Integral Transforms in Science and Engineering

Kurt Bernardo Wolf

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
México D.F., México

Library of Congress Cataloging in Publication Data

Wolf, Kurt Bernardo.
Integral transforms in science and engineering.

(Mathematical concepts and methods in science and engineering; v. 11)
Bibliography: p.
Includes index.
1. Integral transforms. I. Title.
QA432.W64 515'.723 78-12482
ISBN 0-306-39251-8

A Division of Plenum Publishing Corporation
227 West 17th Street, New York, N.Y. 10011

Printed in the United States of America

Contents

PART I. FINITE-DIMENSIONAL VECTOR SPACES AND THE FOURIER TRANSFORM

1. Concepts from Complex Vector Analysis and the Fourier Transform . 3

1.1. N-Dimensional Complex Vector Spaces 3
1.2. Inner Product and Norm in $\mathscr{V}^N$ 5
1.3. Passive and Active Transformations 9
1.4. Unitary Transformations: Permutations and the Fourier Transform 13
1.5. Self-Adjoint Operators 20
1.6. The Dihedral Group 26
1.7. The Axes of a Transformation: Eigenvalues and Eigenvectors . 33

2. The Application of Fourier Analysis to the Uncoupling of Lattices . 43

2.1. Mechanical and Electric Analogies 43
2.2. The Equation of Motion of Coupled Systems and Solution . 49
2.3. Fundamental Solutions, Normal Modes, and Traveling Waves . 56
2.4. Farther-Neighbor Interaction, Molecular and Diatomic Lattices 71
2.5. Energy in a Lattice 85
2.6. Phase Space, Time Evolution, and Constants of Motion . . 88

3. Further Developments and Applications of the Finite Fourier Transform 101

3.1. Convolution: Filters and Windows 101
3.2. Correlation: Signal Detection and Noise. 114
3.3. The Fast Fourier Transform Algorithm 125
3.4. The "Limit" $N \to \infty$: Fourier Series and Integral Transforms . 130

PART II. FOURIER AND BESSEL SERIES

4. Function Vector Spaces and Fourier Series 139
- 4.1. Notions on Function Vector Spaces 139
- 4.2. The Dirichlet Conditions 145
- 4.3. Alternative Representations, Transformations, and Symmetries . 154
- 4.4. Differential Properties and Convergence 162
- 4.5. The Dirac δ and Divergent Series 174
- 4.6. Linear Operators, Infinite Matrices, and Integral Kernels . . 184
- 4.7. Fourier Series for Any Period and the Infinite-Period Limit . 191

5. Fourier Series in Diffusion and Wave Phenomena 195
- 5.1. Heat Diffusion in a Ring 195
- 5.2. The Vibrating String 201
- 5.3. The Infinite Lattice 214

6. Normal Mode Expansion and Bessel Series 221
- 6.1. Eigenfunctions of the Laplacian on Finite Regions: The Rectangular Membrane 221
- 6.2. Laplacian on the Unit Disk: The Circular Membrane . . 229
- 6.3. Sectorial and Annular Membranes 239
- 6.4. Other Series of Orthonormal Functions 246

PART III. FOURIER AND RELATED INTEGRAL TRANSFORMS

7. Fourier Transforms 255
- 7.1. The Fourier Integral Theorem 255
- 7.2. Various Operators and Operations under Fourier Transformation 265
- 7.3. The Dirac δ and the Green's Function for a System. . . 279
- 7.4. Causality and Dispersion Relations 293
- 7.5. Oscillator Wave Functions 307
- 7.6. Uncertainty Relations 328

8. Integral Transforms Related to the Fourier Transform 333
- 8.1. Laplace Transforms 334
- 8.2. Mellin Transforms 344
- 8.3. N-Dimensional Fourier Transforms 352
- 8.4. Hankel Transforms 361
- 8.5. Other Integral Transforms 372

PART IV. CANONICAL TRANSFORMS

9. Construction and Properties of Canonical Transforms . . . 381
- 9.1. Real Linear Canonical Transforms. 381
- 9.2. Complex Linear Transforms and Bargmann Space . . 393
- 9.3. Canonical Transforms by Hyperdifferential Operators . . 405

10. Applications to the Study of Differential Equations 417
10.1. The Diffusion Equation: Similarity Group and Separating Coordinates 417
10.2. Inhomogeneous Linear Canonical Transforms and Parabolic Equations 434

Appendix A. The Gamma Function 445
Appendix B. The Bessel and Related Functions 449
Appendix C. Some Summation Formulas 457

References 459
Notation 469
Index 473

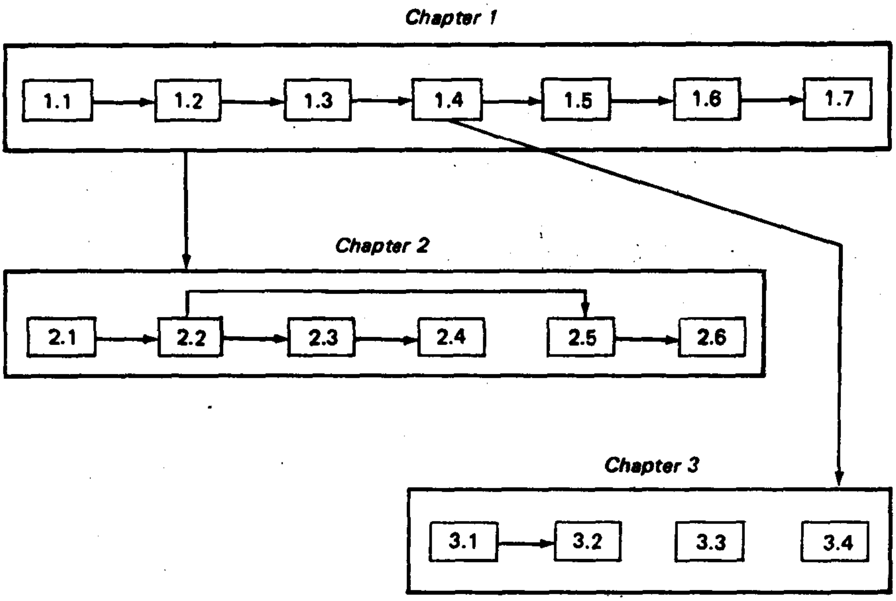
Part I

Finite-Dimensional Vector Spaces and the Fourier Transform

In this part we develop the mathematical framework of *finite-dimensional* Fourier transforms and give the basics of two fields where it has found fruitful application: in the analysis of coupled systems and in communication theory and technology.

Chapter 1 deals with complex vector analysis in N dimensions and leads rather quickly to the tools of Fourier analysis: unitary transformations and self-adjoint operators. The uncoupling of lattices representing one-dimensional crystals and electric *RLC* networks is undertaken in Chapter 2. We examine in detail the fundamental solutions, normal modes, and traveling waves for first-neighbor interactions in simple crystal lattices and extend these to farther-neighbor, molecular, and diatomic crystals. The Fourier formalism is also used to describe the analytical mechanics of these systems: phase space, energy, evolution operators, and other conservation laws. Chapter 3 introduces convolution and correlation, sketching their use in filtering, windowing, and modulation of signals and their detection in the presence of background noise. The workings of the fast Fourier transform (FFT) computation algorithm are given in Section 3.3. Finally, in Section 3.4, some properties of Fourier series and integral transforms (Parts II and III) are put in the form of corresponding properties of the finite Fourier transform on vector spaces whose dimension grows without bound.

Chapters 2 and 3 are independent of each other and can be chosen according to the reader's interest. With the first choice, Sections 1.6 and 1.7 will be particularly needed. The understanding of Chapter 3, on the other hand, does not require basically more than Sections 1.1–1.4. Before going to the following parts in this text, the reader may find Section 3.4 useful. Table 1.1, which gives the main properties of the finite Fourier transform, is placed at the end of Chapter 1.



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

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

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 $\{\boldsymbol{\varepsilon}_n\}_{n=1}^N \coloneqq \{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\}$ 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

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

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 $\boldsymbol{\varepsilon}_{N+n}^I := i\boldsymbol{\varepsilon}_n$, $n = 1, 2, \ldots, N$. (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

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

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

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

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.

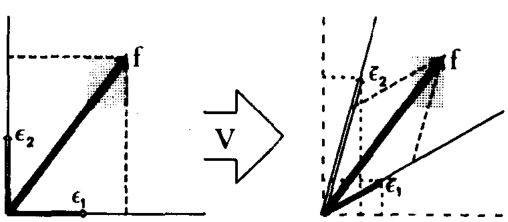


Fig. 1.1. Passive transformation V of a (two-dimensional) vector space. Its description in terms of a basis $\{\boldsymbol{\varepsilon}_i\}$ is replaced by its description in terms of a transformed basis $\{\bar{\boldsymbol{\varepsilon}}_i\}$. The vectors $\mathbf{f}$ in the space are unchanged.

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

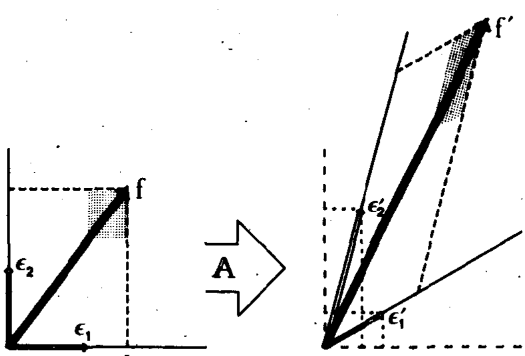


Fig. 1.2. Active transformation $\mathbb{A}$ of a (two-dimensional) vector space. All vectors—basis vectors included—are changed. As the transformation is linear, however, the coordinates of $\mathbf{f}' = \mathbb{A}\mathbf{f}$ in the transformed basis $\{\boldsymbol{\varepsilon}'_i\} = \{\mathbb{A}\boldsymbol{\varepsilon}_i\}$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

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

$$(a\mathsf{A} + b\mathsf{B})\mathbf{f} \coloneqq a\mathsf{A}\mathbf{f} + b\mathsf{B}\mathbf{f}. \tag{1.35}$$

Now let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ be the representing matrices. Then, using (1.31),

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

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

$$(\mathsf{AB})\mathbf{f} \coloneqq \mathsf{A}(\mathsf{B}\mathbf{f}). \tag{1.37}$$

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

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

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 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 A would be described by a different matrix $\bar{\mathbf{A}} = \|\bar{A}_{nm}\|$ whose elements are

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

Exercise 1.8. Show that

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

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

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

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