a) *the spectrum of* $\mathbb{H}$ *is real* and (b) *eigenvectors corresponding to different eigenvalues are orthogonal.* Indeed, consider eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ corresponding to eigenvalues μ_1 and μ_2, which are not necessarily distinct. Then, Eq. (1.102) for $\mathbf{x}_1$ in inner product with $\mathbf{x}_2$ yields

$$\mu_1(\mathbf{x}_2, \mathbf{x}_1) = (\mathbf{x}_2, \mathbb{H}\mathbf{x}_1) = (\mathbb{H}\mathbf{x}_2, \mathbf{x}_1) = \mu_2^*(\mathbf{x}_2, \mathbf{x}_1), \quad (1.106a)$$

i.e.,

$$(\mu_1 - \mu_2^*)(\mathbf{x}_2, \mathbf{x}_1) = 0. \quad (1.106b)$$

Equation (1.106b) for $\mathbf{x}_1 = \mathbf{x}_2$ implies that the eigenvalue μ_1 is real. This then holds for all eigenvalues. Next, if $\mathbf{x}_2 \neq \mathbf{x}_1$ and $\mu_2 \neq \mu_1$, Eq. (1.106b) tells us that $\mathbf{x}_2$ is orthogonal to $\mathbf{x}_1$.

1.7.4. Invariant Subspaces of Multiple Roots

Last, we shall now show that for self-adjoint operators (c) the *m-fold roots* μ_i *of the characteristic polynomial are associated with mutually orthogonal m-dimensional subspaces of* $\mathscr{V}^N$, *each invariant under* $\mathbb{H}$. When all roots are distinct ($m = 1$), we have shown above that $(\mathbf{x}_i, \mathbf{x}_j) = 0$ for $\mu_i \neq \mu_j$, $i, j = 1, 2, \ldots, N$, and the set of eigenvectors of $\mathbb{A}$ is a basis for $\mathscr{V}^N$. When multiplicity occurs in some of the roots μ_i, we can proceed as follows: Consider a first eigenvalue $\mu_1 \neq 0$, a corresponding eigenvector $\mathbf{x}_1$, and the $(N - 1)$-dimensional subspace $\mathscr{V}_{\perp 1}^{N-1}$ orthogonal to $\mathbf{x}_1$. Then, as $\mathbb{H}$ leaves the direction of $\mathbf{x}_1$ invariant, it will also transform $\mathscr{V}_{\perp 1}^{N-1}$ only onto itself: let $\mathbf{y} \in \mathscr{V}_{\perp 1}^{N-1}$ so $(\mathbf{y}, \mathbf{x}_1) = 0$; then, since $\mathbb{H}$ is self-adjoint,

$$0 = \mu_1(\mathbf{y}, \mathbf{x}_1) = (\mathbf{y}, \mathbb{H}\mathbf{x}_1) = (\mathbb{H}\mathbf{y}, \mathbf{x}_1), \tag{1.107a}$$

so that $\mathbb{H}\mathbf{y}$ is still orthogonal to $\mathbf{x}_1$ and thus in $\mathscr{V}_{\perp 1}^{N-1}$. Now, choosing an orthonormal basis in $\mathscr{V}_{\perp 1}^{N-1}$, $\{\boldsymbol{\varepsilon}_1^{(N-1)}, \boldsymbol{\varepsilon}_2^{(N-1)}, \ldots, \boldsymbol{\varepsilon}_{N-1}^{(N-1)}\}$, and scaling $\mathbf{x}_1$ if necessary so that $\|\mathbf{x}_1\| = 1$, we can build a matrix $\mathbf{X}_1$ with columns given by the vectors

$$\mathbf{X}_1 := \|\mathbf{x}_1, \boldsymbol{\varepsilon}_1^{(N-1)}, \boldsymbol{\varepsilon}_2^{(N-1)}, \ldots, \boldsymbol{\varepsilon}_{N-1}^{(N-1)}\|. \tag{1.107b}$$

Now, since any two columns n, m of $\mathbf{X}_1$ are orthogonal,

$$\sum_k (\mathbf{X}_1)_{kn}^* (\mathbf{X}_1)_{km} = \delta_{n,m}, \qquad \mathbf{X}_1^\dagger \mathbf{X}_1 = \mathbf{1}, \tag{1.107c}$$

so that $\mathbf{X}_1$ is a *unitary* matrix. If we multiply the hermitian matrix $\mathbf{H}$ by $\mathbf{X}_1$, the first column of $\mathbf{H}\mathbf{X}_1$ will be $\mu_1 \mathbf{x}_1$, while the other columns will be

$$\mathbf{H}\boldsymbol{\varepsilon}_k^{(N-1)} = \sum_{l=1}^{N-1} H_{lk}^1 \boldsymbol{\varepsilon}_l^{(N-1)} \in \mathscr{V}_{\perp 1}^{N-1}. \tag{1.107d}$$

Hence

$$\mathbf{H}\mathbf{X}_1 = \mathbf{X}_1 \begin{pmatrix} \mu_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^1 \end{pmatrix}, \tag{1.108a}$$

or

$$\mathbf{X}_1^\dagger \mathbf{H} \mathbf{X}_1 = \begin{pmatrix} \mu_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^1 \end{pmatrix}, \tag{1.108b}$$

and adjoining (1.108b), we can see that $\mathbf{H}^1$ is an $(N-1) \times (N-1)$ hermitian matrix. Thus far we have used one eigenvector corresponding to a nonzero eigenvalue to perform a *unitary* transformation on $\mathbf{H}$ and reduce it to a block-diagonal form. Moreover, the characteristic polynomial (1.104) can be written as

$$\begin{aligned} p_N(\mu) &= \det(\mathbf{H} - \mu\mathbf{1}) = \det\big(\mathbf{X}_1^\dagger(\mathbf{H} - \mu\mathbf{1})\mathbf{X}_1\big) = \det(\mathbf{X}_1^\dagger \mathbf{H}\mathbf{X}_1 - \mu\mathbf{1}) \\ &= (\mu_1 - \mu)\det(\mathbf{H}^1 - \mu\mathbf{1}) = (\mu_1 - \mu)p_{N-1}(\mu), \end{aligned} \tag{1.109}$$

where in the term before last $\mathbf{1}$ is the $(N-1) \times (N-1)$ unit matrix. In this way we see that the characteristic polynomial $p_{N-1}(\mu)$ of $\mathbf{H}^1$ has all the roots of $p_N(\mu)$ but μ_1.

The process for $\mathbf{H}$ can now be repeated for $\mathbf{H}^1$ using some other nonzero root μ_2 (which may be equal to μ_1 if this root turns out to be multiple) to successively "extract" root by root. If eigenvectors $\mathbf{x}_j^{(1)} \cdots \mathbf{x}_j^{(m)}$ belong to the same m-fold root μ_j, then any linear combination will also belong to the same eigenvalue, as

$$\mathbf{H} \sum_{r=1}^{m} c_r \mathbf{x}_j^{(r)} = \sum_{r=1}^{m} c_r \mathbf{H}\mathbf{x}_j^{(r)} = \mu_j \sum_{r=1}^{m} c_r \mathbf{x}_j^{(r)}, \tag{1.110}$$

and thus the set $\{\mathbf{x}_j^{(r)}\}_{r=1}^{m}$ spans an m-dimensional space $\mathscr{V}_j^m$, invariant under $\mathbb{H}$ and orthogonal to all other invariant subspaces or axes. If at the end of the recursive process we find m_0 zero roots, these will correspond to the m_0-dimensional subspace $\mathscr{V}_0^{m_0}$ of $\mathscr{V}^N$ orthogonal to all other extracted eigenvectors, and (1.110) holds for $\mathscr{V}_0^{m_0}$ as well. By the Schmidt procedure we can build an orthonormal basis for this subspace.

1.7.5. Diagonalization by Unitary Transformations

In conclusion, we have proven that *if* $\mathbf{H}$ *is a hermitian matrix we can build a unitary matrix* $\mathbf{X} = \mathbf{X}_1\mathbf{X}_2 \cdots \mathbf{X}_k$ (the factor $\mathbf{X}_j$ extracting the jth eigenvector and containing $j - 1$ ones along the diagonal) *such that*

$$\mathbf{X}^\dagger \mathbf{H} \mathbf{X} = \mathbf{H}^D, \tag{1.111}$$

where $\mathbf{H}^D$ *is a diagonal matrix containing along the diagonal the eigenvalues of* $\mathbf{H}$ *and* $\mathbf{X}$ *containing the eigenvectors of* $\mathbf{H}$ *as columns.*

1.7.6. The Second-Difference Operator and Fourier Transformation

As a concrete example, the operator $\mathbb{\Delta}$ of Section 1.5 was shown to be diagonalized by the Fourier transform, i.e., $\mathbf{F}^\dagger \boldsymbol{\Delta} \mathbf{F} = \tilde{\boldsymbol{\Delta}}$ in Eq. (1.61). The eigenvalues of $\mathbb{\Delta}$ are thus the λ_m of Eq. (1.62). Note that $\lambda_m = \lambda_{N-m}$ and $\lambda_N = 0$, so that for odd N, all roots of the characteristic polynomial but one are double, while for even N, $\lambda_{N/2}$ is also simple. Since the eigenvectors of a self-adjoint operator can be made into an orthonormal basis, the matrix representing the operator will be diagonal in that basis. For the $\mathbb{\Delta}$ operator this is precisely the $\boldsymbol{\varphi}$-basis.

1.7.7. Eigenvalues of Functions of Operators

The eigenvalues and eigenvectors of a hermitian operator $\mathbb{H}$ can be used to find those of any function $P(\mathbb{H})$. Indeed, let $\mathbb{H}\mathbf{x} = \mu\mathbf{x}$; then

$$P(\mathbb{H})\mathbf{x} = \sum_{n=0}^{\infty} p_n \mathbb{H}^n \mathbf{x} = \sum_{n=0}^{\infty} p_n \mu^n \mathbf{x} = P(\mu)\mathbf{x}. \tag{1.112}$$

It follows that *if* $\mathbf{x}$ *is an eigenvector of* $\mathbb{H}$ *with eigenvalue* μ *then it will also be an eigenvector of* $P(\mathbb{H})$ *with eigenvalue* $P(\mu)$.

Exercise 1.45. Consider the operators $\mathbb{I}_k$ of the dihedral group represented by the hermitian matrices $\mathbf{I}_k$ in (1.88b). As $\mathbb{I}_k^2 = \mathbb{1}$ and $\mathbb{1}$ has 1 for its sole eigenvalue, show that $\mathbb{I}_k$ can only have eigenvalues ± 1. Find a set of eigenvectors for $\mathbb{I}_k$. For this note that $\mathbb{I}_0 \boldsymbol{\varepsilon}_n = \boldsymbol{\varepsilon}_{N-n}$ and $\mathbb{I}_0 \boldsymbol{\varphi}_m = \boldsymbol{\varphi}_{N-m}$.

1.7.8. Unitary Operators and Their Spectra

The case when $\mathbb{A}$ is a unitary operator $\mathbb{U}$, $\mathbb{U}\mathbb{U}^\dagger = \mathbb{1}$, will now be examined. Of the three main results we proved for self-adjoint operators (on the spectrum, orthogonality, and completeness of eigenvectors) only the first differs for unitary operators. The other two hold *verbatim*. First, note that if $\mathbb{U}\mathbf{x} = \mu\mathbf{x}$ holds, then, multiplying by $\mu^{-1}\mathbb{U}^\dagger$, we find $\mathbb{U}^\dagger\mathbf{x} = \mu^{-1}\mathbf{x}$, so that if $\mathbf{x}$ is an eigenvector of $\mathbb{U}$ with eigenvalue μ, it will also be an eigenvector of $\mathbb{U}^\dagger$ with eigenvalue μ^{-1}. Now consider the analogue of Eq. (1.106) for unitary operators for any two eigenvectors of $\mathbb{U}$ and $\mathbb{U}^\dagger$:

$$\mu_1(\mathbf{x}_2, \mathbf{x}_1) = (\mathbf{x}_2, \mathbb{U}\mathbf{x}_1) = (\mathbb{U}^\dagger\mathbf{x}_2, \mathbf{x}_1) = \mu_2^{-1*}(\mathbf{x}_2, \mathbf{x}_1), \tag{1.112a}$$

i.e.,

$$(\mu_1 - \mu_2^{-1*})(\mathbf{x}_2, \mathbf{x}_1) = 0. \tag{1.112b}$$

For $\mathbf{x}_2 = \mathbf{x}_1$ we thus conclude that the allowed eigenvalues μ must satisfy $\mu^* = \mu^{-1}$, i.e., they can only be complex numbers of unit modulus. Thus *the spectrum of a unitary operator is restricted to lie on the unit circle.* Next, as was the case for hermitian operators from (1.112b), *if* $\mu_1 \neq \mu_2 = \mu_2^{-1*}$, *then the corresponding eigenvectors are orthogonal.* Last, the constructive proof of the statement of Section 1.7.5 can be followed as before with some minor changes. Since $(\mathbb{U}\mathbf{y}, \mathbb{U}\mathbf{x}) = (\mathbf{y}, \mathbf{x}) = (\mathbb{U}^\dagger\mathbf{y}, \mathbb{U}^\dagger\mathbf{x})$, it follows that if $\mathbb{U}$ leaves a subspace of $\mathscr{V}^N$ invariant, so does $\mathbb{U}^\dagger$. The statement stemming from (1.107a) thus also applies for unitary operators. The construction (1.107)–(1.108) can now be followed, replacing hermitian by unitary matrices, and the proof is complete. We do not have to worry about null eigenvalues here.

Exercise 1.46. Consider in $\mathscr{V}^3$ the unitary rotation around the z-axis given by the matrix

$$\mathbf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.113a}$$

Verify that its normalized eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 1/(2)^{1/2} \\ -i/(2)^{1/2} \\ 0 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 1/(2)^{1/2} \\ i/(2)^{1/2} \\ 0 \end{pmatrix} \tag{1.113b}$$

and that they constitute a unitary matrix $\mathbf{X} = \|\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\|$. The first of the eigenvectors is the ordinary axis; the second two are *polarization vectors*. Find the corresponding eigenvalues from (1.113a) by the characteristic equation and check the assignments. Notice that the eigenvalue problem has no complete solution in a purely real space.

Exercise 1.47. The rotation matrices (1.113a) are unitary. Find the hermitian operator which generates the set as (1.74). You can use Eq. (1.79). Verify that the eigenvectors (1.113b) are also eigenvectors of the generating operators. Find the corresponding eigenvalues.

Exercise 1.48. The rotations of the dihedral group D_N in $\mathcal{V}^N$ are unitary. They are represented in the $\boldsymbol{\varepsilon}$-basis by (1.88a), while in the $\boldsymbol{\varphi}$-basis they are represented by (1.91). What are the eigenvalues and eigenvectors?

Exercise 1.49. Consider the *Fourier* transform matrix. Show that its eigenvalues are among the set $\{\pm 1, \pm i\}$. Find eigenvectors for **F**, noting that **F** transforms $\boldsymbol{\varepsilon}_n$ to $\boldsymbol{\varphi}_n$ and $\boldsymbol{\varphi}_n$ to $\boldsymbol{\varepsilon}_{N-n}$.

The reader may ask if any larger class of operators has the property common to self-adjoint and unitary ones: orthogonality and completeness of their eigenvectors. In fact, this is a property of all and only *normal operators*, i.e., those operators $\mathbb{N}$ which commute with their adjoints $\mathbb{N}^\dagger\mathbb{N} = \mathbb{N}\mathbb{N}^\dagger$. The proof of this statement can be seen, for instance, in the book by Fano (1971, Section 2.3). As to the question of whether all operators have eigenvectors which diagonalize their representing matrices, the answer is in the negative. The most one can do in the general case is to achieve a reduction into the *Jacobi canonical form*: a block-diagonal form, one block for each distinct eigenvalue and each block being a matrix with a shifted diagonal of 1's beside the main diagonal. A discussion of this can be found in the book by Gel'fand (1961, Section III).

1.7.9. Eigenbases of Operators with Degenerate Eigenvalues

As we have seen, the eigenvectors of a self-adjoint or unitary operator in $\mathcal{V}^N$ constitute an orthonormal basis for the space called the *spectral basis* or *eigenbasis* of the operator. If this operator describes the time evolution of a system (to be seen in Chapter 2), it is very convenient to use this basis since the coordinate directions defined by the eigenvectors will remain invariant in time and will change only in scale. Moreover, the eigenvectors are conveniently labeled by the eigenvalues of the operator, except, that is, for the ambiguities which may arise when two or more eigenvalues coincide in a multiple root of the characteristic polynomial. Such eigenvalues are said to be *degenerate*. The term is borrowed from quantum mechanics. Our nearest example of degeneracy appears in the eigenvalues of Δ which are equal by pairs: $\lambda_n = \lambda_{N-n}$. To resolve the degeneracy and use eigenvalues to label the basis vectors uniquely we may hope to find one (or more) extra operators whose eigenvalues will specify the eigenvector labels completely within each of the invariant subspaces of the first operator.

1.7.10. Removal of Degeneracy

Specifically, we have $\mathbb{A}\mathbf{x}_i = \mu_i \mathbf{x}_i$ for the first operator, and we need $\mathbb{B}\mathbf{x}_j = \nu_j \mathbf{x}_j$ for the second, so that although some of the μ's and some of the ν's may be degenerate, we hope that the assignment of the *pairs* (μ_i, ν_j) to $\mathbf{x}_{ij}$ will be one to one. We shall prove that *two operators* $\mathbb{A}$ *and* $\mathbb{B}$ *can have simultaneously the same set* $\{\mathbf{x}_i\}_{i=1}^N$ *of eigenvectors if and only if they commute.* If they have the same eigenvector set, they will both be represented in the common spectral basis by diagonal matrices, which commute. If they commute, then once we have found the spectral basis for $\mathbb{A}$ (so that the representing matrix $\bar{\mathbf{A}}$ of $\mathbb{A}$ is diagonal with possibly repeated eigenvalues), the representing matrix $\bar{\mathbf{B}}$ for $\mathbb{B}$ in the same basis can only be block diagonal, each block with the size and position of the sets of degenerate eigenvalues of $\bar{\mathbf{A}}$. Each block in $\bar{\mathbf{B}}$ can be diagonalized as (1.111) without affecting $\bar{\mathbf{A}}$; the result is a final eigenvector basis $\{\bar{\bar{\mathbf{x}}}_i\}_{i=1}^N$ where both $\bar{\mathbf{A}}$ and $\bar{\bar{\mathbf{B}}}$ are fully diagonal.

1.7.11. The Case of the $\mathbb{\Delta}$ Operator

The $\mathbb{\Delta}$ operator, we noted, has doubly degenerate eigenvalues. We can use any of the dihedral operators commuting with $\mathbb{\Delta}$ [Eqs. (1.96)] to complete the labeling. The operator $\mathbb{R}$ seems the wisest choice: in the $\boldsymbol{\varphi}$-basis it is already diagonal [Eq. (1.91) for $k = 1$], and all of its eigenvalues are distinct:

$$\mathbb{R}\boldsymbol{\varphi}_n = \exp(2\pi i n/N)\boldsymbol{\varphi}_n, \qquad n = 1, 2, \ldots, N. \tag{1.114}$$

Indeed, $\mathbb{R}$ could be used alternatively to *define* the $\boldsymbol{\varphi}$-basis vectors uniquely, with no labeling degeneracy. As another choice, the $\mathbb{I}_0$ operator [Eq. (1.92) for $l = 0$] can be used. (The matrix $\tilde{\mathbf{I}}_0$ does not *appear* to be block diagonal since the pairs of degenerate eigenvalue vectors are not placed sequentially in the basis.) If we define

$$\left.\begin{aligned} \boldsymbol{\varphi}_m{}^+ &= 2^{-1/2}(\boldsymbol{\varphi}_m + \boldsymbol{\varphi}_{N-m}), \\ \boldsymbol{\varphi}_m{}^- &= i2^{-1/2}(\boldsymbol{\varphi}_m - \boldsymbol{\varphi}_{N-m}), \end{aligned}\right\} \quad m = 1, 2, \ldots, \begin{cases} \tfrac{1}{2}(N-1), & N \text{ odd}, \\ \tfrac{1}{2}N - 1, & N \text{ even}, \end{cases} \tag{1.115a, b}$$

$$\boldsymbol{\varphi}_0^+ = \boldsymbol{\varphi}_N, \qquad \text{and} \qquad \boldsymbol{\varphi}_{N/2}^+ = \boldsymbol{\varphi}_{N/2} \qquad \text{for } N \text{ even}, \tag{1.115c}$$

we can see that

$$\mathbb{\Delta}\boldsymbol{\varphi}_m{}^\pm = \lambda_m \boldsymbol{\varphi}_m{}^\pm, \tag{1.116a}$$

$$\mathbb{I}_0\boldsymbol{\varphi}_m{}^\pm = \pm\boldsymbol{\varphi}_m{}^\pm, \tag{1.116b}$$

so that (1.115) is the common eigenbasis of $\mathbb{\Delta}$ and $\mathbb{I}_0$.

Exercise 1.50. Verify that the basis (1.115) indeed has N orthonormal vectors.

Exercise 1.51. Find the matrix transforming the $\boldsymbol{\varphi}$-basis to the $\boldsymbol{\varphi}^{\pm}$-basis (1.115).

Exercise 1.52. Write out explicitly the matrices representing $\mathbb{\Delta}$, $\mathbb{I}_0$, and $\mathbb{R}$ in the $\boldsymbol{\varphi}^{\pm}$-basis (1.115). You can do this by using the results of Exercise 1.51 or, for the first two operators, directly from (1.116).

Exercise 1.53. Generalize the choices of basis given by (1.114) and (1.115)–(1.116a): Show that $\mathbb{R}^k$ defines a basis equally well as long as k is not a divisor of N. Regarding $\mathbb{I}_l$, construct eigenbases of $\mathbb{\Delta}$ and $\mathbb{I}_l$ with

$$\boldsymbol{\varphi}_m^{(l)\pm} := \alpha_m{}^{\pm}\boldsymbol{\varphi}_m + \beta_m{}^{\pm}\boldsymbol{\varphi}_{N-m}, \qquad |\alpha_m{}^{\pm}|^2 + |\beta_m{}^{\pm}|^2 = 1, \tag{1.117a}$$

$$\mathbb{I}_l\boldsymbol{\varphi}_m^{(l)\pm} = \pm\boldsymbol{\varphi}_m^{(l)\pm}, \tag{1.117b}$$

where the range of m is the same as in Eqs. (1.115). Show that a good choice is

$$\alpha_m{}^{\pm} = 2^{-1/2}\exp(2\pi ilm/N), \qquad \beta_m{}^{\pm} = \pm 2^{-1/2}\exp(-2\pi ilm/N). \tag{1.117c}$$

Show that this indeed generalizes the $\boldsymbol{\varphi}^{\pm}$-basis in Eqs. (1.115)–(1.116).

Exercise 1.54. Recall that for N even the transformations $\mathbb{K}_l$ [(1.97) and below] also come into play. One can define, in analogy to (1.115),

$$\boldsymbol{\varphi}_m^{1+} := 2^{-1/2}[\exp(i\pi m/N)\boldsymbol{\varphi}_m + \exp(-i\pi m/N)\boldsymbol{\varphi}_{N-m}], \tag{1.118a}$$

$$\boldsymbol{\varphi}_m^{1-} := i2^{-1/2}[\exp(i\pi m/N)\boldsymbol{\varphi}_m - \exp(-i\pi m/N)\boldsymbol{\varphi}_{N-m}], \tag{1.118b}$$

for $m = 1, 2, \ldots, \frac{1}{2}N - 1$, and

$$\boldsymbol{\varphi}_{N/2}^{1+} = \boldsymbol{\varphi}_{N/2}, \qquad \boldsymbol{\varphi}_0^{1+} = \boldsymbol{\varphi}_N. \tag{1.118c}$$

Show that these are eigenfunctions of $\mathbb{K}_0$, i.e.,

$$\mathbb{K}_0\boldsymbol{\varphi}_m^{1\pm} = \pm\boldsymbol{\varphi}_m^{1\pm}. \tag{1.118d}$$

Show that the vectors defined above have *real* components in the $\boldsymbol{\varepsilon}$-basis. One can do the same for the other $\mathbb{K}_l$'s.

Exercise 1.55. Consider in all detail, since it can be done algebraically, the eigenvalues and vectors of 2×2 complex matrices, i.e.,

$$\mathbf{M}\mathbf{v}^{\pm} := \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x^{\pm} \\ y^{\pm} \end{pmatrix} = \lambda^{\pm}\begin{pmatrix} x^{\pm} \\ y^{\pm} \end{pmatrix} = \lambda^{\pm}\mathbf{v}^{\pm}. \tag{1.119}$$

Show that the eigenvalues are

$$\lambda^{\pm} = \tfrac{1}{2}(a + d) \pm \{[\tfrac{1}{2}(a - d)]^2 + bc\}^{1/2}, \tag{1.120}$$

which are real for $\mathbf{M}$ hermitian (a, d real and $b = c^*$). They are also real for $\mathbf{M}$ real and $[\frac{1}{2}(a - d)]^2 + bc \geqslant 0$. Note that $\sum_i \lambda_i = \operatorname{trace} \mathbf{M}$ and $\Pi_i\lambda_i = \det \mathbf{M}$: These are general properties for any dimension.

Exercise 1.56. Examine now the eigenvectors in (1.119). Show that

$$\frac{y^{\pm}}{x^{\pm}} = \frac{\lambda^{\pm} - a}{b} = \frac{c}{\lambda^{\pm} - d}. \tag{1.121}$$

Note that two proper eigenvectors need not exist for arbitrary **M** (for example, when $b = 0$). Show that when $\lambda^{\pm}$ and a are real, the angle between $\mathbf{v}^{+}$ and $\mathbf{v}^{-}$ is (without assuming their normalization)

$$(\mathbf{v}^{+}, \mathbf{v}^{-}) = x^{+*}x^{-}\left[1 - \frac{c}{b^*}\frac{(\lambda^{+} - a)^*}{\lambda^{+} - a}\right]. \tag{1.122}$$

They are thus orthogonal if **M** is hermitian or unitary. The results of the last exercises will be handy later on.

Exercise 1.57. Show that if **x** is an arbitrary normalized vector in $\mathscr{V}^N$, then

$$\mathbf{x}^{I_k\pm} := \tfrac{1}{2}(\mathbb{1} \pm \mathbb{I}_k)\mathbf{x}, \qquad \mathbf{x}^{K_l\pm} := \tfrac{1}{2}(\mathbb{1} \pm \mathbb{K}_l)\mathbf{x} \tag{1.123}$$

are eigenvectors of $\mathbb{I}_k$ and $\mathbb{K}_l$, respectively, with eigenvalues ± 1. The operators $\frac{1}{2}(\mathbb{1} \pm \mathbb{I}_k)$ and $\frac{1}{2}(\mathbb{1} \pm \mathbb{K}_l)$ are *projection* operators onto orthogonal subspaces.

1.7.12. Finding the Fourier Transformation

Exercises 1.58–1.60 show how we can *find* the Fourier transformation **F** as that which diagonalizes the second-difference matrix representative $\boldsymbol{\Delta}$ given by (1.60).

Exercise 1.58. Consider diagonalizing $\boldsymbol{\Delta}$ through an (unknown) unitary matrix **F** as $\boldsymbol{\Delta}\mathbf{F} = \mathbf{F}\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is diagonal with elements λ_n. Show that the $m - n$ element of this equality leads to the recursion relation

$$F_{m+1,n} = (2 + \lambda_n)F_{mn} - F_{m-1,n}. \tag{1.124}$$

The indices in (1.124) are to be considered modulo N. This allows us to write any F_{mn} in terms of F_{1n} and $F_{0n} \equiv F_{Nn}$ as

$$F_{m+1,n} = U_m(x_n)F_{1n} - U_{m-1}(x_n)F_{0n}, \qquad x_n := 1 + \lambda_n/2, \tag{1.125}$$

where the U_m are polynomials in x_n. Combining the two preceding equations shows that they satisfy the recurrence relation

$$U_m(x_n) = 2x_n U_{m-1}(x_n) - U_{m-2}(x_n), \tag{1.126a}$$

$$U_0(x_n) = 1, \qquad U_1(x_n) = 2x_n. \tag{1.126b}$$

Exercise 1.59. Show that the solution to the recurrence relation (1.126) is given by

$$U_m(x) = \sin[(m + 1) \arccos x]/\sin(\arccos x). \tag{1.127}$$

These are the *Chebyshev polynomials of the second kind* of degree m, and (1.126a) is their Christoffel–Darboux formula. [See the mathematical handbook by Abramowitz and Stegun (1964, Chapter 22).] The recurrence relation for the elements of **F** "closes" for $m = N \equiv 0$. From (1.124) for $m = N$ and from (1.125) for $m = N - 2$ and $N - 1$, show that this leads to a pair of homogeneous simultaneous equations in F_{0n} and F_{1n}, the vanishing of whose determinant requires

$$U_{N-1}(x)[U_{N-3}(x) + 2x] - [U_{N-2}(x) + 1]^2 = 0. \qquad (1.128)$$

The roots of this polynomial equation will determine the allowed values of λ_n. By the use of trigonometric identities, (1.128) can be reduced to

$$\sin\theta(1 - \cos N\theta) = 0, \qquad \cos\theta := x = 1 + \lambda/2 \Rightarrow \theta = 2\pi k/N,$$
$$k = 0, \pm 1, \pm 2, \ldots, \qquad (1.129)$$

whose roots yield precisely the values of λ_n given by (1.62b), so the index n can serve to number columns—any other one would just permute the λ_n's in $\mathbf{\Lambda}$. Note, though, that all eigenvalues but λ_N (and $\lambda_{N/2}$ if N is even) are twofold degenerate. See Weinstock (1971).

Exercise 1.60. Substituting the eigenvalues λ_n into (1.125) and letting $F_{1n} = \gamma_n F_{0n}$, find F_{mn}. The requirement of unitarity $\sum_m |F_{mn}|^2 = 1$ will fix $|F_{0n}|$ but leaves a three-dimensional freedom in choosing each complex γ_n and the phase of F_{0n}. The choice $\arg F_{0n} = 0$ and $\gamma_n = \exp(-2\pi i n/N)$ produces *the* Fourier transform matrix (1.48). Examine first the columns F_{mN} and $F_{m,N/2}$ if N is even. There, $U(x_N) = m + 1$ and $U_m(x_{N/2}) = (-1)^m(m + 1)$. Proceed then to the other columns, noting the twofold eigenvalue degeneracy.

Table 1.1. Coordinates in the ε- and φ-Bases of Vectors Subject to Various Operations or Acted upon by Operators[a]

Operation	$\mathbf{f}$	f_n	$\tilde{f}_m$
Linear combination	$a\mathbf{f} + b\mathbf{g}$	$af_n + bg_n$	$a\tilde{f}_m + b\tilde{g}_m$
Inner product	$(\mathbf{f}, \mathbf{g})$	$\sum_n f_n^* g_n$	$\sum_m \tilde{f}_m^* \tilde{g}_m$
Translation (rotation)	$\mathbb{R}^k\mathbf{f}$	f_{n-k}	$\exp(2\pi ikm/N)\tilde{f}_m$
Inversion	$\mathbb{I}_k\mathbf{f}$	f_{N+2k-n}	$\exp(4\pi ikm/N)\tilde{f}_{N-m}$
	$\mathbb{K}_l\mathbf{f}$	$f_{N+2k-n+1}$	$\exp[4\pi i(k+\frac{1}{2})m/N]\tilde{f}_{N-m}$
Second difference	$\mathbb{\Delta}\mathbf{f}$	$f_{n+1} - 2f_n + f_{n-1}$	$-4\sin^2(\pi m/N)\tilde{f}_m$
Complex conjugation		f_n^*	$\tilde{f}_{N-m}^*$
Convolution (Section 3.1)	$\mathbf{f}\,(\varepsilon)\,\mathbf{g}$	$f_n g_n$	$N^{-1/2}\sum_r \tilde{f}_r \tilde{g}_{m-r}$
	$\mathbf{f}\,(\varphi)\,\mathbf{g}$	$N^{-1/2}\sum_s f_s g_{n-s}$	$\tilde{f}_m \tilde{g}_m$
Correlation (Section 3.2)	$\mathbf{f} \circ \mathbf{g}$	$N^{-1/2}\sum_s f_s^* g_{n+s}$	$\tilde{f}_m^* \tilde{g}_m$

[a] In all cases n and m appear mod N.