Computational Continuum Mechanics

Ahmed A. Shabana

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521885690

This page intentionally left blank

COMPUTATIONAL CONTINUUM MECHANICS

This book presents the nonlinear theory of continuum mechanics and demonstrates its use in developing nonlinear computer formulations for large displacement dynamic analysis. Basic concepts used in continuum mechanics are presented and used to develop nonlinear general finite element formulations that can be effectively used in large displacement analysis. The book considers two nonlinear finite element dynamic formulations: a general large-deformation finite element formulation and then a formulation that can efficiently solve small deformation problems that characterize very and moderately stiff structures. The book presents material clearly and systematically, assuming the reader has only basic knowledge in matrix and vector algebra and dynamics. The book is designed for use by advanced undergraduates and first-year graduate students. It is also a reference for researchers, practicing engineers, and scientists working in computational mechanics, bio-mechanics, computational biology, multibody system dynamics, and other fields of science and engineering using the general continuum mechanics theory.

Ahmed A. Shabana is the Richard and Loan Hill Professor of Engineering at the University of Illinois, Chicago. Dr. Shabana received his PhD in mechanical engineering from the University of Iowa. His active areas of research interest are in dynamics, vibration, and control of mechanical systems. He is also the author of other books, including *Theory of Vibration: An Introduction, Vibration of Discrete and Continuous Systems, Computational Dynamics, Railroad Vehicle Dynamics: A Computational Approach*, and *Dynamics of Multibody Systems, Third Edition.*

Computational Continuum Mechanics

Ahmed A. Shabana
Richard and Loan Hill Professor of Engineering
University of Illinois at Chicago

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York
www.cambridge.org
Information on this title: www.cambridge.org/9780521885690

First published in print format 2008

ISBN-13 978-0-511-38059-4 eBook (Adobe Reader)

ISBN-13 978-0-521-88569-0 hardback

Contents

Preface *page* ix

1 **Introduction** 1

1.1 Matrices 2
1.2 Vectors 6
1.3 Summation Convention 12
1.4 Cartesian Tensors 13
1.5 Polar Decomposition Theorem 25
1.6 D'Alembert's Principle 27
1.7 Virtual Work Principle 34
1.8 Approximation Methods 37
1.9 Discrete Equations 40
1.10 Momentum, Work, and Energy 43
1.11 Parameter Change and Coordinate Transformation 45
PROBLEMS 48

2 **Kinematics** 51

2.1 Motion Description 52
2.2 Strain Components 60
2.3 Other Deformation Measures 67
2.4 Decomposition of Displacement 69
2.5 Velocity and Acceleration 71
2.6 Coordinate Transformation 75
2.7 Objectivity 82
2.8 Change of Volume and Area 85
2.9 Continuity Equation 89
2.10 Reynolds' Transport Theorem 90
2.11 Examples of Deformation 92
PROBLEMS 100

3 **Forces and Stresses** 103
3.1 Equilibrium of Forces 103
3.2 Transformation of Stresses 106
3.3 Equations of Equilibrium 107
3.4 Symmetry of the Cauchy Stress Tensor 109
3.5 Virtual Work of the Forces 111
3.6 Deviatoric Stresses 120
3.7 Stress Objectivity 123
3.8 Energy Balance 127
PROBLEMS 129

4 **Constitutive Equations** 131
4.1 Generalized Hooke's Law 132
4.2 Anisotropic Linearly Elastic Materials 134
4.3 Material Symmetry 135
4.4 Homogeneous Isotropic Material 137
4.5 Principal Strain Invariants 144
4.6 Special Material Models for Large Deformations 146
4.7 Linear Viscoelasticity 150
4.8 Nonlinear Viscoelasticity 164
4.9 A Simple Viscoelastic Model for Isotropic Materials 171
4.10 Fluid Constitutive Equations 173
4.11 Navier-Stokes Equations 174
PROBLEMS 175

5 **Plasticity Formulations** 177
5.1 One-Dimensional Problem 179
5.2 Loading and Unloading Conditions 180
5.3 Solution of the Plasticity Equations 181
5.4 Generalization of the Plasticity Theory: Small Strains 190
5.5 J_2 Flow Theory with Isotropic/Kinematic Hardening 197
5.6 Nonlinear Formulation for Hyperelastic–Plastic Materials 214
5.7 Hyperelastic–Plastic J_2 Flow Theory 225
PROBLEMS 230

6 **Finite Element Formulation: Large-Deformation, Large-Rotation Problem** 231
6.1 Displacement Field 233
6.2 Element Connectivity 240
6.3 Inertia and Elastic Forces 243
6.4 Equations of Motion 246
6.5 Numerical Evaluation of the Elastic Forces 250
6.6 Finite Elements and Geometry 256
6.7 Two-Dimensional Euler–Bernoulli Beam Element 263

6.8 Two-Dimensional Shear Deformable Beam Element 267
6.9 Three-Dimensional Cable Element 269
6.10 Three-Dimensional Beam Element 270
6.11 Thin-Plate Element 272
6.12 Higher-Order Plate Element 274
6.13 Element Performance 275
6.14 Other Finite Element Formulations 280
6.15 Updated Lagrangian and Eulerian Formulations 282
PROBLEMS 284

7 **Finite Element Formulation: Small-Deformation, Large-Rotation Problem** 286

7.1 Background 287
7.2 Rotation and Angular Velocity 291
7.3 Floating Frame of Reference 296
7.4 Intermediate Element Coordinate System 297
7.5 Connectivity and Reference Conditions 300
7.6 Kinematic Equations 306
7.7 Formulation of the Inertia Forces 307
7.8 Elastic Forces 311
7.9 Equations of Motion 313
7.10 Coordinate Reduction 314
7.11 Integration of Finite Element and Multibody System Algorithms 317
PROBLEMS 319

References 321
Index 327

Preface

Nonlinear continuum mechanics is one of the fundamental subjects that form the foundation of modern computational mechanics. The study of the motion and behavior of materials under different loading conditions requires understanding of basic, general, and nonlinear, kinematic and dynamic relationships that are covered in continuum mechanics courses. The finite element method, on the other hand, has emerged as a powerful tool for solving many problems in engineering and physics. The finite element method became a popular and widely used computational approach because of its versatility and generality in solving large-scale and complex physics and engineering problems. Nonetheless, the success of using the continuum-mechanics-based finite element method in the analysis of the motion of bodies that experience general displacements, including arbitrary large rotations, has been limited. The solution to this problem requires resorting to some of the basic concepts in continuum mechanics and putting the emphasis on developing sound formulations that satisfy the principles of mechanics. Some researchers, however, have tried to solve fundamental formulation problems using numerical techniques that lead to approximations. Although numerical methods are an integral part of modern computational algorithms and can be effectively used in some applications to obtain efficient and accurate solutions, it is the opinion of many researchers that numerical methods should only be used as a last resort to fix formulation problems. Sound formulations must be first developed and tested to make sure that these formulations satisfy the basic principles of mechanics. The equations that result from the use of the analytically correct formulations can then be solved using numerical methods.

This book is focused on presenting the nonlinear theory of continuum mechanics and demonstrating its use in developing nonlinear computer formulations that can be used in the large displacement dynamic analysis. To this end, the basic concepts used in continuum mechanics are first presented and then used to develop nonlinear general finite element formulations that can be effectively used in the large displacement analysis. Two nonlinear finite element dynamic formulations will be considered in this book. The first is a general large-deformation finite element formulation, whereas the second is a formulation that can be used efficiently to solve small-deformation problems that characterize very and moderately stiff structures.

In this latter case, an elaborate method for eliminating the unnecessary degrees of freedom must be used in order to be able to efficiently obtain a numerical solution. An attempt has been made to present the materials in a clear and systematic manner with the assumption that the reader has only basic knowledge in matrix and vector algebra as well as basic knowledge of dynamics. The book is designed for a course at the senior undergraduate and first-year graduate level. It can also be used as a reference for researchers and practicing engineers and scientists who are working in the areas of computational mechanics, biomechanics, computational biology, multibody system dynamics, and other fields of science and engineering that are based on the general continuum mechanics theory.

In **Chapter 1** of this book, matrix, vector, and tensor notations are introduced. These notations will be repeatedly used in all chapters of the book, and, therefore, it is necessary that the reader reviews this chapter in order to be able to follow the presentation in subsequent chapters. The polar decomposition theorem, which is fundamental in continuum and computational mechanics, is also presented in this chapter. D'Alembert's principle and the principle of virtual work can be used to systematically derive the equations of motion of physical systems. These two important principles are discussed, and the relationship between them is explained. The use of a finite dimensional model to describe the continuum motion is also discussed in Section 8; whereas in Section 9, the procedure for developing the discrete equations of motion is outlined. In Section 10, the principles of momentum and principle of work and energy are presented. In this section, the problems associated with some of the finite element formulations that violate these analytical mechanics principles are discussed. Section 11 of Chapter 1 is devoted to a discussion on the definitions of the gradient vectors that are used in continuum mechanics to define the strain components.

In **Chapter 2**, the general kinematic displacement equations of a continuum are developed. These equations are used to define the strain components. The Green–Lagrange strains and the Almansi or Eulerian strains are introduced. The Green–Lagrange strains are defined in the reference configuration, whereas the Almansi or Eulerian strains are defined in the current deformed configuration. The relationships between these strain components are established and used to shed light on the physical meaning of the strain components. Other deformation measures as well as the velocity and acceleration equations are also defined in this chapter. The important issue of objectivity that must be considered when large deformations and inelastic formulations are used is discussed. The equations that govern the change of volume and area, the conservation of mass, and examples of deformation modes are also presented in this chapter.

Forces and stresses are discussed in **Chapter 3**. Equilibrium of forces acting on an infinitesimal material element is used to define the Cauchy stresses, which are used to develop the partial differential equations of equilibrium. The transformation of the stress components and the symmetry of the Cauchy stress tensor are among the topics discussed in this chapter. The virtual work of the forces due to the change of the shape of the continuum is defined. The deviatoric stresses, stress objectivity, and energy balance equations are also discussed in Chapter 3.

The definition of the strain and stress components is not sufficient to describe the motion of a continuum. One must define the relationship between the stresses and strains using the constitutive equations that are discussed in **Chapter 4**. In Chapter 4, the generalized Hooke's law is introduced, and the assumptions used in the definition of homogeneous isotropic materials are outlined. The principal strain invariants and special large-deformation material models are discussed. The linear and nonlinear viscoelastic material behavior is also discussed in Chapter 4.

In many engineering applications, plastic deformations occur due to excessive forces and impact as well as thermal loads. Several plasticity formulations are presented in **Chapter 5**. First, a one-dimensional theory is used in order to discuss the main concepts and solution procedures used in the plasticity analysis. The theory is then generalized to the three-dimensional analysis for the case of small strains. Large strain nonlinear plasticity formulations as well as the J_2 flow theory are among the topics discussed in Chapter 5. This chapter can be skipped in its entirety because it has no effect on the continuity of the presentation, and the developments in subsequent chapters do not depend on the theory of plasticity in particular.

Nonlinear finite element formulations are discussed in Chapter 6 and 7. Two formulations are discussed in these two chapters. The first is a large-deformation finite element formulation, which is discussed in **Chapter 6**. This formulation, called the absolute nodal coordinate formulation, is based on a continuum mechanics theory and employs displacement gradients as coordinates. It leads to a unique displacement and rotation fields and imposes no restrictions on the amount of rotation or deformation within the finite element. The absolute nodal coordinate formulation has some unique features that distinguish it from other existing large-deformation finite element formulations: it leads to a constant mass matrix; it leads to zero centrifugal and Coriolis forces; it automatically satisfies the principles of mechanics; it correctly describes an arbitrary rigid-body motion including finite rotations; and it can be used to develop several beams, plate, and shell elements that relax many of the assumptions used in classical theorems because this formulation allows for the use of more general constitutive relationships.

Clearly, large-deformation finite element formulations can also be used to solve small deformation problems. However, it is not recommended to use a large-deformation finite element formulation to solve a small-deformation problem. Large-deformation formulations do not exploit some particular features of small-deformation problems, and, therefore, such formulations can be very inefficient in the solution of stiff and moderately stiff systems. It turns out that the development of an efficient small-deformation finite element formulation that correctly describes an arbitrary rigid-body motion requires the use of more elaborate techniques in order to define a local linear problem without compromising the ability of the method to describe large-displacement small-deformation behavior. The finite element floating frame of reference formulation, which is widely used in the analysis of small deformations, is discussed in **Chapter 7** of this book. This formulation allows eliminating high-frequency modes that do not have a significant effect on

the solution, thereby leading to a lower-dimension dynamic model that can be efficiently solved using numerical and computer methods.

I would like to thank many students and colleagues with whom I worked for several years on the subject of flexible body dynamics. I was fortunate to collaborate with excellent students and colleagues who educated me in this important field of computational mechanics. In particular, I would like to thank two of my doctorate students, Bassam Hussein and Luis Maqueda, who provided solutions for several of the examples presented in Chapter 4 and Chapter 5. I am grateful for the help I received from Mr. Peter Gordon, the Engineering Editor, and the production staff of Cambridge University Press. It was a pleasant experience working with them on the production of this book. I would also like to thank my family for their help, patience, and understanding during the time of preparing this book.

Ahmed A. Shabana
Chicago, IL, 2007

1 INTRODUCTION

Matrix, vector, and tensor algebras are often used in the theory of continuum mechanics in order to have a simpler and more tractable presentation of the subject. In this chapter, the mathematical preliminaries required to understand the matrix, vector, and tensor operations used repeatedly in this book are presented. Principles of mechanics and approximation methods that represent the basis for the formulation of the kinematic and dynamic equations developed in this book are also reviewed in this chapter. In the first two sections of this chapter, matrix and vector notations are introduced and some of their important identities are presented. Some of the vector and matrix results are presented without proofs because it is assumed that the reader has some familiarity with matrix and vector notations. In Section 3, the summation convention, which is widely used in continuum mechanics texts, is introduced. This introduction is made despite the fact that the summation convention is rarely used in this book. Tensor notations, on the other hand, are frequently used in this book and, for this reason, tensors are discussed in Section 4. In Section 5, the *polar decomposition theorem*, which is fundamental in continuum mechanics, is presented. This theorem states that any nonsingular square matrix can be decomposed as the product of an orthogonal matrix and a symmetric matrix. Other matrix decompositions that are used in computational mechanics are also discussed. In Section 6, D'Alembert's principle is introduced, while Section 7 discusses the virtual work principle. The finite element method is often used to obtain finite dimensional models of continuous systems that in reality have infinite number of degrees of freedom. To introduce the reader to some of the basic concepts used to obtain finite dimensional models, discussions of approximation methods are included in Section 8. The procedure for developing the discrete equations of motion is outlined in Section 9, while the principle of conservation of momentum and the principle of work and energy are discussed in Section 10. In continuum mechanics, the gradients of the position vectors can be determined by differentiation with respect to different parameters. The change of parameters can lead to the definitions of strain components in different directions. This change of parameters, however, does not change the coordinate system in which the gradient vectors are defined. The effect of the change of parameters on the definitions of the gradients is discussed in Section 11.

1.1 MATRICES

In this section, some identities, results, and properties from matrix algebra that are used repeatedly in this book are presented. Some proofs are omitted, with the assumption that the reader is familiar with the subject of linear algebra.

Definitions An $m \times n$ matrix $\mathbf{A}$ is an ordered rectangular array, which can be written in the following form:

$$\mathbf{A} = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{1.1}$$

where a_{ij} is the ijth element that lies in the ith row and jth column of the matrix. Therefore, the first subscript i refers to the row number, and the second subscript j refers to the column number. The arrangement of Equation 1 shows that the matrix $\mathbf{A}$ has m rows and n columns. If $m = n$, the matrix is said to be *square*, otherwise the matrix is said to be *rectangular*. The *transpose* of an $m \times n$ matrix $\mathbf{A}$ is an $n \times m$ matrix, denoted as $\mathbf{A}^{\mathrm{T}}$, which is obtained from $\mathbf{A}$ by exchanging the rows and columns, that is $\mathbf{A}^{\mathrm{T}} = (a_{ji})$.

A *diagonal matrix* is a square matrix whose only nonzero elements are the diagonal elements, that is, $a_{ij} = 0$ if $i \neq j$. An *identity* or *unit matrix*, denoted as $\mathbf{I}$, is a diagonal matrix that has all its diagonal elements equal to one. The *null* or *zero matrix* is a matrix that has all its elements equal to zero. The *trace* of a square matrix $\mathbf{A}$ is the sum of all its diagonal elements, that is,

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} \tag{1.2}$$

This equation shows that $\mathrm{tr}(\mathbf{I}) = n$, where $\mathbf{I}$ is the identity matrix and n is the dimension of the matrix.

A square matrix $\mathbf{A}$ is said to be *symmetric* if

$$\mathbf{A} = \mathbf{A}^{\mathrm{T}}, \quad a_{ij} = a_{ji} \tag{1.3}$$

A square matrix is said to be *skew symmetric* if

$$\mathbf{A} = -\mathbf{A}^{\mathrm{T}}, \quad a_{ij} = -a_{ji} \tag{1.4}$$

This equation shows that all the diagonal elements of a skew-symmetric matrix must be equal to zero. That is, if $\mathbf{A}$ is a skew-symmetric matrix with dimension n, then $a_{ii} = 0$ for $i = 1, 2, \dots, n$. Any square matrix can be written as the sum of

a symmetric matrix and a skew-symmetric matrix. For example, if $\mathbf{B}$ is a square matrix, $\mathbf{B}$ can be written as

$$\mathbf{B} = \bar{\mathbf{B}} + \tilde{\mathbf{B}} \tag{1.5}$$

where $\bar{\mathbf{B}}$ and $\tilde{\mathbf{B}}$ are, respectively, symmetric and skew-symmetric matrices defined as

$$\bar{\mathbf{B}} = \frac{1}{2}\left(\mathbf{B} + \mathbf{B}^{\mathrm{T}}\right), \quad \tilde{\mathbf{B}} = \frac{1}{2}\left(\mathbf{B} - \mathbf{B}^{\mathrm{T}}\right) \tag{1.6}$$

Skew-symmetric matrices are used in continuum mechanics to characterize the rotations of the material elements.

Determinant The *determinant* of an $n \times n$ square matrix $\mathbf{A}$, denoted as $|\mathbf{A}|$ or $\det(\mathbf{A})$, is a scalar quantity. In order to be able to define the unique value of the determinant, some basic definitions have to be introduced. The *minor* M_{ij} corresponding to the element a_{ij} is the determinant of a matrix obtained by deleting the ith row and jth column from the original matrix $\mathbf{A}$. The *cofactor* C_{ij} of the element a_{ij} is defined as

$$C_{ij} = (-1)^{i+j} M_{ij} \tag{1.7}$$

Using this definition, the determinant of the matrix $\mathbf{A}$ can be obtained in terms of the cofactors of the elements of an arbitrary row j as follows:

$$|\mathbf{A}| = \sum_{k=1}^{n} a_{jk} C_{jk} \tag{1.8}$$

One can show that the determinant of a diagonal matrix is equal to the product of the diagonal elements, and the determinant of a matrix is equal to the determinant of its transpose; that is, if $\mathbf{A}$ is a square matrix, then $|\mathbf{A}| = \left|\mathbf{A}^{\mathrm{T}}\right|$. Furthermore, the interchange of any two columns or rows only changes the sign of the determinant. It can also be shown that if the matrix has linearly dependent rows or linearly dependent columns, the determinant is equal to zero. A matrix whose determinant is equal to zero is called a *singular matrix*. For an arbitrary square matrix, singular or nonsingular, it can be shown that the value of the determinant does not change if any row or column is added or subtracted from another.

It can be shown that the determinant of the product of two matrices is equal to the product of their determinants. That is, if $\mathbf{A}$ and $\mathbf{B}$ are two square matrices, then $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

As will be shown in this book, the determinants of some of the deformation measures used in continuum mechanics are used in the formulation of the energy expressions. Furthermore, the relationship between the volume of a continuum in the undeformed state and the deformed state is expressed in terms of the

determinant of the matrix of position vector gradients. Therefore, if the elements of a square matrix depend on a parameter, it is important to be able to determine the derivatives of the determinant with respect to this parameter. Using Equation 8, one can show that if the elements of the matrix **A** depend on a parameter t, then

$$\frac{d}{dt}|\mathbf{A}| = \sum_{k=1}^{n} \dot{a}_{1k} C_{1k} + \sum_{k=1}^{n} \dot{a}_{2k} C_{2k} + \ldots + \sum_{k=1}^{n} \dot{a}_{nk} C_{nk} \tag{1.9}$$

where $\dot{a}_{ij} = da_{ij}/dt$. The use of this equation is demonstrated by the following example.

EXAMPLE 1.1

Consider the matrix **J** defined as

$$\mathbf{J} = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix}$$

where $J_{ij} = \partial r_i / \partial x_j$, and **r** and **x** are the vectors

$$\mathbf{r}(x_1, x_2, x_3, t) = [\, r_1 \quad r_2 \quad r_3 \,]^{\mathrm{T}}, \quad \mathbf{x} = [\, x_1 \quad x_2 \quad x_3 \,]^{\mathrm{T}}$$

That is, the elements of the vector **r** are functions of the coordinates x_1, x_2, and x_3 and the parameter t. If $J = |\mathbf{J}|$ is the determinant of **J**, prove that

$$\frac{dJ}{dt} = \left(\frac{\partial \dot{r}_1}{\partial r_1} + \frac{\partial \dot{r}_2}{\partial r_2} + \frac{\partial \dot{r}_3}{\partial r_3} \right) J$$

where $\partial \dot{r}_i / \partial r_j = \left(\partial / \partial r_j\right)(dr_i/dt)$, $i, j = 1, 2, 3$.

Solution: Using Equation 9, one can write

$$\frac{dJ}{dt} = \sum_{k=1}^{3} \dot{J}_{1k} C_{1k} + \sum_{k=1}^{3} \dot{J}_{2k} C_{2k} + \sum_{k=1}^{3} \dot{J}_{3k} C_{3k}$$

where C_{ij} is the cofactor associated with element J_{ij}. Note that the preceding equation can be written as

$$\frac{dJ}{dt} = \begin{vmatrix} \dot{J}_{11} & \dot{J}_{12} & \dot{J}_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{vmatrix} + \begin{vmatrix} J_{11} & J_{12} & J_{13} \\ \dot{J}_{21} & \dot{J}_{22} & \dot{J}_{23} \\ J_{31} & J_{32} & J_{33} \end{vmatrix} + \begin{vmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ \dot{J}_{31} & \dot{J}_{32} & \dot{J}_{33} \end{vmatrix}$$

In this equation,

$$\dot{J}_{ij} = \frac{\partial \dot{r}_i}{\partial x_j} = \frac{\partial \dot{r}_i}{\partial r_1}\frac{\partial r_1}{\partial x_j} + \frac{\partial \dot{r}_i}{\partial r_2}\frac{\partial r_2}{\partial x_j} + \frac{\partial \dot{r}_i}{\partial r_3}\frac{\partial r_3}{\partial x_j} = \sum_{k=1}^{3} \frac{\partial \dot{r}_i}{\partial r_k} J_{kj}$$

Using this expansion, one can show that

$$\begin{vmatrix} \dot{J}_{11} & \dot{J}_{12} & \dot{J}_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{vmatrix} = \left(\frac{\partial \dot{r}_1}{\partial r_1}\right) J$$

Similarly, one can show that

$$\begin{vmatrix} J_{11} & J_{12} & J_{13} \\ \dot{J}_{21} & \dot{J}_{22} & \dot{J}_{23} \\ J_{31} & J_{32} & J_{33} \end{vmatrix} = \left(\frac{\partial \dot{r}_2}{\partial r_2}\right) J, \quad \begin{vmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ \dot{J}_{31} & \dot{J}_{32} & \dot{J}_{33} \end{vmatrix} = \left(\frac{\partial \dot{r}_3}{\partial r_3}\right) J$$

Using the preceding equations, it is clear that

$$\frac{dJ}{dt} = \left(\frac{\partial \dot{r}_1}{\partial r_1} + \frac{\partial \dot{r}_2}{\partial r_2} + \frac{\partial \dot{r}_3}{\partial r_3}\right) J$$

This matrix identity is important and is used in this book to evaluate the rate of change of the determinant of the matrix of position vector gradients in terms of important deformation measures.

Inverse and Orthogonality A square matrix $\mathbf{A}^{-1}$ that satisfies the relationship

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \tag{1.10}$$

where $\mathbf{I}$ is the identity matrix, is called the *inverse* of the matrix $\mathbf{A}$. The inverse of the matrix $\mathbf{A}$ is defined as

$$\mathbf{A}^{-1} = \frac{\mathbf{C}_t}{|\mathbf{A}|} \tag{1.11}$$

where $\mathbf{C}_t$ is the *adjoint* of the matrix $\mathbf{A}$. The adjoint matrix $\mathbf{C}_t$ is the transpose of the matrix of the cofactors (C_{ij}) of the matrix $\mathbf{A}$. One can show that the determinant of the inverse $|\mathbf{A}^{-1}|$ is equal to $1/|\mathbf{A}|$.

A square matrix is said to be *orthogonal* if

$$\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I} \tag{1.12}$$

Note that in the case of an orthogonal matrix $\mathbf{A}$, one has

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{-1} \tag{1.13}$$

That is, the inverse of an orthogonal matrix is equal to its transpose. One can also show that if $\mathbf{A}$ is an orthogonal matrix, then $|\mathbf{A}| = \pm 1$; and if $\mathbf{A}_1$ and $\mathbf{A}_2$ are two orthogonal matrices that have the same dimensions, then their product $\mathbf{A}_1\mathbf{A}_2$ is also an orthogonal matrix.

Examples of orthogonal matrices are the 3×3 transformation matrices that define the orientation of coordinate systems. In the case of a right-handed coordinate system, one can show that the determinant of the transformation matrix is +1; this is a *proper orthogonal transformation*. If the right-hand rule is not followed, the determinant of the resulting orthogonal transformation is equal to –1, which is an *improper orthogonal transformation*, such as in the case of a reflection.

Matrix Operations The sum of two matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ is defined as

$$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij}) \tag{1.14}$$

In order to add two matrices, they must have the same dimensions. That is, the two matrices $\mathbf{A}$ and $\mathbf{B}$ must have the same number of rows and same number of columns in order to apply Equation 14.

The product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is another matrix $\mathbf{C}$ defined as

$$\mathbf{C} = \mathbf{AB} \tag{1.15}$$

The element c_{ij} of the matrix $\mathbf{C}$ is defined by multiplying the elements of the ith row in $\mathbf{A}$ by the elements of the jth column in $\mathbf{B}$ according to the rule

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_k a_{ik}b_{kj} \tag{1.16}$$

Therefore, the number of columns in $\mathbf{A}$ must be equal to the number of rows in $\mathbf{B}$. If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix, then $\mathbf{C}$ is an $m \times p$ matrix. In general, $\mathbf{AB} \neq \mathbf{BA}$. That is, matrix multiplication is not commutative. The *associative law* for matrix multiplication, however, is valid; that is, $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{ABC}$, provided consistent dimensions of the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are used.

1.2 VECTORS

Vectors can be considered special cases of matrices. An n-dimensional vector $\mathbf{a}$ can be written as

$$\mathbf{a} = (a_i) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = [a_1 \quad a_2 \quad \ldots \quad a_n]^{\mathrm{T}} \tag{1.17}$$

Therefore, it is assumed that the vector is a column, unless it is transposed to make it a row.

Because vectors can be treated as columns of matrices, the addition of vectors is the same as the addition of column matrices. That is, if $\mathbf{a} = (a_i)$ and $\mathbf{b} = (b_i)$ are two n-dimensional vectors, then $\mathbf{a} + \mathbf{b} = (a_i + b_i)$. Three different types of products, however, can be used with vectors. These are the *dot product*, the *cross product*, and the *outer* or *dyadic product*. The result of the dot product of two vectors is a scalar, the result of the cross product is a vector, and the result of the dyadic product is a matrix. These three different types of products are discussed in the following text.

Dot Product The *dot*, *inner*, or *scalar product* of two vectors **a** and **b** is defined as

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i \tag{1.18}$$

Note that the two vectors **a** and **b** must have the same dimension. The two vectors **a** and **b** are said to be *orthogonal* if

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b} = 0 \tag{1.19}$$

The *norm*, *magnitude*, or *length* of an n-dimensional vector is defined as

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\mathbf{a}^{\mathrm{T}}\mathbf{a}} = \sqrt{\sum_{i=1}^{n} (a_i)^2} \tag{1.20}$$

It is clear from this definition that the norm is always a positive number, and it is equal to zero only when **a** is the zero vector, that is, all the components of **a** are equal to zero.

In the special case of three-dimensional vectors, the dot product of two arbitrary three-dimensional vectors **a** and **b** can be written in terms of their norms as $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \alpha$, where α is the angle between the two vectors. A vector is said to be a *unit vector* if its norm is equal to one. It is clear from the definition of the norm given by Equation 20 that the absolute value of any element of a unit vector must not exceed one. A unit vector $\hat{\mathbf{a}}$ along the vector **a** can be simply obtained by dividing the vector by its norm. That is, $\hat{\mathbf{a}} = \mathbf{a}/|\mathbf{a}|$. The dot product $\mathbf{b} \cdot \hat{\mathbf{a}} = |\mathbf{b}| \cos \alpha$ defines the component of the vector **b** along the unit vector $\hat{\mathbf{a}}$, where α is the angle between the two vectors. The projection of the vector **b** on a plane perpendicular to the unit vector $\hat{\mathbf{a}}$ is defined by the equation $\mathbf{b} - (\mathbf{b} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}}$, or equivalently by $\mathbf{b} - (|\mathbf{b}| \cos \alpha)\hat{\mathbf{a}}$.

Cross Product The vector cross product is defined for three-dimensional vectors only. Let $\mathbf{a}$ and $\mathbf{b}$ be two three-dimensional vectors defined in the same coordinate system. Unit vectors along the axes of the coordinate system are denoted by the vectors $\mathbf{i}_1$, $\mathbf{i}_2$, and $\mathbf{i}_3$. These base vectors are *orthonormal*, that is,

$$\mathbf{i}_i \cdot \mathbf{i}_j = \delta_{ij} \tag{1.21}$$

where δ_{ij} is the *Kronecker delta* defined as

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{1.22}$$

The *cross product* of the two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as

$$\begin{aligned} \mathbf{c} = \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i}_1 & \mathbf{i}_2 & \mathbf{i}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \\ &= (a_2 b_3 - a_3 b_2)\mathbf{i}_1 + (a_3 b_1 - a_1 b_3)\mathbf{i}_2 + (a_1 b_2 - a_2 b_1)\mathbf{i}_3 \end{aligned} \tag{1.23}$$

which can be written as

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{1.24}$$

This equation can be written as

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \tilde{\mathbf{a}}\mathbf{b} \tag{1.25}$$

where $\tilde{\mathbf{a}}$ is the skew-symmetric matrix associated with the vector $\mathbf{a}$ and is defined as

$$\tilde{\mathbf{a}} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{1.26}$$

One can show that the determinant of the skew-symmetric matrix $\tilde{\mathbf{a}}$ is equal to zero. That is, $|\tilde{\mathbf{a}}| = 0$. One can also show that

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a} = -\tilde{\mathbf{b}}\mathbf{a} \tag{1.27}$$

In this equation, $\tilde{\mathbf{b}}$ is the skew-symmetric matrix associated with the vector $\mathbf{b}$. If $\mathbf{a}$ and $\mathbf{b}$ are two parallel vectors, it can be shown that

$$\mathbf{a} \times \mathbf{b} = \mathbf{0} \tag{1.28}$$

That is, the cross product of two parallel vectors is equal to zero.

Dyadic Product Another form of vector product used in this book is the *dyadic* or *outer product*. Whereas the dot product leads to a scalar and the cross product leads to a vector; the dyadic product leads to a matrix. The dyadic product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is written as $\mathbf{a} \otimes \mathbf{b}$ and is defined as

$$\mathbf{a} \otimes \mathbf{b} = \mathbf{a}\mathbf{b}^{\mathrm{T}} \tag{1.29}$$

Note that, in general, $\mathbf{a} \otimes \mathbf{b} \neq \mathbf{b} \otimes \mathbf{a}$. One can show that the dyadic product of two vectors satisfies the following identities:

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{c} = \mathbf{a}(\mathbf{b} \cdot \mathbf{c}), \quad \mathbf{a} \cdot (\mathbf{b} \otimes \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b})\mathbf{c}^{\mathrm{T}} \tag{1.30}$$

In Equation 30, it is assumed that the vectors have the appropriate dimensions. As a special case of the identities of Equation 30, one has $(\mathbf{a} \otimes \mathbf{i}_k)\mathbf{c} = c_k\mathbf{a}$ and $\mathbf{a} \cdot (\mathbf{i}_k \otimes \mathbf{c}) = a_k\mathbf{c}^{\mathrm{T}}$, where $\mathbf{i}_k$, $k = 1, 2, 3, \ldots$, are the base vectors and a_k and c_k are the kth elements of the vectors $\mathbf{a}$ and $\mathbf{c}$, respectively. Similarly, $(\mathbf{a} \otimes \mathbf{b})\mathbf{i}_k = b_k\mathbf{a}$ and $\mathbf{i}_k \cdot (\mathbf{a} \otimes \mathbf{b}) = a_k\mathbf{b}^{\mathrm{T}}$, which show that postmultiplying the dyadic product by one of the kth base vectors defines the kth element of the second vector multiplied by the first vector, whereas premultiplying the dyadic product by the kth base vector defines the kth element of the first vector multiplied by the second vector. The dyadic product satisfies the following additional properties for any arbitrary vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{v}_1$, and $\mathbf{v}_2$ and a square matrix $\mathbf{A}$:

$$\left.\begin{aligned}
&(\mathbf{u} \otimes \mathbf{v})^{\mathrm{T}} = \mathbf{v} \otimes \mathbf{u} \\
&\mathbf{A}(\mathbf{u} \otimes \mathbf{v}) = (\mathbf{A}\mathbf{u} \otimes \mathbf{v}) \\
&(\mathbf{u} \otimes \mathbf{v})\mathbf{A} = \left(\mathbf{u} \otimes \mathbf{A}^{\mathrm{T}}\mathbf{v}\right) \\
&\mathbf{u} \otimes (\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{u} \otimes \mathbf{v}_1 + \mathbf{u} \otimes \mathbf{v}_2
\end{aligned}\right\} \tag{1.31}$$

The second and third identities of Equation 31 show that $(\mathbf{A}\mathbf{u} \otimes \mathbf{A}\mathbf{v}) = \mathbf{A}(\mathbf{u} \otimes \mathbf{v})\mathbf{A}^{\mathrm{T}}$. This result is important in understanding the rule of transformation of the second-order tensors that will be discussed later in this chapter. It is left to the reader as an exercise to verify the identities of Equation 31.

EXAMPLE 1.2

Consider the two vectors $\mathbf{a} = [\,a_1 \quad a_2\,]^{\mathrm{T}}$ and $\mathbf{b} = [\,b_1 \quad b_2 \quad b_3\,]^{\mathrm{T}}$. The dyadic product of these two vectors is given by

$$\mathbf{a} \otimes \mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} [\,b_1 \quad b_2 \quad b_3\,] = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \end{bmatrix}$$

For a given vector $\mathbf{c} = [\,c_1 \quad c_2 \quad c_3\,]^{\mathrm{T}}$, one has

$$\begin{aligned}
(\mathbf{a} \otimes \mathbf{b})\mathbf{c} &= \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \\
&= \begin{bmatrix} a_1 b_1 \\ a_2 b_1 \end{bmatrix} c_1 + \begin{bmatrix} a_1 b_2 \\ a_2 b_2 \end{bmatrix} c_2 + \begin{bmatrix} a_1 b_3 \\ a_2 b_3 \end{bmatrix} c_3 \\
&= \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_1 c_1 + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_2 c_2 + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_3 c_3 = \mathbf{a}(\mathbf{b} \cdot \mathbf{c})
\end{aligned}$$

Also note that the dyadic product $\mathbf{a} \otimes \mathbf{b}$ can be written as

$$\mathbf{a} \otimes \mathbf{b} = \left[\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_1 \quad \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_2 \quad \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_3 \right] = [\mathbf{a} b_1 \quad \mathbf{a} b_2 \quad \mathbf{a} b_3]$$

It follows that if $\mathbf{R}$ is a 2×2 matrix, one has

$$\begin{aligned}
\mathbf{R}(\mathbf{a} \otimes \mathbf{b}) &= \mathbf{R}[\,\mathbf{a} b_1 \quad \mathbf{a} b_2 \quad \mathbf{a} b_3\,] = [\,(\mathbf{R}\mathbf{a}) b_1 \quad (\mathbf{R}\mathbf{a}) b_2 \quad (\mathbf{R}\mathbf{a}) b_3\,] \\
&= (\mathbf{R}\mathbf{a} \otimes \mathbf{b})
\end{aligned}$$

Several important identities can be written in terms of the dyadic product. Some of these identities are valuable in the computer implementation of the dynamic formulations presented in this book because the use of these identities can lead to significant simplification of the computational algorithms. By using these identities, one can avoid rewriting codes that perform the same mathematical operations, thereby saving effort and time by producing a manageable computer code. One of these identities that can be written in terms of the dyadic product is obtained in the following example.

EXAMPLE 1.3

In the computer implementation of the formulations presented in this book, one may require differentiating a unit vector $\hat{\mathbf{r}}$ along the vector $\mathbf{r}$ with respect to the components of the vector $\mathbf{r}$. Such a differentiation can be written in terms of the dyadic product. To demonstrate this, we write

$$\hat{\mathbf{r}} = \frac{1}{\sqrt{\mathbf{r}^{\mathrm{T}}\mathbf{r}}}\mathbf{r} = \frac{1}{|\mathbf{r}|}\mathbf{r}$$

where $|\mathbf{r}| = \sqrt{\mathbf{r}^{\mathrm{T}}\mathbf{r}}$. It follows that

$$\frac{\partial \hat{\mathbf{r}}}{\partial \mathbf{r}} = \frac{1}{\sqrt{\mathbf{r}^{\mathrm{T}}\mathbf{r}}}\left(\mathbf{I} - \frac{1}{\mathbf{r}^{\mathrm{T}}\mathbf{r}}\mathbf{r}\mathbf{r}^{\mathrm{T}}\right)$$

This equation can be written in terms of the dyadic product as

$$\frac{\partial \hat{\mathbf{r}}}{\partial \mathbf{r}} = \frac{1}{\sqrt{\mathbf{r}^{\mathrm{T}}\mathbf{r}}}\left(\mathbf{I} - \frac{1}{\mathbf{r}^{\mathrm{T}}\mathbf{r}}\mathbf{r} \otimes \mathbf{r}\right)$$

Projection If $\hat{\mathbf{a}}$ is a unit vector, the component of a vector $\mathbf{b}$ along the unit vector $\hat{\mathbf{a}}$ is defined by the dot product $\mathbf{b} \cdot \hat{\mathbf{a}}$. The projection of $\mathbf{b}$ along $\hat{\mathbf{a}}$ is then defined as $(\mathbf{b} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}}$, which can be written using Equation 30 as $(\mathbf{b} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}} = (\hat{\mathbf{a}} \otimes \hat{\mathbf{a}})\mathbf{b}$. The matrix $\mathbf{P} = \hat{\mathbf{a}} \otimes \hat{\mathbf{a}}$ defines a *projection matrix*. For an arbitrary integer n, one can show that the projection matrix $\mathbf{P}$ satisfies the identity $\mathbf{P}^n = \mathbf{P}$. This is an expected result because the vector $(\hat{\mathbf{a}} \otimes \hat{\mathbf{a}})\mathbf{b} = \mathbf{P}\mathbf{b}$ is defined along $\hat{\mathbf{a}}$ and has no components in other directions. Other projections should not change this result.

The projection of the vector $\mathbf{b}$ on a plane perpendicular to the unit vector $\hat{\mathbf{a}}$ is defined as $\mathbf{b} - (\mathbf{b} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}}$, which can be written using the dyadic product as $(\mathbf{I} - \hat{\mathbf{a}} \otimes \hat{\mathbf{a}})\mathbf{b}$. This equation defines another projection matrix $\mathbf{P}_p = \mathbf{I} - \hat{\mathbf{a}} \otimes \hat{\mathbf{a}}$, or simply $\mathbf{P}_p = \mathbf{I} - \mathbf{P}$. For an arbitrary integer n, one can show that the projection matrix $\mathbf{P}_p$ satisfies the identity $\mathbf{P}_p^n = \mathbf{P}_p$. Furthermore, $\mathbf{P}\mathbf{P}_p = \mathbf{0}$ and $\mathbf{P} + \mathbf{P}_p = \mathbf{I}$.

EXAMPLE 1.4

Consider the vector $\mathbf{a} = [1 \quad 2 \quad 0]^{\mathrm{T}}$. A unit vector along $\mathbf{a}$ is defined as

$$\hat{\mathbf{a}} = \frac{1}{\sqrt{5}}[1 \quad 2 \quad 0]^{\mathrm{T}}$$

The projection matrix $\mathbf{P}$ associated with this unit vector can be written as

$$\mathbf{P} = \hat{\mathbf{a}} \otimes \hat{\mathbf{a}} = \frac{1}{5}\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

It follows that

$$\mathbf{P}^2 = \frac{1}{25}\begin{bmatrix} 5 & 10 & 0 \\ 10 & 20 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{P}$$

The projection matrix $\mathbf{P}_p$ is defined in this example as

$$\mathbf{P}_p = \mathbf{I} - \hat{\mathbf{a}} \otimes \hat{\mathbf{a}} = \mathbf{I} - \mathbf{P} = \frac{1}{5}\begin{bmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that $\mathbf{P}_p^2 = (\mathbf{I} - \mathbf{P})^2 = \mathbf{I} - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{I} - \mathbf{P} = \mathbf{P}_p$. Successive application of this equation shows that $\mathbf{P}_p^n = \mathbf{P}_p$. The reader can verify this fact by the data given in this example.

1.3 SUMMATION CONVENTION

In this section, another convenient notational method, the *summation convention*, is discussed. The summation convention is used in most books on the subject of continuum mechanics. According to this convention, summation over the values of the indices is automatically assumed if an index is repeated in an expression. For example, if an index j takes the values from 1 to n, then in the summation convention, one has

$$a_{jj} = a_{11} + a_{22} + \ldots + a_{nn} \tag{1.32}$$

and

$$a_{ijj} = a_{i11} + a_{i22} + \ldots + a_{inn} \tag{1.33}$$

The repeated index used in the summation is called the *dummy index*, an example of which is the index j used in the preceding equation. If the index is not a dummy index, it is called a *free index*, an example of which is the index i used in Equation 33. It follows that the trace of a matrix $\mathbf{A}$ can be written using the summation convention as

$$\mathrm{tr}(\mathbf{A}) = a_{ii} \tag{1.34}$$

The dot product between two n-dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ can be written using the summation convention as

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b} = a_i b_i \tag{1.35}$$

The product of a matrix $\mathbf{A}$ and a vector $\mathbf{b}$ is another vector $\mathbf{c} = \mathbf{A}\mathbf{b}$ whose components can be written using the summation convention as

$$c_i = a_{ij} b_j \tag{1.36}$$

It follows that the components of an n-dimensional vector $\mathbf{a} = (a_i)$ defined by the multiplication $\mathbf{a} = \mathbf{R}\mathbf{b}$, where $\mathbf{R} = \left(R_{ij}\right)$ and $\mathbf{b} = (b_i)$, can be written using the summation convention as $a_i = R_{ij}b_j$. Here, i is the free index and j is the dummy index.

The dyadic product between two vectors can also be written using the summation convention. For example, in the case of three-dimensional vectors, one can define the base vectors $\mathbf{i}_k, k = 1, 2, 3$. Any three-dimensional vector can be written in terms of these base vectors using the summation convention as $\mathbf{a} = a_i\mathbf{i}_i = a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + a_3\mathbf{i}_3$. The dyadic product of two vectors $\mathbf{a}$ and $\mathbf{b}$ can then be written as

$$\mathbf{a} \otimes \mathbf{b} = (a_i\mathbf{i}_i) \otimes \left(b_j\mathbf{i}_j\right) = a_ib_j\mathbf{i}_i \otimes \mathbf{i}_j \tag{1.37}$$

For example, if $\mathbf{i}_i = \mathbf{i}_1 = [1 \quad 0 \quad 0]^{\mathrm{T}}$, $\mathbf{i}_j = \mathbf{i}_2 = [0 \quad 1 \quad 0]^{\mathrm{T}}$, and $\mathbf{a}$ and $\mathbf{b}$ are arbitrary three-dimensional vectors, one can show that the dyadic product of the preceding equation can be written in the following matrix form:

$$\mathbf{a} \otimes \mathbf{b} = a_ib_j\mathbf{i}_i \otimes \mathbf{i}_j = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix} \tag{1.38}$$

The dyadic products of the base vectors $\mathbf{i}_i \otimes \mathbf{i}_j$ are called the *unit dyads*. Using this notation, the dyadic product can be generalized to the products of three or more vectors. For example, the *triadic product* of the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ can be written as $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} = (a_i\mathbf{i}_i) \otimes \left(b_j\mathbf{i}_j\right) \otimes (c_k\mathbf{i}_k) = a_ib_jc_k\mathbf{i}_i \otimes \mathbf{i}_j \otimes \mathbf{i}_k$. In this book, the familiar summation sign $\sum$ will be used for the most part, instead of the summation convention.

1.4 CARTESIAN TENSORS

It is clear from the preceding section that a dyadic product is a linear combination of unit dyads. The *second-order Cartesian tensor* is defined as a linear combination of dyadic products. A second-order Cartesian tensor $\mathbf{A}$ takes the following form:

$$\mathbf{A} = \sum_{i,j=1}^{3} a_{ij}\mathbf{i}_i \otimes \mathbf{i}_j \tag{1.39}$$

where a_{ij} are called the components of $\mathbf{A}$. Using the analysis presented in the preceding section, one can show that the second-order tensor can be written in the matrix form of Equation 38. Nonetheless, for a given second-order tensor $\mathbf{A}$, one cannot in general find two vectors $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{A} = \mathbf{a} \otimes \mathbf{b}$. Using Equation 39, one can show that the element a_{ij} can be defined as

$$a_{ij} = \mathbf{i}_i \cdot \mathbf{A}\mathbf{i}_j \tag{1.40}$$

The proof of this equation can be obtained using Equation 30 (Spencer, 1997). To this end, we write

$$\begin{aligned}\mathbf{i}_i \cdot \mathbf{A}\mathbf{i}_j &= \mathbf{i}_i \cdot \left(\sum_{k,l=1}^{3} a_{kl}\mathbf{i}_k \otimes \mathbf{i}_l\right)\mathbf{i}_j = \mathbf{i}_i \cdot \left(\sum_{k,l=1}^{3} a_{kl}\left(\mathbf{i}_l \cdot \mathbf{i}_j\right)\mathbf{i}_k\right) \\ &= \mathbf{i}_i \cdot \left(\sum_{k,l=1}^{3} a_{kl}\delta_{lj}\mathbf{i}_k\right) = \left(\sum_{k,l=1}^{3} a_{kl}\delta_{lj}\delta_{ik}\right) \\ &= a_{ij} \end{aligned} \tag{1.41}$$

The *unit* or *identity tensor* can be written in terms of the base vectors as

$$\mathbf{I} = \sum_{i=1}^{3} \mathbf{i}_i \otimes \mathbf{i}_i \tag{1.42}$$

Using the definition of the second-order tensor as a linear combination of dyadic products, one can show, as previously mentioned, that the components of any second-order tensor can be arranged in the form of a 3×3 matrix. Using this matrix arrangement of the second-order tensor $\mathbf{A}$, another simple proof of Equation 40 can be provided. To this end, we note that the product $\mathbf{A}\mathbf{i}_j$ defines column j of the matrix $\mathbf{A}$ denoted as $\mathbf{A}_j$. The dot product $\mathbf{i}_i \cdot \mathbf{A}_j$ defines element i of the vector $\mathbf{A}_j$, which is the same as a_{ij}.

If the components of $\mathbf{A}$ are defined using a set of base vectors $\bar{\mathbf{i}}_i$ and $\bar{\mathbf{i}}_j$ defined in another coordinate system and are denoted as $\bar{a}_{ij}$, one has

$$\mathbf{A} = \sum_{i,j=1}^{3} a_{ij}\mathbf{i}_i \otimes \mathbf{i}_j = \sum_{i,j=1}^{3} \bar{a}_{ij}\bar{\mathbf{i}}_i \otimes \bar{\mathbf{i}}_j \tag{1.43}$$

Let $\mathbf{R} = \left(R_{ij}\right)$ be the orthogonal matrix of transformation between the two coordinate systems in which the two sets of base vectors are defined, such that

$$\mathbf{i}_i = \mathbf{R}\bar{\mathbf{i}}_i, \quad \bar{\mathbf{i}}_i = \mathbf{R}^{\mathrm{T}}\mathbf{i}_i \tag{1.44}$$

It follows that

$$\bar{\mathbf{i}}_i \otimes \bar{\mathbf{i}}_j = \bar{\mathbf{i}}_i\bar{\mathbf{i}}_j^{\mathrm{T}} = \mathbf{R}^{\mathrm{T}}\mathbf{i}_i\mathbf{i}_j^{\mathrm{T}}\mathbf{R} = \mathbf{R}^{\mathrm{T}}\left(\mathbf{i}_i \otimes \mathbf{i}_j\right)\mathbf{R} \tag{1.45}$$

which is the result obtained earlier in Section 2. Substituting the identity of Equation 45 into Equation 43, one obtains

$$\mathbf{A} = \sum_{i,j=1}^{3} a_{ij}\mathbf{i}_i \otimes \mathbf{i}_j = \mathbf{R}^{\mathrm{T}}\left(\sum_{i,j=1}^{3} \bar{a}_{ij}\mathbf{i}_i \otimes \mathbf{i}_j\right)\mathbf{R} \tag{1.46}$$

Let $\mathbf{A} = (a_{ij})$ and $\bar{\mathbf{A}} = (\bar{a}_{ij})$. It then follows from the preceding equation that

$$\mathbf{A} = (a_{ij}) = \mathbf{R}^{\mathrm{T}}\bar{\mathbf{A}}\mathbf{R} \tag{1.47}$$

The inverse relationship is given by

$$\bar{\mathbf{A}} = \mathbf{R}\mathbf{A}\mathbf{R}^{\mathrm{T}} \tag{1.48}$$

Using Equation 47, one can show that the elements a_{ij} can be written in terms of the elements $\bar{a}_{ij}$ as follows:

$$a_{pq} = \sum_{i,j=1}^{3} R_{ip}R_{jq}\bar{a}_{ij} \tag{1.49}$$

Using matrix notation, it can also be shown that $a_{pq} = \mathbf{R}_p^{\mathrm{T}}\bar{\mathbf{A}}\mathbf{R}_q$, where $\mathbf{R}_k$ is the kth column of the tensor $\mathbf{R}$. That is, $\mathbf{R} = [\,\mathbf{R}_1 \quad \mathbf{R}_2 \quad \mathbf{R}_3\,]$. Equation 47, or equivalently Equation 48, governs the transformation of the second-order tensors. That is, any second-order tensor must obey this transformation rule. In continuum mechanics, the elements of tensors represent physical quantities such as moments of inertia, strains, and stresses. These elements can be defined in any coordinate system. The coordinate systems used depend on the formulation used to obtain the equilibrium equations. It is, therefore, important that the reader understands the rule of the coordinate transformation of tensors and recognizes that such a transformation leads to the definition of the same physical quantities in different frames of reference. We must also distinguish between the transformation of vectors and the change of parameters. The latter does not change the coordinate system in which the vectors are defined. This important difference will be discussed in more detail before concluding this chapter.

A tensor that has the same components in any coordinate system is called an *isotropic tensor*. An example of isotropic tensors is the unit tensor. It can be shown that second-order isotropic tensors take only one form and can be written as $\alpha\mathbf{I}$ where α is a scalar and $\mathbf{I}$ is the unit or the identity tensor. Second-order isotropic tensors are sometimes called *spherical tensors*.

Double Product or Double Contraction If $\mathbf{A}$ is a second-order tensor, the *contraction* of this tensor to a scalar is defined as $\sum_{i=1}^{3} a_{ii} = a_{11} + a_{22} + a_{33} = \mathrm{tr}(\mathbf{A})$, where tr denotes the trace of the matrix (sum of the diagonal elements) (Aris 1962). It can be shown that the trace of a second-order tensor is invariant under orthogonal coordinate transformations. To this end, one can write, using Equation 49, $a_{qq} = \sum_{i,j=1}^{3} R_{iq}R_{jq}\bar{a}_{ij}$, which, using the orthogonality of the columns of the transformation $\mathbf{R} = (R_{ij})$, leads to $\sum_{q=1}^{3} a_{qq} = \sum_{q=1}^{3}\left(\sum_{i,j=1}^{3} R_{iq}R_{jq}\bar{a}_{ij}\right) =$

$\sum_{q,j=1}^{3}\delta_{qj}\bar{a}_{qj} = \sum_{q=1}^{3}\bar{a}_{qq}$. That is, the trace is indeed invariant under coordinate transformation. In addition to the trace, the determinant of $\mathbf{A}$ is the same as the determinant of $\bar{\mathbf{A}}$, that is, $|\mathbf{A}| = |\bar{\mathbf{A}}|$. This important result can also be obtained in the case of second-order tensors using the facts that the determinant of an orthogonal matrix is equal to ±1 and the determinant of the product of matrices is equal to the product of the determinants of these matrices.

If $\mathbf{A}$ and $\mathbf{B}$ are second-order tensors, the *double product* or *double contraction* is defined as

$$\mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) \tag{1.50}$$

Using the properties of the trace, one can show that

$$\mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) = \mathrm{tr}(\mathbf{B}\mathbf{A}^{\mathrm{T}}) = \mathrm{tr}(\mathbf{B}^{\mathrm{T}}\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{B}^{\mathrm{T}}) = \sum_{i,j=1}^{3} a_{ij}b_{ij} \tag{1.51}$$

where a_{ij} and b_{ij} are, respectively, the elements of the tensors $\mathbf{A}$ and $\mathbf{B}$. If $\mathbf{a}$, $\mathbf{b}$, $\mathbf{u}$, and $\mathbf{v}$ are arbitrary vectors and $\mathbf{A}$ is a second-order tensor, one can show that the double contraction has the following properties:

$$\left.\begin{array}{l} \mathrm{tr}(\mathbf{A}) = \mathbf{I} : \mathbf{A} \\ \mathbf{A} : (\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \cdot (\mathbf{A}\mathbf{v}) \\ (\mathbf{a} \otimes \mathbf{b}) : (\mathbf{u} \otimes \mathbf{v}) = (\mathbf{a} \cdot \mathbf{u})(\mathbf{b} \cdot \mathbf{v}) \end{array}\right\} \tag{1.52}$$

It can also be shown that if $\mathbf{A}$ is a symmetric tensor and $\mathbf{B}$ is a skew symmetric tensor, then

$$\mathbf{A} : \mathbf{B} = 0 \tag{1.53}$$

It follows that if $\mathbf{A}$ is a symmetric tensor and $\mathbf{B}$ is an arbitrary tensor, the definition of the double product can be used to show that $\mathbf{A} : \mathbf{B} = \mathbf{A} : \mathbf{B}^{\mathrm{T}} = \mathbf{A} : (\mathbf{B}+\mathbf{B}^{\mathrm{T}})/2$.

If $\mathbf{A}$ and $\mathbf{B}$ are two symmetric tensors, one can show that

$$\mathbf{A} : \mathbf{B} = a_{11}b_{11} + a_{22}b_{22} + a_{33}b_{33} + 2(a_{12}b_{12} + a_{13}b_{13} + a_{23}b_{23}) \tag{1.54}$$

The preceding two equations will be used in this book in the formulation of the elastic forces of continuous bodies. These forces are expressed in terms of the strain and stress tensors. As will be shown in Chapters 2 and 3, the strain and stress tensors are symmetric and are given, respectively, in the following form:

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix}, \quad \boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \tag{1.55}$$

Using Equation 54, one can write the double contraction of the strain and stress tensors as

$$\boldsymbol{\varepsilon} : \boldsymbol{\sigma} = \varepsilon_{11}\sigma_{11} + \varepsilon_{22}\sigma_{22} + \varepsilon_{33}\sigma_{33} + 2(\varepsilon_{12}\sigma_{12} + \varepsilon_{13}\sigma_{13} + \varepsilon_{23}\sigma_{23}) \tag{1.56}$$

Because a second-order symmetric tensor has six independent elements, vector notations, instead of tensor notations, can also be used to define the strain and stress components of the preceding two equations. In this case, six-dimensional strain and stress vectors can be introduced as follows:

$$\left.\begin{aligned} \boldsymbol{\varepsilon}_v &= [\,\varepsilon_{11} \quad \varepsilon_{22} \quad \varepsilon_{33} \quad \varepsilon_{12} \quad \varepsilon_{13} \quad \varepsilon_{23}\,]^{\mathrm{T}} \\ \boldsymbol{\sigma}_v &= [\,\sigma_{11} \quad \sigma_{22} \quad \sigma_{33} \quad \sigma_{12} \quad \sigma_{13} \quad \sigma_{23}\,]^{\mathrm{T}} \end{aligned}\right\} \tag{1.57}$$

where subscript v is used to denote a vector. The dot product of the strain and stress vectors is given by

$$\boldsymbol{\varepsilon} \cdot \boldsymbol{\sigma} = \boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\sigma} = \varepsilon_{11}\sigma_{11} + \varepsilon_{22}\sigma_{22} + \varepsilon_{33}\sigma_{33} + \varepsilon_{12}\sigma_{12} + \varepsilon_{13}\sigma_{13} + \varepsilon_{23}\sigma_{23} \tag{1.58}$$

Note the difference between the results of the double contraction and the dot product of Equations 56 and 58, respectively. There is a factor of 2 multiplied by the term that includes the off-diagonal elements in the double contraction of Equation 56. Equation 56 arises naturally when the elastic forces are formulated, as will be shown in Chapter 3. Therefore, it is important to distinguish between the double contraction and the dot product despite the fact that both products lead to scalar quantities.

Invariants of the Second-Order Tensor Under an orthogonal transformation that represents rotation of the axes of the coordinate systems, the components of the vectors and second-order tensors change. Nonetheless, certain vector and tensor quantities do not change and remain invariant under such an orthogonal transformation. For example, the norm of a vector and the dot product of two three-dimensional vectors remain invariant under a rigid-body rotation.

For a second-order tensor $\mathbf{A}$, one has the following three invariants that do not change under an orthogonal coordinate transformation:

$$\left.\begin{aligned} I_1 &= \mathrm{tr}(\mathbf{A}) \\ I_2 &= \frac{1}{2}\left\{(\mathrm{tr}(\mathbf{A}))^2 - \mathrm{tr}(\mathbf{A}^2)\right\} \\ I_3 &= \det(\mathbf{A}) = |\mathbf{A}| \end{aligned}\right\} \tag{1.59}$$

These three invariants can also be written in terms of the *eigenvalues* of the tensor $\mathbf{A}$. For a given tensor or a matrix $\mathbf{A}$, the eigenvalue problem is defined as

$$\mathbf{A}\mathbf{y} = \lambda\mathbf{y} \tag{1.60}$$

where λ is called the eigenvalue and $\mathbf{y}$ is the *eigenvector* of $\mathbf{A}$. Equation 60 shows that the direction of the vector $\mathbf{y}$ is not affected by multiplication with the tensor $\mathbf{A}$. That is, $\mathbf{Ay}$ can change the length of $\mathbf{y}$, but such a multiplication does not change the direction of $\mathbf{y}$. For this reason, $\mathbf{y}$ is called a principal direction of the tensor $\mathbf{A}$. The preceding eigenvalue equation can be written as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{y} = \mathbf{0} \tag{1.61}$$

For this equation to have a nontrivial solution, the determinant of the coefficient matrix must be equal to zero, that is,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{1.62}$$

This equation is called the *characteristic equation*, and in the case of a second-order tensor it has three roots λ_1, λ_2, and λ_3. Associated with these three roots, there are three corresponding eigenvectors $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$ that can be determined to within an arbitrary constant using Equation 61. That is, for a root $\lambda_i, i = 1, 2, 3$, one can solve the system of homogeneous equations $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{y}_i = \mathbf{0}$ for the eigenvector $\mathbf{y}_i$ to within an arbitrary constant, as demonstrated by the following example.

EXAMPLE 1.5

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

The characteristic equation of this matrix can be obtained using Equation 62 as

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (1 - \lambda)(3 - \lambda)(2 - \lambda) = 0$$

The roots of this characteristic equation define the following three eigenvalues of the matrix $\mathbf{A}$:

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 3$$

Associated with these three eigenvalues, there are three eigenvectors, which can be determined using Equation 61 as

$$(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{y}_i = \mathbf{0}, \quad i = 1, 2, 3$$

or

$$\begin{bmatrix} 1 - \lambda_i & -1 & 2 \\ 0 & 3 - \lambda_i & 1 \\ 0 & 0 & 2 - \lambda_i \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{bmatrix} = \mathbf{0}, \quad i = 1, 2, 3$$

This equation can be used to solve for the eigenvectors associated with the three eigenvalues λ_1, λ_2, and λ_3. For $\lambda_1 = 1$, the preceding equation yields the following system of algebraic equations:

$$\begin{bmatrix} 0 & -1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \end{bmatrix} = \mathbf{0}$$

This system of algebraic equations defines the first eigenvector to within an arbitrary constant as

$$\mathbf{y}_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

For $\lambda_2 = 2$, one has

$$\mathbf{y}_2 = \begin{bmatrix} y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ -1 \end{bmatrix}$$

The eigenvector associated with $\lambda_3 = 3$ can also be determined as

$$\mathbf{y}_3 = \begin{bmatrix} y_{31} \\ y_{32} \\ y_{33} \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

In the special case of a *symmetric tensor*, one can show that the eigenvalues are real and the eigenvectors are orthogonal. Because the eigenvectors can be determined to within an arbitrary constant, the eigenvectors can be normalized as unit vectors. For a symmetric tensor, one can then write

$$\left.\begin{aligned} \mathbf{A}\mathbf{y}_i &= \lambda_i \mathbf{y}_i\,, \quad i = 1, 2, 3 \\ \mathbf{y}_i^{\mathrm{T}} \mathbf{y}_j &= \delta_{ij}\,, \quad i, j = 1, 2, 3 \end{aligned}\right\} \tag{1.63}$$

If $\mathbf{y}_i$, $i = 1, 2, 3$, are selected as orthogonal unit vectors, one can form the orthogonal matrix $\mathbf{\Phi}$ whose columns are the orthonormal eigenvectors, that is,

$$\mathbf{\Phi} = [\,\mathbf{y}_1 \quad \mathbf{y}_2 \quad \mathbf{y}_3\,] \tag{1.64}$$

It follows that

$$\mathbf{A}\mathbf{\Phi} = \mathbf{\Phi}\boldsymbol{\lambda} \tag{1.65}$$

where

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \tag{1.66}$$

Using the orthogonality property of $\boldsymbol{\Phi}$, one has

$$\mathbf{A} = \boldsymbol{\Phi}\boldsymbol{\lambda}\boldsymbol{\Phi}^{\mathrm{T}} = \sum_{i=1}^{3} \lambda_i \mathbf{y}_i \otimes \mathbf{y}_i \tag{1.67}$$

This equation, which defines the *spectral decomposition* of $\mathbf{A}$, shows that the orthogonal transformation $\boldsymbol{\Phi}$ can be used to transform the tensor $\mathbf{A}$ to a diagonal matrix as

$$\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{A}\boldsymbol{\Phi} = \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \tag{1.68}$$

That is, the matrices $\mathbf{A}$ and $\boldsymbol{\lambda}$ have the same determinant and the same trace. This important result is often used in continuum mechanics to study the invariant properties of different tensors.

Let $\mathbf{R}$ be an orthogonal transformation matrix. Using the transformation $\mathbf{y} = \mathbf{R}\mathbf{z}$ in Equation 61 and premultiplying by $\mathbf{R}^{\mathrm{T}}$, one obtains

$$\left(\mathbf{R}^{\mathrm{T}}\mathbf{A}\mathbf{R} - \lambda\mathbf{I}\right)\mathbf{z} = \mathbf{0} \tag{1.69}$$

This equation shows that the eigenvalues of a tensor or a matrix do not change under an orthogonal coordinate transformation. Furthermore, as previously discussed, the determinant and trace of the tensor or the matrix do not change under such a coordinate transformation. One then concludes that the invariants of a symmetric second-order tensor can be expressed in terms of its eigenvalues as follows:

$$\left.\begin{aligned} I_1 &= \mathrm{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \lambda_3 \\ I_2 &= \frac{1}{2}\left\{(\mathrm{tr}(\mathbf{A}))^2 - \mathrm{tr}(\mathbf{A}^2)\right\} = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \\ I_3 &= \det(\mathbf{A}) = \lambda_1\lambda_2\lambda_3 \end{aligned}\right\} \tag{1.70}$$

Some of the material constitutive equations used in continuum mechanics are formulated in terms of the invariants of the strain tensor. Therefore, Equation 70 will be used in later chapters of this book.

For a general second-order tensor $\mathbf{A}$ (symmetric or nonsymmetric), the invariants are $I_1 = \mathrm{tr}(\mathbf{A}), I_2 = \frac{1}{2}\left\{(\mathrm{tr}(\mathbf{A}))^2 - \mathrm{tr}(\mathbf{A}^2)\right\}$, and $I_3 = \det(\mathbf{A})$, as previously presented. One can show that the characteristic equation of a second-order tensor can

be written in terms of these invariants as $\lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0$. Furthermore, by repeatedly multiplying Equation 60 n times by $\mathbf{A}$, one obtains $\mathbf{A}^n\mathbf{y} = \lambda^n\mathbf{y}$. Using this identity after multiplying the characteristic equation $\lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0$ by $\mathbf{y}$, one obtains $\mathbf{A}^3 - I_1\mathbf{A}^2 + I_2\mathbf{A} - I_3\mathbf{I} = \mathbf{0}$, which is the mathematical statement of the *Cayley–Hamilton theorem*, which states that a second-order tensor satisfies its characteristic equation. The simple proof provided here for the Cayley–Hamilton theorem is based on the assumption that the eigenvectors are linearly independent. A more general proof can be found in the literature.

For a second-order skew-symmetric tensor $\mathbf{W}$, one can show that the invariants are given by $I_1 = I_3 = 0$ and $I_2 = w_{12}^2 + w_{13}^2 + w_{23}^2$, where w_{ij} is the ijth element of the tensor $\mathbf{W}$. Using these results, the characteristic equation of a second-order tensor $\mathbf{W}$ can be written as $\lambda^3 + I_2\lambda = 0$. This equation shows that $\mathbf{W}$ has only one real eigenvalue, $\lambda = 0$, whereas the other two eigenvalues are imaginary.

Higher-Order Tensors In continuum mechanics, the stress and strain tensors are related using the constitutive equations that define the material behavior. This relationship can be expressed in terms of a fourth-order tensor whose components are material coefficients. In general, a tensor $\mathbf{A}$ of order n is defined by 3^n elements, which can be written as $a_{ijk\ldots n}$, provided that these elements as the result of a coordinate transformation take the form

$$a_{pq\ldots s} = \sum_{i,j,\ldots,l=1}^{n} R_{ip}R_{jq}\ldots R_{ls}\bar{a}_{ij\ldots l} \tag{1.71}$$

where $\mathbf{R} = (R_{ij})$ is the matrix of coordinate transformation. A lower-order tensor can be obtained as a special case of Equation 71 by reducing the number of indices. A zero-order tensor is represented by a scalar, a first-order tensor is represented by a vector, and a second-order tensor can be represented by a matrix. A tensor of order n is said to be *symmetric* with respect to two indices if the interchange of these two indices does not change the value of the elements of the tensor. The tensor is said to be *antisymmetric* or *skew symmetric* with respect to two indices if the interchange of these two indices changes only the sign of the elements of the tensor.

As in the case of the second-order tensors, higher-order tensors can be defined using outer products. For example, a third-order tensor $\mathbf{T}$ can be defined as the outer product of three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ as follows:

$$\mathbf{T} = (\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}) = \sum_{i,j,k=1}^{3} t_{ijk}\mathbf{i}_i \otimes \mathbf{i}_j \otimes \mathbf{i}_k \tag{1.72}$$

where $\mathbf{i}_l$, $l = i, j, k$ is a base vector. An element of the tensor $\mathbf{T}$ takes the form $u_i v_j w_k$. Roughly speaking, in the case of three-dimensional vectors, one may consider the third-order tensor a linear combination of a new set of unit dyads that consist of 27 elements (three layers, each of which has nine elements). The elements of layer or matrix l, $l = 1, 2, 3$ are given by $w_l(\mathbf{u} \otimes \mathbf{v}) = \mathbf{T}\mathbf{i}_l$. Using this definition of the product