Bruce R. Kusse and Erik A. Westwig

Mathematical Physics

Applied Mathematics for Scientists and Engineers

2nd Edition

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

This Page Intentionally Left Blank

Bruce R. Kusse and Erik A. Westwig
Mathematical Physics

Related Titles

Vaughn, M. T.

Introduction to Mathematical Physics

2006. Approx. 650 pages with 50 figures.

Softcover

ISBN 3-527-40627-1

Lambourne, R., Tinker, M.

Basic Mathematics for the Physical Sciences

2000. 688 pages.

Softcover

ISBN 0-471-85207-4

Tinker, M., Lambourne, R.

Further Mathematics for the Physical Sciences

2000. 744 pages.

Softcover

ISBN 0-471-86723-3

Courant, R., Hilbert, D.

Methods of Mathematical Physics

Volume 1

1989. 575 pages with 27 figures.

Softcover

ISBN 0-471-50447-5

Volume 2

1989. 852 pages with 61 figures.

Softcover

ISBN 0-471-50439-4

Trigg, G. L. (ed.)

Mathematical Tools for Physicists

2005. 686 pages with 98 figures and 29 tables.

Hardcover

ISBN 3-527-40548-8

Bruce R. Kusse and Erik A. Westwig

Mathematical Physics

Applied Mathematics for Scientists and Engineers

2nd Edition

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

The Authors

Bruce R. Kusse
College of Engineering
Cornell University
Ithaca, NY
brk2@cornell.edu

Erik Westwig
Palisade Corporation
Ithaca, NY
ewestwig@palisade.com

For a Solution Manual, lecturers should contact the editorial department at physics@wiley-vch.de, stating their affiliation and the course in which they wish to use the book.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

Printing Strauss GmbH, Mörlenbach
Binding J. Schäffer Buchbinderei GmbH, Grünstadt

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN-13: 978-3-527-40672-2
ISBN-10: 3-527-40672-7

PREFACE

This book is the result of a sequence of two courses given in the School of Applied and Engineering Physics at Cornell University. The intent of these courses has been to cover a number of intermediate and advanced topics in applied mathematics that are needed by science and engineering majors. The courses were originally designed for junior level undergraduates enrolled in Applied Physics, but over the years they have attracted students from the other engineering departments, as well as physics, chemistry, astronomy and biophysics students. Course enrollment has also expanded to include freshman and sophomores with advanced placement and graduate students whose math background has needed some reinforcement.

While teaching this course, we discovered a gap in the available textbooks we felt appropriate for Applied Physics undergraduates. There are many good introductory calculus books. One such example is *Calculus and Analytic Geometry* by Thomas and Finney, which we consider to be a prerequisite for our book. There are also many good textbooks covering advanced topics in mathematical physics such as *Mathematical Methods for Physicists* by Arfken. Unfortunately, these advanced books are generally aimed at graduate students and do not work well for junior level undergraduates. It appeared that there was no intermediate book which could help the typical student make the transition between these two levels. Our goal was to create a book to fill this need.

The material we cover includes intermediate topics in linear algebra, tensors, curvilinear coordinate systems, complex variables, Fourier series, Fourier and Laplace transforms, differential equations, Dirac delta-functions, and solutions to Laplace's equation. In addition, we introduce the more advanced topics of contravariance and covariance in nonorthogonal systems, multi-valued complex functions described with branch cuts and Riemann sheets, the method of steepest descent, and group theory. These topics are presented in a unique way, with a generous use of illustrations and graphs and an informal writing style, so that students at the junior level can grasp and understand them. Throughout the text we attempt to strike a healthy balance between mathematical completeness and readability by keeping the number of formal proofs and theorems to a minimum. Applications for solving real, physical problems are stressed. There are many examples throughout the text and exercises for the students at the end of each chapter.

Unlike many text books that cover these topics, we have used an organization that is fundamentally pedagogical. We consider the book to be primarily a teaching tool, although we have attempted to also make it acceptable as a reference. Consistent with this intent, the chapters are arranged as they have been taught in our two course sequence, rather than by topic. Consequently, you will find a chapter on tensors and a chapter on complex variables in the first half of the book and two more chapters, covering more advanced details of these same topics, in the second half. In our first semester course, we cover chapters one through nine, which we consider more important for the early part of the undergraduate curriculum. The last six chapters are taught in the second semester and cover the more advanced material.

We would like to thank the many Cornell students who have taken the AEP 321/322 course sequence for their assistance in finding errors in the text, examples, and exercises. E.A.W. would like to thank Ralph Westwig for his research help and the loan of many useful books. He is also indebted to his wife Karen and their son John for their infinite patience.

BRUCE R. KUSSE
ERIK A. WESTWIG

Ithaca, New York

CONTENTS

This Page Intentionally Left Blank

1

A REVIEW OF VECTOR AND MATRIX ALGEBRA USING SUBSCRIPT/SUMMATION CONVENTIONS

This chapter presents a quick review of vector and matrix algebra. The intent is not to cover these topics completely, but rather use them to introduce subscript notation and the Einstein summation convention. These tools simplify the often complicated manipulations of linear algebra.

1.1 NOTATION

Standard, consistent notation is a very important habit to form in mathematics. Good notation not only facilitates calculations but, like dimensional analysis, helps to catch and correct errors. Thus, we begin by summarizing the notational conventions that will be used throughout this book, as listed in Table 1.1.

TABLE 1.1. Notational Conventions

Symbol	Quantity
a	A real number
$\underline{B}$	A complex number
A_i	A vector component
$M_{i\cdots j}$	A matrix or tensor element
$[M]$	An entire matrix
$\overline{\mathbf{A}}$	A vector
$\hat{\mathbf{e}}_i$	A basis vector
$\overline{\overline{\mathbf{T}}}$	A tensor
$\mathcal{L}$	An operator

A three-dimensional vector $\overline{\mathbf{V}}$ can be expressed as

$$\overline{\mathbf{V}} = V_x \hat{\mathbf{e}}_x + V_y \hat{\mathbf{e}}_y + V_z \hat{\mathbf{e}}_z, \tag{1.1}$$

where the components (V_x, V_y, V_z) are called the Cartesian components of $\overline{\mathbf{V}}$ and $(\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z)$ are the basis vectors of the coordinate system. This notation can be made more efficient by using ***subscript notation***, which replaces the letters (x, y, z) with the numbers $(1, 2, 3)$. That is, we define:

$$\begin{matrix} \hat{\mathbf{e}}_1 \equiv \hat{\mathbf{e}}_x & & V_1 \equiv V_x \\ \hat{\mathbf{e}}_2 \equiv \hat{\mathbf{e}}_y & \text{and} & V_2 \equiv V_y \\ \hat{\mathbf{e}}_3 \equiv \hat{\mathbf{e}}_z & & V_3 \equiv V_z \end{matrix} . \tag{1.2}$$

Equation 1.1 becomes

$$\overline{\mathbf{V}} = V_1 \hat{\mathbf{e}}_1 + V_2 \hat{\mathbf{e}}_2 + V_3 \hat{\mathbf{e}}_3, \tag{1.3}$$

or more succinctly,

$$\overline{\mathbf{V}} = \sum_{i=1,2,3} V_i \hat{\mathbf{e}}_i. \tag{1.4}$$

Figure 1.1 shows this notational modification on a typical Cartesian coordinate system.

Although subscript notation can be used in many different types of coordinate systems, in this chapter we limit our discussion to Cartesian systems. Cartesian basis vectors are orthonormal and position independent. *Orthonormal* means the magnitude of each basis vector is unity, and they are all perpendicular to one another. *Position independent* means the basis vectors do not change their orientations as we move around in space. Non-Cartesian coordinate systems are covered in detail in Chapter 3.

Equation 1.4 can be compacted even further by introducing the *Einstein summation convention*, which assumes a summation any time a subscript is repeated in the same term. Therefore,

$$\overline{\mathbf{V}} = \sum_{i=1,2,3} V_i \hat{\mathbf{e}}_i \equiv V_i \hat{\mathbf{e}}_i. \tag{1.5}$$

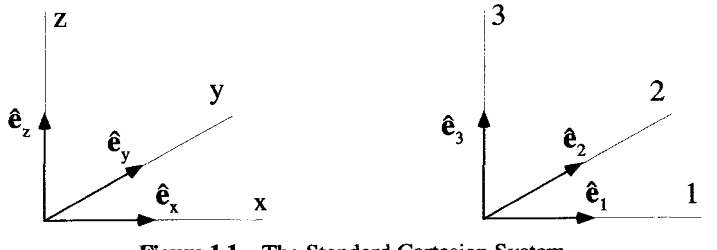

Figure 1.1 The Standard Cartesian System

We refer to this combination of the subscript notation and the summation convention as *subscript/summation notation.*

Now imagine we want to write the simple vector relationship

$$\overline{\mathbf{C}} = \overline{\mathbf{A}} + \overline{\mathbf{B}}. \tag{1.6}$$

This equation is written in what we call *vector notation.* Notice how it does not depend on a choice of coordinate system. In a particular coordinate system, we can write the relationship between these vectors in terms of their components:

$$\begin{aligned} C_1 &= A_1 + B_1 \\ C_2 &= A_2 + B_2 \\ C_3 &= A_3 + B_3. \end{aligned} \tag{1.7}$$

With subscript notation, these three equations can be written in a single line,

$$C_i = A_i + B_i, \tag{1.8}$$

where the subscript i stands for any of the three values $(1, 2, 3)$. As you will see in many examples at the end of this chapter, the use of the subscript/summation notation can drastically simplify the derivation of many physical and mathematical relationships. Results written in subscript/summation notation, however, are tied to a particular coordinate system, and are often difficult to interpret. For these reasons, we will convert our final results back into vector notation whenever possible.

A matrix is a two-dimensional array of quantities that may or may not be associated with a particular coordinate system. Matrices can be expressed using several different types of notation. If we want to discuss a matrix in its entirety, without explicitly specifying all its elements, we write it in *matrix notation* as $[M]$. If we do need to list out the elements of $[M]$, we can write them as a rectangular array inside a pair of brackets:

$$[M] = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1c} \\ M_{21} & M_{22} & \cdots & M_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ M_{r1} & M_{r2} & \cdots & M_{rc} \end{bmatrix}. \tag{1.9}$$

We call this *matrix array notation.* The individual element in the second row and third column of $[M]$ is written as M_{23}. Notice how the row of a given element is always listed first, and the column second. Keep in mind, the array is not necessarily square. This means that for the matrix in Equation 1.9, r does not have to equal c.

Multiplication between two matrices is only possible if the number of columns in the premultiplier equals the number of rows in the postmultiplier. The result of such a multiplication forms another matrix with the same number of rows as the premultiplier and the same number of columns as the postmultiplier. For example, the product between a 3×2 matrix $[M]$ and a 2×3 matrix $[N]$ forms the 3×3 matrix

$[P]$, with the elements given by:

$$\underbrace{\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \\ M_{31} & M_{32} \end{bmatrix}}_{[M]} \underbrace{\begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \end{bmatrix}}_{[N]}$$

$$= \underbrace{\begin{bmatrix} M_{11}N_{11} + M_{12}N_{21} & M_{11}N_{12} + M_{12}N_{22} & M_{11}N_{13} + M_{12}N_{23} \\ M_{21}N_{11} + M_{22}N_{21} & M_{21}N_{12} + M_{22}N_{22} & M_{21}N_{13} + M_{22}N_{23} \\ M_{31}N_{11} + M_{32}N_{21} & M_{31}N_{12} + M_{32}N_{22} & M_{31}N_{13} + M_{32}N_{23} \end{bmatrix}}_{[P]}. \tag{1.10}$$

The multiplication in Equation 1.10 can be written in the abbreviated matrix notation as

$$[M][N] = [P]. \tag{1.11}$$

We can also use subscript/summation notation to write the same product as

$$M_{ij}N_{jk} = P_{ik}, \tag{1.12}$$

with the implied sum over the j index keeping track of the summation. Notice j is in the second position of the M_{ij} term and the first position of the N_{jk} term, so the summation is over the columns of $[M]$ and the rows of $[N]$, just as it was in Equation 1.10. Equation 1.12 is an expression for the ik^{th} element of the matrix $[P]$.

Matrix array notation is convenient for doing numerical calculations, especially when using a computer. When deriving the relationships between the various quantities in physics, however, matrix notation is often inadequate because it lacks a mechanism for keeping track of the geometry of the coordinate system. For example, in a particular coordinate system, the vector $\overline{\mathbf{V}}$ might be written as

$$\overline{\mathbf{V}} = 1\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 + 2\hat{\mathbf{e}}_3. \tag{1.13}$$

When performing calculations, it is sometimes convenient to use a matrix representation of this vector by writing:

$$\overline{\mathbf{V}} \rightarrow [V] = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}. \tag{1.14}$$

The problem with this notation is that there is no convenient way to incorporate the basis vectors into the matrix. This is why we are careful to use an arrow ($\rightarrow$) in Equation 1.14 instead of an equal sign ($=$). In this text, an equal sign between two quantities means that they are perfectly equivalent in every way. One quantity may be substituted for the other in any expression. For instance, Equation 1.13 implies that the quantity $1\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 + 2\hat{\mathbf{e}}_3$ can replace $\overline{\mathbf{V}}$ in any mathematical expression, and vice-versa. In contrast, the arrow in Equation 1.14 implies that $[V]$ can represent $\overline{\mathbf{V}}$, and that calculations can be performed using it, but we must be careful not to directly substitute one for the other without specifying the basis vectors associated with the components of $[V]$.

1.2 VECTOR OPERATIONS

In this section, we investigate several vector operations. We will use all the different forms of notation discussed in the previous section in order to illustrate their differences. Initially, we will concentrate on matrix and matrix array notation. As we progress, the subscript/summation notation will be used more frequently.

As we discussed earlier, a three-dimensional vector $\overline{\mathbf{V}}$ can be represented using a matrix. There are actually two ways to write this matrix. It can be either a (3×1) column matrix or a (1×3) row matrix, whose elements are the components of the vector in some Cartesian basis:

$$\overline{\mathbf{V}} \rightarrow [V] = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \qquad \text{or} \qquad \overline{\mathbf{V}} \rightarrow [V]^\dagger = [\ V_1 \quad V_2 \quad V_3\ ]. \tag{1.15}$$

The standard notation $[V]^\dagger$ has been used to indicate the transpose of $[V]$, indicating an interchange of rows and columns. Remember the vector $\overline{\mathbf{V}}$ can have an infinite number of different matrix array representations, each written with respect to a different coordinate basis.

1.2.1 Vector Rotation

Consider the simple rotation of a vector in a Cartesian coordinate system. This example will be worked out, without any real loss of generality, in two dimensions.

We start with the vector $\overline{\mathbf{A}}$, which is oriented at an angle θ to the 1-axis, as shown in Figure 1.2. This vector can be written in terms of its Cartesian components as

$$\overline{\mathbf{A}} = A_1\hat{\mathbf{e}}_1 + A_2\hat{\mathbf{e}}_2, \tag{1.16}$$

where

$$A_1 = A\cos\theta \qquad A_2 = A\sin\theta. \tag{1.17}$$

In these expressions $A \equiv |\overline{\mathbf{A}}| = \sqrt{A_1^2 + A_2^2}$ is the magnitude of the vector $\overline{\mathbf{A}}$. The vector $\overline{\mathbf{A}}'$ is generated by rotating the vector $\overline{\mathbf{A}}$ counterclockwise by an angle ϕ. This

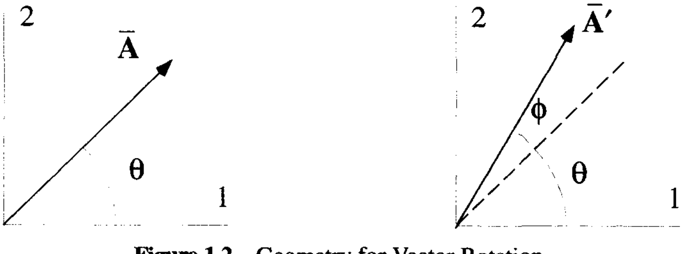

Figure 1.2 Geometry for Vector Rotation

changes the orientation of the vector, but not its magnitude. Therefore, we can write

$$\overline{\mathbf{A}}' = \underbrace{A\cos(\theta+\phi)}_{A_1'}\hat{\mathbf{e}}_1 + \underbrace{A\sin(\theta+\phi)}_{A_2'}\hat{\mathbf{e}}_2. \tag{1.18}$$

The components A_1' and A_2' can be rewritten using the trigonometric identities for the cosine and sine of summed angles as

$$\begin{aligned} A_1' &= A\cos(\theta+\phi) = \underbrace{A\cos\theta}_{A_1}\cos\phi - \underbrace{A\sin\theta}_{A_2}\sin\phi \\ A_2' &= A\sin(\theta+\phi) = \underbrace{A\cos\theta}_{A_1}\sin\phi + \underbrace{A\sin\theta}_{A_2}\cos\phi. \end{aligned} \tag{1.19}$$

If we represent $\overline{\mathbf{A}}$ and $\overline{\mathbf{A}}'$ with column matrices,

$$\overline{\mathbf{A}} \rightarrow [A] = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \qquad \overline{\mathbf{A}}' \rightarrow [A'] = \begin{bmatrix} A_1' \\ A_2' \end{bmatrix}, \tag{1.20}$$

Equations 1.19 can be put into matrix array form as

$$\begin{bmatrix} A_1' \\ A_2' \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}. \tag{1.21}$$

In the abbreviated matrix notation, we can write this as

$$[A'] = [R(\phi)][A]. \tag{1.22}$$

In this last expression, $[R(\phi)]$ is called the rotation matrix, and is clearly defined as

$$[R(\phi)] = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}. \tag{1.23}$$

Notice that for Equation 1.22 to be the same as Equation 1.19, and for the matrix multiplication to make sense, the matrices $[A]$ and $[A']$ must be column matrix arrays and $[R(\phi)]$ must premultiply $[A]$. The result of Equation 1.19 can also be written using the row representations for $\overline{\mathbf{A}}$ and $\overline{\mathbf{A}}'$. In this case, the transposes of $[R]$, $[A]$ and $[A']$ must be used, and $[R]^\dagger$ must postmultiply $[A]^\dagger$:

$$[A']^\dagger = [A]^\dagger\,[R]^\dagger\,. \tag{1.24}$$

Written out using matrix arrays, this expression becomes

$$[\,A_1' \quad A_2'\,] = [\,A_1 \quad A_2\,] \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}. \tag{1.25}$$

It is easy to see Equation 1.25 is entirely equivalent to Equation 1.21.

These same manipulations can be accomplished using subscript/summation notation. For example, Equation 1.19 can be expressed as

$$A_i' = R_{ij}A_j. \tag{1.26}$$

The matrix multiplication in Equation 1.22 sums over the columns of the elements of $[R]$. This is accomplished in Equation 1.26 by the implied sum over j. Unlike matrix notation, in subscript/summation notation the order of A_j and R_{ij} is no longer important, because

$$R_{ij}A_j = A_jR_{ij}. \tag{1.27}$$

The vector $\overline{\mathbf{A}}'$ can be written using the subscript notation by combining Equation 1.26 with the basis vectors

$$\overline{\mathbf{A}}' = R_{ij}A_j\hat{\mathbf{e}}_i. \tag{1.28}$$

This expression demonstrates a "notational bookkeeping" property of the subscript notation. Summing over a subscript removes its dependence from an expression, much like integrating over a variable. For this reason, the process of subscript summation is often called *contracting* over an index. There are two sums on the right-hand side (RHS) of Equation 1.28, one over the i and another over j. After contraction over both subscripts, the are no subscripts remaining on the RHS. This is consistent with the fact that there are no subscripts on the left-hand side (LHS). The only notation on the LHS is the "overbar" on $\overline{\mathbf{A}}'$, indicating a vector, which also exists on the RHS in the form of a "hat" on the basis vector $\hat{\mathbf{e}}_i$. This sort of notational analysis can be applied to all equations. The notation on the LHS of an equals sign must always agree with the notation on the RHS. This fact can be used to check equations for accuracy. For example,

$$\overline{\mathbf{A}}' \neq R_{ij}A_j, \tag{1.29}$$

because a subscript i remains on the RHS after contracting over j, while there are no subscripts at all on the LHS. In addition, the notation indicates the LHS is a vector quantity, while the RHS is not.

1.2.2 Vector Products

We now consider the dot and cross products of two vectors using subscript/summation notation. These products occur frequently in physics calculations, at every level. The dot product is usually first encountered when calculating the work W done by a force $\overline{\mathbf{F}}$ in the line integral

$$W = \int d\overline{\mathbf{r}} \cdot \overline{\mathbf{F}}. \tag{1.30}$$

In this equation, $d\overline{\mathbf{r}}$ is a differential displacement vector. The cross product can be used to find the force on a particle of charge q moving with velocity $\overline{\mathbf{v}}$ in an externally applied magnetic field $\overline{\mathbf{B}}$:

$$\overline{\mathbf{F}} = q(\overline{\mathbf{v}} \times \overline{\mathbf{B}}). \tag{1.31}$$

The Dot Product The dot or inner product of two vectors, $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$, is a scalar defined by

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = |\overline{\mathbf{A}}||\overline{\mathbf{B}}| \cos\theta, \tag{1.32}$$

where θ is the angle between the two vectors, as shown in Figure 1.3. If we take the dot product of a vector with itself, we get the magnitude squared of that vector:

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{A}} = |\overline{\mathbf{A}}|^2. \tag{1.33}$$

In subscript/summation notation, Equation 1.32 is written as

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = A_i \hat{\mathbf{e}}_i \cdot B_j \hat{\mathbf{e}}_j. \tag{1.34}$$

Notice we are using two different subscripts to form $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$. This is necessary to keep the summations independent in the manipulations that follow. The notational bookkeeping is working here, because there are no subscripts on the LHS, and none left on the RHS after contraction over both i and j. Only the basis vectors are involved in the dot product, so Equation 1.34 can be rewritten as

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = A_i B_j (\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j). \tag{1.35}$$

Because we are restricting our attention to Cartesian systems where the basis vectors are orthonormal, we have

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \tag{1.36}$$

The Kronecker delta,

$$\delta_{ij} \equiv \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \tag{1.37}$$

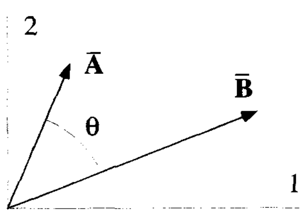

Figure 1.3 The Dot Product

facilitates calculations that involve dot products. Using it, we can write $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$, and Equation 1.35 becomes

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = A_i B_j \delta_{ij}. \tag{1.38}$$

Equation 1.38 can be expanded by explicitly doing the independent sums over both i and j

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = A_1 B_1 \delta_{11} + A_1 B_2 \delta_{12} + A_1 B_3 \delta_{13} + A_2 B_1 \delta_{21} + \cdots. \tag{1.39}$$

Since the Kronecker delta is zero unless its subscripts are equal, Equation 1.39 reduces to only three terms,

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = A_1 B_1 + A_2 B_2 + A_3 B_3 = A_i B_i. \tag{1.40}$$

As one becomes more familiar with the subscript/summation notation and the Kronecker delta, these last steps here are done automatically with the RHS of the brain. Anytime a Kronecker delta exists in a term, with one of its subscripts repeated somewhere else in the same term, the Kronecker delta can be removed, and every instance of the repeated subscript changed to the other subscript of the Kronecker symbol. For example,

$$A_i \delta_{ij} = A_j. \tag{1.41}$$

In Equation 1.38 the Kronecker delta can be grouped with the B_j factor, and contracted over j to give

$$A_i (B_j \delta_{ij}) = A_i B_i. \tag{1.42}$$

Just as well, we could group it with the A_i factor, and sum over i to give an equivalent result:

$$B_j (A_i \delta_{ij}) = B_j A_j. \tag{1.43}$$

This is true for more complicated expressions as well. For example,

$$\begin{gathered} M_{ij}(A_k \delta_{ik}) = M_{ij} A_i \\ \text{or} \\ B_i T_{jk}(\hat{\mathbf{e}}_m \delta_{jm}) = B_i T_{jk} \hat{\mathbf{e}}_j. \end{gathered} \tag{1.44}$$

This flexibility is one of the things that makes calculations performed with subscript/summation notation easier than working with matrix notation.

We should point out that the Kronecker delta can also be viewed as a matrix or matrix array. In three dimensions, this representation becomes

$$\delta_{ij} \to [1] \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.45}$$

This matrix can be used to write Equation 1.38 in matrix notation. Notice the contraction over the index i sums over the rows of the [1] matrix, while the contraction over j sums over the columns. Thus, Equation 1.38 in matrix notation is

$$\begin{aligned}\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} \rightarrow [A]^{\dagger}[1][B] &= [A_1 \quad A_2 \quad A_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} \\ &= [A]^{\dagger}[B]. \end{aligned} \tag{1.46}$$

The Cross Product The cross or vector product between two vectors $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$ forms a third vector $\overline{\mathbf{C}}$, which is written

$$\overline{\mathbf{C}} = \overline{\mathbf{A}} \times \overline{\mathbf{B}}. \tag{1.47}$$

The magnitude of the vector $\overline{\mathbf{C}}$ is

$$|\overline{\mathbf{C}}| = |\overline{\mathbf{A}}||\overline{\mathbf{B}}| \sin\theta, \tag{1.48}$$

where θ is the angle between the two vectors, as shown in Figure 1.4. The direction of $\overline{\mathbf{C}}$ depends on the "handedness" of the coordinate system. By convention, the three-dimensional coordinate systems in physics are usually "right-handed." Extend the fingers of your right hand straight, aligned along the basis vector $\hat{\mathbf{e}}_1$. Now, curl your fingers toward the basis vector $\hat{\mathbf{e}}_2$. If your thumb now points along $\hat{\mathbf{e}}_3$, the coordinate system is right-handed. When the coordinate system is arranged this way, the direction of the cross product follows a similar rule. To determine the direction of $\overline{\mathbf{C}}$ in Equation 1.47, point your fingers along $\overline{\mathbf{A}}$, and curl them to point along $\overline{\mathbf{B}}$. Your thumb is now pointing in the direction of $\overline{\mathbf{C}}$. This definition is usually called the *right-hand rule*. Notice, the direction of $\overline{\mathbf{C}}$ is always perpendicular to the plane formed by $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$. If, for some reason, we are using a left-handed coordinate system, the definition for the cross product changes, and we must instead use a *left-hand rule*. Because the definition of a cross product changes slightly when we move to

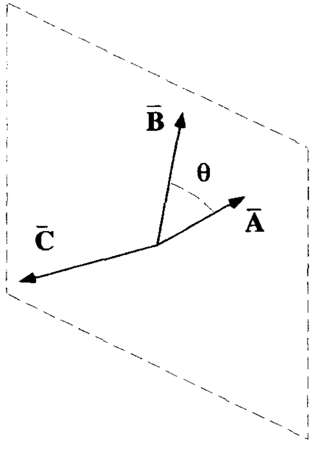

Figure 1.4 The Cross Product

systems of different handedness, the cross product is not exactly a vector, but rather a pseudovector. We will discuss this distinction in more detail at the end of Chapter 4. For now, we will limit our discussion to right-handed coordinate systems, and treat the cross product as an ordinary vector.

Another way to express the cross product is by using an unconventional determinant of a matrix, some of whose elements are basis vectors:

$$\overline{\mathbf{A}} \times \overline{\mathbf{B}} = \begin{vmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}_{det} . \tag{1.49}$$

Expanding the determinant of Equation 1.49 gives

$$\begin{aligned} \overline{\mathbf{A}} \times \overline{\mathbf{B}} = (A_2B_3 &- A_3B_2)\hat{\mathbf{e}}_1 \\ + (A_3B_1 &- A_1B_3)\hat{\mathbf{e}}_2 \\ + (A_1B_2 &- A_2B_1)\hat{\mathbf{e}}_3. \end{aligned} \tag{1.50}$$

This last expression can also be written using subscript/summation notation, with the introduction of the Levi-Civita symbol ϵ_{ijk}:

$$\overline{\mathbf{A}} \times \overline{\mathbf{B}} = A_iB_j\hat{\mathbf{e}}_k\epsilon_{ijk}, \tag{1.51}$$

where ϵ_{ijk} is defined as

$$\epsilon_{ijk} = \begin{cases} +1 & \text{for } (i,j,k) = \text{even permutations of } (1,2,3) \\ -1 & \text{for } (i,j,k) = \text{odd permutations of } (1,2,3) \\ 0 & \text{if two or more of the subscripts are equal} \end{cases} . \tag{1.52}$$

An odd permutation of $(1,2,3)$ is any rearrangement of the three numbers that can be accomplished with an odd number of pair interchanges. Thus, the odd permutations of $(1,2,3)$ are $(2,1,3)$, $(1,3,2)$, and $(3,2,1)$. Similarly, the even permutations of $(1,2,3)$ are $(1,2,3)$, $(2,3,1)$, and $(3,1,2)$. Because the subscripts i, j, and k can each independently take the values $(1,2,3)$, one way to visualize the Levi-Civita symbol is as the $3 \times 3 \times 3$ array shown in Figure 1.5.

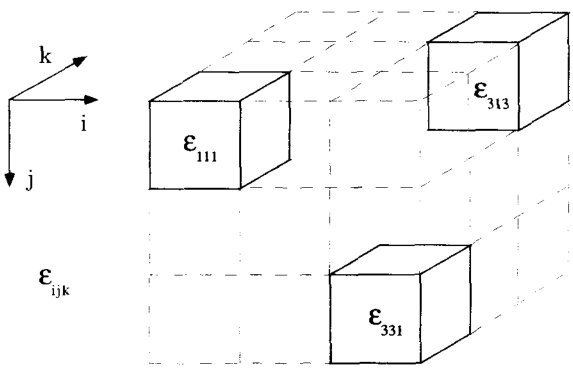

Figure 1.5 The $3 \times 3 \times 3$ Levi-Civita Array

The cross product, written using subscript/summation notation in Equation 1.51, and the dot product, written in the form of Equation 1.38 are very useful for manual calculations, as you will see in the following examples.

1.2.3 Calculations Using Subscript/Summation Notation

We now give two examples to demonstrate the use of subscript/summation notation. The first example shows that a vector's magnitude is unaffected by rotations. The primary function of this example is to show how a derivation performed entirely with matrix notation can also be done using subscript notation. The second derives a common vector identity. This example shows how the subscript notation is a powerful tool for deriving complicated vector relationships.

Example 1.1 Refer back to the rotation picture of Figure 1.2, and consider the products $\overline{\mathbf{A}} \cdot \overline{\mathbf{A}}$ and $\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}'$, first using matrix notation and then using subscript/summation notation. Because $\overline{\mathbf{A}}'$ is generated by a simple rotation of $\overline{\mathbf{A}}$ we know these two dot products, which represent the magnitude squared of the vectors, should be equal.

Using matrices:

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{A}} \rightarrow [A]^\dagger [A] \tag{1.53}$$

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' \rightarrow [A']^\dagger [A']. \tag{1.54}$$

But $[A']$ and $[A']^\dagger$ can be expressed in terms of $[A]$ and $[A]^\dagger$ as

$$[A'] = [R(\phi)][A] \qquad [A']^\dagger = [A]^\dagger [R(\phi)]^\dagger, \tag{1.55}$$

where $R(\phi)$ is the rotation matrix defined in Equation 1.23. If these two equations are substituted into Equation 1.54, we have

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' \rightarrow [A]^\dagger [R(\phi)]^\dagger [R(\phi)][A]. \tag{1.56}$$

The product between the two rotation matrices can be performed,

$$[R(\phi)]^\dagger [R(\phi)] = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{1.57}$$

and Equation 1.56 becomes

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' \rightarrow [A]^\dagger [1][A] = [A]^\dagger [A] \rightarrow \overline{\mathbf{A}} \cdot \overline{\mathbf{A}}. \tag{1.58}$$

Our final conclusion is that

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' = \overline{\mathbf{A}} \cdot \overline{\mathbf{A}}. \tag{1.59}$$

To arrive at this result using matrices, care had to be taken to be sure that the matrix operations were in the proper order and that the row, column, and transpose matrices were all used correctly.

Now let's repeat this derivation using the subscript/summation notation. Equation 1.40 allows us to write

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{A}} = A_i A_i \tag{1.60}$$

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' = A_r' A_r'. \tag{1.61}$$

Notice how we have been careful to use different subscripts for the two sums in Equations 1.60 and 1.61. This ensures the sums will remain independent as they are manipulated in the following steps. The primed components can be expressed in terms of the unprimed components as

$$A_i' = R_{ij} A_j, \tag{1.62}$$

where R_{ij} is the ij^{th} component of the rotation matrix $R[\phi]$. Inserting this expression into Equation 1.61 gives

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' = R_{ru} A_u R_{rv} A_v, \tag{1.63}$$

where again, we have been careful to use the two different subscripts u and v. This equation has three implicit sums, over the subscripts r, u, and v.

In subscript notation, unlike matrix notation, the ordering of the terms is not important, so we rearrange Equation 1.63 to read

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' = A_u A_v R_{ru} R_{rv}. \tag{1.64}$$

Next concentrate on the sum over r, which only involves the $[R]$ matrix elements, in the product $R_{ru}R_{rv}$. What exactly does this product mean? Let's compare it to an operation we discussed earlier. In Equation 1.12, we pointed out the subscripted expression $M_{ij}N_{jk}$ represented the regular matrix product $[M][N]$, because the summed subscript j is in the second position of the $[M]$ matrix and the first position of the $[N]$ matrix. The expression $R_{ru}R_{rv}$, however, has a contraction over the first index of *both* matrices. In order to make sense of this product, we write the first instance of $[R]$ using the transpose:

$$R_{ru} R_{rv} \rightarrow [R]^{\dagger}[R]. \tag{1.65}$$

Consequently, from Equation 1.57,

$$R_{ru} R_{rv} = \delta_{uv}. \tag{1.66}$$

Substituting this result into Equation 1.64 gives

$$\overline{\mathbf{A}}' \cdot \overline{\mathbf{A}}' = A_u A_v \delta_{uv} = A_u A_u = \overline{\mathbf{A}} \cdot \overline{\mathbf{A}}. \tag{1.67}$$

Admittedly, this example is too easy. It does not demonstrate any significant advantage of using the subscript/summation notation over matrices. It does, however, highlight the equivalence of the two approaches. In our next example, the subscript/summation notation will prove to be almost indispensable.

Example 1.2 The subscript/summation notation allows the derivation of vector identities that seem almost impossible using any other approach. The example worked out here is the derivation of an identity for the double cross product between three vectors, $\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}})$. This one example essentially demonstrates all the common operations that occur in these types of manipulations. Other examples are suggested in the problems listed at the end of this chapter.

The expression $\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}})$ is written in vector notation and is valid in any coordinate system. To derive our identity, we will convert this expression into subscript/summation notation in a Cartesian coordinate system. In the end, however, we will return our answer to vector notation to obtain a result that does not depend upon any coordinate system. In this example, we will need to use the subscripted form for a vector

$$\overline{\mathbf{V}} = V_i \hat{\mathbf{e}}_i, \tag{1.68}$$

for a dot product between two vectors

$$\overline{\mathbf{A}} \cdot \overline{\mathbf{B}} = A_i B_i, \tag{1.69}$$

and for a cross product

$$\overline{\mathbf{A}} \times \overline{\mathbf{B}} = A_i B_j \hat{\mathbf{e}}_k \epsilon_{ijk}. \tag{1.70}$$

To begin, let

$$\overline{\mathbf{D}} = \overline{\mathbf{B}} \times \overline{\mathbf{C}}, \tag{1.71}$$

which, written using the Levi-Civita symbol, is

$$\overline{\mathbf{D}} = B_i C_j \hat{\mathbf{e}}_k \epsilon_{ijk}. \tag{1.72}$$

Substituting Equation 1.71 into the expression $\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}})$, and using the Levi-Civita expression again, gives

$$\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}}) = \overline{\mathbf{A}} \times \overline{\mathbf{D}} = A_r D_s \hat{\mathbf{e}}_t \epsilon_{rst}. \tag{1.73}$$

The s^{th} component of $\overline{\mathbf{D}}$ is obtained by dot multiplying both sides of Equation 1.72 by $\hat{\mathbf{e}}_s$ as follows:

$$\begin{aligned} D_s = \hat{\mathbf{e}}_s \cdot \overline{\mathbf{D}} &= \hat{\mathbf{e}}_s \cdot B_i C_j \hat{\mathbf{e}}_k \epsilon_{ijk} \\ &= B_i C_j \epsilon_{ijk} (\hat{\mathbf{e}}_s \cdot \hat{\mathbf{e}}_k) \\ &= B_i C_j \epsilon_{ijk} \delta_{sk} \\ &= B_i C_j \epsilon_{ijs}. \end{aligned} \tag{1.74}$$

Substituting the result of Equation 1.74 into Equation 1.73 gives

$$\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}}) = A_r B_i C_j \epsilon_{ijs} \hat{\mathbf{e}}_t \epsilon_{rst}, \tag{1.75}$$

which we rearrange slightly to read

$$\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}}) = A_r B_i C_j \hat{\mathbf{e}}_t \epsilon_{rst} \epsilon_{ijs}. \tag{1.76}$$

To proceed, some properties of the Levi-Civita symbol need to be developed. First, because of the definition of the Levi-Civita symbol given in Equations 1.52, it is clear that reversing any two of its subscripts just changes its sign, i.e.,

$$\epsilon_{ijk} = -\epsilon_{ikj} = \epsilon_{jki}. \tag{1.77}$$

The second property involves the product of two Levi-Civita symbols that have a common last subscript:

$$\epsilon_{ijk}\epsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}. \tag{1.78}$$

With a considerable amount of effort, it can be shown that the RHS of Equation 1.78 has all the properties described by the product of the two Levi-Civita symbols on the LHS, each governed by Equations 1.52. A proof of this identity is given in Appendix A.

With Equations 1.77 and 1.78 we can return to Equation 1.76, which can now be rewritten

$$\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}}) = A_r B_i C_j \hat{\mathbf{e}}_t (\delta_{rj}\delta_{ti} - \delta_{ri}\delta_{tj}). \tag{1.79}$$

After removing the Kronecker deltas, we obtain

$$\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}}) = A_j B_i C_j \hat{\mathbf{e}}_i - A_i B_i C_j \hat{\mathbf{e}}_j. \tag{1.80}$$

At this point, you can really see the usefulness of the subscript/summation notation. The factors in the two terms on the RHS of Equation 1.80 can now be rearranged, grouped according to the summations, and returned to vector notation in just two lines! The procedure is:

$$\overline{\mathbf{A}} \times (\overline{\mathbf{B}} \times \overline{\mathbf{C}}) = (A_j C_j)(B_i \hat{\mathbf{e}}_i) - (A_i B_i)(C_j \hat{\mathbf{e}}_j) \tag{1.81}$$

$$= (\overline{\mathbf{A}} \cdot \overline{\mathbf{C}})\overline{\mathbf{B}} - (\overline{\mathbf{A}} \cdot \overline{\mathbf{B}})\overline{\mathbf{C}}. \tag{1.82}$$

Equation 1.81 is valid only in Cartesian systems. But because Equation 1.82 is in vector notation, it is valid in any coordinate system.

In the exercises that follow, you will derive several other vector identities. These will illustrate the power of the subscript/summation notation and help you become more familiar with its use.

EXERCISES FOR CHAPTER 1

1. Consider the two matrices:

$$[M] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad [N] = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}.$$

With matrix notation, a product between these two matrices can be expressed as $[M][N]$. Using subscript/summation notation, this same product is expressed as $M_{ij}N_{jk}$.

(a) With the elements of the $[M]$ and $[N]$ matrices given above, evaluate the matrix products $[M][N]$, $[N][M]$ and $[M][M]$, leaving the results in matrix array notation.

(b) Express the matrix products of part (a) using the subscript/summation notation.

(c) Convert the following expressions, which are written in subscript/summation notation, to matrix notation:

i. $M_{jk}N_{ij}$.

ii. $M_{ij}N_{kj}$.

iii. $M_{ji}N_{jk}$.

iv. $M_{ij}N_{jk} + T_{ki}$.

v. $M_{ji}N_{kj} + T_{ik}$.

2. Consider the square 3×3 matrix $[M]$ whose elements M_{ij} are generated by the expression

$$M_{ij} = ij^2 \quad \text{for } c, j = 1, 2, 3,$$

and a vector $\overline{\mathbf{V}}$ whose components in some basis are given by

$$V_k = k \quad \text{for } k = 1, 2, 3.$$

(a) Using a matrix representation for $\overline{\mathbf{V}} \rightarrow [V]$, determine the components of the vector that result from a premultiplication of $[V]$ by $[M]$.

(b) Determine the components of the vector that result from a premultiplication of $[M]$ by $[V]^\dagger$.

3. The matrix

$$[R] = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

represents a rotation. Show that the matrices $[R]^2 = [R][R]$ and $[R]^3 = [R][R][R]$ are matrices that correspond to rotations of 2θ and 3θ respectively.

4. Let $[D]$ be a 2×2 square matrix and $[V]$ a 2×1 row matrix. Determine the conditions imposed on $[D]$ by the requirement that

$$[D][V] = [V]^\dagger[D]$$

for any $[V]$.

5. The trace of a matrix is the sum of all its diagonal elements. Using the subscript/summation notation to represent the elements of the matrices $[T]$ and $[M]$,