Continuum Mechanics of Anisotropic Materials

Stephen C. Cowin

Continuum Mechanics of Anisotropic Materials

Springer

Stephen C. Cowin
Department of Biomedical Engineering
The City College
New York, NY, USA

ISBN 978-1-4614-5024-5 ISBN 978-1-4614-5025-2 (eBook)
DOI 10.1007/978-1-4614-5025-2
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012951630

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my children and grandchildren

Preface

The objective of this book is to describe the methods of formulating continuum models for material behavior. The text is structured in a stepwise hierarchical fashion from basic to more advanced topics. The text begins with an introductory chapter that summarizes some aspects of creating mechanical models of material behavior and also describes different models for material object behavior; the particle, rigid object, lumped parameter, and continuum models. Chapters 2 through 6 contain a development of linear anisotropic continuum mechanics models: a development of basic continuum kinematics is presented in Chap. 2, the continuum formulations of conservation laws is recounted in Chap. 3, the process of modeling material symmetry is explained in Chap. 4, the steps in the formulation of constitutive equations are enumerated in Chap. 5; and four linear continuum theories: flow through rigid porous media, elasticity, viscous fluid theory, and viscoelasticity, are described in Chap. 6. These four continuum models are combined in different ways and applied at the microstructural level in the remainder of the book. Chapter 7 concerns the modeling of material microstructure. Chapters 8 and 9 present developments of the theories of quasi-static and dynamic poroelasticity, respectively, while Chap. 10 presents a mixture theory approach to poroelasticity. The kinematics and mechanics of large elastic deformations are described in Chap. 11. Appendix A on matrices and tensors also contains short reviews of other mathematical topics that occur in the development of the text material. The material in the Appendix has been added to aid the students in remembering what they once knew. It is now presented at the start of the course.

The presentations in the text differ from the customary presentations of these topics in many aspects, two of which are worth pointing out. First, all continuum models are developed for the anisotropic cases rather than the isotropic cases because most tissues are anisotropic in their material properties. Second, a slightly unconventional tensor-matrix notation is employed in this presentation. Its objective is to represent fourth rank tensors as matrices that are composed of tensor components, something that the classical Voigt matrix notation for the anisotropic elasticity tensor does not achieve. In the notation employed here second and fourth rank tensors in three dimensions are represented as vectors and second rank tensors,

respectively, in six dimensions. Transformations in the six-dimensional space, corresponding to three-dimensional transformations, are six-by-six matrix multiplications that are easily entered and quickly computed with symbolic algebra software (Maple, Mathematica, MacSyma, and MatLab). In particular the three-dimensional fourth rank elasticity tensor is represented as a second rank tensor in a space of six dimensions. This notation is described in the Appendix on matrices and tensors.

The material in this text is covered in a Continuum Mechanics course by the author. The course regularly draws students from Chemical, Civil and Mechanical Engineering as well as Biomedical Engineering. The material in the Continuum Mechanics course, in the order covered, is Appendix A, then Chaps. 1 through 7. I cover the material on poroelasticity in a separate course I teach jointly with Luis Cardoso. I would very much appreciate readers communicating to me suggested revisions to this book. In particular any corrections, comments, suggestions of material to be included (or excluded) and suggested problems w/solutions for use as either examples or problems at the end of sections would be appreciated. Please email these materials to sccowin@gmail.com or cowin@ccny.cuny.edu. I will maintain a record of corrections, suggested additions, and suggested (HW) problems w/solutions.

A problem solutions manual is available from the author for instructors using this book in a course. For an instructor to obtain an e-copy, please email a request to sccowin@gmail.com or cowin@ccny.cuny.edu and enclose the name of the instructor, the name of the instructor's institution and course in which the book will be employed.

The contributions from the students who took the courses in which the content of the book was contained in handouts presented have been most helpful. Monte Mehrabadi has made substantial indirect contributions to the text through his 40-year collaboration with the author, as has Luis Cardoso in the last 5 years.

New York, USA Stephen C. Cowin

Contents

Chapter 1
Mechanical Modeling of Material Behavior

1.1 Introduction

The typical types of models that have been employed in the analysis of the mechanical behavior of material objects are described in this chapter. Specifically, descriptions of the various model types employed in mechanics, namely the particle model, the rigid object model, the deformable continuum model, and the lumped parameter model are described. The modeler should view these model types as tools and the task of the modeler is to select the proper tool for the problem at hand.

The content of this chapter is not material that can be learned by rote memorization. It is material that must be thought about and practiced in order to acquire a modeler's skill. In the next section conservation principles, control volumes, and free object diagrams are discussed. In the section that follows, the first problem in modeling, the concept of time, is considered, then that of space. In the next section, the relationship between models and the real physical world is discussed. Four sections that describe particle models, rigid object models, continuum models, and lumped parameter models, respectively, follow a section on the types of models used in mechanics. The final section concerns two philosophical questions related to mechanical modeling, reductionism, and determinism.

1.2 Conservation Principles, Control Volumes, and Free Object Diagrams

There is an aspect of the application of conservation principles of mechanics (those of mass, momentum, angular momentum, energy, etc.) that is an artisan-like skill that requires some experience on the part of the modeler. Most engineering students acquire this skill when they learn to construct "free object diagrams" to apply Newton's laws to solid objects in a course on statics or to draw "control volumes" to apply the conservation principles of mechanics to fluids. These diagrams or

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_1,

volumes are selected to satisfy the criteria of the modeler, which are generally to define the problem by representing the unknown quantities and the known quantities in a way that the conservation principles can be applied to obtain an equation(s) representing the unknown quantities in terms of the known quantities. The construction of free object diagrams or control volumes is an artisan-like skill because there is not a unique way to construct them; the modeler must have insight into the problem. Many such physically correct diagrams can be drawn but it is likely that only a few will yield the modeler the relationship between quantities that he or she sought. This creative aspect of the application of conservation principles of mechanics is prominent because of the great diversity of the situations to which they are applied. These applications range in size from the Nano scale to the Macro scale, from a portion of a protein to a molecule to bridges, airplane structures, and the structure of the universe.

Both free object diagrams and control volumes are drawings made to simplify the application of conservation principles to a particular physical situation. A conservation principle can often be written in the form of an accounting statement:

[The time rate of change of a quantity in a system]
$=$ [the amount of the quantity coming into the system per unit time]
$-$ [the amount of the quantity leaving the system per unit time]
$+$ [the amount of the quantity produced within the system per unit time]
$-$ [the amount of the quantity consumed within the system per unit time].

Thus the application of the conservation of mass (or momentum) of a fluid employs a drawing to balance the net flow or change in the quantity, much like a financial account for an organization is balanced.

Note that the word "body" used in modeling nonbiological applications of Newton's laws, for example "free body diagrams" and "rigid body motion," is replaced in this work by the word "object"; thus reference is made here to "free object diagrams" and "rigid object motions." The reason for this shift in terminology is to avoid the use of the same word for two different meanings in the same phrase, like a "free body diagram" of a body or a "rigid body motion" of a body.

1.3 Models and the Real Physical World

Models have been found to be very effective tools for the analysis of physical problems. The basic elements of these models are Euclidean or classical geometry and the concept of time. The concept of time is intuitive while our mathematical model of time, the real line, that is to say the line representing all real numbers, is abstract. The connection between the mathematical abstraction and our intuitive perception of time is a philosophical matter that should be accepted by the reader;

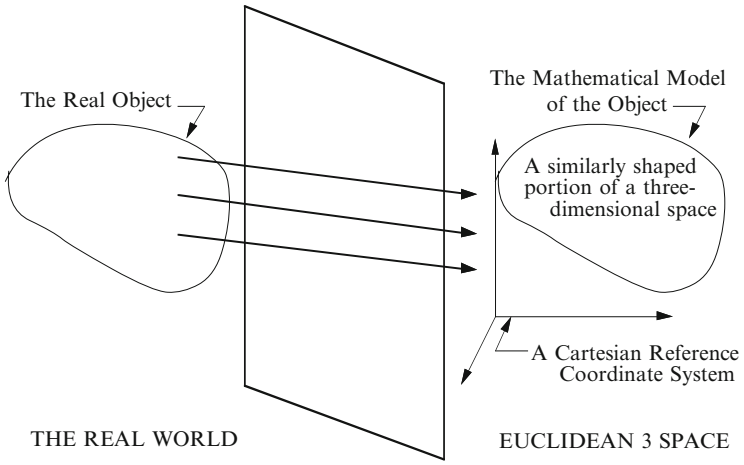


Fig. 1.1 The Euclidean space model of a real object

this relationship is rigorously discussed in a classical book by the German-American mathematician Hermann Weyl (1883–1955) entitled "The Continuum." Once one has accepted the real line as a geometric model for time, the next step of accepting the real line as a geometric model for a one-dimensional structure (such as a string or a fiber) is not difficult. Real lines are used in forming one-, two-, and three-dimensional Cartesian reference coordinate systems for one-, two-, and three-dimensional Euclidean spaces. The real lines used as the reference Cartesian coordinate systems in two and three dimensions are mutually orthogonal.

In the initial stages of model construction for a material object of any size or shape, it is necessary to project the object into Euclidean space (Fig. 1.1). The portion of the three-dimensional Euclidean space that the object occupies is often an exact replica of the space in the real world occupied by the real object. However, the structure of the real object is not carried over into the model unless the modeler makes provision for it. For example, if the object is fibrous, the fibers are not represented in the model unless the modeler explicitly provides for their representation. The advantage of the model is that all the points in the real object now have Cartesian addresses. The triplet of Cartesian coordinates (x_1, x_2, x_3) locates a point in the Euclidean space occupied by a particular point in the real object. This permits the measurement of distance between points in the object, and if the object moves or deforms, the resulting motion can be quantitatively documented. The volume of the object and its centroid may be determined by use of the integral calculus. The greatest advantage, however, is the ability to define functions of physical interest in terms of the reference coordinate system. Thus, for example, the temperature distribution throughout the object can be specified by a function $\theta = \theta(x_1, x_2, x_3)$ defined for all of the Cartesian coordinates (x_1, x_2, x_3) located within the object. If the function is smooth, then a derivative of the function can be taken and the temperature gradient is determined for all of the Cartesian coordinates (x_1, x_2, x_3) located within the object. The point of these remarks is that

the consequence of making a Euclidean space model of an object is that it permits the powerful computational machinery of classical geometry and the integral and differential calculus to be used to calculate quantities of physical or biological interest.

Since the representation of physical phenomena must be independent of the observer, it is necessary to express physical quantities in ways that are independent of coordinate systems. This is because different observers may select different coordinate systems. It therefore becomes a requirement that physical quantities be invariant of the coordinate system selected to express them. On the other hand, in order to work with these physical quantities and evaluate their magnitudes, it is necessary to refer physical quantities to coordinate systems as illustrated in Fig. 1.1. The resolution of this conflict is to express physical quantities as tensors; vectors are tensors of order one and scalars are tensors of order zero. Thus it is no surprise that in classical mechanics the essential concepts of force, velocity, and acceleration are all vectors; hence the mathematical language of classical mechanics is that of vectors. In the mechanics of rigid objects the concepts of position, velocity, and acceleration are all vectors and moments of inertia are second-order tensors. In the mechanics of deformable media the essential concepts of stress, strain, rate of deformation, etc., are all second order tensors; thus, by analogy, one can expect to deal quite frequently with second-order tensors in this branch of mechanics. The reason for this widespread use of tensors is that they satisfy the requirement of invariance of a particular coordinate system on one hand, and yet permit the use of coordinate systems on the other hand. Thus a vector $\mathbf{u}$ represents a quantity that is independent of coordinate system, i.e., the displacement of a point on an object, yet it can be expressed relative to the three-dimensional Cartesian coordinate system with base vectors $\mathbf{e}_\alpha$, $\alpha = \mathrm{I}, \mathrm{II}, \mathrm{III}$, as $\mathbf{u} = \mathrm{u_I}\mathbf{e_I} + \mathrm{u_{II}}\mathbf{e_{II}} + \mathrm{u_{III}}\mathbf{e_{III}}$ and also expressed relative to another three-dimensional Cartesian coordinate system with base vectors $\mathbf{e}_i$, $i = 1, 2, 3$, by $\mathbf{u} = \mathrm{u}_1\mathbf{e}_1 + \mathrm{u}_2\mathbf{e}_2 + \mathrm{u}_3\mathbf{e}_3$. The vector $\mathbf{u}$ and the two coordinate systems are illustrated in Fig. 2.2. These two representations of the components of the vector $\mathbf{u}$ are different and both are correct because a rule, based on the relationship between the two vector bases $\mathbf{e}_\alpha$, $\alpha = \mathrm{I}, \mathrm{II}, \mathrm{III}$, and $\mathbf{e}_i$, $i = 1, 2, 3$, can be derived for calculating one set of components in terms of the other. Thus the vector $\mathbf{u}$ has a physical significance independent of any coordinate system, yet it may be expressed in component form relative to any coordinate system. The property of vectors is shared by all tensors. This is the reason that tensors, as well as vectors and scalars, play a leading role in modeling mechanics phenomena.

1.4 The Types of Models Used in Mechanics

It is possible to divide the discipline of mechanics according to the predominant type of motion an object is considered to be undergoing. The three types of motion are translational, rotational, and deformational. In translational motion all the points of the moving object have the same velocity vector at any instant of time.

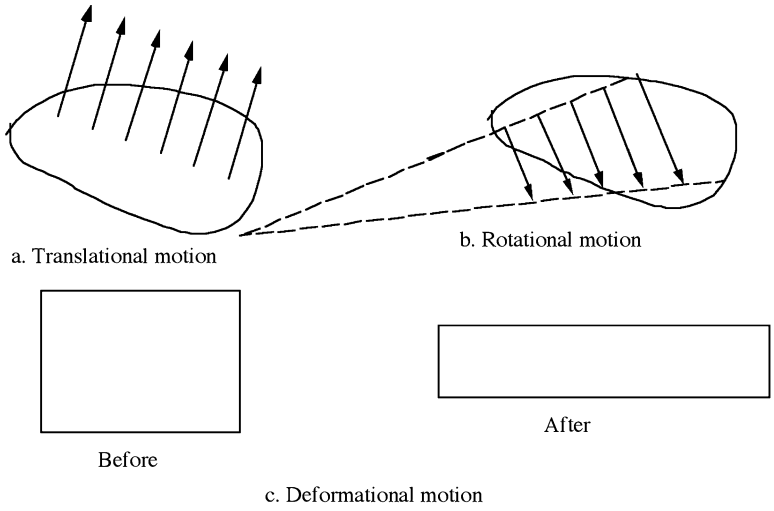


Fig. 1.2 The three types of motion possible for an object

A translational motion is illustrated in Fig. 1.2a. In pure rotational motion the velocities of all the points of the moving object, at any instant, are proportional to the distance of the point from one single fixed axis. This situation is illustrated in Fig. 1.2b in the special case where the fixed axis is perpendicular to the plane of the page. A deformational motion (Fig. 1.2c) is a motion in which some points on the same object move relative to one another. It can be shown that the motion of any object or material system can, at any time, be decomposed into the sum of three motions, a translational motion, a rotational motion, and a deformational motion. If there are many objects (e.g., molecules) and only average properties of the ensemble are sought by statistical methods, the model is said to be one of statistical mechanics.

These motions suggest three of the five types of models of objects used in mechanics: the particle model, the rigid object model, and the deformable continuum model. These models of objects are differentiated from one another on the basis of the type of motion modeled. The particle model only emulates the translational motion of the object; the rigid object model emulates both translational and rotational motion of the object; and the deformable continuum model emulates all three types of motion. The first two models, the particle model and the rigid object model, are described in mechanics books dealing with statics and dynamics of particles and rigid objects. The deformable continuum model is described in books dealing with the mechanics or strength of materials or in books on elasticity and fluid or continuum mechanics.

There is an important fourth category of model used in mechanics that overlaps the first three model types; this category is the lumped parameter model, a very important model type in mechanics. The particle model, rigid object model, deformable continuum model, and the lumped parameter model are discussed in

the next four sections. The fifth categories of mechanics models, statistical models, are not discussed in this volume.

After this chapter, the remainder of the book is an elaboration of the deformable continuum model.

1.5 The Particle Model

The particle model is the simplest model in the hierarchy of models in classical mechanics. This model of an object considers the entire mass of the object as located at the mass center and only the translational motion of the mass center is modeled. Thus the image of the model shown in Euclidean space in Fig. 1.1 shrinks to a mass point located at the mass center, as illustrated in Fig. 1.3. Since the mass center is a point, the particle model is a point model; rotational motions and deformations of the object are neglected. The English natural philosopher Isaac Newton (1642–1727) created the particle model when he took the sun and a planet to be particles and used his universal law of gravitation and his second law to provide an analytical derivation of the three empirical laws of the German astrologer–astronomer Johannes Kepler (1571–1630). In particular, Newton's model showed that the planets moved around the sun in elliptical orbits, a fact previously established by Kepler's observational data. Moments and rotational motions are not considered in the particle model; they are considered in the rigid object model.

The modeling structure described above may be employed as a framework for the statement of Newton's second law. This may be accomplished by letting the vector **p** denote the position of a typical point in the mathematical model of

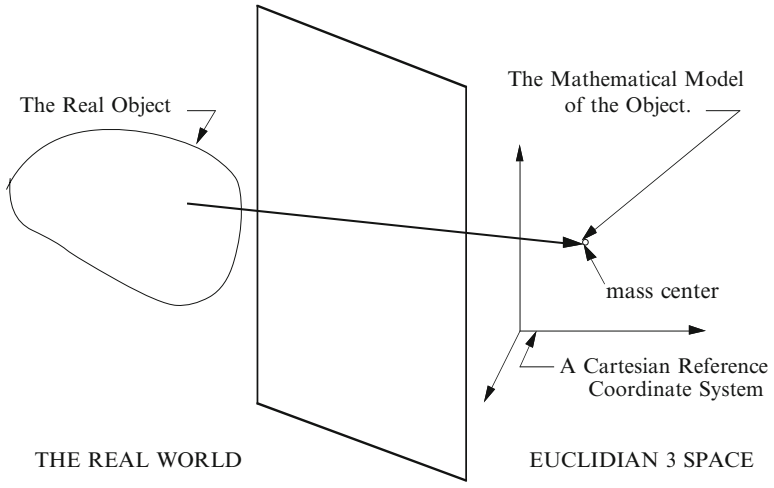


Fig. 1.3 The particle model of a real object

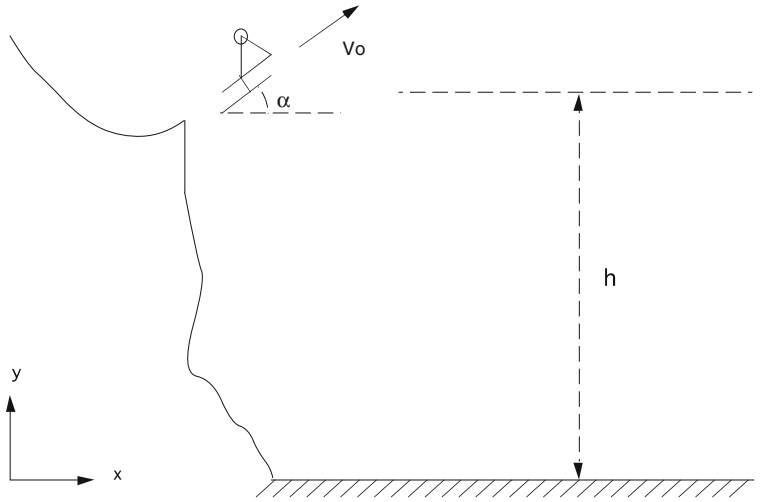


Fig. 1.4 An illustration of the particle model in biomechanics, determining the airborne trajectory of a ski jumper. In the air the ski jumper is acted upon by the attraction of gravity, the drag of the wind and the momentum established in the downhill run before contact with the ground ceased. The ski jumper's trajectory is determined by the solution of Newton's second law with these specified forces

the object, $\mathbf{p} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$, where $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are the Cartesian unit base vectors. The position vector to the mass center of the object is denoted by $\mathbf{p}_{(\mathrm{mc})} = x_{(\mathrm{mc})1}\mathbf{e}_1 + x_{(\mathrm{mc})2}\mathbf{e}_2 + x_{(\mathrm{mc})3}\mathbf{e}_3$. If the object is moving, then the location of the point of the object is changing in the Euclidean space, and the Cartesian coordinates x_1, x_2, x_3 are all continuous, twice differentiable, functions of time, $x_1 = x_1(t)$, $x_2 = x_2(t)$, $x_3 = x_3(t)$, and it is therefore possible to compute the velocity of the mass center of the object, or of any point on the object, as well as its acceleration. The acceleration of the mass center is given by

$$a_{\mathrm{mc}} = (\mathrm{d}^2 x_{(\mathrm{mc})1}/\mathrm{d}t^2)\ \boldsymbol{e}_1 + (\mathrm{d}^2 x_{(\mathrm{mc})2}/\mathrm{d}t^2)\ \boldsymbol{e}_2 + (\mathrm{d}^2 x_{(\mathrm{mc})3}/\mathrm{d}t^2)\ \boldsymbol{e}_3. \tag{1.1}$$

Denoting the total mass of the object by m and the sum of the forces acting on the object by $\mathbf{F}$, a statement of the second law of Newton can then be written in the form

$$\boldsymbol{F} = m\boldsymbol{a}_{\mathrm{mc}}. \tag{1.2}$$

As an illustration of the particle model, consider the question of determining the airborne trajectory of a ski jumper. In the air the ski jumper is acted upon by the attraction of gravity, the drag of the wind, and the momentum established in the downhill run before contact with the ground ceased (Fig. 1.4). The ski jumper's trajectory is determined by the solution of Newton's second law with these specified forces. The trajectory is obtained by an analysis that is completely equivalent to that of an artillery shell or a sub-orbital rocket. While the particle

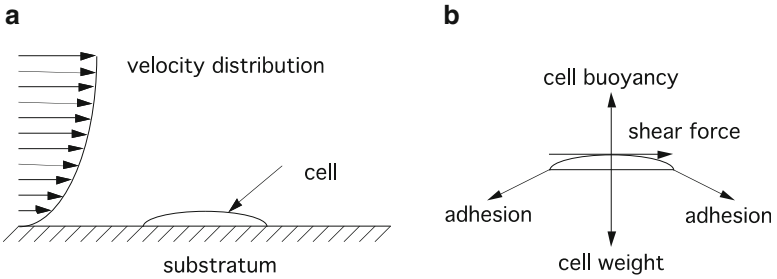


Fig. 1.5 (**a**) A cell adherent to a substratum and subjected to a fluid shearing action. (**b**) An analysis of the forces on a cell of Fig. 1.5 (a), the cell adherent to a substratum and subjected to a fluid shearing action

model is adequate for determining the trajectory of the skier, it is an inadequate model for the transition of the skier's position from the crouch of the downhill run to the erect and forward leaning body posture adopted for flight, and inadequate to deal with the question of impact upon landing.

This traditional method of analysis of macroscopic force systems is also applied at the microscopic level. For example, one method of studying the response of cells to mechanical loading situations is to culture or grow the cells on a surface such as glass and subject the surface to fluid shear stresses as illustrated in Fig. 1.5a. Some of the forces that act on the cell in this flow situation are shown in Fig. 1.5b. The forces that act on a cell include the weight of the cell, W_{cell}, the buoyant force on the cell due to its aqueous environment, W_{buoyant}; the adhesive force of the cell to its substrata, F, the fluid pressure on the cell, the shear force due to fluid flowing over the surface of the cell, forces due to electrical charge or magnetic fields, and self generated forces by the cell. The same forces act on these cells in vivo, but the forces are illustrated in vitro because it is an easier situation to visualize and to draw. The weight of the cell, $W_{\text{cell,}}$ is about 1 piconewton. The cell is subjected to a buoyancy force as a consequence of its immersion in an aqueous environment. The buoyant force on the osteocyte is equal to the weight of the water it displaces, $W_{\text{buoyant}} = 0.9$ piconewton, nine-tenths of a piconewton. The processes of cell adhesion to substrata, as well as the influence of substratum surface properties on cell adhesion, have been studied in recent years by subjecting cells, in vitro, to fluid shear stress. Adhesion of cells to solid substrata is influenced by several substratum surface properties including substratum wettability, surface roughness, and surface charge. The force F of adhesion for a single cell is the surface area A that the cell presents to the flow times the shear stress when the cell is removed from the surface by the fluid shear stress. The shear stress at which cells capable of adhesion detach from a glass substratum is about 400 dyn/cm^2. Assuming an appropriate surface area, the adhesive force on the cell F is about 6,000 piconewtons. Since $W_{\text{cell}} = 1$ piconewton, $F = 6{,}000\ W_{\text{cell}}$. Thus a cell can express adhesive forces that are three to four orders of magnitude larger than the cell weight. For man on the surface of the earth, the largest forces with which we

must cope are generally those due to gravity, that is to say the forces that manifest themselves in the weight of objects. It is quite a different story for the cell because the largest force system they experience is due to adhesion. The adhesive forces on a cell are three to four orders of magnitude larger than the gravitational forces on a cell. More importantly, cells control their adhesive forces; man does not control gravity.

Problem

1.5.1. A ski jumper leaves the ski jump with a velocity v_o in a direction that is an angle α above the horizon (see Fig. 1.4). The final point on the ski jump is an elevation h above the valley floor. If the drag of the wind is neglected, show that the horizontal and vertical velocities, v_x and v_y, respectively, of the skier as he reaches the flat valley floor are given by $v_x = v_o \cos\alpha$ and $v_y = -((v_o \sin\alpha)^2 + 2gh)^{1/2}$. Find the time, t_{touch}, at which the skier touches valley floor as a function of v_o, α, h and g, the acceleration of gravity.

1.6 The Rigid Object Model

The rigid object model differs from the particle model in that the rotational motion as well as the translational motion of the object is considered. Deformations are neglected, hence the adjective "rigid" modifying object. Thus, not only Newton's second law of motion is involved, but also Euler's equations (after their creator, the Swiss mathematician/engineer/physicist Léonard Euler, 1707–1783) for the rotational motion. Euler's equations are special forms of the conservation of angular momentum expressed in a reference coordinate system at the mass center of the rigid object (or at a fixed point of rotation of the object), fixed to the rigid object, and coincident with the principal axes of inertia. If I_{11}, I_{22}, and I_{33} represent the principal moments of inertia (see Appendix section A.8), and M_1, M_2, and M_3 represent the sums of the moments about the three axes, then Euler's equations may be written in the form

$$\begin{aligned} M_1 &= I_{11}(d\omega_1/dt) + \omega_2\omega_3(I_{33} - I_{22}), \\ M_2 &= I_{22}(d\omega_2/dt) + \omega_3\omega_1(I_{11} - I_{33}), \\ M_3 &= I_{33}(d\omega_3/dt) + \omega_1\omega_2(I_{32} - I_{11}), \end{aligned} \tag{1.3}$$

where ω_1, ω_2, and ω_3 are the components of the angular velocity about the respective coordinate axes. In the case when there is only one nonzero component of the angular velocity ω_3 and $\omega_1 = \omega_2 = 0$, then (1.3) reduces to

$$M_3 = I_{33}\alpha_3, \tag{1.4}$$

where $\alpha_3 = d\omega_3/dt$ is the angular acceleration. This is the one-dimensional form of the conservation of angular momentum, or of Euler's equations, a form that usually appears in basic mechanics texts.

If the rigid object restriction is relaxed a bit to allow the moment of inertia I_{33} in (1.4) to vary, then the conservation of angular momentum about an axis may be written in this special case as

$$M_3 = d(I_{33}\omega_3)/dt. \tag{1.5}$$

This is the form of the conservation of angular momentum that is employed to explain why a figure skater spinning at one place on the surface of the ice can increase or decrease his or her angular velocity by extending their arms out from the torso or lowering the arms to the sides of the torso. If the skater is spinning, there is no moment about the axis that is the intersection of the sagittal and frontal or coronal planes, thus $M_3 = 0$ and, from (1.5) above, the product $I_{33}\omega_3$ must be a constant. Since $I_{33}\omega_3 =$ constant, when the skater extends (lowers) the arms, the moment of inertia of the skater increases (decreases) and the angular velocity of the spin must decrease (increase).

For stationary objects, or objects moving with constant velocity, the conservation of linear and angular momentum reduce to the conditions that the sum of the forces and the sum of the moments must be 0. These conditions provide six equations in the case of a three-dimensional problem and three equations in the case of a two-dimensional problem. The application of these equations is the topic of an engineering course on the topic of statics.

Problem

1.6.1. A diver rotates faster when her arms and legs are tucked tightly in so that she is almost like a ball rather than when the limbs are extended in the common diving posture like a straight bar. Consider a diver with a mass of 63 kg, an extended length of 2 m and a tucked length of 1 m. (a) Determine the factor by which her angular velocity in the tucked configuration exceeds her angular velocity in the extended configuration. It is reasonable to approximate the body in the two configurations as cylinders, and to assume that the centroid of the cylinder coincides with the center of mass of the diver. In the extended configuration the cylinder has a length of two meters and an average radius of 0.1 m while in the tucked configuration the cylinder will have a length of 1 m and an average radius of 0.1414 m. Recall that the mass moment of inertia of a cylinder about an axis perpendicular to its long axis and passing through its mass center is $M(3r^2 + h^2)/12$, where r is the radius of the cylinder and h is the height of the cylinder. (b) What is the parameter that predominates in the determination of this ratio?

1.7 The Deformable Continuum Model

The deformable continuum model differs from the particle and rigid object models in that relative movement or motion is permitted between two points in the model. Examples of deformable continua include the elastic solid used, for example, in the

analysis of beam bending and viscous fluid used, for example, in the model of flow through a pipe. The use of the word "continuum" in describing these models stems from the idea illustrated in Fig. 1.1, namely that the model is a domain of Cartesian space in the same shape of the object being modeled. It therefore has all the properties necessary to use the analytical machinery of the calculus. In particular, displacements, strains, velocities, and rates of deformation may be calculated. These developments will be presented in the next and subsequent chapters. The deformable continuum is the focus of this text because it is the primary class of models employed in the study of solid and fluid at both the macro scale, at the nanometer scale, and at scales in between.

1.8 Lumped Parameter Models

Lumped parameter models are extended rigid object models in which some of the elements are assumed not to be rigid, but to respond in simplified specific ways. The word "lumping" is used to imply that not all the properties are modeled exactly, but in a somewhat approximate way. For example, in a lumped parameter model the image of the object in Euclidean space, as shown in Fig. 1.1, need not be an exact model of the object, just a model that contains the features the modeler desires. The mechanical concept of "Coulomb friction" is a "lumped" concept as it occurs in the formula of the French engineer Charles Augustin de Coulomb (1736–1806). The static friction formula of Coulomb is employed to express the force F necessary to cause the motion of weight W resting upon a frictional horizontal surface as $F = \mu W$, where μ represents the coefficient of friction. The sources of what is called "friction" between the surface and the weight W are varied and include, among other things, the effect of surface adhesion, surface films, lubricants, and roughness; these effects are "lumped" together in the concept of Coulomb friction and expressed as a single coefficient, μ.

When linear springs and dashpots are used as elements in a model they are "lumped" representations of an object's stiffness or damping. Their properties describe the constitution of the element and are called constitutive properties. The spring element is also called the Hookian model (Fig. 1.6a) and is characterized by an equation that relates its overall lengthening or shortening, x, to the force applied to the spring, F, by a spring constant, k; thus $F = kx$ (Love, 1927). This model is named after the English natural philosopher Robert Hooke (1635–1703). The dashpot is called the viscous model or damper (Fig. 1.6b) and is characterized by an equation that relates the rate of its overall lengthening or shortening, $\mathrm{d}x/\mathrm{d}t$, to the force applied to the dashpot, F, by a damping constant, η; thus $F = \eta\,(\mathrm{d}x/\mathrm{d}t)$.

Lumped parameter models employing springs and dashpots are used extensively in the study of mechanical systems. Simple forms of these models are used to explain the material response phenomena called creep and stress relaxation. *Creep* is the increasing strain exhibited by a material under constant loading as the time increases. A typical creep experiment on a specimen of material is performed by

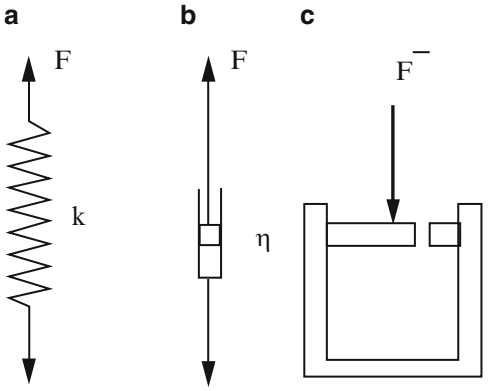


Fig. 1.6 Three lumped parameter models. (**a**) the Hookian spring element, (**b**) the damping element or viscous dashpot, and (**c**) the Darcy or permeability element

placing a constant tensile load on a specimen of the material and measuring the strain as a function of time. The function of time obtained by dividing the resulting strain against time data by a unit load is called the *creep* function. *Stress relaxation* is the reduction or decay of stress in a material under constant strain as the time increases. In a typical stress relaxation experiment on a material a constant tensile strain is applied to a material specimen and the resultant stress is recorded as a function of time. The function of time obtained by dividing the resulting stress against time data by a unit strain is called the *stress-relaxation* function. Equations for representing creep and stress relaxation will be obtained in the discussion of the standard linear solid model below. Materials that exhibit the time-dependent behaviors of creep and stress relaxation are called *viscoelastic*, indicating that they have some properties of both a viscous fluid and an elastic solid (Christensen 1971; Lockett 1972; Pipkin 1972).

As an example of the use of a lumped parameter model, consider the question of determining the stiffness of elastic compliance of a running track so that the runner has the optimal advantage of the elastic rebound of the track surface. This question was answered using the spring and dashpot model of the lower limb shown in Fig. 1.7. While running on soft surfaces like sod is easier on the body than running on hard surfaces like concrete, runners know that they run faster on the harder surface. The question of how hard the surface should be was answered (McMahon & Greene 1978; McMahon & Greene 1979) by tuning the compliance of the track to the compliance of the model of the runner's leg shown in Fig. 1.7. The half dozen or so running tracks that have been constructed on the basis of this model are known to runners as "fast" tracks.

Each of these lumped parameter models is an ideogram for a constitutive idea, e.g., elasticity, damping, or flow through porous media. The Darcy or permeability element is a special lumped parameter model peculiar to porous media. It was developed to explain the flow of fluids though porous media. Specifically, the city engineer of Dijon, France (Henri Philibert Gaspard Darcy, 1803-1858) in the middle of the 1800s developed the model to analyze the flow of water through a packed sand layer in a city fountain (Darcy 1856). A sketch of the type of

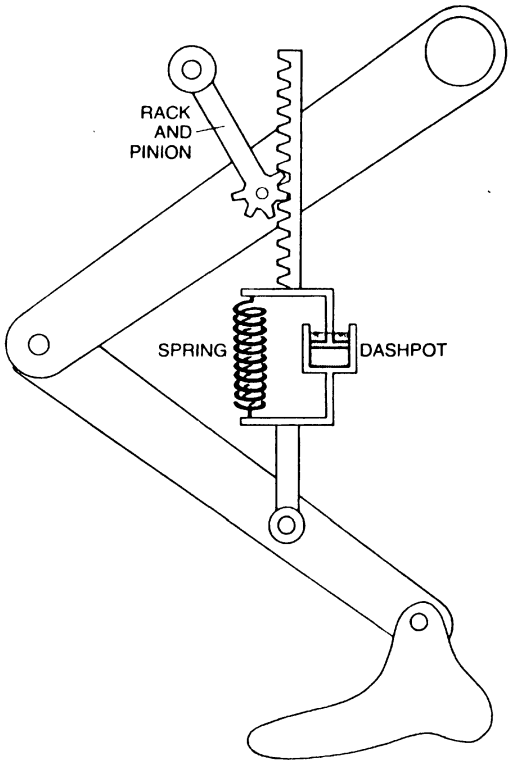


Fig. 1.7 Schematic diagram showing the conceptual model of the leg used to predict the runner's performance on a compliant track. Descending commands from the cortex, brain stem, and spinal centers (acting to crank the rack and pinion) are assumed to be separate from the mechanical properties of the muscles plus local reflexes (*parallel spring* and *dashpot*). From McMahon and Greene (1979)

experiment that Darcy did is shown in Fig. 1.8. A layer of sand of thickness L is supported on a stiff wire mesh with mesh openings larger than the size of the typical fluid passages through the sand layer. On top of the layer of sand is a reservoir of water maintained at a constant height h. The constancy of this head is maintained by providing a continuous supply of water to the reservoir and providing an over flow run-off passage. The domain at the bottom of the sand layer is open to the air again as is the upper surface of the water reservoir; the air pressure is p_o. Thus the decrease in water pressure across the sand layer is from $p + p_o$ where $p = \rho g h$, to p_o. The pressure gradient across the layer is then p/L, where L is the layer thickness. The volume flux of water, that is to say, the volume of water coming out of the sand layer per unit area per unit time, is denoted by q. If the cross-sectional area of the sand layer is denoted by A_o, then the volume per unit time is $A_o(\mathrm{d}w/\mathrm{d}t)$ where $\mathrm{d}w/\mathrm{d}t$ is the rate at which the surface of the water in the catchment basin of cross-sectional area A_o below the sand layer is filling with water. The volume flux q through the layer is then given by $q = (1/A_o)\, A_o(\mathrm{d}w/\mathrm{d}t) = (\mathrm{dw}/\mathrm{d}t)$. Darcy found that the volume flux q through the sand layer was proportional to the pressure gradient across the sand layer, p/L. The constant of proportionality κ is called the permeability and it is calculated in the experiment described by the formula $\kappa = qL/p$. Replacing p/L by the gradient, $(\partial p/\partial L)$, Darcy's law is written

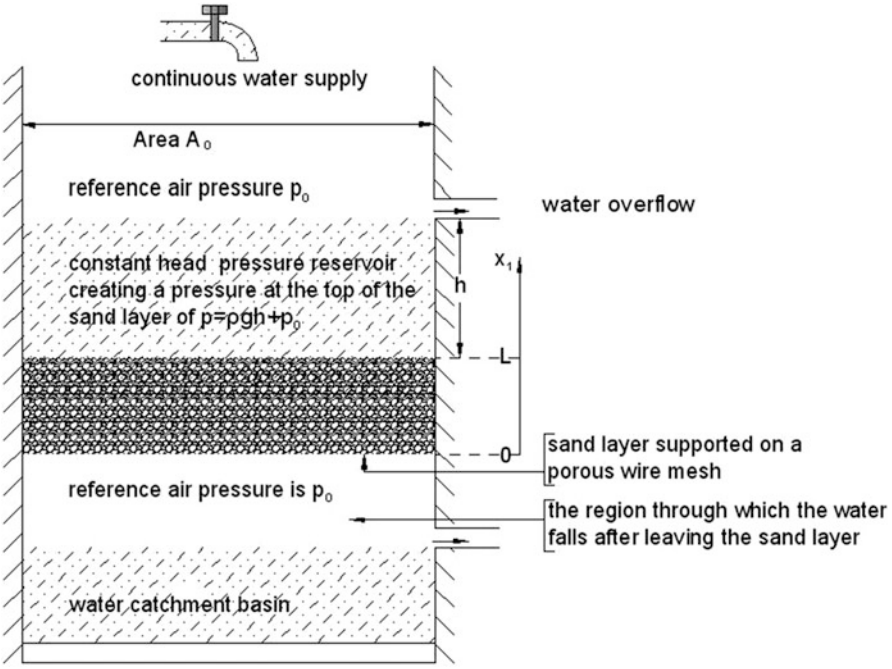


Fig. 1.8 An illustration of an experimental determination of the permeability of a sand layer

$$q = -\kappa \, (\partial p / \partial x), \tag{1.6}$$

where L has been replaced by the coordinate x in the formation of the one-dimensional gradient operator. The minus sign is placed in (1.6) so that the permeability κ is positive. The fluid flow is always from regions of higher fluid pressure to regions of lower fluid pressure, hence the pressure gradient in (1.6) is always negative. The combination of the minus sign in (1.6) and the always-negative pressure gradient mean the volume flow rate q is always positive.

The constitutive idea of the permeability element (Fig. 1.6c) is that of a distributed volumetric resistance to flow throughout the layer thickness L of the porous medium. When a compressive force F is applied to the piston of the permeability element, the water in the chamber under the piston is subjected to a higher pressure and a pressure difference $p + p_o$ is created between the inside of the chamber and the air pressure p_o outside. The water in the chamber then flows from the high-pressure region $p + p_o$ to the low-pressure region p_o and it passes out of the chamber through the hole in the piston. This process continues until all the water has been ejected from the chamber, the piston has moved to the bottom of the cylinder and the chamber no longer exists. The volume flow rate q is the uniform fluid velocity over the cross-section A_o of the orifice in the piston. The pressure

difference between the fluid chamber and the outside air is p, thus $q = \kappa(p/L) = \kappa(\partial p/\partial L)$ where κ is the permeability constant.

Example 1.8.1

Show that, from the viewpoint of the relationship between applied force F and time rate of change of the deflection ($\mathrm{d}x/\mathrm{d}t$), the permeability element is equivalent to a dashpot element characterized by the constant η. An equivalent statement of the problem would be to show that the dashpot constant η of the permeability element (Fig. 1.6c) is related to the permeability κ by $\eta = \frac{A^2}{A_o}\frac{L}{\kappa}$ where A is the area of the piston in the permeability element, L is the thickness of the piston, and A_o is the cross-sectional area of the orifice in the piston.

Solution: To show this, first note that a force balance applied to a free object diagram of the piston of the permeability element (Fig. 1.6c) shows that $F = pA$ where A is the cross-sectional area of the piston. If $\mathrm{d}x/\mathrm{d}t$ denotes the time rate of changes of the downward movement of the piston and q the volume flow rate through the orifice in the piston, then three different representations of the time rate of change of the fluid volume in the cylinder chamber are given by

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -A\frac{\mathrm{d}x}{\mathrm{d}t} = -qA_o$$

where A_o is the cross-sectional area of the piston orifice. It follows that q, the volume flow rate per unit area through the orifice, may be expressed as

$$q = \frac{A}{A_o}\frac{\mathrm{d}x}{\mathrm{d}t}.$$

Combining this result with $F = pA$ and $q = \kappa(p/L)$, a constitutive relation for the dashpot of the form, $F = \eta(\mathrm{d}x/\mathrm{d}t)$, is again obtained but with

$$\eta = \frac{A^2}{A_o}\frac{L}{\kappa}.$$

This result identifies the source of the dashpot-like viscous loss in the permeability element as fluid movement. The velocity of the fluid movement and the fluid pressure are the primary parameters in the permeability element; the applied force and the piston deflection associated with the dashpot element are secondary. It will come as no surprise that the permeability element will often behave just like a dashpot element. However the permeability element is not the same as the dashpot element because the source of the viscous loss is identified as fluid movement in the permeability element and it is unspecified in the dashpot element.

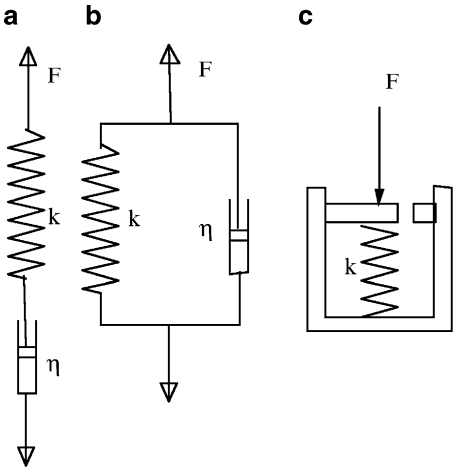


Fig. 1.9 Lumped two-parameter models. (**a**) The Maxwell element; (**b**) the Voigt element; and (**c**) the consolidation or Terzaghi element

Each of the three lumped parameter models described has been characterized by a single constitutive parameter, namely the spring constant, the damping constant, and the permeability constant. The next higher level of lumped parameter models is characterized by combinations of these elementary models with two constitutive parameters, and the level after that by three-parameter models. There are models with more than three parameters, but they are less useful. The three two-parameter models of interest, the Maxwell model, the Voigt model, and the consolidation model are illustrated in Fig. 1.9. James Clerk Maxwell (1831–1879) was a Scottish natural philosopher who first formulated the basic equations of electromagnetism ("The Maxwell Equations"), and Woldemar Voigt (1850–1919) was a German theoretical physicist who wrote a classic volume on crystal physics. The Maxwell model is a combination of a spring and a dashpot in series. When a force applied to a Maxwell model is changed from 0 to a finite value at an instant of time and held constant thereafter, there is an instantaneous initial elastic extension and then there is a continued deformation forever as the damper in the dashpot is drawn through the dashpot cylinder. Thus a Maxwell model exhibits the characteristics of a fluid with an initial elastic response.

Neither the Maxwell model nor the Voigt models are considered to be particularly good models of the force-deformation-time behavior of real materials.

The Terzaghi consolidation model element is a special lumped two-parameter model peculiar to soil mechanics (see, for example (Terzaghi 1943)) that has applications in geomechanics and biomechanics to interstitial water flow in both hard and soft tissues. The model is constructed by combining a spring element and a Darcy or permeability element in parallel (Fig. 1.9c). It was developed by Karl Terzaghi (1883–1963) about 1923 to explain the settlement or consolidation of the soils under the foundations of buildings built on porous, water-saturated soils. A sponge easily illustrates the mechanical principles involved. If the sponge is waterlogged, then its compression under loading can only proceed as fast as the water can drain from the sponge. When a load is initially placed on the piston in

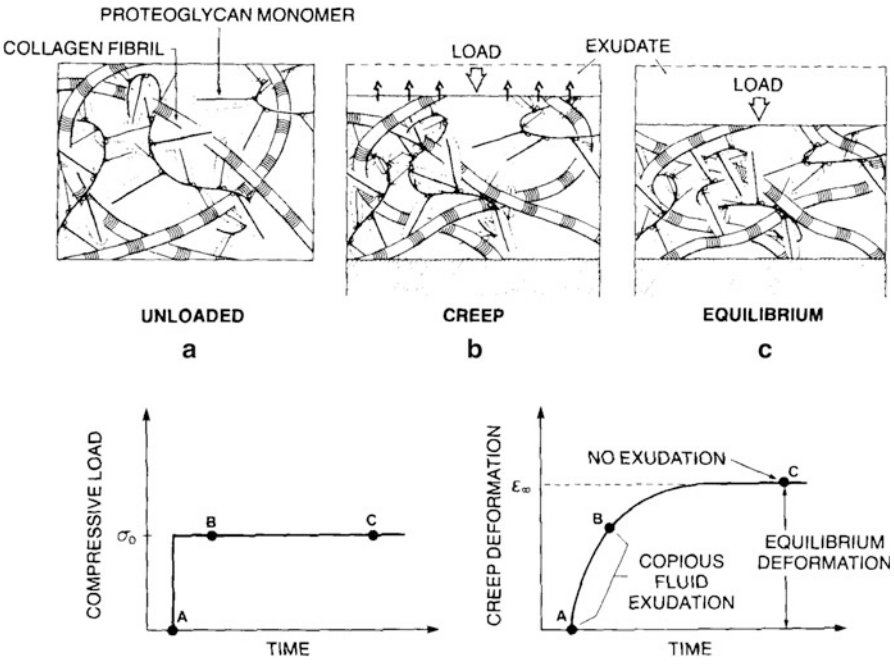


Fig. 1.10 The application of the Terzaghi model to articular cartilage. On the bottom left the applied stress history is illustrated. A constant stress (σ_o) is applied to a sample of articular cartilage at an instant of time and held constant thereafter. The creep response of the sample under this loading is illustrated on the *bottom right*. At the *top* are drawings of the sample illustrating the tissue response. From these illustrations one can see that the creep is accompanied by exudation of the fluid from the sample and that the rate of exudation decreases over time from point A to B to C. At equilibrium the flow ceases and the load is borne entirely by the solid matrix (point C) and the water carries none of the load (due to the fact that any pressure in the water would cause it to flow out of the matrix) (From Nordin and Frankel (1989))

Fig. 1.9c the spring supporting the piston initially takes none of the load because it cannot deflect. It cannot deflect because the piston cannot move due to the fact that the cylinder chamber is filled with a relatively incompressible liquid (water). However, once the water has a chance to exit the cylinder through the orifice in the piston, then it is possible for the piston to begin to move downward under the action of the applied compressive loading. This process of consolidation or settlement proceeds until the spring has deflected an amount sufficient to create a force equal to the applied loading. The relation between the force F applied to the piston of this model and the resulting displacement x of the piston is given by the same constitutive equation that characterizes the Voigt model if the dashpot viscosity η in the Voigt model is replaced by $(A^2/A_o)(L/\kappa)$. The application of this model to articular cartilage is illustrated in Fig. 1.10. This figure illustrates the response of a sample of articular cartilage when a force applied to the sample is changed from 0 to a finite value at an instant of time and held constant thereafter. Like the Voigt

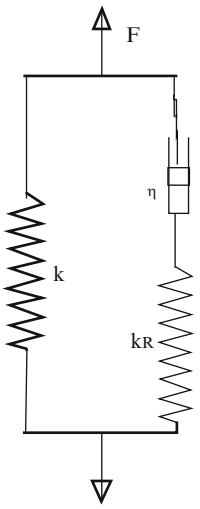


Fig. 1.11 The standard linear solid

element under the same loading, there is no instantaneous deflection, but a creeping deflection begins under the constant applied stress and proceeds asymptotically to a rest value. There is only one lumped three-parameter element of interest, the standard linear solid (SLS, Fig. 1.11).

Example 1.8.2

Using the Fig. 1.11 as a guide, derive the differential equation of the governing force–deflection relationship of the standard linear solid.

Solution: Let F_{L} and F_{R} denote the force in the two branches, left and right, of the standard linear solid; the total force F is then given by $F = F_{\mathrm{L}} + F_{\mathrm{R}}$. Let x denote the overall deflection of the standard linear solid element; the deflections in both branches must be equal; the horizontal cross-bars in spring-dashpot models are not allowed to rotate. The total deflection in the right branch is the sum, x, of the deflection of the dashpot, x_D, and the deflection of the spring, x_S; thus $x = x_D + x_S$. The equations describing the behavior of the three constituent elements are $F_{\mathrm{L}} = k\,x$, $F_{\mathrm{R}} = k_R x_S$ and $F_{\mathrm{R}} = \eta\,(\mathrm{d}x_D/\mathrm{d}t)$, respectively. Note that the force in the two elements on the right branch must be the same. These equations are combined in the following manner. First, note that from $F_{\mathrm{R}} = k_R x_S$ it follows that $(\mathrm{d}x_S/\mathrm{d}t) = (1/k_R)(\mathrm{d}F_{\mathrm{R}}/\mathrm{d}t)$, from $F_R = \eta(\mathrm{d}x_D/\mathrm{d}t)$ it follows that $(\mathrm{d}x_D/\mathrm{d}t) = (1/\eta)F_{\mathrm{R}}$, and from $x = x_D + x_S$ it follows that $(\mathrm{d}x/\mathrm{d}t) = (\mathrm{d}x_D/\mathrm{d}t) + (\mathrm{d}x_S/\mathrm{d}t)$. Combining these results it follows that

$$(\mathrm{d}x/\mathrm{d}t) = (1/\eta)F_R + (1/k_R)(\mathrm{d}F_R/\mathrm{d}t).$$

Now from the equation for the sum of the forces in the two branches ($F = F_{\mathrm{L}} + F_{\mathrm{R}}$) it follows that $F_{\mathrm{R}} = F - F_{\mathrm{L}}$ and since $F_{\mathrm{L}} = k\,x, F_{\mathrm{R}} = F - k\,x$. Substituting this equation for F_{R} in the equation involving F_{R} above it follows that

$$F + (\eta/k_R)(\mathrm{d}F/\mathrm{d}t) = kx + \eta\,(1 + k/k_R)\,(\mathrm{d}x/\mathrm{d}t). \tag{1.7}$$

The physical implications of the constant parameters characterizing the standard linear solid are easier to understand if the constants η and k_R are redefined in terms of time constants. To that end (1.7) is rewritten in the form

$$F + \tau_x \frac{\mathrm{d}F}{\mathrm{d}t} = k\left\{x + \tau_F \frac{\mathrm{d}x}{\mathrm{d}t}\right\}, \tag{1.8}$$

where τ_x and τ_F are material time constants defined by

$$\tau_x = \frac{\eta}{k_R} \quad \text{and} \quad \tau_F = \frac{\eta}{k}\left\{1 + \frac{k}{k_R}\right\} = \frac{\eta}{k} + \tau_x \tag{1.9}$$

These material time constants will be shown to have interpretations as the characteristic relaxation times of the load associated with a steady, constant deflection and the deflection associated with a steady, constant load, respectively. Note that, from the definitions, $\tau_F > \tau_x$ since η, k, and k_R are positive.■

The standard linear solid, characterized by the linear differential equation (1.8), provides a reasonable first model for the phenomena of creep under constant load and stress relaxation under constant deflection, phenomena observed in many materials. The creep function is the increase in time of the deflection $x(t)$ when a unit force is applied to the element at $t = 0$ and held constant forever. The relaxation function is the decrease in time of the force $F(t)$ when a unit deflection is applied to the element at $t = 0$ and held constant forever. The creep and relaxation functions for the standard linear solid are obtained from solving the governing differential equation (1.8). In Sect. A.16 of the Appendix the Laplace transform method (Thomson 1960) for solving the differential equation (1.8) is described; this is the simplest method of solution. In Sect. A.17 of the Appendix the more complicated approach of using direct integration is described. The creep function is the solution of (1.8) for $x(t)$ when $F(t)$ is specified to be the unit step function $h(t)$ and the relaxation function is the solution of (1.8) for $F(t)$ when $x(t)$ is specified to be the unit step function $h(t)$ (see (A219) in the Appendix for a definition of the unit step function). The unit step function $h(t)$ is therefore employed in the representation of deflection history $x(t)$ to obtain the creep function $c(t)$ as well as in the representation of the force history $F(t)$ to obtain the relaxation function $r(t)$. The creep function $c(t)$ for the standard linear solid is given by

$$c(t) = \frac{h(t)}{k}\left\{1 - \left(1 - \frac{\tau_x}{\tau_F}\right)\mathrm{e}^{-t/\tau_F}\right\}. \tag{1.10}$$

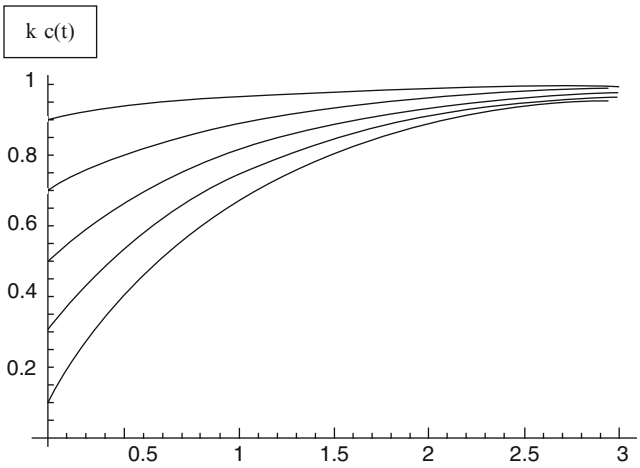


Fig. 1.12 The creep function. This is a plot of the spring constant k times the creep function c(t), on the abscissa, against the dimensionless time ratio t/τ_F on the ordinate. See (1.10). The five curves are for different ratios of τ_x/τ_F. Since $0 < \tau_x/\tau_F < 1$ the plots, from *bottom* to *top* are for values of τ_x/τ_F equal to 0.1, 0.3, 0.5, 0.7 and 0.9

From this result it may be seen that τ_F is indeed the characteristic relaxation time of the deflection at constant load. Plots of the creep function $c(t)$ (multiplied by the spring constant k) against the dimensionless time ratio t/τ_F for different ratios of τ_x/τ_F are shown in Fig. 1.12. Because the values of τ_x/τ_F are restricted by $0 < \tau_x/\tau_F < 1$, the plots in Fig. 1.12 are, from bottom to top, for values of τ_x/τ_F equal to 0.1, 0.3, 0.5, 0.7, and 0.9. The relaxation function for the standard linear solid is given by

$$r(t) = k\left\{1 + \left(\frac{\tau_F}{\tau_x} - 1\right)\mathrm{e}^{-t/\tau_x}\right\}h(t). \tag{1.11}$$

From this result it may be seen that τ_x is indeed the characteristic relaxation time of the load at constant deflection. Plots of the relaxation function $r(t)$, divided by the spring constant k, against the dimensionless time ratio t/τ_x for different ratios of τ_x/τ_F are shown in Fig. 1.13. The values of τ_x/τ_F employed are 0.1, 0.3, 0.5, 0.7, and 0.9; the same values as in Fig. 1.12.

Higher order lumped parameter models are obtained by combining the lower order models described above. There is a strong caveat against the process of combining elementary lumped parameter models to build higher order models. The caveat is that the number of parameters increases and defeats the advantage of simplicity of the lumped parameter model. Thus, the standard linear solid considered is the most reasonable of the spring and dashpot models as a first approximation of the force-deformation-time behavior of real materials.

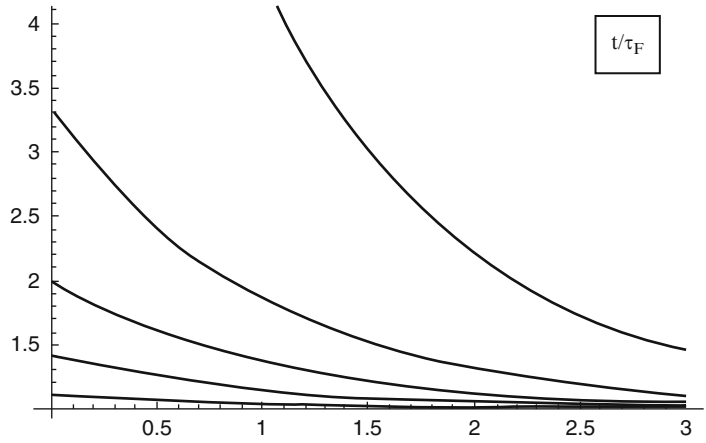


Fig. 1.13 The relaxation function. This is a plot of the relaxation function $r(t)$, divided by the spring constant k, against the dimensionless time ratio on the ordinate. See equation (1.11). The five curves are for different ratios of τ_x/τ_F. Since $0 < \tau_x/\tau_F < 1$ the plots, from the top to the bottom are for values of τ_x/τ_F equal to 0.1, 0.3, 0.5, 0.7 and 0.9. The *top curve*, $\tau_x/\tau_F = 0.1$, will intersect the abscissa a little above the value $r(t)/k = 10$.

Problems

1.8.1. Show that the constitutive relation governing the Maxwell element is $F + (\eta/k)(\mathrm{d}F/\mathrm{d}t) = \eta(\mathrm{d}x/\mathrm{d}t)$.

1.8.2. Show that the constitutive relation governing the Voigt element is $F = k\,x + \eta(\mathrm{d}x/\mathrm{d}t)$.

1.8.3. Show that the creep function for the Maxwell element is $c(t) = [(1/k) + (t/\eta)]h(t)$.

1.8.4. Show that the creep function for the Voigt element is $c(t) = (1/k)[1–\exp(-kt/\eta)]h(t)$.

1.8.5. Show that the relaxation function for the Maxwell element is $r(t) = k[\exp(-kt/\eta)]h(t)$.

1.8.6. Show that the relation between the compressive force F applied to the piston of the Terzaghi or consolidation element and the resulting compressive displacement x of the piston is given by the same constitutive equation that characterizes the Voigt model, namely $F = k\,x + \eta(\mathrm{d}x/\mathrm{d}t)$, if the dashpot viscosity η in the Voigt model is replaced by $(A^2/A_o)(h/\kappa)$ where A is the cross-sectional area of the piston, A_o is the cross-sectional area of the hole in the piston, h the thickness of the piston, and κ is the permeability constant.

1.8.7. Show that the relaxation function for the standard linear solid element is $r(t) = k[1 + ((\tau_F/\tau_x)-1)\exp(-t/\tau_x)]\,h(t)$.

1.8.8. Sketch the deflection of each of the three models shown in Fig. 1.9 to the following loading regime: before $t = 0$ the load is 0, at $t = 0$ the load is

increased to F and held at that value until $t = T$ when it is decreased to 0 again and held at 0 thereafter.

1.8.9. In Example 1.8.1 a calculation was presented that mathematically equated the permeability element and the dashpot element. However, the two elements do not appear to be the same. How do they differ? How are they similar?

1.9 The Limits of Reductionism and Determinism

A basic method of approach in science and engineering is reductionism. The philosophy of this approach is to decompose the object of study into its constituent parts, analyze each part separately, then reconstruct the object and predict its response to stimuli from a knowledge of response of the constituent parts to stimuli. Mathematically, such a prediction of the response of the object from the superposition of the response of its constituent parts generally implies a linear model. An example is the stress analysis of a large bridge that is accomplished by decomposing the bridge into bars, beams, girders, and cables and analyzing the structural capacity of each of these constituents individually. Since reductionism is the philosophy followed in this book, it is appropriate to mention some general and some specific caveats. The prime caveat is that the mathematical model of the system may be non-linear and that superposition of different responses from different stimuli is not a valid assumption.

Determinism is the philosophical proposition that every event, including human cognition and behavior, decision and action, is causally determined by prior occurrences. In the more restricted domain of mechanics, determinism is the idea that future mechanical events are predetermined by previous events. The philosophical concept of determinism is embedded in the representation of the motion of an object employed in mechanics and described in the following chapter. As the concept of determinism is employed in classical mechanics, it is the view of the determinism of the eighteenth century. The quote of the Marquis Pierre-Simon de Laplace (1759–1827) at the beginning of the next chapter captures the nineteenth century idea of determinism. In the nineteenth century the eighteenth century idea was modified by the Heisenberg uncertainty principle, the realization that certain pairs of physical properties, like position and momentum, cannot both be known to arbitrary precision. The greater the precision in measuring one variable limits or compromises the ability to measure the other. The idea of the uncertainty principle is also applied to the situation in which the measurement of a variable in an experimental situation distorts the experimental situation. Laplace's concept of determinism was also eroded in the twentieth century by the discovery of extreme sensitivity to starting or initial conditions for differential equations known as "chaos."

The quote of the Marquis Pierre-Simon de Laplace (1759–1827) at the beginning of the next chapter captures the idea of determinism underlying the representation (2.2). All the material in this book is based on the idealistic deterministic extreme

suggested in the quote of Laplace. The limitations on this idea just stated emphasize a limitation on the character of the models of nature that will be described.

References

Christensen RM (1971) Theory of viscoelasticity. Academic, New York
Darcy H (1856) Les Fontains Publiques de la Ville de Dijon. Dalmont, Paris
Lockett FJ (1972) Nonlinear viscoelastic solids. Academic, London
Love AEH (1927) Elasticity. Dover, New York
McMahon TA, Greene PR (1978) Fast running tracks. Scientific American 239:148–63
McMahon TA, Greene PR (1979) The influence of track compliance upon running. J Biomechanics 12:893–904
Nordin M, Frankel VH (1989) Basic Biomechanics of the Musculoskeletal System, Lea & Febiger, 2nd Ed.
Pipkin AC (1972) Lectures on viscoelasticity theory. Springer, New York
Terzaghi K (1943) Theoretical soil mechanics. Wiley, New York
Thomson WT (1960) Laplace transformation, 2nd edn. Prentice Hall, New York
Weyl H (1987) The continuum. Thomas Jefferson University Press, Kirksville

Chapter 2
Basic Continuum Kinematics

The theme of this chapter was stated with exuberance and in an idealistic deterministic extreme by Marquis Pierre-Simon de Laplace (1759–1827): "Thus, we must consider the present state of the universe as the effect of its previous state and as the cause of those states to follow. An intelligent being which, for a given point in time, knows all the forces acting upon the universe and the positions of the objects of which it is composed, supplied with facilities large enough to submit these data to numerical analysis, would include in the same formula the movements of the largest bodies of the universe and those of the lightest atom. Nothing would be uncertain for it, and the past and future would be known to it."[1]

2.1 The Deformable Material Model, the Continuum

In the deformable material model all types of motion are permitted, but the deformational motions are usually the major concern. Consider the image O of an object in Euclidean space. The object is in a configuration $O(0)$ at $t = 0$ and in a configuration $O(t)$ at time t (Fig. 2.1). The mathematical representation of the motion of a three-dimensional deformable continuum gives a complete history of the motion of each point P on the object $O(0)$, $P \subset O(0)$; in words, $P \subset O(0)$ means all points P contained in ($\subset$) the image of the object, O, at $t = 0$. In order to identify each point P in the object $O(0)$ and to follow the movement of that point in subsequent configurations of the object $O(t)$, each point on an object is given a reference location in a particular coordinate system, called the reference coordinate system. The selection of the reference configuration is the choice of the modeler; here the reference configuration is taken as the configuration of the object at time $t = 0$. To distinguish between the reference location of a point on an object and a location of the same point at a later time, the terminology of "particle" and "place"

[1] Translated by John H. Van Drie (http://www.johnvandrie.com).

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_2,

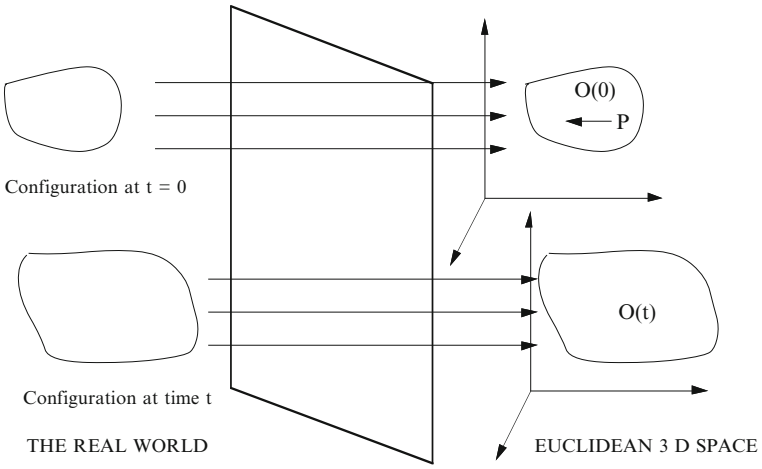


Fig. 2.1 Representation of the motion of an object in Euclidean 2D space

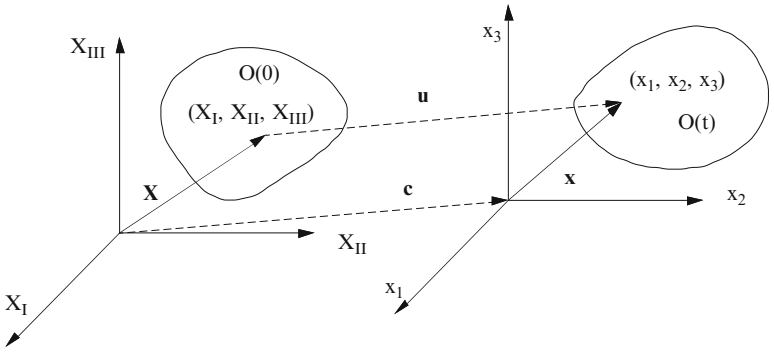


Fig. 2.2 Details of the representation in Euclidean 2D space

of a particle is introduced. Each point $P \subset O(0)$ in the continuum model of the object is labeled by its position in the reference configuration (Fig. 2.2). This procedure assigns a location to each point in the object and such points are called particles. A position vector of a point in a given coordinate system is a vector from the origin of coordinates to that point. In this case the reference configuration is a three-dimensional Cartesian coordinate system with base vectors $\mathbf{e}_\alpha$, $\alpha = \text{I, II, III}$, and coordinates X_α; the position of the particle is described by the vector

$$\mathbf{X} = X_\alpha \mathbf{e}_\alpha.$$

As a simplifying convention, instead of saying the vector $\mathbf{X}$ describes the position of a particle, we define it to be the particle. Thus the notation $\mathbf{X}$ has replaced the notation P and one can speak of all $\mathbf{X} \subset O(0)$ as a complete representation of the object in the reference configuration.

If the motion of one particle $\mathbf{X}$ of an object can be represented, then the motion of all the particles of the object, $\mathbf{X} \subset O(0)$, can be represented. A second coordinate system with axes x_i, $i = 1, 2, 3$, and base vectors $\mathbf{e}_i$, $i = 1, 2, 3$, is introduced to represent the present position of the object $O(t)$; this also represents the present positions of the particles. The triplet (x_1, x_2, x_3), denoted in the shorthand direct notation by $\mathbf{x}$, represents the place at time t of the particle $\mathbf{X}$. The motion of the particle $\mathbf{X}$ is then given by

$$x_1 = \chi_1(X_\mathrm{I}, X_\mathrm{II}, X_\mathrm{III}, t), \quad x_2 = \chi_2(X_\mathrm{I}, X_\mathrm{II}, X_\mathrm{III}, t), \quad x_3 = \chi_3 = (X_\mathrm{I}, X_\mathrm{II}, X_\mathrm{III}, t) \tag{2.1}$$

which is a set of three scalar-valued functions whose arguments are the particle $\mathbf{X}$ and time t and whose values are the components of the place $\mathbf{x}$ at time t of the particle $\mathbf{X}$. Since $\mathbf{X}$ can be any particle in the object, $\mathbf{X} \subset O(0)$, the motion (2.1) describes the motion of the entire object $\mathbf{x} \subset O(t)$ and (2.1) is thus referred to as the *motion* of the object O. In the direct shorthand or vector notation (2.1) is written

$$\mathbf{x} = \chi(\mathbf{X},\ t) \quad \text{for all} \quad \mathbf{X} \subset O(0). \tag{2.2}$$

This is called the *material description of motion* because the material particles $\mathbf{X}$ are the independent variables. Generally the form of the motion, (2.1) or (2.2), is unknown in a problem, and the prime kinematic assumption for all continuum models is that such a description of the motion of an object is possible.

However, if the motion is known, then all the kinematic variables of interest concerning the motion of the object can be calculated from it; this includes velocities, accelerations, displacements, strains, rates of deformation, etc. The present, past, and future configurations of the object are all known. The philosophical concept embedded in the representation (2.2) of a motion is that of determinism. The determinism of the eighteenth century in physical theory was modified by humbler notions of "uncertainty" in the nineteenth century and by the discovery of extreme sensitivity to starting or initial conditions known by the misnomer "chaos" in the twentieth century. The quote of the Marquis Pierre-Simon de Laplace (1759–1827) at the beginning of the chapter captures the idea of determinism underlying the representation (2.2).

A translational rigid object motion is a special case of (2.2) represented by,

$$\mathbf{x} = \mathbf{X} + \mathbf{h}(t) \quad \text{for all} \quad \mathbf{X} \subset O(0), \tag{2.3}$$

where $\mathbf{h}(t)$ is a time-dependent vector. A rotational rigid object motion is a special case of (2.2) represented by

$$\mathbf{x} = \mathbf{Q}(t)\mathbf{X}, \quad \mathbf{Q}(t)\mathbf{Q}(t)^\mathrm{T} = \mathbf{1} \quad \text{for all} \quad \mathbf{X} \subset O(0), \tag{2.4}$$

where $\mathbf{Q}(t)$ is a time-dependent orthogonal transformation. It follows that a general rigid object motion is a special case of (2.2) represented by

$$\mathbf{x} = \mathbf{Q}(t)\mathbf{X} + \mathbf{h}(t),\ \mathbf{Q}(t)\mathbf{Q}(t)^{\mathrm{T}} = \mathbf{1} \quad \text{for all} \quad \mathbf{X} \subset O(0). \tag{2.5}$$

A motion of the form (2.2) is said to be a planar motion if the particles always remain in the same plane. In this case (2.2) becomes

$$x_1 = \chi_1(X_{\mathrm{I}}, X_{\mathrm{II}}, t), \quad x_2 = \chi_2(X_{\mathrm{I}}, X_{\mathrm{II}}, t), x_3 = X_{\mathrm{III}}. \tag{2.6}$$

Another subset of the motion is a deformation of an object from one configuration to another, say from the configuration at $t = 0$ to the configuration at $t = t^*$. In this case the motion (2.2) becomes a deformation

$$\mathbf{x} = \Psi(\mathbf{X}) \quad \text{for all} \quad \mathbf{X} \subset O(0), \tag{2.7}$$

where

$$\Psi(\mathbf{X}) = \chi(\mathbf{X},\ t^*) \quad \text{for all} \quad \mathbf{X} \subset O(0). \tag{2.8}$$

A 3D motion picture or 3D video of the motion of an object may be represented by a subset of the motion (2.2) because a discrete number of images (frames) per second are employed,

$$\mathbf{x} = \chi(\mathbf{X},\ n/\zeta) \quad \text{for all} \quad \mathbf{X} \subset O(0),\ n = 0,\ 1,\ 2, \ldots, \tag{2.9}$$

where ζ is the number of images (frames) per second.

Example 2.1.1
Consider the special case of a planar motion given by

$$x_1 = A(t)X_{\mathrm{I}} + C(t)X_{\mathrm{II}} + E(t), \quad x_2 = D(t)X_{\mathrm{I}} + B(t)X_{\mathrm{II}} + F(t), \quad x_3 = X_{\mathrm{III}}, \tag{2.10}$$

where $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, $F(t)$ are arbitrary functions of time. Further specialize this motion by the selections

$$A(t) = 1 + t, \quad C(t) = t, \quad E(t) = 3t, \quad B(t) = 1 + t, \quad D(t) = t, \quad F(t) = 2t, \tag{2.11}$$

for $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, and $F(t)$. With these selections the motion becomes

$$x_1 = (1 + t)X_{\mathrm{I}} + tX_{\mathrm{II}} + 3t, \quad x_2 = tX_{\mathrm{I}} + (1 + t)X_{\mathrm{II}} + 2t, \quad x_3 = X_{\mathrm{III}}. \tag{2.12}$$

The problem is to find the positions of the unit square whose corners are at the material points $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (0, 0)$, $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (1, 0)$, $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (1, 1)$, $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (0, 1)$ at times $t = 1$ and $t = 2$.

Solution: For convenience let the spatial (x_1, x_2, x_3) and material $(X_{\mathrm{I}}, X_{\mathrm{II}}, X_{\mathrm{III}})$ coordinate systems coincide and then consider the effect of the motion (2.12) on the

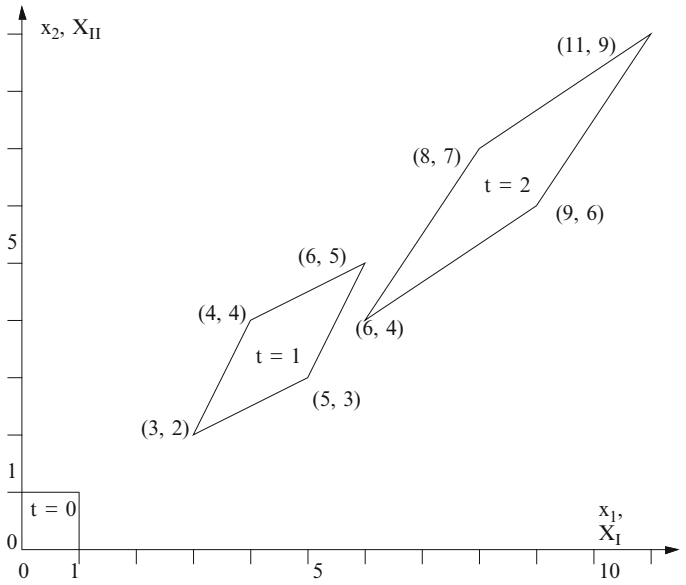


Fig. 2.3 The movement of a square at $t = 0$ due to the motion (2.12)

unit square whose corners are at the material points $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (0, 0)$, $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (1, 0)$, $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (1, 1)$, $(X_{\mathrm{I}}, X_{\mathrm{II}}) = (0, 1)$. At $t = 0$ the motion (2.12) specifies that $x_1 = X_{\mathrm{I}}$, $x_2 = X_{\mathrm{II}}$, and $x_3 = X_{\mathrm{III}}$ so that $t = 0$ has been taken as the reference configuration. The square at $t = 0$ is illustrated in Fig. 2.3. At $t = 1$ the motion (2.12) specifies the places $\mathbf{x}$ of the particles $\mathbf{X}$ as follows:

$$x_1 = 2X_{\mathrm{I}} + X_{\mathrm{II}} + 3, \quad x_2 = X_{\mathrm{I}} + 2X_{\mathrm{II}} + 2, \quad x_3 = X_{\mathrm{III}}.$$

Thus the particles at the four corners of the unit square have the following new places $\mathbf{x}$ at $t = 1$:

$$(3,\ 2) = \chi(0,\ 0),\ (5,\ 3) = \chi(1,\ 0),\ (6,\ 5) = \chi(1,\ 1),\ (4,\ 4) = \chi(0,\ 1).$$

A sketch of the deformed and translated unit square at $t = 1$ is shown in Fig. 2.3. At $t = 2$ the motion (2.12) specifies the places $\mathbf{x}$ of the particles $\mathbf{X}$ as follows:

$$x_1 = 3X_{\mathrm{I}} + 2X_{\mathrm{II}} + 6, \quad x_2 = 2X_{\mathrm{I}} + 3X_{\mathrm{II}} + 4, \quad x_3 = X_{\mathrm{III}}.$$

Thus the particles at the four corners of the unit square have the following new places at $t = 2$:

$$(6, 4) = \chi(0, 0), (9, 6) = \chi(1, 0), (11, 9) = \chi(1, 1), (8, 7) = \chi(0, 1).$$

A sketch of the deformed and translated unit square at $t = 2$ is shown in Fig. 2.3.

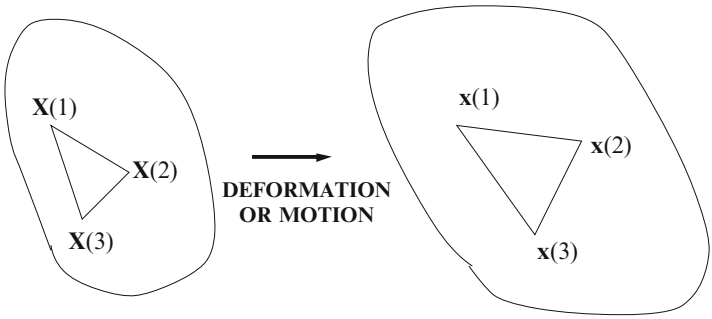


Fig. 2.4 The experimental measurement of a planar homogeneous motion. The reference frame is the laboratory reference frame. The three initial positions ($\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$, $\mathbf{X}^{(3)}$) of the markers are indicated as well as their positions ($\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, $\mathbf{x}^{(3)}$) at time t. In many experiments the markers are attached to a specimen of soft tissue that is undergoing a planar homogeneous motion in order to quantify the motion

Example 2.1.2
An experimental technique in widespread use in the measurement of the planar homogeneous motion of a deformable object is to place three markers (dots or beads) in triangular pattern (so that the markers are not collinear) on the deformable object before a motion. The initial locations of the three markers are recorded relative to a fixed laboratory frame of reference as $(X_{\mathrm{I}}^{(1)}, X_{\mathrm{II}}^{(1)})$, $(X_{\mathrm{I}}^{(2)}, X_{\mathrm{II}}^{(2)})$, and $(X_{\mathrm{I}}^{(3)}, X_{\mathrm{II}}^{(3)})$, Fig. 2.4. If the process is automated a camera is used to follow the motion of the three markers with time and to digitize the data in real time. The instantaneous locations of the three markers at a time t is recorded relative to a fixed laboratory frame of reference as $(x_1^{(1)}(t), x_2^{(1)}(t))$, $(x_1^{(2)}(t), x_2^{(2)}(t))$ and $(x_1^{(3)}(t), x_2^{(3)}(t))$, Fig. 2.4. From these data the experimentalist calculates the time-dependent coefficients $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, and $F(t)$ of the homogeneous planar motion (2.10). Determine the formulas used in the calculation of the time-dependent coefficients $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, and $F(t)$ from the data $(X_{\mathrm{I}}^{(1)}, X_{\mathrm{II}}^{(1)})$, $(X_{\mathrm{I}}^{(2)}, X_{\mathrm{II}}^{(2)})$, $(X_{\mathrm{I}}^{(3)}, X_{\mathrm{II}}^{(3)})$, $(x_1^{(1)}(t), x_2^{(1)}(t))$, $(x_1^{(2)}(t), x_2^{(2)}(t))$, and $(x_1^{(3)}(t), x_2^{(3)}(t))$.

Solution: The data on the motion of each marker provide two equations that may be used for the determination of the time-dependent coefficients. Since there are three markers, a total of six equations is obtained. Three markers are used because it is known that six equations will be needed to solve the linear system of equations for the six unknowns, $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, and $F(t)$. Using the notation for the data and the representation of the homogeneous planar motion (2.10), these six equations are as follows:

$$
\begin{aligned}
x_1^{(1)}(t) &= A(t)X_{\mathrm{I}}^{(1)} + C(t)X_{\mathrm{II}}^{(1)} + E(t), \quad x_2^{(1)}(t) = D(t)X_{\mathrm{I}}^{(1)} + B(t)X_{\mathrm{II}}^{(1)} + F(t),\\
x_1^{(2)}(t) &= A(t)X_{\mathrm{I}}^{(2)} + C(t)X_{\mathrm{II}}^{(2)} + E(t), \quad x_2^{(2)}(t) = D(t)X_{\mathrm{I}}^{(2)} + B(t)X_{\mathrm{II}}^{(2)} + F(t),\\
x_1^{(3)}(t) &= A(t)X_{\mathrm{I}}^{(3)} + C(t)X_{\mathrm{II}}^{(3)} + E(t), \quad x_2^{(3)}(t) = D(t)X_{\mathrm{I}}^{(3)} + B(t)X_{\mathrm{II}}^{(3)} + F(t).
\end{aligned}
$$

The solution to these six equations is

$$A(t) = \frac{X_{\mathrm{II}}^{(1)} x_1^{(2)}(t) - X_{\mathrm{II}}^{(1)} x_1^{(3)}(t) - X_{\mathrm{II}}^{(2)} x_1^{(1)}(t) + X_{\mathrm{II}}^{(2)} x_1^{(3)}(t) + X_{\mathrm{II}}^{(3)} x_1^{(1)}(t) - X_{\mathrm{II}}^{(3)} x_1^{(2)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}},$$

$$B(t) = \frac{X_{\mathrm{I}}^{(2)} x_2^{(1)}(t) - X_{\mathrm{I}}^{(3)} x_2^{(1)}(t) - X_{\mathrm{I}}^{(1)} x_2^{(2)}(t) + X_{\mathrm{I}}^{(3)} x_2^{(2)}(t) + X_{\mathrm{I}}^{(1)} x_2^{(3)}(t) - X_{\mathrm{I}}^{(2)} x_2^{(3)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}},$$

$$C(t) = \frac{-X_{\mathrm{I}}^{(1)} x_1^{(2)}(t) + X_{\mathrm{I}}^{(2)} x_1^{(1)}(t) + X_{\mathrm{I}}^{(1)} x_1^{(3)}(t) - X_{\mathrm{I}}^{(2)} x_1^{(3)}(t) - X_{\mathrm{I}}^{(3)} x_1^{(1)}(t) + X_{\mathrm{I}}^{(3)} x_1^{(2)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}},$$

$$D(t) = \frac{X_{\mathrm{II}}^{(1)} x_2^{(2)}(t) - X_{\mathrm{II}}^{(2)} x_2^{(1)}(t) - X_{\mathrm{II}}^{(1)} x_2^{(3)}(t) + X_{\mathrm{II}}^{(2)} x_2^{(3)}(t) + X_{\mathrm{II}}^{(3)} x_2^{(1)}(t) - X_{\mathrm{II}}^{(3)} x_2^{(2)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}},$$

$$\begin{aligned} E(t) = {} & \frac{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} x_1^{(3)}(t) - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} x_1^{(2)}(t) - X_{\mathrm{I}}^{(1)} X_{\mathrm{II}}^{(2)} x_1^{(3)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}} \\ & + \frac{X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} x_1^{(1)}(t) + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} x_1^{(2)}(t) - X_{\mathrm{I}}^{(2)} X_{\mathrm{II}}^{(3)} x_1^{(1)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}}, \end{aligned}$$

$$\begin{aligned} F(t) = {} & \frac{-X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} x_2^{(2)}(t) + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} x_2^{(1)}(t) + X_{\mathrm{I}}^{(2)} X_{\mathrm{II}}^{(1)} x_2^{(3)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}} \\ & + \frac{-X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} x_2^{(3)}(t) - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)} x_2^{(1)}(t) + X_{\mathrm{I}}^{(1)} X_{\mathrm{II}}^{(3)} x_2^{(2)}(t)}{X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(2)} - X_{\mathrm{II}}^{(1)} X_{\mathrm{I}}^{(3)} - X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(1)} + X_{\mathrm{II}}^{(2)} X_{\mathrm{I}}^{(3)} + X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(1)} - X_{\mathrm{II}}^{(3)} X_{\mathrm{I}}^{(2)}}. \end{aligned}$$

Example 2.1.3
Consider again the experimental technique described in Example 2.1.2, but in this case a deformation rather than a motion, Fig. 2.4. Suppose that the initial locations of the markers are recorded relative to the fixed laboratory frame of reference as $(X_{\mathrm{I}}^{(1)}, X_{\mathrm{II}}^{(1)}) = (0, 0)$, $(X_{\mathrm{I}}^{(2)}, X_{\mathrm{II}}^{(2)}) = (1, 0)$, and $(X_{\mathrm{I}}^{(3)}, X_{\mathrm{II}}^{(3)}) = (0, 1)$. The deformed locations of the three markers relative to the same fixed laboratory frame of reference are $(x_1^{(1)}, x_2^{(1)}) = (1, 2)$, $(x_1^{(2)}, x_2^{(2)}) = (2, 3.25)$, and $(x_1^{(3)}, x_2^{(3)}) = (2.5, 3.75)$. From these data the constant coefficients A, B, C, D, E, and F of the homogeneous planar deformation (2.10) are determined.

Solution: The solution for the motion coefficients $A(t), B(t), C(t), D(t), E(t)$, and $F(t)$ obtained in Example 2.1.2 may be used in the solution to this problem. One simply assigns the time-dependent positions in the formulas for $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, and $F(t)$ to be fixed rather than time dependent by setting $(x_1^{(1)}(t), x_2^{(1)}(t)) = (x_1^{(1)}, x_2^{(1)})$, $(x_1^{(2)}(t), x_2^{(2)}(t)) = (x_1^{(2)}, x_2^{(2)})$, and $(x_1^{(3)}(t), x_2^{(3)}(t)) = (x_1^{(3)}, x_2^{(3)})$. The coefficients

are no longer functions of time so they are denoted by A, B, C, D, E, and F. They are evaluated by substituting the initial and final locations of the set of particles, $(X_{\mathrm{I}}^{(1)}, X_{\mathrm{II}}^{(1)}) = (0,0), (X_{\mathrm{I}}^{(2)}, X_{\mathrm{II}}^{(2)}) = (1,0), (X_{\mathrm{I}}^{(3)}, X_{\mathrm{II}}^{(3)}) = (0,1)$, and $(x_1^{(1)}, x_2^{(1)}) = (1,2), (x_1^{(2)}, x_2^{(2)}) = (2, 3.25), (x_1^{(3)}, x_2^{(3)}) = (2.5, 3.75)$, respectively, into the last set of equations in Example 2.1.2. The values obtained are $A = 1$, $B = 1.75$, $C = 1.5$, $D = 1.25$, $E = 1$, and $F = 2$ and they are obtained by substituting the values for the relevant points given in the statement of the problem above into the last set of equations in Example 2.1.2. The planar homogeneous deformation then has the representation

$$x_1 = 2X_{\mathrm{I}} + 1.5X_{\mathrm{II}} + 1, \quad x_2 = 1.25X_{\mathrm{I}} + 1.75X_{\mathrm{II}} + 2, \quad x_3 = X_{\mathrm{III}},$$

which is a particular case of (2.10) To double check this calculation one can check to see if each marker is mapped correctly from its initial position to its final position.

There are two coordinate systems with respect to which a gradient may be taken, either the spatial coordinate system $\mathbf{x}$, (x_1, x_2, x_3), or the reference material coordinate system $\mathbf{X}$, $(X_{\mathrm{I}}, X_{\mathrm{II}}, X_{\mathrm{III}})$. To distinguish between gradients with respect to these two systems, the usual gradient symbol ∇ will be used to indicate a gradient with respect to the spatial coordinate system $\mathbf{x}$, and the gradient symbol $\nabla_{\mathbf{O}}$ with a subscripted boldface $\mathbf{O}$ will indicate a gradient with respect to the material coordinate system $\mathbf{X}$. The (material) deformation gradient tensor $\mathbf{F}$ is defined by

$$\mathbf{F} = [\nabla_{\mathbf{O}} \otimes \chi(\mathbf{X},\ t)]^{\mathrm{T}} \quad \text{for all } \mathbf{X} \subset O(0). \tag{2.13}$$

The (spatial) inverse deformation gradient tensor $\mathbf{F}^{-1}$ is defined by

$$\mathbf{F}^{-1} = [\nabla \otimes \chi^{-1}(\mathbf{x}, t)]^{\mathrm{T}} \quad \text{for all } \mathbf{x} \subset O(t), \tag{2.14}$$

where

$$\mathbf{X} = \chi^{-1}(\mathbf{x},\ t) \quad \text{for all } \mathbf{x} \subset O(t) \tag{2.15}$$

is the inverse of the motion (2.2). The components of $\mathbf{F}$ and $\mathbf{F}^{-1}$ are

$$\mathbf{F} = \left[\frac{\partial x_i}{\partial X_\alpha}\right] = \begin{bmatrix} \frac{\partial x_1}{\partial X_{\mathrm{I}}} & \frac{\partial x_1}{\partial X_{\mathrm{II}}} & \frac{\partial x_1}{\partial X_{\mathrm{III}}} \\ \frac{\partial x_2}{\partial X_{\mathrm{I}}} & \frac{\partial x_2}{\partial X_{\mathrm{II}}} & \frac{\partial x_2}{\partial X_{\mathrm{III}}} \\ \frac{\partial x_3}{\partial X_{\mathrm{I}}} & \frac{\partial x_3}{\partial X_{\mathrm{II}}} & \frac{\partial x_3}{\partial X_{\mathrm{III}}} \end{bmatrix} \text{ and } \mathbf{F}^{-1} = \left[\frac{\partial X_\alpha}{\partial x_i}\right] = \begin{bmatrix} \frac{\partial X_{\mathrm{I}}}{\partial x_1} & \frac{\partial X_{\mathrm{I}}}{\partial x_2} & \frac{\partial X_{\mathrm{I}}}{\partial x_3} \\ \frac{\partial X_{\mathrm{II}}}{\partial x_1} & \frac{\partial X_{\mathrm{II}}}{\partial x_2} & \frac{\partial X_{\mathrm{II}}}{\partial x_3} \\ \frac{\partial X_{\mathrm{III}}}{\partial x_1} & \frac{\partial X_{\mathrm{III}}}{\partial x_2} & \frac{\partial X_{\mathrm{III}}}{\partial x_3} \end{bmatrix} \tag{2.16}$$

respectively. Using the chain rule for partial derivatives it is easy to verify that $\mathbf{F}^{-1}$ is indeed the inverse of $\mathbf{F}$,

$$\mathbf{F}\mathbf{F}^{-1} = \mathbf{F}^{-1}\mathbf{F} = \mathbf{1}. \tag{2.17}$$

Recall that any motion can be decomposed into a sum of a translational, rotational, and deformational motion. The deformation gradient tensors remove the translational motion as may be easily seen because the translational motion is a separate function of time (cf., e.g., 2.2) that must be independent of the particle $\mathbf{X}$. Thus only the rotational motion and the deformational motion determine $\mathbf{F}$. If $\mathbf{F} = \mathbf{1}$ there are no rotational or deformational motions. If $\mathbf{F} = \mathbf{Q}(t)$, $\mathbf{Q}(t)\mathbf{Q}(t)^{\mathrm{T}} = \mathbf{1}$, it follows from (2.4) that the motion is purely rotational and there is no deformational motion. The *deformation gradient* $\mathbf{F}$ is so named because it is a measure of the deformational motion as long as $\mathbf{F} \neq \mathbf{Q}(t)$. If $\mathbf{F} = \mathbf{Q}(t)$, then the motion is rotational and we replace $\mathbf{F}$ by $\mathbf{Q}(t)$.

The determinant of the tensor of deformation gradients, J, is the Jacobian of the transformation from $\mathbf{x}$ to $\mathbf{X}$, thus

$$J \equiv \mathrm{Det}\ \mathbf{F} = 1/\mathrm{Det}\ \mathbf{F}^{-1}, \tag{2.18}$$

where it is required that

$$0 < J < \infty \tag{2.19}$$

so that a finite continuum volume always remains a finite continuum volume.

If $\mathbf{c}$ represents the position vector of the origin of the coordinate system used for the configuration at time t relative to the origin of the coordinate system used for the configuration at $t = 0$, then the displacement vector $\mathbf{u}$ of the particle $\mathbf{X}$ is given by (Fig. 2.2),

$$\mathbf{u} = \mathbf{x} - \mathbf{X} + \mathbf{c}. \tag{2.20}$$

The displacement vectors $\mathbf{u}$ for all the particles $\mathbf{X} \subset O(0)$ are given by

$$\mathbf{u}(\mathbf{X}, t) = \chi(\mathbf{X}, t) - \mathbf{X} + \mathbf{c}(t), \quad \mathbf{X} \subset O(0), \tag{2.21}$$

or by

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{x} - \chi^{-1}(\mathbf{x}, t) + \mathbf{c}(t), \quad \mathbf{x} \subset O(t). \tag{2.22}$$

Two gradients of the displacement field $\mathbf{u}$ may then be calculated, one with respect to the spatial coordinate system $\mathbf{x}$ denoted by the usual gradient symbol ∇ and one with respect to the material coordinate system $\mathbf{X}$ denoted by the gradient symbol $\nabla_{\mathbf{O}}$, thus

$$[\nabla_{\mathbf{O}} \otimes \mathbf{u}(\mathbf{X}, t)]^{\mathrm{T}} = \mathbf{F}(\mathbf{X}, t) - \mathbf{1} \quad \text{and} \quad [\nabla \otimes \mathbf{u}(\mathbf{x}, t)]^{\mathrm{T}} = \mathbf{1} - \mathbf{F}^{-1}(\mathbf{x}, t), \tag{2.23}$$

when (2.12) and (2.14) are employed. Often the base vectors of the coordinate systems **X** and **x** are taken to coincide, in which case the position vector **c** of the origin of the **x** system relative to the **X** system is zero. The selection of the coordinate system is always the prerogative of the modeler and such selections are usually made to simplify the analysis of the resulting problem.

Example 2.1.4
Compute the deformation gradient and the inverse deformation gradient for the motion given by (2.12). Then compute the Jacobian of the motion and both the spatial and material gradients of the displacement vector.

Solution: The deformation gradients and the inverse deformation gradients for this motion are obtained from (2.16) to (2.12), thus

$$\mathbf{F} = \begin{bmatrix} 1+t & t & 0 \\ t & 1+t & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{F}^{-1} = \frac{1}{1+2t} \begin{bmatrix} 1+t & -t & 0 \\ -t & 1+t & 0 \\ 0 & 0 & 1+2t \end{bmatrix},$$

a result that can be verified using $\mathbf{F}\mathbf{F}^{-1} = \mathbf{1}$ or $\mathbf{F}^{-1}\mathbf{F} = \mathbf{1}$. It is then easy to show that $J = 1 + 2t$. It also follows from (2.22) that

$$[\nabla_{\mathbf{O}} \otimes \mathbf{u}]^{\mathrm{T}} = t \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad [\nabla \otimes \mathbf{u}]^{\mathrm{T}} = \frac{t}{1+2t} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Problems

2.1.1. Sketch the shape and position of the unit square with corners at (0, 0), (1, 0), (1, 1), and (0, 1) subjected to the motion in (2.10) for the seven special cases, (a) through (g) below. The shape and position are to be sketched for each of the indicated values of t.

(a) Translation. $A(t) = 1, B(t) = 1, C(t) = 0, D(t) = 0, E(t) = 2t, F(t) = 2t$ and values of $t = 0, 1, 2$.
(b) Uniaxial extension. $A(t) = 1 + t, B(t) = 1, C(t) = 0, D(t) = 0, E(t) = 0,$ $F(t) = 0$ and values of $t = 0, 1, 2, 3$.
(c) Biaxial extension. $A(t) = 1 + t$, $B(t) = 1 + 2t$, $C(t) = 0$, $D(t) = 0$, $E(t) = 0, F(t) = 0$ and values of $t = 0, 1, 2$.
(d) Simple shearing (R). $A(t) = 1$, $B(t) = 1$, $C(t) = t$, $D(t) = 0$, $E(t) = 0$, $F(t) = 0$ and values of $t = 0, 1, 2$.
(e) Simple shearing (U). $A(t) = 1$, $B(t) = 1$, $C(t) = 0$, $D(t) = t$, $E(t) = 0$, $F(t) = 0$ and values of $t = 0, 1, 2$.
(f) Rigid Rotation (cw). $A(t) = \cos(\pi t/2)$, $B(t) = \cos(\pi t/2)$, $C(t) = \sin(\pi t/2)$, $D(t) = -\sin(\pi t/2)$, $E(t) = 0$, $F(t) = 0$ and values of $t = 0, 1, 2, 3, 4$.

(g) Rigid rotation (ccw). $A(t) = \cos(\pi t/2)$, $B(t) = \cos(\pi t/2)$, $C(t) = -\sin(\pi t/2)$, $D(t) = \sin(\pi t/2)$, $E(t) = 0$, $F(t) = 0$ and values of $t = 0, 1, 2, 3, 4$.

2.1.2. Sketch the shape and position of the square with corners at $(-1, -1)$, $(1, -1)$, $(1, 1)$, and $(-1, 1)$ at times $t = 0, 1, 2, 3, 4$. The square is subjected to the motion in (2.10) with the values of $A(t)$, $B(t)$, $C(t)$, $D(t)$, $E(t)$, and $F(t)$ being those given in 2.1(f), the rigid rotation (clockwise) motion.

2.1.3. For the six motions of the form (2.10) given in Problem 2.1.1, namely 2.1.1 (a) through 2.1.1(f), compute the deformation gradient tensor $\mathbf{F}$, its Jacobian J, and its inverse $\mathbf{F}^{-1}$. Discuss briefly the significance of each of the tensors computed. In particular, explain the form or value of the deformation gradient tensor $\mathbf{F}$ in terms of the motion.

2.1.4. Using the planar homogeneous deformation (2.10), with the values of A, B, C, D, E, and F calculated in Example 2.1.2, show that deformation (2.10) predicts the final positions of the three markers when the initial marker locations $(X_{\mathrm{I}}^{(1)}, X_{\mathrm{II}}^{(1)}) = (0, 0)$, $(X_{\mathrm{I}}^{(2)}, X_{\mathrm{II}}^{(2)}) = (1, 0)$, and $(X_{\mathrm{I}}^{(3)}, X_{\mathrm{II}}^{(3)}) = (0, 1)$ are substituted into it.

2.1.5. In Example 2.1.2 an experimental technique in widespread use in the measurement of the planar homogeneous motion of the deformable object was described and a system of equations was set and solved to determine the time-dependent parameters appearing in the equations describing the planar homogeneous motion. Consider the same problem when the problem is not planar, but three-dimensional. How many markers are necessary in three dimensions and how must the markers be arranged so that they provide the information necessary to determine the time-dependent parameters appearing in the equations describing the three-dimensional homogeneous motion? Explain the process.

2.2 Rates of Change and the Spatial Representation of Motion

The velocity $\mathbf{v}$ and acceleration $\mathbf{a}$ of a particle $\mathbf{X}$ are defined by

$$\mathbf{v} = \dot{\mathbf{x}} = \left.\frac{\partial \chi}{\partial t}\right|_{\mathbf{X}\text{ fixed}}, \quad \mathbf{a} = \ddot{\mathbf{x}} = \left.\frac{\partial^2 \chi}{\partial t^2}\right|_{\mathbf{X}\text{ fixed}}, \tag{2.24}$$

where $\mathbf{X}$ is held fixed because it is the velocity or acceleration of that particular particle that is being determined. The *spatial description of motion* (as opposed to the *material description of motion* represented by (2.2)) is obtained by solving (2.2) for $\mathbf{X}$,

$$\mathbf{X} = \chi^{-1}(\mathbf{x},\ t) \quad \text{for all } \mathbf{X} \subset O(0) \tag{2.25}$$

and substituting the result into the first of the expressions (2.24) for the velocity; thus $\mathbf{v} = \dot{\mathbf{x}} = \dot{\chi}(\mathbf{X}, t)$ becomes

$$\mathbf{v} = \dot{\mathbf{x}} = \dot{\chi}(\chi^{-1}(\mathbf{x}, t), t) = \mathbf{v}(\mathbf{x}, t) \tag{2.26a}$$

or

$$\mathbf{v}(\mathbf{x}, t) = \mathbf{v}({\chi_1}^{-1}(\mathbf{x}, t), {\chi_2}^{-1}(\mathbf{x}, t), {\chi_3}^{-1}(\mathbf{x}, t), t), \tag{2.26b}$$

which emphasizes that the time dependence of the spatial representation of velocity is both explicit and implicit. This representation of the velocity with the places $\mathbf{x}$ as independent variables is called the *spatial representation of motion*. A quantity is said to be in the spatial representation if its independent variables are the places $\mathbf{x}$ and not the particles $\mathbf{X}$. In the material representation the independent variables are the particles $\mathbf{X}$; compare the material description of motion, (2.2), with (2.26). The *material time derivative* is the time derivative following the material particle $\mathbf{X}$; it is denoted by a superposed dot or D/Dt and it is defined as the partial derivative with respect to time with $\mathbf{X}$ held constant. The material time derivative is easy to calculate in the material representation. It is more complicated to calculate in the spatial representation. To determine the acceleration in the spatial representation we must calculate the material time rate of the spatial representation of velocity (2.26). The notation D/Dt introduced above is illustrated using the definitions of (2.24):

$$\mathbf{a} = \left.\frac{\partial^2 \chi}{\partial t^2}\right|_{\mathbf{X}\ \text{fixed}} = \left.\frac{\partial \mathbf{v}}{\partial t}\right|_{\mathbf{X}\ \text{fixed}} \equiv \frac{D\mathbf{v}}{Dt}. \tag{2.27}$$

A formula for $D\mathbf{v}/Dt$ is obtained by observing the explicit and implicit time dependence of the spatial representation of velocity (2.26b) and noting that the time derivative associated with the implicit dependencies may be obtained using the chain rule, thus

$$\frac{D\mathbf{v}}{Dt} = \left.\frac{\partial \mathbf{v}}{\partial t}\right|_{\mathbf{x}\ \text{fixed}} + \left.\frac{\partial \mathbf{v}}{\partial x_i}\frac{\partial x_i}{\partial t}\right|_{\mathbf{X}\ \text{fixed}} = \left.\frac{\partial \mathbf{v}}{\partial t}\right|_{\mathbf{x}\ \text{fixed}} + \frac{\partial \mathbf{v}}{\partial x_i} v_i, \tag{2.28}$$

a result that may be written more simply as

$$\frac{D\mathbf{v}}{Dt} = \left.\frac{\partial \mathbf{v}}{\partial t}\right|_{\mathbf{x}\ \text{fixed}} + \mathbf{v} \cdot \nabla \mathbf{v}. \tag{2.29}$$

The time rate computed by holding the places **x** fixed is called the local time rate. In general, the material time rate is related to the local time rate by the following operator expression that follows from (2.29),

$$\frac{D}{Dt} = \left.\frac{\partial}{\partial t}\right|_{\mathbf{x}\ \text{fixed}} + \mathbf{v}\cdot\nabla, \tag{2.30}$$

where D/Dt is the material time rate of change, $\partial/\partial t$ is the local time rate of change and $\mathbf{v}\cdot\nabla$ determines the convective change of the quantity.

The second order tensor formed by taking the spatial gradient of the velocity field $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ is called the tensor of velocity gradients and is denoted by **L**, thus

$$\mathbf{L} = [\nabla \otimes \mathbf{v}]^{\mathrm{T}} = \left[\frac{\partial v_i}{\partial x_j}\right] = \begin{bmatrix} \dfrac{\partial v_1}{\partial x_1} & \dfrac{\partial v_1}{\partial x_2} & \dfrac{\partial v_1}{\partial x_3} \\ \dfrac{\partial v_2}{\partial x_1} & \dfrac{\partial v_2}{\partial x_2} & \dfrac{\partial v_2}{\partial x_3} \\ \dfrac{\partial v_3}{\partial x_1} & \dfrac{\partial v_3}{\partial x_2} & \dfrac{\partial v_3}{\partial x_3} \end{bmatrix}. \tag{2.31}$$

L is decomposed into a symmetric part **D** called *the rate-of-deformation tensor*, and a skew symmetric part **W** called the *spin tensor*, thus

$$\mathbf{L} = \mathbf{D} + \mathbf{W}, \quad \mathbf{D} = (1/2)(\mathbf{L} + \mathbf{L}^{\mathrm{T}}), \quad \mathbf{W} = (1/2)(\mathbf{L} - \mathbf{L}^{\mathrm{T}}). \tag{2.32}$$

The three nonzero components of **W** can be formed into an axial vector $(1/2)(\nabla \times \mathbf{v})$ which represents the local rotational motion and is called the angular velocity or one-half the vorticity.

The rate-of-deformation tensor **D** defined by the second of (2.32) has the component representation

$$\mathbf{D} = \frac{1}{2}\left[\frac{\partial v_j}{\partial x_i} + \frac{\partial v_i}{\partial x_j}\right] = \frac{1}{2}\begin{bmatrix} 2\dfrac{\partial v_1}{\partial x_1} & \dfrac{\partial v_1}{\partial x_2} + \dfrac{\partial v_2}{\partial x_1} & \dfrac{\partial v_1}{\partial x_3} + \dfrac{\partial v_3}{\partial x_1} \\ \dfrac{\partial v_1}{\partial x_2} + \dfrac{\partial v_2}{\partial x_1} & 2\dfrac{\partial v_2}{\partial x_2} & \dfrac{\partial v_2}{\partial x_3} + \dfrac{\partial v_3}{\partial x_2} \\ \dfrac{\partial v_1}{\partial x_3} + \dfrac{\partial v_3}{\partial x_1} & \dfrac{\partial v_2}{\partial x_3} + \dfrac{\partial v_3}{\partial x_2} & 2\dfrac{\partial v_3}{\partial x_3} \end{bmatrix}. \tag{2.33}$$

The components of **D** along the diagonal are called normal rates of deformation and the components off the diagonal are called shear rates of deformation. The normal rates of deformation, D_{11}, D_{22}, and D_{33}, are measures of instantaneous time rate of change of the material filament instantaneously coincident with the 1, 2, and 3 axes, respectively, and the shear rates of deformation, D_{23}, D_{13}, and D_{12}, are equal to one-half the time rate of decrease in an originally right angle between material

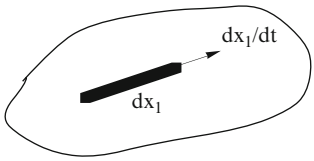


Fig. 2.5 An illustration for the geometric interpretation of the D_{11} component of the rate-of-deformation tensor **D**. A vector of infinitesimal length representing the present position of an infinitesimal material filament coinciding with the x_1 at time t is denoted by dx_1. The instantaneous time rate of change of the material filament instantaneously coincident with dx_1 is $dv_1 = D_{11}dx_1$. The expression $dv_1 = D_{11}dx_1$ shows dv_1 as a linear function of dx_1 at any point **x** and time t. Thus the geometric interpretation of $D_{11} = (d\dot{x}_1/dx_1)$ is that it is the instantaneous time rate of change of dx_1 at time t relative to dx_1 at time t

filaments instantaneously situate upon the 2 and 3 axes, the 1 and 3 axes, and the 1 and 2 axes, respectively.

The rate-of-deformation tensor **D** represents instantaneous rates of change, that is to say how much a quantity is changing compared to its present size. Let dx_1 be a vector of infinitesimal length representing the present position of an infinitesimal material filament coinciding with the x_1 at time t, Fig. 2.5. The instantaneous time rate of change of the material filament instantaneously coincident with dx_1 is $dv_1 = D_{11}dx_1$, a result that follows from the entry in the first column and first row of (2.33). The expression $dv_1 = D_{11}dx_1$ shows dv_1 as a linear function of dx_1 at any point **x** and time t. Thus the geometric interpretation of $D_{11} = (d\dot{x}_1/dx_1)$ is that it is the instantaneous time rate of change of dx_1 at time t relative to dx_1 at time t. Similar geometric interpretations exist for D_{22} and D_{33}.

The geometric interpretation of the normal rate of shearing components D_{11}, D_{22}, and D_{33} is easily extended to obtain a geometric interpretation of the trace of **D** which is also the divergence of the velocity, tr $\mathbf{D} = \nabla \cdot \mathbf{v}$. If dv represents an element of volume in the spatial coordinate system, $dv = dx_1dx_2dx_3$ (Fig. 2.6), the material time rate of change of dv can be computed using the type of formula developed in the previous paragraph; $d\dot{x}_1 = D_{11}dx_1$, $d\dot{x}_2 = D_{22}dx_2$, and $d\dot{x}_3 = D_{33}dx_3$, thus

$$d\dot{v} = \frac{D}{Dt}(dx_1dx_2dx_3) = (D_{11} + D_{22} + D_{33})dx_1dx_2dx_3 = (\text{tr } \mathbf{D})dv \tag{2.34}$$

or, noting from the definition of **D** that

$$\text{tr}\mathbf{D} = D_{11} + D_{22} + D_{33} = \nabla \cdot \mathbf{v} = \text{div } \mathbf{v} \tag{2.35}$$

it follows that

$$\text{tr}\mathbf{D} = \nabla \cdot \mathbf{v} = \frac{d\dot{v}}{dv}. \tag{2.36}$$

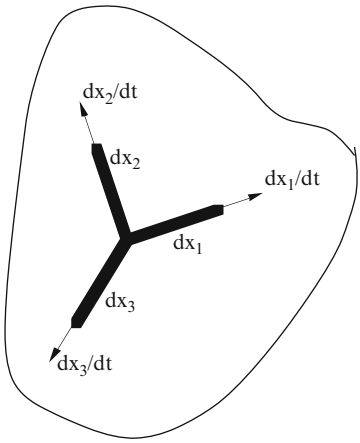


Fig. 2.6 Illustration for the geometric interpretation of the trace of the rate-of-deformation tensor **D** as the instantaneous time rate of change of volume. The material time rate of change of an element of volume in the spatial coordinate system, $dv = dx_1 dx_2 dx_3$, is shown to be $d\dot{v} = \mathrm{tr}\mathbf{D}dv = \nabla \cdot \mathbf{u}dv$, thus the $\nabla \cdot \mathbf{v}$ or tr**D** has the geometric interpretation as the instantaneous time rate of change of material volume

Thus the $\nabla \cdot \mathbf{v}$ or tr**D** have the geometric interpretation as the instantaneous time rate of change of material volume. Another way of viewing this result is to say that the divergence of the velocity field is the time rate of change of a material volume relative to how large it is at the instant (2.35).

The off-diagonal components of the rate-of-deformation tensor, for example D_{12}, represent rates of shearing. D_{12} is equal to one-half the time rate of decrease in an originally right angle between the filaments d**x**(1) and d**x**(2), Fig. 2.7. To see this, note that the dot product of material filaments d**x**(1) and d**x**(2) axes may be written as

$$\mathbf{dx}(1) \cdot \mathbf{dx}(2) = |\mathbf{dx}(1)||\mathbf{dx}(2)| \cos\theta_{12},$$

where θ_{12} is the angle between the two filaments. In the calculation of the material time derivative of the dot product above, $\mathbf{dx}(1) \cdot \mathbf{dx}(2)$, we will employ the formula $dv_i = d\dot{x}_i = L_{ij}dx_j$ that follows from (2.31). The material time derivative of both sides of the equation above is then computed;

$$\begin{aligned}\frac{D}{Dt}(\mathbf{dx}(1) \cdot \mathbf{dx}(2)) &= \mathbf{d\dot{x}}(1) \cdot \mathbf{dx}(2) + \mathbf{dx}(1) \cdot \mathbf{d\dot{x}}(2) \\ &= L_{ij}dx_j(1)dx_2(2) + dx_i(1)L_{ij}dx_j(2) = 2D_{ij}dx_i(1)dx_j(2) \\ &= |\mathbf{d\dot{x}}(1)||\mathbf{dx}(2)| \cos\theta_{12} + |\mathbf{dx}(1)||\mathbf{d\dot{x}}(2)| \cos\theta_{12} \\ &\quad - \dot{\theta}_{12}|\mathbf{dx}(1)||\mathbf{dx}(2)| \sin\theta_{12}\end{aligned}$$

and, since we are interested in the instant that $\theta_{12} = \pi/2$, it follows that

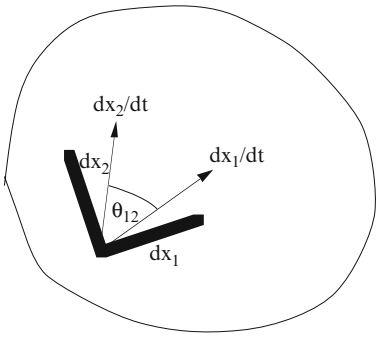


Fig. 2.7 An illustration for the geometric interpretation of the rate of shearing strain component D_{12} of the rate-of-deformation tensor **D**. The *heavy black lines* represent two material filaments of infinitesimal length that are instantaneously perpendicular. The *thin black lines* represent the same two material filaments in the next instant. The instantaneous time rate of change of the angle between the two material filaments, the rate at which the two filaments are coming together or separating is $\dot{\theta}_{12}$. The geometric interpretation of D_{12} is that it is one half the instantaneous time of decrease in an originally right angle between dx_1 and dx_2, $D_{12} = -\dot{\theta}_{12}/2$

$$2D_{ij}dx_i(1)dx_j(2) = -\dot{\theta}_{12}|\mathbf{dx}(1)|\,|\mathbf{dx}(2)|.$$

Finally, if we take $\mathbf{dx}(1) = dx_1\mathbf{e}_1$ and $\mathbf{dx}(2) = dx_2\mathbf{e}_2$ it may be concluded that

$$D_{12} = -\frac{\dot{\theta}_{12}}{2},$$

confirming that D_{12} is equal to one-half the material time rate of decrease in an originally right angle between dx_1 and dx_2. The geometric interpretations of D_{13} and D_{23} are similar. These geometric interpretations of the components of **D** as the instantaneous time rate of change of filaments, angles, and volume are the rationale for calling **D** the *rate of deformation* tensor.

Example 2.2.1
Calculate the velocity and acceleration in the material representation of the motion (2.12) of Example 2.1.1, then determine the spatial representation. Verify that the acceleration computed in the spatial representation is the same as the acceleration computed in the material representation. Calculate the tensor of velocity gradients **L**, the rate of deformation tensor **D**, and the spin tensor **W** for this motion.

Solution: The velocity and acceleration for this motion are given by (2.24) as

$$\dot{x}_1 = X_{\mathrm{I}} + X_{\mathrm{II}} + 3, \quad \dot{x}_2 = X_{\mathrm{I}} + X_{\mathrm{II}} + 2, \quad \dot{x}_3 = 0, \quad \ddot{x}_1 = \ddot{x}_2 = \ddot{x}_3 = 0.$$

In order to find the spatial representation for this motion we must invert the system of equations (2.12) representing the motion, thus

$$X_{\mathrm{I}} = \frac{1}{1+2t}\{(1+t)x_1 - tx_2 - t^2 - 3t\},$$
$$X_{\mathrm{II}} = \frac{1}{1+2t}\{-tx_1 + (1+t)x_2 + t^2 - 2t\}, \quad X_{\mathrm{III}} = x_3$$

then, substituting these expressions into the previous equations for the velocities, the spatial representation of this motion is obtained:

$$v_1 = \frac{1}{1+2t}\{x_1 + x_2 + 3 + t\}, \quad v_2 = \frac{1}{1+2t}\{x_1 + x_2 + 2 - t\}, \quad v_3 = 0.$$

It is known from the first calculation in this example that this motion is one of zero acceleration, $\ddot{x}_1 = \ddot{x}_2 = \ddot{x}_3 = 0$. This may be verified by calculating the acceleration of the spatial representation of the motion above using the material time derivative (2.29), thus

$$\begin{aligned} a_1 &= \frac{\partial v_1}{\partial t} + v_1\frac{\partial v_1}{\partial x_1} + v_2\frac{\partial v_1}{\partial x_2} + v_3\frac{\partial v_1}{\partial x_3} \\ &= \frac{-2}{(1+2t)^2}\{x_1 + x_2 + 3 + t\} + \frac{1}{1+2t} + \frac{1}{(1+2t)^2}\{2(x_1 + x_2) + 5\} = 0, \end{aligned}$$

$$\begin{aligned} a_2 &= \frac{\partial v_2}{\partial t} + v_1\frac{\partial v_2}{\partial x_1} + v_2\frac{\partial v_2}{\partial x_2} + v_3\frac{\partial v_2}{\partial x_3} \\ &= \frac{-2}{(1+2t)^2}\{x_1 + x_2 + 2 - t\} - \frac{1}{1+2t} + \frac{1}{(1+2t)^2}\{2(x_1 + x_2) + 5\} = 0. \end{aligned}$$

The tensor of velocity gradients $\mathbf{L}$ for the motion (2.12) is obtained by substituting the spatial representation for the motion obtained above into (2.31); thus

$$\mathbf{L} = \frac{1}{1+2t}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The rate-of-deformation tensor $\mathbf{D}$ for this motion is equal to $\mathbf{L}$. The spin tensor $\mathbf{W}$ is zero for this motion.

Problems

2.2.1. For the first six motions of the form (2.10) given in Problem 2.1.1, namely 2.1.1(a) through 2.1.1(f), determine the velocity and acceleration in the material (Lagrangian) representation, the velocity and acceleration in the spatial (Eulerian) representation, and the three tensors $\mathbf{L}$, $\mathbf{D}$, and $\mathbf{W}$. Discuss briefly how these algebraic calculations relate to the geometry of the motion.

2.2.2. The motion of a continuum is given by: $x_1 = X_I + X_{II}t + X_{III}t^2$, $x_2 = X_{II} + X_{III}t + X_I t^2$, $x_3 = X_{III} + X_I t + X_{II}t^2$.

(a) Find the inversion of this motion.
(b) Determine the velocity and the acceleration in the material representation.
(c) Find the velocity in the spatial (Eulerian) representation for this motion.
(d) Find the three tensors **L**, **D**, and **W** for this motion.
(e) Find the tensor of deformation gradients **F** for this motion.

2.2.3. The motion of a continuum is given by:

$$x_1 = X_I + X_{II}\sin(\pi t), \quad x_2 = X_{II} - X_I \sin(\pi t), \quad x_3 = X_{III}.$$

(a) Determine the deformation gradient **F** of this deformation.
(b) Determine the instantaneous configuration image of the set of points $(X_I)^2 + (X_{II})^2 = 1$ in the reference configuration.
(c) Describe the geometry of the set of points $(X_I)^2 + (X_{II})^2 = 1$ in the reference configuration and describe what happens to this set of points in the motion of the continuum as time t increases.

2.3 Infinitesimal Motions

The term infinitesimal motion is used to describe the case when the deformation, including rotation, is small. This does not mean that the displacement vector is small; one can have large displacements but small strain infinitesimal motions. Large displacements associated with small strain infinitesimal motions occur in very thin long rods. The criterion for infinitesimal motion is that the square of the gradients of displacement be small compared to the gradients of displacement themselves. Thus, for infinitesimal motions, the squares and products of the nine quantities

$$\frac{\partial u_1}{\partial x_1}, \frac{\partial u_1}{\partial x_2}, \frac{\partial u_1}{\partial x_3}, \frac{\partial u_2}{\partial x_1}, \frac{\partial u_2}{\partial x_2}, \frac{\partial u_2}{\partial x_3}, \frac{\partial u_3}{\partial x_1}, \frac{\partial u_3}{\partial x_2}, \frac{\partial u_3}{\partial x_3} \tag{2.37}$$

must be small compared to their own values. This means, for example, $\{\partial u_2/\partial x_1\}^2$ is required to be much smaller than $\partial u_2/\partial x_1$; each such square and product of these nine quantities is so small that it may be neglected compared to the quantity itself. Using this criterion of smallness, representations of the kinematics variables for infinitesimal motions will be developed in this section.

If the motion is infinitesimal the deformation gradient tensor **F** must not deviate significantly from the unit tensor **1**, the magnitude of the deviation being restricted by the criterion on the deformation gradients stated in the previous paragraph.

The deformation gradient $\mathbf{F}$ may be expressed, using (2.22), in terms of $[\nabla_{\mathbf{O}} \otimes \mathbf{u}]^{\mathrm{T}}$, which is a matrix of components $[\partial u_i/\partial X_\alpha]$, as

$$\mathbf{F} = \mathbf{1} + [\nabla_{\mathbf{O}} \otimes \mathbf{u}]^{\mathrm{T}}. \tag{2.38}$$

Since

$$\nabla_{\mathbf{O}} = \mathbf{F}^{\mathrm{T}} \cdot \nabla \quad \text{or} \quad \frac{\partial}{\partial X_\alpha} = \frac{\partial}{\partial x_j}\frac{\partial x_j}{\partial X_\alpha} = \frac{\partial}{\partial x_j} F_{j\alpha} \tag{2.39}$$

it follows from (2.38) and a result obtained in Appendix A, namely that the transpose of a product of matrices is equal to the product of the transposed matrices in reverse order, $[\mathbf{AB}]^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$, that

$$\mathbf{F} = \mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} \cdot \mathbf{F}. \tag{2.40}$$

This result may be used as a recursion formula for $\mathbf{F}$. In that role this formula for F can be substituted into itself once,

$$\mathbf{F} = \mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} \cdot [\nabla \otimes \mathbf{u}]^{\mathrm{T}} \cdot \mathbf{F} \tag{2.41}$$

and then again and again,

$$\mathbf{F} = \mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} \cdot [\nabla \otimes \mathbf{u}]^{\mathrm{T}} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} \cdot [\nabla \otimes \mathbf{u}]^{\mathrm{T}} \cdot [\nabla \otimes \mathbf{u}]^{\mathrm{T}} + \text{h.o.t.}, \tag{2.42}$$

where h.o.t. stands for "higher order terms." If the terms of second order according to the criterion (2.37) are neglected, the $\mathbf{F}$ is approximated by

$$\mathbf{F} \approx \mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}}. \tag{2.43}$$

From (2.22) it is known that

$$\mathbf{F}^{-1} = \mathbf{1} - [\nabla \otimes \mathbf{u}]^{\mathrm{T}}, \tag{2.44}$$

a formula that is accurate in the approximation because $\mathbf{F}\cdot\mathbf{F}^{-1} = \mathbf{1}$ when terms of second order are neglected.

Two important conclusions may be made from this result. First, for infinitesimal motions the difference between the use of material and spatial coordinates is insignificant, thus $\mathbf{X}$ and $\mathbf{x}$ are equivalent as are the gradient operators $\nabla_{\mathbf{O}}$ and ∇. Concerning these operators note from (2.39) that

$$\nabla_{\mathbf{O}} \otimes \mathbf{u} = \mathbf{F}^{\mathrm{T}} \cdot [\nabla \otimes \mathbf{u}] \tag{2.45}$$

and, substituting for $\mathbf{F}$ using (2.42),

$$\nabla_{\mathbf{O}} \otimes \mathbf{u} = \nabla \otimes \mathbf{u} + [\nabla \otimes \mathbf{u}] \cdot [\nabla \otimes \mathbf{u}]$$

thus, to neglect terms of second order,

$$\nabla_{\mathbf{O}} \otimes \mathbf{u} \approx \nabla \otimes \mathbf{u}. \tag{2.46}$$

For infinitesimal motions, the movement of boundaries due to motion is neglected because the small movement is equivalent to the difference in the use of material and spatial coordinates, which is insignificant. Therefore in all the following considerations of infinitesimal motions the coordinates $\mathbf{x}$ will be used without reference to their material or spatial character, because the result is correct independent of their character. The second important conclusion is that, for infinitesimal motions, $\mathbf{F}$ has the representation

$$\mathbf{F} = \mathbf{1} + [\nabla \otimes \mathbf{u}(\mathbf{x}, t)]^{\mathrm{T}}. \tag{2.47}$$

In the special case when the infinitesimal motion is a rigid object rotation, $\mathbf{F} = \mathbf{Q}$ and $\mathbf{Q} = \mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}}$. The requirement that $\mathbf{Q}$ be orthogonal, $\mathbf{Q}^{\mathrm{T}} \cdot \mathbf{Q} = \mathbf{Q} \cdot \mathbf{Q}^{\mathrm{T}} = \mathbf{1}$, $\mathbf{Q} \cdot \mathbf{Q}^{\mathrm{T}} = (\mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}}) \cdot (\mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}})^{\mathrm{T}} = \mathbf{1} + [\nabla \otimes \mathbf{u}]^{\mathrm{T}} + (\nabla \otimes \mathbf{u}) + (\nabla \otimes \mathbf{u})^{\mathrm{T}} \cdot (\nabla \otimes \mathbf{u}) = \mathbf{1}$, means that

$$(\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u} = 0, \tag{2.48}$$

since $(\nabla \otimes \mathbf{u})^{\mathrm{T}} \cdot (\nabla \otimes \mathbf{u})$ represents terms of the second order terms that are neglected. Defining the symmetric and skew symmetric parts of $\nabla \otimes \mathbf{u}$ as $\mathbf{E}$ and $\mathbf{Y}$,

$$\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u}), \quad \mathbf{Y} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} - \nabla \otimes \mathbf{u}), \tag{2.49}$$

it is seen from (2.48) that $\mathbf{E}$ must be zero when the infinitesimal motion is a rigid object rotation. It may also be seen that the orthogonal rotation $\mathbf{Q}$ characterizing the infinitesimal rigid object rotation is given by

$$\mathbf{Q} = \mathbf{1} + \mathbf{Y}, \tag{2.50}$$

where $\mathbf{Y}$, defined by (2.49), is skew symmetric, $\mathbf{Y} = -\mathbf{Y}^{\mathrm{T}}$, and $\mathbf{Y}\mathbf{Y}^{\mathrm{T}}$ is a second order term, since it is a square of the coefficients (2.37) which are neglected compared to the values of $\mathbf{Y}$.

Returning to the total infinitesimal motion, the definitions (2.49) of $\mathbf{E}$ and $\mathbf{Y}$ may be used to rewrite (2.42) as

$$\mathbf{F} = \mathbf{1} + \mathbf{E} + \mathbf{Y}. \tag{2.51}$$

It has been established that $\mathbf{F}$ represents only the rotational and deformational motion because the translational portion of the motion, being independent of the

coordinates, was removed by spatial or material differentiation. It has also been noted that the special case of $\mathbf{F} = \mathbf{1}$ corresponds to no rotational and no deformational motion. Further it has been shown (2.50) that $\mathbf{Y}$ is associated with pure rigid object rotation. This means that $\mathbf{E}$ must be the tensor representing the deformation. This is indeed the case, as will be shown below. $\mathbf{E}$ is called the *infinitesimal strain tensor* and $\mathbf{Y}$ is called the *infinitesimal rotation tensor*. The representation (2.51) for the tensor of deformation gradients then demonstrates that, for infinitesimal motions, $\mathbf{F} - \mathbf{1}$ may be decomposed into the sum of two terms, $\mathbf{E}$ and $\mathbf{Y}$, which represent the deformational and rigid rotational characteristics of the infinitesimal motion, respectively.

The strain tensor $\mathbf{E}$, defined by the first of (2.49), has the component representation

$$\mathbf{E} = \frac{1}{2}\left[\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right] = \frac{1}{2}\begin{bmatrix} 2\dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} + \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_1}{\partial x_3} + \dfrac{\partial u_3}{\partial x_1} \\ \dfrac{\partial u_1}{\partial x_2} + \dfrac{\partial u_2}{\partial x_1} & 2\dfrac{\partial u_2}{\partial x_2} & \dfrac{\partial u_2}{\partial x_3} + \dfrac{\partial u_3}{\partial x_2} \\ \dfrac{\partial u_1}{\partial x_3} + \dfrac{\partial u_3}{\partial x_1} & \dfrac{\partial u_2}{\partial x_3} + \dfrac{\partial u_3}{\partial x_2} & 2\dfrac{\partial u_3}{\partial x_3} \end{bmatrix}. \tag{2.52}$$

The components of $\mathbf{E}$ along the diagonal are called normal strains and the components off the diagonal are called shear strains. The normal strains, E_{11}, E_{22}, and E_{33}, are measures of change in length per unit length along the 1, 2 and 3 axes, respectively, and the shear strains, E_{23}, E_{13}, and E_{12}, are one-half of the changes in the angle between the 2 and 3 axes, the 1 and 3 axes and the 1 and 2 axes, respectively.

The geometric interpretation of the components of the strain tensor $\mathbf{E}$ stated in the previous paragraph will be analytically developed here. Let dx_1 be a vector of infinitesimal length representing the present position of an infinitesimal material filament coinciding with the x_1 at time t. The displacement of this material filament instantaneously coincident with dx_1 is $du_1 = E_{11}dx_1$, a result that follows from the entry in the first column and first row of (2.52). The expression $du_1 = E_{11}dx_1$ is the change in length of dx_1 as a consequence of the strain as illustrated in Fig. 2.8. Thus the geometric interpretation of $E_{11} = (du_1/dx_1)$ is that it is the change in length per unit length of dx_1. Similar geometric interpretations exist for E_{22} and E_{33}.

The geometric interpretation of the normal strain components E_{11}, E_{22} and E_{33} is easily extended to obtain a geometric interpretation of the trace of the small strain tensor tr $\mathbf{E}$, or equivalently the divergence of the displacement field $\nabla\cdot\mathbf{u}$, tr $\mathbf{E} = \nabla\cdot\mathbf{u}$. If $dv_o = dx_1dx_2dx_3$ represents an undeformed element of volume (Fig. 2.9), the deformed volume is given by $dv = (dx_1 + du_1)(dx_2 + du_2)(dx_3 + du_3)$. Using $du_1 = E_{11}dx_1$, $du_2 = E_{22}dx_2$ and $du_3 = E_{33}dx_3$, the deformed volume is given by $dv = (1 + E_{11})(1 + E_{22})(1 + E_{33})dv_o$. Expanding $dv = (1 + E_{11})(1 + E_{22})(1 + E_{33})dv_o$ and recognizing that the squares of displacement gradients (2.37)

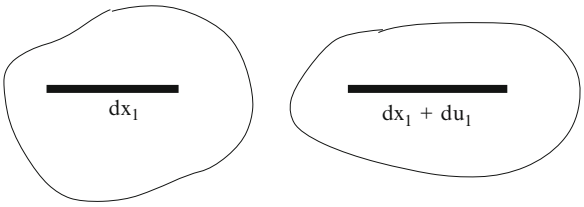


Fig. 2.8 An illustration for the geometric interpretation of the normal strain component E_{11}. The *left* and *right* illustrations of this figure represent the undeformed and deformed configurations, respectively. The *heavy black line* represents the same material filament in the two configurations. E_{11} is equal to the change in length per unit length of the filament between the two configurations. The original length is dx_1 and the change in length due to the deformation is du_1, thus $E_{11} = du_1/dx_1$

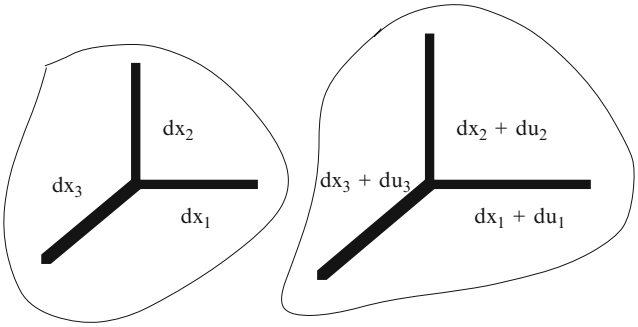


Fig. 2.9 An illustration for the geometric interpretation of the trace of the strain tensor, tr $\mathbf{E}$, or the divergence of the displacement field, $\nabla\cdot\mathbf{u}$, tr$\mathbf{E} = \nabla\cdot\mathbf{u}$. The *left* and *right* illustrations of this figure represent the undeformed and deformed configurations, respectively. The *heavy black lines* represent the same material filaments in the two configurations. The volume element in the undeformed configuration is $dv_o = dx_1dx_2dx_3$ and the deformed volume is given by $dv = (dx_1 + du_1)(dx_2 + du_2)(dx_3 + du_3)$. It may be shown (see text) that $dv = (1 + \text{tr}\mathbf{E})dv_o$. Thus the tr$\mathbf{E} = \nabla\cdot\mathbf{u}$ represents the change in volume per unit volume, $(dv - dv_o)/dv_o$

may be neglected, it follows that $dv = (1 + \text{tr}\mathbf{E})dv_o$. Thus the tr$\mathbf{E}$ represents the change in volume per unit volume, $(dv - dv_o)/dv_o$.

The off-diagonal components of the strain tensor, for example E_{12}, represent the shearing strains. E_{12} is equal to one-half the change in angle that was originally a right angle between the x_1 and x_2 axes. To construct this geometric result algebraically, the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are considered, see Fig. 2.10. After deformation these vectors are $\mathbf{Fe}_1$ and $\mathbf{Fe}_2$, respectively, or, since $\mathbf{F} = \mathbf{1} + \mathbf{E}$, the deformed vectors are given by $\mathbf{e}_1 + \mathbf{Ee}_1$ and $\mathbf{e}_2 + \mathbf{Ee}_2$, respectively. The dot product of the vectors $\mathbf{e}_1 + \mathbf{Ee}_1$ and $\mathbf{e}_2 + \mathbf{Ee}_2$ is $(\mathbf{e}_1 + \mathbf{Ee}_1)\cdot(\mathbf{e}_2 + \mathbf{Ee}_2) = \mathbf{e}_1\cdot\mathbf{e}_2 + \mathbf{e}_1\cdot\mathbf{Ee}_2 + \mathbf{e}_2\cdot\mathbf{Ee}_1 + \mathbf{Ee}_1\cdot\mathbf{Ee}_2$, but since the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal, $\mathbf{e}_1\cdot\mathbf{e}_2 = 0$, and also since $\mathbf{Ee}_1\cdot\mathbf{Ee}_2$ is a higher order term because it contains the squares of the displacement gradients (2.37), this expression reduces to $(\mathbf{e}_1 + \mathbf{Ee}_1)\cdot(\mathbf{e}_2 + \mathbf{Ee}_2) = \mathbf{e}_1\cdot\mathbf{Ee}_2 + \mathbf{e}_2\cdot\mathbf{Ee}_1$. This result is further reduced by noting that $\mathbf{e}_1\cdot\mathbf{Ee}_2 = \mathbf{e}_2\cdot\mathbf{Ee}_1 = E_{12}$,

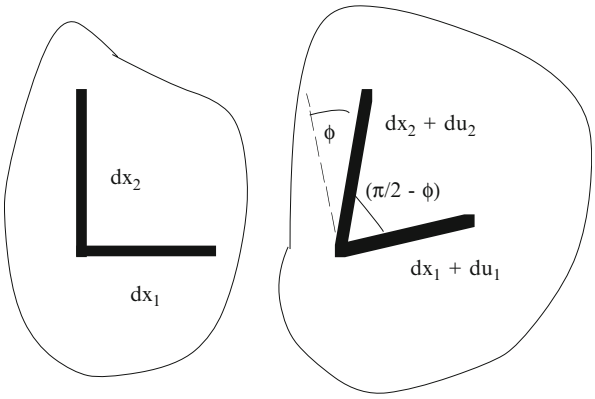


Fig. 2.10 Illustration for the geometric interpretation of the shearing strain E_{12}. The *left* and *right* illustrations of this figure represent the undeformed and deformed configurations, respectively. The *heavy black lines* represent the same material filaments in the two configurations. E_{12} is equal to one-half the change in angle that was originally a right angle between the x_1 and x_2 axes, $\phi/2$ in this figure

thus it follows that $(\mathbf{e}_1 + \mathbf{E}\mathbf{e}_1)\cdot(\mathbf{e}_2 + \mathbf{E}\mathbf{e}_2) = 2E_{12}$. Recalling the formula (A61) for the dot product of two vectors, say $\mathbf{u}$ and $\mathbf{v}$, as equal to the magnitude of the first times the magnitude of the second times the cosine of the angle (say ζ) between them, $\mathbf{u} \cdot \mathbf{v} = u_i v_i = |\mathbf{u}| \cdot |\mathbf{v}| \cos\zeta$, it follows that $2E_{12} = |\mathbf{e}_1 + \mathbf{E}\mathbf{e}_1|\,|\mathbf{e}_2 + \mathbf{E}\mathbf{e}_2| \cos(\pi/2 - \phi)$, where the angle $(\pi/2 - \phi)$ is illustrated in Fig. 2.10. The magnitude of $|\mathbf{e}_1 + \mathbf{E}\mathbf{e}_1|$ is the square root of $(\mathbf{e}_1 + \mathbf{E}\mathbf{e}_1)\cdot(\mathbf{e}_1 + \mathbf{E}\mathbf{e}_1) = \mathbf{e}_1\cdot\mathbf{e}_1 + \mathbf{e}_1\cdot\mathbf{E}\mathbf{e}_1 + \mathbf{e}_1\cdot\mathbf{E}\mathbf{e}_1 + \mathbf{E}\mathbf{e}_1\cdot\mathbf{E}\mathbf{e}_1$, but since $\mathbf{e}_1\cdot\mathbf{e}_1 = 1$ and $\mathbf{E}\mathbf{e}_1\cdot\mathbf{E}\mathbf{e}_1$ is a higher order term, this reduces to the square root of $1 + 2E_{11}$, by a parallel of the arguments used above to obtain the formula for $2E_{12}$. At this point a classical approximation is used. This approximation is that $1 + \varepsilon \approx \sqrt{(1 + 2\varepsilon)}$ if summands of the order ε^2 may be neglected; the proof of this approximation follows easily if one squares it. Then, since the square of E_{11} is a higher order term, the square root of $1 + 2E_{11}$ is given by $1 + E_{11}$, thus $2E_{12} = (1 + E_{11})(1 + E_{22})\cos(\pi/2 - \phi)$ or $2E_{12} = (1 + E_{11})(1 + E_{22})\sin\phi$ or expanding; $2E_{12} = \sin\phi + (E_{11} + E_{22})\sin\phi + E_{11}E_{22}\sin\phi$. Finally, since the angle ϕ is small, $\sin\phi$ is small as are E_{11} and E_{22}, thus the neglect of higher order terms gives $2E_{12} = \phi$, and the interpretation of E_{12} as one-half the change in an angle that was originally a right angle between the x_1 and x_2 axes (Fig. 2.10). These geometric interpretations of the components of $\mathbf{E}$ as the change in the length of filaments, the change in angles and the change in volume deformation between the undeformed and the deformed configurations are the rationale for calling $\mathbf{E}$ the *strain* tensor.

Example 2.3.1

The deformation gradient and the inverse deformation gradient for the motion given by (2.12) were computed in Example 2.1.4. Determine the restriction on the motion given by (2.12) so the motion is infinitesimal. Find the strain tensor $\mathbf{E}$ and the rotation tensor $\mathbf{Y}$ for the infinitesimal motion.

Solution: Comparison of the expressions $\nabla_{\mathbf{O}} \otimes \mathbf{u}(\mathbf{X}, t)$ and $\nabla \otimes \mathbf{u}(\mathbf{x}, t)$ obtained in Example 2.1.4 shows that these two expressions coincide only for very small times t, only if t^2 is much less than t. In this case $\nabla_{\mathbf{O}} \otimes \mathbf{u}(\mathbf{X}, t) = \nabla \otimes \mathbf{u}(\mathbf{x}, t)$ and

$$\nabla \otimes \mathbf{u} = t \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

From this expression for $\nabla \otimes \mathbf{u}$ and (2.49), the rotation tensor is determined to be $\mathbf{Y} = \mathbf{0}$, and the strain tensor $\mathbf{E}$ is given by

$$\mathbf{E} = t \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

as long as t is small.

Problem

2.3.1. For the motions of the form (2.10) given in Problem 2.1.1, namely 2.1.1(a) through 2.1.1(g), determine the conditions under which the motion remains infinitesimal and compute the infinitesimal strain and rotation tensors, $\mathbf{E}$ and $\mathbf{Y}$. Discuss briefly the significance of each of the seven strain tensors computed. In particular, explain the form or value of the strain tensor in terms of the motion.

2.4 The Strain Conditions of Compatibility

Calculating the strain tensor $\mathbf{E}$ given the displacement field $\mathbf{u}$ is a relatively simple matter; one just substitutes the displacement field $\mathbf{u}$ into the formula (2.49) for the strain displacement relations, $\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u})$. Situations occur in which it is desired to calculate the displacement field $\mathbf{u}$ given the strain tensor $\mathbf{E}$. This inverse problem is more difficult because the strain displacement relations, $\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u})$, become a system of first order partial differential equations for the displacement field $\mathbf{u}$. Given the significance of the displacement field $\mathbf{u}$ in an object we generally want to insure that the displacement field $\mathbf{u}$ is continuous and single valued. There are real situations in which the displacement field $\mathbf{u}$ might be discontinuous and multiple valued, but these situations will be treated as special cases. In general it is desired that the integral of the strain–displacement relations, the displacement field $\mathbf{u}$, is continuous and single valued. The conditions of compatibility insure this. The conditions of compatibility are equations that the strain tensor must satisfy so that when the strain–displacement relations are integrated, the resulting displacement field $\mathbf{u}$, is continuous and single valued. The conditions of compatibility may be written in the direct notation as

$$\nabla \times \mathbf{E} \times \nabla = 0 \tag{2.53}$$

or in the index notation as

$$e_{ijk} e_{pmn} \frac{\partial^2 E_{jm}}{\partial x_k \partial x_n} = 0 \tag{2.54}$$

or in scalar form as the following six equations:

$$\begin{aligned}
&\frac{\partial^2 E_{11}}{\partial x_2 \partial x_3} = \frac{\partial}{\partial x_1}\left\{-\frac{\partial E_{23}}{\partial x_1} + \frac{\partial E_{31}}{\partial x_2} + \frac{\partial E_{12}}{\partial x_3}\right\}, \quad 2\frac{\partial^2 E_{12}}{\partial x_1 \partial x_2} = \frac{\partial^2 E_{11}}{\partial x_2^2} + \frac{\partial^2 E_{22}}{\partial x_1^2}, \\
&\frac{\partial^2 E_{22}}{\partial x_3 \partial x_1} = \frac{\partial}{\partial x_2}\left\{-\frac{\partial E_{31}}{\partial x_2} + \frac{\partial E_{12}}{\partial x_3} + \frac{\partial E_{23}}{\partial x_1}\right\}, \quad 2\frac{\partial^2 E_{23}}{\partial x_2 \partial x_3} = \frac{\partial^2 E_{22}}{\partial x_3^2} + \frac{\partial^2 E_{33}}{\partial x_2^2}, \\
&\frac{\partial^2 E_{33}}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_3}\left\{-\frac{\partial E_{12}}{\partial x_3} + \frac{\partial E_{23}}{\partial x_1} + \frac{\partial E_{31}}{\partial x_2}\right\}, \quad 2\frac{\partial^2 E_{31}}{\partial x_3 \partial x_1} = \frac{\partial^2 E_{33}}{\partial x_1^2} + \frac{\partial^2 E_{11}}{\partial x_3^2}.
\end{aligned} \tag{2.55}$$

Equations (2.53) and (2.54) are symmetric second rank tensors in three dimensions and therefore have the six components given by (2.55). It follows that each of the six scalar equations (2.55) must be satisfied in order to insure compatibility. The conditions (2.53) are a direct consequence of the definition of strain, that is to say that $\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u}) = (1/2)(\mathbf{u} \otimes \nabla + \nabla \otimes \mathbf{u})$ implies that $\nabla \times \mathbf{E} \times \nabla = 0$. To see that this is true, consider the result of operating on $\mathbf{E} = (1/2)(\mathbf{u} \otimes \nabla + \nabla \otimes \mathbf{u})$ from the left by $\nabla \times$ and from the right by $\times \nabla$; one obtains the expression

$$2(\nabla \times \mathbf{E} \times \nabla) = \nabla \times \mathbf{u} \otimes \nabla \times \nabla + \nabla \times \nabla \otimes \mathbf{u} \times \nabla. \tag{2.56}$$

The operator $\nabla \times \nabla$, which occurs in both terms on the right hand side of (2.56) is called the "curl grad"; the curl of the gradient applied to a function f is zero, $\nabla \times \nabla f = 0$. In the indicial notation this is easy to see, $\nabla \times \nabla f = e_{ijk}(\partial f / \partial x_j \partial x_k)$ $\mathbf{e}_i = 0$, because of the symmetry of the indices on the partial derivatives and skew-symmetry in the components of the alternator (see Appendix A.8). Both terms on the right hand side of (2.56) contain the operator curl grad, $\nabla \times \nabla$, applied to a function, hence $\nabla \times \mathbf{E} \times \nabla = 0$. It may also be shown that the reverse is true, namely that $\nabla \times \mathbf{E} \times \nabla = 0$ implies that $\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u})$. Thus $\mathbf{E} = (1/2)$ $((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u})$ is a necessary and sufficient condition that $\nabla \times \mathbf{E} \times \nabla = 0$.

In order to both prove and motivate this result consider the two integration paths from the point P^{o} to the point P' in an object (Fig. 2.11). If the result of the integration from the point P^{o} to the point P' is to be the same along all paths chosen between these two points, then the value of the integral around any closed path in the object must be zero. This means that the integrand of the integral must be an exact differential (see Appendix A.15 Exact differentials). Recall the theorem at the

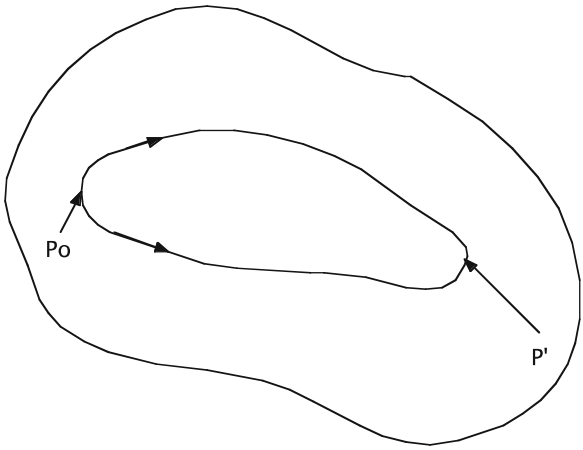


Fig. 2.11 Illustration of two integration paths from the point P^{o} to the point P' in an object. If the result of the integration from the point P^{o} to the point P' is to be the same along all paths chosen between these two points, then the value of the integral around any closed path in the object must be zero. This means that the integrand of the integral must be an exact differential

start of most texts on ordinary differential equations concerning exact differentials: If $M(x, y)$ and $N(x, y)$ are continuous functions and have continuous partial derivatives in a region of the x–y plane, then the expression $M(x, y)dx + N(x, y)dy$ is an exact differential if and only if $\partial M/\partial y = \partial N/\partial x$ throughout the region. This theorem will be applied to prove that the compatibility relations $\nabla \times \mathbf{E} \times \nabla = 0$ are both necessary and sufficient conditions for the continuous and single-valued nature of the displacement field obtained by integration from the strain–displacement relations. If the displacement vector is known at the point P^{o} then integration of $\mathrm{d}\mathbf{u}$ from the point P^{o} to the point P' (Fig. 2.11) will determine $\mathbf{u}(\mathbf{x}')$, thus,

$$\mathbf{u}(\mathbf{x}') = \mathbf{u}^{\mathrm{o}} + \int_{P^{\mathrm{o}}}^{P'} \mathrm{d}\mathbf{u} = \mathbf{u}^{\mathrm{o}} + \int_{P^{\mathrm{o}}}^{P'} (\nabla \otimes \mathbf{u})^{\mathrm{T}} \cdot \mathrm{d}\mathbf{x}. \tag{2.57}$$

Recall from (2.43) and (2.51) that

$$(\nabla \otimes \mathbf{u})^{\mathrm{T}} = \mathbf{E} + \mathbf{Y} \tag{2.58}$$

it follows that

$$\mathbf{u}(\mathbf{x}') = \mathbf{u}^{\mathrm{o}} + \int_{P^{\mathrm{o}}}^{P'} \mathbf{E} \cdot \mathrm{d}\mathbf{x} + \int_{P^{\mathrm{o}}}^{P'} \mathbf{Y} \cdot \mathrm{d}\mathbf{x}. \tag{2.59}$$

The last integral in the previous result may be rewritten as

$$\int_{P^{\mathrm{o}}}^{P'} \mathbf{Y} \cdot \mathrm{d}\mathbf{x} = \int_{P^{\mathrm{o}}}^{P'} \mathbf{Y} \cdot \mathrm{d}(\mathbf{x} - \mathbf{x}') \tag{2.60}$$

and integrated by parts, thus

$$\int_{P^o}^{P'} \mathbf{Y} \cdot d\mathbf{x} = -\mathbf{Y}^o \cdot (\mathbf{x}^o - \mathbf{x}') + \int_{P^o}^{P'} d\mathbf{x} \cdot \nabla \otimes \mathbf{Y} \cdot (\mathbf{x} - \mathbf{x}'). \tag{2.61}$$

Placing the result (2.61) into (2.59) it follows that

$$\mathbf{u}(\mathbf{x}') = \mathbf{u}^o - \mathbf{Y}^o \cdot (\mathbf{x}^o - \mathbf{x}') + \int_{P^o}^{P'} d\mathbf{x} \cdot [\mathbf{E} + \nabla \otimes \mathbf{Y} \cdot (\mathbf{x} - \mathbf{x}')] \tag{2.62}$$

or, in the indicial notation,

$$u_i(\mathbf{x}') = u_i^o - Y_{ik}^o(x_k^o - x_k') + \int_{P^o}^{P'} \left[E_{im} - \frac{\partial Y_{ik}}{\partial x_m}(x_k - x_k') \right] dx_m. \tag{2.63}$$

The relationship between the derivatives of the rotation and strain tensors,

$$\frac{\partial Y_{ik}}{\partial x_m} = \frac{\partial E_{im}}{\partial x_k} - \frac{\partial E_{mk}}{\partial x_i}, \tag{2.64}$$

may easily be verified by substituting the formulas (2.49) relating $\mathbf{E}$ and $\mathbf{Y}$ to the displacement gradients. When the relationship (2.64) is substituted into (2.63) it becomes

$$u_i(\mathbf{x}') = u_i^o - Y_{ik}^o(x_k^o - x_k') + \int_{P^o}^{P'} R_{im}\, dx_m, \tag{2.65}$$

where

$$R_{im} = E_{im} - \left\{ \frac{\partial E_{im}}{\partial x_k} - \frac{\partial E_{mk}}{\partial x_i} \right\} (x_k - x_k'). \tag{2.66}$$

The condition that the integrand in the integral in (2.65) be an exact differential is then expressed as the condition

$$\frac{\partial R_{im}}{\partial x_k} = \frac{\partial R_{ik}}{\partial x_m}. \tag{2.67}$$

When (2.67) is substituted into (2.66), the result

$$0 = \left\{ \frac{\partial^2 E_{mq}}{\partial x_i \partial x_k} + \frac{\partial^2 E_{ik}}{\partial x_q \partial x_m} - \frac{\partial^2 E_{kq}}{\partial x_i \partial x_m} - \frac{\partial^2 E_{im}}{\partial x_q \partial x_k} \right\} (x_q - x'_q) \tag{2.68}$$

is satisfied only when the compatibility conditions (2.54), or equivalently (2.54) or (2.55) or $\nabla \times \mathbf{E} \times \nabla = 0$, hold. Thus $\nabla \times \mathbf{E} \times \nabla = 0$ is a necessary and sufficient condition that the integration of the strain–displacement relations will yield a single-valued and continuous displacement field.

Problems

2.4.1. For the motions of the form (2.10) given in Problem 2.1.1, namely 2.1.1(a) through 2.1.1(g), determine if the infinitesimal strain tensors, **E**, calculated in 2.3.1 satisfy the conditions of compatibility.

2.4.2. Is the following strain state possible for an object in which the displacement field must be continuous and single valued? Justify your answer analytically.

$$\mathbf{e} = c \begin{bmatrix} x_3(x_1^2 + x_2^2) & x_1 x_2 x_3 & 0 \\ x_1 x_2 x_3 & x_3 x_2^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2.4.3. Demonstrate the validity of the formula (2.64) by substituting the formulas relating **E** and **Y** to the displacement gradients (2.49) into (2.64) and show that an identity is obtained. This is more easily done in the indicial notation.

2.4.4. Verify that substitution of the formula (2.67) into (2.66) leads to the result (2.68). This is much more easily done in the indicial notation.

Chapter 3
Continuum Formulations of Conservation Laws

The theme for this chapter is captured by a quote from Herbert Callen's book (Callen 1960) on thermodynamics. Callen introduces the conservation of energy and the concept of internal energy in the following paragraph: "The development of the principle of conservation of energy has been one of the most significant achievements in the evolution of physics. The present form of the principle was not discovered in one magnificent stroke of insight but has been slowly and laboriously developed over two and a half centuries. The first recognition of a conservation principle, by Leibnitz in 1693, referred only to the sum of the kinetic energy ($(1/2)mv^2$) and the potential energy (mgh) of a simple mechanical mass point in the terrestrial gravitational field. As additional types of systems were considered, the established form of the conservation principle repeatedly failed, but in each case it was found possible to revive it by the addition of a new mathematical term - a "new kind of energy." Thus consideration of charged systems necessitated the addition of the Coulomb interaction energy (Q_1Q_2/r) and eventually of the energy of the electromagnetic field. In 1905 Einstein extended the principle to the relativistic region, adding such terms as the relativistic rest-mass energy. In the 1930's Enrico Fermi postulated the existence of a new particle, called the neutrino, solely for the purpose of retaining the energy conservation principle in nuclear reactions. Contemporary research in nuclear physics seeks the form of interaction between nucleons within a nucleus in order that the conservation principle may be formulated explicitly at the subnuclear level. Despite the fact that unsolved problems of this type remain, the energy conservation principle is now accepted as one of the most fundamental, general, and significant principles of physical theory."

3.1 The Conservation Principles

Conservation principles will be cast here in the form of a balance or accounting statement for the time rate of change of a quantity in a system. On the plus side of the accountant's ledger are the amount of the quantity coming into the system and

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_3,

the quantity produced within the system. On the minus side are the amount of the quantity leaving the system and the quantity consumed within the system. The system will be either a fixed material system consisting always of the same set of particles or a fixed spatial (continuum) volume through which material is passing. The quantity will be either mass, linear momentum, angular momentum, or energy.

The focus of this chapter is the development of continuum formulations for the conservation principles of mass, linear momentum, angular momentum, and energy. The statement of conservation of mass is usually the statement that mass cannot be created or destroyed. The conservation of momentum is usually stated in the form of Newton's second law: the sum of the forces acting on an object is equal to the product of the mass of the object and the acceleration of the object. The conservation of angular momentum is the statement that the time rate of change of angular momentum must equal the sum of the applied moments. The conservation of energy is the requirement that the time rate of change of the sum of all the kinetic and internal energies must equal the mechanical power and heat power supplied to the object.

In the next section the continuum formulation of the conservation of mass is developed. In the following section the concept of stress is introduced and its important properties are derived and illustrated. The conservation of momentum, or the second law of Newton, when expressed in terms of stress, is called the *stress equation of motion*. The conservation of angular momentum is employed in the development of the stress equations of motion to show that the stress tensor is symmetric. In the last section the continuum formulation of the conservation of energy is developed.

3.2 The Conservation of Mass

The total mass M at time t of an object O is given by

$$M = \int_O \rho(\mathbf{x}, t)\mathrm{d}v, \tag{3.1}$$

where $\rho(\mathbf{x}, t)$ is the mass density at the place $\mathbf{x}$ within the object at the time t. The statement of mass conservation for the object O is that M does not change with time:

$$\frac{\mathrm{DM}}{\mathrm{Dt}} = \frac{\mathrm{D}}{\mathrm{Dt}} \int_O \rho(\mathbf{x}, t)\mathrm{d}v = 0. \tag{3.2}$$

The material time derivative may be interchanged with the integration over the object O since a fixed material volume is identified as the object,

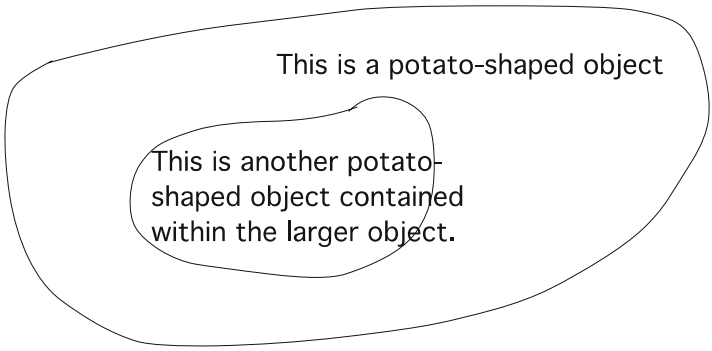


Fig. 3.1 A potato-shaped object and a second potato-shaped object that is fully contained within the first potato-shaped object. All the conservation principles may be applied to both objects separately. Furthermore one may select or define these objects as one chooses

$$\int_O (\dot{\rho}(\mathbf{x},t)\mathrm{d}v + \rho(\mathbf{x},t)\mathrm{d}\dot{v}) = 0. \tag{3.3}$$

Then, using the relationship relating the time rate of change in the volume to the present size of the volume from (2.36), $\mathrm{d}\dot{v} = (\nabla \cdot \mathbf{v})\mathrm{d}v$, it follows that

$$\int_O \{\dot{\rho}(\mathbf{x},t) + \rho(\mathbf{x},t)(\nabla \cdot \mathbf{v})\}\mathrm{d}v = 0. \tag{3.4}$$

The next step in the development of this continuum representation of the conservation of mass is to employ the argument that the integral equation (3.4) over the object O may be replaced by the condition that the integrand in the integral equation (3.4) be identically zero, thus

$$\dot{\rho} + \rho(\nabla \cdot \mathbf{v}) = 0. \tag{3.5}$$

The argument that is used to go from (3.4) and (3.5) is an argument that will be employed three more times in this chapter. The argument requires that the integrand in the integral (3.4) be continuous. The argument is that any part or subvolume of an object O may also be considered as an object and the result (3.4) also holds for that sub-object. In Fig. 3.1 an object and a portion of an object that may be considered as an object itself are illustrated. The argument for the transition (3.4) $\rightarrow$ (3.5) is as follows: suppose it is not true that the integrand of (3.4) is not zero everywhere (i.e., suppose the transition (3.4) $\rightarrow$ (3.5) is not true). If that is the case then there must exist domains of the object in which the integrand is positive and other domains in which the integrand is negative, so that when the integration is accomplished over the entire object the sum is zero. If that is the case

consider a domain of the object in which the integrand of (3.4) is always positive (or negative). Let this domain be an object. For the object chosen in this way the integral on the left-hand side of (3.4) cannot be zero. *This conclusion contradicts* (3.4) because (3.4) must be zero. It may therefore be concluded that the requirement that the integral (3.4) be zero for an object and all sub-objects that can be formed from it means that the integrand must be zero everywhere in the object. Note that a very important transition has occurred in the argument that obtains (3.5) from (3.4). Integral statements such as (3.4) are global statements because they apply to an entire object. However the requirement (3.5) is a local, pointwise condition valid at the typical point (place) in the object. Thus the transition (3.4) $\rightarrow$ (3.5) is from the global to the local or from the object to the point (or particle) in the object. Note also that the converse proof (3.5) and (3.4) is trivial.

Note that (3.5) may be combined with (2.30), the expression decomposing the material time derivative into the sum of a local rate of change and a convective rate of change, to obtain this alternate local statement of mass conservation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{3.6}$$

Another consequence of the conservation of mass is a simple formula for the material time derivative of an integral of the form

$$K = \int_O k(\mathbf{x}, t)\rho(\mathbf{x}, t)\mathrm{d}v, \tag{3.7}$$

where $k(\mathbf{x}, t)$ is a physical quantity (temperature, momentum, etc.) of arbitrary scalar, vector or tensor character and K is the value of the density times quantity $k(\mathbf{x}, t)$ integrated over the entire object O. Since, by (3.2), the material time rate of change of $\rho(\mathbf{x}, t)\mathrm{d}v$ is zero, it follows that

$$\dot{K} = \int_O \dot{k}(\mathbf{x}, t)\rho(\mathbf{x}, t)\mathrm{d}v. \tag{3.8}$$

Problems

3.2.1. For the first six motions of the form (2.10) given in Problem 2.1.1, namely 2.1.1(a) through 2.1.1 (f), determine the ratio of the time rate of change of density $\dot{\rho}$ to the instantaneous density ρ, $\frac{\dot{\rho}}{\rho}$.

3.2.2. In this section it was shown that one could derive the local statement of mass conservation (3.5) from the global statement of mass conservation, DM/Dt $= 0$. Reverse the direction of this derivation, derive the global statement of mass conservation DM/Dt $= 0$ from the fact that the local statement (3.5) or (3.6) is true at all points in an object O.

3.3 The State of Stress at a Point

Stress is a mental construct of humans to represent the internal interactions or internal forces on a material object. The concept of a stress will first be introduced as a vector. Consider the potato-shaped object shown in Fig. 3.2. An imaginary plane Σ characterized by its normal $\mathbf{n}$ divides the object into an upper portion U and a lower portion L. Considering L as a free object, the action of U upon L is statically equivalent to a resultant force $\mathbf{f}$ and couple $\mathbf{m}$. Assuming that the interaction is distributed across Σ, each area element $\Delta A_{(\mathrm{i})}$ of the intersection of Σ and the object may be considered as transmitting a force $\Delta\mathbf{f}_{(\mathrm{i})}$ and a moment $\Delta\mathbf{m}_{(\mathrm{i})}$. The average stress vector, $\mathbf{t}^{A}_{(\mathbf{n})}(P)$, at the point P acting on the plane whose normal is $\mathbf{n}$, is defined as the ratio

$$\mathbf{t}^{A}_{(\mathbf{n})}(P) = \frac{\Delta\mathbf{f}_{(\mathrm{i})}}{\Delta A_{(\mathrm{i})}}. \tag{3.9}$$

The quantity $\mathbf{t}^{A}_{(\mathbf{n})}$ is a vector because the force $\Delta\mathbf{f}_{(\mathrm{i})}$ is a vector and $\Delta A_{(\mathrm{i})}$ is a scalar. The average couple stress vector $\mathbf{c}^{A}_{(\mathbf{n})}(P)$ at the point P acting on the plane whose normal is $\mathbf{n}$ is defined analogously by the ratio

$$\mathbf{c}^{A}_{(\mathbf{n})}(P) = \frac{\Delta\mathbf{m}_{(\mathrm{i})}}{\Delta A_{(\mathrm{i})}}. \tag{3.10}$$

These two definitions are illustrated in Fig. 3.2. The stress vector $\mathbf{t}_{(\mathbf{n})}(P)$ acting at the point P on the plane whose normal is $\mathbf{n}$ is defined as the limit of the average

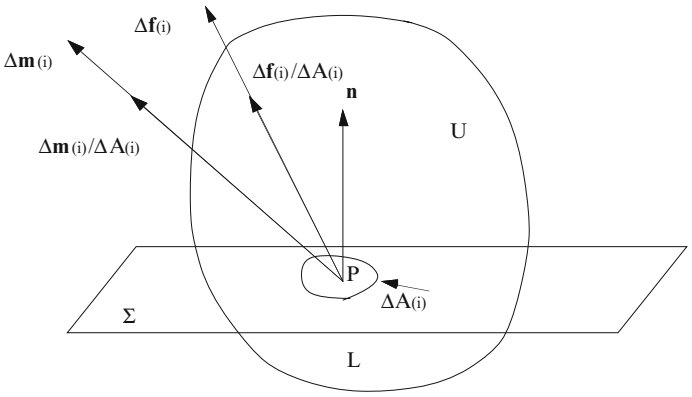


Fig. 3.2 A plane Σ, with normal $\mathbf{n}$, is shown passing through a point P in a potato-shaped object. The force and moment acting across the plane at the point P are indicated

stress vector, $\mathbf{t}^{A}_{(\mathbf{n})}(P)$, as $\Delta A_{(\mathrm{i})}$ tends to zero through a sequence of progressively smaller areas, $\Delta A_{(1)}$, $\Delta A_{(2)}$, $\Delta A_{(3)}$, . . ., $\Delta A_{(\mathrm{N})}$, . . ., all containing the point P,

$$\mathbf{t}_{(\mathbf{n})}(P) = \lim_{\Delta A_{(\mathrm{i})} \to 0} \mathbf{t}^{A}_{(\mathbf{n})}(P). \tag{3.11}$$

When a similar limit is applied to the average couple stress vector we assume that the limit is zero:

$$0 = \lim_{\Delta A_{(\mathrm{i})} \to 0} \mathbf{c}^{A}_{(\mathbf{n})}(P). \tag{3.12}$$

In effect we are assuming that the forces involved are of finite magnitude and that the moment arm associated with $\Delta \mathbf{m}_{(\mathrm{i})}$ vanishes as $\Delta A_{(\mathrm{i})}$ tends to zero. For almost all continuum theories the assumption (3.12) is adequate.

The internal force interaction at a point in an object is adequately represented by the stress vector $\mathbf{t}_{(\mathbf{n})}$ across the plane whose normal is $\mathbf{n}$. However, there is a double infinity of distinct planes with normals $\mathbf{n}$ passing through a single point; thus there is a double infinity of distinct stress vectors acting at each point. The multitude of stress vectors, $\mathbf{t}_{(\mathbf{n})}$, at a point is called the state of stress at the point. The totality of vectors $\mathbf{t}_{(\mathbf{n})}(P)$ at a fixed point P, and for all directions $\mathbf{n}$, is called the *state of stress* at the point P. The representation of the state of stress at a point is simplified by proving that $\mathbf{t}_{(\mathbf{n})}(P)$ must be a linear function of the vector $\mathbf{n}$, as will be done below. The coefficients of this linear relationship will be the stress tensor $\mathbf{T}$. Thus $\mathbf{T}$ will be a linear transformation that transforms $\mathbf{n}$ into $\mathbf{t}_{(\mathbf{n})}$, $\mathbf{t}_{(\mathbf{n})}(P) = \mathbf{T}(P)\cdot\mathbf{n}$. The proof that $\mathbf{T}$ is, in fact, a tensor and the coefficient of a linear transformation will also be provided below. However, even though it has not yet been proved, $\mathbf{T}$ will be referred to as a tensor. The stress tensor $\mathbf{T}$ has components relative to an orthonormal basis that are the elements of the matrix

$$\mathbf{T} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}. \tag{3.13}$$

The component T_{ij} of the stress tensor is the component of the stress vector $\mathbf{t}_{(\mathbf{n})} = \mathbf{t}_{(\mathrm{j})}$ acting on the plane whose normal $\mathbf{n}$ is in the $\mathbf{e}_{\mathrm{j}}$ direction, $\mathbf{n} = \mathbf{e}_{\mathrm{j}}$, projected in the $\mathbf{e}_{\mathrm{i}}$ direction,

$$T_{\mathrm{ij}} = \mathbf{e}_{\mathrm{i}} \cdot \mathbf{t}_{(\mathrm{j})}. \tag{3.14}$$

The nine components of the stress tensor defined by (3.14) and represented by (3.13) are therefore the $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ components of the three stress vectors $\mathbf{t}_{(\mathrm{i})}$, $\mathbf{t}_{(\mathrm{j})}$, and $\mathbf{t}_{(\mathrm{k})}$which act at the point P, on planes parallel to the three mutually perpendicular coordinate planes. This is illustrated in Fig. 3.3. The components T_{11}, T_{22}, T_{33}

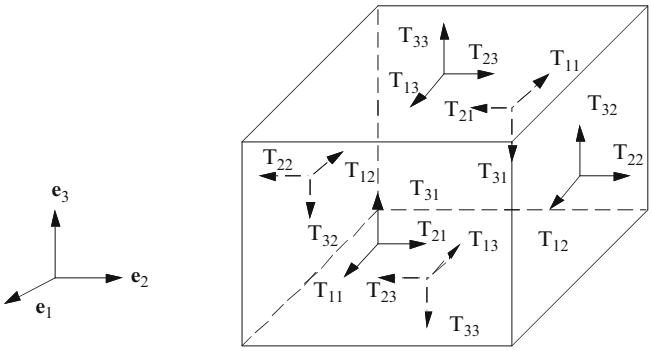


Fig. 3.3 Cartesian components of stress acting on the faces of a small cubic element

are called *normal stresses*, and the remaining components $T_{12}, T_{23}, T_{13}, T_{21}, T_{32}, T_{31}$ are called *shearing stresses*. Each of these components has the dimension of force per unit area.

Recall that stress was defined as the force per unit area that the upper portion U exerts on the lower portion L. From this definition, it follows that if the exterior normal of the object is in the positive coordinate direction, then positive normal and shear stresses will also be in the positive coordinate direction. If the exterior normal of the object is in the negative coordinate direction however, positive normal and shear stresses will point in the negative coordinate direction. The rule for the signs of stress components is as follows: *the stress on a plane is positive if it points in a positive direction on a positive plane, or in a negative direction on a negative plane. Otherwise it is negative.*

It will now be shown, following an 1822 result of Cauchy, that the nine components of stress are sufficient to characterize the entire state of stress at a point. Specifically, it will be shown that *the state of stress at a point P is completely determined if the stress vectors associated with three mutually perpendicular planes are known at P and are continuous in a neighborhood of P*. The stress vector $\mathbf{t}_{(\mathbf{n})}$ acting on any plane whose normal is $\mathbf{n}$ is given by

$$\begin{bmatrix} t_{(\mathbf{n})1} \\ t_{(\mathbf{n})2} \\ t_{(\mathbf{n})3} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \tag{3.15}$$

or

$$\mathbf{t}_{(\mathbf{n})} = \mathbf{T} \cdot \mathbf{n}, t_{(\mathbf{n})\mathrm{i}} = T_{\mathrm{ij}} n_{\mathrm{j}}. \tag{3.16}$$

This means that the stress tensor $\mathbf{T}$ can be considered as a linear transformation that transforms the unit normal $\mathbf{n}$ into the stress vector $\mathbf{t}_{(\mathbf{n})}$ acting on the plane whose normal is $\mathbf{n}$.

To prove this result we consider the tetrahedral element of an object shown in Fig. 3.4 as a free object and apply Newton's second law to the force system acting

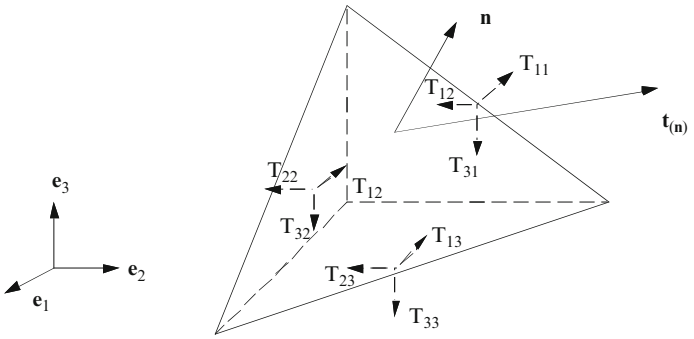


Fig. 3.4 The surface tractions on a tetrahedron

on the tetrahedron. The tetrahedron is selected in such a way that the stress vectors acting on the mutually orthogonal faces are the stress vectors acting on the coordinate planes. Recall that we have represented the components of the stress vectors acting on the planes whose normals are $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ by the components of the stress tensor. The three scalar equations of Newton's second law will suffice for the determination of the unknown components of $\mathbf{t}_{(\mathbf{n})}$ acting on the fourth face. For simplicity we will only derive (3.15) in the $\mathbf{e}_1$ direction. We let A be the area of the inclined face with normal $\mathbf{n}$, and h the perpendicular distance from P to the inclined face. The mean value of $\mathbf{t}_{(\mathbf{n})}$ will be denoted and defined by

$$\bar{\mathbf{t}}_{(\mathbf{n})1} = \frac{1}{A}\int_A \mathbf{t}_{(\mathbf{n})1}\mathrm{d}A = \mathbf{t}_{(\mathbf{n})1}(Q), \tag{3.17}$$

where, as a consequence of the mean value theorem, the point Q lies inside A. Analogously defined mean values of the components of $\mathbf{T}$ over their respective areas will be denoted by $\overline{\mathbf{T}}$. The reason for requiring that the stress components be continuous functions of position in a neighborhood of P is to ensure that the mean values of the stress components actually occur at certain points always within the corresponding areas.

Since the area of the inclined face of the tetrahedron may be represented by the vector $A\mathbf{n}$, where A is the magnitude of the area and $\mathbf{n}$ is the normal to the plane containing the area, the areas of the orthogonal faces are each given by An_1, An_2, An_3. The fact that the areas of four faces of a tetrahedron, where three of the faces of the tetrahedron are orthogonal, are A, An_1, An_2, and An_3 is a result from solid geometry. Summing forces in the $\mathbf{e}_1$ direction and setting the result equal to the mass times the acceleration of the tetrahedron we find that

$$\bar{t}_{(\mathbf{n})1}A - \overline{T}_{11}An_1 - \overline{T}_{12}An_2 - \overline{T}_{13}An_3 + \rho d_1\frac{Ah}{3} = \rho\ddot{x}_1\frac{Ah}{3}, \tag{3.18}$$

where $Ah/3$ is the volume of the tetrahedron, d_1 is the action-at-a-distance force (e.g., gravity) in the $\mathbf{e}_1$ direction and $\ddot{x}_1$ the acceleration of the tetrahedron in that

direction. The next step in this development is to cancel A throughout (3.18) and allow the plane whose normal is $\mathbf{n}$ to approach P, causing the volume of the tetrahedron to vanish as h tends to zero. Before doing this, note that since Q must always lie on A, as h tends to zero, by the mean value theorem:

$$\mathbf{t}_{(\mathbf{n})}(P) = \lim_{h\to 0} \bar{\mathbf{t}}_{(\mathbf{n})1}(P) = \lim_{h\to 0} \mathbf{t}_{(\mathbf{n})1}(Q), \mathbf{T}(P) = \lim_{h\to 0} \overline{\mathbf{T}}(P) = \lim_{h\to 0} \mathbf{T}(Q). \quad (3.19)$$

Canceling A throughout (3.18) and taking the limiting process as h tends to zero, noting that the object force and the acceleration vanish as the volume of the tetrahedron vanishes, it follows from (3.18) and (3.19) that

$$t_{(\mathbf{n})1} = T_{11}n_1 + T_{12}n_2 + T_{13}n_3. \quad (3.20)$$

Repeating this analysis for the $\mathbf{e}_2$ and $\mathbf{e}_3$ directions the result (3.15) is established.

Thus we have shown that the double infinity of possible stress vectors $\mathbf{t}_{(\mathbf{n})}$, which constitutes a state of stress at a point in an object, can be completely characterized by the nine components of $\mathbf{T}$. These nine components are simply the three components of three different stress vectors, one acting on each of the coordinate planes of a reference frame. Thus, in the matrix of tensor components (3.13) the first row consists of the components of the stress vector acting on a plane whose normal is in the $\mathbf{e}_1$ direction. A similar interpretation applies to the second and third rows. When the meaning is not obscured, we will drop the subscript $(\mathbf{n})$ in the equation $\mathbf{t}_{(\mathbf{n})} = \mathbf{T}\cdot\mathbf{n}$ and write it as $\mathbf{t} = \mathbf{T}\cdot\mathbf{n}$, with it being understood that the particular $\mathbf{t}$ depends upon $\mathbf{n}$. The normal stress on a plane is then given by $\mathbf{t}\cdot\mathbf{n} = \mathbf{n}\cdot\mathbf{T}\cdot\mathbf{n}$ and the shear stress in a direction $\mathbf{m}$ lying in the plane whose normal is $\mathbf{n}$, $\mathbf{m}\cdot\mathbf{n} = 0$, is given by $\mathbf{t}\cdot\mathbf{m} = \mathbf{m}\cdot\mathbf{T}\cdot\mathbf{n} = \mathbf{n}\cdot\mathbf{T}\cdot\mathbf{m}$. Note that if a vector $\mathbf{m}'$ is introduced that reverses the direction of $\mathbf{m}$, $\mathbf{m}' = -\mathbf{m}$ then the associated shear stress is given by $\mathbf{t}\cdot\mathbf{m}' = \mathbf{m}'\cdot\mathbf{T}\cdot\mathbf{n} = \mathbf{n}\cdot\mathbf{T}\cdot\mathbf{m}' = -\mathbf{t}\cdot\mathbf{m}$. This shows that, if the unit vector $\mathbf{m}$ is reversed in direction, the opposite value of the shear stress is obtained. Note also that if the unit vector $\mathbf{n}$ is reversed in direction, the opposite value of the shear stress is obtained. If both the unit vectors $\mathbf{n}$ and $\mathbf{m}$ are reversed in direction, the sign of the shear stress is unchanged. These conclusions are all consistent with the definition of the sign of the shear stress.

A short calculation will show that the stress matrix $\mathbf{T}$ is a tensor. Recall that, in order for $\mathbf{T}$ to be a tensor, its components in one coordinate system had to be related to the components in another coordinate system by

$$\mathbf{T}^{(L)} = \mathbf{Q}\cdot\mathbf{T}^{(G)}\cdot\mathbf{Q}^T \quad \text{and} \quad \mathbf{T}^{(G)} = \mathbf{Q}^T\cdot\mathbf{T}^{(L)}\cdot\mathbf{Q}. \text{(A83) repeated}$$

To show that the $\mathbf{T}$ in the relationship (3.15), $\mathbf{t} = \mathbf{T}\cdot\mathbf{n}$, has the tensor property, the equation $\mathbf{t} = \mathbf{T}\cdot\mathbf{n}$ is specified in the Latin coordinate system, $\mathbf{t}^{(L)} = \mathbf{T}^{(L)}\cdot\mathbf{n}^{(L)}$. Then, using the vector transformation law (A77) for $\mathbf{t}$ and $\mathbf{n}$, $\mathbf{t}^{(L)} = \mathbf{Q}\cdot\mathbf{t}^{(G)}$ and $\mathbf{n}^{(L)} = \mathbf{Q}\cdot\mathbf{n}^{(G)}$, respectively, the expression $\mathbf{t}^{(L)} = \mathbf{T}^{(L)}\cdot\mathbf{n}^{(L)}$ is then rewritten as $\mathbf{Q}\cdot\mathbf{t}^{(G)} = \mathbf{T}^{(L)}\cdot\mathbf{Q}\cdot\mathbf{n}^{(G)}$, or $\mathbf{t}^{(G)} = \mathbf{Q}^T\cdot\mathbf{T}^{(L)}\cdot\mathbf{Q}\cdot\mathbf{n}^{(G)}$. Finally, it can be noted from

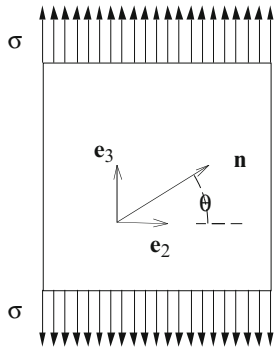


Fig. 3.5 A uniform bar subjected to a uniform stress σ

$\mathbf{t}^{(G)} = \mathbf{Q}^T \cdot \mathbf{T}^{(L)} \cdot \mathbf{Q} \cdot \mathbf{n}^{(G)}$ that since $\mathbf{t}^{(G)} = \mathbf{T}^{(G)} \cdot \mathbf{n}^{(G)}$, it follows that $\mathbf{T}^{(G)} = \mathbf{Q}^T \cdot \mathbf{T}^{(L)} \cdot \mathbf{Q}$. Since this is the transformation rule for a tensor (A83), $\mathbf{T}$ is a tensor.

Example 3.3.1
Determine the stress tensor representing the state of stress at a typical point in the uniform bar subjected to a uniform tensile stress (Fig. 3.5). The applied tensile stress is of magnitude σ and it is assumed that the stress state is the same at all points of the bar. Determine the stress vector $\mathbf{t}$ acting on the plane whose normal is $\mathbf{n}$, where $\mathbf{n}$ is given by $\mathbf{n} = \cos\,\theta\,\mathbf{e}_2 + \sin\,\theta\,\mathbf{e}_3$. Determine the normal stress on the plane whose normal is $\mathbf{n}$ and the shear stress in the direction $\mathbf{m}$, $\mathbf{m}\cdot\mathbf{n} = 0$, $\mathbf{m} = -\sin\,\theta\,\mathbf{e}_2 + \cos\,\theta\,\mathbf{e}_3$, on the plane whose normal is $\mathbf{n}$.

Solution: The components of the stress tensor $\mathbf{T}$ at a typical point in the bar and relative to the coordinate system shown in Fig. 3.5, are given by

$$\mathbf{T} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma \end{bmatrix},$$

thus the only nonzero component of the stress tensor is T_{33}. The stress vector $\mathbf{t}$ acting on the plane whose normal is $\mathbf{n}$ is then given by

$$\mathbf{t} = \mathbf{T}\mathbf{n} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma \end{bmatrix} \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sigma \sin\theta \end{bmatrix},$$

thus this vector has only one nonzero component, namely, $t_3 = \sigma \sin\,\theta$. The normal stress on the plane whose normal is $\mathbf{n}$ is given by $\mathbf{t}\cdot\mathbf{n} = \sigma \sin^2\,\theta$. The shear stress on the plane whose normal is $\mathbf{n}$ in the direction $\mathbf{m}$, $\mathbf{m}\cdot\mathbf{n} = 0$, $\mathbf{m} = -\sin\,\theta\,\mathbf{e}_2 + \cos\,\theta\,\mathbf{e}_3$ is given by $\mathbf{t}\cdot\mathbf{m} = \sigma\,\cos\,\theta\,\sin\,\theta$. (Note that one could choose $\mathbf{m}' = \sin\,\theta\,\mathbf{e}_2 - \cos\,\theta\,\mathbf{e}_3$, where $\mathbf{m}' = -\mathbf{m}$, then $\mathbf{t}\cdot\mathbf{m}' = -\sigma\,\cos\,\theta\,\sin\,\theta = -\mathbf{t}\cdot\mathbf{m}$ and the direction of the shear stress is reversed). From these results it is seen that when the normal to the plane $\mathbf{n}$ coincides with the $\mathbf{e}_3$ direction ($\theta = \pi/2$), the stress component t_3 is

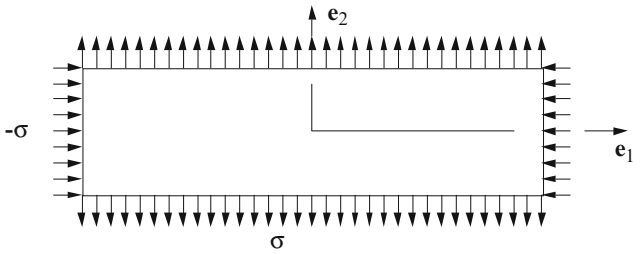


Fig. 3.6 An illustration for problem 3.3.1

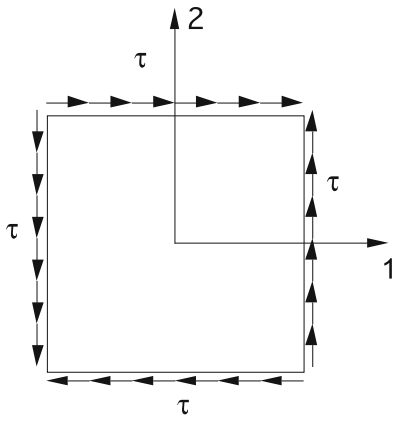


Fig. 3.7 An illustration for problem 3.3.2

equal to σ as one would expect, and when the normal to the plane $\mathbf{n}$ coincides with the $\mathbf{e}_2$ direction, the stress component t_2 is zero. The shear stress $\mathbf{t}\cdot\mathbf{m} = \sigma \cos\theta \sin\theta$ has maxima of $\sigma/2$ at $\theta = \pi/4$ and $3\pi/4$.

Problems

3.3.1. The flat plate shown in Fig. 3.6 has only two stresses acting on it. There is a uniform tensile stress σ in the $\mathbf{e}_2$ direction and a uniform compressive stress—σ in the $\mathbf{e}_1$ direction. Find the stress vector and the components of the stress tensor acting on the planes whose normals are

$$(\mathrm{a})\frac{1}{\sqrt{2}}(\mathbf{e}_1+\mathbf{e}_2), (\mathrm{b})\frac{-1}{\sqrt{2}}(\mathbf{e}_1+\mathbf{e}_2), (\mathrm{c})\frac{1}{\sqrt{2}}(\mathbf{e}_1-\mathbf{e}_2), (\mathrm{d})\frac{1}{\sqrt{2}}(-\mathbf{e}_1+\mathbf{e}_2)$$

3.3.2. The flat plate shown in Fig. 3.7 has shearing stresses τ acting on each of its four faces. Find the stress vector and the stress components acting on the planes whose normals are

$$(\mathrm{a})\mathbf{n} = \frac{1}{\sqrt{2}}(\mathbf{e}_1+\mathbf{e}_2), (\mathrm{b})\mathbf{n} = \frac{1}{\sqrt{2}}(\mathbf{e}_1-\mathbf{e}_2)$$

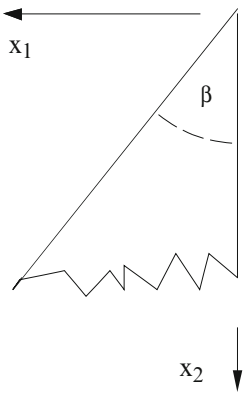


Fig. 3.8 An illustration for problem 3.3.4

3.3.3. Repeat the arguments of this section and show that

$$t_{(\mathbf{n})2} = T_{21}n_1 + T_{22}n_2 + T_{23}n_3$$

List each of the arguments and rules or facts used in the proof.

3.3.4. The roughly triangular-shaped region in Fig. 3.8 with the included angle β represents the upper portion of a dam. The zigzag lines at the bottom of the triangular region are an indication that the dam extends beyond those zigzag lines. Find the stress vectors acting on the face $x_1 = 0$ and on the slanted face of the wedge shown in the Fig. 3.8. The stress matrix at the typical point x_1, x_2, x_3 of the wedge shown in Fig. 3.8 is given by

$$T_{11} = -\gamma x_2, \quad T_{22} = \left(\frac{P}{\tan\beta} - \frac{2\gamma}{\tan^3\beta}\right)x_1 + \left(\frac{\gamma}{\tan^2\beta} - P\right)x_2,$$

$$T_{12} = T_{21} = \frac{-\gamma x_1}{\tan^2\beta}, \quad T_{33} = T_{31} = T_{13} = T_{23} = T_{32} = 0.$$

3.4 The Stress Equations of Motion

In continuum mechanics the stress equations of motion are the most useful form of the principles of balance of linear and angular momentum. The stress equations of motion are statements of Newton's second law (i.e., that force is equal to mass times acceleration) written in terms of stress.

The forces that act on the object in Fig. 3.9 are the surface traction $\mathbf{t}(\mathbf{x}, t)$, which acts at each boundary point, and the action-at-a-distance force $\rho\mathbf{d}$, which represents forces such as gravity and the effect of electromagnetic forces on charges within the

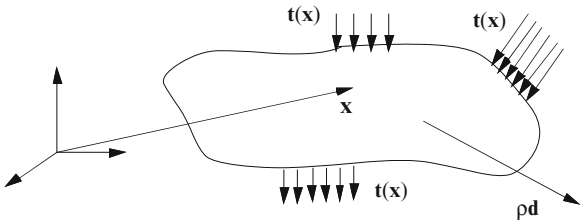


Fig. 3.9 An object acted upon by an action-at-a-distance force $\rho\mathbf{d}$ and a system of surface tractions $\mathbf{t}(\mathbf{x})$

object. For example, at the surface of the earth, in the absence of electromagnetic forces, $\mathbf{d} = -g\mathbf{e}_2$, where $\mathbf{e}_2$ is a positive unit normal to the surface of the earth and g is the acceleration of gravity at the earth's surface. The total force $\Sigma\mathbf{F}$ acting on the object is given by

$$\Sigma\mathbf{F} = \int_{\partial O} \mathbf{t}\mathrm{d}a + \int_{O} \rho\mathbf{d}\mathrm{d}v, \tag{3.21}$$

where ∂O is the surface of the object O. The total moment about the origin of the coordinate system illustrated is

$$\Sigma\mathbf{M} = \int_{\partial O} \mathbf{x} \times \mathbf{t}\mathrm{d}a + \int_{O} \mathbf{x} \times \rho\mathbf{d}\mathrm{d}v, \tag{3.22}$$

where $\mathbf{x}$ is a position vector from the origin. The linear momentum $\mathbf{p}$ and the angular momentum $\mathbf{H}$ of the object in Fig. 3.9 are written as the following integrals over the object O:

$$\mathbf{p} = \int_{O} \rho\dot{\mathbf{x}}\mathrm{d}v, \quad \mathbf{H} = \int_{O} \mathbf{x} \times \rho\dot{\mathbf{x}}\mathrm{d}v. \tag{3.23}$$

The balance of linear momentum requires that the sum of the applied forces equals the time rate of change of the linear momentum, and the balance of angular momentum requires that the sum of the applied moments equals the time rate of change of the angular momentum. In computing the time rates of change of the integrals (3.23) we note that the object O is material so that the material time derivative may be taken inside the integral sign and applied directly to the integrand with the density excluded, (3.8). The material time derivative applied to (3.23) yields

$$\dot{\mathbf{p}} = \int_{O} \rho\ddot{\mathbf{x}}\mathrm{d}v, \dot{\mathbf{H}} = \left[\int_{O} \dot{\mathbf{x}} \times \dot{\mathbf{x}}\rho\mathrm{d}v\right] + \int_{O} \mathbf{x} \times \ddot{\mathbf{x}}\rho\mathrm{d}v = \int_{O} \mathbf{x} \times \ddot{\mathbf{x}}\rho\mathrm{d}v, \tag{3.24}$$

where the term in the square brackets vanishes because $\dot{\mathbf{x}} \times \dot{\mathbf{x}} = 0$. Equating $\dot{\mathbf{p}}$ to $\Sigma\mathbf{F}$ and $\dot{\mathbf{H}}$ to $\Sigma\mathbf{M}$ as required by the conservation of linear and angular momentum, respectively, we obtain from (3.21) and (3.24) that

$$\int_O \rho\ddot{\mathbf{x}}\mathrm{d}v = \int_{\partial O} \mathbf{t}\mathrm{d}a + \int_O \rho\mathbf{d}\mathrm{d}v, \tag{3.25}$$

and from (3.22) and (3.24) that

$$\int_O \mathbf{x} \times \ddot{\mathbf{x}}\rho\mathrm{d}v = \int_{\partial O} \mathbf{x} \times \mathbf{t}\mathrm{d}a + \int_O \mathbf{x} \times \rho\mathbf{d}\mathrm{d}v. \tag{3.26}$$

These integral forms of the balance of linear and angular momentum are the global forms of these principles. The global forms are weaker statements of these balance principles than are the point forms that we will now derive. We say that the point forms are stronger because it must be assumed that the stress is continuously differentiable and that $\rho\ddot{\mathbf{x}}$ and $\rho\mathbf{d}$ are continuous everywhere in the object in order to obtain the point forms from the global forms. The point form of the balance of linear momentum is obtained from (3.25) by first substituting (3.16) into the surface integral in (3.25),

$$\int_O \rho\ddot{\mathbf{x}}\mathrm{d}v = \int_{\partial O} \mathbf{T} \cdot \mathbf{n}\mathrm{d}a + \int_O \rho\mathbf{d}\mathrm{d}v, \tag{3.27}$$

then applying the divergence theorem (A184) to the surface integral in (3.27),

$$\int_O \rho\ddot{\mathbf{x}}\mathrm{d}v = \int_O \nabla \cdot \mathbf{T}^T\mathrm{d}v + \int_O \rho\mathbf{d}\mathrm{d}v, \tag{3.28}$$

The second step in obtaining the point form is to rewrite (3.28) as a single integral,

$$\int_O (\rho\ddot{\mathbf{x}} - \nabla \cdot \mathbf{T}^T - \rho\mathbf{d})\mathrm{d}v = 0, \tag{3.29}$$

and then employ the same argument as was employed in the transition from (3.4) to (3.5); see the discussion following (3.5). It follows then that

$$\rho\ddot{\mathbf{x}} = \nabla \cdot \mathbf{T}^T + \rho\mathbf{d} \tag{3.30}$$

at each point in the object O. There is one more point to make about (3.30) before it is complete. That point is that the stress tensor $\mathbf{T}$ is symmetric and thus the

transpose notation in (3.30) is not necessary. In the next paragraph the conservation of angular momentum is used to show that the stress tensor $\mathbf{T}$ is symmetric.

The arguments to show that the stress tensor $\mathbf{T}$ is symmetric are algebraically simpler if we replace the statement of the conservation of angular momentum given above, (3.26), by the equivalent requirement that the skew-symmetric part of $\mathbf{Z}$,

$$\mathbf{Z} = \int_O \mathbf{x} \otimes \ddot{\mathbf{x}} \rho \mathrm{d}v - \int_{\partial O} \mathbf{x} \otimes \mathbf{t} \mathrm{d}a - \int_O \mathbf{x} \otimes \rho \mathbf{d} \mathrm{d}v, \tag{3.31}$$

be zero:

$$\mathbf{Z} - \mathbf{Z}^T = 0. \tag{3.32}$$

Equation (3.32) is equivalent to (3.26). The rationale for this equivalence is that the components of the cross-product of two vectors, say $\mathbf{a} \times \mathbf{b}$, are equal to the components of the skew-symmetric part of the open product of the two vectors, $\mathbf{a} \otimes \mathbf{b}$. Since (3.32) requires that $\mathbf{Z}$ be symmetric, it follows that the skew-symmetric part of $\mathbf{Z}$ is zero. Equation (3.26) is equivalent to the skew-symmetric part of $\mathbf{Z}$, and the equivalence is established.

The sequence of steps applied to the conservation of linear momentum in the paragraph before last are now applied to the expression (3.31) for $\mathbf{Z}$. The point form of (3.31) is obtained by first substituting (3.16) into the surface integral in (3.31),

$$\mathbf{Z} = \int_O \mathbf{x} \otimes \ddot{\mathbf{x}} \rho \mathrm{d}v - \int_{\partial O} \mathbf{x} \otimes \mathbf{T}\mathbf{n} \mathrm{d}a - \int_O \mathbf{x} \otimes \rho \mathbf{d} \mathrm{d}v, \tag{3.33}$$

then applying the divergence theorem (A184) to the surface integral in (3.33),

$$\mathbf{Z} = \int_O \mathbf{x} \otimes \ddot{\mathbf{x}} \rho \mathrm{d}v - \int_O \{(\nabla \otimes \mathbf{x})\mathbf{T} + (\mathbf{x} \otimes \nabla \cdot \mathbf{T})\} \mathrm{d}v - \int_O \mathbf{x} \otimes \rho \mathbf{d} \mathrm{d}v. \tag{3.34}$$

This result is simplified by observing that $\nabla \otimes \mathbf{x} = \mathbf{1}$ and collecting all the remaining integrals containing $\mathbf{x} \otimes$ together, thus

$$\mathbf{Z} = \int_O \mathbf{T} \mathrm{d}v + \int_O \mathbf{x} \otimes \{\rho \ddot{\mathbf{x}} - \nabla \cdot \mathbf{T}^T - \rho \mathbf{d})\mathrm{d}v. \tag{3.35}$$

The second integral in (3.35) is exactly zero because its integrand contains the point form statement of linear momentum conservation (3.30); thus (3.35) and (3.32) show that

$$\int_O (\mathbf{T} - \mathbf{T}^T) \mathrm{d}v = 0. \tag{3.36}$$

The final step in this development is to employ the same argument as was employed in the transition from (3.4) to (3.5); see the discussion following (3.5). It follows then that the stress tensor is symmetric,

$$\mathbf{T} = \mathbf{T}^T, \tag{3.37}$$

at each point in the object O.

The final form of the stress equations of motion (Truesdell and Toupin, 1960) is obtained from the combination of (3.37) and (3.30), thus

$$\rho\ddot{\mathbf{x}} = \nabla \cdot \mathbf{T} + \rho\mathbf{d}, \mathbf{T} = \mathbf{T}^T. \tag{3.38}$$

This local statement of Newton's second law retains aspects of the original. The mass times acceleration is represented by density times acceleration on the left-hand side. The sum of the forces is represented on the right-hand side by the gradient of the stress tensor and the action-at-a-distance force. The expanded scalar version of the stress equations of motion is

$$\begin{aligned}
\rho\ddot{x}_1 &= \frac{\partial T_{11}}{\partial x_1} + \frac{\partial T_{12}}{\partial x_2} + \frac{\partial T_{13}}{\partial x_3} + \rho d_1, \\
\rho\ddot{x}_2 &= \frac{\partial T_{12}}{\partial x_1} + \frac{\partial T_{22}}{\partial x_2} + \frac{\partial T_{23}}{\partial x_3} + \rho d_2, \\
\rho\ddot{x}_3 &= \frac{\partial T_{13}}{\partial x_1} + \frac{\partial T_{23}}{\partial x_2} + \frac{\partial T_{33}}{\partial x_3} + \rho d_3,
\end{aligned} \tag{3.39}$$

where the symmetry of the stress tensor is expressed in the subscripted indices. For a two-dimensional motion the stress equations of motion reduce to

$$\rho\ddot{x}_1 = \frac{\partial T_{11}}{\partial x_1} + \frac{\partial T_{12}}{\partial x_2} + \rho d_1, \ \rho\ddot{x}_2 = \frac{\partial T_{12}}{\partial x_1} + \frac{\partial T_{22}}{\partial x_2} + \rho d_2. \tag{3.40}$$

Example 3.4.1
The stress tensor in an object is given by

$$\mathbf{T} = \begin{bmatrix} c_1x_1 + c_2x_2 & -c_4x_1 - c_1x_2 & 0 \\ -c_4x_1 - c_1x_2 & c_3x_1 + c_4x_2 + \rho g x_2 & 0 \\ 0 & 0 & c_5 \end{bmatrix},$$

where c_i, $i = 1, \ldots,5$, are constants. This same object is subjected to an action-at a-distance force $\mathbf{d}$ with components [0, −g, 0]. Determine the components of the acceleration vector of this object.

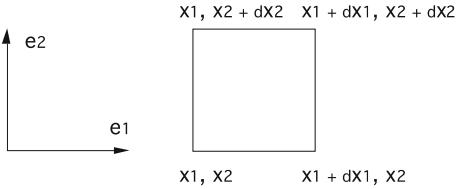


Fig. 3.10 An illustration for problem 3.4.1

Solution: Substitution of the given stress tensor **T** and the action-at-a-distance force **d** into the stress equations of motion (3.39) yields the fact that all the components of the acceleration are zero:

$$\rho\ddot{x}_1 = c_1 - c_1 + 0 = 0, \rho\ddot{x}_2 = -c_4 + c_4 + \rho g - \rho g = 0, \rho\ddot{x}_3 = 0.$$

Problems

3.4.1. Derive the stress equations of motion in two dimensions, (3.40), from the less rigorous argument that consists of applying Newton's second law to the differential element of an object. This element is of length dx_1 and of width dx_2 (see Fig. 3.10). The stresses acting on the element are all referred to the point (x_1, x_2) that is the lower left-hand corner of the element. Use the stresses to calculate the forces acting on the various faces of the element. In order to determine the stresses on opposite faces expand the stresses on one face in a one term Taylor series about the first face. Please keep in mind that, even though the stress tensor is symmetric, the shear stress T_{12} on a face with a normal in a positive direction is opposite in direction from the shear stress T_{12} on a face with a normal in a negative direction.

3.4.2. The components of a stress matrix are

$$\mathbf{T} = \begin{bmatrix} A[x_2^2 + B(x_1^2 - x_2^2)] & -2ABx_1x_2 & 0 \\ -2ABx_1x_2 & A[x_1^2 + B(x_2^2 - x_1^2)] & 0 \\ 0 & 0 & AB(x_1^2 + x_2^2)] \end{bmatrix},$$

where A and B are constants. Does this stress matrix satisfy the stress equations of motion? For what action-at-a-distance force field does it satisfy the stress equations of motion? Assume that the acceleration of the object is zero.

3.4.3. Using the equations of motion to determine the acceleration of an object for which $\mathbf{d} = 5\mathbf{e}_1 + 6\mathbf{e}_2 + 7\mathbf{e}_3$, the density, ρ, is 2 and **T** is given by

$$\mathbf{T} = \begin{bmatrix} 3x_1 & x_3 & 3x_1 \\ x_3 & 4x_2 & 3x_3 \\ 3x_1 & 3x_3 & 7x_3 \end{bmatrix}.$$

3.5 The Conservation of Energy

The idea of energy and its balance or conservation is central in science yet *energy* is not considered to be a precisely defined term. A precise definition would imply that all types of energy are known, and we do not think that they are. In this regard, review the quote from Herbert Callen at the beginning of the chapter. Some energies, such as kinetic energy, are well known and readily identified in any given situation. It is possible to define energy as any member of the set consisting of all energies which are recognized by science, energies such as heat energy, kinetic energy, atomic energy, chemical energy, electromagnetic energy, etc. As science identifies each new energy, it would become a member of this set of energies.

The conservation of energy is therefore viewed here more as a basic method of science rather than as a basic fact in the sense that the charge of an electron is a scientific fact. The conservation of energy is viewed here as a method of checking energetic interactions and discovering new energies. Whenever one approaches a new scientific problem, one tries to select or invent energies such that, by setting their sum equal to a constant, some aspect of the physical phenomenon is correctly described.

In the continuum theories, the known energies will include kinetic energy, heat energy, chemical energy, electromagnetic energy, and so forth. The *total energy* E of a system consists of the sum of all the energies we choose to recognize or define and the remainder of the total energy of a system is said to be the *internal energy* U of the system. That is to say, all the energies that are not singled out and explicitly defined are placed in the category of internal energy. The total energy E of an object consists of a kinetic energy,

$$K = \frac{1}{2}\int_O \rho(\mathbf{v}\cdot\mathbf{v})\mathrm{d}v, \tag{3.41}$$

and an internal energy U,

$$E = K + U, \tag{3.42}$$

where U consists of all energies except kinetic.

The principle of conservation of energy is the statement that the total energy of an object is constant. It is more convenient to reformulate the conservation of energy as a balance of rates: the rate of increase of the total energy of an object is equal to the rate of energy flux into the object. The flux of energy into a object occurs in two ways, first through the mechanical power P of the surface tractions and action-at-a-distance forces and, second, through a direct flow of heat Q into the object. With these definitions and conventions established, the conservation of energy may be written in the form

$$P + Q = \dot{K} + \dot{U} = \dot{E}, \tag{3.43}$$

where $P + Q$ is the rate of energy supply. Equation (3.43) is a global statement of energy conservation and we will need a point form of the principle in continuum mechanics applications. In the point form representation all the variables will be intensive in the conventional thermodynamic use of that word. In thermodynamics an extensive variable is a variable that is additive over the system, e.g., volume or mass, and an intensive variable is a variable that is not additive over the system, e.g., pressure or temperature. To understand these definitions consider adding together two identical masses occupying the same volume at the same temperature and pressure. When two masses have been added together the resulting system has double the volume and double the mass, but it still has the same temperature and pressure. Extensive variables can be made into intensive variables by dividing them by the density of the particle. Thus density or specific volume (the reciprocal of density) is the intensive variable associated with the normally extensive variable mass. The internal energy U, an extensive variable, is represented in terms of the specific internal energy ε, an intensive variable, by the following volume integral

$$U = \int_O \rho \varepsilon \mathrm{d}v. \tag{3.44}$$

Integral representations of the mechanical power P and the non-mechanical or heat power Q supplied to a object O are necessary in order to convert the global form of the energy conservation principle (3.43) to a point form. Heat is transferred into the object at a rate—$\mathbf{q}$ per unit area; the vector $\mathbf{q}$ is called the heat flux vector. The negative sign is associated with $\mathbf{q}$ because of the long-standing tradition in thermodynamics that heat coming out of a system is positive while heat going into a system is negative. The internal sources of heat such as chemical reactions and radiation are represented by the scalar field r per unit mass. Using these representations the total non-mechanical power supplied to an object may be written as the sum of a surface integral and a volume integral,

$$Q = -\int_{\partial O} \mathbf{q} \cdot \mathbf{n} \mathrm{d}a + \int_O \rho r \mathrm{d}v. \tag{3.45}$$

This integral representation for the heat supplied to the object distinguishes between the two possible sources of heat, the internal and the external. Applying the divergence theorem (A184) to the surface integral in (3.45) it is easy to see that Q may also be represented by the volume integral

$$Q = \int_O (\rho r - \nabla \cdot \mathbf{q}) \mathrm{d}v. \tag{3.46}$$

The mechanical power P delivered to the object is represented in integral form by

$$P = \int_{\partial O} \mathbf{t} \cdot \mathbf{v} \mathrm{d}a + \int_{O} \rho \mathbf{d} \cdot \mathbf{v} \mathrm{d}v, \tag{3.47}$$

where $\mathbf{t}$ is the surface traction acting on the surface of the object O, $\mathbf{d}$ is the action-at-a-distance force and $\mathbf{v}$ is the velocity vector. The terms $\mathbf{t}\cdot\mathbf{v}\mathrm{d}a$ and $\rho\mathbf{d}\cdot\mathbf{v}\mathrm{d}v$ both represent the rate at which mechanical work is done on the object, $\mathbf{t}\cdot\mathbf{v}\mathrm{d}a$ is the rate of work of surface forces and $\rho\mathbf{d}\cdot\mathbf{v}\mathrm{d}v$ is the rate of work of action-at-a-distance forces. Substitution of (3.16), $\mathbf{t} = \mathbf{T}\cdot\mathbf{n}$, into (3.47) and subsequent application of the divergence theorem (A184) to the surface integral in the resulting expression yields

$$P = \int_{O} \{\nabla \cdot (\mathbf{T} \cdot \mathbf{v}) + \rho \mathbf{d} \cdot \mathbf{v}\} \mathrm{d}v = \int_{O} \{((\nabla \cdot \mathbf{T}) \cdot \mathbf{v}) + \mathbf{T} : (\nabla \otimes \mathbf{v}) + \rho \mathbf{d} \cdot \mathbf{v}\} \mathrm{d}v,$$

or

$$P = \int_{O} \{(\nabla \cdot \mathbf{T} + \rho \mathbf{d}) \cdot \mathbf{v} + \mathbf{T}{:}\mathbf{L}\} \mathrm{d}v, \tag{3.48}$$

where $\mathbf{L}$ is tensor of velocity gradients defined by (2.31). This result may be further reduced by using the stress equations of motion (3.38) to replace $\nabla\cdot\mathbf{T} + \rho\mathbf{d}$ by $\rho\dot{\mathbf{v}}$, thus

$$P = \int_{O} \{\rho \dot{\mathbf{v}} \cdot \mathbf{v} + \mathbf{T}{:}\mathbf{L}\} \mathrm{d}v. \tag{3.49}$$

Two more manipulations of this expression for P will be performed. First, recall from (2.31) that $\mathbf{L} = [\nabla \otimes \mathbf{v}]^T$ and from (2.32) that $\mathbf{L}$ is decomposed into a symmetric part $\mathbf{D}$ and a skew-symmetric part $\mathbf{W}$ by $\mathbf{L} = \mathbf{D} + \mathbf{W}$. It follows then that

$$\mathbf{T}{:}\mathbf{L} = \mathbf{T}{:}\mathbf{D} + \mathbf{T}{:}\mathbf{W}, \tag{3.50}$$

but $\mathbf{T}{:}\mathbf{W}$ is zero because $\mathbf{T}$ is symmetric by (3.37) and $\mathbf{W}$ is skew-symmetric, hence $\mathbf{T}{:}\mathbf{L} = \mathbf{T}{:}\mathbf{D}$. The second manipulation of (3.49) is to observe that the first integral in (3.49) is the material time rate of change of the kinetic energy K defined by (3.41). To see that the first term in the integral of (3.49) is $\dot{K}$, apply (3.8) to (3.41). With these two changes, the integral expression for P (3.49) now has the form

$$P = \dot{K} + \int_{O} \mathbf{T}{:}\mathbf{D} \mathrm{d}v, \tag{3.51}$$

since, from (3.41),

$$\dot{K} = \int_O \rho(\mathbf{v} \cdot \dot{\mathbf{v}}) \mathrm{d}v.$$

The formula (3.51) shows that the total mechanical power supplied to the object is equal to the time rate of change of kinetic energy plus an integral representing power involved in deforming the object.

The point form statement of the principle of energy conservation will now be obtained by placing the integral representations for U, Q, and P, equations (3.44), (3.46), and (3.51), respectively, into the global statement (3.43), thus

$$\int_O \{\mathbf{T} : \mathbf{D} + \rho r - \nabla \cdot \mathbf{q} - \rho \dot{\varepsilon}\} \mathrm{d}v = 0. \tag{3.52}$$

Now, for the last time in this chapter, the argument employed in the transition from (3.4) to (3.5) is applied here again (see the discussion following (3.5)), thus the integrand of (3.52) must be zero everywhere in the object O;

$$\rho \dot{\varepsilon} = \mathbf{T} : \mathbf{D} + \rho r - \nabla \cdot \mathbf{q}. \tag{3.53}$$

This is the desired point form of the principle of energy conservation. It states that the time rate of change of the specific internal energy ε multiplied by the density ρ is equal to the sum of the stress power, the negative of the divergence of the heat flux, and the internal heat supply term.

Before leaving these considerations of energy, a formula for the quasistatic work done during a loading of an object will be obtained. The mechanical power P delivered to the object, (3.47), is the rate of work. The desired new formula relates to the work done rather that to the rate of doing work or power. A formula identical to formula (3.47) for the rate of work, in every regard except that the velocity is replaced by the displacement, is employed, thus

$$W = \int_{\partial O} \mathbf{t} \cdot \mathbf{u} \mathrm{d}a + \int_O \rho \mathbf{d} \cdot \mathbf{u} \mathrm{d}v. \tag{3.54}$$

W is the mechanical work delivered to the object in a quasistatic loading, $\mathbf{t}$ is the surface traction acting on the surface of the object O, $\mathbf{d}$ is the action-at-a-distance force and $\mathbf{u}$ is the displacement vector. The terms $\mathbf{t}\cdot\mathbf{u}\mathrm{d}a$ and $\rho\mathbf{d}\cdot\mathbf{u}\mathrm{d}v$ both represent the mechanical work done on the object, $\mathbf{t}\cdot\mathbf{u}\mathrm{d}a$ is the work of surface forces and $\rho\mathbf{d}\cdot\mathbf{u}\mathrm{d}v$ is the work of action-at-a-distance forces. Substitution of (3.16), $\mathbf{t} = \mathbf{T}\cdot\mathbf{n}$, into (3.54) and subsequent application of the divergence theorem (A184) to the surface integral in the resulting expression yields

$$W = \int_O \{\nabla \cdot (\mathbf{T} \cdot \mathbf{u}) + \rho \mathbf{d} \cdot \mathbf{u}\} \mathrm{d}v = \int_O \{((\nabla \cdot \mathbf{T}) \cdot \mathbf{u}) + \mathbf{T} : (\nabla \otimes \mathbf{u}) + \rho \mathbf{d} \cdot \mathbf{u}\} \mathrm{d}v,$$

or

$$W = \int_O \{(\nabla \cdot \mathbf{T} + \rho \mathbf{d}) \cdot \mathbf{u} + \mathbf{T} : (\nabla \otimes \mathbf{u})\} \mathrm{d}v. \tag{3.55}$$

This result may be further reduced by using the stress equations of motion (3.38) in the case when $\dot{\mathbf{v}} = 0$ to replace $\nabla \cdot \mathbf{T} + \rho \mathbf{d}$ by 0, thus

$$W = \int_O \mathbf{T} : (\nabla \otimes \mathbf{u}) \mathrm{d}v. \tag{3.56}$$

Recall from (2.49) that $(\nabla \otimes \mathbf{u})^T$ may be decomposed into a symmetric part $\mathbf{E}$ and a skew-symmetric part $\mathbf{Y}$ by $(\nabla \otimes \mathbf{u})^T = \mathbf{E} + \mathbf{Y}$. It follows then that

$$\mathbf{T}:(\nabla \otimes \mathbf{u}) = \mathbf{T}: \mathbf{E} + \mathbf{T}: \mathbf{Y}, \tag{3.57}$$

but $\mathbf{T}$:$\mathbf{Y}$ is zero because $\mathbf{T}$ is symmetric by (3.37) and $\mathbf{Y}$ is skew-symmetric, hence $\mathbf{T}$:$(\nabla \otimes \mathbf{u}) = \mathbf{T}$:$\mathbf{E}$. The work done on the object is then given by

$$W = \int_O \mathbf{T} : \mathbf{E} \mathrm{d}v. \tag{3.58}$$

This means that the local work done is $\mathbf{T}$:$\mathbf{E}$. This result will be of interest in the consideration of elastic objects.

Problem

3.5.1. In terms of the concepts introduced in this section how would one specify a system that was functioning adiabatically globally but not locally? How would one specify a system that was adiabatic locally or point wise? If a system is adiabatic locally what type of energy is the mechanical work done on the object transferred into? If the system is not adiabatic and if the system's internal energy does not change as mechanical work is done on the system, into what type of energy is the mechanical work then converted?

References

Callen HB (1960) Thermodynamics. Wiley, New York

Truesdell CA, Toupin RA (1960) The classical field theories. In: Flugge S (ed) Handbuch der Physik, vol III/1. Springer, Berlin

Chapter 4
Modeling Material Symmetry

"A study of the symmetry of three-dimensional spaces is of great theoretical and practical significance, because symmetrical spaces include crystals (from which, of course, the majority of solids are formed), and all homogeneous fields without exception: electric, magnetic, gravitational, etc. A study of the structures of crystals is unthinkable without a knowledge of the laws governing symmetry of three-dimensional spaces." (Shubnikov and Koptsik 1974)

4.1 Introduction

The variation of material properties with respect to direction at a fixed point in a material is called *material symmetry*. If the material properties are the same in all directions, the properties are said to be *isotropic*. If the material properties are not isotropic, they are said to be *anisotropic*. The type of material anisotropy generally depends upon the size of the representative volume element (RVE). The RVE is the key concept in modeling material microstructure for inclusion in a continuum model. An RVE for a volume surrounding a point in a material is a statistically homogeneous representative of the material in the neighborhood of the point. The RVE concept, described briefly in the following section, is employed in this chapter, which addresses the modeling of material symmetry and, more extensively, in Chapter 7, which addresses the modeling of material microstructure.

The tensors that appear in linear transformations, for example $\mathbf{A}$ in the three-dimensional linear transformation $\mathbf{r} = \mathbf{A} \cdot \mathbf{t}$, (A39), and $\hat{\mathbf{C}}$ in the six-dimensional linear transformation $\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{J}}$, (A160), often represent anisotropic material properties. Several examples of these linear transformations as constitutive equations will be developed in the next chapter. The purpose of this chapter is to present and record representations of $\mathbf{A}$ and $\hat{\mathbf{C}}$ that represent the effects of material symmetry. These results are recorded in Tables 4.3 for $\mathbf{A}$ and Tables 4.4 and 4.5 for $\hat{\mathbf{C}}$, respectively. In these tables the forms of $\mathbf{A}$ and $\hat{\mathbf{C}}$ are given for all eight

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_4,

symmetries. Some background material is provided before the derivation of representations. First, there is a discussion of crystalline and textured materials in Sect. 4.3. Next the only symmetry operation considered here, the plane of mirror symmetry, is introduced in Sect. 4.4, and the symmetries of interest are defined in terms of mirror symmetry planes in Sect. 4.5. The symmetry representations for the forms of $\mathbf{A}$ and $\hat{\mathbf{C}}$ associated with the symmetries of interest are then obtained in Sects. 4.6 and 4.7.

The primary interest here is only in about half of the eight material symmetries admitted by the tensor $\hat{\mathbf{C}}$, however all of them are described for completeness (and because being complete requires little space). The eight material symmetries admitted by the tensor $\hat{\mathbf{C}}$ are triclinic, monoclinic, trigonal, tetragonal, orthotropic, transversely isotropic (or hexagonal), cubic, and isotropic symmetry. The main interest will be in the orthotropic, transversely isotropic, and isotropic symmetries with lesser interest in the triclinic, monoclinic, and trigonal symmetries. Curvilinear and rectilinear anisotropy are described and compared in Sect. 4.8. The representation of the symmetry of a material with chirality (handedness) is considered in Sect. 4.9. Section 4.10 is a short guide to the literature on the subject matter of this chapter.

4.2 The Representative Volume Element

The RVE is a very important conceptual tool for forming continuum models of materials and for establishing restrictions that might be necessary for a continuum model to be applicable. An RVE for a continuum particle $\mathbf{X}$ is a statistically homogeneous representative of the material in the neighborhood of $\mathbf{X}$, that is to say a material volume surrounding $\mathbf{X}$. For purposes of this discussion the RVE is taken to be a cube of side length L_{RVE}; it could be any shape, but it is necessary that it has a characteristic length scale. An RVE is shown in Fig. 4.1; it is a homogenized or average image of a real material volume. Since the RVE image of the material object O averages over the small holes and heterogeneous microstructures, overall it replaces a discontinuous real material object by a smooth continuum model O of the object. The RVE for the representation of a domain of a porous medium by a continuum point is shown in Fig. 4.2. The RVE is necessary in continuum models for all materials; the main question is how large must the length scale L_{RVE} be to obtain a reasonable continuum model. The smaller the value of L_{RVE} the better; in general the value of L_{RVE} should be much less than the characteristic dimension L_{P} of the problem being modeled. On the other hand the L_{RVE} should be much larger than the largest characteristic microstructural dimension L_{M} of the material being modeled, thus $L_{\mathrm{P}} \gg L_{\mathrm{RVE}} \gg L_{\mathrm{M}}$. In wood, for example, this can be a problem because wood has large microstructures and some objects made of wood are small. For low carbon structural steel the bounds on L_{RVE} are much less restrictive, the characteristic size of the problem is greater and the characteristic size of the

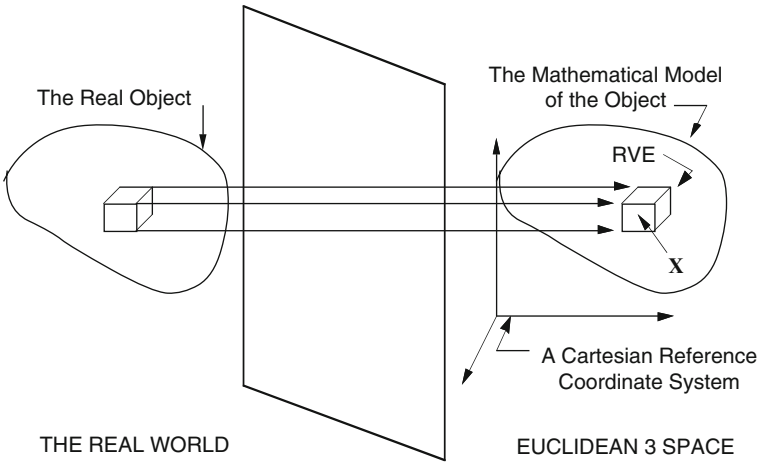


Fig. 4.1 The image of a representative volume element (RVE)

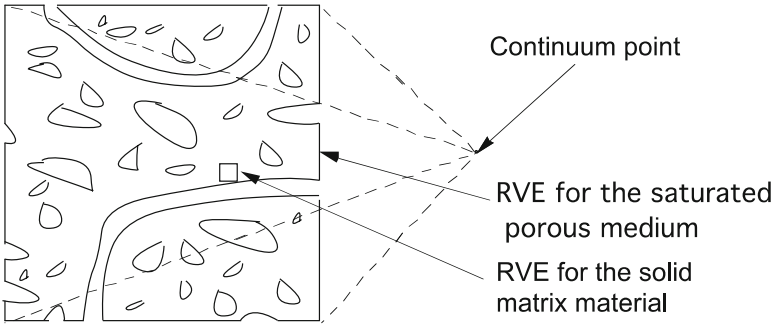


Fig. 4.2 The RVE for the representation of a domain of a porous medium by a continuum point, modified from Cowin (1999)

material microstructure is much less. For many applications the bounds on L_{RVE} are not seriously restrictive, although this is not generally the case with nanomechanical and biological problems. Both in biology and in nanomechanics there are structures that have a significant size range and the modeler must adjust the value of L_{RVE} to the size range of the objects modeled. For example, in biomechanics, continuum models are often made of organs as well as of biological membranes. In the case of the membrane, the L_{RVE} may be less than 0.01 nm while in the case of the organ, the L_{RVE} may be of the order of 0.01 mm or larger. The concept of stress is employed in both cases, with the modeler keeping in mind that the two $L_{\mathrm{RVE}}{}'$s differ by a factor of 1,000,000. The concept of the L_{RVE} may, in this way, be used to justify the application of the concept of stress at different structural levels. The modeler usually does not write down the value of the L_{RVE}

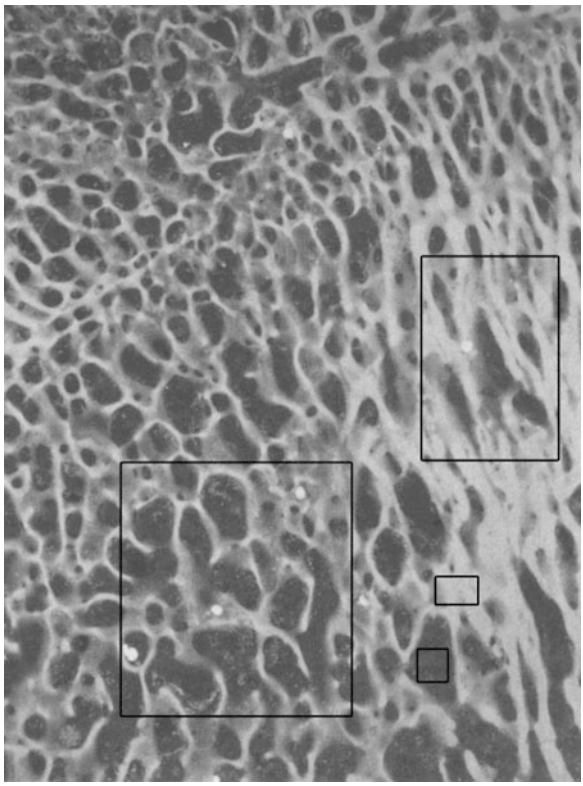


Fig. 4.3 An illustration of a cross-section of trabecular bone. The whiter regions are the bone trabeculae and the darker regions are the spaces occupied by the marrow in the bones of young animals. The rectangular regions represent various RVE's discussed in the text. Adapted from Cowin and Mehrabadi (1989)

in a problem under consideration, hence it is a "hidden" parameter in many applications of continuum models.

The selection of different size RVE's is illustrated in Fig. 4.2. A small RVE will just contain the solid matrix material while a much larger RVE will average over both the pores and the solid matrix. As another illustration of these different RVE sizes relative to a real material, consider a cross-section of trabecular bone shown in Fig. 4.3. The white regions are the bone trabeculae and the darker regions are the pore spaces that are in vivo filled with marrow in the bone of young animals. First consider the small rectangular white region in the lower right quadrant as the first RVE for homogenization. This small rectangular white region is entirely within the trabecular bone domain and thus the properties will be those of trabecular bone. On the other hand, if the small RVE in the darker marrow region is entirely within the bone marrow domain, the properties will be those of the marrow. If the RVE or homogenization domain is taken to be one of the larger rectangles in Fig. 4.3, the properties of the RVE will reflect the properties of both the bone and the marrow, and their values will lie in between these two limits and be proportional to the ratio of the volume of marrow voids to the volume of bone in each rectangle.

4.3 Crystalline Materials and Textured Materials

The difference between the crystalline materials and the textural materials is the difference between the types of force systems that determine the two different types of symmetry. Crystallographic symmetries are determined by the internal force systems that hold the material together in its solid form. These are force systems between atoms or molecules. The lines of action of the attractive forces between lattice points of a crystal determine the crystalline symmetry. On the other hand, the symmetry of textured materials is determined mainly by external, rather than internal, force systems. For example, it is well known that geological materials have material symmetries associated with the stress state experienced by the material during its formative state. Sedimentary deposits are generally organized by the direction of gravity at the time of their formation. Similarly, the material symmetry of structural steel is often determined by the external force systems associated with its method of manufacture (extrusion, rolling, etc.) and not by the fact that it is composed of ferric polycrystals. Man-made composites are generally designed to survive in specific stress states and therefore can generally be considered as having a material symmetry designed for the external force systems they will experience. Plant and animal tissue are known to functionally adapt their local material structure to external loads. In each of these examples the macroscopic material symmetry of the textured material is determined by external force systems, even though at the microscopic level some constituents may have crystalline symmetries determined by internal force systems, as is the case with structural steel and bone tissue.

Crystals have the most clearly defined symmetries of all naturally occurring materials. In crystallography an ideal crystal is defined in terms of a lattice. A *lattice* is an infinite array of evenly spaced points that are all similarly situated. Points are regarded as "similarly situated" if the rest of the lattice appears the same and in the same orientation when viewed from them. An *ideal crystal* is then defined to be an object in which the points, or atoms, are arranged in a lattice. This means that the atomic arrangement appears to be the same and in the same orientation when viewed from all the lattice points, and that the atomic arrangement viewed from any point that is not a lattice point is different from the atomic arrangement viewed from a lattice point. The form and orientation of the lattice are independent of the particular point in the crystal chosen as origin. An ideal crystal is infinite in extent. Real crystals are not only bounded, but also depart from the ideal crystal by possessing imperfections. Forces that act on the lines connecting the lattice points hold crystals together. The force systems that hold a crystal together and give it shape and form are internal force systems.

Most large samples of natural materials are not crystals. They are either not crystalline at all or they are polycrystalline. Polycrystalline materials are composed of small randomly oriented crystalline regions separated by grain boundaries. The material symmetry of these materials is not determined by the crystal structure of their chemical components but by other factors. These factors include optional design for man-made composite materials, growth patterns, and natural selection

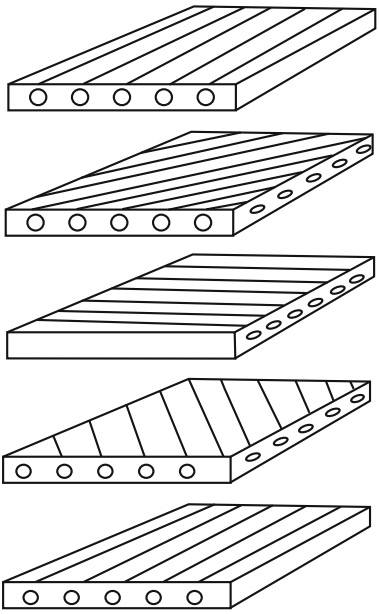


Fig. 4.4 Layers of fibers are alternated to obtain a laminar composite

forces for biological materials, method of formation for geological materials, and method of manufacture for many manufactured materials. These factors are discussed in the following paragraphs.

The selection of material symmetry for a model of a material depends upon the intended application of the model. For example a common application of elasticity theory is to steels that are employed in large objects for structures. In this application steels are conventionally treated as materials with isotropic symmetry. The basis for the isotropic symmetry selection is an RVE of a certain practical size that averages over many grains of the microstructure. Although each crystalline grain is oriented, their orientation is random and their average has no orientation, hence for a large enough RVE the material is isotropic. However if the application of the model is to study the interaction between the crystalline grains, a much smaller RVE will be selected. If the RVE selected is entirely within a single crystalline grain, then this RVE selection would imply cubic material symmetry since ferrous materials are characterized by cubic symmetries. It follows that the selection of different size RVE's can imply different material symmetries. The selection of a particular size RVE is at the discretion of the person making the model, and that selection should be determined by the model's intended use.

Man-made composite materials are often designed to be anisotropic because their intended use is to carry a particular type of loading that requires stiffness and strength in one direction more than in others. While many materials might have the stiffness and strength required, a composite material may have a lesser weight. A unidirectional fiber-reinforced lamina of a composite is illustrated in the top panel of Fig. 4.4. The directions of the fibers in alternate layers can be crossed to

Fig. 4.5 The cylindrical layers associated with each year's growth in wood; note curvilinear orthotropic symmetry

obtain a laminar composite such as that shown in the remaining panels of Fig. 4.4. It is possible to form cylinders and spheres from these laminae. Wood is a natural composite composed of approximately cylindrical layers associated with each year's growth. These growth rings are illustrated in Fig. 4.5. In Fig. 4.6 there is an illustration of the microstructure of a biological material, a three-dimensional view of a nasturtium petiole. In each of these illustrations, it is possible to see how the microstructure of the material will give the material a distinctive anisotropy. Such materials are often called natural composite materials. Bone tissue, bamboo, teeth, and muscles are other examples. These materials evolve their particular microstructures in response to the environmental forces of natural selection.

The method of formation of geological materials generally provides them with a definitive layering that makes them anisotropic. The layered structure is easily seen to be analogous to a layered composite. The deposition of layers is influenced by particle size, because different size particles fall through liquids at different rates. Gravity is the force that gives geological sediments their initial layering. Plate tectonic forces then force these layers in directions other than that in which they were formed, that is why the layers are often viewed in situations where the normal to the plane of the layer is not the direction of gravity.

Macrocomposite man-made materials such as reinforced concrete beams, skis, and helicopter blades are easily seen to be elastically anisotropic. These materials are designed to be anisotropic. In the process of deformation or in the manufacturing process, anisotropy is induced in a material. Anisotropy is also induced in geological and biological materials by deformation. The manufacture of steel by extrusion or rolling induces anisotropy in the steel product as illustrated in Fig. 4.7. Also illustrated, in Fig. 4.8, is the anisotropy induced by deformation.

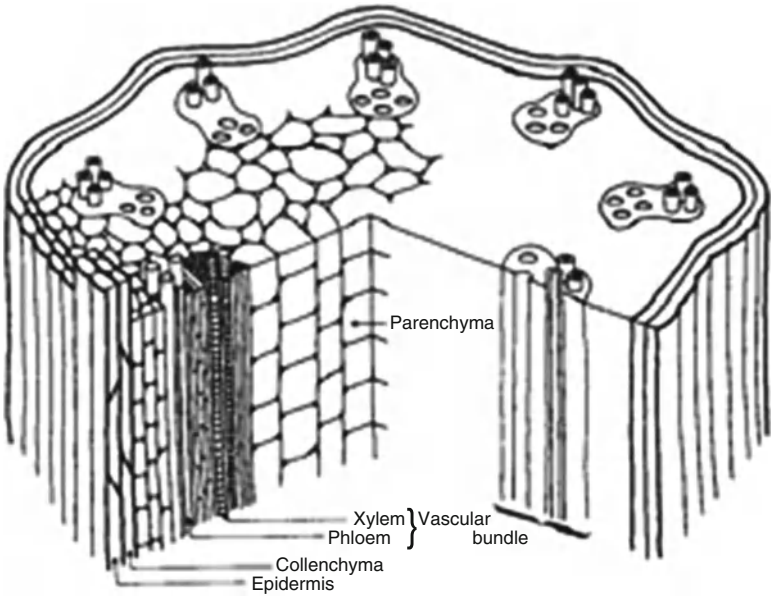


Fig. 4.6 A three-dimensional view of a nasturtium petiole; note curvilinear orthotropic symmetry, reprinted from Wainwright et al. (1976)

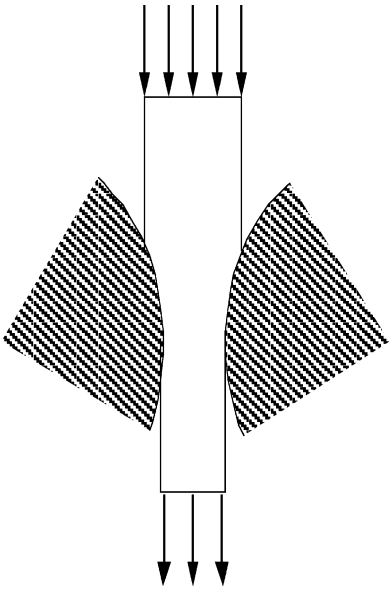


Fig. 4.7 Extrusion or rolling induces anisotropy in a steel product

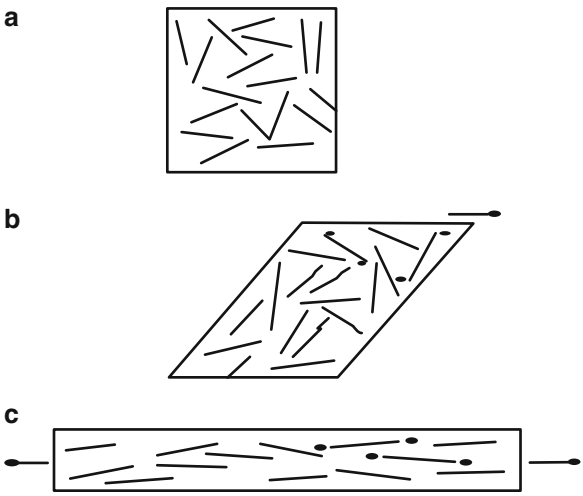


Fig. 4.8 Material anisotropy induced by deformation, from Hull (1981)

The illustration in this figure might represent the fiber deformation in a fibrous composite manufacturing process. However, it could also represent deformation of the collagen fibers in the deformation of a soft tissue.

For noncrystalline materials there are only three material symmetries traditionally considered, orthotropy, transverse isotropy, and isotropy. However, the forms of $\hat{\mathbf{C}}$ for orthotropy and transverse isotropy are the same as the forms of $\hat{\mathbf{C}}$ for the rhombic and hexagonal crystal systems, respectively. Hence when the crystalline and the traditional noncrystalline elastic material symmetries are combined, there are only eight distinct forms of $\hat{\mathbf{C}}$, one for each of the seven crystal systems and isotropy.

4.4 Planes of Mirror Symmetry

Symmetry elements are operations used in the analysis of symmetry. The principal symmetry element of interest here is the plane of mirror or reflective symmetry. We begin with a discussion of congruence and mirror symmetry. Two objects are geometrically congruent if they can be superposed upon one another so that they coincide. The two tetrahedra at the top of Fig. 4.9 are congruent. Congruence of two shapes is a necessary but not sufficient condition for mirror symmetry. A pair of congruent geometric objects is said to have mirror symmetry with respect to a plane if for each point of either object there is a point of the other object such that the pair of points is symmetric with respect to the plane. The two congruent tetrahedra at the bottom of Fig. 4.9 have the special relationship of mirror symmetry with respect to the plane whose end view is indicated by an m. Each congruent geometric object is said to be the reflection of the other. The relationship between the two objects with

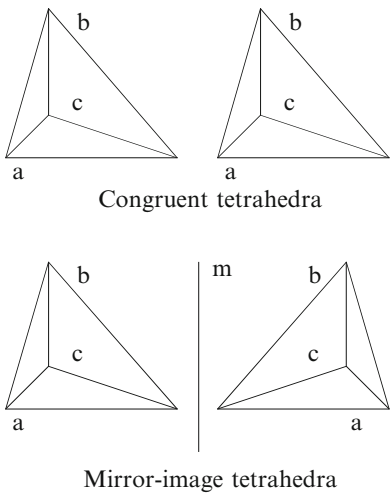


Fig. 4.9 Illustration of mirror symmetry

respect to the symmetry plane is said to be *achiral*, that is to say that the objects are mirror images. If the objects are not reflections of one another, they are said to be *chiral*. The plane with respect to which two objects have mirror symmetry is called their *plane of reflective symmetry*. A material is said to have a *plane of reflective symmetry* or a *mirror plane* at a point in the material if the structure of the material has mirror symmetry with respect to a plane passing through the point.

In the following three sections the concept of a plane of reflective symmetry will now be used to classify the various types of anisotropy possible in the 3-D and 6-D symmetric linear transformations, the $\mathbf{A}$ in the three-dimensional linear transformation $\mathbf{r} = \mathbf{A} \cdot \mathbf{t}$, (A39), and $\hat{\mathbf{C}}$ in the six-dimensional linear transformation $\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{J}}$, (A160). Recall that the matrices $\mathbf{A}$ and $\hat{\mathbf{C}}$ transform according to the rules (A83) and (A162), respectively, when the coordinate system is changed. In order to apply the restrictions of reflective symmetry to (A83) and (A162), it is necessary to have a representation for a plane of reflective symmetry for the orthogonal transformations $\mathbf{Q}$ and $\hat{\mathbf{Q}}$, respectively. To construct such representations let $\mathbf{a}$ be a unit vector representing the normal to a plane of reflective symmetry and let $\mathbf{b}$ be any vector perpendicular to $\mathbf{a}$, then $\mathbf{a}\cdot\mathbf{b} = \mathbf{0}$ for all $\mathbf{b}$. An orthogonal transformation with the properties

$$\mathbf{R}^{(\mathbf{a})} \cdot \mathbf{a} = -\mathbf{a}, \quad \mathbf{R}^{(\mathbf{a})} \cdot \mathbf{b} = \mathbf{b} \tag{4.1}$$

represents a plane of reflective, or mirror, symmetry. The transformation (4.1) carries every vector parallel to the vector $\mathbf{a}$, the normal to the plane of mirror symmetry, into the direction $-\mathbf{a}$ and it carries every vector $\mathbf{b}$ parallel to the plane into itself. The orthogonal transformation with the property (4.1) is given by

$$\mathbf{R}^{(\mathbf{a})} = \mathbf{1} - 2\mathbf{a} \otimes \mathbf{a}, \ (R_{ij}^{(\mathbf{a})} = \delta_{ij} - 2a_i a_j), \tag{4.2}$$

as may be verified directly. The reflective transformation in six dimensions, denoted by $\hat{\mathbf{R}}^{(\mathbf{a})}$, is constructed from $\mathbf{R}^{(\mathbf{a})}$ using (A167), thus

$$\hat{\mathbf{R}}^{(\mathbf{a})} = 2\sqrt{2}$$

$$\begin{bmatrix}
\frac{(1-2a_1^2)^2}{2\sqrt{2}} & \frac{2a_1^2a_2^2}{\sqrt{2}} & \frac{2a_1^2a_3^2}{\sqrt{2}} & 4a_1^2a_2a_3 & 2a_1a_3(2a_1^2-1) & 2a_1a_2(2a_1^2-1) \\
\frac{2a_1^2a_2^2}{\sqrt{2}} & \frac{(1-2a_2^2)^2}{2\sqrt{2}} & \frac{2a_2^2a_3^2}{\sqrt{2}} & 2a_2a_3(2a_2^2-1) & 4a_1a_2^2a_3 & 2a_1a_2(2a_2^2-1) \\
\frac{2a_1^2a_3^2}{\sqrt{2}} & \frac{2a_2^2a_3^2}{\sqrt{2}} & \frac{(1-2a_3^2)^2}{2\sqrt{2}} & 2a_2a_3(2a_3^2-1) & 2a_1a_3(2a_3^2-1) & 4a_1a_2a_3^2 \\
4a_1^2a_2a_3 & 2a_2a_3(2a_2^2-1) & 2a_2a_3(2a_3^2-1) & \frac{2a_1^2+8a_2^2a_3^2-1}{2\sqrt{2}} & \frac{2a_1a_2(4a_3^2-1)}{2\sqrt{2}} & \frac{2a_1a_3(4a_2^2-1)}{2\sqrt{2}} \\
2a_1a_3(2a_1^2-1) & 4a_1a_2^2a_3 & 2a_1a_3(2a_3^2-1) & \frac{2a_1a_2(4a_3^2-1)}{2\sqrt{2}} & \frac{2a_2^2+8a_1^2a_3^2-1}{2\sqrt{2}} & \frac{2a_2a_3(4a_1^2-1)}{2\sqrt{2}} \\
2a_1a_2(2a_1^2-1) & 2a_1a_2(2a_2^2-1) & 4a_1a_2a_3^2 & \frac{2a_1a_3(4a_2^2-1)}{2\sqrt{2}} & \frac{2a_2a_3(4a_1^2-1)}{2\sqrt{2}} & \frac{2a_3^2+8a_2^2a_1^2-1}{2\sqrt{2}}
\end{bmatrix} \tag{4.3}$$

As an example of the application of the result (4.3), the six-dimensional transformations corresponding to planes of mirror symmetry in the $\mathbf{e}_1$ and $\mathbf{e}_2$ directions,

$$\mathbf{R}^{(\mathbf{e}_1)} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R}^{(\mathbf{e}_2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.4}$$

respectively, are given by

$$\hat{\mathbf{R}}^{(\mathbf{e}_1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \quad \hat{\mathbf{R}}^{(\mathbf{e}_2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \tag{4.5}$$

respectively. Other examples are the cases when the normals to the plane of reflective symmetry are vectors in the $\mathbf{e}_1$, $\mathbf{e}_2$ plane, $\mathbf{a} = \cos\theta\ \mathbf{e}_1 + \sin\theta\ \mathbf{e}_2$, or the $\mathbf{e}_2$, $\mathbf{e}_3$ plane, $\mathbf{a} = \cos\theta\,\mathbf{e}_2 + \sin\theta\,\mathbf{e}_3$. In these cases $\mathbf{R}^{(\theta_{12})}$ and $\mathbf{R}^{(\theta_{23})}$ are given by

$$\mathbf{R}^{(\theta_{12})} = \begin{bmatrix} -\cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R}^{(\theta_{23})} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\cos 2\theta & -\sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{bmatrix}, \tag{4.6}$$

respectively, and $\hat{\mathbf{R}}^{(\theta_{12})}$ and $\hat{\mathbf{R}}^{(\theta_{23})}$ are given by

$$\begin{bmatrix} \cos^2 2\theta & \sin^2 2\theta & 0 & 0 & 0 & \sqrt{2}\sin 2\theta\cos 2\theta \\ \sin^2 2\theta & \cos^2 2\theta & 0 & 0 & 0 & -\sqrt{2}\sin 2\theta\cos 2\theta \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & 0 & 0 & -\sin 2\theta & -\cos 2\theta & 0 \\ \sqrt{2}\sin 2\theta\cos 2\theta & -\sqrt{2}\sin 2\theta\cos 2\theta & 0 & 0 & 0 & \sin^2 2\theta - \cos^2 2\theta \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta & \sin^2 2\theta & \sqrt{2}\sin 2\theta\cos 2\theta & 0 & 0 \\ 0 & \sin^2 2\theta & \cos^2 2\theta & -\sqrt{2}\sin 2\theta\cos 2\theta & 0 & 0 \\ 0 & \sqrt{2}\sin 2\theta\cos 2\theta & -\sqrt{2}\sin 2\theta\cos 2\theta & \sin^2 2\theta - \cos^2 2\theta & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos 2\theta & -\sin 2\theta \\ 0 & 0 & 0 & 0 & -\sin 2\theta & -\cos 2\theta \end{bmatrix} \tag{4.7}$$

respectively. The formulas (4.1)–(4.7) provide the 3-D and 6-D orthogonal transformations for a plane of reflective symmetry. These mirror symmetry transformations will be used in Sects. 4.6 and 4.7 to develop the representations in Tables 4.3 and Tables 4.4 and 4.5 for the matrices $\mathbf{A}$ and $\hat{\mathbf{C}}$, respectively, for the various material symmetries.

Problems

4.4.1. Verify that the transformation (4.2) $\mathbf{R}^{(\mathbf{a})} = \mathbf{1} - 2\mathbf{a} \otimes \mathbf{a}$ has the properties (4.1), $\mathbf{R}^{(\mathbf{a})} \cdot \mathbf{a} = -\mathbf{a}$ and $\mathbf{R}^{(\mathbf{a})} \cdot \mathbf{b} = \mathbf{b}$ where $\mathbf{a} \cdot \mathbf{b} = \mathbf{0}$ for all $\mathbf{b}$.

4.4.2. Construct the orthogonal transformations $\mathbf{R}^{(\mathbf{e}_3)}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_3)}$, then verify their orthogonality.

4.4.3. Construct the orthogonal transformations $\mathbf{R}^{(\theta_{13})}$ and $\hat{\mathbf{R}}^{(\theta_{13})}$ associated with the vector $\mathbf{a} = \cos\theta\, \mathbf{e}_1 + \sin\theta\, \mathbf{e}_3$, and verify their orthogonality.

4.4.4. Show that the reflections $\mathbf{R}^{(\theta_{12})}$ and $\hat{\mathbf{R}}^{(\theta_{12})}$ when evaluated at $\theta = 0$ coincide with the reflections $\mathbf{R}^{(\mathbf{e}_1)}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_1)}$, and when evaluated at $\theta = \pi/2$, they coincide with the reflections $\mathbf{R}^{(\mathbf{e}_2)}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_2)}$.

4.4.5. Show that the reflections $\mathbf{R}^{(\theta_{23})}$ and $\hat{\mathbf{R}}^{(\theta_{23})}$ when evaluated at $\theta = 0$ coincide with the reflections $\mathbf{R}^{(\mathbf{e}_2)}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_2)}$, and when evaluated at $\theta = \pi/2$, they coincide with the reflections $\mathbf{R}^{(\mathbf{e}_3)}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_3)}$.

4.4.6. Construct the three- and six-dimensional reflective transformations for each of four normal vectors, $\mathbf{k} = (1/2)(\mathbf{e}_1 + (\sqrt{3})\mathbf{e}_2)$, $\mathbf{p} = (1/2)(-\mathbf{e}_1 + (\sqrt{3})\mathbf{e}_2)$, $\mathbf{m} = (1/2)(\sqrt{3})\mathbf{e}_1 + \mathbf{e}_2)$, $\mathbf{n} = (1/2)\lambda(\sqrt{3})\mathbf{e}_1 - \mathbf{e}_2)$. Show that the set of six vectors $\mathbf{k}$, $\mathbf{p}$, $\mathbf{m}$, $\mathbf{n}$, $\mathbf{e}_1$, and $\mathbf{e}_2$ form a set that makes a pattern. The pattern is such that each vector of the set of six vector points is one of six different directions and makes angles that are each multiples of $\pi/6$ with the other vectors.

4.5 Characterization of Material Symmetries by Planes of Symmetry

In this section the number and orientation of the planes of reflective symmetry possessed by each linear elastic material symmetry will be used to define it. These material symmetries include isotropic symmetry and the seven anisotropic symmetries, triclinic, monoclinic, trigonal, orthotropic, hexagonal (transversely isotropy), tetragonal, and cubic. These symmetries may be classified strictly on the basis of the number and orientation of their planes of mirror symmetry. Figure 4.10 illustrates the relationship between the various symmetries; it is organized such that the lesser symmetries are at the upper left and as one moves to the lower right one sees crystal systems with greater and greater symmetry. The number of planes of symmetry for each material symmetry is given in Table 4.1 and, relative to a selected reference coordinate system, the normals to the planes of symmetry for each material symmetry are specified in Table 4.2. *Triclinic* symmetry has no planes of reflective symmetry so there are no symmetry restrictions for a triclinic material. *Monoclinic* symmetry has exactly one plane of reflective symmetry. *Trigonal* symmetry has three planes of symmetry whose normals all lie in the same plane and make angles of 120° with each other; its threefold character stems from this relative orientation of its planes of symmetry. *Orthotropic* symmetry has three mutually perpendicular planes of reflective symmetry, but the existence of the third plane is implied by the first two. That is to say, if there exist two perpendicular planes of reflective symmetry, there will automatically be a third one perpendicular to both of the first two. *Tetragonal* symmetry has the five planes of symmetry ($\boldsymbol{a}_1$–$\boldsymbol{a}_5$) illustrated in Fig. 4.11; four of the five planes of symmetry have normals that all lie in the same plane and make angles of 45° with each other; its fourfold character stems from this relative orientation of its planes of symmetry. The fifth plane of symmetry is the plane containing the normals to the other four planes of symmetry. *Hexagonal* symmetry has seven planes of symmetry; six of the seven planes of symmetry have normals that all lie in the same plane and make angles of 60° with each other; its sixfold character stems from this relative orientation of its

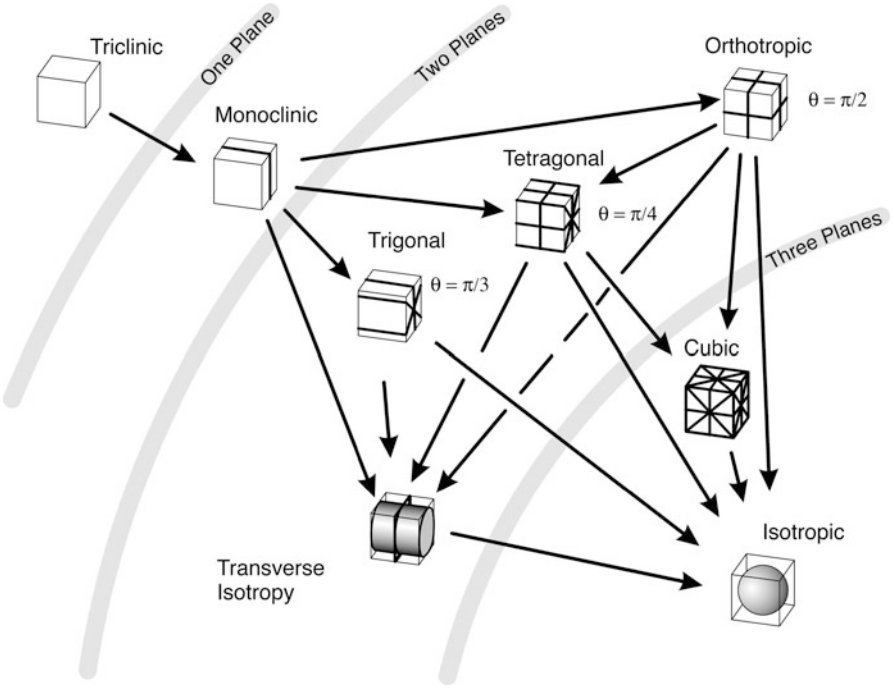


Fig. 4.10 The hierarchical organization of the eight material symmetries of linear elasticity. The figure is organized such that the lower symmetries are at the upper left and as one moves down and across the table to the right one encounters crystal systems with greater and greater symmetry. From Chadwick et al. (2001)

Table 4.1 The distinct symmetries of linear anisotropic elasticity

Triclinic		0	18 (21)
Monoclinic	1	0	12 (13)
Orthotropic or orthorhombic	3	0	9
Tetragonal	5	0	6 (7)
Cubic	9	0	3
Trigonal	3	0	6 (7)
Hexagonal	7	0	5
Transverse isotropy	$1+\infty$	1	5
Isotropy	∞^2	∞^2	2

planes of symmetry. The seventh plane of symmetry is the plane containing the normals to the other six planes of symmetry. The illustration for hexagonal symmetry is similar to that for tetragonal symmetry shown in Fig. 4.11; the difference is that there are six rather than four planes with normals all lying in the same plane and that those normals make angles of 30° rather than 45° with each other. *Cubic* symmetry has the nine planes of symmetry illustrated in Fig. 4.12. The positive octant at the front of Fig. 4.12 is bounded by three of the symmetry planes with normals $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ and contains traces of the six other planes of symmetry.

Table 4.2 The normals to the planes of symmetry of the indicated symmetry group

Type of material symmetry	Normals to the planes of symmetry of the indicated symmetry group
Triclinic	None
Monoclinic	$\mathbf{e}_1$
Orthotropic	$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$
Tetragonal	$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, (1/\sqrt{2})(\mathbf{e}_1 + \mathbf{e}_2)$ and $(1/\sqrt{2})(\mathbf{e}_1 - \mathbf{e}_2)$
Cubic	$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, (1/\sqrt{2})(\mathbf{e}_1 + \mathbf{e}_2), (1/\sqrt{2})(\mathbf{e}_1 - \mathbf{e}_2), (1/\sqrt{2})(\mathbf{e}_1 + \mathbf{e}_3)$ $(1/\sqrt{2})(\mathbf{e}_1 - \mathbf{e}_3), (1/\sqrt{2})(\mathbf{e}_2 + \mathbf{e}_3), (1/\sqrt{2})(\mathbf{e}_2 - \mathbf{e}_3)$
Trigonal	$\mathbf{e}_1, (1/2)(\mathbf{e}_1 + \sqrt{3}\,\mathbf{e}_2)$ and $(1/2)(\mathbf{e}_1 - \sqrt{3}\,\mathbf{e}_2)$
Hexagonal	$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, (1/2)(\sqrt{3}\,\mathbf{e}_1 + \mathbf{e}_2), (1/2)(\sqrt{3}\,\mathbf{e}_1 - \mathbf{e}_2), (1/2)(\mathbf{e}_1 + \sqrt{3}\,\mathbf{e}_2)$ and $(1/2)(\mathbf{e}_1 - \sqrt{3}\,\mathbf{e}_2)$.
Transverse isotropy	$\mathbf{e}_3$ and any vector lying in the $\mathbf{e}_1, \mathbf{e}_2$ plane
Isotropy	Any vector

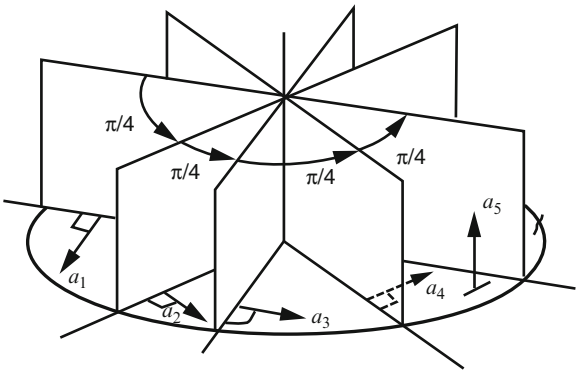


Fig. 4.11 An illustration of the five planes of symmetry characterizing tetragonal symmetry. The normals to the five planes are denoted by $\boldsymbol{a}_1$–$\boldsymbol{a}_4$. Four of the five planes of symmetry have normals that all lie in the same plane and make angles of 45° with each other. The fourfold character of the symmetry stems from the relative orientation of its planes of symmetry. The fifth plane of symmetry is the plane containing the normals to the other four planes of symmetry. Modified from Rovati and Taliercio (2003)

If every vector in a plane is normal to a plane of reflective symmetry, the plane is called a *plane of isotropy*. It can be shown that a plane of isotropy is itself a plane of reflective symmetry. The material symmetry characterized by a single plane of isotropy is said to be *transverse isotropy*. In the case of linear elasticity the $\hat{\mathbf{C}}$ matrix for transversely isotropic symmetry is the same as the $\hat{\mathbf{C}}$ matrix for hexagonal symmetry and so a distinction is not made between these two symmetries. *Isotropic* symmetry is characterized by every direction being the normal to a plane of reflective symmetry, or equivalently, every plane being a plane of isotropy.

In this presentation the reference coordinate system for the elastic symmetries is selected so that there are only 18 distinct components of $\hat{\mathbf{C}}$ for triclinic symmetry, 12 for monoclinic and so that the 7 constant tetragonal and trigonal symmetries are the

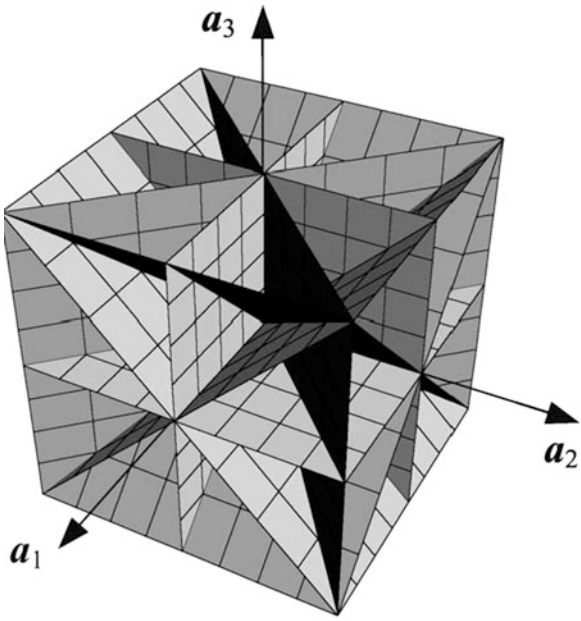


Fig. 4.12 An illustration of the nine planes of symmetry characterizing cubic symmetry. The positive octant at the front of the perspective is bounded by three of the symmetry planes with normals $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ and contains traces of the six other planes of symmetry. From Rovati and Taliercio (2003)

same as the 6 constant tetragonal and trigonal symmetries, respectively. This selection of the coordinate system is always possible without restricting the generality of the matrix representations (Cowin 1995; Fedorov 1968).

This classification of the types of linear elastic material symmetries by the number and orientation of the normals to the planes of material symmetry is fully equivalent to the crystallographic method using group theory (Chadwick et al. 2001).

Problems

4.5.1. Construct diagrams of the number and orientation of the normals to the planes of reflective symmetry for six of the eight material symmetries (diagrams for the other two, tetragonal and cubic, and prose descriptions for all the symmetries are given in the text); the triclinic, monoclinic, orthotropic, hexagonal, transverse isotropic, and isotropic material symmetries.

4.5.2. Construct a diagram for the set of normals to the planes of reflective symmetry given in Problem 4.4.6, with the addition of the normal $\mathbf{e}_3$ so that the set now consists of seven vectors: $\mathbf{k}$, $\mathbf{p}$, $\mathbf{m}$, $\mathbf{n}$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$.

4.5.3. From the images of simplified crystal models shown in Figs. 4.13a, b, identify the appropriate material symmetry for the object in each figure, set up the convenient coordinate system for the object, and list the vectors of the normals to the planes of symmetry characterizing each particular symmetry.

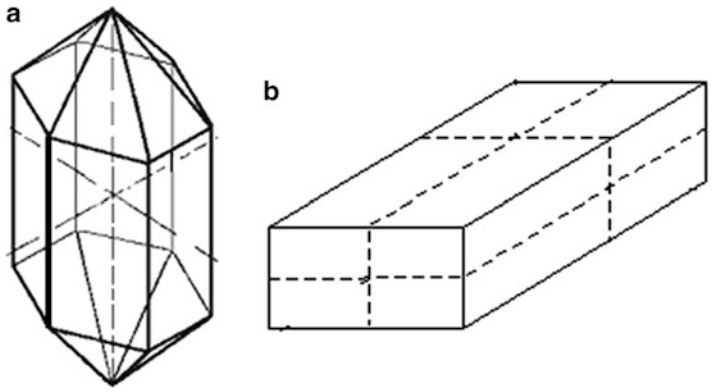


Fig. 4.13 (**a**) See Problem 4.5.3. (**b**) See Problem 4.5.3

4.6 The Forms of the Symmetric Three-Dimensional Linear Transformation A

In this section the definitions of the material symmetries given in Sect. 4.5 are used in conjunction with the orthogonal transformation (4.2) characterizing a plane of reflective symmetry and transformation law (A83) to derive the forms of the three-dimensional linear transformation **A** for the material symmetries of interest. First, since triclinic symmetry has no planes of reflective symmetry, there are no symmetry restrictions for a triclinic material and the linear transformation **A** is unrestricted. This conclusion is recorded in Table 4.3.

Monoclinic symmetry has exactly one plane of reflective symmetry. This means that the material coefficients appearing in **A** must be unchanged by one reflective symmetry transformation. Let $\mathbf{e}_1$ be the normal to the plane of reflective symmetry so that the reflective symmetry transformation is $\mathbf{R}^{(\mathbf{e}_1)}$, given by the first of (4.4). The tensor **A** is subject to the transformation

$$\mathbf{A}^{(L)} = \mathbf{R}^{(\mathbf{e}_1)} \cdot \mathbf{A}^{(G)} \cdot \left[\mathbf{R}^{(\mathbf{e}_1)}\right]^{\mathrm{T}} \tag{4.8}$$

which follows from the first of (A83) by setting $\mathbf{T} = \mathbf{A}$ and $Q = \mathbf{R}^{(\mathbf{e}_1)}$. Substituting for **A** and $\mathbf{R}^{(\mathbf{e}_1)}$ in this equation, one finds that

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

Table 4.3 The forms of the three-dimensional linear transformation **A** for the different material symmetries. Note that these forms are not here required to be symmetric, three of them are and three of them are not. However symmetry will be required for all of them by different and specialized physical arguments in subsequent chapters

Type of material symmetry	Form of linear transformation **A**
Triclinic	$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$
Monoclinic	$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix}$
Orthotropic	$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix}$
Hexagonal, trigonal, tetragonal (some crystal classes)	$\begin{bmatrix} A_{11} & A_{12} & 0 \\ -A_{12} & A_{11} & 0 \\ 0 & 0 & A_{33} \end{bmatrix}$
Transversely isotropic, hexagonal, trigonal, tetragonal (the other crystal classes)	$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{33} \end{bmatrix}$
Isotropic and cubic	$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{11} \end{bmatrix}$

or

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} A_{11} & -A_{12} & -A_{13} \\ -A_{21} & A_{22} & A_{23} \\ -A_{31} & A_{32} & A_{33} \end{bmatrix}. \tag{4.9}$$

The transformation (4.8) or (4.9) is thus seen to leave the tensor **A** unchanged by the reflection only if $A_{12} = A_{21} = A_{13} = A_{31} = 0$. It follows then that the form of the tensor **A** consistent with monoclinic symmetry characterized by a plane of reflective symmetry normal to the $\mathbf{e}_1$ base vector must satisfy the conditions $A_{12} = A_{21} = A_{13} = A_{31} = 0$. This result for monoclinic symmetry is recorded in Table 4.3.

Three mutually perpendicular planes of reflective symmetry characterize orthotropic symmetry, but the third plane is implied by the first two. This means that the material coefficients appearing in the representation of **A** for monoclinic symmetry must be unchanged by one more perpendicular reflective symmetry transformation. Let $\mathbf{e}_2$ be the normal to the plane of reflective symmetry so that

the reflective symmetry transformation is $\mathbf{R}^{(\mathbf{e}_2)}$ given by the second of (4.4). The monoclinic form of the tensor $\mathbf{A}$ is subject to the transformation

$$\mathbf{A}^{(L)} = \mathbf{R}^{(\mathbf{e}_2)} \cdot \mathbf{A}^{(G)} \cdot \left[\mathbf{R}^{(\mathbf{e}_2)}\right]^{\mathrm{T}} \tag{4.10}$$

which follows from the first of (A83) by setting $\mathbf{T} = \mathbf{A}$ and $Q = \mathbf{R}^{(\mathbf{e}_2)}$. Substituting into (4.10) the representation for the monoclinic form for $\mathbf{A}$ and the representation for $\mathbf{R}^{(\mathbf{e}_2)}$, one finds that

$$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

or

$$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & -A_{23} \\ 0 & -A_{32} & A_{33} \end{bmatrix}. \tag{4.11}$$

The transformation (4.10) or (4.11) is thus seen to leave the monoclinic form of the tensor $\mathbf{A}$ unchanged by the reflection only if $A_{32} = A_{23} = 0$. It follows then that the form of the tensor $\mathbf{A}$ consistent with an orthotropic symmetry characterized by planes of reflective symmetry normal to the $\mathbf{e}_1$ and $\mathbf{e}_2$ base vectors must satisfy the conditions $A_{32} = A_{23} = 0$. It is then possible to show that this restriction also permits the existence of a third plane of reflective symmetry perpendicular to the first two. This result for orthotropic symmetry is recorded in Table 4.3.

A transversely isotropic material is one with a plane of isotropy. A plane of isotropy is a plane in which every vector is the normal to a plane of reflective symmetry. This means that the material coefficients appearing in the representation of $\mathbf{A}$ for orthotropic symmetry must be unchanged by any reflective symmetry transformation characterized by any unit vector in a specified plane. Let the plane be the $\mathbf{e}_1$, $\mathbf{e}_2$ plane and let the unit vectors be $\mathbf{a} = \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2$ for any and all values of θ; then the reflective symmetry transformations of interest are $\mathbf{R}^{(\theta_{12})}$ given by the first of (4.6). The orthotropic form of the tensor $\mathbf{A}$ is subject to the transformation

$$\mathbf{A}^{(L)} = \mathbf{R}^{(\theta_{12})} \cdot \mathbf{A}^{(G)} \cdot \left[\mathbf{R}^{(\theta_{12})}\right]^{\mathrm{T}} \tag{4.12}$$

which follows from the first of (A83) by setting $\mathbf{T} = \mathbf{A}$ and $Q = \mathbf{R}^{(\theta_{12})}$. Substitution for $\mathbf{R}^{(\theta_{12})}$ and the orthotropic form for $\mathbf{A}$ into this equation, one finds that

$$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} = \begin{bmatrix} -\cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} \times \begin{bmatrix} -\cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

or

$$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} = \begin{bmatrix} A_{11}\cos^2 2\theta + A_{22}\sin^2 2\theta & \frac{1}{4}\sin 4\theta(A_{11} - A_{22}) & 0 \\ \frac{1}{4}\sin 4\theta(A_{11} - A_{22}) & A_{11}\sin^2 2\theta + A_{22}\cos^2 2\theta & 0 \\ 0 & 0 & A_{33} \end{bmatrix}. \tag{4.13}$$

The transformation (4.12) or (4.13) is thus seen to leave the transversely isotropic form of the tensor $\mathbf{A}$ unchanged by the reflection only if $A_{11} = A_{22}$. It follows then that the transversely isotropic form of the tensor $\mathbf{A}$ consistent with transversely isotropic symmetry characterized by a plane of isotropy in the $\mathbf{e}_1$, $\mathbf{e}_2$ plane must satisfy the conditions $A_{11} = A_{22}$. This result for transversely isotropic symmetry is recorded in Table 4.3. Algebraic procedures identical with those described above may be used to show that the forms of the tensor $\mathbf{A}$ consistent with the trigonal, tetragonal, and hexagonal symmetries are each identical with that for transversely isotropic symmetry.

Isotropic symmetry is characterized by every direction being the normal to a plane of reflective symmetry, or equivalently, every plane being a plane of isotropy. This means that the material coefficients appearing in the representation of $\mathbf{A}$ for transversely isotropic symmetry must be unchanged by any reflective symmetry transformation characterized by any unit vector in any direction. In addition to the $\mathbf{e}_1$, $\mathbf{e}_2$ plane considered as the plane of isotropy for transversely isotropic symmetry, it is required that the $\mathbf{e}_2$, $\mathbf{e}_3$ plane be a plane of isotropy. The second plane of isotropy is characterized by the unit vector $\mathbf{a} = \cos\theta\,\mathbf{e}_2 + \sin\theta\,\mathbf{e}_3$ for any and all values of θ, then the reflective symmetry transformation of interest is $\mathbf{R}^{(\theta_{23})}$ given by the second of (4.6). The transversely isotropic form of the tensor $\mathbf{A}$ must be invariant under the transformation

$$\mathbf{A}^{(L)} = \mathbf{R}^{(\theta_{23})} \cdot \mathbf{A}^{(G)} \cdot \left[\mathbf{R}^{(\theta_{23})}\right]^{\mathrm{T}} \tag{4.14}$$

which follows from the first of (A83) by setting $\mathbf{T} = \mathbf{A}$ and $Q = \mathbf{R}^{(\theta_{23})}$. Substitution for $\mathbf{R}^{(\theta_{23})}$ and the transversely isotropic form for $\mathbf{A}$ into this equation, one finds that

$$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\cos 2\theta & -\sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\cos 2\theta & -\sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{bmatrix}$$

or

$$\begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{33} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{11}\cos^2 2\theta + A_{33}\sin^2 2\theta & \frac{1}{4}\sin 4\theta(A_{11} - A_{33}) \\ 0 & \frac{1}{4}\sin 4\theta(A_{11} - A_{33}) & A_{11}\sin^2 2\theta + A_{33}\cos^2 2\theta \end{bmatrix}, \tag{4.15}$$

>The transformation (4.14) or (4.15) is thus seen to leave the isotropic form of the tensor **A** unchanged by the reflection only if $A_{11} = A_{33}$. It follows then that the isotropic form of the tensor **A** consistent with isotropic symmetry characterized by planes of isotropy in the $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_2, \mathbf{e}_3$ planes must satisfy the conditions $A_{11} = A_{33}$. Actually there are many ways to make the transition from transversely isotropic symmetry to isotropic symmetry other than the method chosen here. Any plane of reflective symmetry added to the plane of isotropy of transversely isotropic symmetry, and not coincident with the plane of isotropy, will lead to isotropic symmetry. This result for isotropic symmetry is recorded in Table 4.3. Algebraic procedures identical with those described above may be used to show that the form of the tensor **A** consistent with cubic symmetry is identical with that for isotropic symmetry.

Problems

4.6.1 Show that the orthotropic form of the tensor **A** given in Table 4.3 is invariant under the transformation $\mathbf{R}^{(\mathbf{e}_3)}$ constructed in Problem 4.4.3.

4.6.2 Show that the isotropic form of the tensor **A** given in Table 4.3 is invariant under the transformation $\mathbf{R}^{(\theta_{13})}$ constructed in Problem 4.4.3.

4.6.3 Construct a representation for **A** that is invariant under the seven reflective transformations formed from the set of normals to the planes of reflective symmetry given in Problem 4.5.2: **k**, **p**, **m**, **n**, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$. Does the result coincide with one of the representations already in Table 4.3? If it does, please explain.

4.7 The Forms of the Symmetric Six-Dimensional Linear Transformation $\hat{\mathbf{C}}$

In this section the definitions of the material symmetries given in Sect. 4.5 are used in conjunction with the six-dimensional orthogonal transformation (4.3) characterizing a plane of reflective symmetry and transformation law (A162) to

Table 4.4 The forms of the six-dimensional linear transformation $\hat{\mathbf{C}}$ for the triclinic, monoclinic, and orthotropic material symmetries. Note that these forms are not here required to be symmetric, three of them are and five of them are not. However symmetry will be required for all of them by different and specialized physical arguments in subsequent chapters

Type of material symmetry	Form of the six-dimensional linear transformation $\hat{\mathbf{C}}$
Triclinic	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & \hat{c}_{14} & \hat{c}_{15} & \hat{c}_{16} \\ \hat{c}_{21} & \hat{c}_{22} & \hat{c}_{23} & \hat{c}_{24} & \hat{c}_{25} & \hat{c}_{26} \\ \hat{c}_{31} & \hat{c}_{32} & \hat{c}_{33} & \hat{c}_{34} & \hat{c}_{35} & \hat{c}_{36} \\ \hat{c}_{41} & \hat{c}_{42} & \hat{c}_{43} & \hat{c}_{44} & \hat{c}_{45} & \hat{c}_{46} \\ \hat{c}_{51} & \hat{c}_{52} & \hat{c}_{53} & \hat{c}_{54} & \hat{c}_{55} & \hat{c}_{56} \\ \hat{c}_{61} & \hat{c}_{62} & \hat{c}_{63} & \hat{c}_{64} & \hat{c}_{65} & \hat{c}_{66} \end{bmatrix}$
Monoclinic	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & \hat{c}_{14} & 0 & 0 \\ \hat{c}_{21} & \hat{c}_{22} & \hat{c}_{23} & \hat{c}_{24} & 0 & 0 \\ \hat{c}_{31} & \hat{c}_{32} & \hat{c}_{33} & \hat{c}_{34} & 0 & 0 \\ \hat{c}_{41} & \hat{c}_{42} & \hat{c}_{43} & \hat{c}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{55} & \hat{c}_{56} \\ 0 & 0 & 0 & 0 & \hat{c}_{65} & \hat{c}_{66} \end{bmatrix}$
Orthotropic	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & 0 & 0 & 0 \\ \hat{c}_{21} & \hat{c}_{22} & \hat{c}_{23} & 0 & 0 & 0 \\ \hat{c}_{31} & \hat{c}_{32} & \hat{c}_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{c}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{c}_{66} \end{bmatrix}$
Trigonal	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & \hat{c}_{14} & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{22} & \hat{c}_{23} & -\hat{c}_{14} & 0 & 0 \\ \hat{c}_{13} & \hat{c}_{23} & \hat{c}_{33} & 0 & 0 & 0 \\ \hat{c}_{41} & -\hat{c}_{41} & 0 & \hat{c}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{44} & \sqrt{2}\hat{c}_{14} \\ 0 & 0 & 0 & 0 & \sqrt{2}\hat{c}_{41} & (\hat{c}_{11} - \hat{c}_{12}) \end{bmatrix}$

derive the forms of the three-dimensional linear transformation $\hat{\mathbf{C}}$ for different material symmetries. The developments in this section parallel those in the previous section step for step. The matrices are different, and the results are different, but the arguments are identical. However, since this topic involves 6×6 rather than 3×3 matrices, a computational symbolic algebra program (e.g., Maple, Mathematica, Matlab, MathCad, etc.) is required to make the calculations simple. One should work through this section with such a program on a computer.

First, since triclinic symmetry has no planes of reflective symmetry, there are no symmetry restrictions for a triclinic material and the linear transformation $\hat{\mathbf{C}}$ is unrestricted. This conclusion is recorded in Tables 4.4 and 4.5. Monoclinic symmetry has exactly one plane of reflective symmetry. This means that the material

coefficients appearing in $\hat{\mathbf{C}}$ must be unchanged by one reflective symmetry transformation. Let $\mathbf{e}_1$ be the normal to the plane of reflective symmetry so that the reflective symmetry transformation is $\hat{\mathbf{R}}^{(\mathbf{e}_1)}$, given by the first of (4.5). The tensor $\hat{\mathbf{C}}$ must be invariant under the transformation

$$\hat{\mathbf{C}}^{(L)} = \hat{\mathbf{R}}^{(\mathbf{e}_1)} \cdot \hat{\mathbf{C}}^{(G)} \cdot [\hat{\mathbf{R}}^{(\mathbf{e}_1)}]^{\mathrm{T}}, \tag{4.16}$$

which follows from the first of (A162) by setting $\hat{\mathbf{Q}} = \hat{\mathbf{R}}^{(\mathbf{e}_1)}$. The pattern of this calculation follows the pattern of calculation in (4.9). That pattern is the substitution for $\hat{\mathbf{C}}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_1)}$ into this equation and the execution of the matrix multiplication. The resulting matrices are not documented here. As mentioned above, they may be easily obtained with any symbolic algebra program. The result is that the tensor $\hat{\mathbf{C}}$ is unchanged by the reflection only if $\hat{c}_{15} = \hat{c}_{51} = \hat{c}_{16} = \hat{c}_{61} = \hat{c}_{25} = \hat{c}_{52} = \hat{c}_{26} = \hat{c}_{62} = \hat{c}_{35} = \hat{c}_{53} = \hat{c}_{36} = \hat{c}_{63} = \hat{c}_{45} = \hat{c}_{54} = \hat{c}_{46} = \hat{c}_{64} = 0$. It follows then that the form of the tensor $\hat{\mathbf{C}}$ consistent with monoclinic symmetry characterized by a plane of reflective symmetry normal to the $\mathbf{e}_1$ base vector must satisfy the conditions $\hat{c}_{15} = \hat{c}_{51} = \hat{c}_{16} = \hat{c}_{61} = \hat{c}_{25} = \hat{c}_{52} = \hat{c}_{26} = \hat{c}_{62} = \hat{c}_{35} = \hat{c}_{53} = \hat{c}_{36} = \hat{c}_{63} = \hat{c}_{45} = \hat{c}_{54} = \hat{c}_{46} = \hat{c}_{64} = 0$. This result for monoclinic symmetry is recorded in Table 4.4.

Three mutually perpendicular planes of reflective symmetry characterize orthotropic symmetry, but the third plane is implied by the first two. This means that the material coefficients appearing in the representation of $\hat{\mathbf{C}}$ for monoclinic symmetry must be unchanged by another perpendicular reflective symmetry transformation. Let the $\mathbf{e}_2$ be the normal to the plane of reflective symmetry so the reflective symmetry transformation is $\hat{\mathbf{R}}^{(\mathbf{e}_2)}$ as given by the second of (4.5). The monoclinic form of the tensor $\hat{\mathbf{C}}$ must be invariant under the transformation

$$\hat{\mathbf{C}}^{(L)} = \hat{\mathbf{R}}^{(\mathbf{e}_2)} \cdot \hat{\mathbf{C}}^{(G)} \cdot [\hat{\mathbf{R}}^{(\mathbf{e}_2)}]^{\mathrm{T}}, \tag{4.17}$$

which follows from the first of (A162) by setting $\hat{\mathbf{Q}} = \hat{\mathbf{R}}^{(\mathbf{e}_2)}$. The pattern of this calculation follows the pattern of calculation in (4.11). That pattern is the substitution of the monoclinic form for $\hat{\mathbf{C}}$ and $\hat{\mathbf{R}}^{(\mathbf{e}_2)}$ into this equation and the execution of the matrix multiplication. The resulting matrices are not documented here; they may be easily obtained with any symbolic algebra program. The result is that the tensor $\hat{\mathbf{C}}$ is unchanged by the reflection only if $\hat{c}_{14} = \hat{c}_{41} = \hat{c}_{24} = \hat{c}_{42} = \hat{c}_{34} = \hat{c}_{43} = \hat{c}_{56} = \hat{c}_{65} = 0$. It follows then that the form of the tensor $\hat{\mathbf{C}}$ consistent with orthotropic symmetry characterized by planes of reflective symmetry normal to the $\mathbf{e}_1$ and $\mathbf{e}_2$ base vectors must satisfy the conditions $\hat{c}_{14} = \hat{c}_{41} = \hat{c}_{24} = \hat{c}_{42} = \hat{c}_{34} = \hat{c}_{43} = \hat{c}_{56} = \hat{c}_{65} = 0$. This result for orthotropic symmetry is recorded in Table 4.4.

A transversely isotropic material is one with a plane of isotropy. This means that the material coefficients appearing in the representation of $\hat{\mathbf{C}}$ for orthotropic

symmetry must be unchanged by any reflective symmetry transformation characterized by any unit vector in a specified plane. Let the designated plane of isotropy be the $\mathbf{e}_1$, $\mathbf{e}_2$ plane and let the unit vector be $\mathbf{a} = \cos\,\theta\,\mathbf{e}_1 + \sin\,\theta\,\mathbf{e}_2$ for any and all values of θ; then the reflective symmetry transformation of interest is $\hat{\mathbf{R}}^{(\theta_{12})}$, given by the first of (4.7). The orthotropic form of the tensor $\hat{\mathbf{C}}$ must be invariant under the transformation

$$\hat{\mathbf{C}}^{(L)} = \hat{\mathbf{R}}^{(\theta_{12})} \cdot \hat{\mathbf{C}}^{(G)} \cdot \left[\hat{\mathbf{R}}^{(\theta_{12})}\right]^{\mathrm{T}}, \tag{4.18}$$

which follows from the first of (A162) by setting $\hat{\mathbf{Q}} = \hat{\mathbf{R}}^{(\theta_{12})}$. The pattern of this calculation follows the pattern of calculation in (4.12). That pattern is the substitution of the orthotropic form for $\hat{\mathbf{C}}$ and $\hat{\mathbf{R}}^{(\theta_{12})}$ into this equation and the execution of the matrix multiplication. The resulting matrices are not documented here. They may be easily obtained with any symbolic algebra program. The result is that the tensor $\hat{\mathbf{C}}$ is unchanged by the reflection only if $\hat{c}_{22} = \hat{c}_{11},\ \hat{c}_{21} = \hat{c}_{12},\ \hat{c}_{23} = \hat{c}_{13},\ \hat{c}_{32} = \hat{c}_{31},\ \hat{c}_{55} = \hat{c}_{44},\ \hat{c}_{66} = \hat{c}_{11} - \hat{c}_{12}$. It follows then that the form of the tensor $\hat{\mathbf{C}}$ consistent with transversely isotropic symmetry characterized by a plane of isotropy whose normal is in $\mathbf{e}_3$ direction must satisfy the conditions $\hat{c}_{22} = \hat{c}_{11},\ \hat{c}_{21} = \hat{c}_{12},\ \hat{c}_{23} = \hat{c}_{13},\ \hat{c}_{32} = \hat{c}_{31},\ \hat{c}_{55} = \hat{c}_{44},\ \hat{c}_{66} = \hat{c}_{11} - \hat{c}_{12}$. This result for transversely isotropic symmetry is recorded in Table 4.5.

Isotropic symmetry is characterized by every direction being the normal to a plane of reflective symmetry, or equivalently, every plane being a plane of isotropy. This means that the material coefficients appearing in the representation of $\hat{\mathbf{C}}$ for transversely isotropic symmetry must be unchanged by any reflective symmetry transformation characterized by any unit vector in any direction. In addition to the $\mathbf{e}_1$, $\mathbf{e}_2$ plane considered for transversely isotropic symmetry it is required that the $\mathbf{e}_2$, $\mathbf{e}_3$ plane be a plane of isotropy. The second plane of isotropy is characterized by the unit vectors $\mathbf{a} = \cos\,\theta\,\mathbf{e}_2 + \sin\,\theta\,\mathbf{e}_3$ for any and all values of θ, then the reflective symmetry transformations of interest are $\hat{\mathbf{R}}^{(\theta_{23})}$ given by the second of (4.7). The form of the tensor $\hat{\mathbf{C}}$ representing transversely isotropic symmetry must be invariant under the transformation

$$\hat{\mathbf{C}}^{(\mathbf{L})} = \hat{\mathbf{R}}^{(\theta_{23})} \cdot \hat{\mathbf{C}}^{(\mathbf{G})} \cdot \left[\hat{\mathbf{R}}^{(\theta_{23})}\right]^{\mathrm{T}}, \tag{4.19}$$

which follows from the first of (A162) by setting $\hat{\mathbf{Q}} = \hat{\mathbf{R}}^{(\theta_{23})}$. The pattern of this calculation follows the pattern of calculation in (4.14) and (4.15). That pattern is the substitution of the transversely isotropic form for $\hat{\mathbf{C}}$ and $\hat{\mathbf{R}}^{(\theta_{23})}$ into this equation and the execution of the matrix multiplication. The resulting matrix is not recorded here; it may be easily obtained with any symbolic algebra program. The result is that the tensor $\hat{\mathbf{C}}$ is unchanged by any of the reflections whose normals lie in the

Table 4.5 The forms of the six-dimensional linear transformation $\hat{\mathbf{C}}$ for transversely isotropic and isotropic materials

Type of material symmetry	Form of the six-dimensional linear transformation $\hat{\mathbf{c}}$
Tetragonal	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & 0 & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{11} & \hat{c}_{13} & 0 & 0 & 0 \\ \hat{c}_{31} & \hat{c}_{31} & \hat{c}_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{c}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{c}_{66} \end{bmatrix}$
Transversely isotropic, hexagonal	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & 0 & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{11} & \hat{c}_{13} & 0 & 0 & 0 \\ \hat{c}_{31} & \hat{c}_{31} & \hat{c}_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{c}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{c}_{11} - \hat{c}_{12} \end{bmatrix}$
Cubic	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{12} & 0 & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{11} & \hat{c}_{12} & 0 & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{12} & \hat{c}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{c}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{c}_{44} \end{bmatrix}$
Isotropic	$\begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{12} & 0 & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{11} & \hat{c}_{12} & 0 & 0 & 0 \\ \hat{c}_{12} & \hat{c}_{12} & \hat{c}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{c}_{11} - \hat{c}_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{c}_{11} - \hat{c}_{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{c}_{11} - \hat{c}_{12} \end{bmatrix}$

plane perpendicular to $\mathbf{e}_1$ only if the transversely isotropic form for $\hat{\mathbf{C}}$ satisfies the additional restrictions: $\hat{c}_{33} = \hat{c}_{11},\ \hat{c}_{31} = \hat{c}_{12},\ \hat{c}_{31} = \hat{c}_{12},\ \hat{c}_{13} = \hat{c}_{12},\ \hat{c}_{44} = \hat{c}_{11} - \hat{c}_{12}$. It follows then that the transversely isotropic form of the tensor $\hat{\mathbf{C}}$ consistent with isotropic symmetry characterized by two perpendicular planes of isotropy must satisfy the conditions $\hat{c}_{33} = \hat{c}_{11},\ \hat{c}_{31} = \hat{c}_{12},\ \hat{c}_{31} = \hat{c}_{12},\ \hat{c}_{13} = \hat{c}_{12},\ \hat{c}_{44} = \hat{c}_{11} - \hat{c}_{12}$. This result for isotropic symmetry is recorded in Table 4.4. Algebraic procedure identical with those described above may be used to obtain the forms of the tensor $\hat{\mathbf{C}}$ consistent with the trigonal, tetragonal, and cubic symmetries listed in Table 4.4.

Problems

4.7.1 Show that the representation for orthotropic symmetry in Table 4.4 is also invariant under the transformation $\hat{\mathbf{R}}^{(\mathbf{e}_3)}$ constructed in Problem 4.4.2.

4.7.2 Show that the representation for isotropic symmetry in Table 4.4 is also invariant under the transformation $\hat{\mathbf{R}}^{(\theta_{13})}$ constructed in Problem 4.4.3.

4.7.3 Construct a representation for $\hat{\mathbf{C}}$ that is invariant under the seven reflective transformations formed from the set of normals to the planes of reflective symmetry given in Problem 4.5.2: $\mathbf{k}$, $\mathbf{p}$, $\mathbf{m}$, $\mathbf{n}$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$. Does the result coincide with one of the representations already in Table 4.4? If it does, please explain.

4.8 Curvilinear Anisotropy

In the case where the type of textured material symmetry is the same at all points in an object, it is still possible for the normals to the planes of mirror symmetry to rotate as a path is traversed in the material. This type of anisotropy is referred to as *curvilinear anisotropy*. The cross-section of a tree illustrated in Fig. 4.5 and the nasturtium petiole in Fig. 4.6 have curvilinear anisotropy. At any point the tree has orthotropic symmetry, but as a path across a cross-section of the tree is followed, the normals to the planes of symmetry rotate. In the cross-section the normals to the planes of symmetry are perpendicular and tangent to the growth rings. Curvilinear anisotropy, particularly curvilinear orthotropy, and curvilinear transverse isotropy are found in many man-made materials and in biological materials. Wood and plane tissue are generally curvilinear orthotropic, as are fiber wound composites. Only textured symmetries can be curvilinear. Crystalline symmetries are rectilinear, that is to say the planes of symmetry cannot rotate as a linear path is traversed in the material.

Curvilinear anisotropies such as those based on the ideal cylindrical and spherical coordinate systems may have mathematical singularities. For example, a curvilinear orthotropy characterized by the ideal cylindrical coordinate system has a singularity at the origin (Tarn 2002). This is due to the fact that the modulus associated with the radial direction is different from that associated with the circumferential or hoop direction (c.f., Fig. 4.5). The singularity at the origin arises due to the fact that the radial direction and the circumferential or hoop direction are indistinguishable at the origin yet they have different moduli. A simple resolution of the mathematical singularity in the model is possible with the proper physical interpretation of its significance in the real material. One such proper physical interpretation of such singular points is to note that a volume element containing such a singular point is *not* a typical RVE and must be treated in a special manner. This is basically the approach of Tarn (2002) who constructs a special volume element, with transversely isotropic symmetry, enclosing the singularity in the cylindrical coordinate system. The mathematical singularity in the model is, in this way, removed and the model corresponds more closely to the real material.

Problems

4.8.1 Sketch the curvilinear nature of the set of three normals to the planes of reflective symmetry that characterize the wood tissue of a tree.

4.9 Symmetries that Permit Chirality

Thus far in the consideration of material symmetries the concern has been with the number and orientation of the planes of material symmetry. In this section the consideration is of those material symmetries that have planes that are *not* normals to planes of reflective symmetry. The triclinic, monoclinic, and trigonal symmetries are the only three of the eight elastic symmetries that permit directions that are not normals to planes of reflective symmetry. Every direction in triclinic symmetry is a direction in which a normal to the plane of material symmetry is *not* permitted. Every direction that lies in the single symmetry plane in monoclinic symmetry is a direction in which a normal to the plane of material symmetry is *not* permitted. The only direction in trigonal symmetry in which a normal to the plane of material symmetry is not permitted is the direction normal to a plane of threefold symmetry. There are *not* other such directions. The triclinic, monoclinic, and trigonal symmetries are also the only three of the eight elastic symmetries that, in their canonical symmetry coordinate system, retain cross-elastic constants connecting normal stresses (strains) to shear strains (stresses) and vice versa. In the $\hat{\mathbf{C}}$ matrices listed in Table 4.4 these cross-elastic constants appear in the lower left and upper right 3×3 sub-matrices for the triclinic, monoclinic, and trigonal symmetries. In the $\mathbf{A}$ matrices listed in Table 4.3 only monoclinic symmetry has a cross-elastic constant. The nonzero cross-elastic constants and the directions that are not normals to planes of reflective symmetry are directly related; such planes disappear when the cross-elastic constants are zero. It is the existence of such planes and associated cross-elastic constants that allow structural gradients and handedness (chirality).

Trigonal symmetry, because it is the highest symmetry of the three symmetries, admits a direction that is not a direction associated with a normal to a plane of reflective symmetry, nor any projected component of a normal to a plane of reflective symmetry. An interesting aspect of trigonal symmetry is the chiral or symmetry-breaking character of the cross-elastic constant $\hat{c}_{14}$. Note that $\hat{c}_{14}$ is not constrained to be of one sign; the sign restriction on $\hat{c}_{14}$ from the positive definiteness of strain energy is

$$-\sqrt{\frac{\hat{c}_{44}(\hat{c}_{11}-\hat{c}_{12})}{2}} < \hat{c}_{14} < \sqrt{\frac{\hat{c}_{44}(\hat{c}_{11}-\hat{c}_{12})}{2}}. \tag{4.20}$$

If $\hat{c}_{14}$ vanishes, the $\hat{\mathbf{C}}$ matrix in Table 4.4 for trigonal symmetry specializes to the $\hat{\mathbf{C}}$ matrix in Table 4.4 for hexagonal or transversely isotropic symmetry. Hexagonal symmetry is a sixfold symmetry with seven planes of mirror symmetry. Six of the normals to these seven planes all lie in the seventh plane and make angles of 30° with one another. A single plane of isotropy characterizes transversely isotropic symmetry. A *plane of isotropy* is a plane of mirror symmetry in which every vector is itself a normal to a plane of mirror symmetry. Since a *plane of isotropy* is also a plane of symmetry, there are an infinity plus one planes of symmetry associated with transverse isotropy.

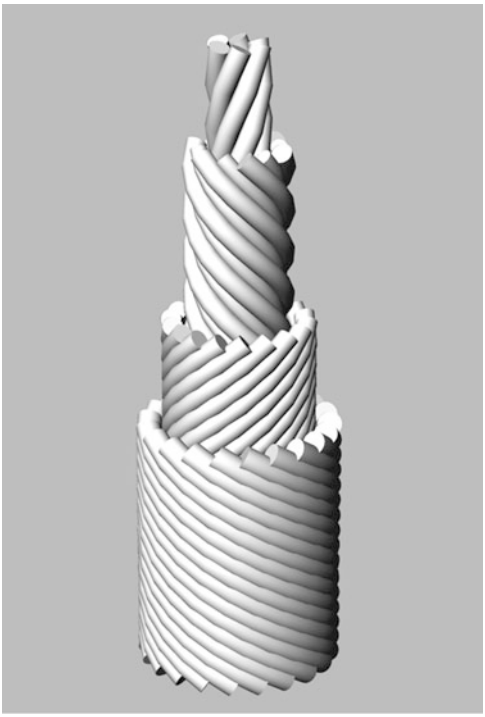


Fig. 4.14 A composite structure composed from sets of left- and right-handed helically wound fibrous laminae that are in the form of concentric coaxial cylinders; the fibers of each lamina or component cylinder are characterized by a different helical angle. The angle of the helices often rotates regularly from one cylinder to the next. This type of structure is called "helicoidal" and described as a cylinder made of "twisted plywood". From Fraldi and Cowin (2002)

A simple thought model is possible for the visualization of the symmetry-breaking character of the elastic constant $\hat{c}_{14}$. This constant could be described as a chiral constant, chiral being a word coined by Kelvin (Thompson 1904) (*I call any geometrical figure, or group of points, chiral, and say it has chirality, if its image in a plane mirror, ideally realized, cannot be brought to coincide with itself.*) and widely used in describing the structure of molecules. It means that a structure cannot be superposed on its mirror image, that the structure has a handedness. For example, helical spirals are chiral; they are either left-handed or right-handed. A composite structure of alternate left- and right-handed helical spirals is illustrated in Fig. 4.14. Consider a composite material constructed of an isotropic matrix material reinforced by only right-handed spiral helices whose long axes are all parallel. These helical spirals may be either touching or separated by a matrix material (Fig. 4.15). Let the helical angle be θ (Fig. 4.15) and let negative values of θ correspond to otherwise similar left-handed helices; the vanishing of θ then corresponds to a straight reinforcement fiber. Assume that when the effective elastic constants for this material are calculated, the sign of $\hat{c}_{14}$ is determined by the sign of θ and vanishes when θ is zero. It is then possible to geometrically visualize the chiral, symmetry-breaking character of $\hat{c}_{14}$ as it passes from positive to negative (or negative to positive) values through zero as the vanishing of a helical angle of one handedness occurs and the initiating of a helical angle of the opposite handedness commences. At the dividing line between the two types of handedness, the reinforcing fibers are straight. In terms of the elastic material

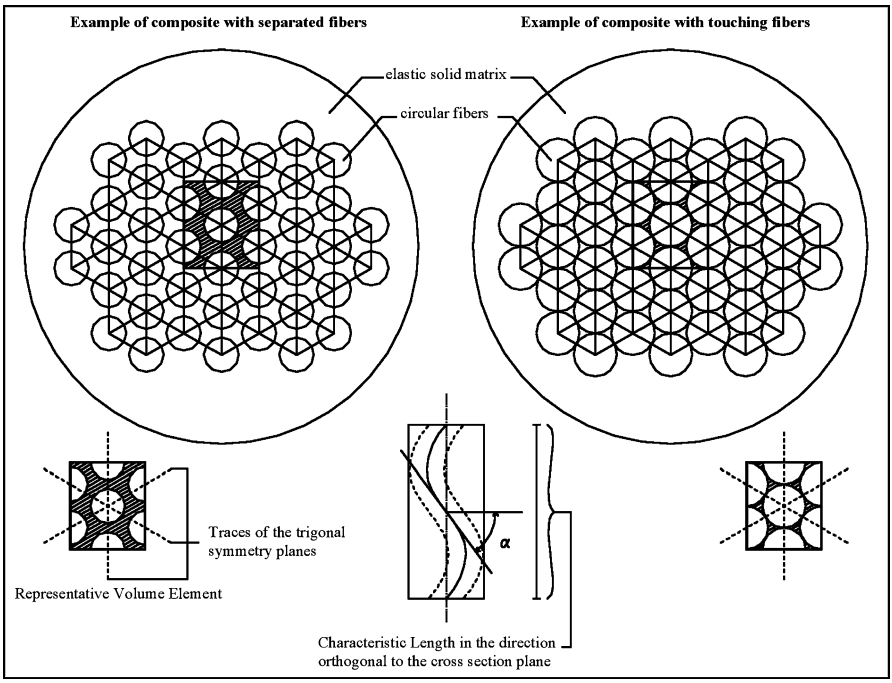


Fig. 4.15 Two examples of composites formed from an elastic matrix and embedded helicoidal circular fibers. The two large circular diagrams illustrate the characteristic threefold trigonal symmetry in the cross-sectional plane. The possible representative volume elements (RVEs) and their characteristic helical angles are also illustrated. From Fraldi and Cowin (2002)

symmetry, as $\hat{c}_{14}$ passes from positive to negative (or negative to positive) values through zero, the elastic material is first a trigonal material of a certain chirality, then a transversely isotropic (or hexagonal) material, and then a trigonal material of an opposite chirality. A RVE of the composite in Fig. 4.14 may be constructed using a set of the helicoidal fibers all having identical circular cross-sections and using the periodicity of the helix (Fig. 4.15, inset). This construction provides an RVE with a material neighborhood large enough to adequately average over the microstructure and small enough to ensure that the structural gradient across it is negligible. An examination of Fig. 4.15 shows that, in the plane orthogonal to the x_3 axis, the threefold symmetry characteristic of trigonal symmetry arises naturally.

This example illustrates how chirality is created in a material with a helical structure. It also demonstrates that the symmetry-breaking chiral elastic constant $\hat{c}_{14}$ in trigonal symmetry is related to the angle of the helical structure of the material, if the material has a helical structure. Further, it again illustrates how different levels of RVE's are associated with different types of material symmetry. In this example the smaller RVE is associated with orthotropic material symmetry and the larger RVE (obtained by volume averaging over the domain of the smaller RVE) is associated with monoclinic symmetry. The result demonstrates that a

material symmetry that permits chirality (i.e., trigonal, monoclinic, or triclinic symmetry) is obtained by averaging over a domain that is characterized by a symmetry that does not permit chirality (i.e., isotropic, cubic, transverse isotropic, tetragonal, and orthotropic). Clearly the result presented depends on the fact that the (non-chiral) orthotropic material symmetry is helically curvilinear. The association of the micro-geometric chiral character of $\hat{c}_{14}$ with a helix is not a unique association. The basic property of $\hat{c}_{14}$ is its symmetry-breaking character, and it may be associated with structural gradients in the material (Cowin 2002; Fraldi and Cowin 2002).

There are many natural and man-made examples of both chiral materials as structures and as local components in globally non-chiral composites. Chiral materials that form chiral structures occur in nature (Neville 1993). Perhaps the most famous is the tusk of the narwhal (in the middle ages the tusk of the narwhal was thought to be the horn of the mythical unicorn). This whale is edentulous except for the upper lateral incisors. The right incisor normally remains embedded in the jaw, but in adult males the left tooth forms a tusk, which can in large specimens reach a length of 2.4 m, and have a diameter of 8 cm at the point of eruption. Normally the tusk is imprinted with the curvature of the bone socket as it erupts or extrudes itself from a bone socket. However, if the tusk slowly twists in the socket as it grows, the imprinted curvature will be neutralized or averaged and the tusk will grow straight with the spiral structure. A second example of a natural chiral structure occurs in trees, both hardwoods and softwoods, due to a combination of genetic and environmental factors. The spiral structure in trees causes a practical problem with telephone and power poles. Changes in the moisture content of the wood of the pole cause the pole to twist after it has been employed as part of a transmission network.

Chiral materials that form chiral and non-chiral structures occur very frequently in nature (Neville 1993). A typical such natural structure is illustrated in Fig. 4.14. The structure is a set of concentric coaxial cylinders, each lamina or cylinder characterized by a different helical angle. The angle of the helices often rotates regularly from one cylinder to the next. This type of structure is called helicoidal and described as a cylinder made of "twisted plywood" (see Figs. 10.33 and 10.34). The helical fibers may or may not be touching, as illustrated in Fig. 4.15. The examples of this structure in nature are numerous and include fish scales and plant stem walls. Man also uses cylinders made of "twisted plywood" to create structures.

4.10 Relevant Literature

A very interesting and perceptive book on symmetry in general, but including all the symmetries of interest in the present text, is the book of Weyl (1952) with the title *Symmetry*. Material symmetry is well explained in the books of Nye (1957) and Fedorov (1968). The treatment of material symmetry in this chapter does not follow the standard treatments of material symmetry contained, for example, in the books

on anisotropic elasticity mentioned in the literature notes at the end of Chapter 6. In the first section of this chapter it was stated that one of the purposes of this chapter was to obtain and record representations of $\mathbf{A}$ and $\hat{\mathbf{C}}$ that represent the effects of material symmetry. This has been done and the results are recorded in Tables 4.3, 4.4 and 4.5 for $\mathbf{A}$ and $\hat{\mathbf{C}}$, respectively. The representations in these tables were developed with the minimum algebraic manipulation invoked and with the minimum rigor but, we think, with the most concise method. The method employed to obtain these representations is new. Cowin and Mehrabadi (1987) first pursued using reflections to characterize the elastic symmetries. Later Cowin and Mehrabadi (1995) developed all the linear elastic symmetries from this viewpoint. Here we have not generally bothered to demonstrate invariance of these representations for all the reflective symmetries that they enjoy. In this presentation we have been content to use the minimum number of reflective symmetries necessary to obtain the symmetry representations. So we do not show the reader that further admissible reflective symmetries will not alter the representation obtained, but it happens to be true.

The established and more general method to obtain these representations is to use the symmetry groups associated with each of these symmetries. The disadvantage of that path in the present text is that it takes too much text space and distracts the reader from the main topic because it requires the introduction of the group concept and then the concept of the material symmetry group, etc. In the method of presentation employed, only the concept of the plane of mirror or reflective symmetry is necessary. The reflective symmetry approach is equivalent to the group theory approach for $\mathbf{A}$ and $\hat{\mathbf{C}}$ tensors (Chadwick et al. 2001), but it may not be so for nonlinear relationships between these tensors. It does, however, not provide all the representations for the second rank $\mathbf{A}$ if $\mathbf{A}$ is not symmetric (Cowin 2003).

References

Chadwick P, Vianello M, Cowin SC (2001) A proof that the number of linear anisotropic elastic symmetries is eight. J Mech Phys Solids 49:2471–2492

Cowin SC (1995) On the number of distinct elastic constants associated with certain anisotropic elastic symmetries. In Casey J, Crochet MJ (ed) Theoretical, experimental, and numerical contributions to the mechanics of fluids and solids—a collection of papers in honor of Paul M. Naghdi, Z Angew Math Phys (Special Issue) 46:S210–S214

Cowin SC (1999) Bone poroelasticity. J Biomech 32:218–238

Cowin SC (2002) Elastic symmetry restrictions from structural gradients. In: Podio-Guidugli P, Brocato M (eds) Rational continua, classical and new—a collection of papers dedicated to Gianfranco Capriz on the occasion of his 75th birthday. Springer Verlag, New York. ISBN 88-470-0157-9

Cowin SC (2003) Symmetry plane classification criteria and the symmetry group classification criteria are not equivalent in the case of asymmetric second order tensors. Chinese J Mech 19:9–14

Cowin SC, Mehrabadi MM (1987) On the identification of material symmetry for anisotropic elastic materials. Quart J Mech Appl Math 40:451–476
Cowin SC, Mehrabadi MM (1989) Identification of the elastic symmetry of bone and other materials. J Biomech 22:503–515
Cowin SC, Mehrabadi MM (1995) Anisotropic symmetries of linear elasticity. Appl Mech Rev 48:247–285
Fedorov FI (1968) Theory of elastic waves in crystals. Plenum Press, New York
Fraldi M, Cowin SC (2002) Chirality in the torsion of cylinders with trigonal material symmetry. J Elasticity 69:121–148
Hull D (1981) An introduction to composite materials. Cambridge University Press, Cambridge
Neville C (1993) Biology of fibrous composites. Cambridge University Press, Cambridge
Nye JF (1957) Physical properties of crystals. Clarendon, Oxford
Rovati M, Taliercio A (2003) Stationarity of the strain energy density for some classes of anisotropic solids. Int J Solids Struct 40:6043–6075
Shubnikov AV, Koptsik VA (1974) Symmetry in science and art. Plenum, New York
Tarn JQ (2002) Stress singularity in an elastic cylinder of cylindrically anisotropic materials. J Elasticity 69:1–13
Thompson W (Lord Kelvin) (1904) Baltimore lectures on molecular dynamics and the wave theory of light. Clay & Sons, London
Wainwright SA, Biggs WD, Currey JD, Gosline JM (1976) Mechanical Design in Organism, Edward Arnold, London
Weyl H (1952) Symmetry. Princeton University Press, Princeton

Chapter 5
Formulation of Constitutive Equations

The theme for this chapter is contained in a quote from Truesdell and Noll's volume on the *Non-Linear Field Theories of Mechanics*: *"The general physical laws in themselves do not suffice to determine the deformation or motion of . an object. subject to given loading. Before a determinate problem can be formulated, it is usually necessary to specify the material of which the. . .object. . .is made. In the program of continuum mechanics, such specification is stated by constitutive equations, which relate the stress tensor and the heat-flux vector to the motion. For example, the classical theory of elasticity rests upon the assumption that the stress tensor at a point depends linearly on the changes of length and mutual angle suffered by elements at that point, reckoned from their configurations in a state where the external and internal forces vanish, while the classical theory of viscosity is based on the assumption that the stress tensor depends linearly on the instantaneous rates of change of length and mutual angle. These statements cannot be universal laws of nature, since they contradict one another. Rather, they are definitions of ideal materials. The former expresses in words the constitutive equation that defines a linearly and infinitesimally elastic material; the latter, a linearly viscous fluid. Each is consistent, at least to within certain restrictions, with the general principles of continuum mechanics, but in no way a consequence of them. There is no reason a priori why either should ever be physically valid, but it is an empirical fact, established by more than a century of test and comparison, that each does indeed represent much of the mechanical behavior of many natural substances of the most various origin, distribution, touch, color, sound, taste, smell, and molecular constitution. Neither represents all the attributes, or suffices even to predict all the mechanical behavior, of any one natural material. No natural. object. is perfectly elastic, or perfectly fluid, any more than any is perfectly rigid or perfectly incompressible. These trite observations do not lessen the worth of the two particular constitutive equations just mentioned. That worth is twofold: First, each represents in ideal form an aspect, and a different one, of the mechanical behavior of nearly all natural materials, and, second, each does predict with considerable though sometimes not sufficient accuracy the observed response of many different natural materials in certain restricted situations."*

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_5,

5.1 Guidelines for the Formulation of Constitutive Equations

The conservation principles of mass, linear momentum, angular momentum, and energy do not yield, in general, a sufficient number of equations to determine all the unknown variables for a physical system. These conservation principles must hold for all materials and therefore they give no information about the particular material of which the system is composed, be it fluid or solid, bone, concrete or steel, blood, oil, honey or water. Additional equations must be developed to describe the material of the system and to complete the set of equations involving the variables of the system so that the set of equations consisting of these additional equations and the conservation equations are solvable for the variables.

Equations that characterize the physical properties of the material of a system are called constitutive equations. Each material has a different constitutive equation to describe each of its physical properties. Thus there is one constitutive equation to describe the mechanical response of steel to applied stress and another to describe the mechanical response of water to applied stress. Constitutive equations are contrasted with conservation principles in that conservation principles must hold for all materials while constitutive equations only hold for a particular property of a particular material. The purpose of this chapter is to present the guidelines generally used in the formulation of constitutive equations, and to illustrate their application by developing four classical continuum constitutive relations, namely Darcy's law for mass transport in a porous medium, Hooke's law for elastic materials, the Newtonian law of viscosity, and the constitutive relations for viscoelastic materials.

5.2 Constitutive Ideas

The basis for a constitutive equation is a constitutive idea, that is to say an idea taken from physical experience or experiment that describes how real materials behave under a specified set of conditions. For example, the constitutive idea of the elongation of a bar being proportional to the axial force applied to the ends of the bar is expressed mathematically by the constitutive equation called Hooke's law. Another example of a constitutive idea is that, in a saturated porous medium the fluid flows from regions of higher pressure to regions of lower pressure; this idea is expressed mathematically by the constitutive equation called Darcy's law for fluid transport in a porous medium. It is not a simple task to formulate a constitutive equation from a constitutive idea. The constitutive idea expresses a notion concerning some aspect of the behavior of real materials, a notion based on the physics of the situation that might be called physical insight. The art of formulating constitutive equations is to turn the physical insight into a mathematical equation. The conversion of insight into equation can never be exact because the equation is precise and limited in the amount of information it can embody while the constitutive idea is embedded in one's entire understanding of the physical situation.

The fine art in the formulation of constitutive equations is to extract the salient constitutive idea from the physical situation under consideration and then to express its essence in an equation.

The four classical constitutive ideas employed here as examples are described in this paragraph. Darcy's law for mass transport in a porous medium may be considered as arising from the idea that, in a saturated porous medium, fluid flows from regions of higher pressure to those of lower pressure. Let ρ_{f} denote the density of the fluid in the pores of the porous medium, ρ_{o} denote a constant reference fluid density, and ϕ denote the porosity of the medium. The velocity of the fluid $\mathbf{v}$ passing through the pores is the velocity relative to the solid porous matrix. This constitutive idea is that the fluid volume flux $\mathbf{q} = \phi\rho_{\mathrm{f}}\mathbf{v}/\rho_{\mathrm{o}}$ through the pores, at a particle $\mathbf{X}$, is a function of the pressure variation in the neighborhood of $\mathbf{X}$, $N(\mathbf{X})$. If $p(\mathbf{X}, t)$ represents the pressure at the particle $\mathbf{X}$ at time t, then this constitutive idea is expressed as

$$\mathbf{q} = \phi\rho_{\mathrm{f}}\mathbf{v}/\rho_{\mathrm{o}} = \mathbf{q}(p(\mathbf{X},\ t), \mathbf{X}), \quad \text{all } \mathbf{X}^{*}\text{in } N(\mathbf{X}) \tag{5.1D}$$

Note that $\mathbf{q}$ has the dimensions of volume flow per unit area, which means it is the volume flow rate of fluid across a certain surface area. The volume flow rate $\mathbf{q}$ is the flow rate relative to the solid porous matrix. The constitutive idea for Fourier's law of heat conduction and Fick's law for diffusion of a solute in a solvent have the same mathematical structure as Darcy's law for mass transport in a porous medium. The constitutive idea for Fourier's law of heat conduction is that heat flows from regions of higher temperature to those of lower temperature. The constitutive idea for Fick's law for diffusion of a solute in a solvent is that a solute diffuses from regions of higher solute concentration to those of lower solute concentration. The developments of the Fourier law and the Fick law are parallel to the development of Darcy's law. For the Fourier law the heat flux vector replaces the volume flux per unit area $\mathbf{q}$ and the temperature replaces the pressure. For Fick's law, for diffusion of a solute in a solvent, the diffusion flux vector replaces the volume flux per unit area $\mathbf{q}$ and the pressure is replaced by the concentration of the solute in the solvent. These substitutions will extend most of what is recorded in this chapter about Darcy's law to the Fourier law and the Fick law.

The D in the equation number above is to indicate that this equation is associated with Darcy's law. In this chapter D, H, N, and V will be used in equation numbers to indicate the constitutive concept that the equation is associated; D is for Darcy, H is for Hooke, N is for Newton, and V is for viscoelastic.

In the development of the remaining constitutive equations, those that assume that stress is a function of different kinematic variables, the stress will be denoted as a vector in six dimensions, $\hat{\mathbf{T}}$, rather than a tensor in three dimensions, $\mathbf{T}$ (see Sect. A.11). The six-dimensional representation has advantages in the formulation of constitutive equations. The main advantage in the present chapter is that all the constitutive ideas to be developed will then have a similar structure except that some will be in three dimensions and the rest in six dimensions. The constitutive

idea for Hooke's law is that of a spring. If a force displaces the end of a spring, there is a relationship between the force and the resulting displacement. Thus, to develop Hooke's law, the stress $\hat{\mathbf{T}}$ at a particle $\mathbf{X}$ is expressed as a function of the variation in the displacement field $\mathbf{u}(\mathbf{X}, t)$ in the neighborhood of $\mathbf{X}$, $N(\mathbf{X})$,

$$\hat{\mathbf{T}} = \hat{\mathbf{T}}(\mathbf{u}(\mathbf{X},\ \mathrm{t}), \mathbf{X}), \quad \text{all } \mathbf{X}^* \text{in } N(\mathbf{X}). \tag{5.1H}$$

The constitutive idea for the Newtonian law of viscosity is that of the dashpot or damper, namely that the force is proportional to the rate at which the deformation is accomplished rather than to the size of the deformation itself. The total stress in a viscous fluid is the sum of the viscous stresses $\mathbf{T}_\mathrm{v}$ plus the fluid pressure p, $\mathbf{T} = -p\mathbf{1} + \mathbf{T}_\mathrm{v}$. The constitutive idea for the Newtonian law of viscosity is that the stress $\hat{\mathbf{T}}_\mathrm{v}$, due to the viscous effects at a particle $\mathbf{X}$, is expressed as a function of the variation in the velocity field $\mathbf{v}(\mathbf{X}, t)$ in the neighborhood of $\mathbf{X}$, $N(\mathbf{X})$. The expression for the total stress in a fluid is the pressure plus the viscous stresses.

$$\hat{\mathbf{T}} = -p\hat{\mathbf{U}} + \hat{\mathbf{T}}_v(\mathbf{v}(\mathbf{X}, t), \mathbf{X}), \quad \text{all } \mathbf{X}^* \text{in} N(\mathbf{X}). \tag{5.1N}$$

Recall that $\hat{\mathbf{U}}$ is the six-dimensional vector with components $\{1, 1, 1, 0, 0, 0\}$; it is the image of the three-dimensional unit tensor $\mathbf{1}$ in six dimensions. Each of the four constitutive ideas described yields the value of a flux or stress at time t due to the variation in a field (temperature, pressure, displacement, velocity) at the particle $\mathbf{X}$ at time t. The constitutive idea for viscoelasticity is different in that the stress at time t is assumed to depend upon the entire history of a field, the displacement field. Thus, while the first three constitutive ideas are expressed as functions, the constitutive idea for viscoelasticity is expressed as a functional of the history of the displacement field. A functional is like a function, but rather than being evaluated at a particular value of its independent variables like a function, it requires an entire function to be evaluated; a functional is a function of function(s). An example of a functional is the value of an integral in which the integrand is a variable function. The constitutive idea for a viscoelastic material is that the stress $\hat{\mathbf{T}}$ at a particle $\mathbf{X}$ is a function of the variation in the history of the velocity field $\mathbf{v}(\mathbf{X}, t)$ in the neighborhood of $\mathbf{X}$, $N(\mathbf{X})$,

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{T}}(\mathbf{v}(\mathbf{X}, t-s),\ \mathrm{s},\ \mathbf{X})\mathrm{d}s, \quad \text{for all } \mathbf{X}^* \text{ in } N(\mathbf{X}), \tag{5.1V}$$

where s is a backward running time variable that is 0 at the present instant and increases with events more distant in the past. Thus the stress $\hat{\mathbf{T}}$ at a particle $\mathbf{X}$ is a function of the entire history of the displacement of the particle; to evaluate the stress, knowledge of the entire history is required. In the sections that follow this one these four constitutive ideas will be developed into linear constitutive equations.

Problem

5.2.1. Record a rigorous definition of the neighborhood of the particle **X**, $N(\mathbf{X})$, using a reference particle X*, a small positive number e and a length measure $\|\mathbf{X} - \mathbf{X}^*\|$. Note that this neighborhood is, in general, a three-dimensional neighborhood.

5.3 Localization

A constitutive equation valid at the particle **X** of a material object can depend upon the behavior of the material in the neighborhood of the particle **X**, $N(\mathbf{X})$, but is unlikely to depend upon the behavior of the material in regions of the object far removed from the particle **X**. The localization guideline for the development of constitutive relations restricts the dependence of constitutive equations valid for a particle **X** to events that occur in $N(\mathbf{X})$. The application of the localization guideline to the four constitutive equations described in the previous section is described in the next paragraph.

The constitutive idea for Darcy's law is considered first. The pressure field $p(\mathbf{X}^*, t)$ at a particle $\mathbf{X}^*$ in $N(\mathbf{X})$ may be related to the pressure field $p(\mathbf{X}, t)$ at a particle **X** by a Taylor series expansion about the point by

$$p(\mathbf{X}^*,\ t) = \mathrm{p}(\mathbf{X},\ t) + (\nabla p(\mathbf{X},\ t)) \cdot (\mathbf{X} - \mathbf{X}^*) + \text{higher order terms}, \tag{5.2}$$

where it is assumed that the pressure field is sufficiently smooth to permit this differentiation. With the Taylor theorem as justification, the $N(\mathbf{X})$ may always be selected sufficiently small so the value of the pressure field $\mathrm{p}(\mathbf{X}^*, t)$ at a particle $\mathbf{X}^*$ in $N(\mathbf{X})$ may be represented by $p(\mathbf{X}, t)$ and $\nabla p(\mathbf{X}, t)$. Thus, by localization, (5.1D) may be rewritten as

$$\mathbf{q} = \phi \rho_{\mathrm{f}} \mathbf{v} / \rho_{\mathrm{o}} = \mathbf{q}(p(\mathbf{X},\ t), \nabla p(\mathbf{X},\ t), \mathbf{X}) \tag{5.3D}$$

Exactly the same argument is applicable to the other three constitutive ideas; thus we have that

$$\hat{\mathbf{T}} = \hat{\mathbf{T}}(\mathbf{u}(\mathbf{X},\ t), \nabla \otimes \mathbf{u}(\mathbf{X},\ t), \mathbf{X}), \tag{5.3H}$$

$$\hat{\mathbf{T}} = -p\hat{\mathbf{U}} + \hat{\mathbf{T}}_{\mathrm{v}}(\mathrm{v}(\mathbf{X},\ t), \nabla \otimes \mathrm{v}(\mathbf{X},\ t), \mathbf{X}), \tag{5.3N}$$

and

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{T}}(\mathrm{v}(\mathbf{X},\ t-s), \nabla \otimes \mathrm{v}(\mathbf{X},\ t-s), \mathbf{X}, s)\mathrm{d}s. \tag{5.3V}$$

Problem

5.3.1. Record a complete statement of Taylor's theorem in the case of three independent variables for the function $f(\mathbf{X})$ using the point $\mathbf{X}_o$ as the point about which the expansion occurs.

5.4 Invariance Under Rigid Object Motions

This guideline for the development of constitutive relations restricts the independent variables and functional dependence of constitutive equations for material behavior by requiring that the constitutive equations be independent of the motions of the object that do not deform the object. The motions of the object that do not deform the object are rigid object motions. This guideline requires that constitutive equations for material behavior be independent of, that is to say unchanged by, superposed rigid object motions. As an illustration consider the object shown in Fig. 5.1. If the object experiences a translation and a rigid object rotation such that the force system acting on the object is also translated and rotated, then the state of stress $\mathbf{T}(\mathbf{X}, t)$ at any particle $\mathbf{X}$ is unchanged. As a second example recall that the volume flow rate $\mathbf{q}$ is the flow rate relative to the solid porous matrix. It follows that the volume flow rate $\mathbf{q}$ in a porous medium is unchanged by (virtual or very slow) superposed rigid object motions.

The application of this guideline of invariance under rigid object motions is illustrated by application to the three constitutive ideas involving stress. The two constitutive ideas involving fluxes automatically satisfy this guideline because the fluxes are defined relative to the material object and the rigid motion does not change the temperature field or the pressure field. The constitutive idea for Hooke's law (5.3H) may be rewritten as

$$\hat{\mathbf{T}} = \hat{\mathbf{T}}(\mathbf{u}(\mathbf{X},\ t), \mathbf{E}(\mathrm{X},\ t), \mathbf{Y}(\mathbf{X},\ t), \mathbf{X}), \tag{5.4H}$$

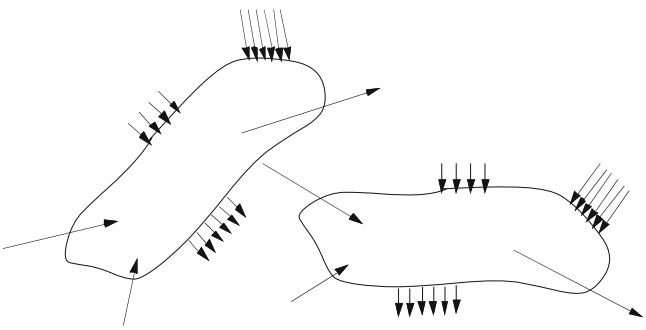


Fig. 5.1 A rigid object rotation of an object, a rotation that includes the force system that acts upon the object

where the tensor of displacement gradients $\nabla \otimes \mathbf{u}(\mathbf{X}, t)$ has been replaced by its symmetric part, the infinitesimal strain tensor $\mathbf{E}(\mathbf{X}, t)$ defined by (2.49), and its skew symmetric part, the infinitesimal rotation tensor $\mathbf{Y}(\mathbf{X}, t)$, also defined by (2.49). In a similar way the tensor of velocity gradients $\nabla \otimes \mathbf{v}(\mathbf{X}, t)$ may be replaced by its symmetric part, the rate of deformation tensor $\mathbf{D}(\mathbf{X}, t)$ defined by (2.32), and its skew symmetric part, the spin tensor $\mathbf{W}(\mathbf{X}, t)$, also defined by (2.32).

$$\hat{\mathbf{T}} = -p\hat{\mathbf{U}} + \hat{\mathbf{T}}_{\mathrm{v}}(\mathrm{v}(\mathbf{X},\ t), D(\mathbf{X},\ t),\ W(\mathbf{X},\ t)\mathbf{X}). \tag{5.4N}$$

Finally, decomposing the tensor of velocity gradients $\nabla \otimes \mathbf{v}(\mathbf{X}, t-\mathrm{s})$ as in the case of the Newtonian law of viscosity, the viscoelastic constitutive relation (5.3V) takes the form

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{T}}(\mathbf{v}(\mathbf{X},\ t-s), \mathbf{D}(\mathbf{X},\ t-s), \mathbf{W}(\mathbf{X},\ t-s), \mathbf{X}, s)\mathrm{d}s. \tag{5.4V}$$

This guideline requires that constitutive equations remain unchanged by superposed rigid object motions, thus measures of translational motion, like the displacement $\mathbf{u}(\mathbf{X}, t)$ and the velocity $\mathbf{v}(\mathbf{X}, t)$, and measures of rotational motion, like the infinitesimal rotation tensor $\mathbf{Y}(\mathbf{X}, t)$ and the spin tensor $\mathbf{W}(\mathbf{X}, t)$, must be excluded from the above equations . Using this guideline the form of these three constitutive ideas is then reduced to

$$\hat{\mathbf{T}} = \hat{\mathbf{T}}(\hat{\mathbf{E}}(X,\ t), X), \tag{5.5H}$$

$$\hat{\mathbf{T}} = -p\hat{\mathbf{U}} + \hat{\mathbf{T}}_{v}(\hat{\mathbf{D}}(X, t), X), \tag{5.5N}$$

and

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{T}}(\hat{\mathbf{D}}(X, t-s), X, s)\mathrm{d}s, \tag{5.5V}$$

where the three-dimensional tensors $\mathbf{E}(\mathbf{X}, t)$ and $\mathbf{D}(\mathbf{X}, t)$ have been replaced by their six-dimensional vector equivalents, $\hat{\mathbf{E}}(\mathbf{X}, t)$ and $\hat{\mathbf{D}}$ $(\mathbf{X}, t)$, respectively.

5.5 Determinism

A constitutive equation valid for a material at a time t must depend upon events that are occurring to the material at the instant t and upon events that have occurred to the material in the past. The constitutive equation cannot depend upon events that

will occur to the material in the future. For example, the constitutive assumption for elastic materials is that the stress depends upon the strain between a previous unstressed reference configuration and the instantaneous configuration of the object. All four of the constitutive equations satisfy this guideline. The first four satisfy it because all the variables entering the relationships are at a time t. The viscoelastic constitutive relation satisfies the guideline by only depending upon past events.

5.6 Linearization

Each of the constitutive ideas considered has been reduced to the form of a vector-valued ($\mathbf{q}$ or $\hat{\mathbf{T}}$) function or functional of another vector (∇p, $\hat{\mathbf{E}}$ or $\hat{\mathbf{D}}$), $\mathbf{X}$, and some scalar parameters. It is assumed that each of these vector-valued functions is linear in the vector argument, thus each may be represented by a linear transformation. For Darcy's law the second order tensor in three dimensions represents the coefficients of the linear transformation and, due to the dependence of the volume flow rate upon pressure, this second order tensor admits the functional dependency indicated:

$$\mathbf{q} = \phi \rho_{\mathrm{f}} \mathbf{v} / \rho_{\mathrm{o}} = -\mathbf{H}(p, \mathbf{X}) \cdot \nabla p(\mathbf{X},\ t). \tag{5.6D}$$

The minus sign was placed in (5.6D) to indicate that the volume fluid flux $\mathbf{q}$ would be directed down the pressure gradient, from domains of higher pore fluid pressure to domains of lower pressure.

For the three constitutive ideas involving the stress vector $\hat{\mathbf{T}}$, second order tensors in six dimensions represent the coefficients of the linear transformation:

$$\hat{\mathbf{T}} = \hat{\mathbf{C}}(\mathbf{X}) \cdot \hat{\mathbf{E}}(\mathbf{x}, t), \tag{5.6H}$$

$$\hat{\mathbf{T}} = -p\hat{\mathbf{U}} + \hat{\mathbf{N}}(\mathbf{X}) \cdot \hat{\mathbf{D}}(\mathbf{X}, t), \tag{5.6N}$$

and

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{G}}(X, s) \cdot \hat{\mathbf{D}}(\mathrm{X}, t - s)\mathrm{d}s. \tag{5.6V}$$

The six-dimensional second order tensors $\hat{\mathbf{C}}(\mathbf{X})$ and $\hat{\mathbf{N}}(\mathbf{X})$are for Hooke's law and the Newtonian law of viscosity, respectively. The six-dimensional second order tensor function $\hat{\mathbf{G}}(\mathbf{X}, s)$represents the viscoelastic coefficients.

Problem

5.6.1. Expand the constitutive relations for the Darcy medium (5.6D) and for Newton's law (5.6N) into their component from relative to a Cartesian 3-D coordinate system.

5.7 Coordinate Invariance

Since the representation of physical phenomena must be independent of the observer, it is necessary to express physical quantities in ways that are independent of coordinate systems. This is because different observers may select different coordinate systems. It therefore becomes a requirement that physical quantities be invariant of the coordinate system selected to express them. On the other hand, in order to work with these physical quantities, it is necessary to refer physical quantities to coordinate systems. In particular, a constitutive equation should be expressed by a relation that holds in all admissible coordinate systems. The admissible coordinate systems may be any coordinate system possible in a Euclidean three-dimensional space. A sufficient condition for the satisfaction of this requirement is to state the constitutive equations in tensorial form since tensors are independent of any particular coordinate system, although their components may be written relative to any particular one. In classical mechanics the essential concepts of force, velocity, and acceleration are all vectors; hence the mathematical language of classical mechanics is that of vectors. In the mechanics of deformable media the essential concepts of stress, strain, rate of deformation, etc., are all second order tensors; thus, by analogy, one can expect to deal quite frequently with second order tensors in this branch of mechanics. The constitutive ideas that are developed in this chapter satisfy the requirement of coordinate invariance by virtue of being cast as tensorial expressions.

5.8 Homogeneous Versus Inhomogeneous Constitutive Models

A material property is said to be *homogeneous* when it is the same at all particles $\mathbf{X}$ in the object, *inhomogeneous* if it varies from particle to particle in an object. Most biological materials are inhomogeneous and many manufactured materials are considered to be homogeneous. Each of the constitutive relations (5.6D), (5.6H), (5.6N), and (5.6V) is presented as inhomogeneous because the tensors representing their material coefficients, $\mathbf{H}(p, \mathbf{X})$, $\hat{\mathbf{C}}(\mathbf{X}), \hat{\mathbf{N}}(\mathbf{X})$, and $\hat{\mathbf{G}}(\mathbf{X},s)$, respectively, are allowed to depend upon the particle $\mathbf{X}$. If the dependence upon $\mathbf{X}$ does not occur, or can be neglected, then the material is homogeneous and the constitutive equations (5.6D), (5.6H), (5.6N), and (5.6V) take the form:

$$\mathbf{q} = \phi\rho_{\mathrm{f}}\mathbf{v}/\rho_{\mathrm{o}} = -\mathbf{H}(p)\cdot\nabla p(\mathbf{x},\ t), \tag{5.7D}$$

$$\hat{\mathbf{T}} = \hat{\mathbf{C}}\cdot\hat{\mathbf{E}}(\mathbf{x},t) \tag{5.7H}$$

$$\hat{\mathbf{T}} = -p\hat{\mathbf{U}} + \hat{\mathbf{N}}\cdot\hat{\mathbf{D}}(\mathbf{x},t) \tag{5.7N}$$

and

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{G}}(s)\cdot\hat{\mathbf{D}}(\mathbf{x},t-s)\mathrm{d}s, \tag{5.7V}$$

Note that, in the above constitutive expressions , not only has the dependence of the material coefficient tensors been removed by eliminating their dependence upon the particle **X**, but also **X** has been replaced by **x** everywhere else. For the two constitutive relations restricted to infinitesimal motions, (5.7H) and (5.7V), the constitutive relations based on a rigid continuum, (5.7D) and Eulerian viscous fluid theory (5.7N), there is no difference between **X** and **x** (see Sect. 2.4), hence **x** could have been used from the beginning of the chapter. For the Newtonian law of viscosity however, the assumption of homogeneity is much more significant because it permits the elimination of **X** from the entire constitutive relation, a constitutive relation that is not restricted to infinitesimal deformations. Thus, even though (5.7N) applies for large deformations, it is independent of **X**. The Newtonian law is different from the other four constitutive relations in another way, as detailed in the next section.

5.9 Restrictions Due to Material Symmetry

The results of the previous chapter are used in this section to further specify the form of the constitutive relations. Isotropy or any type of anisotropy is possible for the three constitutive relations, (5.7D), (5.7H), and (5.7V), that are, or may be, applied to solid or semi-solid materials. The type of anisotropy is expressed in the form of the tensors of material coefficients, $\mathbf{H}$, $\hat{\mathbf{C}}$, and $\hat{\mathbf{G}}(s)$, respectively. Once the type of anisotropy possessed by the solid or semi-solid material to be modeled has been determined, the appropriate form of $\mathbf{H}$ may be selected from Table 4.3 or the form of $\hat{\mathbf{C}}$ or $\hat{\mathbf{G}}(s)$ from Tables 4.4 and 4.5. Thus, for these four constitutive relations any type of material symmetry is possible. In this section and in the first paragraph of the next section the results summarized in Tables 4.3, 4.4 and 4.5 are cited. The derivation of these results is presented in Chap. 4.

The concepts of anisotropy and inhomogeneity of materials are sometimes confused. A constitutive relation is inhomogeneous or homogeneous depending upon whether the material coefficients (i.e., $\mathbf{H}$, $\hat{\mathbf{C}}$, and $\hat{\mathbf{G}}(s)$) depend upon **X** or

not. The type of material symmetry, that is to say either isotropy or the type of anisotropy, characterizing a constitutive relation is reflected in the form of the material coefficient tensors ($\mathbf{H}, \hat{\mathbf{C}}$ or $\hat{\mathbf{G}}(s)$), for example the forms listed in Tables 4.3, 4.4 and 4.5. The tensor $\mathbf{H}$ may have any of the forms in Table 4.3 and the tensor $\hat{\mathbf{C}}$ may have any forms in Tables 4.4 and 4.5. Material symmetry, that is to say the isotropy or type of anisotropy, is the property of a constitutive relation at a particle $\mathbf{X}$, while inhomogeneity or homogeneity of materials relates to how the material properties change from particle to particle. Thus a constitutive relation may be either anisotropic and homogeneous or anisotropic and inhomogeneous. The most mathematically simplifying assumptions are those of an isotropic symmetry and homogeneous material.

The Newtonian law of viscosity, (5.7N), is characterized by these most simplifying assumptions, homogeneity, and isotropy. These assumptions are easily justified when one thinks about the structure of, say, distilled water. Absent gravity, there is no preferred direction in distilled water, and distilled water has the same mechanical and thermal properties at all locations in the volume and in all volumes of distilled water. One can then generalize this thought process to see that fluids are isotropic.

The isotropic form of the Newtonian law of viscosity, (5.7N), is obtained by using the representation for the isotropic form of $\hat{\mathbf{N}}$ obtained from Table 4.5, thus

$$\begin{bmatrix} \hat{T}_1 + p \\ \hat{T}_2 + p \\ \hat{T}_3 + p \\ \hat{T}_4 \\ \hat{T}_5 \\ \hat{T}_6 \end{bmatrix} = \begin{bmatrix} \hat{N}_{11} & \hat{N}_{12} & \hat{N}_{12} & 0 & 0 & 0 \\ \hat{N}_{12} & \hat{N}_{11} & \hat{N}_{12} & 0 & 0 & 0 \\ \hat{N}_{12} & \hat{N}_{12} & \hat{N}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{N}_{11} - \hat{N}_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{N}_{11} - \hat{N}_{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{N}_{11} - \hat{N}_{12} \end{bmatrix} \begin{bmatrix} \hat{D}_1 \\ \hat{D}_2 \\ \hat{D}_3 \\ \hat{D}_4 \\ \hat{D}_5 \\ \hat{D}_6 \end{bmatrix}, \tag{5.8N}$$

This six-dimensional representation is converted to the three-dimensional representation by employing the relations (A163) and introducing the following new notation for the two distinct elements of the 6 by 6 matrix in (5.8N),

$$\hat{N}_{11} = \lambda + 2\mu, \hat{N}_{12} = \lambda \tag{5.9N}$$

then

$$\begin{bmatrix} T_{11} + p \\ T_{22} + p \\ T_{33} + p \\ \sqrt{2}T_{23} \\ \sqrt{2}T_{13} \\ \sqrt{2}T_{12} \end{bmatrix} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\mu \end{bmatrix} \begin{bmatrix} D_{11} \\ D_{22} \\ D_{33} \\ \sqrt{2}D_{23} \\ \sqrt{2}D_{13} \\ \sqrt{2}D_{12} \end{bmatrix} \tag{5.10N}$$

or

$$T_{11} + p = \lambda \text{tr}\mathbf{D} + 2\mu D_{11}, T_{22} + p = \lambda \text{tr}\mathbf{D} + 2\mu D_{22}, T_{33} + p = \lambda \text{tr}\mathbf{D} + 2\mu D_{33},$$
$$T_{23} = 2\mu D_{23}, T_{13} = 2\mu D_{13}, T_{12} = 2\mu D_{12},$$

where λ and μ are viscosity coefficients. It is easy to see that the constitutive relation may be rewritten in three dimensions as

$$\mathbf{T} + p\mathbf{1} = \lambda \text{tr}\mathbf{D})\mathbf{1} + 2\mu\mathbf{D}. \tag{5.11N}$$

This is the form of the constitutive equation for a viscous fluid, the pressure plus the Newtonian law of viscosity, which will be used in the remainder of the text.

Problems

5.9.1. Record the explicit matrix form for the constitutive relation for a Darcy porous medium in an inhomogeneous transversely isotropic material.
5.9.2. Record the explicit matrix form for the constitutive relation for Darcy's law in a homogeneous isotropic material.
5.9.3. Record the explicit matrix form for the constitutive relation for Hooke's law in
5.9.4. Record the explicit matrix form for the constitutive relation for a transversely isotropic, homogeneous viscoelastic material.
5.9.5. Show that the eigenvalues of (5.10N) are $3\lambda + 2\mu$ and 2μ and specify how many times each is repeated.

5.10 The Symmetry of the Material Coefficient Tensors

In this section the question of the symmetry of the matrices of the tensors of material coefficients, $\mathbf{H}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{N}}$and $\hat{\mathbf{G}}(s)$ is considered. Consider first the tensor of material coefficients $\hat{\mathbf{N}}$ for a Newtonian viscous fluid. In the previous section it was assumed that a Newtonian viscous fluid was isotropic, therefore, from Table 4.5, the tensor of material coefficients $\hat{\mathbf{N}}$ is symmetric. In this case the material symmetry implied the symmetry of the tensor of material coefficients. A similar symmetry result emerges for the permeability tensor $\mathbf{H}$ if only orthotropic symmetry or greater symmetry is considered. To see that material symmetry implies the symmetry of the tensor of material coefficients $\mathbf{H}$, if only orthotropic symmetry or greater symmetry is considered, one need only consult Table 4.3. The symmetry of $\mathbf{H}$ is also true for symmetries less than orthotropy, namely monoclinic and triclinic, but the proof will not be given here. Finally, $\hat{\mathbf{G}}(s)$ is never symmetric unless the viscoelastic model is in the limiting cases of $\hat{\mathbf{G}}(0)$ or $\hat{\mathbf{G}}(\infty)$ where the material behavior is elastic.

The symmetry of the tensor of elastic material coefficients $\hat{\mathbf{C}}$ is the only coefficient tensor symmetry point remaining to be demonstrated in this section. In this

development the inverse of (5.7H), the strain–stress relation rather than the stress–strain relation, is employed

$$\hat{\mathbf{E}} = \hat{\mathbf{S}} \cdot \hat{\mathbf{T}}, \quad \hat{\mathbf{S}} = \hat{\mathbf{C}}^{-1}, \tag{5.12H}$$

where $\hat{\mathbf{S}}$ is the compliance tensor of elastic material coefficients. The form and symmetry of $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}$ are identical for any material, and it is easy to show that the symmetry of one implies the symmetry of the other. The symmetry of $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}$ follows from the requirement that the work done on an elastic material in a closed cycle vanish. This requirement stems from the argument that if work can be done on the material in some closed cycle, then the cycle can be reversed and the material can do work in the reversed closed cycle. This would imply that work could be extracted from the material in a closed loading cycle. Thus one would be able to take an inert elastic material and extract work from it. This situation is not logical and therefore it is required that the work done on an elastic material in a closed loading cycle vanish. We express the work done on the material between the strain $\hat{\mathbf{E}}^{(1)}$ and the strain $\hat{\mathbf{E}}^{(2)}$ by

$$W_{12} = \int_1^2 \hat{\mathbf{T}} \cdot d\hat{\mathbf{E}}, \tag{5.13H}$$

and for a closed loading cycle it is required that

$$\oint \hat{\mathbf{T}} \cdot d\hat{\mathbf{E}} = 0. \tag{5.14H}$$

Consider the work done in a closed loading cycle applied to a unit cube of a linear anisotropic elastic material. The loading cycle begins from an unstressed state and contains the following four loading sequences (Fig. 5.2(a)): O → A, the stress is increased slowly from 0 to $\hat{T}_i^{\mathrm{A}}$; A → B, holding the stress state $\hat{T}_i^{\mathrm{A}}$ constant the second stress is increased slowly from $\hat{T}_i^{\mathrm{A}}$ to $\hat{T}_i^{\mathrm{A}} + \hat{T}_i^{\mathrm{B}}, \hat{T}_i^{\mathrm{B}} \neq \hat{T}_i^{\mathrm{A}}$; B → C, holding the second stress state $\hat{T}_i^{\mathrm{B}}$constant the first stress is decreased slowly from $\hat{T}_i^{\mathrm{A}} + \hat{T}_i^{\mathrm{B}}$ to $\hat{T}_i^{\mathrm{B}}$; and C → O, the stress is decreased slowly from $\hat{T}_i^{\mathrm{B}}$ to 0. At the end of this loading cycle the object is again in an unstressed state. The work done in (5.13H) on each of these loading sequences is expressed as an integral in stress:

$$W_{12} = \int_1^2 \hat{\mathbf{T}} \cdot \mathrm{d}\hat{\mathbf{E}} = \int_1^2 \hat{\mathbf{E}} \cdot d\hat{\mathbf{T}}. \tag{5.15H}$$

The integral over the first loading sequence of the cycle, from 0 to $\hat{T}_i^{\mathrm{A}}$, is given by

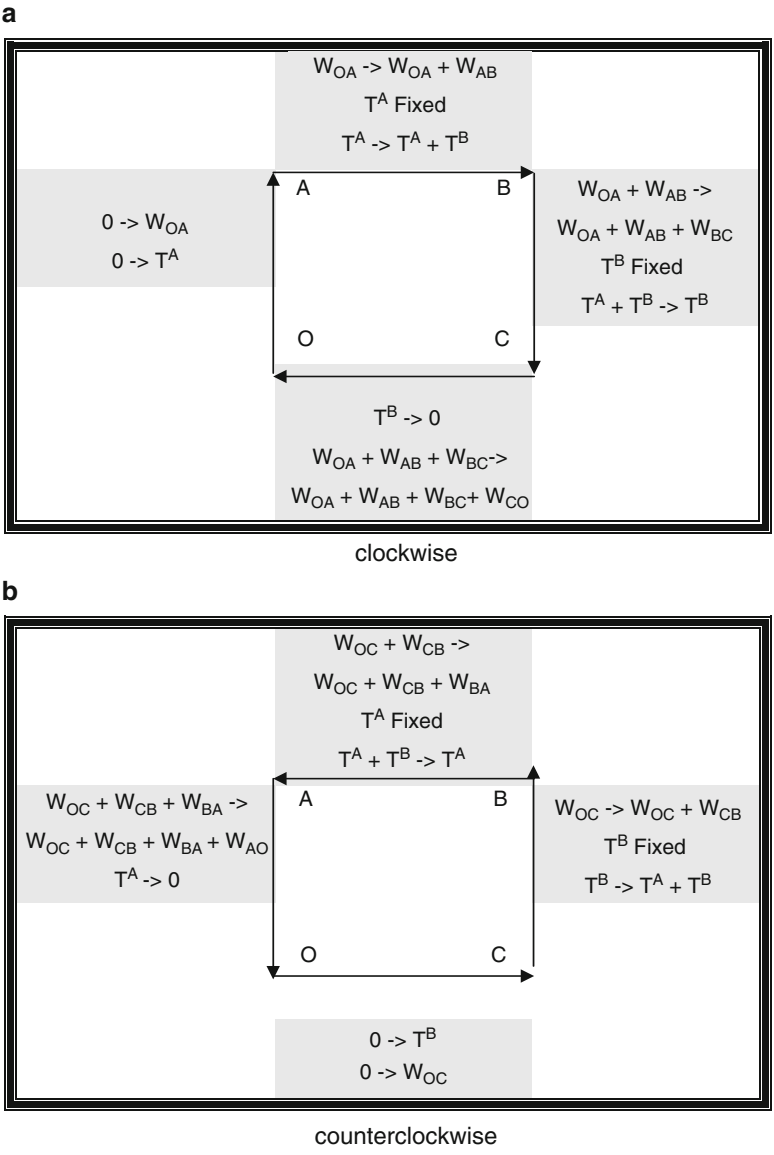


Fig. 5.2 Illustrations of closed loading cycles. (**a**) Clockwise cycle O → A → B → C → O, (**b**) counterclockwise cycle O → C → B → A → O. See the text for further explanation

$$W_{\mathrm{OA}} = \int_{O}^{A} \hat{\mathbf{E}} \cdot d\hat{\mathbf{T}} = \int_{O}^{A} \hat{E}_i d\hat{T}_i = \frac{1}{2} \hat{\mathbf{S}}_{ij} \hat{T}_j^{(\mathrm{A})} \hat{T}_i^{(\mathrm{A})}. \tag{5.16H}$$

The integral over the second loading sequence of the cycle, from $\hat{T}_i^{\mathrm{A}}$ to $\hat{T}_i^{\mathrm{A}} + \hat{T}_i^{\mathrm{B}}$, $\hat{T}_i^{\mathrm{B}} \neq \hat{T}_i^{\mathrm{A}}$ is a bit more complicated because the loading of the object begins from a

state in which it is subjected to the stress $\hat{T}_i^{\mathrm{A}}$, which is held constant during this leg of the cycle. During the second loading sequence the strain is given by

$$\hat{E}_i = \hat{\mathbf{S}}_{ij}(\hat{T}_j^{(\mathrm{A})} + \hat{T}_j), \quad 0 \leq \hat{T}_j \leq \hat{T}_j^{(\mathrm{B})} \tag{5.17H}$$

and the work done from $\hat{T}_i^A$ to $\hat{T}_i^{\mathrm{A}} + \hat{T}_i^{\mathrm{B}}$ is

$$W_{\mathrm{AB}} = \int_{\mathrm{O}}^{\mathrm{B}} \hat{\mathbf{S}}_{ij}(\hat{T}_j^{(\mathrm{A})} + \hat{T}_j)\mathrm{d}\hat{T}_i = \hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{A})}\hat{T}_i^{(\mathrm{B})} + \frac{1}{2}\hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{B})}\hat{T}_i^{(\mathrm{B})} \tag{5.18H}$$

where the factor of one-half does not appear before the first term on the right hand side because $\hat{T}_i^{\mathrm{A}}$ is held fixed during the second loading sequence in this leg of the loading cycle. During the third loading sequence the stress $\hat{T}_i^{\mathrm{B}}$ is held fixed; the strain is given by

$$\hat{\mathbf{E}}_i = \hat{\mathbf{S}}_{ij}(\hat{T}_j^{(\mathrm{B})} + \hat{T}_j), 0 \leq \hat{T}_j \leq \hat{T}_j^{(\mathrm{A})}. \tag{5.19H}$$

The work done during the third loading sequence of the cycle, from $\hat{T}_i^A + \hat{T}_i^B$ to $\hat{T}_i^B$ is then

$$W_{\mathrm{BC}} = \int_{A}^{O} \hat{\mathbf{S}}_{ij}(\hat{T}_j^{(\mathrm{B})} + \hat{T}_j)d\hat{T}_i = -\hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{B})}\hat{T}_i^{(\mathrm{A})} - \frac{1}{2}\hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{A})}\hat{T}_i^{(\mathrm{A})}. \tag{5.20H}$$

The work done during the final loading sequence, from $\hat{T}_i^B$ to 0, is then

$$W_{\mathrm{CO}} = \int_{\mathrm{B}}^{\mathrm{O}} \hat{\mathbf{S}}_{ij}\hat{T}_j d\hat{T}_i = -\frac{1}{2}\hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{B})}\hat{T}_i^{(\mathrm{B})}. \tag{5.21H}$$

The work done in the closed cycle is then the sum $W_{\mathrm{OO}} = W_{\mathrm{OA}} + W_{\mathrm{AB}} + W_{\mathrm{BC}} + W_{\mathrm{CO}}$ given by

$$W_{\mathrm{OO}} = \hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{A})}\hat{T}_i^{(\mathrm{B})} - \hat{\mathbf{S}}_{ij}\hat{T}_j^{(\mathrm{B})}\hat{T}_i^{(\mathrm{A})} = (\hat{\mathbf{S}}_{ij} - \hat{\mathbf{S}}_{ji})\hat{T}_j^{(\mathrm{A})}\hat{T}_i^{(\mathrm{B})}. \tag{5.22H}$$

If the cycle is traversed in reverse (Fig. 5.2(b)), then

$$W_{\mathrm{OO}} = -(\hat{\mathbf{S}}_{ij} - \hat{\mathbf{S}}_{ji})\hat{T}_j^{(\mathrm{A})}\hat{T}_i^{(\mathrm{B})}. \tag{5.23H}$$

This result suggests that, if work is required to traverse the loading cycle in one direction, then work may be extracted by traversing the cycle in the reverse direction. It is common knowledge however, that it is not possible to extract work from an inert material by mechanical methods. If it were, the world would be a different place. To prevent the possibility of extracting work from an inert material, it is required that $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}$ are symmetric,

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}^{\mathrm{T}}, \quad \hat{\mathbf{C}} = \hat{\mathbf{C}}^{\mathrm{T}}. \tag{5.24H}$$

There are further restrictions on the tensors of material coefficients and some of them will be discussed in the next section.

The definition of a linear elastic material includes not only the stress–strain relation $\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{E}}(\mathbf{x}, t)$, but also the symmetry restriction $\hat{\mathbf{C}} = \hat{\mathbf{C}}^{\mathrm{T}}$, (5.7H) and (5.24H), respectively. The symmetry restriction $\hat{\mathbf{C}} = \hat{\mathbf{C}}^{\mathrm{T}}$ is equivalent to the requirement that the work done on an elastic material in a closed loading cycle is zero, (5.14H). The work done is therefore an exact differential (see Sect. A.15). This restriction on the work done allows for the introduction of a potential, the strain energy U. Since the work done on an elastic material in a closed loading cycle is zero, this means that the work done on the elastic material depends only on initial and final states of stress (strain) and not on the path followed from the initial to the final state. From an initial state of zero stress or strain, the strain energy U is defined as the work done (5.15H):

$$U = \int \hat{\mathbf{T}} \cdot \mathrm{d}\hat{\mathbf{E}} = \int \hat{\mathbf{E}} \cdot \mathrm{d}\hat{\mathbf{T}}. \tag{5.25H}$$

The strain energy U may be considered as a function of either $\hat{\mathbf{T}}$ or $\hat{\mathbf{E}}$, $U(\hat{\mathbf{T}})$ or $U(\hat{\mathbf{E}})$. From (5.25H) and the fundamental theorem of the integral calculus, namely that the derivative of an integral with respect to its parameter of integration yields the integrand,

$$\hat{\mathbf{T}} = \frac{\partial U}{\partial \hat{\mathbf{E}}} \text{ and } \hat{\mathbf{E}} = \frac{\partial U}{\partial \hat{\mathbf{T}}} \left(\text{or } \mathbf{T} = \frac{\partial U}{\partial \mathbf{E}} \text{ and } \mathbf{E} = \frac{\partial U}{\partial \mathbf{T}} \right). \tag{5.26H}$$

The following expressions for U are obtained substituting Hooke's law (5.7H) into (5.25H) and then integrating both of the expressions for U in (5.25H), thus

$$U = \frac{1}{2} \hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}} \text{ and } U = \frac{1}{2} \hat{\mathbf{T}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{T}}. \tag{5.27H}$$

It is easy to verify that the linear form of Hooke's law is recovered if the representations (5.27H) for U are differentiated with respect to $\hat{\mathbf{T}}$ and $\hat{\mathbf{E}}$, respectively as indicated by (5.27H). *It then follows that* (5.26H) *or* (5.27H) *constitutes an equivalent definition of a linear elastic material*. The definition of the most

important type of non-linear elastic material model, the hyperelastic material, is that the stress is the derivative of the strain energy with respect to strain, as in (5.26H). In the non-linear case the strain energy is not specified by an expression as simple as (5.27H).

Problems

5.10.1. Consider the work done in a closed loading cycle applied to a unit cube of a linear anisotropic elastic material. The loading cycle this cube will be subjected to begins from an unstressed state and contains the following four loading sequences: O → A, the stress in the x_1 direction is increased slowly from O to T_A; A → B, holding the stress in the x_1 direction, T_A constant, the stress in the x_2 direction is increased slowly from O to T_B; B → C, holding the stress in the x_2 direction constant, the stress in the x_1 direction is decreased slowly from T_A to O; C → O, the stress in the x_2 direction is decreased slowly from T_B to O. At the end of this loading cycle the object is again in an unstressed state. Show that the work done on each of these loading sequences is given by

$$W_{OA} = \frac{1}{2}\hat{\mathbf{S}}_{11}(\hat{T}_1^{(A)})^2,\ W_{AB} = \frac{1}{2}\hat{\mathbf{S}}_{22}(\hat{T}_2^{(B)})^2 + \hat{\mathbf{S}}_{21}T_1^{(A)}\hat{T}_2^{(B)}$$

$$W_{BC} = -\frac{1}{2}\hat{\mathbf{S}}_{11}(\hat{T}_1^{(A)})^2 - \hat{\mathbf{S}}_{12}T_1^{(A)}\hat{T}_2^{(B)},\ W_{OA} = -\frac{1}{2}\hat{\mathbf{S}}_{22}(\hat{T}_2^{(B)})^2$$

and show that the work done around the closed cycle is given by

$$W_{OO} = (\hat{\mathbf{S}}_{21} - \hat{\mathbf{S}}_{12})\hat{T}_1^{(A)}\hat{T}_2^{(B)}.$$

Show that one may therefore argue that $\hat{S}_{12} = \hat{S}_{21}$.

5.10.2 Show that $\hat{\mathbf{S}} = \hat{\mathbf{S}}^T$ implies $\hat{\mathbf{C}} = \hat{\mathbf{C}}^T$.

5.11 Restrictions on the Coefficients Representing Material Properties

In this section other restrictions on the four tensors of material coefficients are considered. Consider first that the dimensions of the material coefficients contained in the tensor must be consistent with the dimensions of the other terms occurring in the constitutive equation. The constitutive equation must be invariant under changes in gauge of the basis dimensions as would be affected, for example, by a change from SI units to the English foot-pound system.

It will be shown here that all the tensors of material coefficients are positive definite as well as symmetric except for the viscoelastic tensor function $\hat{\mathbf{G}}(s)$. To see that the permeability tensor $\mathbf{H}(p)$ is positive definite let $\nabla p = \mathbf{n}(\partial p/\partial w)$ where

w is a scalar length parameter in the $\mathbf{n}$ direction. The volume flow rate $\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o$ projected in the $\mathbf{n}$ direction, $\phi\rho_f\mathbf{v}/\rho_o$, is then given by

$$\phi\rho_f\mathbf{v}\cdot\mathbf{n}/\rho_o = -(\partial p/\partial w)\mathbf{n}\cdot\mathbf{H}\cdot\tilde{\mathbf{n}} \tag{5.28D}$$

In order for the volume flux per unit area $\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o$ in the $\mathbf{n}$ direction to be pointed in the direction of decreasing pressure, it is necessary to require that

$$\mathbf{n}\cdot\mathbf{H}\cdot\mathbf{n}>0 \text{ for all unit vectors } \mathbf{n}. \tag{5.29D}$$

If the fluid flowed the other way, all the mass of the fluid would concentrate itself at the highest-pressure location and we know that that does not happen. The condition (5.29D) is the condition that the symmetric tensor $\mathbf{H}$ be positive definite. This condition is satisfied if all the eigenvalues of $\mathbf{H}$ are positive.

The tensor of material coefficients for the Newtonian law of viscosity is positive definite also. To see this, the local stress power $\mathrm{tr}(\mathbf{T}\cdot\mathbf{D}) = \mathbf{T}:\mathbf{D}$ is calculated using the constitutive equation (5.11N) and the decomposition (A18) of the rate of deformation tensor,

$$\mathbf{D} = (1/3)(\mathrm{tr}\mathbf{D})\mathbf{1} + \ \mathrm{dev}\mathbf{D},\ \mathrm{dev}\mathbf{D} = \mathbf{D} - (1/3)(\mathrm{tr}\mathbf{D})\mathbf{1}. \tag{5.30N}$$

The stress due to viscous stresses may be recast in the form

$$\mathbf{T} + \mathrm{p}\mathbf{1} = ((3\dot{\lambda}2\tilde{\mu})3)(\mathrm{tr}\mathbf{D})\mathbf{1} + \ 2\mu\,\mathrm{dev}\mathbf{D}, \tag{5.31N}$$

using (5.11N) and (5.30N). Calculation of the viscous stress power $\mathrm{tr}\{(\mathbf{T} + \mathrm{p}\mathbf{1})\mathbf{D}\}$ using the two equations above then yields

$$(\mathbf{T} + \ \mathrm{p}\mathbf{1}) : \mathbf{D} = ((3\dot{\lambda}2\,\tilde{\mu}\,3(\mathrm{tr}\mathbf{D})^2 + \ 2\mu\,\mathrm{trdev}\mathbf{D})^2. \tag{5.32N}$$

Note that the terms in (5.32N) involving $\mathbf{D}$, and multiplying the expressions $3\dot{\lambda}2\mu$ and 2μ are squared; thus if the viscous stress power $\mathrm{tr}\{(\mathbf{T} + \mathrm{p}\mathbf{1})\cdot\mathbf{D}\} = (\mathbf{T} + \mathrm{p}\mathbf{1}):\mathbf{D}$ is to be positive it is necessary that

$$3\dot{\lambda}2\,\mu\tilde{\geq}\mu\tilde{\geq}. \tag{5.33N}$$

The viscous stress power $(\mathbf{T} + p\mathbf{1}):\mathbf{D}$ must be positive for an inert material as the world external to the material is working on the inert material, not the reverse. The inequalities restricting the viscosities (5.33N) also follow for the condition that the 6 by 6 matrix (5.10N) be positive definite. Finally, to see that the tensor of elastic coefficients is positive definite, the local form of the work done expressed in terms of stress and strain, $\mathbf{T}:\mathbf{E} = \hat{\mathbf{T}}\cdot\hat{\mathbf{E}}$ is employed. Since $\hat{\mathbf{T}} = \hat{\mathbf{C}}\cdot\hat{\mathbf{E}}$ it follows that $\mathbf{T}:\mathbf{E} = \hat{\mathbf{T}}\cdot\hat{\mathbf{E}} = \hat{\mathbf{E}}\cdot\hat{\mathbf{C}}\cdot\hat{\mathbf{E}}$ thus from the requirement that the local work done on an inert material be positive, $\mathbf{T}:\mathbf{E} = \hat{\mathbf{T}}\cdot\hat{\mathbf{E}} > 0$, it follows that

$$\hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}} > 0 \quad \text{for all non} - \text{zero strains } \hat{\mathbf{E}} \tag{5.34H}$$

Thus it follows that $\hat{\mathbf{C}}$ is a positive definite symmetric tensor, $\hat{\mathbf{C}} = \hat{\mathbf{C}}^{\mathrm{T}}$ and $\hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}} > 0$ for all non-zero strains $\hat{\mathbf{E}}$.

5.12 Summary of Results

In this chapter a progressive development of four constitutive relations has been presented. Beginning with the constitutive idea, restrictions associated with the notions of localization, invariance under rigid object motions, determinism, coordinate invariance, and material symmetry were imposed. In the development the constitutive equations were linearized and the definition of homogeneous versus inhomogeneous constitutive models was reviewed. Restrictions due to material symmetry, the symmetry of the material coefficient tensors, and restrictions on the coefficients representing material properties were developed. The results of these considerations are the following constitutive equations

$$\mathbf{q} = \phi \rho_{\mathrm{f}} \mathbf{v} / \rho_{\mathrm{o}} = -\mathbf{H}(p) \cdot \nabla p(\mathbf{x},\ t), \mathbf{H}(p)\ = \mathbf{H}^{\mathrm{T}}(p), \tag{5.36D}$$

$$\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{E}},\ \text{where } \hat{\mathbf{C}} = \hat{\mathbf{C}}^{T} \tag{5.36H}$$

where $\mathbf{H}(p)$ and $\hat{\mathbf{C}}$ are positive definite, and

$$\mathbf{T} = -\mathrm{p}\mathbf{1} + \lambda(\mathrm{tr}\mathbf{D})\mathbf{1} + 2\mu\mathbf{D}, \tag{5.36N}$$

where p is the fluid pressure and λ and μ are viscosity coefficients ($3\lambda\ 2\mu \gtrsim 2\mu \gtrsim$), and

$$\hat{\mathbf{T}} = \int_{s=0}^{\infty} \hat{\mathbf{G}}(s) \cdot \hat{\mathbf{D}}(\mathbf{x}, t-s)\mathrm{d}s, \tag{5.36V}$$

where there are no symmetry restrictions on $\hat{\mathbf{G}}(s)$. All the constitutive equations developed in this chapter, including Darcy's law, can be developed from many different arguments. Darcy's law can also be developed from experimental or empirical results for seepage flow in non-deformable porous media; all of the other constitutive equations in this chapter have experimental or empirical basis. Analytical arguments for these constitutive equations are presented so that it is understood by the reader that they also have an analytical basis for their existence. Darcy's law is a form of the balance of linear momentum and could include a body force term; however, such a body force would normally be a constant, and since it is

only the divergence of $\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o$ that appears in the theory, such a body force would not appear in the final theory. In particular, Darcy's law could also be developed from the conservation of linear momentum, or from the Navier Stokes equations that, as will be shown in Chap. 6, is a combination of the stress equations of motion and the Newtonian law of viscosity (5.6D).

References

Christensen RM (1971) Theory of viscoelasticity. Academic Press, New York and London
Darcy H (1856) Les Fontains Publiques de la Ville de Dijon. Dalmont, Paris
Gurtin ME (1972) The linear theory of elasticity. In: Flugge S (ed) Handbuch der Physik, vol 2. Springer Verlag, Berlin
Lockett FJ (1972) Nonlinear viscoelastic solids. Academic Press, New York and London
Sokolnikoff IS (1956) Mathematical theory of elasticity. McGraw Hill, New York
Truesdell CA, Toupin RA (1960) The classical field theories. In: Flugge S (ed) Handbuch der Physik, vol III/1. Springer Verlag, Berlin
Truesdell CA, Noll W (1965) The non-linear field theories of mechanics. In: Flugge S (ed) Handbuch der Physik, vol III/3. Springer Verlag, Berlin

Chapter 6
Four Linear Continuum Theories

Four linear theories are considered in this chapter. Each has a distinctive and interesting history. Each one of the theories was originally formulated between 1820 and 1860. Representative of the theme of this chapter are the opening lines of the Historical Introduction in A. E. H. Love's Theory of Elasticity (original edition, 1892): "*The Mathematical Theory of Elasticity is occupied with an attempt to reduce to calculation the state of strain, or relative displacement, within a solid .. object.. which is subject to the action of an equilibrating system of forces, or is in a state of slight internal relative motion, and with endeavours to obtain results which shall be practically important in applications to architecture, engineering, and all other useful arts in which the material of construction is solid*."

6.1 Formation of Continuum Theories

Four linear continuum theories are developed in this chapter. These are the theories of fluid flow through rigid porous media, of elastic solids, of viscous fluids, and of viscoelastic materials. There are certain features that are common in the development of each of these theories: they all involve at least one conservation principle and one constitutive equation and for each, it is necessary to specify boundary or initial conditions to properly formulate boundary value problems. Some of the continuum theories involve more than one conservation principle, more than one constitutive equation, and some kinematics relations. Thus this chapter draws heavily upon the material in the previous chapters and serves to integrate the kinematics, the conservation principles, and the constitutive equations into theories that may be applied to physical situations to explain physical phenomena. This is generally accomplished by the solution of partial differential equations in the context of specific theories.

The differential equations that are formulated from these linear theories are usually the familiar, fairly well-understood differential equations, and they

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_6,

represent each of the three major types of second-order partial differential equations—the elliptic, the parabolic, and the hyperbolic. These three types of differential equations are characteristic of three different types of physical situations. Elliptic partial differential equations typically occur in equilibrium or steady-state situations where time does not enter the problem. Laplace's equation,

$$\nabla^2 f(x, y, z) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0, \tag{6.1}$$

and Poisson's equation, $\nabla^2 f(x, y, z) = g(x, y, z)$, are typical elliptic partial differential equations. Parabolic partial differential equations,

$$\nabla^2 f(x, y, z) = \frac{\partial f}{\partial t}, \tag{6.2}$$

typically occur in diffusion problems, such as thermal diffusion in a heat conducting material or fluid pressure diffusion in a rigid porous medium as will be seen in the following section. Hyperbolic partial differential equations,

$$\nabla^2 f(x, y, z) = \frac{\partial^2 f}{\partial t^2}, \tag{6.3}$$

often characterize dynamic situations with propagating waves, and are called the *wave equations*. A *boundary value problem* is the problem of finding a solution to a differential equation or to a set of differential equations subject to certain specified boundary and/or initial conditions. The theories developed in this chapter all lead to boundary value problems. Thus, in this chapter, conservation principles, constitutive equations, and some kinematics relations from the previous chapters are employed to formulate continuum theories that lead to physically motivated and properly formulated boundary value problems.

6.2 The Theory of Fluid Flow Through Rigid Porous Media

The theory of fluid flow through rigid porous media reduces to a typical diffusion problem; the diffusion of the pore fluid pressure through the porous medium. The solid component of the continuum is assumed to be porous, rigid, and stationary, thus the strain and rate of deformation are both zero. The pores in the solid cannot be closed,—most of them must be open and connected; so it is possible for the pore fluid to flow about the medium. The assumption of an immobile rigid porous continuum is not necessary as the constitutive equation for the porous medium may be combined with the equations of elasticity to form a theory for poroelastic materials (Chap. 8). Recall that ρ_f denotes the density of the fluid in the pores of the

porous medium, ρ_o a reference value of density, $\mathbf{v}$ the velocity of the fluid passing through the pore, ϕ the porosity of the medium, and $\mathbf{q}$ the volume flux of fluid per unit area, $\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o$, through the pores. This constitutive idea is that the fluid volume flux $\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o$ at a particle $\mathbf{X}$ is a function of the pressure variation in the neighborhood of $\mathbf{X}$, $N(X)$. The constitutive relation for the rigid porous continuum is Darcy's law. Darcy's law relates the fluid volume flow rate, $\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o$, to the gradient ($\nabla p$) of the pore pressure p,

$$\mathbf{q} = \phi\rho_f\mathbf{v}/\rho_o = -\mathbf{H}(p)\cdot\nabla p(\mathbf{x},t), \mathbf{H}(p) = \mathbf{H}^T(p), \quad (5.36D)\ \textit{repeated}$$

where the symmetry in material tensor $\mathbf{H}$ has been shown in Chap. 5 to hold for material symmetries greater than monoclinic. The conservation law that is combined with Darcy's law is the conservation of mass (3.6) in a slightly rearranged form. In (3.6) the density ρ is replaced by the product of the porosity and the fluid density, $\phi\rho_f$, in order to account for the fact that the fluid is only in the pores of the medium, and the resulting mass balance equation is divided throughout by ρ_o, thus

$$\frac{1}{\rho_o}\frac{\partial\phi\rho_f}{\partial t} + \nabla\cdot(\phi\rho_f\mathbf{v}/\rho_o) = 0. \tag{6.4}$$

In the case of compressible fluids it is reasonable to assume that fluid is barotropic, that is to say that the fluid density ρ_ϕ is a function of pressure, $\rho_\phi = \rho_\phi(p)$, in which case (6.4) may be written as

$$\frac{\phi}{\rho_o}\frac{\partial\rho_f}{\partial p}\frac{\partial p}{\partial t} + \nabla\cdot\mathbf{q} = 0, \tag{6.5}$$

where $\phi\rho_f\mathbf{v}/\rho_o$ has been replaced by $\mathbf{q}$ and where it has been assumed that the porosity ϕ is not a function of time. Substituting (5.36D) into (6.5), and multiplying through by the inverse of the factor multiplying the partial derivative of the pressure p with respect to time, a differential equation for the pore pressure is obtained,

$$\frac{\partial p}{\partial t} = \left(\frac{\rho_o}{\phi}\frac{\partial p}{\partial\rho_f}\right)\nabla\cdot(\mathbf{H}\cdot\nabla p). \tag{6.6}$$

If it is assumed that $\mathbf{H}$ and $\frac{\partial p}{\partial\rho_f}$ are constants, and if the viscosity μ of the pore fluid is introduced by the substitution

$$\mathbf{H} \equiv \frac{1}{\mu}\mathbf{K}, \tag{6.7}$$

then

$$\frac{\partial p}{\partial t} = \frac{1}{\tau}\mathbf{K} : (\nabla\otimes\nabla)p, \tag{6.8}$$

where

$$\tau = \mu \frac{\phi}{\rho_o} \frac{\partial \rho_{\mathrm{f}}}{\partial p} \tag{6.9}$$

is a material constant of dimension time. The time constant τ depends upon the porosity of the medium and the viscosity and barotropic character of the pore fluid. The constant intrinsic permeability tensor $\mathbf{K}$ is of dimension length squared; it depends only upon the arrangement and size of the pores in the medium and, in particular, is independent of the pore fluid properties. The differential equation (6.8) is a typical diffusion equation for an anisotropic medium; in the case of an isotropic medium the differential equation (6.8) becomes

$$\frac{\partial p}{\partial t} = \frac{k}{\tau} \nabla^2 p, \tag{6.10}$$

where $\mathbf{K} = k\mathbf{1}$. Mathematically equivalent differential equations for anisotropic and isotropic heat conductors are obtained from the Fourier law of heat conduction and the conservation of energy. The pore fluid density ρ_ϕ satisfies the same diffusion equation (6.8) as the pore fluid pressure p,

$$\frac{\partial \rho_{\mathrm{f}}}{\partial t} = \frac{1}{\tau} \mathbf{K} : (\nabla \otimes \nabla) \rho_{\mathrm{f}}, \tag{6.11}$$

a result that follows from the assumed barotropic character, $\rho_\phi = \rho_\phi(p)$, of the pore fluid, and the assumption that $\frac{\partial p}{\partial \rho_{\mathrm{f}}}$ is constant.

The boundary conditions on the pore pressure field customarily employed in the solution of the differential equation (6.8) are (1) that the external pore pressure p is specified at the boundary (a lower pressure on one side of the boundary permits flow across the boundary), (2) that the pressure gradient ∇p at the boundary is specified (a zero pressure gradient permits no flow across the boundary), (3) that some linear combination of (1) and (2) is specified. The complete theory for the flow of a fluid through a rigid porous medium consists of the differential equation (6.8) specified for an object O and boundary and initial conditions. The boundary conditions include the prescription of some combination of the pressure and the mass flux normal to the boundary ∂O as a function of time, $\mathbf{n}\cdot\mathbf{q} = \phi\ \rho_{\mathrm{f}}\ \mathbf{n}\cdot\mathbf{v}(\boldsymbol{x}^*, t)/\rho_{\mathrm{o}} = -\ (1/\mu)\ \mathbf{n}\cdot\mathbf{K}\cdot\nabla\mathbf{p}(\boldsymbol{x}^*, t)$, $\boldsymbol{x}^* \subset \partial O$, thus

$$-\ (c_1/\mu)\mathbf{n} \cdot \mathbf{K} \cdot \nabla \mathbf{p}(x^*, t) + c_2 \mathbf{p}(x^*, t) = f(x^*, t), x^* \subset \partial O \tag{6.12}$$

where c_1 and c_2 are constants and $f(\boldsymbol{x}^*, t)$ is a function specified on ∂O. In the case when c_1 is zero, this condition reduces to a restriction of the boundary pressure and, in the case when c_2 is zero, it is a restriction on the component of the mass flux vector normal to the boundary.

The basic pore fluid flow problem described by (6.8) and (6.12) may be simplified by specifying the type of material symmetry. If the material has orthotropic symmetry, employing a representation from Table 4.3 and the principal coordinate system for orthotropic symmetry, (6.8) and (6.12) become

$$\frac{\partial p}{\partial t} = \frac{K_{11}}{\tau}\frac{\partial^2 p}{\partial x_1^2} + \frac{K_{22}}{\tau}\frac{\partial^2 p}{\partial x_2^2} + \frac{K_{33}}{\tau}\frac{\partial^2 p}{\partial x_3^2} \text{ in } O, \tag{6.13}$$

$$-\frac{c_1}{\mu}\left(n_1 K_{11}\left\{\frac{\partial p}{\partial x_1}\right\}_{\mathbf{x}*} + n_2 K_{22}\left\{\frac{\partial p}{\partial x_2}\right\}_{\mathbf{x}*} + n_3 K_{33}\left\{\frac{\partial p}{\partial x_3}\right\}_{\mathbf{x}^*}\right) + c_2 p(\mathbf{x}^*, t)$$
$$= f(\mathbf{x}^*, t), x^* \subset \partial O, \tag{6.14}$$

and, if the material is isotropic, it follows from a representation in Table 4.3 that (6.13) reduces to (6.10) and (6.14) may be specialized as follows:

$$-(c_1/\mu)K\mathbf{n}\cdot\nabla\mathbf{p}(x^*, t) + c_2\mathbf{p}(x^*, t) = f(x^*, t), x^* \subset \partial O \tag{6.15}$$

The coordinate system may be rescaled so that the differential equation and boundary conditions are those of distorted heat conduction objects with isotropic material symmetry. To accomplish this the coordinates x_1, $x_{2,}$ and x_3 are rescaled by

$$x = \left(\sqrt{\frac{k}{K_{11}}}\right)x_1, y = \left(\sqrt{\frac{k}{K_{22}}}\right)x_2, z = \left(\sqrt{\frac{k}{K_{33}}}\right)x_3, \text{ where}$$
$$k^3 = K_{11}K_{22}K_{33}. \tag{6.16}$$

Then the differential equation (6.10) for an isotropic medium applies in the *distorted or stretched* O, and the boundary conditions (6.14) are

$$-\frac{c_1}{\mu}\left(n_1\sqrt{kK_{11}}\left\{\frac{\partial p}{\partial x}\right\}_{\mathbf{x}^*} + n_2\sqrt{kK_{22}}\left\{\frac{\partial p}{\partial y}\right\}_{\mathbf{x}^*} + n_3\sqrt{kK_{33}}\left\{\frac{\partial p}{\partial z}\right\}_{\mathbf{x}*}\right) + c_2 p(\mathbf{x}^*, t)$$
$$= f(\mathbf{x}^*, t), x^* \subset \text{distorted } \partial O. \tag{6.17}$$

In this restatement of the anisotropic problem one trades a slightly more complicated differential equation (6.8) for a simpler one (6.10) and obtains the slightly more complicated boundary condition above.

Example 6.2.1

A layer of thickness L of a rigid porous material is between two fluid reservoirs both containing the same fluid at the same pressure p_o, as illustrated in Fig. 1.8. Let x_1 be a coordinate that transverses the perpendicular distance between two layers; one reservoir is located at $x_1 = L$ and the opening to the other reservoir is located at $x_1 = 0$, although the fluid level in the second reservoir is below $x_1 = 0$ to maintain

the exit pressure of p_o. At $t = 0$ the pore fluid pressure in the reservoir at $x_1 = L$ is raised from p_o to $\rho gh + p_o$ and held at $\rho gh + p_o$ for all subsequent times, $\infty > t > 0$. The pressure at the entrance to the second reservoir at $x_1 = 0$, or the exit from the layer, is held at p_o for all times, $\infty > t > -\infty$. (a) Show that the steady-state solution (that is to say after the startup effects of the pressure changes at $t = 0$ have vanished) is a linear variation in the pore fluid pressure from $\rho gh + p_o$ at $x_1 = L$ to p_o at $x_1 = 0$. (b) How does the pressure $p(x_1, t)$ in the layer evolve in time to the linear long-term steady-state solution?

Solution: The problem is one-dimensional in the direction of x_1. The one-dimensional form of the differential equation (6.8) is

$$\frac{\partial p}{\partial t} = \frac{K_{11}}{\tau} \frac{\partial^2 p}{\partial x_1^2}. \qquad \text{(u)unsteady}$$

In the special case of steady states it reduces to

$$\frac{\partial^2 p}{\partial x_1^1} = 0. \qquad \text{(s)steady}$$

The solution to the steady-state equation (s) subject to the condition that $p = p_o$ at $x_1 = 0$ and $p = \rho gh + p_o$ at $x_1 = L$ for all t for all times, $\infty > t > > 0$ is $p(x_1) = \rho gh(x_1/L) + p_o$. This result represents a linear variation in the pore fluid pressure from $\rho gh + p_o$ at $x_1 = L$ to p_o at $x_1 = 0$. In the case of unsteady flow a solution to the differential equation (u) for the unsteady situation is sought, subject to the conditions that $p = p_o$ everywhere in the medium and on its boundaries for $t < 0$, that $p = p_o$ at $x = 0$ for all times $\infty > t > 0$ and $p = \rho gh + p_o$ at $x_1 = L$ for all $\infty > t > 0$. Before solving the differential equation (u), it is first rendered dimensionless by introducing the dimensionless pressure ratio P, the dimensionless coordinate X, and the dimensionless time parameter T, thus

$$P = \frac{p - p_o}{\rho gh}, X = \frac{x_1}{L}, \text{and } T = \frac{K_{11}}{L^2} \frac{t}{\tau},$$

respectively. These equations are solved for p, x_1, and t,

$$p = \rho ghP + p_o, x_1 = LX, \text{and } t = \frac{L^2 \tau}{K_{11}} T,$$

and substituted into the differential equation (u) for the unsteady situation which then converts to the dimensionless version of this differential equation

$$\frac{\partial P}{\partial T} = \frac{\partial^2 P}{\partial X^2}.$$

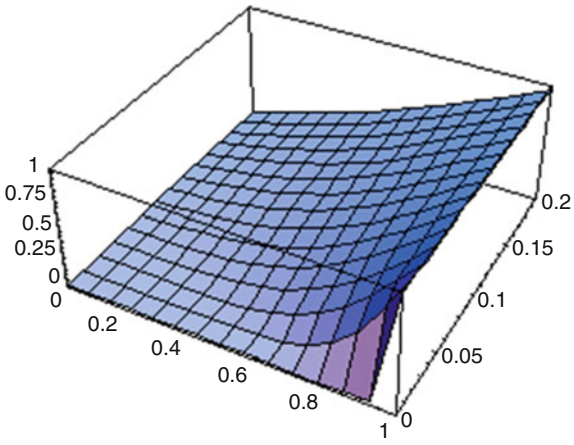


Fig. 6.1 The temporal evolution of the pressure distribution in a layer of rigid porous material. The vertical scale is the dimensionless pressure P with a range of 0 to 1. The front horizontal scale is the dimensionless coordinate X, also with a range of 0 to 1; it traverses the porous layer. The dimensionless time constant T is plotted from 0 to 0.2 in the third direction. The purpose of the plot is to illustrate the evolution of the steady state linear distribution of pressure across the layer of rigid porous material

Note that the steady-state solution $p(x_1) = \rho g h(x_1/L) + p_o$ becomes $P = X$ in the notation of the dimensionless variables. The solution to this dimensionless differential equation for $t \geq 0$ subject to these boundary and initial conditions is

$$P = X - \frac{2}{\pi} \sum_{n=1}^{n=\infty} \frac{1}{n} (-1)^{n-1} \mathrm{e}^{-n^2 \pi^2 T} \sin(n\pi X).$$

Substitution of this solution back into the differential equation above may be used to verify that the solution is a solution to the differential equation. Note that the steady-state solution $P = X$ is recovered as t tends to infinity. The temporal evolution of the steady-state linear distribution of pore fluid pressure across the layer of rigid porous material is illustrated in Fig. 6.1. The vertical scale is the dimensionless pressure P with a range of 0 to 1. The front horizontal scale is the dimensionless coordinate X, also with a range of 0 to 1; it traverses the porous layer. The dimensionless time constant T is plotted from 0 to 0.2 in the third direction.

Problems

6.2.1 Show that the pore fluid density ρ_ϕ satisfies the same differential equation for diffusion, equation (6.11), as the pore fluid pressure p (6.8).

6.2.2 Verify that the form of the rescaled equations (6.10) and (6.17) follow from (6.13) and (6.14). Describe the shape of a homogeneous orthotropic material object O that is in the form of a cube after it is rescaled and distorted so that the differential equation is isotropic.

6.2.3 Show that the solution to the dimensionless differential equation $\frac{\partial P}{\partial T} = \frac{\partial^2 P}{\partial X^2}$ and the prescribed boundary and initial conditions that $p = p_o$ everywhere in the medium and on its boundaries for $t < 0$, that $p = p_o$ at $x = 0$ for for $t \in [-\infty, \infty]$ and $p = p_1$ at $L = h$ for $t \in [0, \infty]$, is given by

$$P = X - \frac{2}{\pi} \sum_{n=1}^{n=\infty} \frac{1}{n} (-1)^{n-1} \mathrm{e}^{-n^2 \pi^2 T} \sin(n\pi X).$$

6.2.4 Record the explicit matrix form for the constitutive relation for the Darcy medium (5.6D).

6.2.5 Record the explicit matrix form for the constitutive relation for a Darcy porous medium in an inhomogeneous transversely isotropic material.

6.2.6 Record the explicit matrix form for the constitutive relation for Darcy's law in a homogeneous isotropic material.

6.3 The Theory of Elastic Solids

An overview of the theory of linear elastic solids can be obtained by considering it as a system of fifteen equations in fifteen scalar unknowns. The fifteen scalar unknowns are the six components of the stress tensor $\mathbf{T}$, the six components of the strain tensor $\mathbf{E}$, and the three components of the displacement vector $\mathbf{u}$. The parameters of an elasticity problem are the tensor of elastic coefficients $\hat{\mathbf{C}}$, the density ρ, and the action-at-a-distance force $\mathbf{d}$, which are assumed to be known. The system of fifteen equations consists of stress–strain relations from the anisotropic Hooke's law,

$$\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{E}}, \text{where } \hat{\mathbf{C}} = \hat{\mathbf{C}}^{\mathrm{T}}. (5.7\mathrm{H}) \text{ and } (5.24\mathrm{H}) \text{ repeated}$$

the six strain–displacement relations,

$$\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u}), (2.49) \text{ repeated}$$

and the three stress equations of motion,

$$\rho \ddot{\mathbf{u}} = \nabla \cdot \mathbf{T} + \rho \mathbf{d}, \mathbf{T} = \mathbf{T}^{\mathrm{T}}. \tag{6.18}$$

This form of the stress equations of motion differs from (3.38) only in notation: the acceleration is here represented by $\ddot{\mathbf{u}}$ rather than $\ddot{\mathbf{x}}$, a result that follows from (2.20) by taking the time derivative twice and assuming that the material and spatial reference frames are not accelerating relative to one another. The system of fifteen equations in fifteen scalar unknowns may be reduced to a system of three equations in three scalar unknowns by accomplishing the following algebraic steps: (1) substitute the strain–displacement relations (2.49) into the stress–strain relations

(5.7H), then (2) substitute the resulting expression relating the stress to the first derivatives of the displacement into the three stress equations of motion (6.18). The result is a system of three equations in three scalar unknowns, the three components of the displacement vector. This algebraic simplification will be accomplished in the case of an isotropic material after Hooke's law for isotropic materials is developed in the next paragraph.

The stress–strain and strain–stress relations of the anisotropic Hooke's law, (5.7H) and (5.24H), respectively, are developed next. It has been shown that the tensor of elastic material coefficients $\hat{\mathbf{C}}$ is symmetric and positive definite, as is its inverse $\hat{\mathbf{S}}$, the compliance tensor of elastic material coefficients. The strain–stress relations (as opposed to the stress–strain relations) are

$$\hat{\mathbf{E}} = \hat{\mathbf{S}} \cdot \hat{\mathbf{T}}, \hat{\mathbf{S}} = \hat{\mathbf{C}}^{-1}, (5.12\text{H}) \text{ repeated}$$

The form and symmetry of $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}$ are identical for any material symmetry and, for the material symmetries of interest, the form appropriate to the material symmetry is given in Tables 4.4 and 4.5. The notation employed thus far for $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}$, the notation that allows their representation as second order tensors in six dimensions, is not the traditional notation. To obtain the traditional notation, Hooke's law (5.7H) is expressed in its matrix format,

$$\begin{bmatrix} \hat{T}_1 \\ \hat{T}_2 \\ \hat{T}_3 \\ \hat{T}_4 \\ \hat{T}_5 \\ \hat{T}_6 \end{bmatrix} = \begin{bmatrix} \hat{c}_{11} & \hat{c}_{12} & \hat{c}_{13} & \hat{c}_{14} & \hat{c}_{15} & \hat{c}_{16} \\ \hat{c}_{21} & \hat{c}_{22} & \hat{c}_{23} & \hat{c}_{24} & \hat{c}_{25} & \hat{c}_{26} \\ \hat{c}_{31} & \hat{c}_{32} & \hat{c}_{33} & \hat{c}_{34} & \hat{c}_{35} & \hat{c}_{36} \\ \hat{c}_{41} & \hat{c}_{42} & \hat{c}_{43} & \hat{c}_{44} & \hat{c}_{45} & \hat{c}_{46} \\ \hat{c}_{51} & \hat{c}_{52} & \hat{c}_{53} & \hat{c}_{54} & \hat{c}_{55} & \hat{c}_{56} \\ \hat{c}_{61} & \hat{c}_{62} & \hat{c}_{63} & \hat{c}_{64} & \hat{c}_{65} & \hat{c}_{66} \end{bmatrix} \begin{bmatrix} \hat{E}_1 \\ \hat{E}_2 \\ \hat{E}_3 \\ \hat{E}_4 \\ \hat{E}_5 \\ \hat{E}_6 \end{bmatrix}, \tag{6.19}$$

and then converted to the traditional three-dimensional component representation by employing the relations (A163),

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ \sqrt{2}T_{23} \\ \sqrt{2}T_{13} \\ \sqrt{2}T_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \sqrt{2}c_{14} & \sqrt{2}c_{15} & \sqrt{2}c_{16} \\ c_{12} & c_{22} & c_{23} & \sqrt{2}c_{24} & \sqrt{2}c_{25} & \sqrt{2}c_{26} \\ c_{13} & c_{23} & c_{33} & \sqrt{2}c_{34} & \sqrt{2}c_{35} & \sqrt{2}c_{36} \\ \sqrt{2}c_{14} & \sqrt{2}c_{24} & \sqrt{2}c_{34} & 2c_{44} & 2c_{45} & 2c_{46} \\ \sqrt{2}c_{15} & \sqrt{2}c_{25} & \sqrt{2}c_{35} & 2c_{45} & 2c_{55} & 2c_{56} \\ \sqrt{2}c_{16} & \sqrt{2}c_{26} & \sqrt{2}c_{36} & 2c_{46} & 2c_{56} & 2c_{66} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \\ \sqrt{2}E_{23} \\ \sqrt{2}E_{13} \\ \sqrt{2}E_{12} \end{bmatrix}, \tag{6.20}$$

and introducing the matrix coefficients c_{ij}, $i, j = 1, \ldots, 6$, defined in Table 6.1. It is easy to verify that the matrix equation (6.20) may be rewritten as

Table 6.1 The elasticity and compliance in different notations

1	2	3	1	2	3
C_{1111}	c_{11}	$\hat{c}_{11}$	S_{1111}	s_{11}	$\hat{S}_{11}$
C_{2222}	c_{22}	$\hat{c}_{22}$	S_{2222}	s_{22}	$\hat{S}_{22}$
C_{3333}	c_{33}	$\hat{c}_{33}$	S_{3333}	s_{33}	$\hat{S}_{33}$
C_{1122}	c_{12}	$\hat{c}_{12}$	S_{1122}	s_{12}	$\hat{S}_{12}$
C_{1133}	c_{13}	$\hat{c}_{13}$	S_{1133}	s_{13}	$\hat{S}_{13}$
C_{2233}	c_{23}	$\hat{c}_{23}$	S_{2233}	s_{23}	$\hat{S}_{23}$
C_{2323}	c_{44}	$\frac{1}{2}\hat{c}_{44}$	S_{2323}	$\frac{1}{4}s_{44}$	$\frac{1}{2}\hat{S}_{44}$
C_{1313}	c_{55}	$\frac{1}{2}\hat{c}_{55}$	S_{1313}	$\frac{1}{4}s_{55}$	$\frac{1}{2}\hat{S}_{55}$
C_{1212}	c_{66}	$\frac{1}{2}\hat{c}_{66}$	S_{1212}	$\frac{1}{4}s_{66}$	$\frac{1}{2}\hat{S}_{66}$
C_{1323}	c_{54}	$\frac{1}{2}\hat{c}_{54}$	S_{1323}	$\frac{1}{4}s_{54}$	$\frac{1}{2}\hat{S}_{54}$
C_{1312}	c_{56}	$\frac{1}{2}\hat{c}_{56}$	S_{1312}	$\frac{1}{4}s_{56}$	$\frac{1}{2}\hat{S}_{56}$
C_{1223}	c_{64}	$\frac{1}{2}\hat{c}_{64}$	S_{1223}	$\frac{1}{4}s_{64}$	$\frac{1}{2}\hat{S}_{64}$
C_{2311}	c_{41}	$\frac{1}{\sqrt{2}}\hat{c}_{41}$	S_{2311}	$\frac{1}{2}s_{41}$	$\frac{1}{\sqrt{2}}\hat{S}_{41}$
C_{1311}	c_{51}	$\frac{1}{\sqrt{2}}\hat{c}_{51}$	S_{1311}	$\frac{1}{2}s_{51}$	$\frac{1}{\sqrt{2}}\hat{S}_{51}$
C_{1211}	c_{61}	$\frac{1}{\sqrt{2}}\hat{c}_{61}$	S_{1211}	$\frac{1}{2}s_{61}$	$\frac{1}{\sqrt{2}}\hat{S}_{61}$
C_{2322}	c_{42}	$\frac{1}{\sqrt{2}}\hat{c}_{42}$	S_{2322}	$\frac{1}{2}s_{42}$	$\frac{1}{\sqrt{2}}\hat{S}_{42}$
C_{1322}	c_{52}	$\frac{1}{\sqrt{2}}\hat{c}_{52}$	S_{1322}	$\frac{1}{2}s_{52}$	$\frac{1}{\sqrt{2}}\hat{S}_{52}$
C_{1222}	c_{62}	$\frac{1}{\sqrt{2}}\hat{c}_{62}$	S_{1222}	$\frac{1}{2}s_{62}$	$\frac{1}{\sqrt{2}}\hat{S}_{62}$
C_{2333}	c_{43}	$\frac{1}{\sqrt{2}}\hat{c}_{43}$	S_{2333}	$\frac{1}{2}s_{43}$	$\frac{1}{\sqrt{2}}\hat{S}_{43}$
C_{1333}	c_{53}	$\frac{1}{\sqrt{2}}\hat{c}_{53}$	S_{1333}	$\frac{1}{2}s_{53}$	$\frac{1}{\sqrt{2}}\hat{S}_{53}$
C_{1233}	c_{63}	$\frac{1}{\sqrt{2}}\hat{c}_{63}$	S_{1233}	$\frac{1}{2}s_{63}$	$\frac{1}{\sqrt{2}}\hat{S}_{63}$

Column 1 illustrates the Voigt notation of these quantities as fourth order tensor components in a three-dimensional Cartesian space. Column 2 represents the Voigt matrix or double index notation. Column 3 illustrates the Kelvin-inspired notation for these quantities as second order tensor components in a six-dimensional Cartesian space

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ T_{23} \\ T_{13} \\ T_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \\ 2E_{23} \\ 2E_{13} \\ 2E_{12} \end{bmatrix}. \tag{6.21}$$

This matrix represents the classical notation of Voigt (1910) for the anisotropic stress–strain relations. Unfortunately the matrix $\mathbf{c}$ appearing in (6.21) does not represent the components of a tensor, while symmetric matrices $\hat{\mathbf{C}}$ and $\hat{\mathbf{S}}$ do represent the components of a second-order tensor in a 6-dimensional space.

A chart relating the component notations of the matrix **c** (and its inverse **s**) to the component notations for $\hat{C}$ and $\hat{S}$ is given in Table 6.1. Table 6.1 also relates these coefficients to the traditional notation for the representation of these tensors as fourth-order tensors in a three-dimensional space, a notation that is not employed in this text. The various components in Table 6.1 are either equal or differ by multiples of $\sqrt{2}$ from each other. In the case of orthotropic symmetry it follows from Table 4.4 and (6.19) through (6.21) that

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ T_{23} \\ T_{13} \\ T_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{22} & c_{23} & 0 & 0 & 0 \\ c_{13} & c_{23} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \\ 2E_{23} \\ 2E_{13} \\ 2E_{12} \end{bmatrix}, \tag{6.22a}$$

or

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{12} & c_{22} & c_{23} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \end{bmatrix}, \tag{6.22b}$$

$$T_{23} = 2c_{44}E_{23}, T_{13} = 2c_{55}E_{13}, T_{12} = 2c_{66}E_{12},$$

and, in the case of isotropic symmetry, it follows again from Table 4.4 and (6.19) through (6.21) that

$$\begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ T_{23} \\ T_{13} \\ T_{12} \end{bmatrix} = \begin{bmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \\ 2E_{23} \\ 2E_{13} \\ 2E_{12} \end{bmatrix}, \tag{6.23}$$

where the coefficients c_{11} and c_{12} are expressed in terms of the Lame´ moduli of elasticity, λ and μ; $c_{11} = \lambda + 2\mu$ and $c_{12} = \lambda$ (note that $\hat{c}_{11} = \lambda + 2\mu$ and $\hat{c}_{12} = \lambda$). In equation (5.11N) the Greek letters λ and μ are also used to denote the viscosity coefficients. This dual use for these Greek letters will continue throughout the text as they are traditional notations in elasticity theory and in viscous fluid theory. The reader should keep in mind that the significance of λ and μ will depend upon context, viscous fluid or elastic solid. Developing the six scalar equations that come from the matrix equation (6.23) in algebraic analogy with the transition from (5.8N)

Table 6.2 The isotropic elastic constants expressed in terms of certain pairs of other isotropic elastic constants

	λ	μ	E	v	k
λ, μ	■	■	$\frac{\mu(3\lambda+2\mu)}{\lambda+\mu}$	$\frac{\lambda}{2(\lambda+\mu)}$	$\frac{3\lambda+2\mu}{3}$
λ, v	■	$\frac{\lambda(1-2v)}{3v}$	$\frac{\lambda(1-v)(1-2v)}{v}$	■	$\frac{\lambda(1-v)}{3v}$
λ, k	■	$\frac{3(k-\lambda)}{2}$	$\frac{9k(k-\lambda)}{3k-\lambda}$	$\frac{\lambda}{3k-\lambda}$	■
μ, E	$\frac{(2\mu-E)\mu}{(E-3\mu)}$	■	■	$\frac{(E-2\mu)}{2\mu}$	$\frac{\mu E}{3(3\mu-E)}$
μ, v	$\frac{2\mu v}{1-2v}$	■	$2\mu(1+v)$	■	$\frac{2\mu(1+v)}{3(1-2v)}$
μ, k	$\frac{3k-2\mu}{3}$	■	$\frac{9k\mu}{3k+\mu}$	$\frac{3k-2\mu}{6k+2\mu}$	■
E, v	$\frac{vE}{(1+v)(1-2v)}$	$\frac{E}{2(1+v)}$	■	■	$\frac{E}{3(1-2v)}$
E, k	$\frac{3k(3k-E)}{9k-E}$	$\frac{3kE}{9k-E}$	■	$\frac{(3k-E)}{6k}$	■
v, k	$3kv/(1+v)$	$\frac{3k(1-2v)}{2(1+v)}$	$3k(1-2v)$	■	■

to (5.11N), a result algebraically equivalent to (5.11N) with **D** replaced by **E** then follows:

$$\mathbf{T} = \lambda(\mathrm{tr}\mathbf{E})1 + 2\mu\mathbf{E}. \tag{6.24}$$

For an isotropic linear elastic material there are just two independent elastic constants. These two constants are represented, for example, by the Lamé moduli λ and μ. Another set of isotropic elastic constants in common use are the Young's modulus E, the shear modulus G $(=\mu)$, and Poisson's ratio v, where the three constants are related by $2G(1+v) = E$ so that only two are independent. Any single isotropic elastic constant can be expressed in terms of any two other isotropic elastic constants as documented by Table 6.2, which contains expressions for most of the usual isotropic elastic constants in terms of different pairs of the other isotropic elastic constants. A frequently employed isotropic elastic constant is the bulk modulus k, which represents the ratio of an applied mean hydrostatic stress, $-p = (\mathrm{tr}\mathbf{T})/3$, to a volumetric strain. Recall that $\mathrm{tr}\mathbf{E}$ represents the volumetric strain per unit volume. The relationship between volumetric strain per unit volume and the mean hydrostatic stress, $-p = (\mathrm{tr}\mathbf{T})/3$, is obtained by taking the trace of (6.24), thus $-3p = (3\lambda + 2\mu)\,\mathrm{tr}\mathbf{E}$. The bulk modulus k is then given by the following different representations,

$$k \equiv \frac{-p}{\mathrm{tr}\mathbf{E}} = \lambda + \frac{2\mu}{3} = \frac{E}{3(1-2v)}; \tag{6.25}$$

the equivalence of the last equality for the bulk modulus k may be seen from the last column of Table 6.2. The isotropic strain–stress relations are obtained from (6.24), thus

$$\mathbf{E} = \frac{1}{E}\{(1+v)\mathbf{T} - v(tr\mathbf{T})1\}. \tag{6.26}$$

Reverting briefly to the case of orthotropic symmetry, the strain–stress relations may be written in the form

$$\begin{bmatrix} E_{11} \\ E_{22} \\ E_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{E_1} & \frac{-v_{21}}{E_2} & \frac{-v_{31}}{E_3} \\ \frac{-v_{12}}{E_1} & \frac{1}{E_2} & \frac{-v_{32}}{E_3} \\ \frac{-v_{13}}{E_1} & \frac{-v_{23}}{E_2} & \frac{1}{E_3} \end{bmatrix} \begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \end{bmatrix} \tag{6.27}$$

$$E_{23} = \frac{T_{23}}{2G_{23}}, E_{13} = \frac{T_{13}}{2G_{13}}, E_{12} = \frac{T_{12}}{2G_{12}},$$

where E_1, E_2, E_3 represent the Young's moduli in the x_1, x_2, x_3 directions; G_{23}, G_{31}, G_{12} represent the shear moduli about the x_1, x_2, x_3 axes and v_{23}, v_{31}, v_{12}, v_{32}, v_{13}, and v_{21} are Poisson's ratios. The Poisson ratio v_{21} represents the strain in the x_1 direction due to the normal strain in the x_2 direction and where symmetry of the compliance tensor requires that

$$\frac{-v_{13}}{E_1} = \frac{-v_{31}}{E_3}, \frac{-v_{13}}{E_1} = \frac{-v_{31}}{E_3}, \frac{-v_{23}}{E_2} = \frac{-v_{32}}{E_3}. \tag{6.28}$$

Biomedical Historical Note: Thomas Young (1773–1829) was a child prodigy, a well educated physician, a physicist and a student of languages who attempted to decipher Egyptian hieroglyphics and who translated the Rosetta Stone. Although he is well known for his concept of the modulus of elasticity, he did significant work in explaining how the eye functioned. He argued that the lens of an eye changed shape to focus light as necessary. He suggested that the retina responded to three principal colors that combined to form all the other colors. More generally he considered the nature of light and discovered the principle of interference of light.

As noted in the introductory paragraph of this section, the system of 15 equations in 15 unknowns can be reduced to a set of three equations in three unknowns or, equivalently, to a single vector equation in three dimensions. The resulting equations are known as the Navier equations of elasticity and they are similar in form to the Navier–Stokes equations of viscous fluid theory developed in the following section. To obtain these equations for an isotropic material one substitutes (2.49) into (6.24) and then places the modified (6.24) for the stresses

in terms of the components of the displacement vector into the equations of motion (6.18), thus

$$\rho\ddot{\mathbf{u}} = (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \mu\nabla^2\mathbf{u} + \rho\mathbf{d}. \tag{6.29}$$

This is the Navier equation. If the Navier equation is solved for the displacement field $\mathbf{u}$, then the strain field $\mathbf{E}$ can be determined from (2.49) and the stress field $\mathbf{T}$ from (6.24). The resulting stress field will satisfy the stress equations of motion (6.18) because (6.29) is an alternate statement of the equations of motion.

Consider now the special case when one chooses the stress (or the strain) as the unknown and the displacement is to be calculated last from the determined strain tensor. In this case one must consider the strain–displacement relations (2.49) to be a system of first-order partial differential equations to be solved for the components of the displacement vector $\mathbf{u}$ given the components of strain tensor $\mathbf{E}$ (see Sect. 2.4). The *conditions of compatibility* in terms of strain, (2.53), (2.54), or (2.55), are a set of necessary and sufficient conditions that the first-order partial differential equations (2.49) must satisfy in order that (2.49) have a single valued and continuous solution $\mathbf{u}$.

The general problem associated with the basic system of fifteen equations is to find the fields $\mathbf{T}(\mathbf{x}, t)$, $\mathbf{E}(\mathbf{x}, t)$, and $\mathbf{u}(\mathbf{x}, t)$ for all $\mathbf{x} \in O$ and $t \in [0, t]$ given a particular object O of density ρ and elastic coefficients $\hat{C}$ (or, in the case of isotropy, λ and μ) acted upon by an action-at-a-distance force $\mathbf{d}$ and some surface loading or specified displacements at the boundary during a specified time interval $[0, t]$. Such problems are called the *initial-boundary value problems* of the theory of elasticity and they are classified in several ways. First, they are classified as either elastostatic, elastoquasi-static, or elastodynamic. The *elastostatic boundary value problems* are those in which $\mathbf{T}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\mathbf{u}(\mathbf{x})$ are independent of time, and the inertia term in the stress equations of motion, $\rho\ddot{\mathbf{u}}$, is zero. The *elastoquasi-static problems* are those in which $\mathbf{T}(\mathbf{x}, t)$, $\mathbf{E}(\mathbf{x}, t)$, and $\mathbf{u}(\mathbf{x}, t)$ are time dependent, but the inertia term in the stress equations of motion, $\rho\ddot{\mathbf{u}}$, is small enough to be neglected. The *elastodynamic initial-boundary value problems* are those in which $\mathbf{T}(\mathbf{x}, t)$, $\mathbf{E}(\mathbf{x}, t)$, and $\mathbf{u}(\mathbf{x}, t)$ are time dependent and the inertia term in the stress equations of motion, $\rho\ddot{\mathbf{u}}$, is neither zero nor negligible.

The formulation of boundary value problems is considered next. The boundary value problems are classified as displacement, traction, and mixed or mixed–mixed boundary value problems. An object O with boundary ∂O is illustrated in Fig. 6.2. The total boundary of the object O is divided into the sum of two boundaries, the displacement boundary $\partial O_{\mathbf{u}}$ over which the boundary conditions are specified in terms of displacement and the traction boundary $\partial O_{\mathbf{t}}$ over which the boundary conditions are specified in terms of the surface tractions $\mathbf{t}$. Note that $\partial O_{\mathbf{t}} \cap \partial O_{\mathbf{u}} = \emptyset$ and $\partial O_{\mathbf{t}} \cup \partial O_{\mathbf{u}} = \partial O$. It is required for some problems to further subdivide the boundaries to include the situation in which the normal tractions and transverse displacements, or transverse tractions and normal displacements, are specified over portions of the object boundary, but that is not done here. The rigid wall indicated in

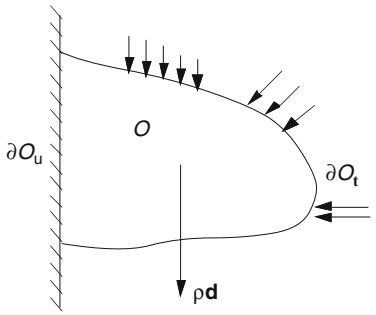


Fig. 6.2 An object denoted by O whose total boundary is divided into traction and displacement boundaries denoted by $\partial O_{\mathbf{t}}$ and $\partial O_{\mathbf{u}}$, $\partial O_{\mathbf{t}} \cap \partial O_{\mathbf{u}} = \emptyset$, $\partial O_{\mathbf{t}} \cup \partial O_{\mathbf{u}} = \partial O$, respectively

Fig. 6.2 is intended to suggest a boundary upon which the boundary condition on the displacement is specified and the other surface is intended to suggest a boundary on which the (zero or nonzero) surface tractions are specified.

In the *displacement boundary value problem* the continuous surface displacement $\mathbf{u}(\mathbf{x}^*, t)$ is specified on the boundary ∂O for $[0, t]$, $\mathbf{x}^* \in \partial O$, where $[0, t]$ is the time interval for which a solution is desired, and the continuous initial displacement field $\mathbf{u}^o(\mathbf{x})$ is specified for all $\mathbf{x} \in O$. The displacement boundary value problem is the following; given a particular object O of density ρ and elastic coefficients $\hat{\mathbf{C}}$ (or, in the case of isotropy, λ and μ) acted upon by an action-at-a-distance force $\mathbf{d}$, determine the fields $\mathbf{u}(\mathbf{x}, t)$, $\mathbf{T}(\mathbf{x}, t)$, $\mathbf{E}(\mathbf{x}, t)$ which satisfy the system of equations (2.49), (6.18) and some form of Hooke's law (5.7H), the initial conditions

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}(\mathbf{x}), \dot{\mathbf{u}}(\mathbf{x}, 0) = \dot{\mathbf{u}}(\mathbf{x}), \mathbf{x} \in O, \tag{6.30}$$

and the displacement boundary condition

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{u}^*(\mathbf{x}^*, t), \mathbf{x}^* \in \partial O, t \in [0, t]. \tag{6.31}$$

In the *traction boundary value problem* the specification of surface displacement (6.31) is replaced by the specification of the surface traction $\mathbf{t}(\mathbf{x}^*, t)$ for all $\mathbf{x}^* \in \partial O$ and $t \in [0, t]$, thus

$$\mathbf{t}(\mathbf{x}^*, t) = \mathbf{T}(\mathbf{x}^*, t)\mathbf{n}, \mathbf{n} \perp \partial O, \mathbf{x}^* \in \partial O, t \in [0, t], \tag{6.32}$$

where $\mathbf{n}$ is the unit exterior normal to the boundary. In a *mixed boundary value problem* there is a portion of the boundary on which the displacements are specified and a portion of the boundary on which the surface tractions are specified. These portions of the boundary are denoted in Fig. 6.2 by $\partial O_{\mathbf{u}}$ and $\partial O_{\mathbf{t}}$ and they are nonempty, non-intersecting portions of the boundary whose union is the entire boundary ∂O, $\partial O_{\mathbf{t}} \cap \partial O_{\mathbf{u}} = \emptyset$, $\partial O_{\mathbf{t}} \cup \partial O_{\mathbf{u}} = \partial O$. The typical mixed-boundary value problem must satisfy the condition (6.31) on $\partial O_{\mathbf{u}}$ and the surface traction condition (6.32) on $\partial O_{\mathbf{t}}$. The mixed–mixed boundary value problem of elasticity is characterized by boundaries where the two types of boundary conditions appear on the same portion of the boundary. For example the normal displacement is

specified and the shear stress is specified on the same portion of the boundary, or the normal stress and the tangential displacement are specified on the same portion of the boundary. The mixed–mixed boundary value problems are complicated and are not considered here.

Any solution to an elasticity problem is unique. That is to say that, for a specific object of a specified material acted upon by a specified action-at-a-distance force and subject to specific boundary conditions, there is one and only one solution to the set of 15 equations in 15 unknowns. The common strategy for proving uniqueness theorems is to assume non-uniqueness, that is to say assume there are two, and then prove that the two must be equal. The uniqueness theorem for linear elasticity is proved using this strategy. Assume that there are two solutions $\mathbf{u}^{(1)}(\mathbf{x}, t)$, $\mathbf{T}^{(1)}(\mathbf{x}, t)$, $\mathbf{E}^{(1)}(\mathbf{x}, t)$, and $\mathbf{u}^{(2)}(\mathbf{x}, t)$, $\mathbf{T}^{(2)}(\mathbf{x}, t)$, $\mathbf{E}^{(2)}(\mathbf{x}, t)$, to the same elasticity problem, that is to say a problem in which the object, the material, the action-at-a-distance force, and the boundary conditions to which the object is subjected, are all specified. The linearity of the system of equations for linear elasticity permits one problem solution for a specified object and material, action-at-a-distance force, traction boundary conditions, and displacement boundary conditions, to be superposed upon a second solution for the same specified object, material and displacement boundary conditions, but for a different action-at-a-distance force, different traction boundary conditions and different displacement boundary conditions. Thus, for example, two solutions, $\mathbf{u}^{(1)}(\mathbf{x}, t)$, $\mathbf{T}^{(1)}(\mathbf{x}, t)$, $\mathbf{E}^{(1)}(\mathbf{x}, t)$ and $\mathbf{u}^{(2)}(\mathbf{x}, t)$, $\mathbf{T}^{(2)}(\mathbf{x}, t)$, $\mathbf{E}^{(2)}(\mathbf{x}, t)$, for the same specified object and material, but different traction boundary conditions, displacement boundary conditions and action-at-a-distance forces, may be added together, $\mathbf{u}^{(1)}(\mathbf{x}, t) + \mathbf{u}^{(2)}(\mathbf{x}, t)$, $\mathbf{T}^{(1)}(\mathbf{x}, t) + \mathbf{T}^{(2)}(\mathbf{x}, t)$, $\mathbf{E}^{(1)}(\mathbf{x}, t) + \mathbf{E}^{(2)}(\mathbf{x}, t)$, to obtain the solution for specified object and material, for the traction boundary conditions and displacement boundary conditions and action-at-a-distance force that are the sum of the two sets of traction boundary conditions, displacement boundary conditions and action-at-a-distance forces. In the proof of uniqueness the principle of superposition is used to define the difference problem obtained by subtracting the two (possibly different) solutions $\mathbf{u}^{(1)}(\mathbf{x}, t)$, $\mathbf{T}^{(1)}(\mathbf{x}, t)$, $\mathbf{E}^{(1)}(\mathbf{x}, t)$, and $\mathbf{u}^{(2)}(\mathbf{x}, t)$, $\mathbf{T}^{(2)}(\mathbf{x}, t)$, $\mathbf{E}^{(2)}(\mathbf{x}, t)$, to the same elasticity problem. The difference problem to which the fields $\mathbf{u}^{(1)}(\mathbf{x}, t) - \mathbf{u}^{(2)}(\mathbf{x}, t)$, $\mathbf{T}^{(1)}(\mathbf{x}, t) - \mathbf{T}^{(2)}(\mathbf{x}, t)$, and $\mathbf{E}^{(1)}(\mathbf{x}, t) - \mathbf{E}^{(2)}(\mathbf{x}, t)$ are a solution is thus a problem for the same specified object and material but for a zero action-at-a-distance force and for zero stress boundary conditions on $\partial O_{\mathbf{t}}$ and zero displacement boundary conditions on $\partial O_{\mathbf{u}}$. The objective is to obtain the solution to this difference problem by considering the work done on a linearly elastic object by the surface tractions and the action-at-a-distance force does this most efficiently. The relation between the work done by the surface tractions and by the "action-at-a-distance force" on the object may be expressed as an integral over the object of the local work done per unit volume, tr**T:E** (see (3.53) and (3.57))**:**

$$\int_{\partial O} \mathbf{t} \cdot \mathbf{u} \, \mathrm{d}a + \int_O \rho \mathbf{d} \cdot \mathbf{u} \, \mathrm{d}v = \int_O \mathrm{tr}\{\mathbf{T} \cdot \mathbf{E}\} \, \mathrm{d}v = \int_O \mathbf{T} : \mathbf{E} \, \mathrm{d}v = \int_O \hat{\mathbf{T}} \cdot \hat{\mathbf{E}} \, \mathrm{d}v. \tag{6.33}$$

The local work done is tr$\mathbf{T}\cdot\mathbf{E}$, which also has the representation $\mathbf{T} : \mathbf{E} = \hat{\mathbf{T}} \cdot \hat{\mathbf{E}}$ (see (3.57) and (A164)). Partitioning the two types of boundaries being considered here, (6.33) may be rewritten as

$$\int_{\partial O_t} \mathbf{t} \cdot \mathbf{u} \, da + \int_{\partial O_u} \mathbf{t} \cdot \mathbf{u} \, da + \int_O \rho \mathbf{d} \cdot \mathbf{u} \, dv = \int_O \hat{\mathbf{T}} \cdot \hat{\mathbf{E}} \, dv. \tag{6.34}$$

Each of the three integrals on the left hand side of (6.34) is zero. The first one is zero because the difference in boundary surface tractions is zero for the difference problem. The second one is zero because the difference in boundary surface displacements is zero for the difference problem. The third one is zero because the difference in "action-at-a-distance" forces is zero for the difference problem. Since the left hand side of (6.34) is zero and since Hooke's law relates the stress and strain by $\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{E}}$, it follows that $\text{tr}(\mathbf{T} \cdot \mathbf{E}) = \hat{\mathbf{T}} \cdot \hat{\mathbf{E}} = \hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}}$; thus for the special case of the difference solution,

$$\int_O (\hat{\mathbf{E}} \cdot \hat{C} \cdot \hat{\mathbf{E}}) \, dv = 0. \tag{6.35}$$

The final step in this proof of uniqueness is to recall that $\hat{\mathbf{C}}$ is positive definite; $\hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}} > 0$ for all nonzero strains $\hat{\mathbf{E}}$, (5.34H). There is a contradiction between (6.35) and the requirement that $\hat{\mathbf{C}}$ be positive definite, $\hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}} > 0$, everywhere in the object unless $\hat{\mathbf{E}} = 0$. This shows that the strain, and therefore the stress, in the difference solution is zero. It does not show that the displacement is zero however. If fact, the displacement may represent any rigid object motion. Thus a solution to a linear elasticity problem is only unique up to a rigid object motion.

The uniqueness theorem of linear elasticity theory is a very important tool in the solution of elasticity problems. From it, the elasticity problem solver knows that if a candidate solution satisfies all the boundary conditions as well as all the 15 elasticity equations, then the candidate solution is a unique solution to the problem. In particular, it allows the elasticity problem solver to guess candidate solutions to elasticity problems, or to make partial guesses. The literature of elasticity does not describe these guessing or semi-guessing methods as guessing, rather it uses more dignified terminology like "the semi-inverse method." The following example is an illustration.

Example 6.3.1

Consider the problem of pure bending of a beam of orthotropic elastic material. The long axis of the beam coincides with the x_3 direction and the bending moment is applied about the x_1 axis (Fig. 6.3). The coordinate system has its origin at the centroid of the cross-sectional area and the 1, 2 axes coincide with the principal axes of the area moment of inertia. Determine formulas for the stress components in terms of the applied moment M_1 and the geometric properties of the cross-section.

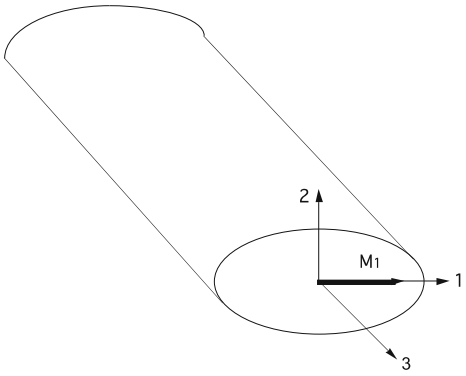


Fig. 6.3 A beam subjected to a pure bending loading. The bending moment M_1 is applied about the 1-axis. There are no other loads. The origin of coordinates is at the centroid of the cross-section. The coordinate axes coincide with the principal axes of the area moment of inertia

Solution: To obtain this solution, it is necessary to account only for the bending moment about the x_1 axis, M_1, and the fact that the lateral surfaces of this beam, that is to say the surfaces other than the ends normal to the x_3 axis, are unloaded, that is to say that there are no surface tractions applied. Because the lateral boundaries are unloaded we are going to guess that stresses that could act on the lateral boundaries are zero everywhere in the object, thus

$$T_{11} = T_{22} = T_{12} = T_{13} = T_{23} = 0.$$

The only nonzero stress is then the axial stress T_{33}. The moment M_1 must then be balanced by a distribution of the axial stress T_{33} in the beam. In general $T_{33} = T_{33}(x_1, x_2, x_3)$; however this may be reduced to $T_{33} = T_{33}(x_1, x_2)$ by observing that T_{33} must be independent of x_3. The argument for T_{33} being independent of x_3 is a physical one. Consider a free object diagram at any location along the length of the beam in Fig. 6.3. The reactive force system at that (any) location must be equal to the moment M_1 applied at the end. Thus the stress distribution must be the same along the entire length of the beam. The only nonzero strains due to T_{33} are then computed from (6.27),

$$E_{11}(x_1, x_2) = \frac{-\nu_{13}}{E_1} T_{33}(x_1, x_2), E_{22}(x_1, x_2) = \frac{-\nu_{23}}{E_2} T_{33}(x_1, x_2), E_{33}(x_1, x_2)$$
$$= \frac{1}{E_3} T_{33}(x_1, x_2).$$

When the compatibility equations (2.54) are applied to these three strains, the following differential equations for T_{33} are obtained:

$$\frac{\partial^2 T_{33}}{\partial x_1^2} = \frac{\partial^2 T_{33}}{\partial x_2^2} = \frac{\partial^2 T_{33}}{\partial x_1 \partial x_2} = 0.$$

The solution to this system of differential equations is $T_{33} = c_o + c_1 x_1 + c_2 x_2$, where c_o, c_1, and c_2 are constants. The next step in the solution to this problem is to

require that the stress distribution T_{33} satisfy the conditions that the axial load on the beam is zero, that the bending moment about the x_1 axis is equal to M_1 and that the bending moment about the x_2 axis is equal to 0. The respective integrals in the 1, 2 plane over the cross-sectional area are given by

$$\int_A T_{33}\ \mathrm{d}A = 0, \int_A T_{33}x_2\ \mathrm{d}A = M_1, \int_A T_{33}x_1\ \mathrm{d}A = 0.$$

The second integral equates M_1 to the couple generated by the integral of the moment of the stress T_{33} times the area patch at a location x_2. The positive direction is determined by the right hand rule. Substituting $T_{33} = c_\mathrm{o} + c_1x_1 + c_2x_2$ into these three integrals, it follows that

$$c_\mathrm{o}\int_A \mathrm{d}A + c_1\int_A x_1\ \mathrm{d}A + c_2\int_A x_2\ \mathrm{d}A = 0, c_\mathrm{o}\int_A x_2\ \mathrm{d}A + c_1\int_A x_1x_2\ \mathrm{d}A + c_2\int_A x_2^2\ \mathrm{d}A = M_1$$

$$c_\mathrm{o}\int_A x_1\ \mathrm{d}A + c_1\int_A x_1^2\ \mathrm{d}A + c_2\int_A x_1x_2\ \mathrm{d}A = 0.$$

These results are simplified by noting the cross-sectional area A, the x_i, $i = 1, 2$, centroid of the cross-sectional area, denoted by $\overline{x}_i$, and the components of the area moment of inertia tensor (A134), I_{11}, I_{22}, I_{12}:

$$A = \int_A \mathrm{d}A, \overline{x}_i = \frac{1}{A}\int_A x_i\mathrm{d}A, I_{11} = \int_A x_2^2\mathrm{d}A, I_{22} = \int_A x_1^2\ \mathrm{d}A, I_{12} = -\int_A x_1x_2\ \mathrm{d}A.$$

Since the origin of coordinates was selected at the centroid, it follows that $\overline{x}_i$ are zero, thus $c_\mathrm{o} = 0$. Then, since the coordinate system has been chosen to be the principal axes of the area moment of inertia, it follows that the product of inertia I_{12} vanishes, thus $c_2 = M_1/I_{11}$ and $c_1 = 0$. It follows that $T_{33} = M_1x_2/I_{11}$.

The solution of the stresses is then $T_{11} = T_{22} = T_{12} = T_{13} = T_{23} = 0$ and $T_{33} = M_1x_2/I_{11}$. The solution for the strains is

$$E_{11} = \frac{-\nu_{13}M_1x_2}{E_1I_{11}}, E_{22} = \frac{-\nu_{23}M_1x_2}{E_2I_{11}}, E_{33} = \frac{M_1x_2}{E_3I_{11}}, E_{12} = E_{13} = E_{23} = 0,$$

and the solution for the displacements may be obtained from the solution for the strains and integration of the strain–displacement relations (2.49). This solution satisfies each of the 15 equations of elasticity and the boundary conditions specified for this problem; thus by the uniqueness theorem, it is the unique solution to this bending problem.

Example 6.3.2

For the problem considered in Example 6.3.1, determine the solution for the displacement field $\mathbf{u}(\mathbf{x}, t)$ from the solution for the strains,

$$E_{11} = \frac{-\nu_{13} M_1 x_2}{E_1 I_{11}}, E_{22} = \frac{-\nu_{23} M_1 x_2}{E_2 I_{11}}, E_{33} = \frac{M_1 x_2}{E_3 I_{11}}, E_{12} = E_{13} = E_{23} = 0,$$

by integration of the strain–displacement relations (2.49).

Solution: From this strain solution and the strain–displacement relations (2.49), a system of six first-order partial differential equations for the components of the displacement vector are obtained:

$$\frac{\partial u_1}{\partial x_1} = \frac{-\nu_{13} M_1 x_2}{E_1 I_{11}}, \frac{\partial u_2}{\partial x_2} = \frac{-\nu_{23} M_1 x_2}{E_2 I_{11}}, \frac{\partial u_3}{\partial x_3} = \frac{M_1 x_2}{E_3 I_{11}}, \frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} = 0, \frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} = 0, \frac{\partial u_3}{\partial x_2} + \frac{\partial u_2}{\partial x_3} = 0.$$

Integration of the first three of these equations yields

$$u_1 = \frac{-\nu_{13} M_1 x_2 x_1}{E_1 I_{11}} + w_1(x_2, x_3), u_2 = \frac{-\nu_{23} M_1 x_2^2}{2 E_2 I_{11}} + w_2(x_1, x_3), u_3 = \frac{M_1 x_2 x_3}{E_3 I_{11}} + w_3(x_1, x_2).$$

This representation for the components of **u** can then be substituted into the second set of three equations above, thus

$$\frac{\partial w_1(x_2, x_3)}{\partial x_2} + \frac{\partial w_2(x_1, x_3)}{\partial x_1} = \frac{\nu_{13} M_1 x_1}{E_1 I_{11}}, \frac{\partial w_1(x_2, x_3)}{\partial x_3} + \frac{\partial w_3(x_1, x_2)}{\partial x_1} = 0$$
$$\frac{\partial w_3(x_1, x_2)}{\partial x_2} + \frac{\partial w_1(x_1, x_3)}{\partial x_3} = -\frac{M_1 x_3}{E_1 I_{11}}.$$

The problem of determining the displacement field **u** is now the problem of determining the functions $w_1(x_2, x_3)$, $w_2(x_1, x_3)$, and $w_3(x_1, x_2)$. Differentiation of the first equation above with respect to x_3, the second equation above with respect to x_2, and the third equation above with respect to x_1, one obtains the following:

$$\frac{\partial^2 w_1(x_2, x_3)}{\partial x_2 \partial x_3} = -\frac{\partial^2 w_2(x_1, x_3)}{\partial x_1 \partial x_3}, \frac{\partial^2 w_1(x_1, x_3)}{\partial x_2 \partial x_3} = -\frac{\partial^2 w_3(x_1, x_2)}{\partial x_1 \partial x_2}$$
$$\frac{\partial^2 w_3(x_1, x_2)}{\partial x_1 \partial x_2} = -\frac{\partial^2 w_2(x_1, x_3)}{\partial x_1 \partial x_3},$$

from which we conclude that $\frac{\partial^2 w_1(x_1, x_3)}{\partial x_2 \partial x_3} = \frac{\partial^2 w_2(x_1, x_3)}{\partial x_1 \partial x_3} = \frac{\partial^2 w_3(x_1, x_2)}{\partial x_1 \partial x_2} = 0$. It follows that each component of **w** must be the sum of two functions, each of a different single variable, thus.

$$w_1(x_2, x_3) = f_{12}(x_2) + f_{13}(x_3), w_2(x_1, x_3) = f_{21}(x_1) + f_{23}(x_3),$$
$$w_3(x_1, x_2) = f_{31}(x_1) + f_{32}(x_2).$$

The substitution of these representations for $\mathbf{w}$ back into the differential equations above yields the following:

$$\frac{\partial f_{12}(x_2)}{\partial x_2}+\frac{\partial f_{21}(x_1)}{\partial x_1}=\frac{\nu_{13}M_1x_1}{E_1I_{11}},\frac{\partial f_{13}(x_3)}{\partial x_3}+\frac{\partial f_{31}(x_1)}{\partial x_1}=0,$$
$$\frac{\partial f_{31}(x_1)}{\partial x_2}+\frac{\partial f_{23}(x_3)}{\partial x_3}=-\frac{M_1x_3}{E_3I_{11}}.$$

A study of the equations above shows that certain combinations of terms are constants, denoted by c_{12}, c_{13}, and c_{32},

$$\frac{\partial f_{12}(x_2)}{\partial x_2}=-Y^{\mathrm{o}}_{12}=-\frac{\partial f_{21}(x_1)}{\partial x_1}+\frac{\nu_{13}M_1x_1}{E_1I_{11}},\frac{\partial f_{13}(x_3)}{\partial x_3}=-Y^{\mathrm{o}}_{13}=-\frac{\partial f_{31}(x_1)}{\partial x_1},$$

$$\frac{\partial f_{32}(x_2)}{\partial x_2}=Y^{\mathrm{o}}_{23}=-\frac{\partial f_{23}(x_3)}{\partial x_3}-\frac{M_1x_3}{E_1I_{11}},$$

thus

$$f_{12}(x_2)=-Y^{\mathrm{o}}_{12}x_2+c_{12},f_{21}(x_1)=Y^{\mathrm{o}}_{12}x_1+\frac{\nu_{13}M_1x_1^2}{2E_1I_{11}}+c_{21},$$
$$f_{13}(x_3)=-Y_{13}x_3+c_{13},f_{31}(x_1)=Y^{\mathrm{o}}_{13}x_1+c_{13},$$
$$f_{32}(x_2)=Y^{\mathrm{o}}_{23}x_2+c_{32},f_{23}(x_3)=-Y^{\mathrm{o}}_{23}x_3-\frac{M_1x_3^2}{2E_3I_{11}}+c_{23}.$$

Substitution of these results into the formulas for $\mathbf{w}$,

$$w_1(x_2,x_3)=-Y^{\mathrm{o}}_{12}x_2-Y^{\mathrm{o}}_{13}x_3+c_{12}+c_{13},$$
$$w_2(x_1,x_3)=\frac{\nu_{13}M_1x_1^2}{2E_1I_{11}}-\frac{M_1x_3^2}{2E_3I_{11}}+Y^{\mathrm{o}}_{12}x_1-Y^{\mathrm{o}}_{23}x_3+c_{21}+c_{23},$$
$$w_3(x_1,x_2)=Y^{\mathrm{o}}_{23}x_2+Y^{\mathrm{o}}_{13}x_1+c_{13}+c_{32},$$

and then into the results of the first integration above, the components of $\mathbf{u}$ are obtained:

$$u_1=\frac{-\nu_{13}M_1x_2x_1}{E_1I_{11}}-Y^{\mathrm{o}}_{12}x_2-Y^{\mathrm{o}}_{13}x_3+u^{\mathrm{o}}_1,$$
$$u_2=-\frac{M_1}{2I_{11}}\left(\frac{x_3^2}{E_3}-\frac{\nu_{23}x_2^2}{E_2}+\frac{\nu_{13}x_1^2}{E_1}\right)+Y^{\mathrm{o}}_{23}x_1-Y^{\mathrm{o}}_{23}x_3+u^{\mathrm{o}}_2,$$
$$u_3=\frac{M_1x_2x_3}{E_1I_{11}}+Y^{\mathrm{o}}_{23}x_2+Y^{\mathrm{o}}_{13}x_1+u^{\mathrm{o}}_3,$$

where $u_1^o=c_{12}+c_{13},u_2^o=c_{21}+c_{23},u_3^o=c_{13}+c_{32}$. The vector $\mathbf{u}^{\mathrm{o}}$ represents a superposed rigid object translation as one can see from the fact that $\mathbf{u}^{\mathrm{o}}$ are the

constant components of **u**; they are the values of **u** at the origin. The rigid object rotation **Y** may be determined from the displacement field **u** using (2.49), thus

$$Y_{12} = Y_{12}^{\mathrm{o}}, Y_{13} = Y_{13}^{\mathrm{o}}, Y_{23} = \frac{M_1 x_3}{I_{11}}\left(\frac{1}{E_1} + \frac{1}{E_3}\right) + Y_{23}^{\mathrm{o}}.$$

The constants $Y_{12}^{\mathrm{o}}, Y_{13}^{\mathrm{o}}, Y_{23}^{\mathrm{o}}$ therefore represent the superposed rigid object rotation, the values of the rigid object rotation at the origin. If the superposed rigid object translation and rotation are zero, then the displacement field **u** is given by

$$u_1 = \frac{-v_{13} M_1 x_2 x_1}{E_1 I_{11}}, u_2 = -\frac{M_1}{2I_{11}}\left(\frac{x_3^2}{E_3} - \frac{v_{23} x_2^2}{E_2} + \frac{v_{13} x_1^2}{E_1}\right), u_3 = \frac{M_1 x_2 x_3}{E_1 I_{11}},$$

and the rigid object rotation **Y** has only one nonzero component,

$$Y_{23} = \frac{M_1 x_3}{I_{11}}\left(\frac{1}{E_1} + \frac{1}{E_3}\right),$$

which represents the rotation along the beam as the distance increases from the beam end at the origin of coordinates. The total rotation between the two ends of the bent beam of length L is then given by

$$Y_{23}^{\text{Total end-to-end rotation}} = \frac{M_1 L}{I_{11}}\left(\frac{1}{E_1} + \frac{1}{E_3}\right).$$

Example 6.3.3
In mechanics of materials the deflection curve for a beam is considered to be the deflection curve for the neutral axis, the neutral axis being by definition the curve that coincides with the centroid of the cross-section at each cross-section. Using the results of the problem considered in Examples 6.3.1 and 6.3.2, determine the formula for the deflection curve for a beam subject to pure bending.

Solution: The displacement of the neutral axis of a beam subjected to pure bending may be determined from the formulas for the displacement field given in Example 6.3.2 above,

$$u_1 = \frac{-v_{13} M_1 x_2 x_1}{E_1 I_{11}}, u_2 = -\frac{M_1}{2I_{11}}\left(\frac{x_3^2}{E_3} - \frac{v_{23} x_2^2}{E_2} + \frac{v_{13} x_1^2}{E_1}\right), u_3 = \frac{M_1 x_2 x_3}{E_1 I_{11}}.$$

The centroid of the beam's cross-section in Examples 6.3.1 was set at the origin of coordinates in the planar cross-section, thus, for the centroid, x_1 and x_2 are zero and

$$u_1 = 0, u_2 = -\frac{M_1 x_3^2}{2I_{11}E_3}, u_3 = 0.$$

The displacement u_2 represents the displacement curve. In the case of isotropy, and in the notation customary of mechanics of materials, the deflection curve is given by

$$y(x) = -\frac{Mx^2}{2EI}.$$

Example 6.3.4
This problem concerns the propagation of elastic plane waves. The problem is to determine the wave speed of a plane wave propagating in a symmetry direction in an orthotropic material. Two kinds of plane waves are to be considered, one in which the oscillating displacement component varies in the direction of propagation and the other in which it varies in a direction perpendicular to the direction of propagation. Let the direction of propagation be the x_1 direction and let the direction perpendicular to the direction of propagation be the x_2 direction. Both the x_1 and the x_2 directions are directions of the symmetry axes of the orthotropic material. The displacement that varies in the direction of its propagation is $u_1 = u_1(x_1, t)$ and represents an time varying longitudinal motion of axial compression or tension or some combination of the two. The displacement that varies in the perpendicular direction is $u_1 = u_1(x_2, t)$ and represents a time varying shearing motion. The motion $u_1 = u_1(x_1, t)$ is called the longitudinal (L) motion and $u_1 = u_1(x_2, t)$, the shear (S) motion. Neglect the action-at-a-distance force.

Solution: The differential equations governing these two motions are obtained from the governing set of elasticity equations, the first three equations in this section. First, from the strain–displacement relations (2.49) it follows that all the strain components but one is zero for both the longitudinal and shearing motions and that the nonzero components are given by

$$E_{11} = \frac{\partial u_1}{\partial x_1} \text{ for } (L) \text{ and } E_{12} = \frac{1}{2}\frac{\partial u_1}{\partial x_2} \text{ for } (S).$$

Second, using these two results in the stress–strain relations for orthotropic materials (6.22), it follows that

$$T_{11} = c_{11}\frac{\partial u_1}{\partial x_1}, T_{22} = c_{12}\frac{\partial u_1}{\partial x_1}, T_{33} = c_{13}\frac{\partial u_1}{\partial x_1}, \text{for } (L) \text{ and } T_{12} = c_{66}\frac{\partial u_1}{\partial x_2} \text{ for } (S).$$

Third, upon substitution of these stresses and the functional form of the two motions, $u_1 = u_1(x_1, t)$ and $u_1 = u_1(x_2, t)$, in the stress equations of motion (6.18) one obtains the differential equations

$$\frac{\partial^2 u_1}{\partial t^2} = c_{\mathrm{L}}^2 \frac{\partial^2 u_1}{\partial x_1^2} \text{ for } (L) \text{ and } \frac{\partial^2 u_1}{\partial t^2} = c_{\mathrm{S}}^2 \frac{\partial^2 u_1}{\partial x_2^2} \text{ for } (S),$$

where the notations

$$c_{\mathrm{L}}^2 = \frac{c_{11}}{\rho} \text{ for } (L) \text{ and } c_{\mathrm{S}}^2 = \frac{c_{66}}{\rho} \text{ for } (S),$$

have been introduced. The solution of these equations is so very similar that only the longitudinal case is considered. These are differential equations whose solutions can be explicitly calculated. Consider the differential equation for the longitudinal wave and let $\xi = x_1 + c_{\mathrm{L}}t$ and $\eta = x_1 - c_{\mathrm{L}}t$. The second derivatives that appear in the differential equation are obtained using the chain rule, thus

$$\frac{\partial^2 u_1}{\partial x_1^2} = \frac{\partial^2 u_1}{\partial \xi^2} + 2\frac{\partial^2 u_1}{\partial \xi \partial \eta} + \frac{\partial^2 u_1}{\partial \eta^2} \quad \text{and} \quad \frac{\partial^2 u_1}{\partial t^2} = c_{\mathrm{L}}^2 \left[\frac{\partial^2 u_1}{\partial \xi^2} - 2\frac{\partial^2 u_1}{\partial \xi \partial \eta} + \frac{\partial^2 u_1}{\partial \eta^2} \right].$$

When these expressions are substituted into the differential equation for the longitudinal wave it reduces to

$$\frac{\partial^2 u_1}{\partial \xi \partial \eta} = 0,$$

which has the solution

$$u_1(x_1, t) = u_1(\xi, \eta) = p(\xi) + q(\eta) = p(x_1 + c_{\mathrm{L}}t) + q(x_1 - c_{\mathrm{L}}t).$$

The transformation $\xi = x_1 + c_{\mathrm{L}}t$ represents a translation of the coordinate system in the $+x_1$ direction by the amount $c_{\mathrm{L}}t$. Since this translation is proportional to the time, a point $\xi = x_1 + c_{\mathrm{L}}t$ held constant means $\frac{\partial x_1}{\partial t} = -c_{\mathrm{L}}$, thus $\xi =$ constant moves in the $-x_1$ direction with speed c_{L}. A solution of the form $u_1(x_1, t) = p(\xi) = p(x_1 + c_{\mathrm{L}}t)$ represents a wave traveling with velocity $-c_{\mathrm{L}}$ without changing its shape. For example $u_1(x_1, t) = \sin(x_1 + c_{\mathrm{L}}t)$ represents a sine wave traveling with velocity $-c_{\mathrm{L}}$. Similarly $u_1(x_1, t) = q(\eta) = q(x_1 - c_{\mathrm{L}}t)$ represents a wave traveling with velocity $+c_{\mathrm{L}}$ without changing its shape. Thus the solution $u_1(x_1, t) = u_1(\xi, \eta) = p(\xi) + q(\eta) = p(x_1 + c_{\mathrm{L}}t) + q(x_1 - c_{\mathrm{L}}t)$ of the differential equation for the longitudinal wave is the sum of a wave traveling to the left with velocity $-c_{\mathrm{L}}$ and one traveling to the right with velocity $+c_{\mathrm{L}}$. Since the two waves travel in opposite directions, the shape of $u_1(x_1, t)$ will in general change with time.

The initial-boundary value problem is composed of the differential equation for the longitudinal wave and the initial conditions $u_1(x_1, 0) = f(x_1)$ and $\frac{\partial u_1}{\partial t}(x_1, 0) = g(x_1)$ for $0 < x_1 < \infty$. These initial conditions determine the form of the functions p and q in the solution. From the solution $u_1(x_1, t) = p(x_1 + c_{\mathrm{L}}t) + q(x_1 - c_{\mathrm{L}}t)$ and the chain rule it follows that

$$p(x_1) + q(x_1) = f(x_1), c_{\mathrm{L}}\frac{\partial p}{\partial \xi}(x_1) - c_{\mathrm{L}}\frac{\partial q}{\partial \eta}(x_1) = g(x_1) \text{ for } 0 < \mathrm{x}_1 < \infty.$$

Combining these two equations, it follows that $2c_{\mathrm{L}}\frac{\partial p}{\partial \xi}(x_1) = c_{\mathrm{L}}\frac{\partial f}{\partial \eta}(x_1) + g(x_1)$ and; integration of these two equations yields

$$p(\xi) = \frac{1}{2}f(\xi) + \frac{1}{2c_{\mathrm{L}}}\int_0^{\xi} g(\zeta)\,\mathrm{d}\zeta + \text{constant} \quad \text{and}$$
$$q(\eta) = \frac{1}{2}f(\eta) - \frac{1}{2c_{\mathrm{L}}}\int_0^{\eta} g(\zeta)\,\mathrm{d}\zeta - \text{constant}$$

Substitution of these results for the functions $p(\xi)$ and $q(\eta)$ into $u_1(x_1,t) = p(x_1 + c_{\mathrm{L}}t) + q(x_1 - c_{\mathrm{L}}t)$ yields

$$\begin{aligned} u_1(x_1,t) &= p(x_1 + c_{\mathrm{L}}t) + q(x_1 - c_{\mathrm{L}}t) \\ &= \frac{1}{2}[f(x_1 + c_{\mathrm{L}}t) + f(x_1 - c_{\mathrm{L}}t)] \\ &\quad + \frac{1}{2c_{\mathrm{L}}}\left[\int_0^{x_1+c_{\mathrm{L}}t} g(\zeta)\,\mathrm{d}\zeta - \int_0^{x_1-c_{\mathrm{L}}t} g(\zeta)\,\mathrm{d}\zeta\right] \\ &= \frac{1}{2}[f(x_1 + c_{\mathrm{L}}t) + f(x_1 - c_{\mathrm{L}}t)] + \frac{1}{2c_{\mathrm{L}}}\int_{x_1-c_{\mathrm{L}}t}^{x_1+c_{\mathrm{L}}t} g(\zeta)\,\mathrm{d}\zeta, \end{aligned}$$

where the constant terms in this expression that are inherited from the functions $p(\xi)$ and $q(\eta)$ that are zero due to the initial condition $u_1(x_1,0) = f(x_1)$. A similar result holds for the S wave; one has only to change the c_{L} to c_{S} and the subscript on x_1 from 1 to 2 to obtain the S result. The difference between the two results is that the shear or S wave is a propagating shearing motion, as opposed to a propagating compression/tension motion, and that it travels at a different wave speed. If the same results were obtained for an isotropic elastic material, the wave speeds would be

$$c_{\mathrm{L}}^2 = \frac{\lambda + 2\mu}{\rho} \text{ for } (L) \text{ and } c_{\mathrm{S}}^2 = \frac{2\mu}{\rho} \text{ for } (S).$$

Problems

6.3.1. Verify that (6.21) may be determined directly from (6.20).
6.3.2. Record the form of (6.22) for transversely isotropic materials.
6.3.3. Calculate (6.24) from (6.23).
6.3.4. Record the explicit matrix form for the constitutive relation for Hooke's law (5.6H).
6.3.5. Record the explicit matrix form for the constitutive relation for Hooke's law in a homogeneous orthotropic material.
6.3.6. Record the explicit matrix form for the constitutive relation for a transversely isotropic, homogeneous viscoelastic material.
6.3.7. Verify the Navier equation (6.29) for isotropic linear elasticity. Accomplish this by substituting (2.49) into (6.24) and then place the modified (6.24) for

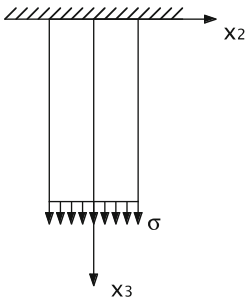


Fig. 6.4 See Problem 6.3.11

the stresses in terms of the components of the displacement vector into the equations of motion (6.18).

6.3.8. Use (6.23) and (5.34H) to determine the restrictions on the coefficients λ and μ so that $\hat{\mathbf{E}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{E}} > 0$ for all nonzero strains $\hat{\mathbf{E}}$.

6.3.9. Use the result of the previous problem and Table 6.2 to obtain restrictions on the values of the isotropic Poisson's ratio ν.

6.3.10. Show that, in an isotropic elastic material the principal axes of stress and strain always coincide. {Hint: Recall that the principal axes of stress (strain) are characterized by the vanishing of the shearing stresses (shearing strains)}.

6.3.11. Prove that the principal axes of stress and the principal axes of strain cannot coincide in a triclinic material.

6.3.12. Under what conditions do the principal axes of stress and the principal axes of strain not coincide for an orthotropic elastic material?

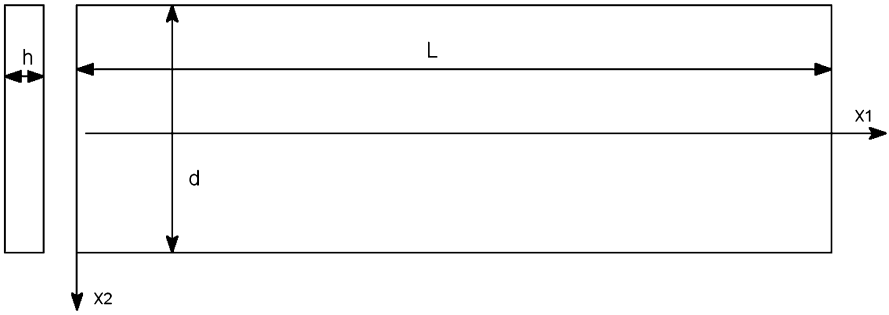
6.3.13. Verify that the conditions of compatibility in terms of strain (2.53) or (2.54) are identities by substituting the components of (2.49) or (2.52) into (2.54).

6.3.14. The bar shown in Fig. 6.4 is made of an orthotropic material. It is fixed to a rigid surface at $x_3 = 0$; the origin of the coordinate system is positioned at the center of the bar where the bar is attached to the rigid surface. It is subjected to constant stress σ acting in the x_3 direction along its lower surface. The orthotropic elastic constants appear in the strain–stress equations (6.27) and (6.28).

Assume a stress state in the bar and then calculate the strain state in the bar. Next calculate the displacement field. If this calculation takes more time than you have, write out the steps you would take to find the solution and guess what the solution is for the displacement field.

6.3.15. As an extension of Example 6.3.1, show that if the beam in Fig. 6.3 is bent about both the x_1 and the x_2 axes, and if is also subjected to an axial tensile load of magnitude P, then $T_{33} = P/A + M_1 x_2/I_{11} - M_2 x_1/I_{22}$.

6.3.16. This problem is a plane stress problem, which means that the stressed domain is a thin plate of thickness h under the action of forces applied at the boundary and organized so that their directions all lie in the plane of the plate. The domain is a rectangular region of length L and width d.



With respect the coordinate system shown the displacements in the 1 and 2 directions are given by

$$u_1 = -\frac{A}{2}x_1^2 x_2 + \frac{A}{3}\left(1+\frac{v}{2}\right)x_2^3 + \frac{A}{2}\left[L^2 - (1+v)\frac{d^2}{2}\right]x_2,$$
$$u_2 = \frac{vA}{2}x_1 x_2^2 + \frac{A}{6}x_1^3 - \frac{AL^2}{2}x_1 + \frac{AL^3}{3}.$$

In the case of plane stress the strain–stress relations reduce to the following:

$$E_{11} = \frac{1}{E}(T_{11} - vT_{22}), E_{22} = \frac{1}{E}(T_{22} - vT_{11}), E_{12} = \frac{(1+v)}{E}T_{12}.$$

(a) Calculate the strain field in the rectangular region of length L and width d.
(b) Calculate the stress field in the rectangular region of length L and width d.
(c) Calculate the values of the stress field at the surfaces $x_2 = \pm\frac{d}{2}$ of the rectangular region of length L and width d.
(d) Calculate the stress applied on the surfaces $x_2 = \pm\frac{d}{2}$ of the rectangular region of length L and width d.
(e) Calculate the stress applied on the surface $x_1 = 0$ of the rectangular region of length L and width d.
(f) Name the equation that was employed in the two previous calculations.

6.3.17 Show that the differential equation $\frac{\partial^2 u_1}{\partial t^2} = c_S^2 \frac{\partial^2 u_1}{\partial x_2^2}$ and the initial conditions that require that the displacement and velocity at time $t = 0$ to be given by $u_1(x_2, 0)$ and $(\partial u_1/\partial t)(x_2, 0)$ are identically satisfied by the solution

$$u_1(x_2,t) = \frac{1}{2}[u_1(x_2 + c_S t, 0) + u_1(x_2 - c_S t, 0)] + \frac{1}{2c_S}\int_{x_2 - c_S t}^{x_2 + c_S t} \frac{\partial u_1}{\partial t}(\xi, 0)\, \mathrm{d}\xi.$$

6.4 The Theory of Viscous Fluids

An overview of the theory of viscous fluids, without temperature effects, can be obtained by considering it as a system of seventeen equations in seventeen scalar unknowns. The seventeen scalar unknowns are the six components of the stress tensor $\mathbf{T}$, the fluid pressure p, the fluid density ρ, the six components of the rate-of-deformation tensor $\mathbf{D}$, and the three components of the velocity vector $\mathbf{v}$. The parameters of a viscous fluid problem are the viscosity coefficients λ and μ ($3\lambda + 2\mu > 0, 2\mu > 0$) and the action-at-a-distance force $\mathbf{d}$, which are assumed to be known. The system of seventeen equations consists of a constitutive equation relating the density ρ to the pressure p, $\rho = \rho(p)$ (and, in thermal-viscous problems, to the temperature), the six equations of the Newtonian law of viscosity,

$$\mathbf{T} = -p1 + \lambda(\mathrm{tr}\mathbf{D})1 + 2\mu\mathbf{D}, \qquad (5.11\mathrm{N}) \text{ repeated}$$

the six rate-of-deformation-velocity relations,

$$\mathbf{D} = (1/2)((\nabla \otimes \mathbf{v})^{\mathrm{T}} + \nabla \otimes \mathbf{v}), \qquad (2.32) \text{ repeated}$$

the one equation of the conservation of mass,

$$\dot{\rho} + \rho(\nabla \cdot \mathbf{v}) = 0, \qquad (3.5) \text{ repeated}$$

and the three stress equations of motion,

$$\rho\dot{\mathbf{v}} = \nabla \cdot \mathbf{T} + \rho\mathbf{d}, \mathbf{T} = \mathbf{T}^{\mathrm{T}}. \qquad (6.36)$$

This form of the stress equations of motion differs from (3.37) and (6.18) only in notation: the acceleration is here represented by $\dot{\mathbf{v}}$ rather than $\ddot{\mathbf{x}}$ or $\ddot{\mathbf{u}}$, respectively, a result that follows from (2.20) and (2.24). The system of seventeen equations in seventeen scalar unknowns may be reduced to a system of four equations in four scalar unknowns, the pressure p and the three components of the velocity $\mathbf{v}$, by accomplishing the following algebraic steps: (1) substitute the rate-of-deformation-velocity (2.32) into the stress–strain relations (5.11N), then (2) substitute the resulting expression relating the stress to the first derivatives of the velocity into the three stress equations of motion (6.36). The result is a system of three equations in three scalar unknowns, the three components of the velocity vector $\mathbf{v}$:

$$\rho\dot{\mathbf{v}} = -\nabla p + (\lambda + \mu)\nabla(\nabla \cdot \mathbf{v}) + \mu\nabla^2\mathbf{v} + \rho\mathbf{d}. \qquad (6.37)$$

This is the Navier–Stokes equation for viscous fluid flow. Substituting the barotropic relation $\rho = \rho(p)$ into the conservation of mass (3.5) yields the fourth equation in the set of four equations for the four unknowns, p and the components of $\mathbf{v}$,

$$\frac{\partial \rho}{\partial p}\dot{p} + \rho(p)(\nabla \cdot \mathbf{v}) = 0. \tag{6.38}$$

This system of equations will become more complicated if thermal effects are considered. They may also be simplified in different ways. An easy simplification is to assume that the viscosity of the fluid is small and can be neglected (that is to say the fluid is assumed to be inviscid), then λ and μ are set equal to zero, (6.37) becomes

$$\rho\dot{\mathbf{v}} = -\nabla p + \rho\mathbf{d}, \tag{6.39}$$

and (6.38) is unchanged. The system of equations (6.37) and (6.38) may also be simplified using the assumption of incompressibility of the fluid. An incompressible material is one which is not permitted to have changes in its volume, $\mathrm{tr}\mathbf{D} = \nabla\cdot\mathbf{v} = 0$. If the volume cannot change, the density ρ of the fluid cannot change. It follows that the barotropic relationship $\rho = \rho(p)$ is not appropriate and the pressure is no longer determined by the density. The pressure field p in an incompressible material is a Lagrange multiplier (see Example 6.4.1) that serves the function of maintaining the incompressibility constraint, $\nabla\cdot\mathbf{v} = 0$. Because the volume of the fluid cannot change, p does no work on the fluid; it is a function of $\mathbf{x}$ and t, $p(\mathbf{x}, t)$, to be determined by the solution of the system of differential equations and boundary/initial conditions. The reduced Navier–Stokes equation for viscous fluid flow and incompressibility constraint now becomes a system of four equations

$$\rho\dot{\mathbf{v}} = -\nabla p + \mu\nabla^2\mathbf{v} + \rho\mathbf{d}, \nabla \cdot \mathbf{v} = 0 \tag{6.40}$$

for the four unknown fields, the three components of $\mathbf{v}(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$. The typical boundary condition applied in viscous fluid theory is the "no slip" condition. This condition requires that a viscous fluid at a solid surface must stick to the surface and have no velocity, $\mathbf{v}(\mathbf{x}^*, t) = 0$ for $\mathbf{x}^* \in \partial O_s$, where ∂O_s stands for the solid boundary of the fluid domain.

Example 6.4.1 Pressure as a Lagrange Multiplier in Incompressible Fluids
The constraint of incompressibility is imposed using a Lagrange multiplier. In order to describe what this means and how it is accomplished, Lagrange's method of calculating extrema in problems in which there is a constraint summarized briefly. Lagrange's method is for the solution of a type of problem in which one must find the extremal values of a function $f(x, y, z)$ subject to the constraint $g(x, y, z) = c$. The solution to the problem is obtained by forming the function $q(x, y, z) = f(x, y, z) + \Lambda g(x, y, z)$ where Λ is a constant, called the Lagrange multiplier, whose value is to be determined. Treating x, y, and z as independent variables four independent conditions $\partial q/\partial x = 0$, $\partial q/\partial y = 0$, $\partial q/\partial z = 0$, and $g(x, y, z) = c$ are available to find the four unknowns, x, y, z, and Λ. As an example of the application of Lagrange's method, consider the problem of finding the maximum or minimum distance from the point (a, b, c) to a point on the surface of the unit sphere, $x^2 + y^2 + z^2 = 1$. The function that is to be extremized is the square of the distance between

(x, y, z) and (a, b, c), $f(x, y, z) = (x-a)^2 + (y-b)^2 + (z-c)^2$. This function is to be extremized subject to the constraint that (x, y, z) be a point on the surface of the unit sphere, $x^2 + y^2 + z^2 = 1$, thus $q(x, y, z) = (x-a)^2 + (y-b)^2 + (z-c)^2 + \Lambda(x^2 + y^2 + z^2-1)$. Setting the derivatives of $q(x, y, z)$ with respect to x, y, and z equal to zero it follows that $x = a/(\Lambda + 1)$, $y = b/(\Lambda + 1)$, $z = c/(\Lambda + 1)$. From the constraint condition $x^2 + y^2 + z^2 = 1$, it follows that $\Lambda + 1 = \pm\sqrt{(a^2 + b^2 + c^2)}$. In the special case when $(a, b, c) = (2, 0, 0)$, $(\Lambda + 1) = \pm 2$, and $(x, y, z) = (1, 0, 0)$ or $(x, y, z) = (-1, 0, 0)$; thus the point on the unit sphere closest to $(2, 0, 0)$ is $(1, 0, 0)$ and the point on the unit sphere furthest from $(2, 0, 0)$ is $(-1, 0, 0)$.

With this background the problem is to employ Lagrange's method to minimize the dissipation due to the rate of volume change, as opposed to the dissipation due to the shearing motion, in a viscous fluid. From equation (5.23N) the stress power or dissipation is given by $\mathbf{T} : \mathbf{D} = -p(\mathrm{tr}\mathbf{D}) + (3\lambda + 2\tilde{\mu}3)(\mathrm{tr}\mathbf{D})^2 + 2\mu\mathrm{tr}(\mathrm{dev}\mathbf{D})^2$. The constraint condition is that $\mathrm{tr}\mathbf{D} = 0$. The problem is to show that when this constraint is imposed, the pressure becomes a constant Lagrange multiplier.

Solution: In this case $q(\mathrm{tr}\mathbf{D}) = -p(\mathrm{tr}\mathbf{D}) + ((3\lambda + 2\mu)/3)(\mathrm{tr}\mathbf{D})^2 + 2\mu\mathrm{tr}(\mathrm{dev}\mathbf{D})^2 + \Lambda(\mathrm{tr}\mathbf{D})$; thus from $\partial q/\partial(\mathrm{tr}\mathbf{D}) = (\Lambda-p) + (2(3\lambda + 2\mu)/3)(\mathrm{tr}\mathbf{D}) = 0$, $\partial\mathrm{tr}(\mathrm{dev}\mathbf{D})^2/\partial(\mathrm{tr}\mathbf{D}) = 0$, and the constraint condition $\mathrm{tr}\mathbf{D} = 0$, it follows that $\Lambda = p$.

In a paper of 1883 Sir Osborne Reynolds showed that the transition between laminar flow governed by the Newtonian law of viscosity, and the form of the Navier–Stokes equations considered here, and the chaotic flow called turbulence depended upon a dimensionless number that is now called the Reynolds number. The Reynolds number R is equal to $\rho Vd/\mu$ where ρ and μ are the density and viscosity of the fluid, respectively, and V and d are a representative velocity and a representative length of the problem under consideration, respectively. Only laminar flows of viscous fluids are considered in this book, hence there is always a certain value of a Reynolds number for which the solution no longer describes the physical situation accurately.

Example 6.4.2 Couette Flow

Consider an incompressible viscous fluid of viscosity μ in the domain between two infinite flat solid plates at $x_2 = \pm h/2$. Action-at-a-distance forces are not present and the plate at $x_2 = h/2$ is moving at constant velocity V in the positive x_1 direction. There is no pressure gradient. Determine the velocity distribution and the stress that must be applied to the top plate to maintain its motion.

Solution: Assume that the only nonzero velocity component is in the x_1 direction and that it depends only upon the x_2 coordinate, $v_1 = v_1(x_2)$. This velocity field automatically satisfies the incompressibility condition, $\nabla\cdot\mathbf{v} = 0$, and the reduced Navier-Stokes equations are

$$\frac{\partial v_1}{\partial x_2^2} = 0 \text{ and } \nabla p = 0.$$

The solution to the first of these differential equations, which is in fact an ordinary differential equation, is given by $v_1(x_2) = c_1x_2 + c_2$. The no-slip boundary conditions are that $v_1(-h/2) = 0$ and $v_1(h/2) = V$, thus $c_1 = V/h$ and $c_2 = V/2$, thus $v_1(x_2) = (V/h)(x_2 + h/2)$. The solution to the second of these differential equations is that p is a constant. The stress at the top plate is given by the T_{12} component of (5.11N),

$$T_{12} = \mu \frac{\partial v_1}{\partial x_2} = \frac{\mu V}{h}.$$

Example 6.4.3 Plane Poiseuille Flow
Consider an incompressible viscous fluid of viscosity μ in the domain between two infinite flat solid plates at $x_2 = \pm h/2$. Action-at-a-distance forces are not present and the plates are not moving. The flow is steady and there is a constant pressure drop in the x_1 direction given by $\partial p/\partial x_1$. Determine the velocity distribution.

Solution: As in the previous example, assume that the only nonzero velocity component is in the x_1 direction and that it depends only upon the x_2 coordinate, $v_1 = v_1(x_2)$. This velocity field automatically satisfies the incompressibility condition, $\nabla \cdot \mathbf{v} = 0$, and the reduced Navier-Stokes equation is

$$\mu \frac{\partial^2 v_1}{\partial x_2^2} = \frac{\partial p}{\partial x_1}.$$

The solution to this differential equation, which is again an ordinary differential equation, is given by

$$v_1 = \frac{1}{2\mu} \frac{\partial p}{\partial x_1} x_2^2 + c_3 x_2 + c_4.$$

The boundary conditions are that $v_1(\pm h/2) = 0$, thus $c_3 = 0$ and $c_4 = -\frac{1}{2\mu} \frac{\partial p}{\partial x_1} \left(\frac{h}{2}\right)^2$, and it follows that the velocity profile is parabolic in shape,

$$v_1 = \frac{1}{2\mu} \left(-\frac{\partial p}{\partial x_1}\right) \left\{ \left(\frac{h}{2}\right)^2 - x_2^2 \right\},$$

which is written with a minus sign in front of the pressure gradient to emphasize that the pressure gradient is negative in the direction of flow or, equivalently, the pressure is dropping in the direction of flow. The volume flow rate per unit length Q is given by

$$Q = \int_{-h/2}^{h/2} v_1 \, \mathrm{d}x_2 = \frac{1}{2\mu} \left(-\frac{\partial p}{\partial x_1}\right) \int_{-h/2}^{h/2} \left\{ \left(\frac{h}{2}\right)^2 - x_2^2 \right\} \mathrm{d}x_2 = \left(-\frac{\partial p}{\partial x_1}\right) \frac{h^3}{12\mu}.$$

In the case of pipe flow under a steady pressure gradient the differential equation corresponding to $\mu \frac{\partial^2 v_1}{\partial x_2^2} = \frac{\partial p}{\partial x_1}$ is, in cylindrical coordinates r and z (see Sect. A.14), is given by $\frac{\mu}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right) = \frac{\partial p}{\partial z}$ where the pressure gradient is assumed to be a constant. The solution to this equation in cylindrical coordinates subject to the "no slip" boundary condition at the pipe wall is a similar parabolic profile to the one obtained above

$$v_z = \left(-\frac{\partial p}{\partial z}\right)\frac{r_o^2}{4\mu}\left\{1 - \frac{r^2}{r_o^2}\right\},$$

where r_o is the radius of the pipe and r and z are two of the three cylindrical coordinates. The volume flow rate is given by

$$Q = 2\pi \int_0^{r_o} r v_z \, dr = \overline{v}\pi r_o^2 = \left(-\frac{\partial p}{\partial z}\right)\frac{\pi r_o^4}{8\mu},$$

where $\overline{v}$ is the mean velocity.

Historical Note: The solution to the Navier–Stokes equations for steady flow in a pipe is called Poiseuille flow after Jean-Louis-Marie Poiseuille (1799–1869), a Parisian physician and physiologist interested in the flow of blood. Poiseuille received his medical degree in 1828 and established his practice in Paris. He developed an improved method for measuring blood pressure. He also is believed to be the first to have used the mercury manometer to measure blood pressure. In the 1840s Poiseuille experimentally determined the basic properties of steady laminar pipe flow using water as a substitute for blood. The formula for Q above is rewritten below with the negative pressure gradient expressed as the change in pressure Δp along the entire length of pipe divided by the pipe length L, thus

$$Q = \overline{v}\pi r_o^2 = \frac{\Delta p \pi r_o^4}{8\mu L}.$$

This formula had not been derived when Poiseuille did his very careful experimental work, which demonstrated its principal features using capillary tubes of glass (models of the blood capillary vessels). Poiseuille showed the volume flow rate Q was proportional to the pressure drop along the pipe Δp, to the fourth power of the radius r_o of the pipe and inversely proportional to the length of the pipe, L. In honor of Poiseuille the unit of viscosity is call the poise. The poise has the symbol P and it is equal to one (dyne-second)/(centimeter)^2 or 0.1 Pa−s.

Problems

6.4.1. Verify the calculation of the Navier–Stokes equations (6.37) by (1) substituting the rate-of-deformation-velocity (2.32) into the stress–strain relations (5.11N), then (2) substituting the resulting expression relating the stress to the first derivatives of the velocity into the three stress equations of motion (6.36).

6.4.2. Find the constant c in $\frac{\partial \mathrm{tr}\mathbf{D}^2}{\partial \mathrm{tr}\mathbf{D}} = \frac{\partial D_{ij}D_{ji}}{\partial D_{kk}} = c\,\mathrm{tr}\mathbf{D}$

6.4.3. Prove $\frac{\partial \mathrm{tr}(\mathrm{dev}\mathbf{D})^2}{\partial \mathrm{tr}\mathbf{D}} = 0$.

6.4.4. Determine the shear stress acting on the lower plate in example 6.4.2.

6.4.5. Determine a formula for the stress exerted upon the upper plate during plane Poiseuille flow (example 6.4.3).

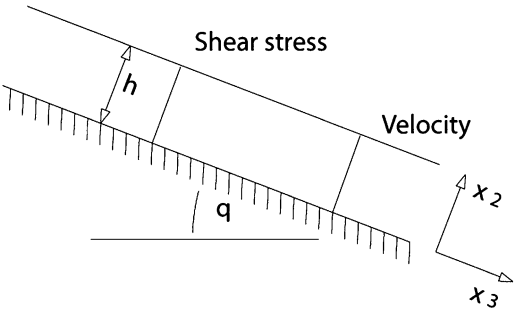
6.4.6. Determine the volume flux per unit length in the direction of flow in example 6.4.3

6.4.7. Consider a viscous fluid layer between two parallel flat plates. A reference coordinate system with x_1 in the plane of the plates and x_2 as the direction perpendicular to the plane of the plates is to be employed. Relative to this coordinate system the plates are located at $x_2 = \pm h/2$. For Couette flow the velocity distribution is $v_1 = (V/h)(x_2 + h/2)$ and for plane Poiseuille flow the velocity distribution is

$$v_1 = \frac{1}{2\mu}\left(-\frac{\partial p}{\partial x_1}\right)\left\{\left(\frac{h}{2}\right)^2 - x_2^2\right\}.$$

(a) Determine formulas for the shear stress in the fluid for Couette flow and for Poiseuille flow.
(b) Plot the shear stress in the fluid as a function of x_2 for Couette flow. In this case let the units of shear stress on the graph be multiples of mV/h.
(c) Plot the shear stress in the fluid as a function of x_2 for Poiseuille flow. In this case let the units of shear stress on the graph be multiples of $-\frac{h}{2}\frac{\partial p}{\partial x_1}$.

6.4.8 The figure below shows a slope making an angle of theta with the horizontal and a layer of viscous fluid of thickness h flowing down the sloping plane.



The velocity field is given by

$$v_2 = \frac{\rho g \sin\theta}{2\mu}(2hx_3 - x_3^2).$$

The coordinate direction x_3 is parallel to the sloping plane and points down slope and the coordinate direction x_2 is perpendicular to the sloping plane.

(a) Determine the rate-of-strain tensor for this flow.
(b) Is this flow volume increasing or decreasing?
(c) What is the stress field in this flow?

(d) What is the stress on the surface $x_3 = h$? Please describe these stresses in prose.
(e) What is the stress on the surface $x_3 = 0$? Please describe these stresses in prose.
(f) In the portion of the figure containing the words "Shear stress" there is a line drawn across the flow domain. Please sketch in the shear stress distribution across the flow field using this line as the zero shear stress reference.
(g) In the portion of the figure containing the word "Velocity" there is a line drawn across the flow domain. Please sketch in the velocity distribution across the flow field using this line as the zero velocity reference.

6.5 The Theory of Viscoelastic Materials

An algebraic overview of the theory of linear viscoelastic materials can be obtained by considering it as a system of fifteen equations in fifteen unknown functions of time. The fifteen unknown functions of time are the six components of the stress tensor $\mathbf{T}$, the six components of the strain tensor $\mathbf{E}$, and the three components of the displacement vector $\mathbf{u}$. The parameter functions of a viscoelasticity problem are the tensor of viscoelastic coefficients $\hat{G}(s)$, call the tensor of *relaxation functions*, the density $\rho(s)$ and the action-at-a-distance force $\mathbf{d}(s)$, which are assumed to be known. The system of fifteen equations consists of viscoelastic stress–strain relations,

$$\begin{aligned}\hat{\mathbf{T}}(\mathbf{x},t) &= \int_{s=-\infty}^{s=t} \hat{G}(t-s)\cdot\hat{\mathbf{D}}(\mathbf{x},s)\,\mathrm{d}s \\ &= \int_{s=-\infty}^{s=t} \hat{G}(t-s)\cdot\frac{D\hat{\mathbf{E}}}{Ds}(\mathbf{x},s)\,\mathrm{d}s,\end{aligned} \qquad (5.36\mathrm{V})\ \text{repeated}$$

the six strain–displacement relations,

$$\mathbf{E} = (1/2)((\nabla\otimes\mathbf{u})^{\mathrm{T}} + \nabla\otimes\mathbf{u}), \qquad (2.49)\ \text{repeated}$$

and the three stress equations of motion,

$$\rho\ddot{\mathbf{u}} = \nabla\cdot\mathbf{T} + \rho\mathbf{d}, \mathbf{T} = \mathbf{T}^{\mathrm{T}}, \qquad (6.18)\ \text{repeated}$$

Interchanging the roles of stress and strain the current strain is expressed as a function of the past history of stress by

$$\hat{\mathbf{E}}(\mathbf{x},t) = \int_{s=-\infty}^{s=t} \hat{\mathbf{J}}(t-s)\cdot\frac{D\hat{\mathbf{T}}}{Ds}(\mathbf{x},s)\,\mathrm{d}s, \qquad (6.41)$$

where $\hat{\mathbf{J}}(s)$is called the tensor of creep function.

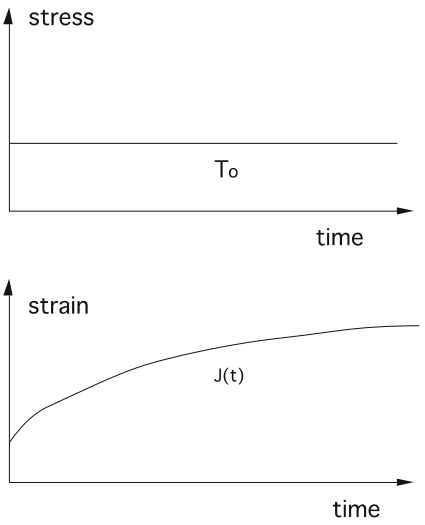


Fig. 6.5 The creep function $J(t)$. The creep function is the strain response to a step input of stress at $t = 0$

The physical significance of the creep and relaxation functions, $\hat{\mathbf{J}}(s)$ and $\hat{\mathbf{G}}(s)$, are best illustrated by one-dimensional examples; uniaxial one-dimensional examples will be used here. A uniaxial tension specimen of homogeneous material is subjected to step increase in stress $\hat{T}_1 = T_o h(t)$, where $h(t)$ is the unit step function, $h(t) = 0$ for $t < 0$ and $h(t) = 1$ for $t > 0$. The loading function $\hat{T}_1 = T_o h(t)$ is plotted in the top panel of Fig. 6.5. The creep function $\hat{J}_{11}(s)$is the uniaxial strain-vs.-time response of the uniaxially stressed specimen to the step increase in stress, $\hat{T}_1 = T_o h(t)$. The function $\hat{J}_{11}(s)$is illustrated in the lower panel of Fig. 6.5. The strain response is given by $\hat{E}_1 = T_o \hat{J}_{11}(s)$. The creep function defined and measured in this way may be used to predict the creep response to a more complicated stress history. Suppose, for example, the stress history is a long series of step jumps rather than just one step jump (Fig. 6.5). The creep strain response to this new stress history may be built up by repeated application of the basic result $\hat{E}_1 = T_o \hat{J}_{11}(s)$ to each step, and subsequent summation of strains associated with each step,

$$\hat{E}_1(t) = \hat{J}_{11}(t)\Delta\hat{T}_1(0) + \hat{J}_{11}(t - t_1)\Delta\hat{T}_1(t_1) + \cdots + \hat{J}_{11}(t - t_n)\Delta\hat{T}_1(t_n) + \cdots . \tag{6.42}$$

The multiple step plot of Fig. 6.6 is familiar from the introductory presentations to the process of integration in which a horizontal axis is divided into segments and the curve is approximated by different level steps drawn horizontally for each segment so that the curve is approximated by the series of various sized steps. It follows that any curve representing a stress history could be approximated arbitrarily closely by a set of various sized steps like those illustrated in Fig. 6.6 and represented analytically by an equation of the type (6.42). In preparation for a passage to the limit of the type used in the integral calculus, (6.42) is rewritten as

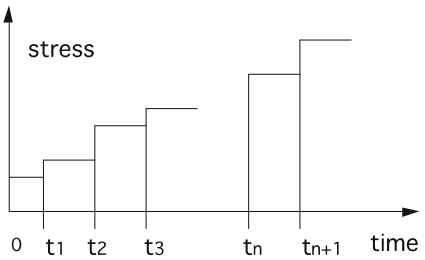


Fig. 6.6 A plot of stress increments applied at specific time increments

$$\hat{E}_1(t) = \sum_{i=1}^{n} \hat{J}_{11}(t - t_i)\Delta\hat{T}_1(t_i). \tag{6.43}$$

In the passage to the limit the series of times t_i, $i = 1, \ldots, n$ will be replaced by the continuous parameter s, and the step in the stress by $\Delta\hat{T}_1(t_i)$ by $\frac{\mathrm{d}\hat{T}_1}{\mathrm{d}s}\mathrm{d}s$, thus

$$\hat{E}_1(t) = \int_0^t \hat{J}_{11}(t - t_i)\frac{\mathrm{d}\hat{T}_1}{\mathrm{d}s}\mathrm{d}s. \tag{6.44}$$

Finally, going back to the beginning of this development there was nothing special about applying the loading at $t = 0$; it could have been started at any time, thus we set it back to the beginning of time,

$$\hat{E}_1(t) = \int_{s=-\infty}^{s=t} \hat{J}_{11}(t - s)\frac{\mathrm{d}\hat{T}_1}{\mathrm{d}s}\mathrm{d}s. \tag{6.45}$$

It is now clear that (6.45) is a special case of (6.41). The physical significance of the creep function $\hat{J}_{11}(s)$ is the following: it is the uniaxial normal strain $\hat{E}_1(t)$ vs. time response of a specimen subjected to a step increase in the normal stress in the same direction, $\hat{T}_1 = T_o h(t)$.

The physical significance of the relaxation function $\hat{G}_{11}(s)$ can be developed in a similar manner by reversing the roles of stress and strain used in the case of the creep function. A uniaxial tension specimen of homogeneous material is subjected to a step increase in strain, $\hat{E}_1 = E_o h(t)$. The strain loading function $\hat{E}_1 = E_o h(t)$ is plotted in the top panel of Fig. 6.7. The relaxation function $\hat{G}_{11}(s)$is the uniaxial stress vs. time response of the uniaxially strained specimen to the step increase in strain, $\hat{E}_1 = E_o h(t)$. The function $\hat{G}_{11}(s)$ is illustrated in the lower panel of Fig. 6.7. The stress response is given by $\hat{T}_1 = \hat{G}_{11}(t)\hat{E}_o$. The relaxation function defined and measured in this way may be used to predict the relaxation response to a more complicated strain history as described above for the creep function. Following completely analogous steps one finds that

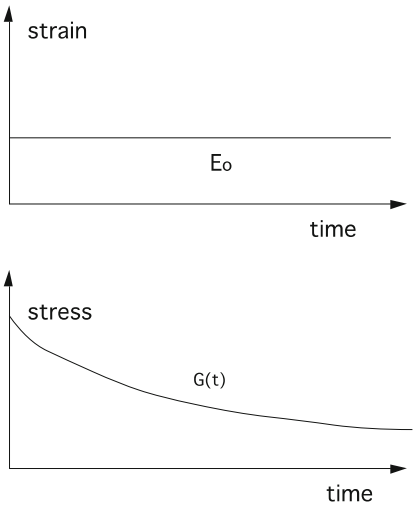


Fig. 6.7 The relaxation function $G(t)$. The relaxation function is the stress response to a step input of strain at $t = 0$

$$\hat{T}_1(t) = \int_{s=-\infty}^{s=t} \hat{G}_{11}(t-s)\frac{\mathrm{d}\hat{E}_1}{\mathrm{d}s}\mathrm{d}s. \tag{6.46}$$

The result (6.46) is a special case of (5.36V). The physical significance of the relaxation function $\hat{G}_{11}(s)$ is the following: it is the uniaxial normal stress $\hat{T}_1(t)$ vs. time response of a specimen subjected to a step increase in the normal strain in the same direction, $\hat{E}_1 = E_o h(t)$.

The creep and relaxation functions, $\hat{\mathbf{J}}(s)$ and $\hat{\mathbf{G}}(s)$, are hypothesized to have a property called *fading memory*. The fading memory hypothesis for $\hat{\mathbf{J}}(s)(\hat{\mathbf{G}}(s))$ is that the strain (stress) depends more strongly upon the recent history than it does upon the distant history of the value of stress (strain). Mathematically this is the requirement that the functions $\hat{\mathbf{J}}(s)$ and $\hat{\mathbf{G}}(s)$ are continuously decreasing functions of the backward running time parameter s. This then decreases the influence of the more distant events. For each component of the viscoelastic strain–stress relations to possess this fading memory behavior, it is sufficient that the magnitude of the slope of each component of the creep function tensor be a continuously decreasing function of time, thus

$$\left|\frac{D\hat{J}_{ij}(s)}{Ds}\right|_{s=s_1} < \left|\frac{D\hat{J}_{ij}(s)}{Ds}\right|_{s=s_2} \text{ for all } i,j = 1,2,\ldots 6, \text{ and for } s_1 > s_2 > 0. \tag{6.47}$$

For each component of the viscoelastic stress–strain relations a similar fading memory hypothesis holds, thus

$$\left|\frac{D\hat{G}_{ij}(s)}{Ds}\right|_{s=s_1} < \left|\frac{D\hat{G}_{ij}(s)}{Ds}\right|_{s=s_2} \quad \text{for all } i,j = 1,2,\ldots 6, \text{ and for } s_1 > s_2 > 0. \tag{6.48}$$

The isotropic form of the viscoelastic stress–strain relations (5.36V) is obtained by using the representation for the isotropic form of $\hat{G}(s)$ obtained from Table 5.4, thus

$$\hat{G}(s) = \begin{bmatrix} \hat{G}_{11}(s) & \hat{G}_{12}(s) & \hat{G}_{12}(s) & 0 & 0 & 0 \\ \hat{G}_{12}(s) & \hat{G}_{11}(s) & \hat{G}_{12}(s) & 0 & 0 & 0 \\ \hat{G}_{12}(s) & \hat{G}_{12}(s) & \hat{G}_{11}(s) & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{G}_{11}(s) - \hat{G}_{12}(s) & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{G}_{11}(s) - \hat{G}_{12}(s) & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{G}_{11}(s) - \hat{G}_{12}(s) \end{bmatrix}$$

This six-dimensional representation is converted to the three-dimensional representation by employing the relations in Table 6.1 and introducing the following new notation for the two distinct elements of this 6-by-6 matrix, thus

$$\hat{G}_{11}(s) = \frac{k_{\text{tr}}(s) + 2G_{\text{dev}}(s)}{3}, \hat{G}_{12}(s) = \frac{k_{\text{tr}}(s) - G_{\text{dev}}(s)}{3}. \tag{6.49}$$

The isotropic form of the viscoelastic stress–strain relations (5.36V) may then be rewritten in three dimensions as

$$\text{tr}\mathbf{T}(\mathbf{x},t) = \int_{s=-\infty}^{s=t} k_{\text{tr}}(t-s)\frac{D}{Ds}\{\text{tr}\mathbf{E}(\mathbf{x},s)\}\,\text{d}s \tag{6.50}$$

and

$$\text{dev}\mathbf{T}(\mathbf{x},t) = \int_{s=-\infty}^{s=t} G_{\text{dev}}(t-s)\frac{D}{Ds}\{\text{dev}\mathbf{E}(\mathbf{x},s)\}\,\text{d}s, \tag{6.51}$$

where $k_{\text{tr}}(s)$ and $G_{\text{dev}}(s)$ represent independent relaxation functions. In a similar set of arguments it may be shown that the isotropic form of the viscoelastic strain–stress relations (6.41) may be expressed in terms of two isotropic creep functions, $j_{\text{tr}}(s)$ and $J_{\text{dev}}(s)$, thus

$$\text{tr}\mathbf{E}(\mathbf{x},t) = \int_{s=-\infty}^{s=t} j_{\text{tr}}(t-s)\frac{D}{Ds}\{\text{tr}\mathbf{T}(\mathbf{x},s)\}\,\text{d}s \tag{6.52}$$

and

$$\text{dev}\mathbf{E}(\mathbf{x},t) = \int_{s=-\infty}^{s=t} J_{\text{dev}}(t-s)\frac{D}{Ds}\{\text{dev}\mathbf{T}(\mathbf{x},s)\}\,\text{d}s. \tag{6.53}$$

Viscoelastic materials have properties characteristic of both fluids and solids and it is sometimes important to distinguish between viscoelastic fluids and viscoelastic

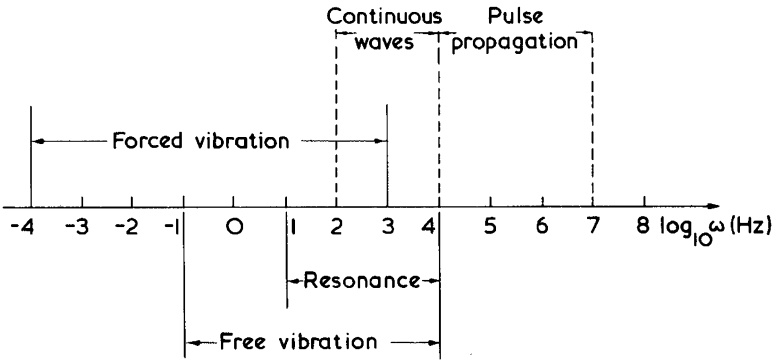


Fig. 6.8 Frequency ranges that are most effective for the various types of dynamical mechanical tests. From Lockett (1972), Fig. 4 on page 50

solids. One aspect of fluid behavior is that a fluid is always isotropic while a solid may be anisotropic. A second characteristic of a fluid, a characteristic that aids in distinguishing isotropic viscoelastic solids from isotropic viscoelastic fluids, is the asymptotic value of $G_{\text{dev}}(t)$ as t tends to infinity. A necessary and sufficient condition that an isotropic viscoelastic material be a solid is that $G_{\text{dev}}(t)$ tends to a nonzero constant as t tends to infinity. A necessary, but not sufficient condition, that an isotropic viscoelastic material be a fluid is that $G_{\text{dev}}(t)$ tends to zero as t tends to infinity. It is convenient to define

$$G^{\text{o}}_{\text{dev}} \equiv \lim_{t\to\infty} G_{\text{dev}}(t), \tag{6.54}$$

then

$$G_{\text{dev}}(t) \equiv G^{\text{o}}_{\text{dev}} + G^{\infty}_{\text{dev}}(t), \tag{6.55}$$

where

$$\lim_{t\to\infty} G^{\text{o}}_{\text{dev}}(t) = 0. \tag{6.56}$$

Thus if $G^{\text{o}}_{\text{dev}} \neq 0$, the material is a viscoelastic solid and if $G^{\text{o}}_{\text{dev}} = 0$, it may be a fluid.

Viscoelastic materials differ from the other materials considered in this chapter because their material properties are determined by material functions rather than material constants. The material functions are functions of time, or of time transformed, that is to say frequency. Viscoelastic materials are probed with various dynamical test systems in order to evaluate the forms of the material functions, their peaks and valleys. Different dynamical test systems are effective for determining the material functions for different frequency ranges, Fig. 6.8. Both free and resonant

vibrations and pulse and harmonic wave propagation techniques have the ranges of effectiveness indicated in Fig. 6.8. One of the easiest probing tests for a viscoelastic material is to subject the material to forced steady state oscillations. This testing method is effective over a wide range of frequencies, Fig. 6.8.

In the special case of forced steady state oscillations, special forms of the stress–strain relations emerge. As an example we consider the case of the deviatoric part of the isotropic stress–strain relations (6.51). It is assumed that the material is subjected to a forced deviatoric strain specified as a harmonic function of time,

$$\mathrm{dev}\mathbf{E}(t) = \{\mathrm{dev}\mathbf{E}_{\mathrm{o}}\}\mathrm{e}^{i\omega t} = \{\mathrm{dev}\mathbf{E}_{\mathrm{o}}\}\{\cos\omega t + i\sin\omega t\}. \tag{6.57}$$

Upon substitution of the strain (6.57) into the viscoelastic stress–strain relation (6.51), along with the decomposition of $G_{\mathrm{dev}}(t)$ given by (6.55), it follows that

$$\mathrm{dev}\mathbf{T}(\mathbf{x},t) = G^{\mathrm{o}}_{\mathrm{dev}}\,\mathrm{dev}\mathbf{E}_{\mathrm{o}}\mathrm{e}^{its} + i\omega\mathrm{dev}\mathbf{E}_{\mathrm{o}}\int_{s=-\infty}^{s=t} G^{\infty}_{\mathrm{dev}}(t-s)\mathrm{e}^{i\omega s}\mathrm{d}s. \tag{6.58}$$

Now, making a change of variable, $t-s=\eta$,

$$\mathrm{dev}\mathbf{T}(\mathbf{x},t) = \left\{G^{\mathrm{o}}_{\mathrm{dev}} + \omega\int_0^{\infty}\sin\omega\eta\, G^{\infty}_{\mathrm{dev}}(\eta)\,\mathrm{d}\eta + i\omega\int_0^{\infty}\cos\omega\eta\, G^{\infty}_{\mathrm{dev}}(\eta)\,\mathrm{d}\eta\right\}\mathrm{dev}\mathbf{E}_{\mathrm{o}}\mathrm{e}^{i\omega t}, \tag{6.59}$$

or

$$\mathrm{dev}\mathbf{T}(\mathbf{x},t) = G^{*}_{\mathrm{dev}}(i\omega)\mathrm{dev}\mathbf{E}_{\mathrm{o}}\mathrm{e}^{i\omega t}, \tag{6.60}$$

where

$$G^{*}_{\mathrm{dev}}(i\omega) = G^{'}_{\mathrm{dev}}(\omega) + iG^{''}{}_{\mathrm{dev}}(\omega), \tag{6.61}$$

and

$$G^{'}_{\mathrm{dev}}(\omega) = G^{\mathrm{o}}_{\mathrm{dev}} + \omega\int_0^{\infty}\sin\omega\eta\, G^{\infty}_{\mathrm{dev}}(\eta)\,\mathrm{d}\eta, G^{''}_{\mathrm{dev}}(\omega) = \omega\int_0^{\infty}\cos\omega\eta\, G^{\infty}_{\mathrm{dev}}(\eta)\,\mathrm{d}\eta. \tag{6.62}$$

The functions $G^{'}_{\mathrm{dev}}(\omega)$ and $G^{''}_{\mathrm{dev}}(\omega)$ are called the storage and loss moduli, respectively. The formulas (6.62) show that the real and complex parts of the complex modulus $G^{*}_{\mathrm{dev}}(i\omega)$, the only material function in the specialized steady-state oscillatory viscoelastic stress–strain relation (6.60), are determined by the relaxation function, $G_{\mathrm{dev}}(t)$. The stress–strain relations (6.60) may also be written as

$$\mathrm{dev}\mathbf{T}(\mathbf{x},t) = |G^{*}_{\mathrm{dev}}(i\omega)|\mathrm{dev}\mathbf{E}_{\mathrm{o}}\mathrm{e}^{i(\omega t+\varphi_{\mathrm{dev}})}, \tag{6.63}$$

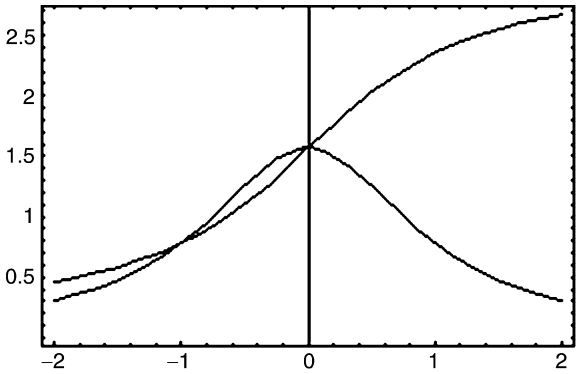


Fig. 6.9 The complex modulus $G^*(i\omega)$. The vertical scale represents both the real and imaginary parts of $G^*(i\omega)$ which have the dimensions of one upon stress. The horizontal scale is the log of the frequency, log ω. The monotonically increasing curve represents $G'(\omega)$, the real part of $G^*(i\omega)$, and the curve with a peak represents $G''(\omega)$, the imaginary part of $G^*(i\omega)$.

where the phase angle $\varphi_{\text{dev}}(\omega)$ is given by

$$\varphi_{\text{dev}}(\omega) = \tan^{-1}\left\{\frac{G''_{\text{dev}}(\omega)}{G'_{\text{dev}}(\omega)}\right\}. \tag{6.64}$$

The quantity $\tan\,\varphi_{\text{dev}}(\omega)$ is called the loss tangent. The steady-state harmonic strain lags behind the stress by the phase angle $\varphi_{\text{dev}}(\omega)$. Typical plots of the storage and loss moduli, $G'_{\text{dev}}(\omega)$ and $G''_{\text{dev}}(\omega)$, respectively as a function of ω are shown in Fig. 6.9. However these curves for real material seldom look exactly like these examples.

Example 6.5.1
An isotropic viscoelastic material is subjected to a step loading in shear strain E_{12}. The magnitude of the step loading is E_{o}. The unit step function $h(t)$ is used to represent the step loading, $E_{12} = E_{\text{o}}h(t)$. Recall that $h(t)$ is a function that is defined as 0 for $t < 0$ and as 1 for $t > 0$. The derivative of the unit step function is the delta function $\delta(t)$, $(d/\mathrm{d}t)(h(t)) = \delta(t)$, where the delta function has the property that

$$f(t*) = \int_{-\infty}^{\infty} \delta(t - t^*) f(t)\,\mathrm{d}t.$$

Determine the stress response to the strain loading $E_{12} = E_{\text{o}}h(t)$.

Solution: Substitution of strain loading $E_{12} = E_{\text{o}}h(t)$ into the appropriate stress–strain relation (6.51) yields the following simple formula,

$$T_{12} = \int_{s=-\infty}^{s=t} G_{\text{dev}}(t-s)E_{\text{o}}\delta(s)\,\mathrm{d}s = G_{\text{dev}}(t)E_{\text{o}}.$$

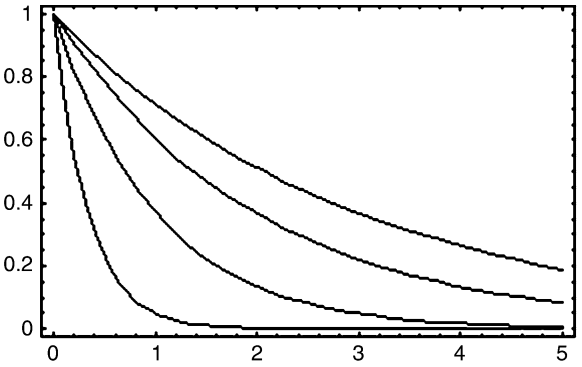


Fig. 6.10 Four typical relaxation functions $G(t)$. The vertical scale is the relaxation function $G(t)$ that has the dimensions of stress. The horizontal scale is time

It follows then that the relaxation function is just the value of the resulting stress divided by E_o,

$$G_{dev}(t) = (1/E_o)T_{12}.$$

This result can then be used as the basis of an experiment to determine the relaxation function $G_{dev}(t)$. Typical relaxation functions are of the form of decaying functions of time like that shown in Fig. 6.10. The decaying with time is consistent with the hypothesis of fading memory.

Example 6.5.2
The Maxwell model is a lumped viscoelastic model (see section 1.8) that is a combination of a spring and a dashpot in series (Fig. 1.9a). When a force applied to a Maxwell model is changed from zero to a finite value at an instant of time and held constant thereafter, there is an instantaneous initial elastic extension and then there is a continued deformation forever as the damper in the dashpot is drawn through the dashpot cylinder. Thus a Maxwell model exhibits the characteristics of a fluid with an initial elastic response. In this model, in general, the deflection represents the strain and the force represents the stress. In this particular illustration the deflection represents the shearing strain and the force represents the shearing stress. The differential equation for a Maxwell model is formed by adding together the rate of strain of the spring, obtained by differentiating Hooke's law in isotropic shear, $E_{12} = (1/G_o)T_{12}$, with respect to time, $(d/dt)E_{12} = (1/G_o)(d/dt)T_{12}$, and the dashpot $(d/dt)E_{12} = (1/\tau_r G_o)T_{12}$ where $\tau_r G_o$ represents the viscosity of the dashpot, thus

$$\frac{dE_{12}}{dt} = \frac{1}{\tau_r G_o} T_{12} + \frac{1}{G_o}\frac{dT_{12}}{dt}.$$

The problem is to determine the response of the Maxwell model to a step loading in shearing strain E_{12}. The magnitude of the step loading is E_o. This example is formally similar to the previous example, the only change being in the model. The solution of the problem requires the solution of a first-order ordinary differential

equation subject to the case when the strain history is given by $E_{12} = E_o h(t)$ and no other loading is applied to the model. Determine the stress response to this strain history loading.

Solution: The general solution to the first-order ordinary differential equation

$$\frac{dy}{dt} + f(t)y - g(t)$$

is (see Sect. A.17)

$$y(t) = e^{-F}\left(c + \int_{t_o}^{t} g(s)e^{F}ds\right), \text{where } F(t) = \int_{t_o}^{t} f(s)\,ds,$$

and where c is a constant of integration. In the differential equation of interest $y(t) = T_{12}(t)$, $F(t) = (t/\tau_r)$, $g(s) = (d/dt)E_{12} = (d/dt)(E_o h(s))$ and, due to the initial condition $c = 0$, thus

$$T_{12}(y) = e^{-t/\tau_r}\left(\int_0^t G_o E_o \frac{dh(s)}{ds} e^{s/\tau_r} ds\right) = e^{-t/\tau_r}\left(\int_0^t G_o E_o \delta(s) e^{s/\tau_r} ds\right)$$
$$= G_o E_o h(t) e^{-t/\tau_r}.$$

The special form of the relaxation function for the Maxwell model may then be identified as

$$G_{dev}(t) = (1/E_o)T_{12} = G_o h(t) e^{-t/\tau_r}.$$

Problems

6.5.1 Derive (6.50) and (6.51) from (5.36V).

6.5.2 Derive (6.52) and (6.53) from (6.41).

6.5.3 Determine the strain response to this stress loading for an isotropic viscoelastic material subjected to a step loading in shear stress T_{12}. The magnitude of the step loading is T_o. This problem is formally similar to the Example 6.5.1, the only change being that the step loading is now in shear stress rather than shear strain. The step loading is represented by $T_{12} = T_o h(t)$. Show that the creep function is just the value of the resulting strain divided by T_o. Note that typical creep functions are of the form of increasing functions of time like that shown in Fig. 6.11.

6.5.4 The Voigt model is a lumped viscoelastic model that is a combination of a spring and a dashpot in parallel (Fig. 1.10b). When a force applied to a Voigt model is changed from zero to a finite value at an instant of time and then held constant thereafter, extension occurs only after the dashpot begins to move. Thus the Voigt model is initially rigid; then it begins to creep asymptotically under the constant applied load to a rest value. In parallel models like the

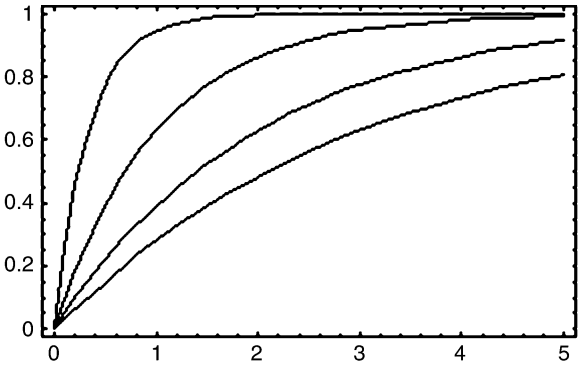


Fig. 6.11 Four typical creep functions $J(t)$. The vertical scale is the creep function $J(t)$ that has the dimensions of one upon stress. The horizontal scale is time

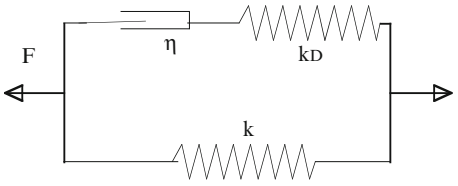


Fig. 6.12 A three-parameter viscoelastic model

Voigt model, the transverse or horizontal bars or connecting elements are required not to rotate, so the two parallel elements are constrained to always have the same extension in both elements at any instant of time. In these spring and dashpot models, in general, the deflection represents the strain and the force represents the stress. In this particular illustration the deflection represents the shearing strain and the force the shearing stress. The differential equation for a Voigt model is formed by adding together the stress in the two parallel branches, $T_{12} = (1/J_o)E_{12}$ for the spring and, $T_{12} = (\tau_r/J_o)(d/dt)(E_{12})$ for the dashpot, thus

$$\frac{dE_{12}}{dt} = \frac{1}{\tau_r G_o} T_{12} + \frac{1}{G_o} \frac{dT_{12}}{dt}.$$

The problem is to determine the response of the Voigt model to a step loading in shearing stress T_{12}. The magnitude of the step loading is T_o. This problem is formally similar to the Example 6.5.2, the only change being in the model. The solution of the problem requires the solution of a first order ordinary differential equation subject to the case when the stress history is given by $T_{12} = T_o h(t)$ and no other loading is applied to the model. Determine the stress response to this strain history loading and show that, for this model, the creep function is given by $J(t) = (1/T_o)E_{12} = J_o h(t)\left(1 - e^{-t/\tau_c}\right)$. Typical creep functions are shown in Fig. 6.11.

6.5.5 A three-parameter viscoelastic model is shown in the Fig. 6.12. The model consists of two branches, the lower branch with a spring and the upper branch

with a spring and a dashpot. These two branches are connected by two transverse elements that are indicated by vertical lines. The rules for formulating the overall force-deformation ($F-\delta$) relation for these spring and dashpot models are: (1) the transverse elements that connect the two branches remain parallel as the system is deformed; (2) the springs follow the rule that $F = k\delta$, and the dashpots follow the rule that $F = \eta(\mathrm{d}\delta/\mathrm{d}t)$. Let F_{L} and F_{U} denote the forces in the upper and lower branches, respectively, δ_{D} the deflection of the dashpot and δ_{S} the deflection of the spring in series with the dashpot. Construct the overall force-deformation ($F-\delta$) relation for this spring and dashpot model. List each step of your argument and explain its rationale.

6.6 Relevant Literature

This chapter presents brief descriptions of four theories of material behavior. There are many volumes written on each of these theories and there are entire periodicals devoted to publishing recent results in each of these theories. The purpose of this section is to mention some of the general literature associated with each of these four theories, particularly that was drawn upon in the writing of this chapter.

In heat conduction there are the texts of Carslaw and Jaeger (1959), Ozisik (1980) and Jiji (2000). The thermoelasticity book of Boley and Weiner (1960) also contains much related material. The journals include the ASME *Journal of Heat Transfer* and the *International Journal of Heat and Mass Transfer*.

Flow through porous media is addressed in books by Bear (1972), Carman (1956), Scheidegger (1960) and others. Several of the books on poroelasticity mentioned in Chap. 8 contain much related material. The related journals include *Transport in Porous Media, Water Resources Research*, and the *Journal of Hydrology*.

The theory of elastic solids is described in the classical text of Love (1927), in Timoshenko and Goodier (1951), Sokolnikoff (1956), Saada (1974) and in Gurtin (1972). Anisotropic elasticity is the subject of books by Hearmon (1961), Lekhnitskii (1963), and Fedorov (1968). Current publications on elasticity theory may be found in the *Journal of Elasticity* and in many related journals such as the *International Journal of Solids and Structures*, the *Journal of the Mechanics and Physics of Solids*, and the ASME *Journal of Applied Mechanics*.

The theories associated with fluid behavior are described in the classical books of Lamb (1932), Prandtl and Tietjens (1934a, b), and Schlichting (1960); Langois (1964) is an introduction to slow viscous flow. The current widely used text is Batchleor (2000). There are many journals specializing in fluid mechanics; the most broad and most successful is the *Journal of Fluid Mechanics*.

The books that deal with viscoelasticity theory include Christensen (1971), Lakes (1999), Lockett (1972), Pipkin (1972), and the contemporary work of Wineman and Rajagopal (2000). Journals in this area include *Rheologica Acta*, *Journal of Rheology* and *Biorheology*.

References

Theory of Rigid Porous Media

Boley BA, Weiner JH. (1960) Theory of Thermal Stresses, Wiley, New York
Carslaw HS, Jaeger JC (1959) Conduction of Heat in Solids, 2nd ed., Oxford Clarendon Press
Jiji LM (2010) Heat Convection, 2nd ed., Springer, New York
Özişik MN (1993), *Heat Conduction*, 2nd Ed, Wiley, New York
Bear J (1972) Dynamics of fluids in porous media. Elsevier, New York
Carman PC (1956) Flow of gases through porous media. Butterworths, London
Scheidegger AE (1960) The physics of flow through porous media, 2nd edn. University of Toronto Press, Toronto

Classical Elasticity Theory

Gurtin ME (1972) The linear theory of elasticity. In: Flugge S (ed) Handbuch der Physik, V1a/2ed. Springer, Berlin
Love AEH (1927) Elasticity. Dover, New York
Saada AS (1974) Elasticity theory and applications. Pergamon, New York
Sokolnikoff IS (1956) Mathematical theory of elasticity. McGraw Hill, New York
Timoshenko SP, Goodier JN (1951) Theory of elasticity. McGraw Hill, New York

Anisotropic Elasticity Theory

Fedorov FI (1968) Theory of elastic waves in crystals. Plenum Press, New York
Hearmon RFS (1961) An introduction to applied anisotropic elasticity. Oxford University Press, Oxford
Lekhnitskii SG (1963) Theory of elasticity of an anisotropic elastic body. Holden Day, San Francisco
Ting TCT (1996) Anisotropic elasticity—theory and applications. Oxford University Press, New York

Classical Fluid Theory

Batchleor GK (2000) An introduction to fluid mechanics. Cambridge, Cambridge
Lamb H (1932) Hydrodynamics. Dover, New York
Langois WE (1964) Slow viscous flow. Macmillan, New York

Prandtl L, Tietjens OG (1934a) Fundamentals of hydro- and aeromechanics. McGraw Hill, New York
Prandtl L, Tietjens OG (1934b) Applied hydro- and aeromechanics. McGraw Hill, New York
Schlichting H (1960) Boundary layer theory. McGraw Hill, New York

Viscoelasticity Theory

Christensen RM (1971) Theory of viscoelasticity. Academic Press, New York
Lakes RS (1999) Viscoelastic solids. CRC Press, Boca Raton
Lockett FJ (1972) Nonlinear viscoelastic solids. Academic Press, London
Pipkin AC (1972) Lectures on viscoelasticity theory. Springer, New York
Wineman AS, Rajagopal KR (2000) Mechanical response of polymers. Cambridge University Press, Cambridge

Chapter 7
Modeling Material Microstructure

Continuum mechanics deals with idealized materials consisting of material points and material neighborhoods. It assumes that the material distribution, the stresses, and the strains within an infinitesimal material neighborhood of a typical particle (or a material element) can be regarded as essentially uniform. On the microscale, however, the infinitesimal material neighborhood, in general, is not uniform, consisting of various constituents with differing properties and shapes, i.e., an infinitesimal material element has its own complex and, in general, evolving microstructure. Hence, the stress and strain fields within the material element likewise are not uniform at the microscale level. One of the main objectives of micromechanics is to express in a systematic and rigorous manner the continuum quantities associated with an infinitesimal material neighborhood in terms of the parameters that characterize the microstructure and properties of the microconstituents of the material neighborhood. From Nemat Nasser and Hori (1999), page 11

7.1 Introduction

The heterogeneity of materials is obvious at all levels of their hierarchical structure, including the smallest level. At the smallest level there are spaces between the atoms or molecules that constitute holes in the material making the material discontinuous. One purpose of this chapter is to reconcile this fact with the continuity assumption of the continuum models described in the preceding chapters so that one will understand how these continuum models are applied to material objects. A second purpose is to relate the effective material parameters and material symmetry used in continuum models to the material microstructure.

The key concept in modeling material microstructure, as well as for modeling material symmetry, for inclusion in a continuum model is the representative volume element (RVE). The discussion of the RVE contained in Chap. 4, particularly Sect. 4.2, is further developed in Sect. 7.2. A material composed of two or more distinct constituent materials is called a composite material. Most natural materials are

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_7,

composite materials. Many think that the first man-made composite material was the reinforced brick constructed by using straw to reinforce the clay of the brick. Dried clay is satisfactory in resisting compression, but not very good in tension. The straw endows the brick with the ability to sustain greater tensile forces. Most structural soft tissues in animals can carry tensile forces adequately but do not do well with compressive forces. In particular, due to their great flexibility they may deform greatly under compressive forces. The mineralization of the collagenous tissues provides those tissues with the ability to resist compressive forces; thus bone and teeth are composites of an organic phase, primarily collagen, and an inorganic or mineral phase. Effective material parameters for composite materials, defined in Sect. 7.3, are generally determined by expressions that depend upon the phase or constituent-specific material parameters and their geometries. Examples of effective elastic constants and effective permeabilities are developed in Sects. 7.4 and 7.5, respectively. Restrictions on the RVE for the case of a gradient in its material properties are considered in Sect. 7.6. The continuum modeling of material microstructures with vectors and tensors is described in Sect. 7.7, with a particular emphasis on the fabric tensor, a measure of local microstructure in a material with more than a single constituent. The stress–strain–fabric relation is developed in Sect. 7.8. Some of the relevant literature is described in Sect. 7.9.

7.2 The Representative Volume Element

Recall from Sect. 4.2 that, for this presentation, the RVE is taken to be a cube of side length L_{RVE}; it could be any shape, but it is necessary that it have a characteristic length scale (Fig. 4.1). The RVE for the representation of a domain of a porous medium by a continuum point was illustrated in Fig. 4.2. We begin here by picking up the question of how large must the length scale L_{RVE} be to obtain a reasonable continuum model. The L_{RVE} should be much larger than the largest characteristic microstructural dimension L_{M} of the material being modeled and smaller than the characteristic dimension of the problem to be addressed L_{P}, thus $L_{\mathrm{P}} \gg L_{\mathrm{RVE}} \gg L_{\mathrm{M}}$.

The question of the size of L_{RVE} can also be posed in the following way: How large a hole is no hole? The value of L_{RVE} selected determines what the modeler has selected as too small a hole, or too small an inhomogeneity or microstructure, to influence the result the modeler is seeking. An interesting aspect of the RVE concept is that it provides a resolution of a paradox concerning stress concentrations around circular holes in elastic materials. The stress concentration factor associated with the hole in a circular elastic plate in a uniaxial field of otherwise uniform stress is three times the uniform stress (Fig. 7.1). This means that the stress at certain points in the material on the edge of the hole is three times the stress five or six hole diameters away from the hole. The hole has a concentrating effect of magnitude 3. The paradox is that the stress concentration factor of 3 is independent of the size of the hole. Thus, no matter how small the hole, there is a stress concentration factor of 3 associated with the hole in a field of uniaxial tension or compression. One way to

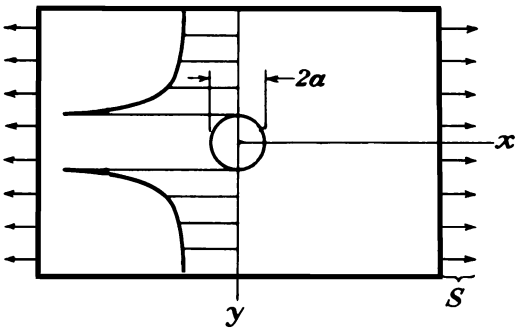


Fig. 7.1 An illustration of the stress concentration (by a factor of 3) associated with the hole in a circular plate in a uniaxial field of otherwise uniform stress, S. From Cowin (2002)

resolve the paradox described above is to observe that the modeler has decided how big a hole is no hole by choosing to recognize a hole of a certain size and selecting a value of L_{RVE} to be much less. Let the size of the largest hole in the plate to be d_{L} and assume that the plate has a dimension of 100 d_{L}, thus $d_{\mathrm{L}} \geq L_{\mathrm{RVE}} \geq L_{\mathrm{M}}$. Let d_{S} denote the dimension of the largest of the other holes in the plate. Thus any hole whose dimension is less than $L_{\mathrm{M}} = d_{\mathrm{S}}$ will not appear in the model although it is in the real object. The interpretation of the solution to the problem illustrated by Fig. 7.1 is that there is only one hole in the model of radius d_{L}, no holes of a size less than d_{L} and greater than $L_{\mathrm{M}} = d_{\mathrm{S}}$, and all holes in the real object of a size less than $L_{\mathrm{M}} = d_{\mathrm{S}}$ have been "homogenized" or averaged over.

The macro or continuum properties that are employed in continuum models are micro material properties that have been averaged over an RVE. Let ρ denote the microdensity field and $\mathbf{T}$ the micro stress field; then the average or macrodensity field $\langle\rho\rangle$ and the macro stress tensor $\langle\mathbf{T}\rangle$ field are obtained by volume averaging over the microscale. The averaging integral operator $\langle f\rangle$ of the micro field f is given by

$$\langle f\rangle = \frac{1}{V_{\mathrm{RVE}}} \oint_{V} f dv. \tag{7.1}$$

Equation (7.1) represents the *homogenization* of the local or micro material parameter fields. That is to say, in the volume V_{RVE}, the average field $\langle f\rangle$ replaces the inhomogeneous field f in the RVE. The length scale over which the homogenization is accomplished is L_{RVE} or the cube root of the volume V_{RVE}, which is intended to be the largest dimension of the unit cell over which the integration (7.1) is accomplished. A plot of the values of the macro density $\langle\rho\rangle$ as a function of the size of the RVE is sketched in Fig. 7.2. Note that as the size of the volume V_{RVE} or the L_{RVE} is decreased, the value of the density $\langle\rho\rangle$ begins to oscillate because the small volume of dense solid material in the volume V_{RVE} is greatly influenced by the occurrence of small voids. On the other hand, as the size of the volume V_{RVE} or the L_{RVE} is increased, the value of the density tends to a constant, stable value.

As another illustration of these ideas consider a cross-section of trabecular bone shown in Fig. 4.3. As was noted in Sect. 4.2, the white regions are the bone

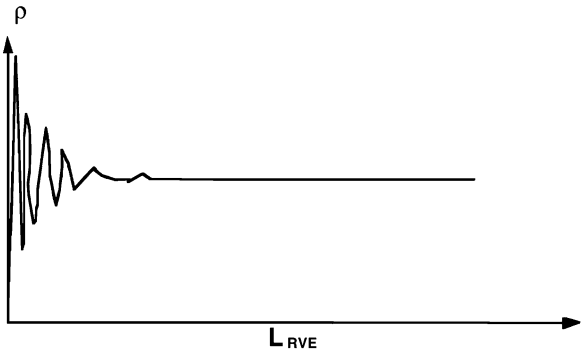


Fig. 7.2 A plot of the values of the macro density as a function of the size L_{RVE} of the RVE. Note that as the size of the volume $V_{\mathrm{RVE}} = (L_{\mathrm{RVE}})^3$ is decreased, the value of the density begins to oscillate because the amount of dense solid material in the volume V_{RVE} is greatly influenced by the occurrence of small voids. On the other hand, as the size of the volume $V_{\mathrm{RVE}} = (L_{\mathrm{RVE}})^3$ is increased, the value of the density tends to a constant, stable value. The concept for this plot is taken from Prandtl and Tietjens (1934). Prandtl used this type of diagram to illustrate density in the transition from rarefied gases in the upper atmosphere to denser air

trabeculae and the darker regions are the pore spaces that are in vivo filled with marrow in the bone of young animals. First consider the small rectangular white region in the lower left quadrant as the first RVE for homogenization. This small rectangular white region is entirely within the trabecular bone domain and thus the global or macro density $\langle\rho\rangle$ and stress tensor $\langle\mathbf{T}\rangle$ obtained by volume averaging over the microscale density ρ and stress tensor $\mathbf{T}$ will be those for trabecular bone. On the other hand, if the small RVE in the darker marrow region is entirely within the whole domain, the global or macro density $\langle\rho\rangle$ and stress tensor $\langle\mathbf{T}\rangle$ obtained by volume averaging will be those associated with the marrow. If the RVE or homogenization domain is taken to be one of the larger rectangles in Fig. 4.3, the global or macro density $\langle\rho\rangle$ and stress tensor $\langle\mathbf{T}\rangle$ obtained by volume averaging over the microscale will be different from the microscale density ρ and stress tensor $\mathbf{T}$ for both the bone and the marrow, and their values will lie in between these two limits and be proportional to the ratio of the volume of marrow voids to the volume of bone in each rectangle.

Problems

7.2.1 Find the average density $\langle\rho\rangle$ of a cube of volume b^3 if one-half the cube has a density of ρ_{o} and the other half of the cube has a density $\omega\rho_{\mathrm{o}}$.

7.2.2 Find the average density $\langle\rho\rangle$ of a sphere of radius whose density ρ is given by $\rho = k\rho_{\mathrm{o}}r$ where r is the radial spherical coordinate.

7.2.3 Find the average stress $\langle\mathbf{T}\rangle$ in a heterogeneous cube of volume b^3 if each face of the cube is subjected to a different pressure. The pressure on the face normal to $\mathbf{e}_1$ and $-\mathbf{e}_1$ is p_1, $\mathbf{e}_2$ and $-\mathbf{e}_2$ is p_2, $\mathbf{e}_3$ and $-\mathbf{e}_3$ is p_3. Find the average stress $\langle\mathbf{T}\rangle$ in a heterogeneous cube when all the surfaces are subjected to the same pressure.

7.3 Effective Material Parameters

One of the prime objectives in the discipline of composite materials, a discipline that has developed over the last half-century, is to evaluate the effective material parameters of a composite in terms of the material parameters and configurational geometries of its constituent components or phases. The purpose of this section is to show that the effective material properties may be expressed in terms of integrals over the surface of the RVE. The conceptual strategy is to average the heterogeneous properties of a material volume and to conceptually replace that material volume with an equivalent homogenous material that will provide exactly the same property volume averages as the real heterogeneous material, allowing the calculation of the material properties of equivalent homogeneous material. The material volume selected for averaging is the RVE and the material properties of equivalent homogeneous material are then called the effective properties of the RVE. The calculational objective is to compute the effective material properties in terms of an average of the real constituent properties. This is accomplished by requiring that the integrals of the material parameters over the bounding surface of the RVE for the real heterogeneous material equal those same integrals obtained when the RVE consists of the equivalent homogeneous material. Thus we seek to express the physical fields of interest for a particular RVE in terms of RVE boundary integrals.

In order to construct an effective anisotropic Hooke's law it is necessary to represent the global or macro stress and strain tensors, $\langle \mathbf{E} \rangle$ and $\langle \mathbf{T} \rangle$, respectively as integrals over the boundary of the RVE. To accomplish this for the stress tensor we begin by noting the easily verified identity (see the Appendix, especially problem A.3.3)

$$(\nabla \otimes \mathbf{x})^T = \mathbf{1}, \quad \frac{\partial x_i}{\partial x_j} e_i \otimes e_j = \delta_{ij} e_i \otimes e_j. \tag{7.2}$$

Using the identity (7.2) a second identity involving the stress is constructed,

$$\mathbf{T} = \mathbf{1} \cdot \mathbf{T} = (\nabla \otimes \mathbf{x})^{\mathrm{T}} \cdot \mathbf{T} = \{\nabla \cdot (\mathbf{T} \otimes \mathbf{x})\}^{\mathrm{T}}. \tag{7.3}$$

The derivation of the last equality in (7.3) employs the fact that the divergence of the stress tensor is assumed to be zero, $\Delta \cdot \mathbf{T} = 0$. This restriction on the stress tensor follows from the stress equation of motion (3.37) when the acceleration and action-at-a-distance forces are zero. The term $\nabla \cdot (\mathbf{T} \otimes \mathbf{x})$in (7.3) is written with the $\mathbf{T}$ first in the parenthesis to indicate that the divergence operator is to be applied to $\mathbf{T}$ and not to $\mathbf{x}$; the latter case would be written as $\nabla \cdot (\mathbf{x} \otimes \mathbf{T})$. The expansion of $\nabla \cdot (\mathbf{T} \otimes \mathbf{x})$is

$$\nabla \cdot (\mathbf{T} \otimes \mathbf{x}) = \nabla \cdot \mathbf{T} \otimes \mathbf{x} + \mathbf{T} \cdot (\nabla \otimes \mathbf{x}) = \mathbf{T} \tag{7.4a}$$

since $\nabla \cdot \mathbf{T} = 0$and $\nabla \otimes \mathbf{x} = \mathbf{1}$. In the indicial notation this development is written

$$\frac{\partial}{\partial x_j}(T_{jk}x_i) = x_i\frac{\partial T_{jk}}{\partial x_j} + T_{jk}\frac{\partial x_i}{\partial x_j} = T_{ik} \tag{7.4b}$$

since $\frac{\partial T_{jk}}{\partial x_j} = 0$, $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$. The expansion of $\nabla \cdot (\mathbf{x} \otimes \mathbf{T})$ is shown below:

$$\nabla \cdot (\mathbf{x} \otimes \mathbf{T}) = (\nabla \cdot \mathbf{x})\mathbf{T} + \mathbf{x} \cdot (\nabla \otimes \mathbf{T}) = 3\mathbf{T} + \mathbf{x} \cdot (\nabla \otimes \mathbf{T}) \tag{7.5}$$

Substitution of the identity (7.3) into the definition for $\langle \mathbf{T} \rangle$ yields

$$\langle \mathbf{T} \rangle = \frac{1}{V_{\text{RVE}}} \oint_{\text{V}} \mathbf{T} dv = \frac{1}{V_{\text{RVE}}} \oint_{\text{V}} \{\nabla \cdot (\mathbf{T} \otimes \mathbf{x})\}^{\text{T}} dv, \tag{7.6}$$

and subsequent application of the divergence theorem (A184) converts the last volume integral in (7.6) to the following surface integral:

$$\langle \mathbf{T} \rangle = \frac{1}{V_{\text{RVE}}} \oint_{\partial V} \{\mathbf{n} \cdot (\mathbf{T} \otimes \mathbf{x})\}^{T} ds. \tag{7.7}$$

Finally, employing the Cauchy relation (3.16), $\mathbf{t} = \mathbf{T} \cdot \mathbf{n}$, provides a relationship between $\langle \mathbf{T} \rangle$ and the integral over the boundary of the RVE depending only upon the stress vector $\mathbf{t}$ acting on the boundary:

$$\langle \mathbf{T} \rangle = \frac{1}{V_{\text{RVE}}} \oint_{\partial V} \mathbf{x} \otimes \mathbf{t} ds. \tag{7.8}$$

This is the desired relationship because it expresses $\langle \mathbf{T} \rangle$ in terms only of boundary information, the surface tractions $\mathbf{t}$ acting on the boundary.

It is even easier to construct a similar representation of $\langle \mathbf{E} \rangle$ as an integral over the boundary of the RVE. To accomplish this set the f in (7.1) equal to $\nabla \otimes \mathbf{u}$ and then employ the divergence theorem to obtain

$$\langle \nabla \otimes \mathbf{u} \rangle = \frac{1}{V_{\text{RVE}}} \oint_{V} (\nabla \otimes \mathbf{u}) dv = \frac{1}{V_{\text{RVE}}} \oint_{\partial V} (\mathbf{n} \otimes \mathbf{u}) ds. \tag{7.9}$$

This form of the divergence theorem employed above is obtained by setting $\mathbf{T} = \mathbf{u} \otimes \mathbf{c}$ in (A184), where $\mathbf{u}$ is the displacement vector and $\mathbf{c}$ is a constant vector. The divergence theorem (A184) may then be written in the form

$$\oint_{V} \mathbf{c} \cdot (\nabla \otimes \mathbf{u}) dv = \oint_{\partial V} \mathbf{c} \cdot (\mathbf{n} \otimes \mathbf{u}) ds, \tag{7.10}$$

and since this must hold for all constant vectors $\mathbf{c}$, it follows that

$$\oint_V \nabla \otimes \mathbf{u} dv = \oint_{\partial V} \mathbf{n} \otimes \mathbf{u} ds. \tag{7.11}$$

The final result in the representation for $\langle \mathbf{E} \rangle$ is achieved immediately by recalling the definition of the small strain tensor $\mathbf{E} = (1/2)((\nabla \otimes \mathbf{u})^{\mathrm{T}} + \nabla \otimes \mathbf{u})$, thus from (7.9)

$$\langle \mathbf{E} \rangle = \frac{1}{V_{\mathrm{RVE}}} \oint_{\partial V} \frac{1}{2} (\mathbf{n} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{n}) ds. \tag{7.12}$$

With the representations (7.8) and (7.12) for the global or macro stress and strain tensors, $\langle \mathbf{E} \rangle$ and $\langle \mathbf{T} \rangle$, respectively, the effective anisotropic elastic constants $\widehat{\mathbf{C}}^{eff}$ are defined by the relation

$$\langle \widehat{\mathbf{T}} \rangle \equiv \widehat{\mathbf{C}}^{eff} \langle \widehat{\mathbf{E}} \rangle. \tag{7.13}$$

This formula provides the tool for the evaluation of the effective material elastic constants of a composite in terms of the material parameters of its constituent components or phases and the arrangement and geometry of the constituent components. In the next section results obtained using this formula are recorded in the cases of spherical inclusions in a matrix material and aligned cylindrical voids in a matrix material.

As a second example of this averaging process for material parameters the permeability coefficients in Darcy's law are considered. In this case the vectors representing volume averages of the mass flow rates $\langle \mathbf{q} \rangle$ and the pressure gradient $\langle \nabla p \rangle$ have to be expressed in terms of surface integrals over the RVE. Obtaining such a formula for the pressure gradient $\langle \nabla p \rangle$ is straightforward. Substitute $\mathbf{r} = p\mathbf{c}$, where $\mathbf{c}$ is a constant vector, into the divergence theorem in the form (A183), then remove the constant vector from the integrals, thus

$$\mathbf{c} \cdot \left(\oint_V \nabla p dv - \oint_{\partial V} (\mathbf{n} p dv \right) = 0; \tag{7.14}$$

then, since (7.14) must hold for all vectors $\mathbf{c}$, it follows from (7.14) and (7.1) that

$$\langle \nabla p \rangle = \frac{1}{V_{\mathrm{RVE}}} \oint_{\partial V} \mathbf{n} p ds. \tag{7.15}$$

Obtaining a surface integral representation for the mass flow rates $\langle \mathbf{q} \rangle$ is slightly more complicated. To start, consider the expression $\nabla \cdot (\mathbf{q} \otimes \mathbf{x})$ and recall the discussion of the special form of the divergence theorem following (7.9) in which it was indicated that $\nabla \cdot (\mathbf{q} \otimes \mathbf{x})$ would imply that the divergence operator would be applied to $\mathbf{q}$, not to $\mathbf{x}$, thus $\nabla \cdot (\mathbf{q} \otimes \mathbf{x}) = (\nabla \cdot \mathbf{q})\mathbf{x} + \mathbf{q}$. Then if the second rank tensor in the divergence theorem in the form (A184) is set equal to $\mathbf{x} \otimes \mathbf{q}$; the divergence of $\mathbf{x} \otimes \mathbf{q}$ is then equal to $\mathbf{x}(\nabla \cdot \mathbf{q}) + \mathbf{q}$, and (A184) yields

$$\oint_V \mathbf{q} + \mathbf{x}(\nabla \cdot \mathbf{q}) dv = \oint_{\partial V} (\mathbf{q} \cdot \mathbf{n})\mathbf{x} \, dv. \tag{7.16}$$

A second integral formula involving $\mathbf{q}$ is obtained by setting $\mathbf{r}$ in the divergence theorem in the form (A183) equal to $\mathbf{q}$, thus

$$\oint_V \nabla \cdot \mathbf{q} \, dv = \oint_{\partial V} \mathbf{q} \cdot \mathbf{n} \, dv. \tag{7.17}$$

Now, if it is assumed that there are no sources or sinks in the volume V and that there is no net flow across the surfaces ∂V, both of the integrals in (7.17) are zero. Then, employing the argument that is used to go from (3.4) to (3.5), it follows that $\Delta \cdot \mathbf{q} = 0$ in the region and using (7.1) and (7.16) one may conclude that

$$\langle \mathbf{q} \rangle = \frac{1}{V_{\text{RVE}}} \oint_{\partial V} (\mathbf{n} \cdot \mathbf{q})\mathbf{x} \, ds. \tag{7.18}$$

Using the representations (7.18) and (7.15) for the volume averages of the mass flow rates $\langle \mathbf{q} \rangle$ and the pressure gradient $\langle \nabla p \rangle$, respectively, the effective anisotropic permeability constants H^{eff} are defined by the relation

$$\langle \mathbf{q} \rangle \equiv -H^{eff} \langle \nabla p \rangle. \tag{7.19}$$

This formula provides the tool for the evaluation of the effective permeability of a porous material in terms of the porous architecture of the solid phase and the properties of the fluid in those pores. In the section after next the result (7.19) is used to evaluate the effective permeability in a simple uniaxial model with multiple aligned cylindrical channels.

Although it is frequently not stated, all continuum theories employ local effective constitutive relations such as those defined by (7.13) and (7.19). This is necessarily the case because it is always necessary to replace the real material by a continuum model that does not contain the small-scale holes and inhomogeneities the real material contains, but which are not relevant to the concerns of the modeler. In the presentations of many continuum theories the substance of this modeling procedure is incorporated in a shorthand statement to the effect that a continuum model is (or will be) employed.

This approach is reasonable because, for many continuum theories, the averaging arguments are intuitively justifiable. This is generally not the case for biological tissues and nanomechanics in general. In almost all continuum theories the notation for RVE averaging is not employed, but it is implied. Thus in continuum models the notation of the typical field variable f really means $\langle f \rangle$, the RVE average of f. In particular, in any continuum theory involving the use of the stress tensor $\mathbf{T}$, it is really the RVE averaged $\langle \mathbf{T} \rangle$ that is being represented, even though an RVE has not been specified.

Problem

7.3.1 Find the average stress $\langle \mathbf{T} \rangle$ in a heterogeneous cube of volume b^3 if each face of the cube is subjected to a different pressure. The pressure on the face normal to $\mathbf{e}_i$ and $-\mathbf{e}_i$ is p_i for $i = 1, 2, 3$. Also, find the average stress $\langle \mathbf{T} \rangle$ in a heterogeneous cube when all surfaces are subjected to the same pressure.

7.4 Effective Elastic Constants

As a first example of the effective Hooke's law (7.13), consider a composite material in which the matrix material is isotropic and the inclusions are spherical in shape, sparse in number (dilute), and of a material with different isotropic elastic constants. In this case the effective elastic material constants are also isotropic and the bulk and shear moduli, K^{eff} and G^{eff}, are related to the matrix material bulk and shear moduli, K_{m} and G_{m}, and Poisson's ratio ν_{m} and to the inclusion bulk and shear moduli, K_{i} and G_{i}, by (Christensen, 1971; pp. 46–47)

$$\frac{K^{\text{eff}}}{K_{\text{m}}} = 1 + \frac{\left(\frac{K_i}{K_{\text{m}}} - 1\right)\phi_s}{1 + \frac{K_{\text{i}} - K_{\text{m}}}{K_{\text{m}} + (4/3)G_{\text{m}}}}, \quad \frac{G^{\text{eff}}}{G_{\text{m}}} = 1 - \frac{15(1 - \nu_{\text{m}})\left(1 - \frac{G_{\text{i}}}{G_{\text{m}}}\right)\phi_s}{7 - 5\nu_{\text{m}} + 2(4 - 5\nu_{\text{m}})\frac{G_{\text{i}}}{G_{\text{m}}}}, \tag{7.20}$$

where ϕ_{s} is the porosity associated with the spherical pores. Thus if the porosity ϕ_{s} and the matrix and inclusion constants K_{m}, G_{m}, K_{i} and G_{i} are known, the formulas (7.20) may be used to determine the effective bulk and shear moduli, K^{eff} and G^{eff}, recalling that for an isotropic material the Poisson's ratio ν_{m} is related to K_{m} and G_{m} by $\nu_{\text{m}} = (3K_{\text{m}} - 2G_{\text{m}})/(6K_{\text{m}} + 2G_{\text{m}})$, see Table 6.2.

As a simple example of these formulas consider the case when $\nu_{\text{m}} = 1/3$ and, since there are only two independent isotropic elastic constants, G_{m} and K_{m} are related. A formula from Table 6.2 may be used to show that $G_{\text{m}} = (3/8)K_{\text{m}}$. Substituting $\nu_{\text{m}} = 1/3$ and $G_{\text{m}} = (3/8)K_{\text{m}}$ in (7.20), they simplify as follows:

$$\frac{K^{\text{eff}}}{K_{\text{m}}} = 1 + \frac{\left(\frac{K_{\text{i}}}{K_{\text{m}}} - 1\right)\phi_{\text{s}}}{1 + \frac{2}{3}\left(\frac{K_{\text{i}}}{K_{\text{m}}} - 1\right)}, \quad \frac{G^{\text{eff}}}{G_{\text{m}}} = 1 - \frac{15\left(1 - \frac{G_{\text{i}}}{G_{\text{m}}}\right)\phi_s}{8 + 7\frac{G_{\text{i}}}{G_{\text{m}}}}. \tag{7.21}$$

The ratio of the effective bulk modulus to the matrix bulk modulus, in the limiting cases when as the ratio of inclusion bulk modulus to the matrix bulk modulus tends to zero and infinity, are given by

$$\lim_{K_i/K_m \to 0} \frac{K^{\text{eff}}}{K_m} = 1 - 3\phi_s \quad \text{and} \quad \lim_{K_i/K_m \to \infty} \frac{K^{\text{eff}}}{K_m} = 1 + \frac{3\phi_s}{2}, \tag{7.22}$$

respectively. The ratio of effective shear modulus to the matrix shear modulus, in the limiting cases when as the ratio of inclusion shear modulus to the matrix shear modulus tends to zero and infinity, are given by

$$\lim_{G_i/G_m \to 0} \frac{G^{\text{eff}}}{G_m} = 1 - \frac{15\phi_s}{8} \quad \text{and} \quad \lim_{\frac{G_i}{G_m} -> \infty} \frac{G^{\text{eff}}}{G_m} = 1 + \frac{15\phi_s}{7}, \tag{7.23}$$

respectively. These results illustrate certain intuitive properties of effective moduli. As the moduli of the inclusion decrease (increase) relative to the moduli of the matrix material, the effective elastic constants decrease (increase) relative to the elastic constants of the matrix material. If the inclusions are voids, the formulas (7.20) simplify to:

$$\frac{K^{\text{eff}}}{K_m} = 1 - \frac{\phi_s}{1 - \frac{K_m}{K_m + (4/3)G_m}}, \quad \frac{G^{\text{eff}}}{G_m} = 1 - \frac{15(1 - \nu_m)\phi_s}{7 - 5\nu_m}. \tag{7.24}$$

In the case when $\nu_m = 1/3$ these two formulas reduce to the first of (7.22) and the first of (7.23) respectively. The first equation of (7.24) is given by Nemat-Nasser and Hori (1999) as their equation (5.2.6b).

If the material of the inclusion is a fluid, (7.20) simplifies to the following:

$$\frac{K^{\text{eff}}}{K_m} = 1 + \frac{\left(\frac{K_f}{K_m} - 1\right)\phi_s}{1 + \frac{K_f - K_m}{K_m + (4/3)G_m}}, \quad \frac{G^{\text{eff}}}{G_m} = 1 - \frac{15(1 - \nu_m)\phi_s}{7 - 5\nu_m}, \tag{7.25}$$

where K_f represents the bulk modulus of the fluid.

Example Exercise 7.4.1

Problem: Calculate the effective bulk modulus K^{eff}, shear modulus G^{eff}, and Young's modulus E^{eff}, for a composite material consisting of a steel matrix material and spherical inclusions. The spherical inclusions are made of magnesium, have a radius r, and are contained within unit cells that are cubes with a dimension of $5r$. The Young's modulus of steel (magnesium) is 200 GPa (45 GPa) and the shear modulus of steel (magnesium) is 77 GPa (16 GPa).

Solution: The isotropic bulk modulus K of a material may be determined from the Young's modulus E and the shear modulus G by use of the formula

$K = EG/(9G - 3E)$ given in Table 6.2. The bulk modulus of steel (magnesium) is 166 GPa (80 GPa). The volume fraction of the spherical inclusions is the ratio of the volume of one sphere, $(4/3)\pi r^3$, to the volume of the unit cell, $(5r)^3$, thus

$$\phi_s = \frac{(4/3)\pi r^3}{125 r^3} = \frac{(4/3)\pi}{125} = \frac{4}{375}\pi = 0.0335.$$

Substitution of the accumulated information into the formulas (7.20) one finds that K^{eff} and G^{eff},

$$\frac{K^{\text{eff}}}{166} = 1 + \frac{\left(\frac{80}{166} - 1\right).335}{1 + \frac{80-166}{166+(4/3)77}}, \quad \frac{G^{\text{eff}}}{77} = 1 - \frac{15(1-.3)\left(1 - \frac{16}{77}\right).0335}{7 - 5(.3) + 2(4 - 5(.3))\frac{16}{77}},$$

are given by 161 GPa and 74 GPa, respectively. The effective Young's modulus E^{eff} is determined to be 103 GPa from the formula $E = 9KG/(3K + G)$ given in Table 6.2.

As a second example of the effective Hooke's law (7.13), consider a composite made up of a linear elastic homogenous, isotropic solid matrix material containing cylindrical cavities aligned in the x_3 direction (Fig. 7.3). Although the matrix material is assumed to be isotropic, the cylindrical cavities aligned in the x_3 direction require that the material symmetry of the composite be transverse isotropy (Chap. 4, Table 4.5). The matrix of tensor compliance components $\hat{\mathbf{S}}$ for the effective transversely isotropic engineering elastic constants is in the following form:

$$\hat{\mathbf{S}} = \begin{bmatrix} \frac{1}{E_1^{\text{eff}}} & \frac{-\nu_{12}^{\text{eff}}}{E_1^{\text{eff}}} & \frac{-\nu_{13}^{\text{eff}}}{E_3^{\text{eff}}} & 0 & 0 & 0 \\ \frac{-\nu_{12}^{\text{eff}}}{E_1^{\text{eff}}} & \frac{1}{E_1^{\text{eff}}} & \frac{-\nu_{13}^{\text{eff}}}{E_3^{\text{eff}}} & 0 & 0 & 0 \\ \frac{-\nu_{13}^{\text{eff}}}{E_3^{\text{eff}}} & \frac{-\nu_{13}^{\text{eff}}}{E_3^{\text{eff}}} & \frac{1}{E_3^{\text{eff}}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2G_{23}^{\text{eff}}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2G_{23}^{\text{eff}}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1+\nu_{12}^{\text{eff}}}{E_1^{\text{eff}}} \end{bmatrix}.$$

The effective elastic constants are expressed in terms of the matrix elastic constants and the volume fraction of cylindrical cavities, which is denoted by ϕ_c. The volume fraction ϕ_c is assumed to be small and the distribution of cavities dilute and random. Terms proportional to the square and higher orders of ϕ_c are neglected. When ϕ_c is this small, several different averaging methods show, using a plane

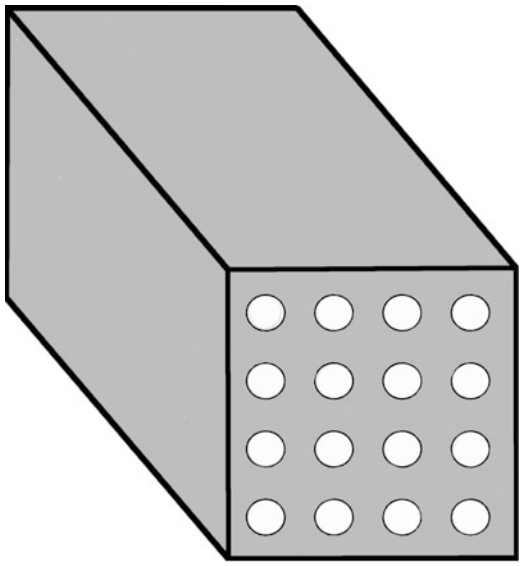


Fig. 7.3 An illustration of a block of a homogeneous material containing cylindrical pores all aligned in one direction

stress assumption in the x_1, x_2 plane, that the effective elastic constants are given by (Nemat-Nasser and Hori 1999; Chap. 2; (5.1.18a, b, c, d) and (5.1.27a, c)):

$$\begin{aligned} &\frac{G_{13}^{\text{eff}}}{G_{\text{m}}} = \frac{G_{23}^{\text{eff}}}{G_{\text{m}}} = 1 - \frac{4\phi_c}{1+\nu_{\text{m}}}, \ \frac{E_1^{\text{eff}}}{E_{\text{m}}} = \frac{E_2^{\text{eff}}}{E_{\text{m}}} = 1 - 3\phi_c, \ \frac{E_3^{\text{eff}}}{E_{\text{m}}} = 1 - \phi_c, \\ &\frac{\nu_{12}^{\text{eff}}}{\nu_{\text{m}}} = \frac{\nu_{21}^{\text{eff}}}{\nu_{\text{m}}} = 1 - \left(3 - \frac{1}{\nu_{\text{m}}}\right)\phi_c, \ \frac{\nu_{31}^{\text{eff}}}{\nu_{\text{m}}} = \frac{\nu_{32}^{\text{eff}}}{\nu_{\text{m}}} = 1, \ \frac{\nu_{13}^{\text{eff}}}{\nu_{\text{m}}} = \frac{\nu_{23}^{\text{eff}}}{\nu_{\text{m}}} = 1 - 2\phi_{\text{c}}. \end{aligned} \tag{7.26}$$

As noted, the cylindrical cavities aligned in the x_3 direction change that material symmetry but the isotropic character of the plane perpendicular to the x_3 direction is retained. The material in the plane perpendicular to the x_3 direction is isotropic; all the elastic constants associated with that plane will be isotropic, as shown in the following exercise.

Example Exercise 7.4.2

Problem: The effective elastic constants (7.23) for the composite composed of an isotropic matrix material containing cylindrical cavities aligned in the x_3 direction are isotropic in the plane perpendicular to the x_3 direction. Verify that this is the case showing that if the matrix material satisfies the isotropy relationship $2G_{\text{m}} = E_{\text{m}}/(1 + \nu_{\text{m}})$, the effective elastic constants $G_{13}^{\text{eff}} = G_{23}^{\text{eff}}$, $E_1^{\text{eff}} = E_2^{\text{eff}}$ and $\nu_{12}^{\text{eff}} = \nu_{21}^{\text{eff}}$ also satisfy the isotropy relationship. However, due to the notation, there is a multitude of equivalent forms: $2G_{13}^{\text{eff}} = 2G_{23}^{\text{eff}} = E_1^{\text{eff}}/(1 + \nu_{12}^{\text{eff}}) = E_2^{\text{eff}}/(1 + \nu_{12}^{\text{eff}}) = E_1^{\text{eff}}/(1 + \nu_{21}^{\text{eff}}) = E_2^{\text{eff}}/(1 + \nu_{21}^{\text{eff}})$.

Solution: The first formula of (7.26) is rewritten as

$$2G_{13}^{\text{eff}} = 2G_{23}^{\text{eff}} = 2G_{\text{m}} - \frac{8G_{\text{m}}\phi_c}{1+\nu_{\text{m}}},$$

then from $2G_{\text{m}} = E_{\text{m}}/(1+\nu_{\text{m}})$ it follows that

$$2G_{13}^{\text{eff}} = 2G_{23}^{\text{eff}} = \frac{E_{\text{m}}(1+\nu_{\text{m}}-4\phi_{\text{c}})}{(1+\nu_{\text{m}})^2}.$$

From the second formula of (7.26) one can see that $E_{\text{m}} = \frac{E_1^{\text{eff}}}{1-3\phi_c}$, thus,

$$2G_{13}^{\text{eff}} = 2G_{23}^{\text{eff}} = \frac{E_1^{\text{eff}}(1+\nu_{\text{m}}-4\phi_{\text{c}})}{(1-3\phi_{\text{c}})(1+\nu_{\text{m}})^2},$$

or to the neglect of squared terms in ϕ_c,

$$2G_{13}^{\text{eff}} = 2G_{23}^{\text{eff}} = \frac{E_1^{\text{eff}}}{(1+\nu_{\text{m}})^2}(1+\nu_{\text{m}}-(3\nu_{\text{m}}-1)\phi_{\text{c}}).$$

From the fourth formula of (7.26) one can write $\nu_{12}^{\text{eff}} = \nu_{21}^{\text{eff}} = \nu_m - (3\nu_m - 1)\phi_c$, thus obtaining the desired result

$$2G_{13}^{\text{eff}} = 2G_{23}^{\text{eff}} = E_1^{\text{eff}}\left(1+\nu_{12}^{\text{eff}}\right).$$

In the special case when ν_{m} is 1/3, the expressions for the effective constants (7.26) simplify to

$$\begin{aligned}
&\frac{G_{13}^{\text{eff}}}{G_{\text{m}}} = \frac{G_{23}^{\text{eff}}}{G_{\text{m}}} = \frac{E_1^{\text{eff}}}{E_{\text{m}}} = \frac{E_2^{\text{eff}}}{E_{\text{m}}} = 1-3\phi_c,\\
&\frac{E_3^{\text{eff}}}{E_{\text{m}}} = 1-\phi_c,\ \frac{\nu_{12}^{\text{eff}}}{\nu_{\text{m}}} = \frac{\nu_{21}^{\text{eff}}}{\nu_{\text{m}}} = 1,\ \frac{\nu_{31}^{\text{eff}}}{\nu_{\text{m}}} = \frac{\nu_{32}^{\text{eff}}}{\nu_{\text{m}}} = 1,\ \frac{\nu_{13}^{\text{eff}}}{\nu_{\text{m}}} = \frac{\nu_{23}^{\text{eff}}}{\nu_{\text{m}}} = 1-2\phi_c.
\end{aligned} \tag{7.27}$$

From these results it is apparent that the Young's modulus and the shear modulus in the transverse plane are more severely reduced than the axial Young's modulus as the porosity increases. In the example below, the second of (7.27), $E_3^{\text{eff}}/E_{\text{m}} = 1-\phi_{\text{c}}$, is constructed from a mechanics of materials argument.

Example Exercise 7.4.3

Problem: Recall the mechanics of materials formula for the deflection $\delta = PL/AE$ of a bar of cross-sectional area A, length L, and modulus E subjected to an axial force P. Apply this formula to the bar with axially aligned cylindrical cavities illustrated in Fig. 7.3 to show that $E_3^{\text{eff}}/E_{\text{m}} = 1-\phi_{\text{c}}$. The bar and the cylindrical cavities are aligned in the x_3 direction.

Solution: The formula $\delta = PL/AE$ will be applied in two ways to the bar with axially aligned cylindrical cavities illustrated in Fig. 7.3 and subjected to an axial force P. First apply the formula to the matrix material imagining a reduced cross-sectional area A_m not containing the voids, thus $E_m = PL/A_m\delta$. Next apply the formula to the entire bar containing the voids, thus $E_3^{eff} = PL/A\delta$. The relationship between the cross-sectional total area A and the cross-sectional area A_m occupied by the matrix material is $A_m = A(1 - \phi_c)$. The desired result is established by eliminating P, L, δ, A, and A_m between these three formulas.

Problems

7.4.1 Calculate the effective bulk modulus, K^{eff}, shear modulus, G^{eff}, and Young's modulus, E^{eff}, for a composite material consisting of a steel matrix material with spherical inclusions that contain water. The spherical inclusions, which have a radius r, are contained within unit cells that are cubes with a dimension of $4r$. The Young's modulus of steel is 200 GPa and the shear modulus of steel is 77 GPa. The bulk modulus of water is 2.3 GPa.

7.4.2 Calculate the effective bulk modulus, K^{eff}, shear modulus, G^{eff}, and Young's modulus, E^{eff}, for a composite material consisting of a magnesium matrix material and spherical inclusions. The spherical inclusions are made of steel, have a radius r, and are contained within unit cells that are cubes with a dimension of $5r$. The Young's modulus of steel (magnesium) is 200 GPa (45 GPa) and the shear modulus of steel (magnesium) is 77 GPa (16 GPa).

7.4.3 Calculate the effective moduli for a composite material consisting of a steel matrix material and cylindrical voids of radius r contained in unit squares $8r$ by $8r$. The Young's modulus of steel is 77 GPa and its Poisson's ratio is 0.33.

7.4.4 Considering the effective moduli for a composite material consisting of a matrix material and aligned cylindrical voids, show that for small values of the volume fraction the in-plane effective shear moduli and the in-plane effective Young's moduli decrease more rapidly with increasing porosity than the out-of-plane effective Young's moduli. For simplicity consider the case when ν_m is 1/3.

7.5 Effective Permeability

In this section the effective axial permeability of the bar with axially aligned cylindrical cavities illustrated in Fig. 7.3 is calculated. The method used is the simplest one available to show that Darcy's law is a consequence of the application of the Newtonian law of viscosity to a porous medium with interconnected pores. The Navier–Stokes equations (6.37) are a combination of the Newtonian law of viscosity (5.36 N) and the stress equations of motion (3.37), as was shown in the previous chapter. Thus one can say that Darcy's law is a consequence of the application of either the Newtonian law of viscosity, or the Navier–Stokes equations, to a porous medium with interconnected pores. There are a number of

more general, more rigorous, and more elegant proofs of this result (Ene and Sanchez-Palencia 1975; Sanchez-Palencia 1980; Burridge and Keller 1981) but the one presented below suffices to make the point.

Consider the bar with axially aligned cylindrical cavities illustrated in Fig. 7.3 as a porous medium, the pores being the axially aligned cylindrical cavities. Each pore is identical and can be treated as a pipe for the purpose of determining the fluid flow through it. In the case of pipe flow under a steady pressure gradient $\partial p/\partial x_3$, the velocity distribution predicated by the Navier–Stokes equations is a parabolic profile (compare Example 6.4.3),

$$v_3 = -\left(\frac{\partial p}{\partial x_3}\right)\frac{r_o^2}{4\mu}\left(1-\frac{r^2}{r_o^2}\right), \tag{7.28}$$

where μ is the viscosity, r_o is the radius of the pipe and r and x_3 are two of the three cylindrical coordinates. The volume flow rate through the pipe is given by

$$Q = 2\pi\int_0^{r_\mathrm{o}} r v_3 dr = -\left(\frac{\partial p}{\partial x_3}\right)\frac{\pi r_o^4}{8\mu}, \tag{7.29}$$

which, multiplied by the total number of axially aligned cylindrical cavities per unit area, n_c, gives the average volume flow rate per unit area along the bar,

$$\langle \mathbf{q}\rangle = \frac{1}{V_\mathrm{RVE}}\oint_{\partial V}(\mathbf{n}\cdot\mathbf{q})\mathbf{x}\mathrm{d}s = \left(-\frac{\partial p}{\partial x_3}\right)n_c\frac{\pi r_o^4}{8\mu}. \tag{7.30}$$

The pressure gradient $\partial p/\partial x_3$ is a constant in the bar, hence the average of the pressure gradient over the bar is given by the constant value;

$$\langle \nabla p\rangle = \frac{1}{V_\mathrm{RVE}}\oint_V \nabla p \mathrm{d}v = \frac{\partial p}{\partial x_3}, \tag{7.31}$$

a result that, combined with (7.28), yields

$$\langle \mathbf{q}\rangle = -n_c\frac{\pi r_o^4}{8\mu}\langle \nabla p\rangle. \tag{7.32}$$

A comparison of this representation for $\langle \mathbf{q}\rangle$ with that of (7.19) yields the representation for the hydraulic permeability H_{33}^{eff} in the x_3 direction,

$$H_{33}^{\mathrm{eff}} = n_\mathrm{c}\frac{\pi r_o^4}{8\mu}. \tag{7.33}$$

This derivation has demonstrated that the hydraulic permeability depends upon the viscosity of the fluid μ and the geometry of the pores. The intrinsic permeability is defined as the hydraulic permeability times the viscosity of the fluid in the pores, $K_{33}^{\text{eff}} = \mu H_{33}^{\text{eff}} = n_{\text{c}} \frac{\pi r_o^4}{8}$. Since n_{c}, the total number of axially aligned cylindrical cavities per unit area has the dimension of one over length squared, the intrinsic permeability $K_{33}^{\text{eff}} = n_{\text{c}} \frac{\pi r_o^4}{8}$ is of dimension length squared. The intrinsic permeability is independent of the type of fluid in the pores and dependent only upon the size and geometrical arrangement of the pores in the medium. The comments above concerning the connection between, and the relative properties of, the intrinsic permeability and the hydraulic permeability are general and not tied to the particular model used here to calculate the effective permeability H_{33}^{eff}, thus $\mathbf{K}^{eff} = \mu \mathbf{H}^{eff}$ in general.

7.6 Structural Gradients

A material containing a structural gradient, such as increasing/decreasing porosity is said to be a gradient material. Figure 7.4 is an illustration of an example material with a layered structural gradient. Spheres of varying diameters and one material type are layered in a matrix material of another type. As a special case, the spheres may be voids. Figure 7.5 is an illustration of a material with a structural gradient that is not formed by layering. Spheres of varying diameters and one material type are graded in a size distribution in a material of another type. Again, as a special case, the spheres may be voids. Gradient materials may be man-made, but they also occur in nature. Examples of natural materials with structural gradients include cancellous bone and the growth rings of trees.

The RVE plays an important role in determining the relationship between the structural gradient and the material symmetry. In a material with a structural gradient, if it is not possible to select an RVE so that it is large enough to adequately

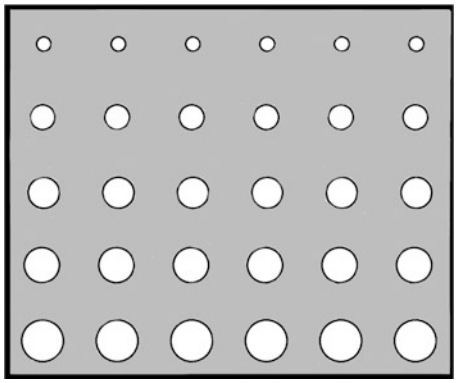


Fig. 7.4 An illustration of a material with a layered structural gradient. Spheres of varying diameters and one material type are layered in a material of another type. As a special case, the spheres may be voids. From Cowin (2002)

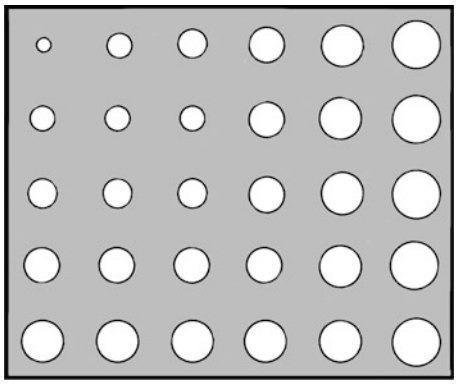


Fig. 7.5 An illustration of a material with a unlayered structural gradient. Spheres of varying diameters and one material type are graded in a size distribution in a material of another type. As a special case, the spheres may be voids. From Cowin (2002)

average over the microstructure and *also* small enough to insure that the structural gradient across the RVE is negligible, then it is necessary to restrict the material symmetry to accommodate the gradient. However, in a material with a structural gradient, if an RVE may be selected so that it is large enough to adequately average over the microstructure and small enough to insure that the structural gradient across the RVE is negligible, then it is not necessary to restrict the material symmetry to accommodate the gradient.

It is easy to see that a structural gradient is incompatible with a plane of symmetry. If a given direction is the direction of a structural gradient it cannot also be the direction of a normal to a plane of symmetry, nor any projected component of a normal to a plane of reflective symmetry, because the increasing or decreasing structure with increasing distance from the reference plane, such as the layered spherical inclusions in Fig. 7.4, would be increasing on one side of the plane and decreasing on the other side of the plane violating mirror symmetry. Gradient materials are thus a type of chiral material described in Sect. 4.9. Gradient materials are chiral materials because they require that the material symmetries possess at least one direction that is not a direction associated with a normal to a plane of reflective symmetry, nor any projected component of a normal to a plane of reflective symmetry. The point is that, on the same scale, the normal to a plane of symmetry and a material structural gradient are incompatible unless they are perpendicular. This incompatibility restricts the type of linear elastic symmetries possible for gradient materials to the same symmetries that are possible for chiral materials (Sect. 4.9), namely, trigonal, monoclinic, and triclinic symmetries (c.f., Fig. 4.10).

The normal to a plane of material symmetry can only be perpendicular to the direction of a uniform structural gradient. The argument for this conclusion is a purely geometrical one. First note that the direction of a normal to a plane of material symmetry cannot be coincident with the direction of the structural gradient because the structural gradient is inconsistent with the reflective structural symmetry required by a plane of mirror symmetry. Next consider the case when the normal to a plane of material symmetry is inclined, but not perpendicular, to the direction

of the structural gradient. In this case, the same situation prevails because the structural gradient is still inconsistent with the reflective structural symmetry required by a plane of mirror symmetry. The only possibility is that the normal to a plane of material symmetry is perpendicular to the direction of the structural gradient. Thus, it is concluded that the only linear elastic symmetries permitted in a material containing a structural gradient are those symmetries characterized by having all their normals to their planes of mirror or reflective symmetry perpendicular to the structural gradient. The caveat to this conclusion is that the structural gradient and the material symmetry are at the same structural scale in the material. Only the three linear elastic symmetries, triclinic, monoclinic, and trigonal, satisfy the condition that they admit a direction perpendicular to all the normals to their planes of mirror or reflective symmetry. Trigonal symmetry has the highest symmetry of the three symmetries and admits a direction that is not a direction associated with a normal to a plane of reflective symmetry, nor any projected component of a normal to a plane of reflective symmetry.

In summary, in a material with a structural gradient, if an RVE may be selected so that is large enough to adequately average over the microstructure and small enough to insure that the structural gradient across the RVE is negligible, then it is not necessary to restrict the material symmetry to accommodate the gradient. However, in a material with a structural gradient, if an RVE cannot be selected such that the structural gradient across an adequately sized RVE is negligible, then it is not necessary to restrict the material symmetry to accommodate the structural gradient.

7.7 Tensorial Representations of Microstructure

The description and measurement of the microstructure of a material with multiple distinct constituents is called quantitative stereology (Underwood 1969) or texture analysis (Bunge 1982) or, in the case of biological tissues, it becomes part of histology. The concern here is primarily with the modeling of the material microstructure and only secondarily with techniques for its measurement.

It is recognized that the volume fraction of a constituent material is the primary geometric measure of local material structure in a material with multiple distinct constituents. This means that in the purely geometric kinematic description of the arrangement of the microstructure the volume fraction of a constituent material is the primary parameter in the geometric characterization of the microstructure.

The volume fraction of a constituent in a multiconstituent material does not provide information on the arrangement or architecture of microstructure of the multiconstituent material, only information on the volume of the constituent present. The second best measure of local material microstructure depends upon the type of material microstructure being modeled and the objective of the modeler. One approach to the modeling of material microstructures is to use vectors and tensors to characterize the microstructural architecture. For example, in theories for

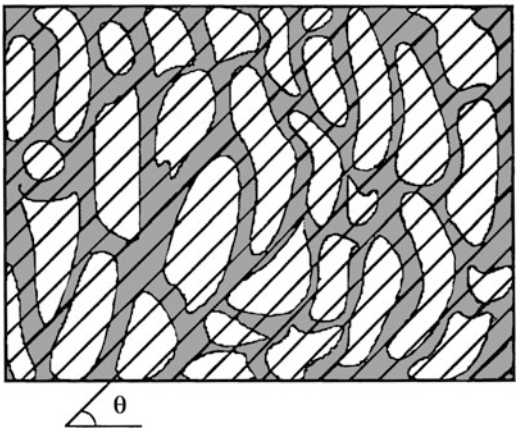


Fig. 7.6 Test lines superimposed on a cancellous bone specimen. The test lines are oriented at the angle θ. The mean intercept length measured at this angle is denoted $L(\theta)$. From Cowin and Mehrabadi (1989)

the flow of liquid crystals a vector is often used to characterize the long axis of the liquid crystal. In early liquid crystal theories the formation of the constitutive equation for the liquid crystal follows the development outlined in Chap. 5 for the constitutive equation for the Newtonian law of viscosity up to equation (5.4 N). In the early liquid crystal theory the equation equivalent to (5.4 N) was assumed to also depend upon the unit vector **n**, a vector coincident with the long axis of the liquid crystal (de Gennes and Prost 1993).

The modeling of the microstructural architecture of a material with two distinct constituents, one dispersed in the other, has been accomplished using a second rank tensor called the fabric tensor. Fabric tensors may be defined in a number of ways; it is required only that the fabric tensor be a positive definite tensor that is a quantitative stereological measure of the microstructural architecture, a measure whose principal axes are coincident with the principal microstructural directions and whose eigenvalues are proportional to the distribution of the microstructure in the associated principal direction. The fabric tensor is a continuum point property (as usual its measurement requires a finite test volume or RVE) and is therefore considered to be a continuous function of position in the material.

One type of fabric tensor is the mean intercept length (MIL) tensor. The MIL in a material is the average distance, measured along a particular straight line, between two interfaces of the two phases or constituents (Fig. 7.6). The value of the mean intercept length is a function of the slope of the line, θ, along which the measurement is made in a specified plane. A grid of parallel lines is overlaid on the plane through the specimen of the binary material and the distance between changes of phase, first material to second material or second material to first material, are counted. The average of these lengths is the mean intercept length at the angle θ, the angle characterizing the orientation of the set of parallel lines. Figure 7.6 illustrates such measurements. It is frequently observed that when the mean intercept lengths measured in the selected plane in the specimen are plotted in a polar diagram as a function of θ, producing a closed curve in the plane. If the test lines are rotated

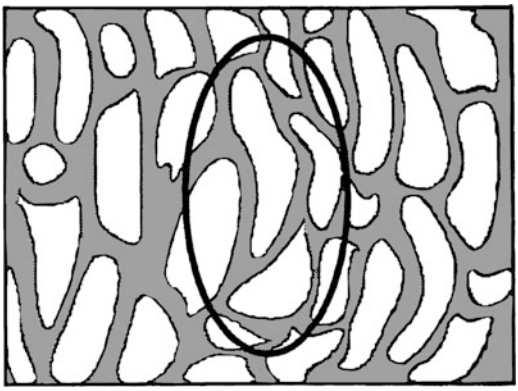


Fig. 7.7 An ellipse is constructed by making a polar pole of the average value of the mean intercept length $L(\theta)$ for each value of the angle θ. From Cowin and Mehrabadi (1989)

through several values of θ and the corresponding values of mean intercept length $L(\theta)$ are measured, the data are often found to fit the equation for an ellipse very closely,

$$(1/L^2(\theta)) = M_{11}\cos^2\theta + M_{22}\sin^2\theta + 2M_{12}\ \sin\theta\cos\theta, \tag{7.34}$$

where M_{11}, M_{22} and M_{12} are constants when the reference line from which the angle θ is measured is constant. The subscripts 1 and 2 indicate the axes of the x_1, x_2 coordinate system to which the measurements are referred. The ellipse is shown superposed on the binary microstructure it represents in Fig. 7.7. The mean intercept lengths in all directions in a three-dimensional binary microstructure structure would be represented by an ellipsoid and would therefore be equivalent to a positive definite second rank tensor. The constants M_{11}, M_{22} and M_{12} introduced in the foregoing are then the components in a matrix representing the tensor M which are related to the mean intercept length $L(\mathbf{n})$, where $\mathbf{n}$ is a unit vector in the direction of the test line, by $(1/L^2(\mathbf{n})) = \mathbf{n}\cdot\mathbf{M}\cdot\mathbf{n}$.

The fabric tensor is commonly computed from data obtained by using stereographic or image analysis methods (Odgaard 1997) such as Mean Intercept Length (MIL) (Whitehouse, 1974), Volume Orientation (VO) (Odgaard et al., 1990), Star Length Distribution (SLD) (Smit et al. 1998), or Intercept Segment Deviation (Chiang et al. 2006). The experimental procedure for the fabric measurement of cancellous bone is described by Whitehouse (1974), Harrigan and Mann (1984), Turner and Cowin (1987) and Turner et al. (1990). A number of ways of constructing a fabric tensor for a material with two distinct constituents are described by Odgaard (1997, 2001) for a particular porous material, cancellous bone. These methods are applicable to any material with at least two distinct constituents and include the stereological methods known as the mean intercept length method, the volume orientation method and the star volume distribution method. In multiphase materials such as cellular materials, foams, and cancellous bone, the unit vectors may represent the orientation of the interface surface area or

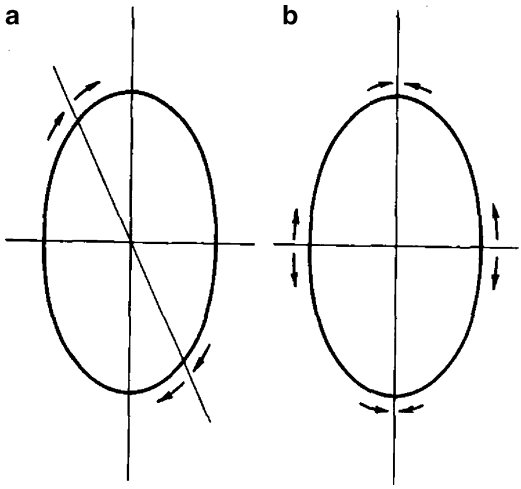


Fig. 7.8 The mean intercept length ellipse as an indicator of the normals to planes of mirror symmetry. (**a**) The direction of the skewed line cannot be a normal to a plane of mirror symmetry because the direction of increasing mean intercept length has an anti-mirror symmetry. (**b**) The directions of the major and minor axes of the ellipse are normals to planes of mirror symmetry. From Cowin and Mehrabadi (1989)

the inclusion volume orientation (Hilliard 1967). As pointed out by Odgaard et al. (1997), for the same microstructural architecture, these different fabric tensor definitions each lead to representations of the data that are, effectively, the same.

The existence of a mean intercept or fabric ellipsoid for an anisotropic porous material suggests elastic orthotropic symmetry. To visualize this result, consider an ellipse, illustrated in Fig. 7.8, that is one of the three principal planar projections of the mean intercept length ellipsoid. The planes perpendicular to the major and minor axes of the ellipse illustrated in Fig. 7.8(b) are planes of mirror symmetry because the increasing or decreasing direction of the mean intercept length, indicated by the arrow heads, is the same with respect to either of these planes. On the other hand, if one selects an arbitrary direction such as that illustrated in Fig. 7.8(a), it is easy to see that the selected direction is not a normal direction for a plane of mirror symmetry because the direction of increasing mean intercept length is reversed from its appropriate mirror image position. Therefore, there are only two planes of mirror symmetry associated with the ellipse. Considering the other two ellipses that are planar projections of the mean intercept ellipsoid, the same conclusion is reached. Thus, only the three perpendicular principal axes of the ellipsoids are normals to planes of mirror symmetry. This means that, if the matrix materials involved are isotropic the material symmetry will be determined only by the fabric ellipsoid and that the material symmetry will be orthotropy or a greater symmetry. If two of the principal axes of the mean intercept length ellipsoid were equal (i.e. an ellipsoid of revolution or, equivalently, a spheroid, either oblate or prolate), then the elastic symmetry of the material is transversely isotropic. If the fabric ellipsoid degenerates to a sphere, the elastic symmetry of the material is isotropic.

The fabric tensor employed here is denoted by $\mathbf{F}$ and is related to the mean intercept length tensor $\mathbf{M}$ by $\mathbf{F} = \mathbf{M}^{-1/2}$. The positive square root of the inverse of $\mathbf{F}$ is well defined because $\mathbf{M}$ is a positive definite symmetric tensor. The principal axes

of **F** and **M** coincide, only the shape of the ellipsoid changes. **M** is a positive definite symmetric tensor because it represents an ellipsoid. The fabric tensor or mean intercept ellipsoid can be measured using the techniques described above for a cubic specimen. On each of three orthogonal faces of a cubic specimen of cancellous bone an ellipse will be determined from the directional variation of mean intercept length on that face. The mean intercept length tensor or the fabric tensor can be constructed from these three ellipses that are the projections of the ellipsoid on three perpendicular planes of the cube.

7.8 The Stress–Strain–Fabric Relation

If the porous medium may be satisfactorily represented as an orthotropic, linearly elastic material, the associated elasticity tensor $\hat{\mathbf{C}}$ will depend upon the porous architecture represented by the fabric tensor **F**. The six-dimensional second rank elasticity tensor $\hat{\mathbf{C}}$ relates the six-dimensional stress vector to the six-dimensional infinitesimal strain vector in the linear anisotropic form of Hooke's law, (4.36 H). The elasticity tensor $\hat{\mathbf{C}}$ completely characterizes the linear elastic mechanical behavior of the porous medium. If it is assumed that all the anisotropy of the porous medium is due to the anisotropy of its solid matrix pore structure, that is to say that the matrix material is itself isotropic, then a relationship between the components of the elasticity tensor $\hat{\mathbf{C}}$ and **F** can be constructed. From previous studies of porous media it is known that the medium's elastic properties are strongly dependent upon its apparent density or, equivalently, the solid volume fraction of matrix material. This solid volume fraction is denoted by ν and is defined as the volume of matrix material per unit bulk volume of the porous medium. Thus $\hat{\mathbf{C}}$ will be a function of ν as well as **F**. A general representation of $\hat{\mathbf{C}}$ as a function of ν and **F** was developed based on the assumption that the matrix material of the porous medium is isotropic and that the anisotropy of the porous medium itself is due only to the geometry of the microstructure represented by the fabric tensor **F**. The mathematical statement of this notion is that the stress tensor **T** is an isotropic function of the strain tensor **E** and the fabric tensor **F** as well as the porosity φ. Thus the tensor valued function

$$\mathbf{T} = \mathbf{T}(\varphi, \mathbf{E}, \mathbf{F}), \tag{7.35}$$

has the property that

$$\mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathrm{T}} = \mathbf{T}(\varphi, \mathbf{Q}\mathbf{E}\mathbf{Q}^{\mathrm{T}}, \mathbf{Q}\mathbf{F}\mathbf{Q}^{\mathrm{T}}), \tag{7.36}$$

for all orthogonal tensors **Q**. The most general form of the relationship between the stress tensor and the strain and the fabric tensors consistent with the isotropy assumption (7.37) is

$$\begin{aligned}\mathbf{T} = {} & a_1\mathbf{1}\,\mathrm{tr}\,\mathbf{E} + a_2\,(\mathbf{F}[\mathrm{tr}\mathbf{E}] + \mathbf{1}\mathrm{tr}\,[\mathbf{E}\cdot\mathbf{F}]) + a_3(\mathbf{1}\mathrm{tr}[\mathbf{F}\cdot\mathbf{E}\cdot\mathbf{F}] + (\mathrm{tr}\mathbf{E})\mathbf{F}^2) \\ & + b_1\mathbf{F}\mathrm{tr}[\mathbf{F}\cdot\mathbf{E}] + b_2(\mathbf{F}\mathrm{tr}[\mathbf{F}\cdot\mathbf{E}\cdot] + (\mathrm{tr}\mathbf{E}\cdot\mathbf{F})\mathbf{F}^2) + b_3(\mathbf{F}^2\mathrm{tr}[\mathbf{F}\cdot\mathbf{E}\cdot\mathbf{F}]) \\ & + 2c_1\mathbf{E} + 2c_2(\mathbf{F}\cdot\mathbf{E} + \mathbf{E}\cdot\mathbf{F}) + 2c_3(\mathbf{F}^2\cdot\mathbf{E} + \mathbf{E}\cdot\mathbf{F}^2)\end{aligned} \tag{7.37}$$

where a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, and c_3 are functions of φ, tr**F**, tr$\mathbf{F}^2$ and tr$\mathbf{F}^3$. This representation has been used to represent the elastic behavior of highly porous cancellous bone tissue. The form of the functional dependence of the elasticity tensor C_{ijkm} (recall the stress–strain form of the anisotropic Hooke's law, $T_{ij} = C_{ijkm}E_{km}$) upon fabric is given by

$$\begin{aligned}C^d_{ijkm} = {} & a^c_1\delta_{ij}\delta_{km} + a^c_2(F_{ij}\delta_{km} + \delta_{ij}F_{km}) + a^c_3(\delta_{ij}F_{kq}F_{qm} + \delta_{km}F_{iq}F_{qj}) \\ & + b^c_1F_{ij}F_{km} + b^c_2(F_{ij}F_{kq}F_{qm} + F_{km}F_{iq}F_{qj}) + b^c_3F_{is}F_{sj}F_{kq}F_{qm} \\ & + c^c_1(\delta_{ki}\delta_{mj} + \delta_{mi}\delta_{kj}) + c^c_2(F_{ki}\delta_{mj} + F_{kj}\delta_{mi} + F_{im}\delta_{kj} + F_{mj}\delta_{ki}) \\ & + c^c_3(F_{ir}F_{rk}\delta_{mj} + F_{kr}F_{rj}\delta_{mi} + F_{ir}F_{rm}\delta_{kj} + F_{mr}F_{rj}\delta_{ik})\end{aligned} \tag{7.38}$$

where, as in (7.38) the a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, and c_3 with the superscript c are functions of φ, tr**F**, tr$\mathbf{F}^2$ and tr$\mathbf{F}^3$. The fourth-rank elastic compliance tensor S_{ijkm} (recall $E_{ij} = S_{ijkm}T_{km}$) for the strain–stress relation is

$$\begin{aligned}S_{ijkm} = {} & a^s_1\delta_{ij}\delta_{km} + a^s_2(F_{ij}\delta_{km} + \delta_{ij}F_{km}) + a^s_3(\delta_{ij}F_{kq}F_{qm} + \delta_{km}F_{iq}F_{qj}) \\ & + b^s_1F_{ij}F_{km} + b^s_2(F_{ij}F_{kq}F_{qm} + F_{km}F_{iq}F_{qj}) + b^s_3F_{is}F_{sj}F_{kq}F_{qm} \\ & + c^s_1(\delta_{ki}\delta_{mj} + \delta_{mi}\delta_{kj}) + c^s_2(F_{ki}\delta_{mj} + F_{kj}\delta_{mi} + F_{im}\delta_{kj} + F_{mj}\delta_{ki}) \\ & + c^s_3(F_{ir}F_{rk}\delta_{mj} + F_{kr}F_{rj}\delta_{mi} + F_{ir}F_{rm}\delta_{kj} + F_{mr}F_{rj}\delta_{ik})\end{aligned} \tag{7.39}$$

where a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2 and c_3 with the superscript s are functions of ϕ and the invariants of **F** and porosity φ or solid volume fraction (1-φ) of the material. The least elastic material symmetry for which the representation holds is orthotropy. It therefore holds for transverse isotropy and isotropy as well as orthotropy.

Problems

7.8.1 Specialize the representation (7.39) for the components of the fourth order compliance tensor to the case of isotropic symmetry. Relate the 9 coefficients in your result to the two elastic constants in equation (6.26), Young's modulus and Poisson's ratio. Upon what parameter(s) do the Young's modulus and Poisson's ratio obtained depend?

7.8.2 Specialize the representation (7.38) for the components of the fourth order elasticity tensor to the case of isotropic symmetry. Relate the 9 coefficients in your result to the two Lamé moduli in equation (6.23) or (6.24). The resulting Lamé moduli are functions of which parameter(s)?

7.9 Relevant Literature

Good discussions of the RVE concept appear in the books by Christensen (1979) and Nemat Nasser and Hori (1999). The inaugural quote of the next Chapter, Chap. 8, describes a very early application of the RVE concept by Maurice Biot (1941) to poroelasticity.

The books on composite materials by Jones (1975), Tsai and Hahn (1980), Hull (1981) and Halpin (1984) tend to specialize in laminated composites, that is to composites constructed in the manner illustrated in Fig. 7.7. The early book of Spencer (1972), which develops the theory of fiber reinforced materials, has been used in developing mathematical models for several biological tissues. Early rigorous summaries of composite material modeling were given by Sendeckyj (1974) and Christensen (1979); a contemporary rigorous and exhaustive development of elastic composite theory is contained in Nemat Nasser and Hori (1999). Torquato (2002) presents a general, contemporary, rigorous and exhaustive development of the quantification and modeling of microstructure.

Wainwright et al. (1976) and Neville (1993) are two interesting books on biological composites, both plant and animal natural composites. Neville's book explores nature's wide use of the helicoidal structure as a composite building geometry. The book of Vincent (1982) is of more restricted scope. Wood is the most widely analyzed natural composite material and its mechanics and microstructure are described in an abundant literature, for example Kollmann and Cote (1968) and Bodig and Jayne (1982). The book by Niklas (1992) on plant biomechanics covers a wider range of the plant world, but it also deals with mechanism and microstructure.

The quantification and measurement of material microstructure is not addressed in the present book, but it is addressed in books on histology, metallurgy, and materials science. The book of Underwood (1969) on quantitative stereology, that of Bunge (1982) on texture analysis in material science and any standard textbook on histology are examples of this literature.

References

Biot MA (1941) General theory of three-dimensional consolidation. J Appl Phys 12:155–164

Bodig J, Jayne BA (1982) Mechanics of wood and wood composites. Van Nostrand, New York

Bunge HJ (1982) Texture analysis in materials science. Butterworths, London

Burridge R, Keller JB (1981) Poroelasticity equations derived from microstructure. J Acoust Soc Am 70:1140–1146

Christensen RM (1979) Mechanics of composite materials. Wiley, New York

Chiang MYM, Wang X, Landis FA, Dunkers J, Snyder CR (2006) Quantifying the directional parameter of structural anisotropy in porous media. Tissue Eng 12(6):1597–1606

Cowin SC, Mehrabadi MM (1989) Identification of the elastic symmetry of bone and other materials. J Biomech 22:503–515

Cowin SC (2002) Elastic symmetry restrictions from structural gradients. In: Podio-Guidugli P, Brocato M (eds) Rational continua, classical and new—a collection of papers dedicated to Gianfranco Capriz on the occasion of his 75th birthday. Springer Verlag, Milano. ISBN 88-470-0157-9
de Gennes PG, Prost J (1993) The physics of liquid crystals. Clarendon, Oxford
Ene HJ, Sanchez-Palencia E (1975) Equations et phénoménes di surface pour l'écoulement dans unmodèle de milieu poreux. J de Mécanique 14:73–81
Halpin JC (1984) Primer on composite materials: analysis. Technomic Publ, Lancaster, PA
Hilliard JE (1967) Determination of structural anisotropy. Stereology—Proceedings of the 2nd International Congress for Stereology, Chicago. Springer, Berlin, p 219
Hull D (1981) An introduction to composite materials. Cambridge Univ Press, Cambridge
Jones RM (1975) Mechanics of composite materials. Scripta, Washington
Kollmann FFP, Cote WA (1968) Principles of wood science and technology. Springer, New York
Nemat-Nasser S, Hori M (1999) Micromechanics: overall properties of heterogeneous materials. Elsevier, Amsterdam
Neville C (1993) Biology of fibrous composites. Cambridge Univ. Press, Cambridge
Niklas KJ (1992) Plant biomechanics. Univ. of Chicago Press, Chicago
Odgaard A (1997) Three-dimensional methods for quantification of cancellous bone architecture. Bone 20:315–328
Odgaard A (2001) Quantification of cancellous bone architecture. In: Cowin SC (ed) Bone mechanics handbook. CRC Press, Boca Raton, FL
Prandtl L, Tietjens OG (1934) Fundamentals of hydro- and aeromechanics. Dover, New York
Sanchez-Palencia E (1980) Non-homogeneous media and vibration theory. Springer, Berlin
Sendeckyj GP (1974) Elastic behavior of composites. In: Sendeckyj GP (ed) Mechanics of composite materials, vol 2. Academic Press, New York and London, pp 46–83
Smit TH, Schneider E, Odgaard A (1998) Star length distribution: a volume-based concept for the characterization of structural anisotropy. J Microsc (Oxford) 191:249–257
Spencer AJM (1972) Deformations of fibre-reinforced materials. Clarendon, Oxford
Torquato S (2002) Random heterogeneous materials. Springer, New York
Tsai SW, Hahn HT (1980) Introduction to composite materials. Technomic Publ, Westport, CT
Underwood EE (1969) Quantitative stereology. Addison-Wesley, Reading, MA
Vincent JFV (1982) Structural biomaterials. Wiley, New York
Wainwright SA, Biggs WD, Currey JD, Gosline JM (1976) Mechanical design in organisms. Edward Arnold, London

Chapter 8
Quasistatic Poroelasticity

The theme for this chapter is contained in a quote from the 1941 paper of M. A. Biot that clearly describes an RVE: "Consider a small cubic element of soil, its sides being parallel with the coordinate axes. This element is taken to be large enough compared to the size of the pores so that it may be treated as homogeneous, and at the same time small enough, compared to the scale of the macroscopic phenomena in which we are interested, so that it may be considered as infinitesimal in the mathematical treatment."

8.1 Poroelastic Materials

Thinking about, or experimenting with, a fluid-saturated sponge can develop an intuitive sense of the response of a saturated elastic porous medium to mechanical loading. If a fluid-saturated sponge is compressed, fluid will flow from the sponge. If the sponge is in a fluid reservoir and compressive pressure is subsequently removed, the sponge will reimbibe the fluid and expand. The volume of the sponge will also increase if its exterior openings are sealed and the external pore fluid pressure is increased. However, the volume of the sponge will also decrease if its exterior openings are sealed and the external pore fluid pressure is decreased. The basic ideas underlying the theory of poroelastic materials are that the pore fluid pressure contributes to the total stress in the porous matrix medium and that the pore fluid pressure alone can strain the porous matrix medium. There is fluid movement in a porous medium due to differences in pore fluid pressure created by different pore volume strains associated with the mechanical loading of the porous medium.

This chapter contains a presentation of the quasistatic theory of poroelasticity. Quasistatic means that inertia terms, that is to say mass times acceleration terms, are neglected. Thus quasistatic means that processes are either static or moving so slowly that the inertia terms are much smaller than other terms in the balance equations for linear and angular momentum. The following chapter contains a presentation of the dynamic theory of poroelasticity, the theory that deals with

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_8,

nonzero inertia terms and focuses on wave propagation. These two aspects of the theory of poroelasticity are treated separately because both the physics and the mathematical models are different. These differences stem from the physical differences in the pore fluid behavior relative to the solid porous matrix material. In the quasistatic theory the pore fluid flows through the pores, from pore to pore. In the dynamic theory the wave passing through the porous medium travels at a speed that does generally not permit the time for the pore fluid to flow from pore to pore. A consequence of this difference in the pore fluid behavior leads to two theories that have very different characteristics.

Poroelastic theory is a useful model in many geological and biological materials because almost all of these materials have an interstitial fluid in their pores. In biological tissues the interstitial fluid has many functions, one important function being to transport nutrients from the vasculature to the cells in the tissue and to transport waste products away. In some tissues the pore fluid pressure creates a turgor or osmotic pressure that stiffens a soft tissue structure and in other tissues it is part of the intercellular communication system. In tissues, the quasistatic deformation of the porous medium has a significant effect on the movement of pore fluid, although the fluid pressure and fluid movement generally have only a small effect on the deformation of the porous medium; the porous medium deformation generally pushes the fluid around, not vice-versa. The pore fluid pressure stays relatively low in tissues because a higher pressure could collapse blood vessels and render the tissue ischemic. However the pore fluid pressure does not have to stay low in tissues like articular cartilage that are avascular. One could speculate that the reason articular cartilage is avascular is to avoid the collapse of blood vessels. Many of these effects have been, or will be, modeled using poroelasticity theory.

Quasistatic fluid movement in rocks and soils has many features of importance to human populations. Water supply is one of the most important, removing oil and gas from the ground is another. High water content can make soil masses unstable; previously stable slopes may be caused to flow over human developments. Earthquakes may cause fluid-saturated soil masses to liquefy and the buildings situated upon them to sink into the soil mass. The theory of poroelasticity has its origins in addressing the problem of buildings settling or sinking into water saturated soil masses due to their own weight, not the liquefaction of the soil mass.

The quasistatic poroelastic theory developed in this chapter is a combination and modification of three of the four theories, those for elastic solids, viscous fluids and flow through porous media, developed in Chap. 6. The development of the theory is strongly dependent on the microstructural modeling concepts described in Chap. 7. In fact much of the material that has preceded this chapter has been laying the foundation for this chapter. Laying this foundation began with the discussion of the Terzaghi and Darcy lumped parameter models in Chap. 1. Terzaghi was one of the first engineers to address the problem of buildings settling or sinking into water saturated soil masses due to their own weight. Darcy investigated the flow of water through sand layers as a factor in the design of fountains in his hometown of Dijon.

Two cases associated with the effects of fluid pressure are considered for the purpose of measuring the elastic constants in a poroelastic material, the drained and the undrained. In the *drained* case the fluid pressure in the pores is zero; draining all the pores before the test or executing the test very slowly achieves this because the fluid in the pores will drain from the negligible fluid pressure associated with slow fluid movement. In a porous medium the pores are assumed to be connected; there are no unconnected pores that prevent the flow of fluid through them. In the *undrained* case the pores that permit fluid to exit the test specimen are sealed; thus pressure will build in the specimen when other forces load it, but the pressure cannot cause the fluid to move out of the specimen. Paraphrasing the opening quote of this chapter, the representative volume element (RVE) for a porous medium (Fig. 4.2) is considered as a small cube (Fig. 4.1) that is large enough, compared to the size of the largest pores, that it may be treated as homogeneous, and at the same time small enough, compared to the scale of the macroscopic phenomena which are of interest, that it may be considered as infinitesimal in the mathematical treatment. The creator of the poroelastic theory advanced this description of the RVE for a saturated porous medium before the terminology RVE came into wide usage. Maurice Anthony Biot (1905–1985), a Belgian-American engineer who made many theoretical contributions to mechanics, developed poroelastic theory. It is his 1941 paper (Biot 1941) that is the basis of the isotropic form of the theory described here.

There are three sets of elastic constants employed in this poroelasticity theory, the drained, $\hat{\mathbf{S}}^{\mathrm{d}}$, the undrained, $\hat{\mathbf{S}}^{\mathrm{u}}$, and those of the matrix material, $\hat{\mathbf{S}}^{\mathrm{m}}$. The RVE associated with the large box in Fig. 4.2 is used for the determination of the drained and the undrained elastic constants while the RVE associated with the smallest box is used for the characterization of the matrix elastic constants. Unlike the large box RVE, the small box RVE contains no pores. The elastic compliance matrices $\hat{\mathbf{S}}^{\mathrm{u}}$, $\hat{\mathbf{S}}^{\mathrm{d}}$, and $\hat{\mathbf{S}}^{\mathrm{m}}$ for these materials have a similar representation:

$$\hat{\mathbf{S}}^{\mathrm{x}} = \begin{bmatrix} \hat{S}^x_{11} & \hat{S}^x_{12} & \hat{S}^x_{13} & \hat{S}^x_{14} & \hat{S}^x_{15} & \hat{S}^x_{16} \\ \hat{S}^x_{12} & \hat{S}^x_{22} & \hat{S}^x_{23} & \hat{S}^x_{24} & \hat{S}^x_{25} & \hat{S}^x_{26} \\ \hat{S}^x_{13} & \hat{S}^x_{23} & \hat{S}^x_{33} & \hat{S}^x_{34} & \hat{S}^x_{35} & \hat{S}^x_{36} \\ \hat{S}^x_{14} & \hat{S}^x_{24} & \hat{S}^x_{34} & \hat{S}^x_{44} & \hat{S}^x_{45} & \hat{S}^x_{46} \\ \hat{S}^x_{15} & \hat{S}^x_{25} & \hat{S}^x_{35} & \hat{S}^x_{45} & \hat{S}^x_{55} & \hat{S}^x_{56} \\ \hat{S}^x_{16} & \hat{S}^x_{26} & \hat{S}^x_{36} & \hat{S}^x_{46} & \hat{S}^x_{56} & \hat{S}^x_{66} \end{bmatrix},$$

where $x = d$, u, or m stands for drained, undrained, and matrix, respectively. The special forms of $\hat{\mathbf{S}}^{\mathrm{x}}$ associated with particular elastic symmetries are listed in Tables 4.4 and 4.5.

There are seven scalar stress variables and seven scalar strain variables in poroelasticity. The seven scalar stress variables are the six components of the stress tensor $\mathbf{T}$ and fluid pressure p in the pores. The seven scalar strain variables are the components of the strain tensor $\mathbf{E}$ and the variation in fluid content ζ, a dimensionless measure of the fluid mass per unit volume of the porous material. The variation

in fluid content ζ may be changed in two ways, either the density of the fluid may change, or the porosity ϕ may change. This set of variables may be viewed as the conjugate pairs of stress measures ($\mathbf{T}$, p) and strain measures ($\mathbf{E}$, ζ) appearing in the following form in an expression for work done on the poroelastic medium: $\mathrm{d}W = \mathbf{T}$: $\mathrm{d}\mathbf{E} + p\ \mathrm{d}\zeta$. The linear stress–strain-pore pressure constitutive relation $\mathbf{E} = \mathbf{E}(\mathbf{T}, p)$, consisting of six scalar equations, is described in Sect. 8.2 and the single-scalar-linear equation fluid content-stress-pore pressure constitutive relation $\zeta = \zeta\ (\mathbf{T}, p)$ is described in Sect. 8.3. Thus the seven scalar stress variables will be linearly related to the seven scalar strain variables. The poroelastic theory considered here is fully saturated, which means that the volume fraction of fluid is equal to the porosity, ϕ, of the solid matrix. Darcy's law, which relates the gradient of the pore pressure to the mass flow rate, is the subject of Sect. 8.4. The special form of poroelasticity theory in the case when both the matrix material and the pore fluid are considered to be incompressible is developed in Sect. 8.5. Formulas for the undrained elastic coefficients $\hat{\mathbf{S}}^{\mathrm{u}}$ as functions of ϕ, the bulk modulus of the pore fluid, K_{f}, and drained and matrix elastic constants, $\hat{\mathbf{S}}^{\mathrm{d}}$ and $\hat{\mathbf{S}}^{\mathrm{m}}$, respectively, $\hat{\mathbf{S}}^{\mathrm{u}} = \hat{\mathbf{S}}^{\mathrm{u}}(\hat{\mathbf{S}}^{\mathrm{d}}, \hat{\mathbf{S}}^{\mathrm{m}}, \phi, K_{\mathrm{f}})$ are developed in Sect. 8.6. The objective of Sects. 8.2, 8.3, 8.4, 8.5, 8.6, and 8.7 is to develop the background for the basic system of equations for poroelasticity recorded in Sects. 8.8 and 8.9. If the reader would like to have a preview of where the Sects. 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, and 8.8 are leading, they can peruse Sect. 8.9 and observe how the various elements in Sects. 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, and 8.8 are combined with the conservation of mass (3.6) and the conservation of momentum (3.37) to form poroelasticity theory. Examples of the solutions to poroelastic problems are given in Sect. 8.10 and a summary of the literature, with references, appears in Sect. 8.11.

8.2 The Stress–Strain-Pore Pressure Constitutive Relation

The basic hypothesis is that the average strain $\hat{\mathbf{E}}$ in the RVE of the saturated porous medium is related, not only to the average stress $\hat{\mathbf{T}}$ in the RVE, but also to the fluid pressure p in the fluid-filled pores. Thus the stress–strain-pore pressure constitutive relation for a saturated porous medium linearly relates the strain $\hat{\mathbf{E}}$ in the saturated porous medium not only to the stress $\hat{\mathbf{T}}$, but also to the fluid pressure p in the fluid-filled pores; this is expressed as the strain–stress-pore pressure relation

$$\hat{\mathbf{E}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}} + \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} p \tag{8.1}$$

or the stress-pore pressure strain relation

$$\hat{\mathbf{T}} + \hat{\mathbf{A}} p = \hat{\mathbf{C}}^{\mathrm{d}} \cdot \hat{\mathbf{E}}, \tag{8.2}$$

where $\hat{\mathbf{S}}^{\mathrm{d}}$ represents the drained anisotropic compliance elastic constants of the saturated porous medium and $\hat{\mathbf{C}}^{\mathrm{d}}$ is its reciprocal, the drained anisotropic elasticity tensor. This constitutive equation is a modification of the elastic strain-stress relation

(4.12H) to include the effect of the pore pressure p. The six-dimensional vector (three-dimensional symmetric second order tensor) $\hat{\mathbf{A}}$ is called the Biot effective stress coefficient tensor. When $p = 0$ the stress–strain-pore pressure relation (8.1) coincides with the elastic strain-stress relation (5.12H). The porous elastic material is treated as a composite of an elastic solid and a pore fluid. The Biot effective stress coefficient vector $\hat{\mathbf{A}}$ is related not only to the drained effective elastic constants of porous matrix material $\hat{\mathbf{S}}^{\mathrm{d}}$, but also to the elastic constants of the solid matrix material $\hat{\mathbf{S}}^{\mathrm{m}}$. The solid matrix material elastic constants of the porous material are based on an RVE that is so small that it contains none of the pores (Fig. 4.2). The Biot effective stress coefficient tensor $\hat{\mathbf{A}}$ is related to the difference between effective drained elastic constants $\hat{\mathbf{S}}^{\mathrm{d}}$ and the solid matrix material elastic compliance tensor $\hat{\mathbf{S}}^{\mathrm{m}}$ by the formula

$$\hat{\mathbf{A}} = (\hat{\mathbf{1}} - \hat{\mathbf{C}}^{\mathrm{d}} \cdot \hat{\mathbf{S}}^{\mathrm{m}}) \cdot \hat{\mathbf{U}}, \tag{8.3}$$

where $\hat{\mathbf{U}} = [1, 1, 1, \ 0, \ 0, 0]^{\mathrm{T}}$ is the six-dimensional vector representation of the three-dimensional unit tensor $\mathbf{1}$. A derivation of (8.3) is given at the end of this section. Note that the symbol $\hat{\mathbf{U}}$ is distinct from the unit tensor in six dimensions that is denoted by $\hat{\mathbf{1}}$; $\hat{\mathbf{U}}$ is described in further detail in Sect. A.11, just before (A.165). The components of $\hat{\mathbf{A}}$ depend upon both the matrix and drained elastic constants. The assumption concerning the symmetry of $\hat{\mathbf{S}}^{\mathrm{d}}$ is interesting and complex. If the symmetry of $\hat{\mathbf{S}}^{\mathrm{d}}$ is less than transversely isotropic and/or its axis of symmetry is not coincident with the transversely isotropic axis of symmetry of $\hat{\mathbf{C}}^{\mathrm{d}}$, then the 6D vector $\hat{\mathbf{C}}^{\mathrm{d}}$ in expression (8.3) has, in general, six nonzero components and the solution to problems is more complicated. However if both $\hat{\mathbf{C}}^{\mathrm{d}}$ and $\hat{\mathbf{S}}^{\mathrm{d}}$ are transversely isotropic with respect to a common axis, then

$$\hat{\mathbf{A}} = [\hat{A}_1, \hat{A}_1, \hat{A}_3, 0, 0, 0]^{\mathrm{T}}, \tag{8.4}$$

where

$$\hat{A}_1 = 1 - (\hat{C}^{\mathrm{d}}_{11} + \hat{C}^{\mathrm{d}}_{12})(\hat{S}^{\mathrm{m}}_{11} + \hat{S}^{\mathrm{m}}_{12} + \hat{S}^{\mathrm{m}}_{13}) - \hat{C}^{\mathrm{d}}_{13}(2\hat{S}^{\mathrm{m}}_{13} + \hat{S}^{\mathrm{m}}_{33}),$$
$$\hat{A}_3 = 1 - 2\hat{C}^{\mathrm{d}}_{13}(\hat{S}^{\mathrm{m}}_{11} + \hat{S}^{\mathrm{m}}_{12} + \hat{S}^{\mathrm{m}}_{13}) - \hat{C}^{\mathrm{d}}_{33}(2\hat{S}^{\mathrm{m}}_{13} + \hat{S}^{\mathrm{m}}_{33}).$$

In the case when both $\hat{\mathbf{C}}^{\mathrm{d}}$ and $\hat{\mathbf{S}}^{\mathrm{d}}$ are isotropic, it follows that $\hat{\mathbf{A}}$ is given by $\alpha\hat{\mathbf{U}}$ ($\alpha\mathbf{1}$ in 3D),

$$\hat{\mathbf{A}} = \alpha\hat{\mathbf{U}} \quad \text{where} \quad \alpha = [1 - (\mathrm{K}^{\mathrm{d}}/\mathrm{K}^{\mathrm{m}})], \tag{8.5}$$

where α is called the isotropic effective stress coefficient. In most situations there appears to be little disadvantage in assuming that $\hat{\mathbf{S}}^{\mathrm{d}}$ is isotropic. The assumption does not mean that the real matrix material is actually isotropic, it only means that there is little error in assuming that it is isotropic because the principal determinant of

the symmetry of the drained elastic constants, $\hat{\mathbf{C}}^{\mathrm{d}}$, is the arrangement of pores; the symmetry of the material surrounding the pores, $\hat{\mathbf{S}}^{\mathrm{m}}$, has only a minor effect. These low consequences of the isotropy assumption for the symmetry of $\hat{\mathbf{S}}^{\mathrm{m}}$ have been discussed by several authors (Shafiro and Kachanov 1997; Kachanov 1999; Sevostianov and Kachanov 2001; Cowin 2004). The *Biot effective stress coefficient tensor* $\hat{\mathbf{A}}$ is so named because it is employed in the definition of the effective stress $\hat{\mathbf{T}}^{\mathrm{eff}}$:

$$\hat{\mathbf{T}}^{\mathrm{eff}} = \hat{\mathbf{T}} + \hat{\mathbf{A}}p. \tag{8.6}$$

This definition of effective stress reduces the stress–strain-pressure relation (8.1) to the same form as (4.12H), thus

$$\hat{\mathbf{E}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}}^{\mathrm{eff}}. \tag{8.7}$$

The advantage of the representation (8.7) is that the fluid-saturated porous material may be thought of as an ordinary elastic material, but one subjected to the "effective stress" $\hat{\mathbf{T}}^{\mathrm{eff}}$ rather than an (ordinary) stress $\hat{\mathbf{T}}$.

The matrix elastic compliance tensor $\hat{\mathbf{S}}^{\mathrm{m}}$ may be evaluated from knowledge of the drained elastic compliance tensor $\hat{\mathbf{S}}^{\mathrm{d}}$ using composite or effective medium theory described in Chap. 7. For example, if the matrix material is isotropic and the pores are dilute and spherical in shape, then the drained elastic material is isotropic and the bulk and shear moduli, K^{d} and G^{d}, are related to the matrix bulk and shear moduli, K^{m} and G^{m}, and Poisson's ratio ν^{m} by

$$K^{\mathrm{d}} = K^{\mathrm{m}} - \frac{\phi K^{\mathrm{m}}}{1 - K^{\mathrm{m}}/(K^{\mathrm{m}} + (4/3)G^{\mathrm{m}})}, \quad \frac{G^{\mathrm{d}}}{G^{\mathrm{m}}} = 1 - \frac{15(1-\nu^{\mathrm{m}})\phi}{7 - 5\nu^{\mathrm{m}}}, \tag{8.8}$$

where ϕ is the porosity associated with the spherical pores. Problem 8.2.2 at the end of this section derives the expressions (8.8) from the formulas (7.16). If the porosity ϕ and the drained constants K^{d} and G^{d} are known, the formulas (8.8) may be used to determine the matrix bulk and shear moduli, K^{m} and G^{m}, recalling that for an isotropic material the Poisson's ratio ν^{m} is related to K^{m} and G^{m} by $\nu^{\mathrm{m}} = (3K^{\mathrm{m}} - 2G^{\mathrm{m}})/(6K^{\mathrm{m}} + 2G^{\mathrm{m}})$, see Table 6.2. As an example, if $K^{\mathrm{d}} = 11.92$ GPa, $G^{\mathrm{d}} = 4.98$ GPa and $\phi = 0.5$, then $K^{\mathrm{m}} = 16.9$ GPa and $G^{\mathrm{m}} = 5.48$ GPa.

The final consideration in this section is the derivation of the formula (8.3) for the *Biot effective stress coefficient tensor* $\hat{\mathbf{A}}$. The material in this paragraph follows the derivation of the formula by Carroll (1979). In his proof, which generalized the elegant proof of Nur and Byerlee (1971) from the isotropic case to the anisotropic case, Carroll (1979) begins by noting that the response of the fluid-saturated porous material may be related to that of the drained material by considering the loading

$$\mathbf{t} = \mathbf{T} \cdot \mathbf{n} \text{ on } \mathrm{O_o}, \quad \mathbf{t} = -p\mathbf{n} \text{ on } \mathrm{O_p}, \text{ loading} \tag{8.9}$$

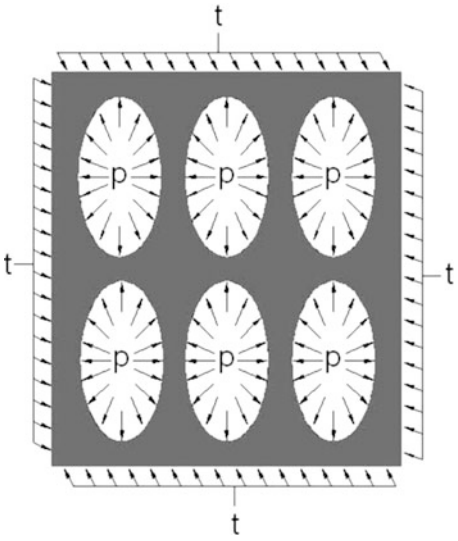


Fig. 8.1 Cartoon of the total loading for a cube of material (only a cross-section is visible) representing a mechanically loaded portion of a saturated anisotropic compressible poroelastic medium. The pores in this porous material are represented by ellipsoids which appear as ellipses in the cross-section; however pore shape and pore connectivity is unrestricted. The pressure on the walls of the ellipse is indicated by the *arrows* perpendicular to the walls, $\mathbf{t} = -p\mathbf{n}$ on O_p, where O_p represents the boundaries of the pores. The tractions on the exterior boundary, $\mathbf{t} = \mathbf{T} \cdot \mathbf{n}$ on O_o, where O_o represents the outer boundary of the porous medium, are indicated by the *arrows* slanted with respect to the lines forming the boundary of the square and acting on that boundary

where O_o and O_p represent the outer boundary of the porous medium and the pore boundary, respectively. This loading is illustrated for a cube of material (only a cross-section is visible) in Fig. 8.1. The pores in this porous material are represented by ellipsoids which appear as ellipses in the cross-section of Fig. 8.1. The pressure on the walls of the ellipse is indicated by the arrows perpendicular to the walls, $\mathbf{t} = -p\mathbf{n}$ on O_p.

The tractions on the exterior boundary, $\mathbf{t} = \mathbf{T} \cdot \mathbf{n}$ on O_o, are indicated by the arrows slanted with respect to the lines forming the boundary of the square and acting on that boundary.

The first key to this proof is to treat the loading (8.9) as the superposition of two separate loadings:

$$\mathbf{t} = -p\mathbf{n} \text{ on } O_o, \quad \mathbf{t} = -p\mathbf{n} \text{ on } O_p, \text{ loading } (8.10) \tag{8.10}$$

and

$$\mathbf{t} = \mathbf{T} \cdot \mathbf{n} + p\mathbf{n} \text{ on } O_o, \quad \mathbf{t} = 0 \text{ on } O_p, \text{ loading } (8.11). \tag{8.11}$$

The loading (8.10) is illustrated in Fig. 8.2a; this loading creates a uniform hydrostatic pressure p in the matrix material and, consequently, a uniform strain if

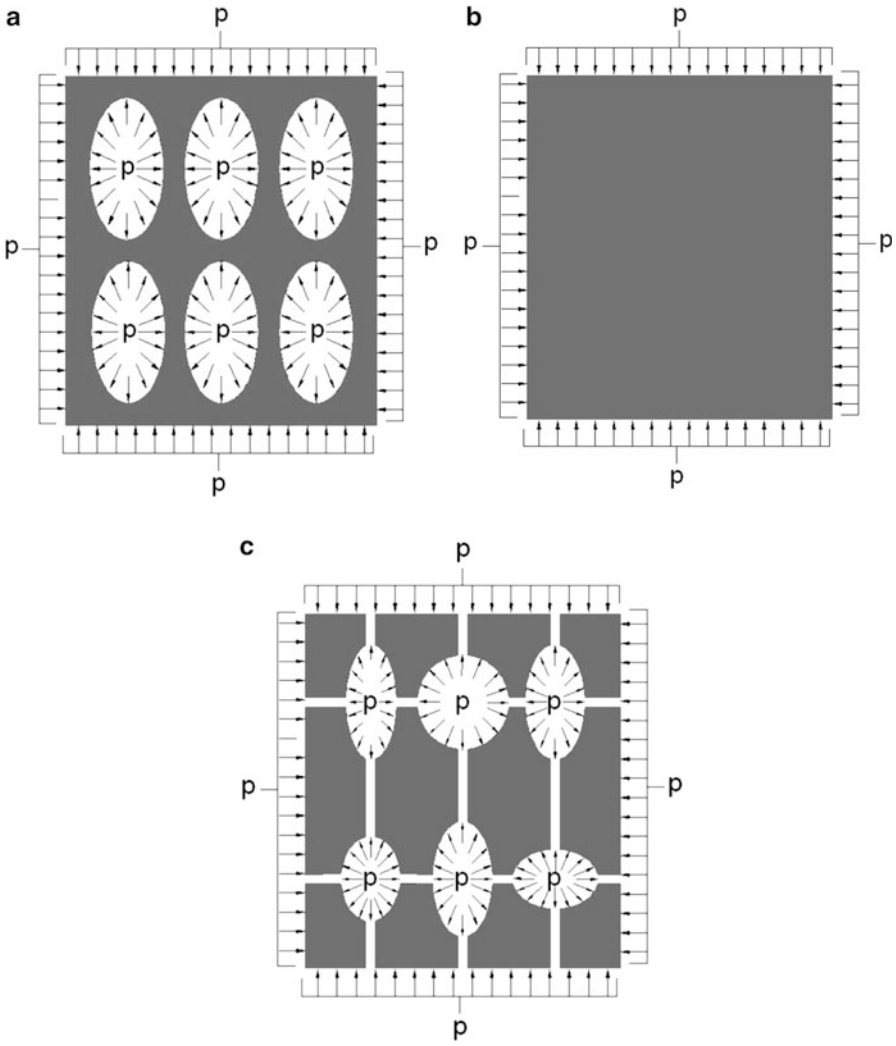


Fig. 8.2 (**a**) Cartoon of the loading (11) for a cube of material (only a cross-section is visible) representing a mechanically loaded portion of a saturated anisotropic compressible poroelastic medium. This loading creates a uniform hydrostatic pressure p in the matrix material and, consequently, a uniform strain if the porous material is homogeneous. The strain in the porous material is then the same as that in the matrix material; in effect, uniform straining of the matrix material results in the same straining of the pore space. Nur and Byerlee (1971) illustrate this clearly by pointing out that the loading (8.10) of the solid is achieved by filling the pores with the matrix material, as illustrated in (**b**). (**b**) Cartoon of the loading (8.10) for a cube of material (only a cross-section is visible) equivalent to the cartoon of the loading (8.10) in (**a**). The fact that (**a**)–(**c**) are mechanically equivalent is one of the keys to understanding this proof. The loading (8.10) creates a uniform hydrostatic pressure p in the matrix material and, consequently, a uniform strain if the porous material is homogeneous. The strain in the porous material is then the same as that in the matrix material; in effect, uniform straining of the matrix material results in the same straining

the porous material is homogeneous. The strain in the porous material is then the same as that in the matrix material; in effect, uniform straining of the matrix material results in the same straining of the pore space. Nur and Byerlee (1971) illustrate this clearly by pointing out that the loading (8.10) of the solid is achieved by filling the pores with the matrix material. This is illustrated in Fig. 8.2b. Filling the pores with the matrix material has created in the cube a uniform material in which the pressure everywhere is the same. Thus there is no difference in the pressure and strain fields in Fig. 8.2a from those in Fig. 8.2b. Both Fig. 8.2a, b have the same pressure p acting everywhere and the same homogeneous strain. With a little reflection the reader will see that the conclusion that has just been drawn is independent of the shape, size, and connectivity between the pores. Thus, rather than ellipsoids, the pores of Fig. 8.2a could all be of arbitrary shape and size and they could all be connected as shown in Fig. 8.2c, but the same pressure p acts everywhere as well as the same homogeneous strain, just as in all three Fig. 8.2a–c. The resulting strain $\hat{\mathbf{E}}^{(8.10)}$ for loading (8.10) in the homogeneous matrix material is then uniform, and it is given by

$$\hat{\mathbf{E}}^{(8.10)} = -p\hat{\mathbf{S}}^{\mathrm{m}} \cdot \hat{\mathbf{U}}, \tag{8.12}$$

where the tensor equation $\mathbf{T} = -p\mathbf{1}$ in 3D has been written in 6D as $\hat{\mathbf{T}} = -p\hat{\mathbf{U}}$.

Since the pore pressure in loading (8.11) is zero, the exterior surface loading ($\mathbf{t} = \mathbf{T} \cdot \mathbf{n} + \mathbf{pn}$ on O_{o}) may be considered as being applied to a drained elastic material, Fig. 8.3. The resulting strain $\hat{\mathbf{E}}^{(8.11)}$ is given by

$$\hat{\mathbf{E}}^{(8.11)} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot (\hat{\mathbf{T}} + p\hat{\mathbf{U}}) \tag{8.13}$$

Fig. 8.2 (continued) of the pore space. Nur and Byerlee (1971) illustrate this clearly by pointing out that the loading (8.10) of the solid is achieved by filling the pores in (**a**) with the matrix material, as illustrated in this figure. It is also important to note that this conclusion is independent of the shape of the pores and their connectivity with each other and with the external environment as long as the external environment is at a pressure p; see (**c**). (**c**) Cartoon of the loading (8.10) for a cube of material (only a cross-section is visible) equivalent to the cartoon of the loading (8.10) in (**a**)–(**c**). The fact that (**a**)–(**c**) are mechanically equivalent is one of the keys to understanding this proof. The loading (8.10) creates a uniform hydrostatic pressure p in the matrix material and, consequently, a uniform strain if the porous material is homogeneous. The strain in the porous material is then the same as that in the matrix material; in effect, uniform straining of the matrix material results in the same straining of the pore space. Nur and Byerlee (1971) illustrate this clearly by pointing out that the loading (8.10) of the solid is achieved by filling the pores in (**a**) with the matrix material, as illustrated in (**b**). As this figure illustrates, it is also important to note that this conclusion is independent of the shape of the pores and their connectivity with each other and with the external environment as long as the external environment is at a pressure p

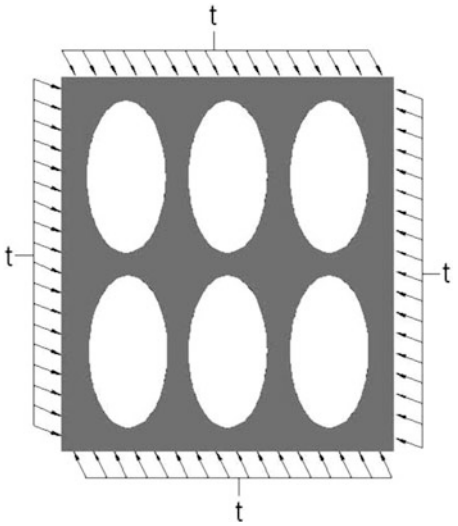


Fig. 8.3 Cartoon of the loading (8.11) for a cube of material (only a cross-section is visible). Observe that the loadings (8.10) and (8.11) may be superposed to produce the total loading (8.9), illustrated in Fig. 8.1. This loading consists of a surface traction that is the sum of the actual surface traction $\mathbf{t} = \mathbf{T} \cdot \mathbf{n}$ and the superposed pore pressure, $\mathbf{t} = \mathbf{T} \cdot \mathbf{n} + p\mathbf{n}$ on O_o, the external surface. The surface traction on the pores is zero

and the total strain due to the loadings (8.10) plus (8.11) is the total strain due to the loading (8.9)

$$\hat{\mathbf{E}}^{(8.9)} = \hat{\mathbf{E}}^{(8.10)} + \hat{\mathbf{E}}^{(8.11)} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}} + p(\hat{\mathbf{S}}^{\mathrm{d}} - \hat{\mathbf{S}}^{\mathrm{m}}) \cdot \hat{\mathbf{U}}. \tag{8.14}$$

Figures 8.1, 8.2a, and 8.3 correspond to the loadings (8.9), (8.10), and (8.11) respectively, and the strains $\hat{\mathbf{E}}^{(8.9)}$, $\hat{\mathbf{E}}^{(8.10)}$, and $\hat{\mathbf{E}}^{(8.11)}$ have the same respective correspondence. It is easy to see that (8.14) may be rewritten as

$$\hat{\mathbf{E}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot (\hat{\mathbf{T}} + (\hat{\mathbf{1}} - \hat{\mathbf{C}}^{\mathrm{d}} \cdot \hat{\mathbf{S}}^{\mathrm{m}}) \cdot \hat{\mathbf{U}}p), \tag{8.15}$$

where $\hat{\mathbf{E}} = \hat{\mathbf{E}}^{(8.9)}$ is the total strain. Comparing this result with (8.1) it may be seen that the *Biot effective stress coefficient tensor* $\hat{\mathbf{A}}$ is given by (8.3). Note that this proof is based on superposition that is a characteristic of linear systems and thus applies to all the considerations of linear compressible poroelasticity, isotropic or anisotropic, but fails when the deformations are no longer infinitesimal.

Problems

8.2.1. Show that the representation (8.4) for $\hat{\mathbf{A}}$ in the case when both $\hat{\mathbf{C}}^{\mathrm{d}}$ and $\hat{\mathbf{S}}^{\mathrm{d}}$ are transversely isotropic with respect to a common axis reduces to (8.5) when both $\hat{\mathbf{C}}^{\mathrm{d}}$ and $\hat{\mathbf{S}}^{\mathrm{d}}$ are isotropic.

8.2.2. Derive the expressions (8.8) for K^{d} and G^{d} in terms of the volume fraction ϕ, matrix bulk and shear moduli, K^{m} and G^{m}, and Poisson's ratio ν^{m} from the formulas (7.20).

8.3 The Fluid Content-Stress-Pore Pressure Constitutive Relation

The basic field variables for poroelasticity are the total stress $\hat{\mathbf{T}}$, the pore pressure p, the strain in the solid matrix $\hat{\mathbf{E}}$, and the variation in (dimensionless) fluid content ζ. The variation in fluid content ζ is the variation of the fluid mass per unit volume of the porous material due to diffusive fluid mass transport. In terms of the volume fraction of fluid ϕ (ϕ is also the porosity in a fluid-saturated porous medium), the fluid density ρ_{f}, and initial values of these two quantities ρ_{fo} and ϕ_{o} respectively, are defined as

$$\zeta \equiv \frac{\rho_{\mathrm{f}}\phi - \rho_{\mathrm{fo}}\phi_{\mathrm{o}}}{\rho_{\mathrm{fo}}} = \frac{\rho_{\mathrm{f}}}{\rho_{\mathrm{fo}}}\phi - \phi_{\mathrm{o}}. \tag{8.16}$$

It is important to note that the variation in fluid content ζ may be changed in two ways, either the density of the fluid ρ_{f} may change from its reference value of ρ_{fo}, or the porosity ϕ may change from its reference value of ϕ_{o}. This set of variables may be viewed as the conjugate pairs of stress measures ($\hat{\mathbf{T}}$, p) and strain measures ($\hat{\mathbf{E}}$, ζ) appearing in the following form in an expression for work done on the poroelastic medium: $\mathrm{dW} = \hat{\mathbf{T}} \cdot \mathrm{d}\hat{\mathbf{E}} + \mathrm{p}\,\mathrm{d}\zeta$. Thus the pressure p is viewed as another component of stress, and the variation in fluid content ζ is viewed as another component of strain; the one conjugate to pressure is the expression for work. It follows that the variation in fluid content ζ is linearly related to both the stress $\hat{\mathbf{T}}$ and the pore pressure p,

$$\zeta = \hat{\mathbf{A}} \cdot \hat{\mathcal{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}} + C^{\mathrm{d}}_{\mathrm{eff}} p \quad \text{where } \mathrm{C}^{\mathrm{d}}_{\mathrm{eff}} = \frac{1}{\mathrm{K}^{\mathrm{d}}_{\mathrm{Reff}}} - \frac{1}{\mathrm{K}^{\mathrm{m}}_{\mathrm{Reff}}} + \phi\left(\frac{1}{\mathrm{K}^{\mathrm{f}}} - \frac{1}{\mathrm{K}^{\mathrm{m}}_{\mathrm{Reff}}}\right) \tag{8.17}$$

or related, using (8.2) to both the strain $\hat{\mathbf{E}}$ and the pressure p, by

$$\zeta = \hat{\mathbf{A}} \cdot \hat{\mathbf{E}} + \Lambda p, \quad \Lambda = C^{\mathrm{d}}_{\mathrm{eff}} - \hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} \tag{8.18}$$

and where the various super- and subscripted K's are different bulk moduli; for example, K^{f} is the bulk modulus of the pore fluid. Before introducing formulas for the other two bulk moduli in (8.17), note that the isotropic elastic compliance tensor $\hat{\mathbf{S}}$ twice contracted with $\hat{\mathbf{U}}$, the six-dimensional vector representation of the three-dimensional unit tensor $\mathbf{1}$, $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{U}}$ (see (A.166)), is equal to $3(1 - 2\nu)^{-1}$, which, in turn, is the reciprocal of the bulk modulus (see Table 7.2),

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{U}} = \frac{3(1-2\nu)}{E} \equiv \frac{1}{K}. \tag{8.19}$$

Hill (1952) showed that $K_{\text{Reff}} = (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{U}})^{-1}$, where $\hat{\mathbf{S}}$ is now the anisotropic elastic compliance tensor, is the Reuss lower bound on the effective (isotropic) bulk modulus of the anisotropic elastic material $\hat{\mathbf{S}}$ and that the Voigt effective bulk modulus of an anisotropic elastic material, $K_{\text{Veff}} = (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{-1} \cdot \hat{\mathbf{U}})/9$, is the upper bound,

$$(\hat{\mathbf{U}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{U}})^{-1} = K_{\text{Reff}} \leq K_{\text{eff}} \leq K_{\text{Veff}} = (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{-1} \cdot \hat{\mathbf{U}})/9. \tag{8.20}$$

In the case of isotropy the two bounds coincide with the isotropic bulk modulus, K, thus:

$$\begin{aligned} \frac{1}{K_{\text{Reff}}} = \frac{1}{K_{\text{R}}} = \frac{1}{K_{\text{eff}}} = \frac{1}{K_{\text{Veff}}} = \frac{1}{K_{\text{V}}} = (\hat{\mathbf{U}} \cdot \hat{S} \cdot \hat{\mathbf{U}})^{-1} = \frac{\hat{\mathbf{U}} \cdot \hat{S}^{-1} \cdot \hat{\mathbf{U}}}{9} \\ = \frac{3(1-2\nu)}{E} \equiv \frac{1}{K}, \end{aligned} \tag{8.21}$$

where E is the isotropic Young's modulus and ν is the Poisson's ratio and where the subscript eff is no longer applicable because these K's are the actual bulk moduli rather than the effective bulk moduli. The Reuss effective bulk modulus of an anisotropic elastic material K_{Reff} occurs naturally in anisotropic poroelastic theory as shown, but not noted, by Thompson and Willis (1991). In analogy with this result (8.19), $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{U}}$ is defined as the inverse of the effective bulk modulus ($1/K$) when the material is not isotropic. Thus, for the orthotropic drained elastic compliance tensor $\hat{\mathbf{S}}^{\text{d}}$, and the matrix of the orthotropic elastic compliance tensor $\hat{\mathbf{S}}^{\text{d}}$, the following definitions, which were employed in (8.17), are introduced:

$$\frac{1}{K^{\text{d}}_{\text{eff}}} = \hat{\mathbf{U}} \cdot \hat{S}^{\text{d}} \cdot \hat{\mathbf{U}} = \frac{1}{E^{\text{d}}_1} + \frac{1}{E^{\text{d}}_2} + \frac{1}{E^{\text{d}}_3} - \frac{2\nu^{\text{d}}_{23}}{E^{\text{d}}_2} - \frac{2\nu^{\text{d}}_{31}}{E^{\text{d}}_3} - \frac{2\nu^{\text{d}}_{12}}{E^{\text{d}}_1}, \tag{8.22}$$

$$\frac{1}{K^{\text{m}}_{\text{eff}}} = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\text{d}} \cdot \hat{\mathbf{U}} = \frac{1}{E^{\text{m}}_1} + \frac{1}{E^{\text{m}}_2} + \frac{1}{E^{\text{m}}_3} - \frac{2\nu^{\text{m}}_{23}}{E^{\text{m}}_2} - \frac{2\nu^{\text{m}}_{31}}{E^{\text{m}}_3} - \frac{2\nu^{\text{m}}_{12}}{E^{\text{m}}_1}. \tag{8.23}$$

Using these notations the formula for Λ, (8.18), may be recast in terms of the effective bulk moduli. Using (8.3) note that,

$$\hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\text{d}} \cdot \hat{\mathbf{A}} = \frac{1}{K^{\text{d}}_{\text{Reff}}} - \frac{2}{K^{\text{m}}_{\text{Reff}}} + \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\text{d}} \cdot \hat{\mathbf{C}}^{\text{d}} \cdot \hat{\mathbf{S}}^{\text{d}} \cdot \hat{\mathbf{U}}, \tag{8.24}$$

where (8.19) and (8.22) have been employed. Substituting this result into (8.18) and employing (8.17) to remove $C^{\text{d}}_{\text{eff}}$, it follows that

$$\Lambda = \frac{1}{K^{\text{m}}_{\text{Reff}}} + \phi\left(\frac{1}{K^{\text{f}}} - \frac{1}{K^{\text{m}}_{\text{Reff}}}\right) - \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\text{d}} \cdot \hat{\mathbf{C}}^{\text{d}} \cdot \hat{\mathbf{S}}^{\text{d}} \cdot \hat{\mathbf{U}}. \tag{8.25}$$

Problems

8.3.1. Prove that $\hat{\mathbf{U}} \cdot \hat{\mathbf{T}} = \text{tr}\,\mathbf{T}$.

8.3.2. Show that the isotropic elastic compliance tensor $\hat{S}$, twice contracted with $\hat{\mathbf{U}}$, the six-dimensional vector representation of the three-dimensional unit tensor $\mathbf{1}$, $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}} \cdot \hat{\mathbf{U}}$, is given by (8.19).

8.3.3. Show that, in the case of isotropy, $\hat{\mathbf{A}}$ is given by $\alpha\hat{\mathbf{U}}$, where α is the isotropic effective stress coefficient given by $\alpha = [1 - (K^{\mathrm{d}}/K^{\mathrm{m}})]$.

8.3.4. In the case of isotropy show that $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{C}}^{\mathrm{d}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}}$ is equal to $K^{\mathrm{d}}_{\mathrm{Reff}}/(K^{\mathrm{m}}_{\mathrm{Reff}})^2$, and thus $\hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}}$ is given by $\hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} = \frac{1}{K^{\mathrm{d}}} - \frac{\alpha}{K^{\mathrm{m}}}$ and $\Lambda = \frac{1}{K^{\mathrm{m}}_{\mathrm{Reff}}} + \phi(\frac{1}{K^{\mathrm{f}}} - \frac{1}{K^{\mathrm{m}}}) - \frac{K^{\mathrm{d}}}{(K^{\mathrm{m}})^2}$.

8.4 Darcy's Law

The constitutive equations of poroelasticity developed thus far are the strain–stress-pore pressure relations (8.1) and the fluid content-stress-pore pressure relation (8.17). The other constitutive relation of poroelasticity is Darcy's law, which relates the fluid mass flow rate, $\rho_{\mathrm{f}}\mathbf{v}$, to the gradient ($\nabla p$) of the pore pressure p,

$$\mathbf{q} = (\phi\rho_{\mathrm{f}}/\rho_{\mathrm{fo}})\mathbf{v} = -\mathbf{H}(p)\nabla p(\mathbf{x}, t), \mathbf{H}(p) = \mathbf{H}^{\mathrm{T}}(p), \quad \text{(4.36D) repeated}$$

where the symmetry in $\mathbf{H}$ has been shown to hold for material symmetries greater than monoclinic (Chap. 4). In this equation ρ_{f} is the fluid density and ρ_{fo} is a reference value of the fluid density. It will be assumed that $\mathbf{H}$ is independent of pore fluid pressure and that $\mathbf{H} = \mathbf{K}/\mu$, where $\mathbf{K}$ is the intrinsic Darcy's law permeability tensor, and μ is the fluid viscosity. The intrinsic permeability tensor $\mathbf{K}$ has units of length squared and is a function of the porous structure only, not the fluid in the pores; thus Darcy's law takes the form

$$\mathbf{q} = (\phi\rho_{\mathrm{f}}/\rho_{\mathrm{fo}})\mathbf{v} = -(1/\mu)\mathbf{K}\nabla p(\mathbf{x}, t), \mathbf{K} = \mathbf{K}^{\mathrm{T}}, \tag{8.26}$$

where it has been shown that the symmetry in $\mathbf{K}$ holds for material symmetries greater than monoclinic and where the volume flux $\mathbf{q}$ has the dimension of velocity because it is the volume flow rate per unit area. In the case of isotropy, Darcy's law is written in the form

$$\mathbf{q} = (\phi\rho_{\mathrm{f}}/\rho_{\mathrm{fo}})\mathbf{v} = -(k/\mu)\nabla p(\mathbf{x}, t). \tag{8.27}$$

Recall from Sect. 1.8 the lumped parameter model for Darcy's law and consider that idea again in the present context. Also, recall from Sect. 7.5 that it was shown

how the Darcy permeability law could be obtained from a model for the flow of a viscous fluid through a rigid porous solid. These ideas, plus the example of the compressed, fluid-saturated sponge described in Sect. 8.1, convey the physical meaning of Darcy's law in three different models.

8.5 Matrix Material and Pore Fluid Incompressibility Constraints

The two incompressibility constituent-specific constraints for poroelasticity are that matrix material and the pore fluid is incompressible. These two incompressibility constraints are kinematic in nature, requiring that both materials experience no volume change at any stress level. Constraints of this type introduce an indeterminate pressure in both the fluid and the matrix material that must be equal in both materials at any location from the requirement of local force equilibrium. Thus the two assumptions are compatible.

The requirement that the fluid be incompressible is implemented by requiring that the reciprocal of the bulk modulus of the fluid tends to zero as the instantaneous density tends to the initial density, $\lim_{\rho_f \to \rho_{fo}} \frac{1}{K^f} \to 0$, or by imposing the condition that the fluid density ρ_f be a constant; thus ρ_f and its initial value ρ_fo are equal. Recall that, in the case of a compressible fluid above, the value of the fluid content ζ can change if the fluid density changes or if the porosity changes. However, in the case of the incompressible fluid since $\rho_\mathrm{f} = \rho_{fo}$, the fluid content ζ (8.16) can change only if the porosity changes,

$$\zeta = \phi - \phi_\mathrm{o}. \tag{8.28}$$

A similar change occurs for the Darcy's law (8.17) when $\rho_\mathrm{f} = \rho_\mathrm{fo}$,

$$\mathbf{q} = \phi \mathbf{v} = -(1/\mu)\mathbf{K}\nabla p(\mathbf{x}, t), \mathbf{K} = \mathbf{K}^\mathrm{T}. \tag{8.29}$$

The requirement that the matrix material be incompressible involves a slightly longer development. Hooke's law for the matrix material is written as

$$\hat{\mathbf{E}}^\mathrm{m} = \hat{\mathbf{S}}^\mathrm{d}\hat{\mathbf{T}}^\mathrm{m} \tag{8.30}$$

and the incompressibility constraint for the matrix material is the requirement that the dilatational strain,

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{E}}^\mathrm{m} = \mathrm{tr}\, \mathbf{E}^\mathrm{m} = E^\mathrm{m}_{kk} = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^\mathrm{d} \cdot \hat{\mathbf{T}}^\mathrm{m} = \hat{\mathbf{T}}^\mathrm{m} \cdot \hat{\mathbf{S}}^\mathrm{d} \cdot \hat{\mathbf{U}} \tag{8.31}$$

vanish for all possible stress states $\hat{\mathbf{T}}^\mathrm{m}$, thus

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{E}}^\mathrm{m} = 0 \Rightarrow \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^\mathrm{d} = \hat{\mathbf{S}}^\mathrm{d} \cdot \hat{\mathbf{U}} = 0. \tag{8.32}$$

The components of the vector $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}}$ are given by

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = \left[\frac{1}{K_1^{\mathrm{m}}}, \frac{1}{K_2^{\mathrm{m}}}, \frac{1}{K_3^{\mathrm{m}}}, \frac{1}{K_4^{\mathrm{m}}}, \frac{1}{K_5^{\mathrm{m}}}, \frac{1}{K_6^{\mathrm{m}}}\right], \tag{8.33a}$$

where

$$\frac{1}{K_\alpha^{\mathrm{m}}} = S_{\alpha 1}^{\mathrm{m}} + S_{\alpha 2}^{\mathrm{m}} + S_{\alpha 3}^{\mathrm{m}} \quad (\alpha = 1, \ldots, 6) \tag{8.33b}$$

in the case of no (triclinic) symmetry, and by

$$\begin{aligned}
\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{m}} &= \left[\frac{1}{K_1^{\mathrm{m}}}, \frac{1}{K_2^{\mathrm{m}}}, \frac{1}{K_3^{\mathrm{m}}}, 0, 0, 0\right], \quad \frac{1}{K_1^{\mathrm{m}}} = \frac{1 - \nu_{12}^{\mathrm{m}} - \nu_{13}^{\mathrm{m}}}{E_1^{\mathrm{m}}}, \\
&\frac{1}{K_2^{\mathrm{m}}} = \frac{1 - \nu_{21}^{\mathrm{m}} - \nu_{23}^{\mathrm{m}}}{E_2^{\mathrm{m}}}, \quad \frac{1}{K_3^{\mathrm{m}}} = \frac{1 - \nu_{13}^{\mathrm{m}} - \nu_{32}^{\mathrm{m}}}{E_3^{\mathrm{m}}}, \\
\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{m}} &= \left[\frac{1}{K_1^{\mathrm{m}}}, \frac{1}{K_1^{\mathrm{m}}}, \frac{1}{K_3^{\mathrm{m}}}, 0, 0, 0\right], \quad \frac{1}{K_3^{\mathrm{m}}} = \frac{1 - 2\nu_{31}^{\mathrm{m}}}{E_3^{\mathrm{m}}}, \quad \frac{1}{K_1^{\mathrm{m}}} = \frac{1}{K_2^{\mathrm{m}}} \\
&= \frac{1 - \nu_{12}^{\mathrm{m}} - \nu_{13}^{\mathrm{m}}}{E_1^{\mathrm{m}}} = \frac{1 - \nu_{21}^{\mathrm{m}} - \nu_{23}^{\mathrm{m}}}{E_2^{\mathrm{m}}}, \\
\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{m}} &= \frac{1}{K^{\mathrm{m}}}[1, 1, 1, 0, 0, 0], \quad \frac{1}{K^{\mathrm{m}}} = \frac{1 - 2\nu^{\mathrm{m}}}{E^{\mathrm{m}}}, \quad \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{m}} \cdot \hat{\mathbf{U}} = \frac{3}{K_1^{\mathrm{m}}} = \frac{1}{K^{\mathrm{m}}} \\
&= \frac{3(1 - 2\nu^{\mathrm{m}})}{E^{\mathrm{m}}}
\end{aligned} \tag{8.34}$$

in the cases of orthotropic, transversely isotropic, and isotropic symmetries, respectively. As may be seen from (8.33) the incompressibility condition $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} = 0$ requires that $\hat{\mathbf{S}}^{\mathrm{d}}$ be singular, Det $\hat{\mathbf{S}}^{\mathrm{d}} = 0$. From (8.32) and (8.34) the incompressibility condition $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} = 0$ is expressed in terms of Poisson's ratios for the orthotropic, transversely isotropic and isotropic symmetries by

$$\nu_{12}^m = \frac{E_1^m E_3^m}{E_2^m (E_1^m + E_3^m)}, \nu_{13}^m = 1 - \frac{E_1^m E_3^m}{E_2^m (E_1^m + E_3^m)}, \nu_{23}^m = \frac{E_1^m}{E_1^m + E_3^m}, \tag{8.35}$$

$$\nu_{12}^{\mathrm{m}} = 1 - \frac{E_1^{\mathrm{m}}}{2E_3^{\mathrm{m}}}, \quad \nu_{13}^{\mathrm{m}} = \frac{E_1^{\mathrm{m}}}{2E_3^{\mathrm{m}}}, \quad \nu_{31}^{\mathrm{m}} = \frac{1}{2} \tag{8.36}$$

and $\nu^{\mathrm{m}} = 1/2$, respectively.

The significance of these incompressibility results is illustrated here by simple geometric considerations. Consider an incompressible isotropic material for which one must have $\nu^{\mathrm{m}} = 1/2$. If a cube of this material with a volume $V_{\mathrm{o}} = a^3$ is extended by a uniform tensile stress in one direction, the length in that extension direction becomes $a + \Delta_{\mathrm{EX}} a$ and the lengths of the other two faces become $a + \Delta a$.

The new volume is then $V = (a + \Delta_{EX}a)(a + \Delta a)^2$, which when higher order terms are neglected, becomes $V = V_o + a^2(\Delta_{EX}a + 2\Delta a)$. It follows that if there is to be no volume change, $V = V_o$, then the two transverse sides actually contract by an amount $\Delta a = -(1/2)\Delta_{EX}a$. Dividing both sides of $\Delta a = -(1/2)\Delta_{EX}a$ by a it follows that the strain in the transverse direction ε is equal to one-half the strain in the tension direction, $\varepsilon = -(1/2)\varepsilon_T$. Recalling the definition of Poisson's ratio, it follows that $\nu^m = 1/2$ for the material in this example.

Only two of three sets of poroelastic constitutive equations described in the previous section are influenced by these two incompressibility constraints. Darcy's law is unchanged because in both cases it is based on the assumption that the movement of the boundaries of the pores is a higher order term that is negligible and thus the law has the same form in the compressible and incompressible cases as it would have in a rigid porous material. The stress–strain-pore pressure and the fluid content-stress-pore pressure are modified in the case of incompressibility from their forms in the compressible case. The stress–strain-pore pressure relation (8.1) for the incompressible case is given by

$$\hat{\mathbf{E}} = \hat{\mathbf{S}}^{\mathrm{d}} \cdot (\hat{\mathbf{T}} + \hat{\mathbf{U}}p) \tag{8.37}$$

for $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = 0$ since the Biot effective stress coefficient tensor $\hat{\mathbf{A}}$ given by (8.3) takes the form

$$\hat{\mathbf{A}} = \hat{\mathbf{U}} \tag{8.38}$$

for $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = 0$. The definition of the effective stress $\hat{\mathbf{T}}^{\mathrm{eff}}$

$$\hat{\mathbf{T}}^{\mathrm{eff}} = \hat{\mathbf{T}} + \hat{\mathbf{U}}p \tag{8.39}$$

changes for the incompressible case and the Hooke's law (8.7) holds in the incompressible case with this revised definition of $\hat{\mathbf{T}}^{\mathrm{eff}}$.

Since the reciprocals of $K^{\mathrm{m}}_{\mathrm{Reff}}$ and K^{f} vanish in the case of incompressibility, the fluid content-stress-pore pressure relation (8.17) is modified in the case of incompressibility to the form

$$\zeta = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}} + \frac{p}{K^{\mathrm{d}}_{\mathrm{Reff}}} \quad \text{where } C^{\mathrm{d}}_{\mathrm{eff}} = \frac{1}{K^{\mathrm{d}}_{\mathrm{Reff}}} = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} \tag{8.40}$$

thus, from (8.25), (8.38), (8.40) and the vanishing of the reciprocals of the bulk moduli of the fluid and the matrix material, it follows that

$$\zeta = \hat{\mathbf{U}} \cdot \hat{\mathbf{E}} \quad \text{and} \quad \Lambda = 0. \tag{8.41}$$

The specialization of (8.32) to the case of incompressibility recovers the obvious consequence of the assumption of material matrix incompressibility,

$$\frac{1}{K^{\mathrm{m}}_{\mathrm{Reff}}} = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} = 0 \tag{8.42}$$

for $\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} = 0$. The pore pressure is given by (8.18) as

$$p = \frac{1}{\Lambda}[\zeta - (\hat{\mathbf{A}} \cdot \hat{\mathbf{E}})] \tag{8.43}$$

and one can observe from the preceding that, for incompressibility,

$$\Lambda \to 0 \quad \text{and} \quad [\zeta - (\hat{\mathbf{A}} \cdot \hat{\mathbf{E}})] \to 0 \tag{8.44}$$

and it follows that the pressure p given by (8.43) becomes indeterminate in the formula (8.43) as the porous medium constituents become incompressible. A Lagrange multiplier is then introduced (Example 6.4.1), thus (8.39) above now applies. The convention that there are two very different meanings associated with the symbol for pore pressure p is maintained here. In the compressible case p is a thermodynamic variable determined by an equation of state that includes the temperature and the specific volume of the fluid as variables, but in the incompressible case p is a Lagrange multiplier whose value is determined by the boundary conditions independent of the temperature and the specific volume of the fluid.

Problems

8.5.1. Show that (8.17) reduces to $\zeta = \hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}} + (\frac{1}{K^{\mathrm{d}}_{\mathrm{Reff}}} - \frac{1}{K^{\mathrm{m}}_{\mathrm{Reff}}})p$ where $C^{\mathrm{d}}_{\mathrm{eff}} = \frac{1}{K^{\mathrm{d}}_{\mathrm{Reff}}} - \frac{1}{K^{\mathrm{m}}_{\mathrm{Reff}}}$ when the bulk moduli of the matrix material and the fluid are equal.

8.5.2. Show that (8.17) reduces to (8.40) when the matrix material and the fluid are assumed to be incompressible.

8.6 The Undrained Elastic Coefficients

If no fluid movement in the poroelastic medium is possible, then the variation in fluid content ζ, that is to say the variation of the fluid volume per unit volume of the porous material due to diffusive fluid mass transport, is zero. In this situation (8.17) may be solved for p; thus the pore pressure is related to the solid stress $\hat{\mathbf{T}}$ by $p = -\hat{\mathbf{B}} \cdot \hat{\mathbf{T}}$, where

$$\hat{\mathbf{B}} = \frac{1}{C^{\mathrm{d}}_{\mathrm{eff}}} \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} = \frac{1}{C^{\mathrm{d}}_{\mathrm{eff}}} (\hat{\mathbf{S}}^{\mathrm{d}} - \hat{\mathbf{S}}^{\mathrm{d}}) \cdot \hat{\mathbf{U}}. \tag{8.45}$$

Skempton (1954) first introduced the components of the tensor $\hat{\mathbf{B}}$, thus it is reasonable to call the tensor $\hat{\mathbf{B}}$ the Skempton compliance difference tensor. In the case of incompressibility, it follows from (8.45), (8.40), and (8.38) that the Skempton compliance difference tensor in the incompressible case is given by

$$\hat{\mathbf{B}} = K^{\mathrm{d}}_{\mathrm{Reff}} \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} \tag{8.46}$$

from which it follows, with use of (8.31), that

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{B}} = K^{\mathrm{d}}_{\mathrm{Reff}} \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} = 1 \tag{8.47}$$

in the incompressible case. For the isotropic compressible case the Skempton compliance difference tensor has the form,

$$\hat{\mathbf{B}} = \frac{S}{3}\hat{\mathbf{U}}, \quad \text{or} \quad \hat{\mathrm{B}}_1 = \hat{\mathrm{B}}_2 = \hat{\mathrm{B}}_3 = \frac{\mathrm{S}}{3}, \quad \hat{\mathrm{B}}_4 = \hat{\mathrm{B}}_5 = \hat{\mathrm{B}}_6 = 0, \tag{8.48}$$

where S is the Skempton parameter,

$$S = \frac{\alpha}{C^{\mathrm{d}} K^{\mathrm{d}}}. \tag{8.49}$$

The subscript eff is removed from $C^{\mathrm{d}}_{\mathrm{eff}}$ as well as all the K's in (8.17) because these K's are the actual bulk moduli for the isotropic material rather than the effective isotropic bulk moduli of an anisotropic material. It follows from $p = -\hat{\mathbf{B}} \cdot \hat{\mathbf{T}}$ that, in case of an isotropic and compressible medium, $p = -(S/3)\hat{\mathbf{U}} \cdot \hat{\mathbf{T}} = -(S/3)\mathrm{tr}\,\mathbf{T}$. Note that if $K^{\mathrm{f}} = K^{\mathrm{m}}$, then $C^{\mathrm{d}} = \alpha/K^{\mathrm{d}}$ and $S = 1$ whether or not these constituents are incompressible. In the isotropic compressible case (8.45) may be used to show that $\hat{\mathbf{A}} = SK^{\mathrm{d}}C^{\mathrm{d}}\hat{\mathbf{U}}$. In the isotropic incompressible case the Skempton parameter S is equal to 1, thus $p = -(1/3)\hat{\mathbf{U}} \cdot \hat{\mathbf{T}} = -(1/3)\mathrm{tr}\,\mathbf{T}$. It also follows in the isotropic incompressible case that $K^{\mathrm{d}} = 1/C^{\mathrm{d}}$ and $\hat{\mathbf{A}} = \hat{\mathbf{U}}$, as shown by (8.38).

In the case of compressibility, the undrained elastic coefficients $\hat{S}^{\mathrm{u}}$ are related to the drained elastic constants $\hat{\mathbf{S}}^{\mathrm{d}}$ and the tensor $\hat{\mathbf{A}}$ by

$$\hat{\mathbf{S}}^{\mathrm{u}} = \hat{\mathbf{S}}^{\mathrm{d}} - \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} \otimes \hat{B} = \hat{\mathbf{S}}^{\mathrm{d}} - \frac{1}{C^{\mathrm{d}}_{\mathrm{eff}}} (\hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} \otimes \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}}). \tag{8.50}$$

In the case of incompressibility, the undrained elastic coefficients $\hat{\mathbf{S}}^{\mathrm{u}}$ are related to the drained elastic constants $\hat{\mathbf{S}}^{\mathrm{d}}$ by

$$\hat{\mathbf{S}}^{\mathrm{u}} = \hat{\mathbf{S}}^{\mathrm{d}} - K^{\mathrm{d}}_{\mathrm{Reff}} (\hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} \otimes \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}}) \tag{8.51}$$

a result that follows from (8.50) with the application of (8.40). When (8.51) is dotted with $\hat{\mathbf{U}}$ and (8.40) again employed, it follows that the undrained elastic coefficients are also incompressible in the case of assumed incompressibility of the matrix material and the fluid, as one would anticipate:

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{u}} = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} - K^{\mathrm{d}}_{\mathrm{Reff}} (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} \otimes \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}}) = \hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} - \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} = 0. \tag{8.52}$$

In the isotropic compressible case (8.49) reduces to formulas for the undrained bulk modulus K^{u} (where $K^{\mathrm{u}} = (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{u}} \cdot \hat{\mathbf{U}})^{-1}$) and the undrained Poisson's ratio ν^{u} in terms of ν^{d}, K^{d}, K^{f}, K^{m}, and ϕ, thus

$$K^{\mathrm{u}} = K^{\mathrm{d}} + \frac{K^{\mathrm{f}}(1 - (K^{\mathrm{d}}/K^{\mathrm{m}}))^2}{K^{\mathrm{f}}/K^{\mathrm{m}}(1 - (K^{\mathrm{d}}/K^{\mathrm{m}})) - \phi + \phi} \quad \text{and}$$
$$\nu^{\mathrm{u}} = \frac{3\nu^{\mathrm{d}} + S(1 - 2\nu^{\mathrm{d}})(1 - (K^{\mathrm{d}}/K^{\mathrm{m}}))}{3 - S(1 - 2\nu^{\mathrm{d}})(1 - (K^{\mathrm{d}}/K^{\mathrm{m}}))}. \tag{8.53}$$

In the isotropic incompressible case (8.53) reduces to $1/K^{\mathrm{u}} = 0$ and $\nu^{u} = 1/2$, consistent with the general result for incompressibility for the undrained constant set. It follows that $E^{\mathrm{u}} = 3G$.

Problems

8.6.1. Expand the second equality in (8.50) as a six-by-six matrix equation.

8.6.2. Expand the second equality in (8.50) as a six-by-six matrix equation for the special case of transversely isotropic symmetry using the technical elastic constants, i.e., Young's moduli and Poisson's ratios.

8.6.3. Show that (8.53) are a consequence of the two equations (7.20). Do this by working backward; start with (8.53) and substitute K^{d} and G^{d} from (8.8), $\alpha = [1 - (K^{\mathrm{d}}/K^{\mathrm{m}})]$ and employ the relations between the elastic isotropic constants listed in Table 7.2. The mechanics of solving this problem is straightforward algebraic substitution. However, it can become a task if one is not careful to keep the algebraic objective in view. Try using a symbolic algebra program.

8.7 Expressions of Mass and Momentum Conservation

The conservation of mass is expressed by the equation of continuity,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (3.6) \text{ repeated}$$

The form of the mass conservation equation (3.6) is altered to apply to the pore fluid volume by first replacing ρ by $\phi\rho_{\mathrm{f}}$ in (3.6) and then dividing the equation through by ρ_{fo}, thus

$$\frac{1}{\rho_{\mathrm{fo}}}\frac{\partial \phi \rho_{\mathrm{f}}}{\partial t}+\frac{1}{\rho_{\mathrm{fo}}}\nabla\cdot(\phi\rho_{\mathrm{f}}\mathbf{v})=0. \tag{8.54}$$

The form of (8.54) is changed by use of the time derivative of (8.16),

$$\frac{\partial \zeta}{\partial t}=\frac{1}{\rho_{\mathrm{fo}}}\frac{\partial \phi \rho_{\mathrm{f}}}{\partial t} \tag{8.55}$$

and (8.26), thus

$$\partial\zeta/\partial\tau+\nabla\cdot\mathbf{q}=0. \tag{8.56}$$

In the case of incompressibility, $\rho_{\mathrm{f}}=\rho_{\mathrm{fo}}$, and (8.54) becomes

$$\frac{\partial \phi}{\partial t}+\nabla\cdot(\phi\mathbf{v})=0. \tag{8.57}$$

The stress equations of motion in three dimensions,

$$\rho\ddot{\mathbf{u}}=\nabla\cdot\mathbf{T}+\rho\mathbf{d},\quad \mathbf{T}=\mathbf{T}^{\mathrm{T}},\quad (3.37)\text{ repeated}$$

have no simple representation in 6D vector notation, and the conventional notation is employed; $\ddot{\mathbf{u}}$ represents the acceleration and $\mathbf{d}$ the action-at-a-distance force.

Problems

8.7.1. Evaluate each of the following formulas in the limit as $\phi\to 0$ (note that $\phi\to 0$ implies $\phi_{\mathrm{o}}\to 0$): (a) (8.8), (b) (8.13), (c) (8.17), (d) (8.22), (e) (8.27), (f) (8.49), (g) (8.50), (h) (8.53), (i) (8.54).

8.7.2. Evaluate each of the following formulas in the limit as $\phi\to 1$ (note that $\phi\to 1$ implies $\phi_{\mathrm{o}}\to 1$): (a) (8.3), (b) (8.2), (c) (8.17), (d) (8.18). The last two results requires the easily justified restriction that $1/K_{\mathrm{eff}}^{\mathrm{m}}\to 0$ as $\phi\to 1$.

8.8 The Basic Equations of Poroelasticity

An overview of the theory of poroelastic materials can be obtained by considering it as a system of 18 equations in 18 scalar unknowns. This system of equations and unknowns, a combination of conservation principles and constitutive equations, is described in this section. The 18 scalar unknowns are the six components of the stress tensor $\mathbf{T}$, the fluid pressure p, the fluid density ρ_{f}, the variation in fluid content ζ, the six components of the strain tensor $\mathbf{E}$ and the three components of the displacement vector $\mathbf{u}$. The 18 scalar equations of the theory of poroelastic solids are the six equations of the strain–stress-pressure relation (8.1), the six strain displacement relations (2.49),

$2\mathbf{E} = ((\Delta \otimes \mathbf{u})^{\mathrm{T}} + \Delta \otimes \mathbf{u})$, the three equations of motion (3.37), the one fluid content-stress-pressure relation (8.17) (or the one fluid content-strain-pressure relation (8.18)) and the one mass conservation equation (8.56) and a relation between the fluid pressure and the density $p = p(\rho_{\mathrm{f}})$ which is not specified here. The parameters of a poroelasticity problem are the drained effective elastic constants of porous matrix material $\hat{\mathbf{S}}^{\mathrm{d}}$, the Biot effective stress coefficients $\hat{\mathbf{A}}$, $C^{\mathrm{d}}_{\mathrm{eff}}$, the fluid viscosity μ, the intrinsic permeability tensor $\mathbf{K}$, and the action-at-a-distance force $\mathbf{d}$, which are all assumed to be known. If the displacement vector $\mathbf{u}$ is taken as the independent variable, no further equations are necessary. However, if it is not, use of the compatibility equations (2.54) is necessary to insure that the displacements are consistent.

There are many methods of approach to the solution of poroelastic problems for compressible media. The method selected depends upon the information that is provided and the fields that are to be calculated. One approach that has been effective is to solve for the variation in fluid content ζ if the stress or the strain field is known or may be calculated without reference to the variation in fluid content ζ. The diffusion equation for the variation in volume fraction is obtained by first substituting Darcy's law (8.17) into the expression (8.56) for the conservation of mass and subsequently eliminating the pore pressure by use of (8.43), thus

$$\frac{\partial \zeta}{\partial t} - \frac{1}{\mu\Lambda}\hat{\mathbf{K}} \cdot \hat{\mathbf{O}}\zeta = -\frac{1}{\mu\Lambda}\hat{\mathbf{K}} \cdot \hat{\mathbf{O}}[\hat{\mathbf{A}} \cdot \hat{\mathbf{E}}]. \tag{8.58}$$

This shows that the time rate of change of the fluid content ζ is due either to fluid flux or to volume changes caused by the strain field. It is possible to replace the strain on the right hand side of (8.58) by stress in which case (8.58) becomes

$$\frac{\partial \zeta}{\partial t} - \frac{1}{\mu C^{\mathrm{d}}_{\mathrm{eff}}}\hat{\mathbf{K}} \cdot \hat{\mathbf{O}}\zeta = -\frac{1}{\mu C^{\mathrm{d}}_{\mathrm{eff}}}\hat{\mathbf{K}} \cdot \hat{\mathbf{O}}[\hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{T}}]. \tag{8.59}$$

Diffusion equations for the pressure field are also employed in the solution of poroelastic problems. The first diffusion equation for the pore pressure field is obtained by substituting Darcy's law (8.17) into the expression (8.56) for the conservation of mass and subsequently eliminating the variation in fluid content ζ by use of (8.17), thus

$$\frac{\partial p}{\partial t} - \frac{1}{\mu\Lambda}\hat{\mathbf{K}} \cdot \hat{\mathbf{O}}p = -\frac{1}{\Lambda}\hat{\mathbf{A}} \cdot \frac{\partial \hat{\mathbf{E}}}{\partial t}. \tag{8.60}$$

The alternative diffusion equation for the pore pressure field is obtained by replacing the strain on the right hand side of (8.60) by stress, thus

$$\frac{\partial p}{\partial t} - \frac{1}{\mu C^{\mathrm{d}}_{\mathrm{eff}}}\hat{\mathbf{K}} \cdot \hat{\mathbf{O}}p = \frac{1}{C^{\mathrm{d}}_{\mathrm{eff}}}\left(\hat{\mathbf{U}} \cdot (\hat{\mathbf{S}}^{\mathrm{m}} - \hat{\mathbf{S}}^{\mathrm{d}}) \cdot \frac{\partial \hat{\mathbf{T}}}{\partial t}\right). \tag{8.61}$$

For an orthotropic material equation (8.61) may be written in the following form:

$$
\begin{aligned}
&C_{\text{eff}}\frac{\partial p}{\partial t}-\frac{1}{\mu}K_{11}\frac{\partial^2 p}{\partial x_1^2}-\frac{1}{\mu}K_{22}\frac{\partial^2 p}{\partial x_2^2}-\frac{1}{\mu}K_{33}\frac{\partial^2 p}{\partial x_3^2}\\
&\quad=-\left\{\left(\frac{1}{E_1^{\text{d}}}-\frac{\nu_{12}^{\text{d}}}{E_1^{\text{d}}}-\frac{\nu_{31}^{\text{d}}}{E_3^{\text{d}}}-\frac{1}{E_1^{\text{m}}}+\frac{\nu_{12}^{\text{m}}}{E_1^{\text{m}}}+\frac{\nu_{31}^{\text{m}}}{E_3^{\text{m}}}\right)\frac{\partial T_{11}}{\partial t}\right.\\
&\qquad+\left(\frac{1}{E_2^{\text{d}}}-\frac{\nu_{12}^{\text{d}}}{E_1^{\text{d}}}-\frac{\nu_{23}^{\text{d}}}{E_2^{\text{d}}}-\frac{1}{E_2^{\text{m}}}+\frac{\nu_{12}^{\text{m}}}{E_1^{\text{m}}}+\frac{\nu_{23}^{\text{m}}}{E_2^{\text{m}}}\right)\frac{\partial T_{22}}{\partial t}\\
&\qquad\left.+\left(\frac{1}{E_3^{\text{d}}}-\frac{\nu_{23}^{\text{d}}}{E_2^{\text{d}}}-\frac{\nu_{31}^{\text{d}}}{E_3^{\text{d}}}-\frac{1}{E_3^{\text{m}}}+\frac{\nu_{23}^{\text{m}}}{E_2^{\text{m}}}+\frac{\nu_{31}^{\text{m}}}{E_3^{\text{m}}}\right)\frac{\partial T_{33}}{\partial t}\right\}.
\end{aligned}
\tag{8.62}
$$

The boundary conditions on the pore pressure field customarily employed in the solution of this differential equation are (1) that the external pore pressure p is specified at the boundary (a lower pressure permits flow across the boundary), (2) that the pressure gradient ∇p at the boundary is specified (a zero pressure gradient permits no flow across the boundary), (3) that some linear combination of (1) and (2) is specified.

Problem

8.8.1. Verify that (8.62) follows from (8.61) once the assumption of orthotropy is made. Record the form of (8.62) for transversely isotropic symmetry.

8.9 The Basic Equations of Incompressible Poroelasticity

In this section the development of compressible poroelasticity as a system of eighteen equations in 18 scalar unknowns presented in the previous section is specialized to the case of incompressibility. The result is a system of 17 equations in 17 scalar unknowns because the fluid density ρ_{f} is a constant, $\rho_{\text{f}}=\rho_{\text{fo}}$, and no longer an unknown, and an equation relating the fluid pressure to the fluid density $p=p(\rho_{\text{f}})$ does not exist. The other 17 equations in 17 scalar unknowns are the same except for the constraint of matrix material incompressibility. It is important to note that only algebraic coefficients of terms are changed by the transition to incompressible components; the order of the differential equations is unchanged. The diffusion equation (8.61) makes an easy transition to the incompressible case. For incompressibility it follows from (8.40) that $C_{\text{eff}}^{\text{d}}=1/K_{\text{Reff}}^{\text{d}}$ and from (8.32) that $\hat{\mathbf{U}}\cdot\hat{\mathbf{S}}^{\text{d}}=\hat{\mathbf{S}}^{\text{d}}\cdot\hat{\mathbf{U}}=0$, thus (8.61) reduces to

$$
\frac{1}{K_{\text{eff}}^{\text{d}}}\frac{\partial p}{\partial t}-\frac{1}{\mu}\hat{\mathbf{K}}\cdot\hat{\mathbf{O}}p=-\left(\hat{\mathbf{U}}\cdot\hat{\mathbf{S}}^{\text{d}}\cdot\frac{\partial\hat{\mathbf{T}}}{\partial t}\right)
\tag{8.63}
$$

and for an orthotropic material equation (8.63) may be written in the following form:

$$\frac{1}{K_{\text{eff}}^{\text{d}}}\frac{\partial p}{\partial t}-\frac{1}{\mu}K_{11}\frac{\partial^2 p}{\partial x_1^2}-\frac{1}{\mu}K_{22}\frac{\partial^2 p}{\partial x_2^2}-\frac{1}{\mu}K_{33}\frac{\partial^2 p}{\partial x_3^2}$$
$$=-\left\{\left(\frac{1}{E_1^{\text{d}}}-\frac{\nu_{12}^{\text{d}}}{E_1^{\text{d}}}-\frac{\nu_{31}^{\text{d}}}{E_3^{\text{d}}}\right)\frac{\partial T_{11}}{\partial t}+\left(\frac{1}{E_2^{\text{d}}}-\frac{\nu_{12}^{\text{d}}}{E_1^{\text{d}}}-\frac{\nu_{23}^{\text{d}}}{E_2^{\text{d}}}\right)\frac{\partial T_{22}}{\partial t}+\left(\frac{1}{E_3^{\text{d}}}-\frac{\nu_{23}^{\text{d}}}{E_2^{\text{d}}}-\frac{\nu_{31}^{\text{d}}}{E_3^{\text{d}}}\right)\frac{\partial T_{33}}{\partial t}\right\}. \tag{8.64}$$

The boundary conditions on the pore pressure field are coincident with those described at the end of the previous section.

8.10 Some Example Isotropic Poroelastic Problems

Example 8.10.1
Formulate the differential equations governing the problem of determining the vertical surface settlement of a layer of poroelastic material resting on a stiff impermeable base subjected to a constant surface loading. The layer, illustrated in Fig. 8.4, is in the x_1, x_2 plane and the x_3 positive coordinate is in the thickness direction and it is pointed downward in Fig. 8.4. The surface is subjected to an applied compressive stress $T_{33} = -P(t)$, the only nonzero strain component is E_{33}. The free surface of the layer permits the passage of fluid out of the layer.
Solution: First, since the free surface of the layer permits the passage of fluid and the supporting base of the layer is impermeable, the boundary conditions on the pore pressure field are $p = 0$ at $x_3 = 0$, $\partial p/\partial x_3 = 0$ at $x_3 = L$. Next, using the fact that the only nonzero strain component is E_{33} ($E_{33} = \partial u_3/\partial x_3$ from (3.52)) and that the applied compressive stress $T_{33} = -P(t)$ is uniform throughout the layer, the strain–stress-pressure relations (8.1) specialize to the following:

$$0=\frac{1}{E^{\text{d}}}\{(1+\nu^{\text{d}})T_{11}-\nu^{\text{d}}\,\text{tr}\mathbf{T}+(1-2\nu^{\text{d}})\alpha\text{p}\},$$
$$0=\frac{1}{E^{\text{d}}}\{(1+\nu^{\text{d}})T_{22}-\nu^{\text{d}}\,\text{tr}\mathbf{T}+(1-2\nu^{\text{d}})\alpha\text{p}\}$$
$$\frac{\partial u_3}{\partial x_3}=\frac{1}{E^{\text{d}}}\{-(1+\nu^{\text{d}})P(t)-\nu^{\text{d}}\,\text{tr}\mathbf{T}+(1-2\nu^{\text{d}})\alpha\text{p}\}.$$

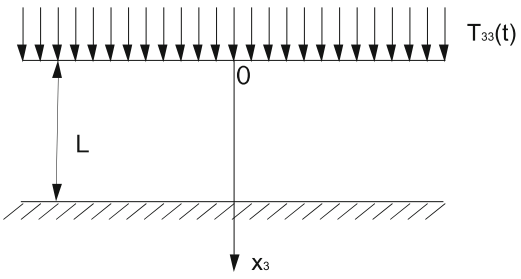


Fig. 8.4 Illustration of a layer of poroelastic material resting on a stiff impermeable base subjected to a uniform time varying surface loading

The solutions of these equations for T_{11}, T_{22}, tr$\mathbf{T}$, and T_{33} are

$$T_{11} = T_{22} = -\frac{\nu^{\mathrm{d}}}{(1-\nu^{\mathrm{d}})}P(t) - \frac{(1-2\nu^{\mathrm{d}})}{(1-\nu^{\mathrm{d}})}\alpha p,$$
$$\mathrm{tr}\mathbf{T} = \frac{E^{\mathrm{d}}}{(1-2\nu^{\mathrm{d}})}\frac{\partial u_3}{\partial x_3} - 3\alpha p = -\frac{(1+\nu^{\mathrm{d}})}{(1-\nu^{\mathrm{d}})}P(t) - \frac{2(1-2\nu^{\mathrm{d}})}{(1-\nu^{\mathrm{d}})}\alpha p$$

and

$$T_{33} = -P(t) = 3K^{\mathrm{d}}\frac{(1-\nu^{\mathrm{d}})}{(1+\nu^{\mathrm{d}})}\frac{\partial u_3}{\partial x_3} - \alpha p$$

and the single strain component E_{33}, is given by

$$E_{33} = \frac{\partial u_3}{\partial x_3} = \frac{(1+\nu^{\mathrm{d}})}{3K^{\mathrm{d}}(1-\nu^{\mathrm{d}})}(-P(t) + \alpha p).$$

The stress equations of motion ((3.37) or (3.38)), or equilibrium in this case, reduce to the condition that the derivative of the T_{33} stress component with respect to x_3 must vanish, thus from the equation directly above,

$$\frac{\partial}{\partial x_3}\left\{3K^{\mathrm{d}}\frac{(1-\nu^{\mathrm{d}})}{(1+\nu^{\mathrm{d}})}\frac{\partial u_3}{\partial x_3} - \alpha p\right\} = 0.$$

Since tr$\mathbf{E} = E_{33}$, substitution of E_{33} into the pressure diffusion equation (8.60), and use of (8.38) for both material and drained constants yields

$$\frac{\partial p}{\partial t} - c\frac{\partial^2 p}{\partial x_3^2} = W\frac{\mathrm{dP}}{\mathrm{dt}}, \tag{8.65}$$

where

$$c = \frac{c_{\mathrm{I}}W}{\alpha}, \quad W = \frac{\alpha(1+\nu^{\mathrm{d}})}{3\Lambda K^{\mathrm{d}}(1-\nu^{\mathrm{d}}) + \alpha^2(1+\nu^{\mathrm{d}}))}$$

and c_{I} represents the value of the constant c when the matrix material and the pore fluid are incompressible,

$$c_{\mathrm{I}} = \frac{K_{11}3K^{\mathrm{d}}(1-\nu^{\mathrm{d}})}{\mu(1+\nu^{\mathrm{d}})}.$$

In the special case when the matrix material and the fluid are incompressible, $c = c_{\mathrm{I}}$and $W = 1$. The two following examples examine special solutions of these equations that may then be specialized to these more special assumptions by the appropriate selection of c and W.

Example 8.10.2
Determine the vertical surface settlement of a layer of poroelastic material resting on a stiff impermeable base subjected to a constant surface loading $T_{33} = -P(t) = -p_o h(t)$. The layer, illustrated in Fig. 8.4, is in the x_1, x_2 plane and the x_3 positive coordinate direction is downward. The conditions for the drainage of the layer are described in Example 8.10.1.

Solution: Since the applied stress at the surface is constant for times greater than zero, it follows that equation (8.65) becomes

$$\frac{\partial p}{\partial t} = c\frac{\partial^2 p}{\partial x_3^2},$$

where the initial pore pressure p^{I} is obtained from the formula for the pore pressure in an undrained isotropic and compressible medium, $p = -(S/3)\hat{\mathbf{U}} \cdot \hat{\mathbf{T}} = -(S/3)\mathrm{tr}\mathbf{T}$ where

$$S = \frac{\alpha}{C^{\mathrm{d}} K^{\mathrm{d}}} \tag{8.49}$$

thus

$$p(x_3, 0) = p^{\mathrm{I}} = \frac{\alpha P_{\mathrm{o}}}{C^{\mathrm{d}} K^{\mathrm{d}}}$$

in one of the initial conditions on the pressure. The other is that $\partial p/\partial x_3 = 0$ at $x_3 = L$.

Assuming separation of variables, $p = X(x_3)T(t)$, it follows from $\partial p/\partial t = c(\partial^2 p/\partial x_3^2)$ that

$$\frac{1}{cT}\frac{\partial T}{\partial t} = \frac{1}{X}\frac{\partial^2 X}{\partial x_3^2} = -\beta^2$$

or

$$\frac{1}{T}\frac{\partial T}{\partial t} = \frac{\partial \log T}{\partial t} = -c\beta^2, \quad \frac{\partial^2 X}{\partial x_3^2} + \beta^2 X = 0$$

thus performing the integration,

$$T(t) = C\mathrm{e}^{-c\beta^2 t}, \quad X(x_3) = A\cos\beta x_3 + B\sin\beta x_3$$

and

$$p(x_3, t) = (A\cos\beta x_3 + B\sin\beta x_3)\mathrm{e}^{-c\beta^2 t}.$$

The condition that $\partial p/\partial x_3 = 0$ at $x_3 = L$ requires that $\partial p/\partial x_3|_{x_3=L} = \beta(-A\sin\beta L + B cos\beta L)\mathrm{e}^{-c\beta^2} = 0$ which is satisfied by setting $A = 0$, and $cos\beta L = 0$ is satisfied by setting $\beta = (1+2n)\pi/2L$ where $n = 0, 1, 2, \ldots$. Substituting these results back into $p(x_3,t) = (A\cos\eta x_3 + B\sin\eta x_3)\mathrm{e}^{-c\eta^2 t}$, and summing over all possible values of n, one obtains the representation

$$p(x_3,t) = \sum_{n=0}^{\infty}\left(B_n \sin\frac{(1+2n)\pi}{2L}x_3\right)\mathrm{e}^{-c((1+2n)\pi/2L)^2 t}.$$

The condition that $p(x_3,0) = p^{\mathrm{I}} = \alpha P_{\mathrm{o}}/C^{\mathrm{d}}K^{\mathrm{d}}$ then yields

$$p(x_3,0) = p^{\mathrm{I}} = \frac{\alpha P_{\mathrm{o}}}{C^{\mathrm{d}}K^{\mathrm{d}}} = \sum_{n=0}^{\infty} B_n \sin\frac{(1+2n)\pi}{2L}x_3.$$

If we multiply both sides of the previous equation by $\sin((1+2m)\pi x_3/2L)$, integrate the result from 0 to L, and recall the orthogonality relations

$$\int_0^L \sin\left\{\frac{(1+2n)\pi x_3}{2L}\right\}\sin\left\{\frac{(1+2m)\pi x_3}{2L}\right\}\mathrm{d}x_3 = \frac{\mathrm{L}}{2}\delta_{\mathrm{nm}}$$

it follows that

$$B_n = \frac{\alpha P_{\mathrm{o}}}{C^{\mathrm{d}}K^{\mathrm{d}}}\int_0^L \sin\left\{\frac{(1+2n)\pi x_3}{2L}\right\}\mathrm{d}x_3 = \frac{2\mathrm{L}\alpha\mathrm{P_o}}{\mathrm{C^dK^d}(1+2\mathrm{n})}$$

and the solution for the pressure field is

$$p(x_3,t) = \frac{2L\alpha P_{\mathrm{o}}}{C^{\mathrm{d}}K^{\mathrm{d}}}\sum_{n=0}^{\infty}\left(\frac{1}{(1+2n)}\sin\frac{(1+2n)\pi}{2L}x_3\right)\mathrm{e}^{-c((1+2n)\pi/2L)^2 t}.$$

Note that this result satisfies the initial conditions (Figs. 8.5 and 8.6).

Example 8.10.3
Determine the vertical surface settlement of a layer of poroelastic material resting on a stiff impermeable base subjected to a harmonic surface loading $T_{33} = -P(t) = -P_{\mathrm{o}}\mathrm{e}^{\mathrm{i}\omega\mathrm{t}}$. The layer, illustrated in Fig. 8.4, is in the x_1, x_2 plane and the x_3 positive coordinate direction is downward. The conditions for the drainage of the layer are described in Example 8.10.1.

Solution: Substituting the surface loading $P(t) = P_{\mathrm{o}}\mathrm{e}^{\mathrm{i}\omega\mathrm{t}}$ into the pressure diffusion equation in Example 8.10.1, the diffusion takes the form

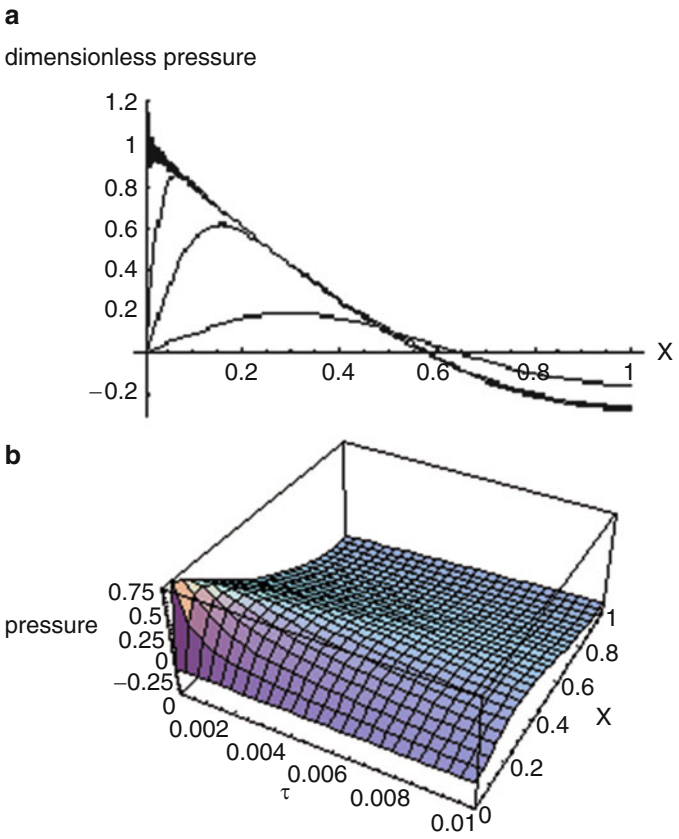


Fig. 8.5 Illustration of the decay of the pressure with time in a layer of poroelastic material resting on a stiff impermeable base subjected to a constant surface loading. (**a**) The dimensionless pressure $p(X, \tau)/Wp_o$ is plotted (ordinate) against the entire range of the dimensionless layer coordinate $X = x_3/L$ from 0 to 1 (abscissa) for dimensionless time t values of 0, 0.0001, 0.001, 0.01, 0.1 and 1. The top curve with the sinusoidal oscillations is the curve for $\tau = 0$. The sinusoidal oscillations arise because only a finite number of terms (200) of the Fourier series were used to determine the plot. It is important to note that this curve begins at the origin and very rapidly rises to the value 1, then begins the (numerically oscillating) decay that is easily visible. The curves for values of t of 0.0001, 0.001 and 0.01 are the first, second and third curves below and to the right of the one for $\tau = 0$. The curves for values of τ of 0.1 and 1 both appear as $p(X, \tau)/Wp_o = 0$ in this plot. (**b**) The dimensionless pressure $p(X, \tau)/Wp_o$ is plotted against the entire range of the dimensionless layer coordinate $X = x_3/L$ from 0 to 1 (abscissa) and the range of dimensionless time τ from 0 to 0.01

$$\frac{\partial p}{\partial t} - c\frac{\partial^2 p}{\partial x_3^2} = \mathrm{i}\omega \mathrm{W} \mathrm{P_o} \mathrm{e}^{\mathrm{i}\omega \mathrm{t}},$$

where the boundary conditions are the same as those in the example above, namely that $p = 0$ at $x_3 = 0$ and $\partial p/\partial x_3 = 0$ at $x_3 = L$. The solution is obtained by assuming that the pressure p is of the form $p(x_3, t) = f(x_3)\mathrm{e}^{\mathrm{i}\omega \mathrm{t}}$, thus the partial differential equation above reduces to an ordinary differential equation,

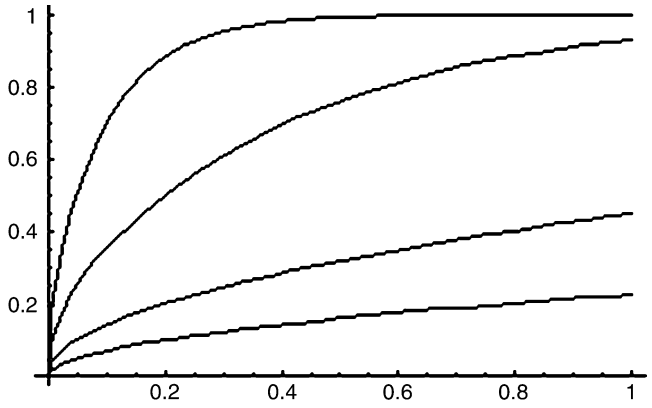


Fig. 8.6 Illustration of the vertical settlement of a layer of poroelastic material resting on a stiff impermeable base subjected to a constant surface loading. This is a plot of the function $g(t)$ against time. The curve underneath the top curve is for a layer twice as thick as the layer of the top curve. The two curves below are for layers that are five and ten times as thick, respectively, as the layer associated with the top curve. Compare with Fig. 1.11c

$$\mathrm{i}\omega \mathrm{f}(\mathrm{x}_3) - \mathrm{c}\frac{\mathrm{d}^2\mathrm{f}}{\mathrm{dx}_3^2} = \mathrm{i}\omega \mathrm{W}\mathrm{p}_\mathrm{o}$$

after dividing through by $\mathrm{e}^{\mathrm{i}\omega \mathrm{t}}$. The solution to this ordinary differential equation is

$$f(x_3) = Wp_\mathrm{o} + A\sinh\frac{x_3}{L}\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}} + B\cosh\frac{x_3}{L}\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}}.$$

The boundary conditions of the previous example, namely that the pressure is zero at $x_3 = 0$ and that the pressure gradient is zero at $x_3 = L$ yield the following solution to the original partial differential equation above, thus

$$p(x_3,t) = Wp_\mathrm{o}\left[1 + \tanh\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}}\sinh\frac{x_3}{L}\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}} - \cosh\frac{x_3}{L}\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}}\right]\mathrm{e}^{\mathrm{i}\omega \mathrm{t}}.$$

The solution to the problem of semi-infinite domain of poroelastic material subjected to a harmonic surface loading $T_{33} = -P(t) = -p_\mathrm{o}\mathrm{e}^{\mathrm{i}\omega \mathrm{t}}$ may also be obtained without difficulty. This problem is also illustrated in Fig. 8.4 if the stiff impermeable base is removed. The solution to the original differential equation above with the boundary conditions that the pressure is zero at $x_3 = 0$ and that the pressure gradient tends to zero as x_3 becomes large is

$$p(x_3,t) = Wp_\mathrm{o}a(x_3)\mathrm{e}^{\mathrm{i}\omega \mathrm{t}}, \quad \text{where} \quad \mathrm{a}(\mathrm{x}_3) = [1 - \mathrm{e}^{-\mathrm{x}_3\sqrt{\mathrm{i}\omega/\mathrm{c}}}].$$

Fig. 8.7 A plot of the function $|a(x_3)|$ against x_3. This plot has a peak of 1.06702 at $x_3 = 3.33216$ and then it is asymptotic to the value one as x_3 tends to infinity

The derivation of this solution is a problem at the end of the section. The real part of $a(x_3)$ is plotted in Fig. 8.7. This plot has a peak of 1.06702 at $x_3 = 3.33216$ and then it is asymptotic to the value one as x_3 tends to infinity. This indicates that the depth beyond $\delta = 3.33216\sqrt{c/\omega}$ from the surface, where $a(x_3)$ has an almost constant value of one, the pore pressure p is in phase with the surface loading and proportional to it by the factor W. The interpretation of this result is that the departure of the pore pressure fluctuations from the undrained solution is confined to a boundary layer of thickness δ. The semi-infinite domain solution is therefore applicable to the finite layer problem under consideration provided $\delta < L$.

The desired settlement of the free surface, $u_3(0, t)$ is calculated following the method of Example 8.10.2 thus

$$u(0,t) = \frac{\alpha(1+\nu^{\mathrm{d}})WLp_{\mathrm{o}}}{3K^{\mathrm{d}}(1-\nu^{\mathrm{d}})}\sqrt{\frac{c}{\mathrm{i}\omega \mathrm{L}^2}}\tanh\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}}\Bigg]\mathrm{e}^{\mathrm{i}\omega t}.$$

A plot against frequency of the absolute value of the function

$$\sqrt{\frac{c}{\mathrm{i}\omega \mathrm{L}^2}}\tanh\sqrt{\frac{\mathrm{i}\omega \mathrm{L}^2}{c}}$$

determining the amplitude of the settlement of the free surface $u_3(0, t)$ is shown in Fig. 8.8. At very large frequency the poroelastic layer behaves as if it were undrained, that is to say $|u(0, t)| \to 1$ as $\omega \to \infty$. At low frequencies it behaves as if it were drained, that is to say $|u(0, t)| \to 0$ as $\omega \to 0$.

Problems

8.10.1. Using the assumptions of Example 8.10.1 and the equation (8.50) derive the pressure diffusion equation

$$\frac{\partial p}{\partial t} - c\frac{\partial^2 p}{\partial x_3^2} = W\frac{\partial P}{\partial t},$$

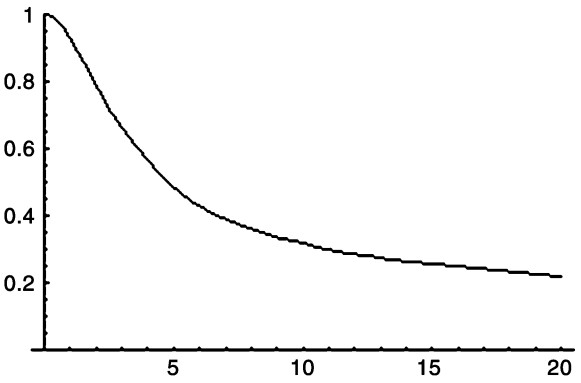


Fig. 8.8 Illustration of the vertical settlement of a layer of poroelastic material resting on a stiff impermeable base subjected to a harmonic surface loading. A plot of the absolute value of the function $\sqrt{\frac{c}{i\omega L^2}}\tanh\sqrt{\frac{i\omega L^2}{c}}$ against frequency. See Example 8.8.3

where c and W are given by

$$c = c_{\mathrm{I}} W/\alpha,\ W = \alpha(1+\nu^{\mathrm{d}})/(3\Lambda K^{\mathrm{d}}(1-\nu^{\mathrm{d}}) + \alpha^2(1+\nu^{\mathrm{d}}))$$

and c_{I} represents the value of the constant c when the matrix material and the pore fluid are incompressible,

$$c_{\mathrm{I}} = \frac{K_{11} 3K^{\mathrm{d}}(1-\nu^{\mathrm{d}})}{\mu(1+\nu^{\mathrm{d}})}.$$

8.10.2. Verify that the solution to the pressure diffusion differential equation in Example 8.10.2 satisfies the specified form of the differential equation and the appropriate boundary and initial conditions.

8.10.3. Determine the flux q_3 of the pore fluid from out of the top surface of the layer in Example 8.10.2.

8.10.4. Verify that the solution to the pressure diffusion differential equation in Example 8.10.3 satisfies the specified form of the differential equation and the appropriate boundary conditions.

8.10.5. Determine the pressure distribution in a semi-infinite domain of poroelastic material subjected to an harmonic surface loading $T_{33} = -P(t) = -P_{\mathrm{o}}\mathrm{e}^{\mathrm{i}\omega t}$. The surface of the domain is in the x_1, x_2 plane and the x_3 positive coordinate direction is downward. Drainage of the semi-infinite domain is only allowed at the surface. The solution to the pressure diffusion equation,

$$\partial p/\partial t - c(\partial^2 p/\partial x_3^2) = W(\partial P/\partial t)$$

for the semi-infinite domain will determine the pressure field.

8.10.6. Determine the flux q_3 of the pore fluid from out of the top surface of the layer in Example 8.10.3.

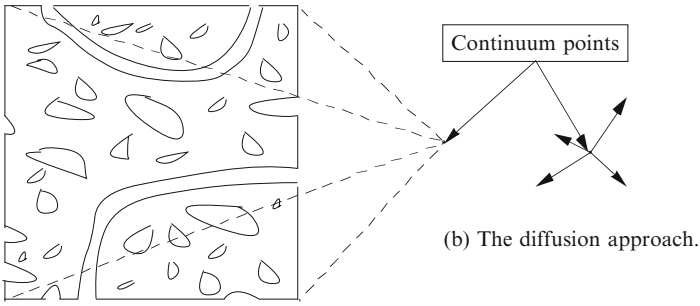
8.11 Three Approaches to Poroelasticity

There are three different major approaches to the development of the same basic equations for the theory of poroelasticity. Each approach is rigorous to its own hypotheses and stems from well-established mathematical and/or physical models for averaging material properties. The averaging processes are the difference between the three approaches.

The first approach, the effective medium approach, originates in the solid mechanics tradition and the averaging process involved is the determination of effective material parameters from a RVE, as discussed in the work of Hashin and Shtrikman (1961), Hill (1963) and others. Standard contemporary references are the books of Christensen (1979) and Nemat-Nasser and Hori (1993). The effective medium (parameter) approach appears in a primitive form in the early Biot work (1941) and grows in sophistication with time through the work of Nur and Byerlee (1971), Rice and Cleary (1976), Carroll (1979), Rudnicki (1985) and Thompson and Willis (1991). The development of effective moduli/parameter theory over the last 50 years occurred almost in parallel with the increasing sophistication and refinement of the original Biot formulation.

The second approach is called the mixture theory approach; mixture theory is developed in Chap. 10. Mixture theory is based on diffusion models and has a very different philosophy and a longer history than the RVE approach. It stems from a fluid mechanics and thermodynamical tradition and goes back to the last century. Fick and Stefan suggested (Truesdell and Toupin 1960, section 158) that each place in a fixed spatial frame of reference might be occupied by several different particles, one for each constituent. This is a Eulerian approach in that the flux of the various species toward and away from a fixed spatial point is considered. Truesdell (1957) assigned to each constituent of a mixture in motion a density, a body force density, a partial stress, a partial internal energy density, a partial heat flux, and a partial heat supply density. He postulated equations of balance of mass, momentum, and energy for each constituent and derived the necessary and sufficient conditions that the balance of mass, momentum, and energy for the mixture be satisfied. Bowen (1967) summarized the formative years of this subject. For subsequent developments see Bowen (1976, 1980, 1982) and Müller (1968, 1985). An advantage of mixture theory approach over the other approaches appears when a number of different fluid species are present and in relative motion. This advantage is attractively illustrated in the theory for the swelling and deformation behavior of articular cartilage by Lai et al. (1991), Gu et al. (1993) and Huyghe and Janssen (1997).

The key difference between the effective parameter approach and the mixture approach to poroelastic models is the averaging process employed. The effective parameter approach illustrated in Fig. 8.9a is a schematic version of the viewpoint described in Biot (1941). A small but finite volume of the porous medium is used for the development of constitutive equations for the fluid-infiltrated solid. These constitutive equations are then assumed to be valid at a point in the continuum. This is an early form of the RVE approach used in composite material theory today.



(b) The diffusion approach.

(a) The effective parameter approach.

Fig. 8.9 Illustrations of the effective medium approach and the mixture theory approach

The length or size of the RVE is assumed to be many times larger than the length scale of the microstructure of the material, say the size of a pore. The length of the RVE is the length of the material structure over which the material microstructure is averaged or "homogenized" in the process of forming a continuum model. Biot's presentation is consistent with the notion of an RVE, although the RVE terminology did not exist when his theory was formulated. The homogenization approach is illustrated in Fig. 8.9a by the dashed lines from the four corners of the RVE to the continuum point. The material parameters or constants associated with the solid phase are more numerous and difficult to evaluate compared to those associated with the fluid phase. The Biot—effective modulus approach provides a better understanding of the effective solid mechanical parameters like effective solid moduli and constituent compressibility than does the mixture theory approach.

The averaging process for the mixture approach is illustrated in Fig. 8.9b. This is a Eulerian approach in that the flux of the various species toward and away from a fixed spatial point is considered. The fixed spatial point is shown in Fig. 8.9b and the vectors represent the velocities of various species passing through the fixed spatial point. It is important to note that, for mixture theory, the averaging is density weighted on the basis of the density of each species in the mixture, instead of being averaged over a finite volume of the porous solid as in the Biot approach. This is the key difference between the Biot and the mixture theory approach. In neither approach is a length scale specified, but an averaging length is implied in the Lagrangian nor material, Biot—effective modulus, approach because a finite material volume is employed as the domain to be averaged over. On the other hand the mixture theory is Eulerian and considers a fixed spatial point through which different materials pass and, as with the Biot approach, no length scale is suggested. It is difficult to imagine a length scale for the mixture theory approach other than one based on the mean free paths associated with the constituents. The significantly different averaging lengths in the two approaches reflect the difference in the averaging methods. Cowin and Cardoso (2012) suggest a method for developing the Biot- effective modulus approach in mixture theory that requires small modifications of mixture theory.

Bowen (1982) showed that it was possible to recover the Biot constitutive equations from the mixture theory approach. In particular Bowen derived equations (his 8.23 and 8.24) that have the same form as equations (7) and (8) of Rice and Cleary (1976); these equations are (3) and (5) in the text below. (Actually the second summand on the right hand side of Bowen's equation (8.24) is missing a factor of 3; this is probably a typo or an algebraic slip.) Earlier Mow and Lai (1980, footnote page 291) indicated how the mixture theory-based (incompressible) biphasic theory for articular cartilage could be linearized to the incompressible Biot theory.

In Chap. 10 the difference between the RVE approach of Biot, represented by Fig. 8.9a and the mixture theory approach represented by Fig. 8.9b is narrowed by minor changes in the mixture theory approach.

The third major approach to the development of the poroelastic equations is due to Burridge and Keller (1981). These authors rederived the dynamic form of the same basic set of equations using a two-space method of homogenization. This method provides a systematic method for deriving macroscopic equations that govern the behavior of the medium on the microscale. Thus, at the continuum point, the three rigorous theoretical developments lead to the same set of equations, and the difference between Biot (RVE), the mixture-theory-based, and the homogenization derivations is the method of averaging. The nature of the equations is better understood because there are three approaches. The Biot approach provides better insight into the nature of the parameters associated with the solid phase, the mixture theory approach provides the mechanism for averaging over different fluid phases, and the homogenization approach illuminates the dynamical (wave propagation) characteristics of the theory.

The basic equations of quasistatic poroelasticity developed here are extended to include a dependence the fabric tensor (introduced at the end of Chap. 7) in Cowin (2004).

8.12 Relevant Literature

There is only one text on poroelasticity (Coussy 1995), but there is much related material in Bear (1972). There is also a book on the very interesting historical development of the theory, de Boer (2000), which also presents a noteworthy presentation of the theory. The presentation of poroelasticity in this chapter was taken from the following papers: Biot (1941, 1955, 1956a, b, 1962a, b), Rice and Cleary (1976), Rudnicki (1985), Thompson and Willis (1991), Cowin (2003), Cowin and Mehrabadi (2007), Cowin and Cardoso (2009) and the excellent summary of Detournay and Cheng (1993). In these papers the proofs omitted in this chapter, as well as some technical restrictions or assumptions underlying those proofs, are documented. The exception to the inclusion of proofs occurred at the end of Sect. 8.2. The proof included was the derivation of the formula (8.3)

for the *Biot effective stress coefficient tensor* $\hat{\mathbf{A}}$. This formula has an interesting history. It was suggested by Geertsma (1957) and by Skempton (1961) then proved with increasing generality by Nur and Byerlee (1971), Carroll (1979) and Thompson and Willis (1991). The material in this paragraph follows the elegant derivation of the formula by Nur and Byerlee (1971) in the isotropic case and by Carroll (1979) for the anisotropic case. The problems described in Sect. 8.10 were taken from Detournay and Cheng (1993).

References

Bear J (1972) Dynamics of fluids in porous media. Elsevier, New York
Biot MA (1941) General theory of three-dimensional consolidation. J Appl Phys 12:155–164
Biot MA (1955) Theory of elasticity and consolidation for a porous anisotropic solid. J Appl Phys 26:182–185
Bowen RM (1967) Toward a thermodynamics and mechanics of mixtures. Arch Rat Mech Anal 24:370–403
Bowen RM (1976) Theory of mixtures. In Eringen AC (ed) Continuum physics, vol III, Mixtures and EM field theories. Academic, New York, pp 1–127
Bowen RM (1980) Incompressible porous media models by use of the theory of mixtures. Int J Eng Sci 18:1129–1148
Bowen RM (1982) Compressible porous media models by use of the theory of mixtures. Int J Eng Sci 20:697–735
Burridge R, Keller JB (1981) Poroelasticity equations derived from microstructure. J Acoust Soc Am 70:1140–1146
Carroll MM (1979) An effective stress law for anisotropic elastic deformation. J Geophys Res 84:7510–7512
Christensen RM (1979) Mechanics of composite materials. Wiley, New York
Coussy O (1995) Mechanics of porous continua. Wiley, England
Cowin SC (2003) A recasting of anisotropic poroelasticity in matrices of tensor components. Transport Porous Med 50:35–56
Cowin SC (2004) Anisotropic poroelasticity: fabric tensor formulation. Mech Mater 36:665–677
Cowin SC, Mehrabadi MM (2007) Compressible and incompressible constituents in anisotropic poroelasticity: the problem of unconfined compression of a disk. J Mech Phys Solids 55:161–193
Cowin SC, Cardoso (2012) Mixture theory-based poroelasticity as a model of interstitial tissue growth, Mech Mat 44:47-57
de Boer R (2000) Theory of porous media. Springer, Berlin
Detournay E, Cheng H-DA (1993) Fundamentals of poroelasticity. In: Hudson JA (ed) Comprehensive rock engineering: principles, practice & projects. Pergamum Press, Oxford, pp 113–171
Geertsma J (1957) The effect of fluid pressure decline on volumetric changes of porous rocks. Trans AIME 210:333–340
Gu WY, Lai WM, Mow VC (1993) Transport of fluid and ions through a porous-permeable charged-hydrated tissue, and streaming potential data on normal bovine articular cartilage. J Biomech 26:709–723
Hashin Z, Shtrikman S (1961) Note on the variational approach to the theory of composite elastic materials. J Franklin Inst 271:336–341
Hill R (1952) The elastic behavior of crystalline aggregate. Proc Phys Soc A 65:349–354

Hill R (1963) Elastic properties of reinforced solids: some theoretical principles. J Mech Phys Solids 11:357–372
Huyghe JM, Janssen JD (1997) Quadriphasic mechanics of swelling incompressible porous media. Int J Eng Sci 35:793–802
Kachanov M (1999) Solids with cracks and non-spherical pores: proper parameters of defect density and effective elastic properties. Int J Fract 97:1–32
Lai WM, Hou JS, Mow VC (1991) A triphasic theory for the swelling and deformation behaviors of articular cartilage. J Biomech Eng 113:245–258
Mow VC, Lai WL (1980) Recent developments in synovial joint mechanics. SIAM Rev 22:275–317
Müller I (1968) A thermodynamic theory of mixtures. Arch Rat Mech Anal 28:1–39
Müller I (1985) Thermodynamics. Pitman, Boston
Nemat-Nasser S, Hori M (1993) Micromechanics: overall properties of heterogeneous materials. North-Holland, Amsterdam
Nur A, Byerlee JD (1971) An exact effective stress law for elastic deformation of rock with fluids. J Geophys Res 76:6414
Rice JR, Cleary MP (1976) Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents. Rev Geophys Space Phys 14:227–241
Rudnicki JW (1985) Effect of pore fluid diffusion on deformation and failure of rock. In: Bazant Z (ed) Mechanics of geomaterials. Wiley, New York, pp 315–347
Sevostianov I, Kachanov M (2001) Author's response. J Biomech 34:709–710
Shafiro B, Kachanov M (1997) Materials with fluid filled pores of various shapes: effective elastic properties and fluid pressure polarization. Int J Solids Struct 34:3517–3540
Skempton AW (1954) The pore pressure coefficients A and B. Géotechnique 4:143–147
Skempton AW (1961) Effective stress in soils, concrete and rock. In Pore pressure and suction in soils. Butterworth, London
Thompson M, Willis JR (1991) A reformation of the equations of anisotropic poroelasticity. J Appl Mech 58:612–616
Truesdell CA (1957) Sulle basi della termomeccania. Rend Lincei 22(33–38):1158–1166
Truesdell CA, Toupin RA (1960) The classical field theories. In: Flugge S (ed) Handbuch der Physik. Springer, Berlin

Chapter 9
Dynamic Poroelasticity

The modern theory of the motion of the water in the soil based on Darcy's law does not take into account the fact that the particles of the soil can be elastically compressed and extended, assuming that the external forces and the hydrostatic pressure act on the liquid filling these pores only. This simplifying assumption necessitates a correction even in the case of problems on the steady flow of soil water under the influence of given external forces. It becomes, however, wholly untenable in the case of such questions as the propagation of elastic vibrations in the soil (Frenkel 1944).

9.1 Poroelastic Waves

Dynamic poroelasticity is very different from the static and quasistatic poroelasticity considered in the previous chapter. The concern here shifts from pore fluid movements over longer distances and longer time spans to pore fluid movements over shorter distances and shorter time spans associated with the passage of waves. In poroelasticity there are the usual shear waves in an elastic solid media (see Example 6.3.4), but there are two types of compressive waves called the fast and slow waves. The fast wave is the usual compressive wave characteristic of elastic solid, slightly modified by its pore fluid, and it moves mainly through the solid matrix. The slow wave is unusual, and peculiar to poroelasticity, in that the relative movement of the pore fluid and the solid matrix is its generator.

The intended primary application of the dynamic poroelastic results is to water-saturated geological and biological materials. The primary objective of the application of dynamic poroelasticity to geological materials is to understand and interpret the consequences of waves associated with earthquakes and to access oil producing potential of geological structures. The primary objective of the application of dynamical poroelasticity to biological tissue is to understand its use as a model for ultrasound in noninvasive clinical tool to evaluate the state of internal

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_9,

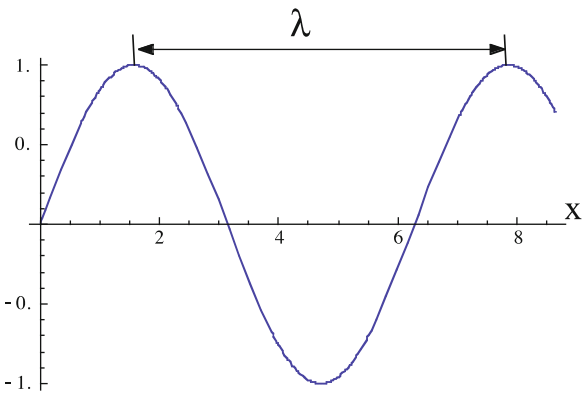


Fig. 9.1 An harmonic plane progressive poroelastic wave considered to be moving in the x direction in medium is illustrated. The relationship between the wave number of a harmonic wave k, the wave length of the wave λ, the wave velocity v, the frequency f (in radians per second) and the frequency ω (in cycles per second, Hertz or Hz) is given by $k = 2\pi/\lambda = 2\pi f/v = \omega/v$

tissue structures. Most techniques for the noninvasive evaluation tissues generate ionizing radiation; ultrasound does not.

This chapter concerns harmonic plane progressive poroelastic waves. Harmonic describes the sinusoidal character of the amplitude of the wave, plane refers to the consideration that the wave propagates as a planar surface and progressive means that it moves through the medium, in this case a poroelastic medium. The amplitude and the wavelength of an harmonic plane progressive poroelastic wave are illustrated in Fig. 9.1. Such a wave is represented mathematically in several different ways, for example a displacement vector $\mathbf{u}(x, t)$ may be written as

$$\mathbf{u}(x,t) = \mathbf{a}\mathrm{e}^{\mathrm{i}\omega((\mathbf{n}\cdot\mathbf{x}/v)-t)} \quad \text{or} \quad \mathbf{u}(x,t) = \mathbf{a}\mathrm{e}^{\mathrm{i}(k(\mathbf{n}\cdot\mathbf{x})-\omega t)}.$$

In these representations $x = \mathbf{x}\cdot\mathbf{n}$ represents distance along the propagation path, $\mathbf{n}$ is the direction in which the wave is propagating, $\mathbf{a}$ is the amplitude or polarization of the wave and t represents time. If the directions of $\mathbf{a}$ and $\mathbf{n}$ coincide the wave is said to be a pure longitudinal wave or pressure (P) wave. If the directions of $\mathbf{a}$ and $\mathbf{n}$ are perpendicular the wave is said to be a pure shear (S) wave. The plane that is propagating in the direction $\mathbf{n}$ is the plane perpendicular to $\mathbf{n}$. The wave number of a harmonic wave, k, the wave length of the wave, λ, the wave velocity, v, the frequency f (in radians per second) and the frequency ω (in cycles per second, Hertz or Hz) are related by $k = 2\pi/\lambda = 2\pi f/v = \omega/v$. From these relationships one can see that $v = \omega/k$, which is actually called the phase velocity of the wave and denoted by some as $v_\mathrm{p} = \omega/k$ to distinguish it from the group velocity of the wave, $v_\mathrm{g} = \mathrm{d}\omega/\mathrm{d}k$. The phase velocity is the speed of a crest belonging to the average wave number k. The group velocity is the velocity of the wave's modulating envelope.

Problems

9.1.1. There is a characteristic length L_RVE associated with the representative volume element (RVE) employed in poroelasticity. In order for a wave

passing through this poroelastic medium to obtain an appropriate average of properties in the poroelastic RVE, what condition must be satisfied between the wavelength of a wave λ and the characteristic length L_{RVE} associated with the volume element employed in the poroelastic model?

9.1.2. What parameter of an experiment may the experimentalist control to insure that the criterion that is the answer to question 9.1.1 above is satisfied?

9.2 Historical Backgrounds and the Relationship to the Quasistatic Theory

The formulation of the theory of wave motions in the context of poroelastic theory presented here is consistent with the presentations of Biot (1941, 1955, 1956a, b, 1962a, b), Plona and Johnson (1983), Sharma (2005, 2008) and many others. The origins of this analysis appear in the work of Frenkel (1944). Unchanged by the addition of anisotropy is the fact that the total elastic volumetric response in poroelasticity described in the previous chapter is due to a combination of the elastic volumetric response of the matrix material of the porous solid, the volumetric elastic response of the pore fluid, and the pore volume changes in the porous medium. The poroelastic constitutive equations are described in this and the following section follow Biot's (1956a, b, 1962a, b) formulation of the appropriate two coupled wave equations. In the following section the coupled wave equations ((9.16) and (9.17)) for the propagation of plane waves in an anisotropic, saturated porous medium are developed and, in the section after that, the relationships between the material coefficients and the fabric are recorded. The algebra associated with the representation of plane waves is developed in Sect. 9.4, and the fabric dependence of the coefficients is recorded in Sect. 9.5. The propagation of plane waves in a principal direction material symmetry and the direction that is not a principal direction of material symmetry, are recorded in Sects. 9.6 and 9.7, respectively.

In his 1956 papers on wave propagation Biot (1956a, b) let $\mathbf{u}$ represent the displacement vector of the solid matrix phase as has been done in this chapter, and $\mathbf{U}$ represent the displacement vector of the fluid phase, which is not done in this chapter. These were the two basic kinematic quantities employed in Biot (1956a, b). In Biot (1962a) the displacement vector of the fluid phase $\mathbf{U}$ was replaced by the displacement vector $\mathbf{w}$ of the fluid relative to the solid, thus

$$\mathbf{w} = \phi(\mathbf{U} - \mathbf{u}). \tag{9.1}$$

The present development follows Biot (1962a, b) and the two basic kinematic fields are considered to be the displacement vectors $\mathbf{u}$ and $\mathbf{w}$. The relative velocity of the fluid and solid components is, from (9.1),

$$\dot{\mathbf{w}} = \phi(\dot{\mathbf{U}} - \dot{\mathbf{u}}). \tag{9.2}$$

The variation in fluid content ζ may then be defined in terms of $\mathbf{w}$:

$$\zeta = -\nabla \cdot \mathbf{w}. \tag{9.3}$$

Recall that the variation in fluid content ζ is the variation of the fluid volume per unit volume of the porous material; it is defined as the difference between the strain of the pore space and the strain of the fluid volume in the pore space and is dimensionless (see (8.16)).

In addition to the slight changes in kinematic notation noted above, Biot (1962a, b) employed slightly different notation for the poroelastic constitutive relations. As an introduction to these changes consider the inverse of (8.17),

$$p = M(\zeta - \hat{\mathbf{A}} \cdot \hat{\mathbf{E}}), \quad (p = M(\zeta - A_{ij}E_{ij})),$$

where M is the inverse of Λ defined by (8.18),

$$M = 1/\Lambda \tag{9.4}$$

and the representation of stress $\hat{\mathbf{T}}$ as a function of the strain $\hat{\mathbf{E}}$ and the variation in fluid content ζ, rather than as a function of the strain and the pore pressure p, is

$$\hat{\mathbf{T}} = \{\hat{\mathbf{C}}^{\mathrm{d}} + M(\hat{\mathbf{A}} \otimes \hat{\mathbf{A}})\} \cdot \hat{\mathbf{E}} - M\hat{\mathbf{A}}\zeta, \quad (T_{ij} = (C^{\mathrm{d}}_{ijkm} + MA_{km}A_{ij})E_{km} - MA_{ij}\zeta). \tag{9.5}$$

Biot (1962a, equations (9.11)) employed a slightly different notation for the two previous expressions, namely

$$\hat{\mathbf{T}} = \hat{\mathbf{Z}}^{\mathrm{d}} \cdot \hat{\mathbf{E}} - \zeta\hat{\mathbf{M}}, \quad (T_{ij} = Z^{\mathrm{d}}_{ijkm}E_{km} - \zeta M_{ij}) \tag{9.6}$$

and

$$p = -\hat{\mathbf{M}} \cdot \hat{\mathbf{E}} + \zeta M, \quad (p = -M_{km}E_{km} + \zeta M), \tag{9.7}$$

where Biot's parameters M, $\hat{\mathbf{M}}$ and $\hat{\mathbf{Z}}^{\mathrm{d}}$ are related to $C^{\mathrm{d}}_{\mathrm{eff}}$, $\hat{\mathbf{A}}$ and $\hat{\mathbf{C}}^{\mathrm{d}}$ above by

$$\begin{aligned} &\hat{\mathbf{M}} = M\hat{\mathbf{A}}, \quad (M_{ij} = MA_{ij}), \quad \hat{\mathbf{Z}}^{\mathrm{d}} = \hat{\mathbf{C}}^{\mathrm{d}} + M(\hat{\mathbf{A}} \otimes \hat{\mathbf{A}}), \\ &(Z^{\mathrm{d}}_{ijkm} = C^{\mathrm{d}}_{ijkm} + MA_{ij}A_{km}). \end{aligned} \tag{9.8}$$

It is interesting to note that Biot's elasticity tensor $\hat{\mathbf{Z}}^{\mathrm{d}}$ differs from the drained elasticity tensor $\hat{\mathbf{C}}^{\mathrm{d}}$ by the term $M(\hat{\mathbf{A}} \otimes \hat{\mathbf{A}})$, which is M times the open product of Biot effective stress coefficient vector $\hat{\mathbf{A}}$ with itself. Equations (9.6) and (9.7) take the following forms when the strain–displacement relations (2.49) and (9.3) are employed:

$$T_{ij} = Z_{ijkm} u_{k,m} + M_{ij} w_{k,k} \tag{9.9}$$

and

$$p = -M_{km} u_{k,m} - M w_{k,k}. \tag{9.10}$$

The balance of momentum in the form of the dynamical stress equations of motion

$$\rho \ddot{\mathbf{x}} = \nabla \cdot \mathbf{T} + \rho \mathbf{d}, \quad \mathbf{T} = \mathbf{T}^{\mathrm{T}} \quad \text{(3.37 repeated)}$$

will now be applied twice, once to the solid phase and once to the fluid phase. In both cases of its application the action-at-a-distance force $\mathbf{d}$ is neglected. The application to the solid phase involves the mass times acceleration terms for the fluid saturated solid phase $\phi \rho_{\mathrm{f}} \ddot{\mathbf{U}} + (1 - \phi) \rho_{\mathrm{s}} \ddot{\mathbf{u}}$ and may be reduced to $\rho \ddot{\mathbf{u}} + \rho_{\mathrm{f}} \ddot{\mathbf{w}}$ when (9.2) is used as well as the definition of ρ,

$$\rho = (1 - \phi) \rho_{\mathrm{s}} + \phi \rho_{\mathrm{f}}, \tag{9.11}$$

where ρ_{s} represents the density of the solid matrix material, thus

$$\nabla \cdot \mathbf{T} = \rho \ddot{\mathbf{u}} + \phi \rho_{\mathrm{f}} \ddot{\mathbf{w}}. \tag{9.12}$$

In the application of the balance of momentum to the fluid phase the mass times acceleration term may first be written as $\rho_{\mathrm{f}} \ddot{\mathbf{U}}$ and may be rendered in the form $\rho_{\mathrm{f}}(\ddot{\mathbf{u}} + (\ddot{\mathbf{w}}/\phi))$ by use of (9.2). However Biot (1962a, b) extends this formulation of this mass times acceleration term of include $\mathbf{J}$, the micro–macro velocity average tensor, thus $\rho_{\mathrm{f}}(\ddot{\mathbf{u}} + (\ddot{\mathbf{w}}/\phi))$ becomes $\rho_{\mathrm{f}}(\ddot{\mathbf{u}} + \mathbf{J} \cdot \ddot{\mathbf{w}})$ where the newly defined $\mathbf{J}$ incorporates the factor ϕ^{-1}. The micro–macro velocity average tensor $\mathbf{J}$ functions like a density distribution function that relates the relative micro-solid–fluid velocity to its bulk volume average $\dot{\mathbf{w}}$. In introducing this concept Biot was clearly viewing and modeling the poroelastic medium as hierarchical. The use of $\rho_{\mathrm{f}}(\ddot{\mathbf{u}} + \mathbf{J} \cdot \ddot{\mathbf{w}})$ yields a balance of linear momentum for the pore fluid phase in the form

$$-\nabla p = \rho_{\mathrm{f}}(\ddot{\mathbf{u}} + \mathbf{J} \cdot \ddot{\mathbf{w}}). \tag{9.13}$$

The next step in the development of the coupled wave equations is to substitute the expression (9.9) for the solid stress tensor and the expression (9.10) for the pore fluid pressure into the conservation of momentum for the solid phase (9.12) and the conservation of momentum for the fluid phase (9.13), respectively. However the analysis to this point does not include the force on the fluid phase set up by the drag of the fluid moving over the solid surface. To account for this force the flow-resistivity tensor $\mathbf{R}$, defined as the inverse of the intrinsic permeability tensor $\mathbf{K}$ (see (8.26)), is introduced:

$$\mathbf{R} = \mathbf{K}^{-1}. \tag{9.14}$$

The viscous resistive force $\mu\mathbf{R} \cdot \dot{\mathbf{w}}$ represents the effect of the fluid–solid interaction on the fluid phase. Formally it should be subtracted from the left hand side of (9.13), but we add it to the right hand side, to the mass times acceleration terms, thus

$$-\nabla p = \rho_{\mathrm{f}}(\ddot{\mathbf{u}} + \mathbf{J} \cdot \ddot{\mathbf{w}}) + \mu\mathbf{R} \cdot \dot{\mathbf{w}}. \tag{9.15}$$

Finally, accomplishing the substitutions mentioned above into (9.12) and (9.13), but using (9.15) instead of (9.13), one obtains

$$Z_{ijkm}\frac{\partial^2 u_k}{\partial x_m \partial x_j} + M_{ij}\frac{\partial^2 w_k}{\partial x_k \partial x_j} = \rho\ddot{u}_i + \rho_{\mathrm{f}}\ddot{w}_i, \tag{9.16}$$

$$M_{km}\frac{\partial^2 u_k}{\partial x_m \partial x_i} + M\frac{\partial^2 w_k}{\partial x_k \partial x_i} = \phi\rho_{\mathrm{f}}(\ddot{u}_i + J_{ij}\ddot{w}_j) + \mu R_{ij}\dot{w}_j. \tag{9.17}$$

Equations (9.16) and (9.17) are two coupled wave equations for the solid displacement field $\mathbf{u}$ and the displacement field $\mathbf{w}$ of the fluid relative to the solid.

Problems

9.2.1. Explain the differences between the quasistatic formulation and the dynamic formulation of the theory of poroelasticity.

9.2.2. How does the composite elasticity tensor $\hat{\mathbf{Z}}^{\mathrm{d}}$, $\hat{\mathbf{Z}}^{\mathrm{d}} = \hat{\mathbf{C}}^{\mathrm{d}} + M(\hat{\mathbf{A}} \otimes \hat{\mathbf{A}})$, change when the porosity of the porous medium vanishes?

9.2.3. Show that the coupled system of equations (9.16) and (9.17) reduce to the wave equation for an anisotropic elastic continuum with no porosity.

9.2.4. Using the indicial notation substitute the expression (9.9) for the solid stress tensor and the expression (9.10) for the pore fluid pressure into the conservation of momentum for the solid phase (9.12) and the conservation of momentum for the fluid phase (9.15), respectively, and derive the coupled wave equations (9.16) and (9.17).

9.3 Fabric Dependence of the Material Coefficients in the Coupled Wave Equations

The form of the functional dependence of the drained elasticity tensor C^{d}_{ijkm} upon fabric is that given by (7.38) where the superscript c on the coefficients in (7.38) is replaced by double superscript, *cd*, the *c* indicating the elasticity coefficients and the *d* indicating that they are the drained coefficients. All of these coefficients are scalar-valued functions of ϕ, II, and III. Recalling that the Biot effective stress coefficient tensor A_{ij} is related to the difference between effective drained elastic constants C^{d}_{ijkm} and the solid matrix material elastic compliance tensor S^{m}_{ijkm} by the formula (8.3), where C^{d}_{ijkm} is expressed in terms of the fabric tensor by (7.38) above and S^{m}_{ijkm} is not a function of the fabric tensor because it represents the elastic constants of the matrix material. Recall also that the result (7.39) was based on the assumption that the matrix material is isotropic and that the anisotropy of the solid porous material is determined by the fabric tensor, thus we express the isotropic form of S^{m}_{ijkm} in terms of the bulk modulus and the shear modulus, K^{m} and G, respectively:

$$S^{\mathrm{m}}_{ijkm} = \frac{1}{2G}\left(\delta_{ik}\delta_{jm} - \frac{1}{3}\delta_{ij}\delta_{km}\right) + \frac{1}{9K^{\mathrm{m}}}\delta_{ij}\delta_{km}. \tag{9.18}$$

The form of S^{m}_{ijkm} that appears in (8.3) is S^{m}_{kmqq} and it is given by (9.18) as

$$S^{\mathrm{m}}_{kmqq} = \frac{1}{3K^{\mathrm{m}}}\delta_{km}. \tag{9.19}$$

Substituting (9.19) and (7.38) into (8.3) and simplifying, one finds that the Biot effective stress coefficient tensor A_{ij} is related to the fabric tensor $\mathbf{F}$ by

$$A_{ij} = \delta_{ij} - \frac{1}{3K^{\mathrm{m}}}\{a^{cd}_{o}\delta_{ij} + a^{cd}_{\mathrm{I}}F_{ij} + a^{cd}_{\mathrm{II}}F_{iq}F_{qj}\}, \tag{9.20}$$

where

$$\begin{aligned} a^{cd}_{\mathrm{o}} &= 3a^{cd}_{1} + a^{cd}_{2} + a^{cd}_{3}(1 - 2\mathrm{II}) + 2c^{cd}_{1}, \\ a^{cd}_{\mathrm{I}} &= 3a^{cd}_{2} + b^{cd}_{1} + b^{cd}_{2}(1 - 2\mathrm{II}) + 4c^{cd}_{2}, \\ a^{cd}_{\mathrm{II}} &= 3a^{cd}_{3} + b^{cd}_{2} + b^{cd}_{3}(1 - 2\mathrm{II}) + 4c^{\mathrm{cd}}_{3} \end{aligned} \tag{9.21}$$

and where II is the second invariant of $\mathbf{F}$. Biot's parameters M_{ij} and Z_{ijkm} are related to $C^{\mathrm{d}}_{\mathrm{eff}}$, A_{ij}, and C^{d}_{ijkm} above by (9.8). Formulas relating M_{ij} and Z_{ijkm} directly to the fabric tensor $\mathbf{F}$ will now be obtained by using the formula (7.38) above expressing C^{d}_{ijkm} in terms of the fabric tensor and the expression (9.20) relating A_{ij} to fabric, thus

$$M_{ij} = M\delta_{ij} - \frac{M}{3K^{\mathrm{m}}}\{a_o^{cd}\delta_{ij} + a_{\mathrm{I}}^{cd}F_{ij} + a_{\mathrm{II}}^{cd}F_{iq}F_{qj}\} \tag{9.22}$$

and

$$\begin{aligned} Z_{ijkm} = &\left(a_1^{cd} + M\frac{(3K^{\mathrm{m}} - a_o^{\mathrm{d}})^2}{(3K^{\mathrm{m}})^2}\right)\delta_{ij}\delta_{km} + \left(a_2^{cd} - \frac{M(3K^{\mathrm{m}} - a_o^{\mathrm{d}})a_{\mathrm{I}}^{\mathrm{d}}}{(3K^{\mathrm{m}})^2}\right)(F_{ij}\delta_{km} + \delta_{ij}F_{km}) \\ &+\left(a_3^{cd} - M\frac{(3K^{\mathrm{m}} - a_o^{cd})a_{\mathrm{II}}^{cd}}{(3K^{\mathrm{m}})^2}\right)(\delta_{ij}F_{kq}F_{qm} + \delta_{km}F_{iq}F_{qj}) + \left(b_1^{cd} + M\frac{(a_{\mathrm{I}}^{cd})^2}{(3K^{\mathrm{m}})^2}\right)F_{ij}F_{km} \\ &+\left(b_2^{cd} + M\frac{a_{\mathrm{I}}^{cd}a_{\mathrm{II}}^{cd}}{(3K^{\mathrm{m}})^2}\right)(F_{ij}F_{kq}F_{qm} + F_{km}F_{iq}F_{qj}) + \left(b_3^{cd} + M\frac{(a_{\mathrm{II}}^{cd})^2}{(3K^{\mathrm{m}})^2}\right)F_{is}F_{sj}F_{kq}F_{qm} \\ &+c_1^{cd}(\delta_{ki}\delta_{mj} + \delta_{mi}\delta_{kj}) + c_2^{cd}(F_{ki}\delta_{mj} + F_{kj}\delta_{mi} + F_{im}\delta_{kj} + F_{mj}\delta_{ki}) \\ &+c_3^{cd}(F_{ir}F_{rk}\delta_{mj} + F_{kr}F_{rj}\delta_{mi} + F_{ir}F_{rm}\delta_{kj} + F_{mr}F_{rj}\delta_{ik}). \end{aligned} \tag{9.23}$$

Problem

9.3.1. Calculate $\mathrm{tr}\mathbf{A}^2$, where the Biot effective stress coefficient tensor $\mathbf{A}$ is given by (9.20).

9.3.2. Derive the formulas (9.22) and (9.23) relating the Biot's parameters M_{ij} and Z_{ijkm} to fabric.

9.4 Plane Waves

The propagation of a harmonic plane wave is represented kinematically by a direction of propagation, denoted by $\mathbf{n}$ which is a unit normal to the wave front, and $\mathbf{a}$ or $\mathbf{b}$, which are the directions of displacement for the wave fronts associated with $\mathbf{u}$ and $\mathbf{w}$, respectively. These two harmonic plane waves are represented by

$$\mathbf{u}(\mathbf{x},t) = \mathbf{a}\exp\left[\mathrm{i}\omega\left(\frac{\mathbf{n}\cdot\mathbf{x}}{v} - t\right)\right], \quad \mathbf{w}(\mathbf{x},t) = \mathbf{b}\exp\left[\mathrm{i}\omega\left(\frac{\mathbf{n}\cdot\mathbf{x}}{v} - t\right)\right], \tag{9.24}$$

where v is the wave velocity in the direction $\mathbf{n}$, $\mathbf{x}$ is the position vector, ω is the frequency and t is time. The slowness vector $\mathbf{s}$ is defined as $\mathbf{s} = (1/v)\mathbf{n}$, and the wave speed v may be complex. As in elastic solid wave propagation (see Example 6.3.4), a transverse wave is characterized by $\mathbf{a}\cdot\mathbf{n} = 0$, a longitudinal wave by $\mathbf{a}\cdot\mathbf{n} = 1$. Substituting the representations (9.24) for the plane waves into the field equations (9.16) and (9.17) one obtains equations that are in Biot (1962a, b) and Sharma (2005, 2008) and many other places,

$$(Q_{ik} - \rho\delta_{ik}v^2)a_k + (C_{ik} - \rho_f\delta_{ik}v^2)b_k = 0, \tag{9.25}$$

$$(C_{ki} - \rho_f\delta_{ik}v^2)a_k + \left(Mn_kn_i - \rho_f J_{ik}v^2 - \frac{i\mu}{\omega}R_{ik}v^2\right)b_k = 0, \tag{9.26}$$

where the following notation has been introduced:

$$Q_{ik} = Z^{\mathrm{d}}_{ijkm}n_m n_j, \quad C_{ik} = M_{ij}n_j n_k. \tag{9.27}$$

Q is the acoustical tensor from elastic wave propagation and **C** represents the interaction of the displacement fields **u** and **w**. Rewritten in matrix notation equations (9.25) and (9.26) take the form

$$(\mathbf{Q} - \rho v^2\mathbf{1}) \cdot \mathbf{a} + (\mathbf{C} - \rho_f v^2\mathbf{1}) \cdot \mathbf{b} = 0, \tag{9.28}$$

$$(\mathbf{C}^{\mathrm{T}} - \rho_f v^2\mathbf{1}) \cdot \mathbf{a} + \left(M\mathbf{n} \otimes \mathbf{n} - \left(\rho_f\mathbf{J} + \frac{i\mu}{\omega}\mathbf{R}\right)v^2\right) \cdot \mathbf{b} = 0. \tag{9.29}$$

These equations represent an eigenvalue problem, the squares of the wave speeds v^2 representing the eigenvalues and the vectors **a** and **b** representing the eigenvectors. Rewriting (9.28) and (9.29) as a scalar 6 by 6 matrix formed from the four 3 by 3 matrices that appear in (9.28) and (9.29) and also representing the two 3D vectors **a** and **b** as one 6D vector, the following representation is obtained:

$$\begin{bmatrix} \mathbf{Q} - \rho v^2\mathbf{1} & \mathbf{C} - \rho_f v^2\mathbf{1} \\ \mathbf{C}^{\mathrm{T}} - \rho_f v^2\mathbf{1} & M\mathbf{n} \otimes \mathbf{n} - \left(\rho_f\mathbf{J} + \frac{i\mu}{\omega}\mathbf{R}\right)v^2 \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = 0. \tag{9.30}$$

Please note that the scalar 6 by 6 matrix operating on the vectors **a** and **b** is symmetric; the 3 by 3 matrices along the diagonal are symmetric and, even though **C** is not symmetric, having the transpose of **C** is the lower left 3 by 3 matrix and the 3 by 3 matrix **C** itself in the upper right makes the 6 by 6 matrix symmetric. Since the right hand side of this linear system of equations is a zero 6D vector, it follows from Cramer's rule that, in order to avoid the trivial solution, it is necessary to set the determinant of the 6 by 6 matrix equal to zero, thus

$$\begin{vmatrix} \mathbf{Q} - \rho v^2\mathbf{1} & \mathbf{C} - \rho_f v^2\mathbf{1} \\ \mathbf{C}^{\mathrm{T}} - \rho_f v^2\mathbf{1} & M\mathbf{n} \otimes \mathbf{n} - \left(\rho_f\mathbf{J} + \frac{i\mu}{\omega}\mathbf{R}\right)v^2 \end{vmatrix} = 0. \tag{9.31}$$

This condition will provide six (four nonzero) values of the possible squares of the wave speeds v^2 in the direction **n**. In each direction there will be four nontrivial wave speeds, two representing shear waves and one each representing the Biot fast and slow waves. For each value of a squared wave speed v^2 substituted back into (9.30), two 3D vectors **a** and **b** will be determined subject to the condition that they are both unit vectors. The ease with which these calculations are described does not

convey their algebraic complexity. However contemporary symbolic algebraic software solves this type of eigenvalue problem in a few keystrokes.

Problems

9.4.1. Using the indicial notation substitute the representations (9.24) for the plane waves into the field equations (9.16) and (9.17) and derive the coupled equations (9.28) and (9.29) for the velocity and amplitudes of the two waves in the fixed direction **n**.

9.4.2. How do you use your favorite symbolic algebraic software to solve (9.30) in the case where the parameters are all specified numerically.

9.5 Fabric Dependence of the Tensors Q, C, J, and R

The governing equations for anisotropic poroelasicity for quasi-static and dynamic poroelasticity were developed and extended to include the dependence of the constitutive relations upon a pore structure fabric tensor **F** as well as the porosity (Cowin 1985, 2003, 2004; Cowin and Cardoso 2011; Cardoso and Cowin 2011, 2012). Formulas relating the acoustic tensor **Q**, the flow-resistivity tensor **R** and the tensor **C**, representing the interaction of the velocity fields **u** and **w**, to the fabric tensor **F** are obtained in this subsection. The dependence of the elastic acoustic tensor **Q** upon the fabric tensor **F** is obtained by substituting (9.23) into the first of (9.27),

$$\begin{aligned}\mathbf{Q} = &(c_1^{cd} + c_2^{cd}\mathrm{tr}\{\mathbf{F}\cdot\mathbf{n}\otimes\mathbf{n}\} + c_3^{cd}\mathrm{tr}\{\mathbf{F}^2\cdot\mathbf{n}\otimes\mathbf{n}\})\mathbf{1}\\ &+q_1\mathbf{n}\otimes\mathbf{n} + c_2^{cd}\mathbf{F} + q_2(\mathbf{F}\cdot\mathbf{n}\otimes\mathbf{n} + \mathbf{n}\otimes\mathbf{n}\cdot\mathbf{F})\\ &+c_3^{cd}\mathbf{F}^2 + q_3(\mathbf{F}^2\cdot\mathbf{n}\otimes\mathbf{n} + \mathbf{n}\otimes\mathbf{n}\cdot\mathbf{F}^2) + q_4\mathbf{F}\cdot\mathbf{n}\otimes\mathbf{F}\cdot\mathbf{n}\\ &+q_5(\mathbf{F}\cdot\mathbf{n}\otimes\mathbf{F}^2\cdot\mathbf{n} + \mathbf{F}^2\cdot\mathbf{n}\otimes\mathbf{F}\cdot\mathbf{n}) + q_6\mathbf{F}^2\cdot\mathbf{n}\otimes\mathbf{F}^2\cdot\mathbf{n},\end{aligned} \tag{9.32}$$

where

$$\begin{aligned}&q_1 = c_1^{cd} + a_1^{cd} + M\frac{(3K^{\mathrm{m}} - a_o^{\mathrm{d}})^2}{(3K^{\mathrm{m}})^2}, \quad q_2 = c_2^{cd} + a_2^{cd} - \frac{M(3K^{\mathrm{m}} - a_o^{\mathrm{d}})a_{\mathrm{I}}^{\mathrm{d}}}{(3K^{\mathrm{m}})^2},\\ &q_3 = c_3^{cd} + a_3^{cd} - M\frac{(3K^{\mathrm{m}} - a_o^{\mathrm{d}})a_{\mathrm{II}}^{\mathrm{d}}}{(3K^{\mathrm{m}})^2}, \quad q_4 = b_1^{cd} + M\frac{(a_{\mathrm{I}}^{\mathrm{d}})^2}{(3K^{\mathrm{m}})^2},\\ &q_5 = b_2^{cd} + M\frac{a_{\mathrm{I}}^{\mathrm{d}}a_{\mathrm{II}}^{\mathrm{d}}}{(3K^{\mathrm{m}})^2}, \quad q_6 = b_3^{cd} + M\frac{(a_{\mathrm{II}}^{\mathrm{d}})^2}{(3K^{\mathrm{m}})^2}.\end{aligned} \tag{9.33}$$

The six quantities defined in (9.33) are scalar-valued functions of ϕ, II, and III.

The formula for the tensor **C** is obtained by substituting (9.8) into the second of (9.27) and then employing (9.20), thus

$$[\mathbf{C} = M\{\mathbf{n} - \frac{1}{3K^m}\{a_0^{cd}\mathbf{n} + a_I^{cd}\mathbf{F}\cdot\mathbf{n} + a_{II}^{cd}\mathbf{F}^2\cdot\mathbf{n}\}\otimes\mathbf{n}. \tag{9.34}$$

by

$$\mathbf{C} = (\mathbf{M} \cdot \{\mathbf{n} - \frac{1}{3K^m}\{a_0^{cd}\mathbf{n} + a_I^{cd}\mathbf{F} \cdot \mathbf{n} + a_{II}^{cd}\mathbf{F}^2 \cdot \mathbf{n}\}) \otimes \mathbf{n}. \tag{9.34}$$

The micro–macro velocity average tensor $\mathbf{J}$ is related to the fabric by

$$\mathbf{J} = \mathbf{Q} \cdot (j_1\mathbf{1} + j_2\mathbf{F} + j_3\mathbf{F}^2), \tag{9.35}$$

where $\mathbf{Q}$ represents a rotation matrix associated with the transformation between the principal axes of $\mathbf{F}$ and the reference coordinate system used for $\mathbf{J}$ and j_1, j_2, and j_3 are functions of ϕ, II and III. Similarly, The flow-resistivity tensor $\mathbf{R}$, is related to the fabric by

$$R = r_1\mathbf{1} + r_2\mathbf{F} + r_3\mathbf{F}^2, \tag{9.36}$$

where r_1, r_2, and r_3 are functions of ϕ, II and III, and $\mathbf{R}$ is equivalent to the inverse of the second-rank intrinsic permeability tensor $\mathbf{K}$.

Problems

9.5.1. Express the direct notation equation (9.32) for $\mathbf{Q}$ in the indicial notation.
9.5.2. Express the equation (9.32) for $\mathbf{Q}$ in the matrix notation in the principal coordinate system of the fabric tensor.

9.6 Propagation of Waves in a Principal Direction of Material Symmetry

In this section the solution is developed for waves that propagate in the direction of a principal axis of material symmetry. The direction of propagation is selected to be the one direction, thus $\mathbf{n}$ are given by the vector $\mathbf{n} = [1, 0, 0]^{\mathrm{T}}$. The solution to the problem is the solution of the 6 by 6 system of equations given by (9.23), thus the values of the tensors $\mathbf{J}$, $\mathbf{R}$, $\mathbf{C}$, and $\mathbf{Q}$ in the coordinate system of the principal axes of material symmetry and at the vector $\mathbf{n} = [1, 0, 0]^{\mathrm{T}}$ are determined first. Under these conditions $\mathbf{J}$ is determined from (9.35) with $\mathbf{Q} = \mathbf{1}$, $\mathbf{R}$ from (9.36), $\mathbf{C}$ from (9.34) and $\mathbf{Q}$ from (9.32) are given by

$$\mathbf{J} = \begin{bmatrix} J_{11} & 0 & 0 \\ 0 & J_{22} & 0 \\ 0 & 0 & J_{33} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} R_{11} & 0 & 0 \\ 0 & R_{22} & 0 \\ 0 & 0 & R_{33} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} C_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{9.37}$$

and

$$\mathbf{Q} = \begin{bmatrix} Q_{11} & 0 & 0 \\ 0 & Q_{22} & 0 \\ 0 & 0 & Q_{33} \end{bmatrix}, \tag{9.38}$$

In the coordinate system of the principal axes of material symmetry and at the vector $\mathbf{n} = \{1, 0, 0\}$, the four 3 by 3 submatrices that form the 6 by 6 matrix in equation (9.30) are given by

$$\mathbf{Q} - \rho v^2 \mathbf{1} = \begin{bmatrix} Q_{11} - \rho v^2 & 0 & 0 \\ 0 & Q_{22} - \rho v^2 & 0 \\ 0 & 0 & Q_{33} - \rho v^2 \end{bmatrix}, \tag{9.39}$$

$$\mathbf{C} - \rho_{\mathrm{f}} v^2 \mathbf{1} = \mathbf{C}^{\mathrm{T}} - \rho_{\mathrm{f}} v^2 \mathbf{1} = \begin{bmatrix} C_{11} - \rho_{\mathrm{f}} v^2 & 0 & 0 \\ 0 & -\rho_{\mathrm{f}} v^2 & 0 \\ 0 & 0 & -\rho_{\mathrm{f}} v^2 \end{bmatrix}, \tag{9.40}$$

$$\begin{aligned} &M\mathbf{n} \otimes \mathbf{n} - \left(\rho_{\mathrm{f}} \mathbf{J} + \frac{i\mu}{\omega} \mathbf{R}\right) v^2 \\ &= \begin{bmatrix} M - \rho_{\mathrm{f}} v^2 J_{11} - \frac{i\mu v^2}{\omega} R_{11} & 0 & 0 \\ 0 & -\rho_{\mathrm{f}} v^2 J_{22} - \frac{i\mu v^2}{\omega} R_{22} & 0 \\ 0 & 0 & -\rho_{\mathrm{f}} v^2 J_{33} - \frac{i\mu v^2}{\omega} R_{33} \end{bmatrix}. \end{aligned} \tag{9.41}$$

Substitution of the four 3 by 3 matrices above into the 6 by 6 determinant (9.31) reveals that result may be expressed as three 2 by 2 matrices for the three sets of components, $\{a_1, b_1\}$, $\{a_2, b_2\}$, and $\{a_3, b_3\}$;

$$\begin{aligned} \begin{bmatrix} Q_{11} - \rho v^2 & C_{11} - \rho_{\mathrm{f}} v^2 \\ C_{11} - \rho_{\mathrm{f}} v^2 & M - \left(\rho_{\mathrm{f}} J_{11} + \frac{i\mu}{\omega} R_{11}\right) v^2 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = 0, \\ \begin{bmatrix} Q_{22} - \rho v^2 & -\rho_{\mathrm{f}} v^2 \\ -\rho_{\mathrm{f}} v^2 & -\left(\rho_{\mathrm{f}} J_{22} + \frac{i\mu}{\omega} R_{22}\right) v^2 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = 0, \\ \begin{bmatrix} Q_{33} - \rho v^2 & -\rho_{\mathrm{f}} v^2 \\ -\rho_{\mathrm{f}} v^2 & -\left(\rho_{\mathrm{f}} J_{33} + \frac{i\mu}{\omega} R_{33}\right) v^2 \end{bmatrix} \begin{bmatrix} a_3 \\ b_3 \end{bmatrix} = 0. \end{aligned} \tag{9.42}$$

Requiring that the determinants of these 2 by 2 matrices vanish yields four nontrivial solutions for the squared wave speed v^2. The vanishing of the first of the determinants of these 2 by 2 matrices provides two roots of a quadratic equation that represent the fast and the slow squared longitudinal wave speeds,

$$v^2 = \frac{N}{2\rho_{\mathrm{f}} L_{11}} \pm \sqrt{\left[\frac{N}{2\rho_{\mathrm{f}} L_{11}}\right]^2 + \frac{\rho_{\mathrm{f}}(C_{11}^2 - MQ_{11})}{\rho_{\mathrm{f}} L_{11}}}, \tag{9.43}$$

where

$$L_{ij} = \rho J_{ik} + \frac{i\mu}{\rho_{\mathrm{f}}\omega} R_{ik} - \rho_{\mathrm{f}}\delta_{ij}, \quad N = M\rho + \rho_{\mathrm{f}}\left(Q_{11}\left(J_{11} + \frac{i\mu}{\rho_{\mathrm{f}}\omega}R_{11}\right) - 2C_{11}\right), \tag{9.44}$$

or

$$v^2 = v_o^2 \pm c_o^2, \tag{9.45}$$

where the notation

$$v_o^2 = \frac{N}{2\rho_{\mathrm{f}}L_{11}}, \quad c_o^2 = \sqrt{\left[\frac{N}{2\rho_{\mathrm{f}}L_{11}}\right]^2 + \frac{\rho_{\mathrm{f}}(C_{11}^2 - MQ_{11})}{\rho_{\mathrm{f}}L_{11}}} \tag{9.46}$$

has been employed. The vanishing of the second and third of the determinants of the 2 by 2 matrices in (9.42) provide a zero root and a nonzero root from each determinant. The two nonzero roots are the squared shear wave speeds

$$v^2 = \frac{Q_{22}}{L_{22}}\left(J_{22} + \frac{i\mu}{\rho_{\mathrm{f}}\omega}R_{22}\right) \quad \text{and} \quad v^2 = \frac{Q_{33}}{L_{33}}\left(J_{33} + \frac{i\mu}{\rho_{\mathrm{f}}\omega}R_{33}\right). \tag{9.47}$$

The vectors **a** and **b** for the fast and slow waves are given by

$$\mathbf{a} = \{a_1, 0, 0\}, \quad \mathbf{b} = \{b_1, 0, 0\}, \tag{9.48}$$

where a_1 and b_1 are related by the following two equivalent expressions for the fast wave

$$a_1 = \frac{C_{11} - \rho_{\mathrm{f}}(v_o^2 + c_o^2)}{\rho(v_o^2 + c_o^2) - Q_{11}}, \quad b_1 = \frac{M - (\rho_{\mathrm{f}}J_{11} + (i\mu/\omega)R_{11})(v_o^2 + c_o^2)}{\rho_{\mathrm{f}}(v_o^2 + c_o^2) - C_{11}}b_1 \tag{9.49}$$

and the next two equivalent expressions for the slow wave,

$$a_1 = \frac{C_{11} - \rho_{\mathrm{f}}(v_o^2 - c_o^2)}{\rho(v_o^2 - c_o^2) - Q_{11}}, \quad b_1 = \frac{M - (\rho_{\mathrm{f}}J_{11} + (i\mu/\omega)R_{11})(v_o^2 - c_o^2)}{\rho_{\mathrm{f}}(v_o^2 - c_o^2) - C_{11}}b_1 \tag{9.50}$$

and for the two shear waves by

$$\begin{aligned} &\mathbf{a} = \{0, a_2, 0\}, \quad \mathbf{b} = \{0, b_2, 0\}, \quad a_2 = -\left(J_{22} + \frac{i\mu}{\omega}R_{22}\right)b_2 \quad \text{and} \\ &\mathbf{a} = \{0, 0, a_3\}, \quad \mathbf{b} = \{0, 0, b_3\}, \quad a_3 = -\left(J_{33} + \frac{i\mu}{\omega}R_{33}\right)b_3, \end{aligned} \tag{9.51}$$

respectively.

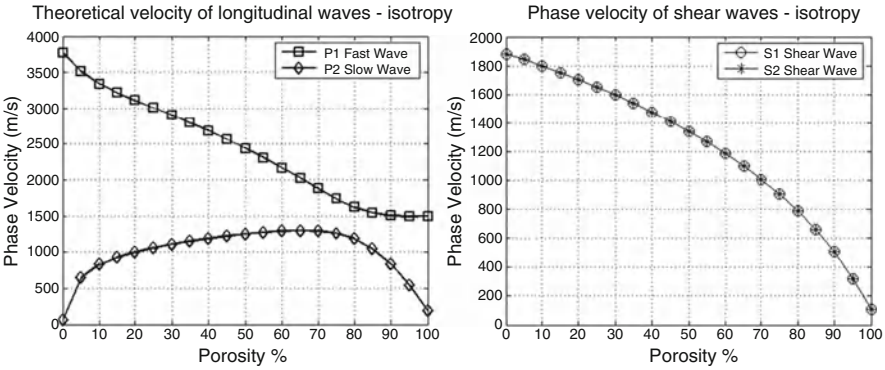


Fig. 9.2 A plot of the ultrasound wave velocities as a function of the porosity from the theoretical model. P1 is the fast wave, P2 the slow wave, S1 and S2 are the two propagating shear waves, while S3 and S4 are non-propagating shear waves with null velocity. These results are plotted for a medium with isotropic fabric. The velocity of propagation for all directions is the same

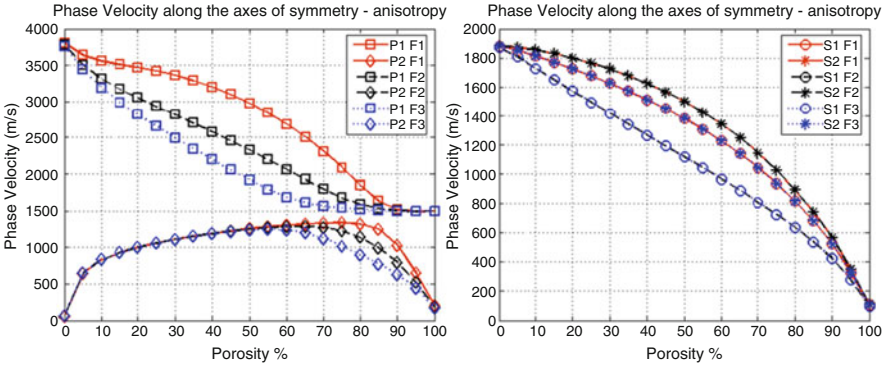


Fig. 9.3 A similar plot to Fig. 9.2, but for a medium with anisotropic fabric. Propagation along the three different principal directions is shown

The six-ultrasound wave velocities as a function of the porosity from the model are shown in Fig. 9.2 for an isotropic material. These six possible wave velocities are the six roots of (9.42). In the figure the fast wave is denoted by P1 (P for pressure), the slow wave by P2, S1, and S2 denote the two propagating shear waves (S for shear), while S3 and S4 are non-propagating shear waves with zero velocity. These results are plotted for a medium for the isotropic case, thus the velocity of propagation is the same for all directions. In Fig. 9.3, the six-ultrasound wave velocities are plotted again as a function of porosity, but for an anisotropic medium. Propagation along all three principal directions are shown for each wave as a solid, dashed, and dotted line. In these two figures one can see that the response of the compression or P waves is the most varied. The fast wave has a high velocity at low porosity and that velocity drops to the velocity of sound in water, about 1,500 m/s,

at very high porosity. The slow wave is at or near zero velocity at very high and very low porosities and climes to a peak near 60% porosity. The nonzero shear wave velocities are highest at the lowest porosity and decrease linearly to zero at the highest porosity.

The experimental identification of the fast and slow waves is illustrated in Figs. 9.4 and 9.5. Figure 9.4 contains four panel figures with plots of amplitude vs. time for a wave passing thought different media. Figure 9.5 contains four panel figures with plots the frequency spectrum vs. time for a wave passing thought the same four different media. For both figures, panels 1 and 4 the media in the container is water, but the placement of the emitting and receiving transducers is different. In panel 2 there is a fluid saturated porous specimen and in panel 3 there is the porous specimen but no water. In each figure these various combinations of test media and transducer setups are illustrated in the small cartoons to the left or right of each plot of amplitude vs. time. In each cartoon the emitting and receiving transducers are indicated as tubes (bottom and top of the cartoon) applied to the specimen in the water or air filled container. In panel 1 the shape of the wave or frequency spectrum is determined by its passage through the small space between transducers. In panel 2 the shape of the wave or frequency spectrum is determined by its passage through a fluid saturated porous specimen between transducers. In panel 3 there is no water, just the porous specimen, so the shape of the wave or frequency spectrum is determined by its passage through the unsaturated porous specimen between transducers. Panel 4 is the same as panel 1 except the space between the transducers is the same as if the specimen were there. Notice that, in this case, the shape of the wave or frequency spectrum is almost the same as in panel 1 but it is displaced to a greater time. It is only in the second panel that the waveform shape and frequency spectrum are determined by passage through the saturated porous specimen. In this panel one can see that the first part of the waveform is similar to that of the fast wave and the last part is similar to that of the slow wave. Panel 4 shows the passage of only the wave in the fluid because there is no porous specimen in the container; this situation is similar to that of the slow wave in the porous media shown in the second panel. Panel 3 shows the passage of a wave propagating within an unsaturated porous specimen; this wave is similar to the fast wave shown in the second panel. These data suggest that under the tested conditions of porosity, the propagation of the fast wave is mainly related to the solid phase of the medium, while the slow wave is characteristic of the fluid phase (Cardoso et al. 2003).

A plot of the fast (top three curves) and slow (lower three curves) wave speeds as a function of frequency for different degrees of anisotropy at a porosity of 50 % in cancellous bone tissue is shown in Fig. 9.6. In ultrasonic measuring systems the viscous effects of the pore fluid damp out the genesis of the slow wave and its potential observation below the critical frequency. The amplitude damping of both waves also occurs at frequencies above the viscous frequency, which is 10^4 times the critical frequency, making the observation of both waves above the viscous frequency challenging (Cardoso et al. 2008).

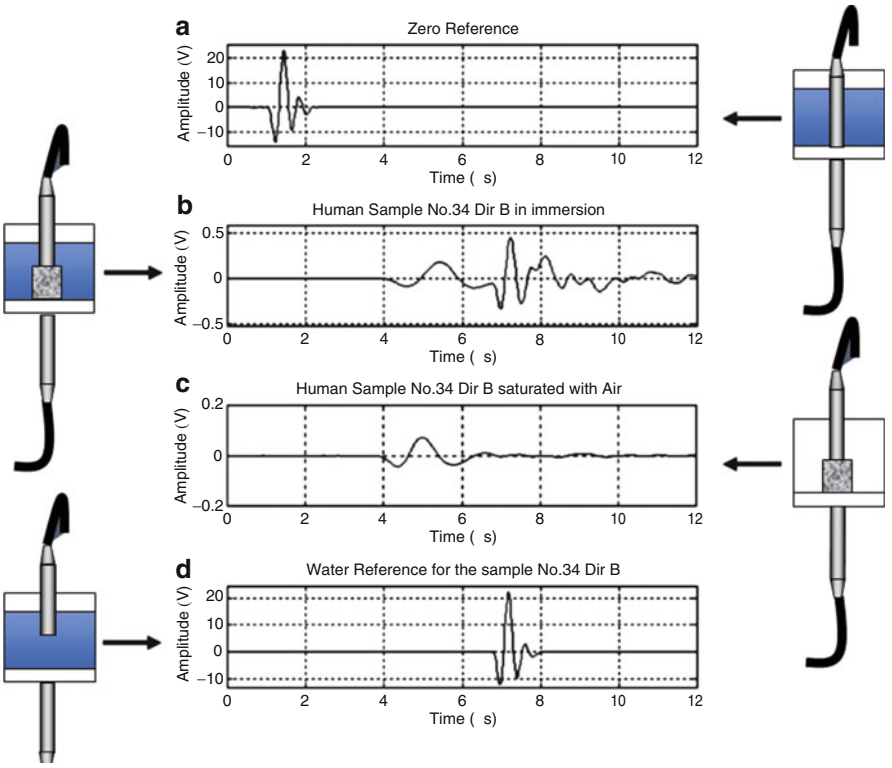


Fig. 9.4 Each panel of this four-panel figure contains a plot of amplitude vs. time for a wave passing thought different media. In panels 1 and 4 the media is water, but the placement of the emitting and receiving transducers is different. In panel 2 there is a fluid saturated porous specimen and in panel 3 there is the porous specimen but no water. These various combinations of test media and transducer setups are illustrated in the small cartoons to the left or right of each plot of amplitude vs. time. In each cartoon the emitting and receiving transducers are indicated as tubes (bottom and top of the cartoon) applied to the specimen in the water or air filled container. In panel 1 the signal detected corresponds to the wave at the bottom of the measurement cell. This is the ultrasound wave that would excite the porous media under the conditions shown in panels 2 and 3. In panel 2 the shape of the wave is determined by its passage through a fluid saturated porous specimen between transducers. In panel 3 there is no water, just the porous specimen, so the shape of the wave is determined by its passage through the unsaturated porous specimen between transducers. Panel 4 is the same as panel 1 except the space between the transducers is the same as if the specimen were there. Notice that, in this case, the shape of the wave is the same as in panel 1 but it is displaced to a greater time. It is only in panel 2 that the waveform shape is determined by the passage through the saturated porous specimen and one can see that the first part of the waveform is similar to that of the wave propagating in the unsaturated porous medium and the last part of the waveform is similar to that of the wave propagating in the fluid. Therefore, the waveform in panel 2 is shown to be composed by two waves; the first wave is identified as the fast wave of poroelastic wave propagation theory (shorter arrival time), and the second wave is the slow wave (larger arrival time). In this example, the propagation of the fast wave is closely related to the solid phase of the medium, while the slow wave is mostly related to the fluid saturating the pores

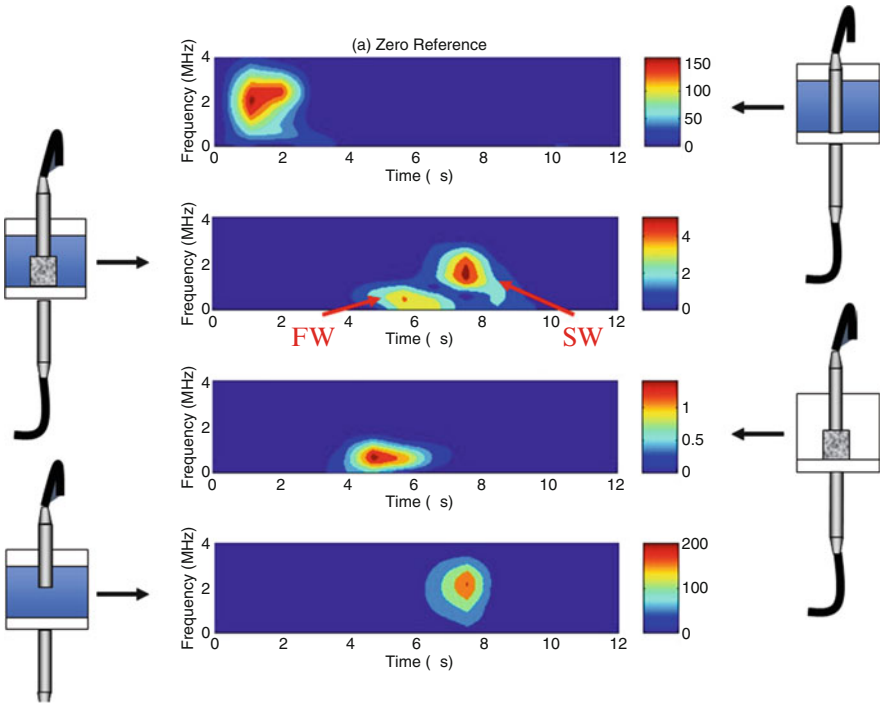


Fig. 9.5 The four panels in this figure represent different data from the same four experiments illustrated in Fig. 9.3, thus the description of the experiment and the small cartoons for each panel have the same significance. The difference is the plot of amplitude vs. time for a wave passing thought different media have been replaced by the frequency spectrum of waves vs. time for the wave passing thought the different media. The frequency spectrum for the wave in water alone, panels 1 and 4, is seen to be the same general shape; this shape is missing from panel 3 where there is no water, but is identified in panel 2 as the frequency spectrum representing the energy associated with the slow wave (SW). The frequency spectrum representing the energy associated with the fast wave (FW) is shown in panel 3 where the wave is passed through the unsaturated porous specimen

Example 9.6.1: Wave Propagation in a Principal Direction of Symmetry
Determine the wave velocities and the polarization vectors or eigenvectors for cancellous bone associated with a harmonic wave propagating along an axis of material symmetry. Let the axis of material symmetry be $\mathbf{e}_1$. The properties of the specified cancellous bone are a porosity $\phi = 0.5$, a density of the matrix material of the porous structure of 2,000 kg/m^3, and a density of the pore fluid of the porous structure of 1,000 kg/m^3. From Yang et al. (1999) the values of the compliance and elasticity tensors for orthotropic elastic coefficients are approximately given by

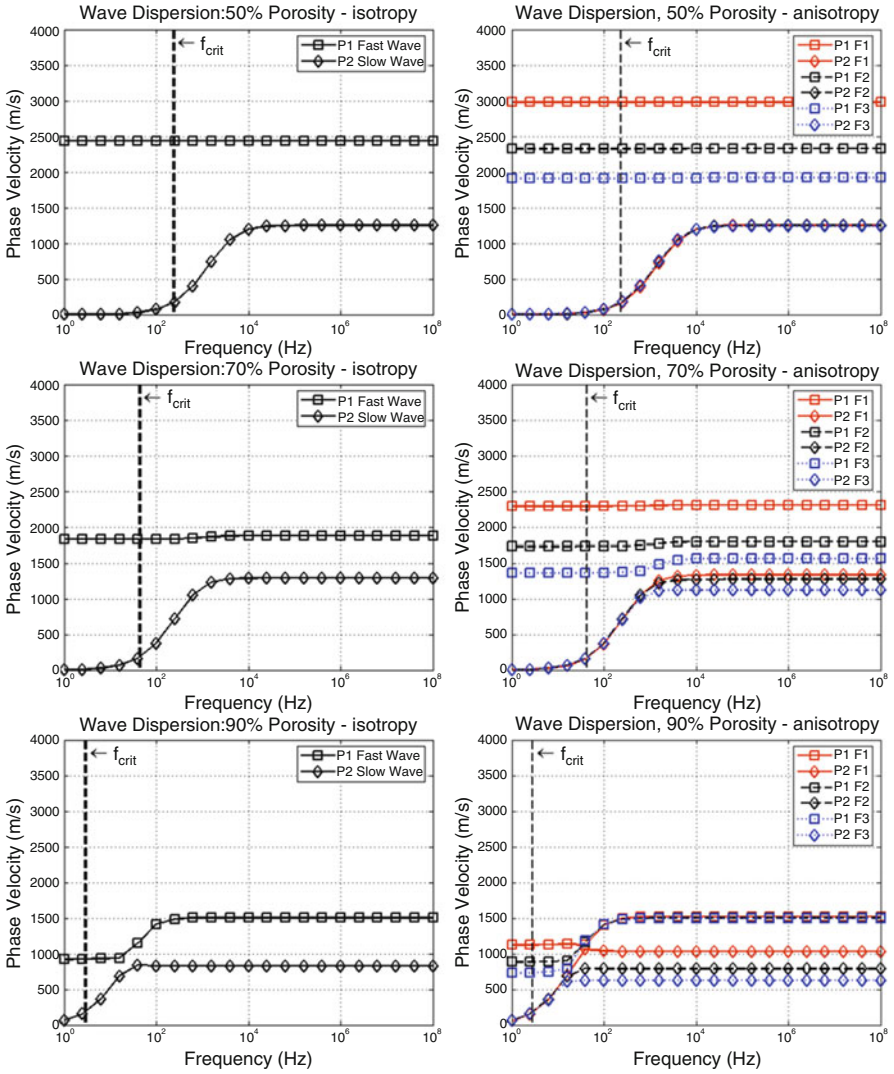


Fig. 9.6 A plot the fast (top three curves) and slow (lower three curves) wave speeds as a function of frequency for different degrees of anisotropy at a porosities of 50, 70 and 90 % in cancellous bone tissue. In ultrasonic measuring systems the viscous effects of the pore fluids damp out potential observations of these waves below the critical frequency and the same damping occurs at frequencies above the viscous frequency, which is 10^4 times the critical frequency

$$\hat{\mathbf{S}}^{\mathrm{d}} = \frac{1}{E_{\mathrm{t}}(1-\phi)^2}\begin{bmatrix} \frac{1}{1240} & \frac{-1}{7045} & \frac{-1}{3924} & 0 & 0 & 0 \\ \frac{-1}{7045} & \frac{1}{885} & \frac{-1}{3457} & 0 & 0 & 0 \\ \frac{-1}{3924} & \frac{-1}{3457} & \frac{1}{528.8} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5333.3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{633.3} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{927.6} \end{bmatrix} \tag{9.52}$$

and

$$\hat{\mathbf{C}}^{\mathrm{d}} = E_{\mathrm{t}}(1-\phi)^2\begin{bmatrix} 1348 & 225 & 216 & 0 & 0 & 0 \\ 225 & 958 & 177 & 0 & 0 & 0 \\ 216 & 177 & 585 & 0 & 0 & 0 \\ 0 & 0 & 0 & 533.3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 633.3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 927.6 \end{bmatrix}, \tag{9.53}$$

where we take $E_{\mathrm{t}} = 10$ MPa. The value of the fluid viscosity μ is 1,000 Pa s. The flow-resistivity tensor $\mathbf{R}$ and the micro–macro velocity tensor $\mathbf{J}$ are assigned the following values to simplify the calculation:

$$R = \mathbf{0}, \quad \mathbf{J} = \phi^{-1}\mathbf{1} = 2\mathbf{1}. \tag{9.54}$$

Solution: First note that the values of the matrices $\hat{S}^{\mathrm{m}}$ and $\hat{\mathbf{C}}^{\mathrm{m}}$ are given by the formulas for $\hat{\mathbf{S}}^{\mathrm{d}}$ and $\hat{\mathbf{C}}^{\mathrm{d}}$ when the porosity is zero, thus

$$\hat{\mathbf{C}}^{\mathrm{d}} \cdot \hat{\mathbf{S}}^{\mathrm{m}} = (1-\phi)^2\mathbf{1}$$

and it follows from (9.2), $\hat{\mathbf{A}} = (\hat{\mathbf{1}} - \hat{\mathbf{C}}^{\mathrm{d}} \cdot \hat{\mathbf{S}}^{\mathrm{m}}) \cdot \hat{\mathbf{U}}$, that

$$\hat{\mathbf{A}} = \phi(2-\phi)\hat{\mathbf{U}}.$$

To calculate the value of the parameter Λ specified by (8.18), $= C^d_{eff} - \hat{A} \cdot \hat{S} \cdot \hat{A}$, we first evaluate $\hat{A} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{A}$ and C^d_{eff}. Using $\hat{A} = \phi(2-\phi)\hat{U}$, it follows that

$$\hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{A}} = \phi^2(2-\phi)^2\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}} = \frac{\phi^2(2-\phi)^2}{K^d_{eff}}.$$

To evaluate K^d_{Reff} and K^m_{Reff} in the formula (8.18) for C^d_{eff} we note that

$$K^d_{Reff} = (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{\mathbf{U}})^{-1}, \quad K^m_{Reff} = (\hat{\mathbf{U}} \cdot \hat{\mathbf{S}}^{\mathrm{m}} \cdot \hat{\mathbf{U}})^{-1}$$

thus it follows that for the drained elastic constants,

$$K^{\mathrm{d}}_{\mathrm{Reff}} = 407 E_{\mathrm{t}}(1-\phi)^2 = 1.018 \text{ GPa}$$

and the result for the matrix material is obtained by setting $\phi = 0$,

$$K^{\mathrm{m}}_{\mathrm{Reff}} = 407 E_{\mathrm{t}} = 4.073 \text{ GPa}.$$

Then from (8.17), noting that value of the bulk fluid modulus K^{f} is 2.25 GPa, C^d_{eff} is given by

$$C^d_{eff} = \frac{1}{K^d_{Reff}} - \frac{1}{K^{\mathrm{m}}_{Reff}} + \phi\left(\frac{1}{K^f} - \frac{1}{K^m_{Reff}}\right) = \frac{0.836}{\text{GPa}}$$

and the value of Λ is then

$$\Lambda = C^d_{eff} - \hat{\mathbf{A}} \cdot \hat{\mathbf{S}}^{\mathrm{d}} \cdot \hat{A} = \frac{0.2836}{\text{GPa}}.$$

Using the values of the fluid density ρ^{f} (1,000 kg/m^3) and the solid density ρ^{s} (2,000 kg/m^3), ρ given by (9.11) has the value

$$\rho = \{2(1-\phi) + \phi\} 1,000(\mathrm{kg/m}^3) = 1,500(\mathrm{kg/m}^3).$$

These parameter values, in addition to the formulas above and the observation from (9.4) that the scalar M is simply ϕ times the inverse of the parameter Λ,

$$M = \frac{\phi}{\Lambda} = \frac{\text{GPa}}{0.5672} = 1.76 \text{ GPa}.$$

Thus from (9.8)

$$\hat{\mathbf{M}} = M\hat{\mathbf{A}} = M\phi(2-\phi) = \hat{\mathbf{U}} = 1.76 \text{ GPa}\left(\frac{3}{4}\right)\hat{\mathbf{U}} = 1.322 \text{ GPa } \hat{\mathbf{U}}.$$

The results above will first be specialized to the propagation of waves in a principal direction of material symmetry $\mathbf{n} = [1, 0, 0]^{\mathrm{T}}$ and then in a direction that

will generate quasi-waves, $\mathbf{n} = (1/\sqrt{3})[1, 1, 1]^{\mathrm{T}}$. Recall from (9.8) the formula for $\hat{Z}^d$, thus

$$\hat{\mathbf{Z}}^d = \hat{\mathbf{C}}^{\mathrm{d}} + M(\hat{\mathbf{A}} \otimes \hat{\mathbf{A}}) = \hat{\mathbf{C}}^{\mathrm{d}} + 0.992 \text{ GPa } \hat{\mathbf{U}} \otimes \hat{\mathbf{U}}.$$

Proceed with the first case, $\mathbf{n} = [1,0,0]^{\mathrm{T}}$, from (9.27) we note that for $\mathbf{n} = [1,0,0]^{\mathrm{T}}$,

$$Q_{ik} = Z^{\mathrm{d}}_{i1k1}, \quad C_{ik} = M_{i1}n_k \quad or$$

$$\mathbf{Q} = \begin{bmatrix} Z^{\mathrm{d}}_{1111} & Z^{\mathrm{d}}_{1121} & Z^{\mathrm{d}}_{1131} \\ Z^{\mathrm{d}}_{1121} & Z^{\mathrm{d}}_{2121} & Z^{\mathrm{d}}_{2131} \\ Z^{\mathrm{d}}_{1131} & Z^{\mathrm{d}}_{2131} & Z^{\mathrm{d}}_{3131} \end{bmatrix} = \begin{bmatrix} \hat{Z}^{\mathrm{d}}_{11} & \hat{Z}^{\mathrm{d}}_{16} & \hat{Z}^{\mathrm{d}}_{15} \\ \hat{Z}^{\mathrm{d}}_{16} & \hat{Z}^{\mathrm{d}}_{66} & \hat{Z}^{\mathrm{d}}_{56} \\ \hat{Z}^{\mathrm{d}}_{15} & \hat{Z}^{\mathrm{d}}_{56} & \hat{Z}^{\mathrm{d}}_{55} \end{bmatrix} = \begin{bmatrix} \hat{Z}^{\mathrm{d}}_{11} & 0 & 0 \\ 0 & \hat{Z}^{\mathrm{d}}_{66} & 0 \\ 0 & 0 & \hat{Z}^{\mathrm{d}}_{55} \end{bmatrix}$$

$$= \begin{bmatrix} 4.32 & 0 & 0 \\ 0 & 2.32 & 0 \\ 0 & 0 & 1.58 \end{bmatrix} \text{GPa},$$

$$C = \begin{bmatrix} M_{11} & 0 & 0 \\ M_{21} & 0 & 0 \\ M_{31} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1.322 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{GPa}.$$

The solution to the problem is the solution of the 6 by 6 system of equations given by (9.30), thus the values of the tensors **J**, **R**, **C**, and **Q** in the coordinate system of the principal axes of material symmetry and at the vector $\mathbf{n} = [1, 0, 0]^{\mathrm{T}}$ are determined first. The specified numerical values of **J**, **R**, **C**, and **Q**, the four 3 by 3 submatrices that form the 6 by 6 matrix in equation (9.30), are given by (compare (9.39) to (9.41))

$$\mathbf{Q} - \rho v^2 \mathbf{1} = 1,000 \begin{bmatrix} 3.48 \times 10^6 - 1.5v^2 & 0 & 0 \\ 0 & 2.32 \times 10^6 - 1.5v^2 & 0 \\ 0 & 0 & 1.58 \times 10^6 - 1.5v^2 \end{bmatrix},$$

where ρ is calculated above to be 1,500 kg/m^3,

$$C - \rho_{\mathrm{f}} v^2 \mathbf{1} = C^{\mathrm{T}} - \rho_{\mathrm{f}} v^2 \mathbf{1} = 1,000 \begin{bmatrix} 1.322 \times 10^6 - v^2 & 0 & 0 \\ 0 & -v^2 & 0 \\ 0 & 0 & -v^2 \end{bmatrix},$$

$$M\mathbf{n} \otimes \mathbf{n} - v^2 \left\{ \rho_{\mathrm{f}} \mathbf{J} + \frac{i\mu}{\omega} \mathbf{R} \right\} = 1,000 \begin{bmatrix} 1.76 \times 10^6 - 2v^2 & 0 & 0 \\ 0 & -2v^2 & 0 \\ 0 & 0 & -2v^2 \end{bmatrix}.$$

These results may be assembled into a rather large 6 by 6 determinant,

$$1{,}000\begin{vmatrix} 3.48\times10^6-1.5v^2 & 0 & 0 & 1.322\times10^6-v^2 & 0 & 0 \\ 0 & 2.32\times10^6-1.5v^2 & 0 & 0 & -v^2 & 0 \\ 0 & 0 & 1.58\times10^6-1.5v^2 & 0 & 0 & -v^2 \\ 1.322\times10^6-v^2 & 0 & 0 & 1.76\times10^6-\mathbf{2}v^2 & 0 & 0 \\ 0 & -v^2 & 0 & 0 & -\mathbf{2}v^2 & 0 \\ 0 & 0 & -v^2 & 0 & 0 & -\mathbf{2}v^2 \end{vmatrix}=0,$$

but the division of this into three 2 by 2 determinants is more manageable. Substitution of the four 3 by 3 matrices above into the 6 by 6 determinant (9.31) reveals that this result may be expressed as three 2 by 2 matrices for the three sets of components, $\{a_1, b_1\}$, $\{a_2, b_2\}$, and $\{a_3, b_3\}$ (as accomplished in (9.42));

$$1{,}000\begin{bmatrix} 3.48\times10^6-1.5v^2 & 1.322\times10^6-v^2 \\ 1.322\times10^6-v^2 & 1.76\times10^6-\mathbf{2}v^2 \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}=0,$$

$$1{,}000\begin{bmatrix} 2.32\times10^6 & -v^2 \\ -v^2 & -\mathbf{2}v^2 \end{bmatrix}\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}=0,$$

$$1{,}000\begin{bmatrix} 1.58\times10^6 & -v^2 \\ -v^2 & -\mathbf{2}v^2 \end{bmatrix}\begin{bmatrix} a_3 \\ b_3 \end{bmatrix}=0.$$

Requiring that the determinants of the three 2 by 2 matrices above vanish yields four nontrivial solutions for the squared wave speed v^2. The vanishing of the first of the determinants of these 2 by 2 matrices provides two roots of a quadratic equation that represent the fast and the slow squared longitudinal wave speeds. The fast and the slow squared longitudinal wave speeds are given by,

$$v=1{,}629\ \text{m/s},\quad v=909\ \text{m/s},$$

and the vanishing of the second and third of the determinants of the 2 by 2 matrices above provides a zero root and a nonzero root from each determinant. The two nonzero roots are the squared shear wave speeds

$$v=1{,}523\ \text{m/s}\quad\text{and}\quad v=1{,}258\ \text{m/s}.$$

The vectors **a** and **b** for the fast and slow waves are given by

$$\mathbf{a}=\{-2.66b_1,0,0\},\quad \mathbf{b}=\{b_1,\,0,\,0\}$$

and

$$\mathbf{a}=\{-0.211b_1,0,0\},\quad \mathbf{b}=\{b_1,0,0\}$$

and for the two shear waves by

$$\mathbf{a} = \{0, -2b_2, 0\}, \quad \mathbf{b} = \{0, b_2, 0\}, \quad \text{and} \quad \mathbf{a} = \{0, 0, -2b_3\}, \quad \mathbf{b} = \{0, 0, b_3\},$$

respectively.

Problems

9.6.1. Which wave modes propagate in a porous medium saturated with a very low-density gas (i.e. He)? How many wave modes of each type propagate?
9.6.2. Which wave modes propagate in a porous medium when the porosity of the porous medium is zero? How many wave modes of each type propagate?
9.6.3. Using Fig. 9.6 please explain what is meant by dispersion of fast and slow waves plotted vs. frequency.
9.6.4. Explain the rise and fall of dispersion described in problem 9.6.3 in terms the properties of the fluid (specify frequency range).
9.6.5. Using symbolic algebraic software show that the vanishing of the first of the determinants (9.42) provides the two roots given by (9.45) and (9.46).
9.6.6. Using symbolic algebraic software show that the vanishing of the second and third of the determinants (9.42) provides the two roots given by (9.47) and two zero roots as well.
9.6.7. Explain how the amplitudes **a** and **b** of the two waves are determined once the wave velocity is calculated in a typical problem of the type of problems 9.4.1 or 9.4.2.
9.6.8. Determine the wave velocities and the polarization vectors or eigenvectors associated with a harmonic wave propagating along an axis of material symmetry $\mathbf{e}_1$ in cancellous bone with a porosity of 0.2. Use the properties specified for cancellous bone in Example 9.6.1.
9.6.9. Determine the wave velocities and the polarization vectors or eigenvectors associated with a harmonic wave propagating along an axis of material symmetry $\mathbf{e}_1$ in cancellous bone with a porosity of 0.5. Use the properties specified for cancellous bone in Example 9.6.1.

9.7 Propagation of Waves in a Direction That Is Not a Principal Direction of Material Symmetry; Quasi-Waves

In this section the theoretical framework for poroelastic waves is extended to the propagation of waves along a general direction in orthotropic porous media, not waves propagating the in the specific direction of a material symmetry axis considered in the previous section, which are called pure waves to distinguish them from the wave types considered in the present section. The kind of waves under consideration here are called quasi-longitudinal waves or quasi-shear waves as their amplitudes or polarization vectors are neither parallel nor perpendicular to the

direction of propagation **n**, in contrast to the case with pure longitudinal waves or pure shear waves. The waves under consideration here are said to be *quasi-longitudinal waves* if their amplitudes or polarization vectors make an angle of 45° or less with the direction of propagation **n**, or to be *quasi-shear waves* their amplitudes or polarization vectors make an angle of more than 45° with the direction of propagation **n**. The theory developed in the previous sections is applicable to quasi-waves as well as to pure waves, but the algebra is much simpler for pure waves.

Example 9.5.1: Wave Propagation in a Direction That Is Not a Principal Direction of Symmetry

Determine the wave velocities and the polarization vectors or eigenvectors for cancellous bone associated with a harmonic wave propagating in a direction that is not an axis of material symmetry. Let the axis of material symmetry be $\mathbf{e}_1$ and the direction that is not an axis of material symmetry be $\mathbf{n} = \{1,\ 1,\ 1\}(1/\sqrt{3})$. The properties of the specified cancellous bone those of Example 9.6.1.

Using the data from Example 9.6.1 and $\mathbf{n} = (1/\sqrt{3})[1,\ 1,\ 1]^{\mathrm{T}}$, the 6 by 6 determinant of the wave speeds (9.31) in this case is given by

$$1{,}000\begin{vmatrix} 2.46\times10^6-1.5v^2 & 9.6\times10^6 & 7.08\times10^6 & 4.407\times10^6-v^2 & 0.44\times10^6 & 0.44\times10^6 \\ 9.6\times10^6 & 2.053\times10^6-1.5v^2 & 0.59\times10^6 & 0.44\times10^6 & 0.44\times10^6-v^2 & 0.44\times10^6 \\ 0.708\times10^6 & 0.59\times10^6 & 1.50\times10^6-1.5v^2 & 0.44\times10^6 & 0.44\times10^6 & 0.44\times10^6-v^2 \\ 0.44\times10^6-v^2 & 0.44\times10^6 & 0.44\times10^6 & 0.59\times10^6-\mathbf{2}v^2 & 0.59\times10^6 & 0.59\times10^6 \\ 0.44\times10^6 & 0.44\times10^6-v^2 & 0.44\times10^6 & 0.59\times10^6 & 0.59\times10^6-\mathbf{2}v^2 & 0.59\times10^6 \\ 0.44\times10^6 & 0.44\times10^6 & 0.44\times10^6-v^2 & 0.59\times10^6 & 0.59\times10^6 & 0.59\times10^6-\mathbf{2}v^2 \end{vmatrix}=0$$

and the wave speeds are 1,692, 1,127, 1,039, and 893 m/s. For the 1,692 m/s wave speed the polarization vectors are

$$\mathbf{a} = \{1, 0.74, 0.40\}a_1, \quad \mathbf{b} = \{-0.42, -0.29, -0.122\}a_1.$$

This vector **a** makes an angle of 19° with the direction of propagation **n**. This vector **b** makes an angle of 156° with the direction of propagation **n**. Thus this is the quasi-longitudinal fast wave. For the 1,127 m/s wave speed the polarization vectors are

$$\mathbf{a} = \{1, -1.30, 0.007\}a_1, \quad \mathbf{b} = \{-0.56, 0.59, 0.06\}a_1.$$

This vector **a** makes an angle of 96° with the direction of propagation **n**. This vector **b** makes an angle of 91° with the direction of propagation **n**. Thus this is a quasi-shear wave. For the 1,039 m/s wave speed the polarization vectors are

$$\mathbf{a} = \{1, 0.94, -3.52\}a_1, \quad \mathbf{b} = \{-1.09, -0.106, 1.17\}a_1.$$

This vector **a** makes an angle of 104° with the direction of propagation **n**. This vector **b** makes an angle of 107° with the direction of propagation **n**. Thus this is a quasi-shear wave. For the 893 m/s wave speed the polarization vectors are

$$\mathbf{a} = \{1, -2.30, -11.9\}a_1, \quad \mathbf{b} = \{11.08, 12.73, 17.55\}a_1.$$

This vector **a** makes an angle of 129° with the direction of propagation **n**. This vector **b** makes an angle of 110° with the direction of propagation **n**. Thus this is a quasi-longitudinal slow wave.

Problems

9.7.1. What is the word used to describe waves composed of a mixture of longitudinal and wave modes?

9.7.2. For the mixed waves of the previous question, what is the particle's polarization vector orientation relative to the vector representing the direction of wave propagation?

9.7.3. Determine the wave velocities and the polarization vectors or eigenvectors associated with a harmonic wave propagating in a direction that is not an axis of material symmetry in cancellous bone with a porosity of 0.2. Let the direction that is not an axis of material symmetry be $\mathbf{n} = \{1, 1, 1\}(1/\sqrt{3})$. Use the properties specified for cancellous bone in Example 9.4.1.

References

Biot MA (1941) General theory of three-dimensional consolidation. J Appl Phys 12:155–164

Biot MA (1955) Theory of elasticity and consolidation for a porous anisotropic solid. J Appl Phys 26:182–185

Biot MA (1956a) Theory of propagation of elastic waves in a fluid saturated porous solid I low frequency range. J Acoust Soc Am 28:168–178

Biot MA (1956b) Theory of propagation of elastic waves in a fluid saturated porous solid II higher frequency range. J Acoust Soc Am 28:179–191

Biot MA (1962a) Mechanics of deformation and acoustic propagation in porous media. J Appl Phys 33:1482–1498

Biot MA (1962b) Generalized theory of acoustic propagation in porous dissipative media. J Acoust Soc Am 28:1254–1264

Cardoso L, Cowin SC (2011) Fabric dependence of quasi-waves in anisotropic porous media. J Acoust Soc Am 129(5):3302–3316

Cardoso L, Cowin SC (2012) Role of the structural anisotropy of biological tissues on poroelastic wave propagation. Mech Mater 44:174–188

Cardoso L, Teboul F, Sedel L, Meunier A, Oddou C (2003) In vitro acoustic waves propagation in human and bovine cancellous bone. J Bone Miner Res 18(10):1803–1812

Cardoso L, Meunier A, Oddou C (2008) In vitro acoustic wave propagation in human and bovine cancellous bone as predicted by the Biot's theory. J Mech Med Biol 8(2):1–19

Cowin SC (1985) The relationship between the elasticity tensor and the fabric tensor. Mechanics of Materials 4: 137–47.

Cowin SC (2003) A recasting of anisotropic poroelasticity in matrices of tensor components. Trans Porous Med 50:35–56

Cowin SC (2004) Anisotropic poroelasticity: fabric tensor formulation. Mech Mater 36:665–677

Cowin S, Cardoso L (2011) Fabric dependence of poroelastic wave propagation in anisotropic porous media. Biomech Model Mechanobiol 10:39–65

Frenkel J (1944) On the theory of seismic and seismo-electric phenomena in moist soils. J Phys USSR 8:230–241

Plona TJ, Johnson DL (1983) Acoustic properties of porous systems: I. phenomenological description. In: Johnson DL, Sen PN (eds) Physics and chemistry of porous media. AIP Conference Proceedings No. 107, pp 89–104

Sharma MD (2005) Propagation of inhomogeneous plane waves in dissipative anisotropic poroelastic solids. Geophys J Int 163:981–990

Sharma MD (2008) Propagation of harmonic plane waves in a general anisotropic porous solid. Geophys J Int 172:982–994

Yang G, Kabel J, van Rietbergen B, Odgaard A, Huiskes R, Cowin SC (1999) The anisotropic Hooke's law for cancellous bone and wood. J Elast 53:125–146.

Chapter 10
Mixture Theory-Based Poroelasticity and the Second Law of Thermodynamics

Here is an appropriate interpretation of the second law of thermodynamics: Dissipation principle: *For all thermodynamic processes that are admissible for a given constitutive assumption, the entropy production must be positive or zero*. The decisive word in this postulate is the quantifier all it makes the postulate a restrictive condition on the internal constitutive assumptions that can be imposed on systems of the type under consideration. Indeed, if internal constitutive assumptions are laid down at will and without restriction, the entropy production can be expected to be positive or zero only for some but not for all admissible processes. Thus, the second law is not a restriction on the kind of processes that can occur in nature, but a restriction on the kind of material properties that physical systems occurring in nature can have. Walter Noll, 8th International Congress on Thermal Stresses, 2009

10.1 Introduction

The title of this chapter makes reference to three topics; mixture theory, poroelasticity, and the second law of thermodynamics. In Sect. 8.11 it was noted that there are different major approaches to the development of the same basic equations for the theory of poroelasticity. The first approach, the effective medium approach, was the subject of Chap. 8. The second approach, the mixture theory approach, is one of the subjects of this chapter. It would be helpful to read or reread the paragraphs in Sect. 8.11 that deal with this second approach as much of the development of the subject described there will be detained in this Chapter. A mixture is a material with two or more ingredients, the particles of which are separable, independent, and uncompounded with each other. If the distinct phases of a mixture retain their identity, the mixture is said to be immiscible; if they lose their identity, the mixture is said to be miscible. Mixture theory provides a basis upon which the poroelasticity model of Chaps. 8 and 9 may be extended to multicomponent mixtures. The possible constituents of a mixture include a porous solid, solvents and solutes and possibly other constituents. The theory of mixtures is

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_10,

based on diffusion models that originated in a fluid mechanics and thermodynamics tradition and was formulated in the century before last.

The modeling of pore fluid flow in porous materials whose fluid and matrix constituents are solid components, solvents and solutes, and which may be modeled as incompressible, is the main subject of this chapter. The modeling of pore fluid flow in geological and biological materials whose components are compressible was dealt with in Chap. 8 on quasistatic poroelasticity. In Chap. 8 it was shown how the incompressible case may be approached from the compressible and thus demonstrating that it is possible to treat the compressible and incompressible cases jointly, but it is easier to address them separately because the interaction of the interstitial fluid flow and the solid matrix in these two tissue types is significantly different. Chap. 9 concerned wave propagation in poroelasticity and necessarily deal with the case when all constituents were compressible.

The porous medium behavior of hard biological tissues is similar to the behavior of saturated porous rocks, marble and granite, while the porous medium behavior of soft tissues is similar to the behavior of saturated soils, the sort of geological deposits one might call "swampy soils." Although both are saturated porous media, their detailed modeling and physical behavior are quite different. The term unsaturated porous media generally refers to cases when the matrix pores are filled with a fluid and a gas as, for example, the soil near the roots of a plant may contain water, air, and soil solids. In the case of hard tissues and saturated porous rocks, the fact that the bulk stiffness of the matrix material is large compared to the bulk stiffness of water means that (1) only a fraction of the hydrostatic stress in the matrix material is transferred to the pore fluid, and (2) the strains levels in many practical problems of interest are small. In the case of soft tissues and the saturated porous soils, (1) the strains can be large (however, only small strains are considered in this chapter), and (2) the bulk stiffness of the matrix material is about the same as the bulk stiffness of water which means that almost all of the hydrostatic stress in the matrix material is transferred to the pore fluid. The effective Skempton parameter defined by (8.49) or (8.45) is a measure of the fraction of the hydrostatic stress in the matrix material that is transferred to the pore fluid. For soft tissues and saturated soils the Skempton parameter approaches one, indicating that almost 100% of the hydrostatic stress in the matrix material is transferred to the pore fluid. As a consequence of this fact that the response to volumetric deformation of the fluid and the solid matrix in soft tissue is much stiffer than the deviatoric response, it is reasonable to assume that the soft tissues and the contained fluid phase are incompressible. Thus soft tissues are "hard" with respect to hydrostatic deformations and soft with respect to shearing or deviatoric deformations. This fact is the justification for the assumption of incompressibility of both the matrix material and the pore fluid made in the development of porous media models for soft tissues. The assumption of incompressibility is not correct for marble, granite, and hard tissue. For marble and granite the Skempton parameter is between 0.5 and 0.6 and for the lacunar canalicular porosity of bone it is between 0.4 and 0.5. This means that only about 50% of the hydrostatic stress in the matrix material is transferred to the pore fluid in these materials with a stiff bulk modulus.

10.2 The Present Mixture Theory-Based Poroelasticity Approach

Since mixture theory was first presented by Truesdell (1957) its relationship to the previously established Biot's poroelasticity theory (1941) has been a subject of discussion. In this Chapter the overlap in the two theories is increased. In several important ways the mixture model of saturated porous media is more general than the Biot (1941) model of poroelasticity; Bowen (1980, 1982) recovered the model of Biot (1941) from the mixture theory approach. The most important way in which the mixture model is more general than the Biot poroelastic model is that the mixture model admits the possibility of following many solid and fluid constituents and it admits the possibility of having chemical reactions occurring. Thus some constituents might vanish and others might be created. The contrast with Biot theory is that Biot theory considers the single solid and fluid components to be chemically inert. In several important ways the poroelastic model of Biot (1941, 1956a,b, 1962) offers better conceptual mechanisms for relating the elements of the physical situation to their mathematical representations, a principal example being in the distinction between the matrix, the drained and the undrained elastic constants. It is the objective of this contribution to transfer the selected Biot conceptual mechanisms to a mixture theory formulation of poroelasticity, thus combining the advantages of Biot's ideas with mixture theory.

The mixture theory-based poroelasticity presented in this chapter is augmented from the usual presentation by the addition of two poroelastic concepts developed by Biot and described in the two previous chapters. The first of these is the use of the larger RVE in Fig. 8.9a rather that the Eulerian point often employed when the mixture consist only of fluids and solutes, Fig. 8.9b. The second is the subRVE-RVE velocity average tensor $\mathbf{J}$, which Biot called the micro–macro velocity average tensor and which is related to pore structure fabric by (9.35). These two poroelastic concepts are developed in Cowin and Cardoso (2012). Traditional mixture theory allows for the possibility constituents to be open systems, but the entire mixture is a closed system. In this development the mixture is also considered to be an open system. The velocity of a solid constituent is employed as the main reference velocity in preference to the mean velocity (10.25) concept in the early formulations of mixture theory. The mean velocity concept is avoided in the mixture theory-based poroelasticity because the averaging of the solid velocity and the fluid velocity is seldom a quantity of physical interest as it is in mixture in which all constituents are fluids. The standard development of statements of the conservation principles and entropy inequality employed in mixture theory are modified to account for these kinematic changes and to allow for supplies of mass, momentum, and energy to each constituent and to the mixture as a whole.

This presentation of the theory of mixtures is restricted to the situation in which all the mixture constituents are incompressible, immiscible and all are at the same temperature θ. It is assumed that terms proportional to the square of diffusion velocities will be negligible. Bowen (1976, p. 27) considers the case where they

are not negligible. It is also assumed that the stress tensor associated with each constituent is symmetric and that there are no action-at-a-distance couples, as there would be, for example, if the material contained electric dipoles and was subjected to an electrical field. The restrictions associated with each of these assumptions may be removed.

10.3 The Second law of Thermodynamics

The development of constitutive equations for theories of mixtures cannot proceed without such a formal algebraic statement of the irreversibility principle, the second law of thermodynamics and that is why that topic is introduced in this chapter. Thus far in the development of the subjects of this book it has not been necessary to formulate a specific equation restricting the direction of development or evolution of material processes. In Chap. 6, where the linear continuum theories of heat conduction, elastic solids, viscous fluids, and viscoelastic materials were developed, direct physical arguments about irreversible processes could be made, without invoking the second law of thermodynamics. These arguments, which were in fact special applications of the second law, influenced only the signs of material coefficients and stemmed from intuitively acceptable statements like "heat only flows from hot to cold." In this chapter a statement of the entropy inequality (the second law) is introduced and its use is developed as a method of restricting constitutive functions to physically acceptable processes using the arguments introduced by Coleman and Noll (1963). The basis of that argument is summarized in the quote from Walter Noll in 2009 repeated at the top of this Chapter.

10.4 Kinematics of Mixtures

In formulation of mixture theory-based poroelasticity presented here, the Eulerian point used as a model of the continuum point (Fig. 8.9b) for a mixture whose constituents are all fluids is replaced by a larger RVE introduced by Biot (1962) as the model of the poroelastic continuum point (Fig. 8.9a). Further, Biot's concept of the RVE level representation of the fluid velocity as a function of the pore fluid velocities in the sub RVE pores is employed. Biot related the components of the relative microvelocity field $\dot{\mathbf{w}}^{\text{micro}}$ in the sub RVE pores to the RVE level fluid velocity vector $\mathbf{v}$ by a linear transformation or second order tensor denoted here as $\mathbf{J}$,

$$\dot{\mathbf{w}}^{\text{micro}} = \mathbf{J} \cdot \mathbf{v}.$$

Biot noted that $\mathbf{J}$ depended on the coordinates in the pores and the pore geometry. The formula (9.35) relates $\mathbf{J}$ to the fabric tensor of the RVE. Thus, the mixture

theory concept of the mean velocity (10.25) of the solid and fluid constituents is replaced by reference to the velocity of a selected constituent $\mathbf{v}_{(s)}$ and the diffusion velocities relative to a selected constituent, the $\mathbf{v}$ introduced above.

In the theory of mixtures each place $\mathbf{x}$ in a fixed spatial frame of reference might be occupied by several different particles, one for each constituent $\mathbf{X}_{(a)}$, $a = 1, 2,\ldots,N$, of the mixture. This representation is a direct generalization of the single constituent continuum considered in Chap. 2, thus the material description of motion (2.2) is generalized to a description that recognizes all the constituents of the mixture:

$$\mathbf{x} = \chi(\mathbf{X}_{(a)}, t) \quad \text{for all } \mathbf{X}_{(a)} \subset \mathrm{O}_{(a)}(0). \tag{10.1}$$

The inverse of the motion (10.1) is then, in analogy with the relationship between (2.2) and (2.15), given by

$$\mathbf{X}_{(a)} = \chi_{(a)}^{-1}(\mathbf{x}, t) \quad \text{for all } \mathbf{X}_{(a)} \subset \mathrm{O}_{(a)}(0). \tag{10.2}$$

Similar generalizations to multicomponent mixtures of the formulas for the deformation gradient and its inverse, (2.13) and (2.14), are straightforward. The deformation gradient tensor for the ath constituent $\mathbf{F}_{(a)}$ is defined by

$$\mathbf{F}_{(a)} = [\nabla_{(a)} \otimes \chi(\mathbf{X}_{(a)}, t)]^{\mathrm{T}} \quad \text{for all } \mathbf{X}_{(a)} \subset \mathrm{O}_{(a)}(0), \tag{10.3}$$

and the inverse deformation gradient tensor $\mathbf{F}_{(a)}^{-1}$ is, from (10.2), defined by

$$\mathbf{F}_{(a)}^{-1} = \left[\nabla \otimes \chi_{(a)}^{-1}(\mathbf{x}, t)\right]^{\mathrm{T}} \quad \text{for all } \mathbf{x} \subset \mathrm{O}_{(a)}(t). \tag{10.4}$$

The determinant of the tensor of deformation gradient for the ath constituent, $\mathrm{J}_{(a)}$, is the Jacobian of the transformation from $\mathbf{x}$ to $\mathbf{X}_{(a)}$, thus

$$\mathrm{J}_{(a)} \equiv \mathrm{Det}\,\mathbf{F}_{(a)} = \frac{1}{\mathrm{Det}\,\mathbf{F}_{(a)}^{-1}}, \tag{10.5}$$

where it is required that

$$0 < \mathrm{J}_{(a)} < \infty \tag{10.6}$$

so that a finite continuum volume always remains a finite continuum volume.

The velocity $\mathbf{v}_{(a)}$ and acceleration $\frac{\mathrm{D}^{(a)}\mathbf{v}_{(a)}}{\mathrm{D}t}$ of a particle of the ath constituent, $\mathbf{X}_{(a)}$, are defined by formulas that are generalizations of the definitions (2.24) for the velocity $\mathbf{v}$ and acceleration $\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t}$ in a single component continuum,

$$\mathbf{v}_{(a)} = \dot{\mathbf{x}}_{(a)} = \frac{\mathrm{D}^a \mathbf{x}_{(a)}}{\mathrm{D}t} = \left.\frac{\partial \boldsymbol{\chi}_{(a)}(\mathbf{X}_{(a)}, t)}{\partial t}\right|_{X_{(a)}\text{fixed}}, \tag{10.7}$$

$$\mathbf{a}_{(a)} = \ddot{\mathbf{x}}_{(a)} = \frac{\mathrm{D}^a v_{(a)}}{\mathrm{D}t} = \left.\frac{\partial^2 \chi_{(a)}(\mathbf{X}_{(a)}, t)}{\partial t^2}\right|_{X_{(a)}\text{ fixed}}, \tag{10.8}$$

where $\mathbf{X}_{(a)}$ is held fixed because it is the velocity or acceleration $\mathbf{X}_{(a)}$ that is being determined. The *spatial description of motion* of the particle $\mathbf{X}_{(a)}$ (as opposed to the *material description of motion* of the particle $\mathbf{X}_{(a)}$ represented by (10.1)) is obtained by substituting (10.2) into the expressions (10.7) for the velocity; thus $\mathbf{v}_{(a)} = \frac{\mathrm{D}^a}{\mathrm{D}t}\chi_{(a)}$ $(\mathbf{X}_{(a)}, t)$ becomes

$$\mathbf{v}_{(a)} = \frac{\mathrm{D}^a}{\mathrm{D}t}\chi_{(a)}\left(\chi_{(a)}^{-1}(\mathbf{x}, t), t\right) = \mathbf{v}_{(a)}(\mathbf{x}, t), \tag{10.9}$$

which is a generalization of (2.26). The *material time derivative* of the ath constituent is the time derivative following the material particle $\mathbf{X}_{(a)}$; it is denoted by $\mathrm{D}^a/\mathrm{D}t$ and is defined as the partial derivative with respect to time with $\mathbf{X}_{(a)}$ held constant. If $\Gamma(\mathbf{x}, t)$ represents a function of $\mathbf{x}$ and t, the material time derivative of the ath constituent is given by:

$$\frac{\mathrm{D}^a}{\mathrm{D}t} = \left.\frac{\partial \Gamma\left(\chi_{(a)}(X_{(a)}, t), t\right)}{\partial t}\right|_{\mathbf{X}_{(a)}\text{ fixed}}. \tag{10.10}$$

This definition is simply a generalization of (10.7) to an arbitrary function $\Gamma(\mathbf{x}, t)$. It then follows from (2.29) that

$$\frac{\mathrm{D}^a \Gamma}{\mathrm{D}t} = \frac{\partial \Gamma(x,t)}{\partial t} + \mathbf{v}_{(a)} \cdot \nabla[\Gamma(\mathbf{x}, t)]. \tag{10.11}$$

The modeler may select one component of the mixture as special because, from the viewpoint of the modeler, that constituent serves as a key reference relative to which the movement of all the other constituents may be referred. This constituent of the mixture is denoted by s. The motion of the selected constituent is, from (10.1), given by $\mathbf{x} = \chi(\mathbf{X}_s, t)$ for all $\mathbf{X}_s \subset O_s(0)$. The material time derivative following the selected constituent is given by (10.11) with the label a replaced by the label s. A relationship between the time derivative following the selected constituent s and the time derivative following the generic constituent a is obtained by subtracting the two formulas for the time derivatives:

$$\frac{\mathrm{D}^a \Gamma}{\mathrm{D}t} = \frac{D^s \Gamma}{\partial t} + \mathbf{v}_{(a/s)} \cdot \nabla[\Gamma(\mathbf{x}, t)], \tag{10.12}$$

where

$$\mathbf{v}_{(a/s)} = \mathbf{v}_{(a)} - \mathbf{v}_{(s)}, \tag{10.13}$$

represents the diffusion velocity of the ath constituent relative to the s constituent.

The tensor of velocity gradients for the ath constituent $\mathbf{L}_{(a)}$ is formed by taking the spatial gradient of the velocity field for the ath constituent $\mathbf{v}_{(a)} = \mathbf{v}_{(a)}(\mathbf{x}, t)$, thus

$$\mathbf{L}_{(a)} = \nabla \otimes \mathbf{v}_{(a)}. \tag{10.14}$$

Please note that this definition is completely analogous to the definition of the tensor of velocity gradients for a single constituent material, $\mathbf{L}$, given by (2.31). Using the chain rule it is easy to show that $\mathbf{L}_{(a)}$ also has the representation

$$\mathbf{L}_{(a)} = \frac{\mathrm{D}^a \mathbf{F}_{(a)}}{\mathrm{D}t} \cdot \mathbf{F}_{(a)}^{-1}. \tag{10.15}$$

If $\rho_{(a)}$ denotes the density of the ath constituent, then the density of the mixture may be defined by

$$\rho(\mathbf{x}, t) = \sum_{a=1}^{N} \rho_{(a)}(\mathbf{x}, t). \tag{10.16}$$

Physically, $\rho_{(a)}$ represents the mass of the ath constituent per unit volume of the mixture. *The* true material density for the ath constituent is denoted by $\gamma_{(a)}$ and represents the mass of the ath constituent per unit volume of the ath constituent. The quantity $\rho_{(a)}$ is sometimes called the *bulk* density as opposed to the *true material density,* $\gamma_{(a)}$. The *volume fraction* of the ath constituent, $\phi_{(a)}$, that is to say the volume of the ath constituent per unit volume of the mixture, is defined by

$$\phi_{(a)}(\mathbf{x}, t) = \frac{\rho_{(a)}(\mathbf{x}, t)}{\gamma_{(a)}(\mathbf{x}, t)}, \tag{10.17}$$

which may be viewed a factoring the bulk density into two components.

$$\rho_{(a)}(\mathbf{x}, t) = \phi_{(a)}(\mathbf{x}, t)\gamma_{(a)}(\mathbf{x}, t). \tag{10.18}$$

It is assumed that the sum of all volume fractions divided by the total volume is equal to one,

$$\sum_{a=1}^{N} \phi_{(a)}(\mathbf{x}, t) = 1. \tag{10.19}$$

The porosity of the ath constituent is 1-$\phi_{(a)}$. If the ath constituent is incompressible, then $\gamma_{(a)}$ is a constant. Observe from (10.18) that the bulk density $\rho_{(a)}$ need not be

constant even if the ath constituent is incompressible since the volume fraction $\phi_{(a)}$ may change. The mixture is only incompressible when all the $\gamma_{(a)}$, $a = 1, 2,...,N$, are incompressible. If less than N constituents are incompressible, (10.19) is a constraining relationship between the densities. Note that the mixture density ρ given by (10.19) may be variable even when all the constituents are incompressible, that is to say all the $\gamma_{(a)}$'s are constant, because the volume fraction (10.17) of the constituents present at a point $\mathbf{x}$ is variable.

10.5 The Conservation Laws for Mixtures

In this section equations are postulated equations for the balance of mass, momentum, and energy for each constituent and then the necessary and sufficient conditions are obtained so that the usual global balance of mass, momentum, and energy for the entire mixture is satisfied. In order to postulate equations for the balance of mass, momentum, and energy for each constituent $\mathbf{X}_{(a)}$, $a = 1, 2,...,N$, each constituent of the mixture is assigned a density $\rho_{(a)}$, an action-at-a-distance force density $\mathbf{d}_{(a)}$, a partial stress $\mathbf{T}_{(a)}$, a partial internal energy density $\varepsilon_{(a)}$, a partial heat flux, and a partial heat supply density, $r_{(a)}$.

The local statement of mass conservation for a single constituent continuum,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \qquad \text{(3.6 repeated)}$$

may be written for each constituent $a = 1, 2,...,N$ as

$$\frac{\partial \rho_{(a)}}{\partial t} + \nabla \cdot \left(\rho_{(a)} \mathbf{v}_{(a)}\right) = \breve{s}_{(a)}. \qquad (10.20)$$

where $\breve{s}_{(a)}$ represents the mass supply to a constituent from other constituents and from external sources. The local statement of mass conservation for a single constituent continuum rewritten in terms of the selected velocity becomes

$$\frac{\partial \rho_{(a)}}{\partial t} + \nabla \cdot \rho_{(a)} \left(\mathbf{v}_{(s)} + \mathbf{v}_{(a/s)}\right) = \breve{s}_{(a)}. \qquad (10.21)$$

The sum of all mass supplies to a constituent from other constituents is denoted by $\breve{s}$, thus

$$\sum_{a=1}^{N} \breve{s}_{(a)}(t) = \breve{s}(t). \qquad (10.22)$$

The summation of (10.21) over all constituents and the use of (10.22) yields

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left(\rho \mathbf{v}_{(s)} + \sum_{a=1}^{N} \rho_{(a)} \mathbf{v}_{(a/s)} \right) = \breve{s}(t). \tag{10.23}$$

When the selected point (s) for velocity reference is the point where the velocity is equal to the mean velocity, the statement of the conservation of mass above reduces to the traditional formula below involving the mean velocity,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \breve{s}(t). \tag{10.24}$$

The constituent form of mass balance (10.20) summed over all the constituents produces the continuum statement (3.6) if the definitions (10.16) for the density of the continuum mixture and the mean mixture velocity $\mathbf{v}$,

$$\mathbf{v} = \frac{1}{\rho} \sum_{a=1}^{N} \rho_{(a)} \mathbf{v}_{(a)} \tag{10.25}$$

are employed. Recall from the introductory paragraph of this section that this presentation of mixture-based poroelasticity will not employ this mean velocity concept. An exception to this non-use is note the fact that some results simplify when the velocity of a selected constituent $\mathbf{v}_{(s)}$ is set equal to the mean velocity (10.25). The conservation of momentum for a single constituent continuum,

$$\rho \dot{\mathbf{v}} = \nabla \cdot \mathbf{T} + \rho \mathbf{d}, \tag{3.29 repeated}$$

may be written as

$$\rho_{(a)} = \frac{\mathrm{D}^{(a)} \mathbf{v}_{(a)}}{\mathrm{D}t} = \nabla \cdot \mathbf{T}_{(a)} + \rho_{(a)} \mathbf{d}_{(a)} + \breve{\mathbf{p}}_{(a)}, \tag{10.26}$$

where $\mathbf{T}_{(a)}$ is the partial stress, $\mathbf{d}_{(a)}$ is the action-at-a-distance force density and $\breve{\mathbf{p}}_{(a)}$ is the momentum supply associated with constituent a. The momentum supply $\breve{\mathbf{p}}_{(a)}$ is the only term that is not directly associated with a term in (3.29); it represents the transfer of momentum from the other constituents to constituent a. In this presentation it is assumed that all the partial stress tensors $\mathbf{T}_{(a)}$ are symmetric. The assumption is consistent with the mixture theory applications that are to be considered here, but it is an assumption that may be avoided if necessary (Bowen 1976, 1980). The conservation of energy for constituent a is a similar generalization of the single constituent continuum result (3.52),

$$\rho \dot{\varepsilon} = \mathbf{T} : \mathbf{D} - \nabla \cdot \mathbf{q} + \rho r, \tag{3.52 repeated}$$

thus

$$\rho_{(a)} \frac{\mathrm{D}^{a} \varepsilon_{(a)}}{\mathrm{D}t} = \mathbf{T}_{(a)} : \mathbf{D}_{(a)} - \nabla \cdot \mathbf{q}_{(a)} + \rho_{(a)} r_{(a)} + \breve{\varepsilon}_{(a)}, \tag{10.27}$$

where $\varepsilon_{(a)}$ is the partial internal energy density, $\mathbf{q}_{(a)}$ is the partial heat flux vector, $r_{(a)}$ is the heat supply density, $\breve{\varepsilon}_{(a)}$ is the energy supply and $\mathbf{D}_{(a)} = (1/2)\,(\mathbf{L}_{(a)} + (\mathbf{L}_{(a)})^{\mathrm{T}})$ by the extension of (2.32) to each constituent. The energy supply $\breve{\varepsilon}_{(a)}$ is the only term that is not directly associated with a term in the single constituent continuum form of energy conservation (3.52); it represents the transfer of energy from the other constituents to constituent a.

If the sum of all mass supplies to a constituent from other constituents, denoted by $\breve{s}$ and defined by (10.22) is zero, then the summation of the forms of the balance of mass (10.20), the balance of momentum (10.26) and the balance of energy (10.27) for each constituent over all the constituents is required to produce again the single constituent continuum forms of the balance of mass (3.6), the balance of momentum (3.29) and the balance of energy (3.52), respectively. In the case when the summation is over the density-weighted time derivatives of specific quantities following the generic constituent as, for example, on the left hand side of (10.27), the result is difficult to interpret. Thus a formula relating the time derivative of the selected component to the sum of the density-weighted time derivatives has been developed. Let the constituent-specific quantity per unit mass be denoted by $\varpi_{(a)}$ and its density-weighted sum by $\rho\varpi$, thus

$$\varpi = \frac{1}{\rho}\sum_{a=1}^{N}\rho_{(a)}\varpi_{(a)}. \tag{10.28}$$

The desired formula relating the sum of the density-weighted, constituent-specific, time derivatives to the time derivative following the selected component

$$\sum_{a=1}^{N}\rho_{(a)}\frac{D^a\varpi_{(a)}}{Dt} = \rho\frac{D^s\varpi}{Dt} + \sum_{b=1}^{N-s}\nabla\cdot\left(\varpi_{(b)}\rho_{(b)}v_{(b/s)}\right) - \varpi\nabla\cdot\sum_{b=1}^{N-s}\rho_{(b)}\,\mathbf{v}_{(b/s)} + \left\{\varpi\breve{s}(t) - \sum_{a=1}^{N}\varpi_{(a)}\breve{s}_{(a)}(t)\right\}, \tag{10.29}$$

where $\mathbf{v}_{(a/s)}$ is the diffusion velocity relative to the selected component defined by (10.13).

The derivation of (10.29) is given in the Appendix to this Chapter. The derivation involves the following relationship that follows from (10.25) and (10.16) with the use of (10.13) to note that $\mathbf{v}_{(s/s)}$ must be zero:

$$\rho(\mathbf{v} - \mathbf{v}_{(s)}) = \sum_{a=1}^{N}\rho_{(a)}\,\mathbf{v}_{(a/s)} = \sum_{b=1}^{N-s}\rho_{(b)}\,\mathbf{v}_{(b/s)}, \tag{10.30}$$

The derivation of (10.29) involves the expressions for the time derivatives (10.11) and (10.12), the constituent-specific mass balance (10.20) and the definition of the density-weighted sum $\rho\varpi$ in terms of the constituent-specific quantity per unit mass denoted by $\varpi_{(a)}$, (10.28). When the result (10.30) is incorporated in (10.29) it takes the form

$$\sum_{a=1}^{N} \rho_{(a)} \frac{D^a \varpi_{(a)}}{Dt} = \rho \frac{D^s \varpi}{Dt} + \sum_{b=1}^{N-s} \nabla\cdot(\varpi_{(b)}\rho_{(b)} v_{(b/s)}) - \varpi \nabla \cdot \rho(\mathbf{v} - \mathbf{v}_{(s)}) + \left\{ \varpi \breve{s}(t) - \sum_{a=1}^{N} \varpi_{(a)} \breve{s}_{(a)}(t) \right\} \tag{10.31}$$

Note that (10.30) reduces to

$$\sum_{a=1}^{N} \rho_{(a)} \frac{D^a \varpi_{(a)}}{Dt} = \rho \frac{D^s \varpi}{Dt} + \sum_{b=1}^{N-s} \nabla \cdot (\varpi_{(b)}\, \rho_{(b)}\, v_{(b/s)}) + \left\{ \varpi \breve{s}(t) - \sum_{a=1}^{N} \varpi_{(a)} \breve{s}_{(a)}(t) \right\} \tag{10.32}$$

when $\mathbf{v}_{(s)} = \mathbf{v}$ and to

$$\sum_{a=1}^{N} \rho_{(a)} \frac{D^a \varpi_{(a)}}{Dt} = \rho \frac{D\varpi}{Dt} \tag{10.33}$$

when $\varpi_{(a)}$ is assumed not to have a dependence upon the index a. Application of the formula (10.31) relating the sum of the density-weighted, constituent-specific, time derivatives to the time derivative following the selected component to the special case of the velocity $\mathbf{v}_{(a)}$ yields the following representation:

$$\sum_{a=1}^{N} \rho_{(a)} \frac{D^a \mathbf{v}_{(a)}}{Dt} = \rho \frac{D^s \mathbf{v}}{Dt} + \sum_{a=1}^{N} \nabla \cdot [\rho_{(a)} \mathbf{v}_{(a/s)} \otimes \mathbf{v}_{(a/s)}] - \frac{1}{\rho} \sum_{a=1}^{N} \nabla \cdot (\rho_{(a)} \mathbf{v}_{(a/s)}) \left\{ \sum_{a=1}^{N} \rho_{(a)} \mathbf{v}_{(a/s)} \right\} + \left\{ \mathbf{v} \breve{s}(t) - \sum_{a=1}^{N} \mathbf{v}_{(a)} \breve{s}_{(a)}(t) \right\} \tag{10.34}$$

which reduces to

$$\sum_{a=1}^{N} \rho_{(a)} \frac{D^a \mathbf{v}_{(a)}}{Dt} = \rho \frac{D^s \mathbf{v}}{Dt} + \left\{ \mathbf{v} \breve{s}(t) - \sum_{a=1}^{N} \mathbf{v}_{(a)} \breve{s}_{(a)}(t) \right\} \tag{10.35}$$

where terms of order of the diffusion velocity $\mathbf{v}_{(a/s)}$ squared have been neglected or when $\mathbf{v} = \mathbf{v}_{(s)}$.

With the results (10.31) and (10.34) in hand it is now possible to return to the development of the sums of the constituent-specific balance equations. Recall that it is required that the summation of the forms of the balance of mass (10.20), the balance of momentum (10.26) and the balance of energy (10.27) for each constituent over all the constituents is required to produce again the single constituent continuum forms of the balance of mass (3.6), the balance of momentum (3.29) and the balance of energy (3.52), respectively, to within the supply terms for each constituent and the supply term for the mixture. The summation of the component-specific form of the conservation of linear momentum (10.26), employing the representation (10.35) for the sum of the density-weighted, component-specific time derivatives of the component-specific velocities, one obtains a result that is similar to the single component form (3.29),

$$\rho \frac{D^s \mathbf{v}}{Dt} = \nabla \cdot \mathbf{T} + \rho \mathbf{d} + \breve{\mathbf{p}}(t), \tag{10.36}$$

if the squares of the diffusion velocities are neglected. The total stress $\mathbf{T}$ is defined by

$$\mathbf{T} = \sum_{a=1}^{N} \mathbf{T}_{(a)}, \tag{10.37}$$

and the sum of the action-at-a-distance forces by

$$\mathbf{d} = \frac{1}{\rho} \sum_{a=1}^{N} (\rho_{(a)} \mathbf{d}_{(a)}), \tag{10.38}$$

and the sum of the constituent momentum supplies $\breve{\mathbf{p}}_{(a)}$ is denoted by $\breve{\mathbf{p}}$,

$$\sum_{a=1}^{N} \{\breve{\mathbf{p}}_{(a)} + \mathbf{v}_{(a)} \breve{s}_{(a)}\} - \mathbf{v}\breve{s}(t) = \breve{\mathbf{p}}. \tag{10.39}$$

If the velocity of the selected component is equal to the mean velocity of the mixture, $\mathbf{v}_{(s)} = \mathbf{v}$, the results (10.35) through (10.39) will coincide with results that appear in Bowen (1967, 1976, 1980, 1982).

The summation of the constituent-specific form of the balance of energy (10.27) over all the constituents, and subsequently employing the formula (10.29) with $\varpi_{(a)}$ replaced by $\varepsilon_{(a)}$, yields

$$\begin{aligned} \rho \frac{D^s \varepsilon}{Dt} = \rho r + \breve{\varepsilon}(t) + \sum_{a=1}^{N} \mathbf{T}_{(a)} \\ : \mathbf{D}_{(a)} - \nabla \cdot \sum_{a=1}^{N} \left\{ \mathbf{q}_{(a)} + (\varepsilon_{(a)} - \varepsilon) \rho_{(a)} \mathbf{v}_{(a/s)} \right\}, \end{aligned} \tag{10.40}$$

where ε is the specific internal energy density for the mixture and r is the heat supply density for the mixture given by

$$\varepsilon = \frac{1}{\rho}\sum_{a=1}^{N} \rho_{(a)}\,\varepsilon_{(a)},\ r = \frac{1}{\rho}\sum_{a=1}^{N} \rho_{(a)}\,r_{(a)}, \tag{10.41}$$

and where the sum of the energy supplies $\breve{\varepsilon}_{(a)}$ is denoted by $\breve{\varepsilon}(t)$,

$$\breve{\varepsilon}(t) = \sum_{a=1}^{N} \left\{ \breve{\varepsilon}_{(a)} + \varepsilon_{(a)}\breve{s}_{(a)}(t) \right\} - \varepsilon\breve{s}(t). \tag{10.42}$$

The key results of this section are the statements of the conservation of mass, momentum, and energy for each constituent and the summation of these component forms to yield statements of these conservation principles for the mixture. The kinematic identities (10.29) and (10.34) form the other important result; its derivation is a suggested problem below.

Problems

10.5.1 Show that the constituent form of mass balance (10.18) summed over all the constituents will produce the continuum statement (3.6).

10.5.2 Derive the formula relating the sum of the density-weighted, constituent-specific, time derivatives to the time derivative following the selected component, (10.29). In the course of this derivation you will likely employ (10.11), (10.12), and (10.28).

10.5.3 Derive the formula (10.34) from (10.29) by setting $\boldsymbol{\varpi}_{(a)}$ equal to $\mathbf{v}_{(a)}$.

10.6 A Statement of Irreversibility in Mixture Processes

Many physical quantities can be considered as influencing the specific internal energy density ε of a material object. (Recall that the specific internal energy density ε was introduced in the section on the conservation of energy, Sect. 3.5). These include, for example, the specific volume, the components of a tensor measuring deformation or strain, the densities, or the concentrations of the constituents of the mixture, and so on. In the mixtures of interest here the set of parameters characterizing the thermodynamic substate of a particle $\mathbf{X}$ of the mixture, which actually represents an RVE, will be the infinitesimal strain tensor $\mathbf{E}$ given by (2.52) and each of the densities of the constituents, $\rho_{(a)}$. The notation $\{\mathbf{E}, \rho_{(a)}\}$ is introduced for these parameters. The set of parameters $\{\mathbf{E}, \rho_{(a)}\}$ is said to characterize the thermodynamic substate of a particle $\mathbf{X}$ in an object (i.e. a thermodynamic system).

Knowledge of the thermodynamic substate $\{\mathbf{E}, \rho_{(a)}\}$ does not, however, completely characterize the thermodynamic properties of a thermodynamic system

because it does not describe the relationship between two substates. Given two thermodynamic substates $\{\mathbf{E}, \rho_{(a)}\}$ and $\{\mathbf{E}^*, \rho_{(a)}{}^*\}$, the knowledge is not sufficient to specify whether a transition $\{\mathbf{E}, \rho_{(a)}\} \rightarrow \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ is possible or not. This ordering of thermodynamic substates is accomplished by the introduction of entropy; the steps of this introduction we develop in the following paragraph.

A transition $\{\mathbf{E}, \rho_{(a)}\} \rightarrow \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ between thermodynamic substates is simply an ordered pair $\{\mathbf{E}, \rho_{(a)}\}, \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ of thermodynamic substates. If the transitions $\{\mathbf{E}, \rho_{(a)}\} \rightarrow \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ and $\{\mathbf{E}^*, \rho_{(a)}{}^*\} \rightarrow \{\mathbf{E}, \rho_{(a)}\}$ are both possible, the transition is said to be *reversible*. If the transition $\{\mathbf{E}, \rho_{(a)}\} \rightarrow \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ is possible, but $\{\mathbf{E}^*, \rho_{(a)}{}^*\} \rightarrow \{\mathbf{E}, \rho_{(a)}\}$ is not, then the transition is said to be *irreversible*. The directionality of transitions may be expressed in the following axiom of ordering: Given two thermodynamic substates $\{\mathbf{E}, \rho_{(a)}\}$ and $\{\mathbf{E}^*, \rho_{(a)}{}^*\}$, it is possible to decide whether the transition $\{\mathbf{E}, \rho_{(a)}\} \rightarrow \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ is possible or not.

The existence of an ordering of substates is analogous to the existence of the "greater than" relation ">" for real numbers. For example, given two distinct real numbers a and a^* we have either $a > a^*$, $a^* > a$ or $a = a^*$ while no such ordering holds for, say, the complex numbers. This permits the construction of a homeomorphism between the ordering of thermodynamic states and the ordering of real numbers. This is done by introducing a real-valued function η which assigns a real number to each thermodynamic substate $\eta = \eta(\mathbf{E}, \rho_{(a)})$, in such a way that $\eta(\mathbf{E}^*, \rho_{(a)}{}^*) > \eta(\mathbf{E}, \rho_{(a)})$ if the transition $\{\mathbf{E}, \rho_{(a)}\} \rightarrow \{\mathbf{E}^*, \rho_{(a)}{}^*\}$ is irreversible. The substates are thereby ordered and labeled by means of the real-valued function of the substate, η. In any particular physical situation such a function is empirically determinable, and if η is one such function, then so is $f(\eta)$, where f is a monotonically increasing function of its argument. Assuming a particular function η to have been chosen, it is called the empirical entropy of the thermodynamic system.

The thermodynamic state of a particle $\mathbf{X}$ in an object is completely specified by the thermodynamic substate $\{\mathbf{E}, \rho_{(a)}\}$and the entropy η of the (RVE associated with the) particle. The basic assumption of thermodynamics is that the thermodynamic state completely determines the (specific) internal energy ε independent of time, place, motion and stress, thus $\varepsilon = \varepsilon(\eta, \mathbf{E}, \rho_{(a)}, \mathbf{X})$. Choice of the exact functional form of ε defines different thermodynamic substances. If $\mathbf{X}$ does not appear in the form of ε chosen, the substance is said to be thermodynamically simple.

In order to develop a rationale for a differential equation in entropy production η, the time rate of change of internal energy ε in the conservation of energy (3.52) is expressed in two different ways:

$$\rho = \frac{\mathrm{D}^s \varepsilon}{\mathrm{D}t} = P_E + Q_E = P_I + Q_I, \tag{10.43}$$

where P_E and Q_E denote the contribution of the external mechanical and nonmechanical power and P_I and Q_I denote the division into internal mechanical and nonmechanical power, respectively. The external quantities are defined, using (3.50) and (3.52), by

$$P_E = \mathbf{T} : \mathbf{D},\ Q_E = \rho r - \nabla \cdot \mathbf{h}, \tag{10.44}$$

where $\mathbf{h}$ is the heat flux vector. The internal mechanical power is represented by the power generated by the time rate of change of the internal parameters $\{\mathbf{E}, \rho_{(a)}\}$ contracted with the corresponding internal force systems to form scalar quantities with the dimension of energy. The internal nonmechanical power is represented by the product of the time rate of change of the specific entropy with the density and the temperature θ, thus

$$Q_I = \rho\theta \frac{\mathrm{D}^s \eta}{\mathrm{D}t}. \tag{10.45}$$

From (10.43) and (10.45) it follows that

$$\rho\,\theta \frac{D^s \eta}{Dt} = P_E - P_I + Q_E; \tag{10.46}$$

thus the production of specific entropy is the difference between the external and internal mechanical work plus the external nonmechanical power.

Let N denote the entropy of an entire object and let η denote the entropy per unit mass or specific entropy at the particle $\mathbf{X}$ of the object, thus

$$N = \int_o \rho\eta \,\mathrm{d}v. \tag{10.47}$$

The time rate of change of the N and η are related by

$$\frac{D^s N}{Dt} = \int_o \rho \frac{D^s \eta}{Dt} \mathrm{d}v \tag{10.48}$$

which follows from (3.8). The total entropy production $\frac{\mathrm{D}^s N}{\mathrm{D}t}$ is written as the sum of the external production $\frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t}$ and the internal production $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t}$, thus

$$\frac{\mathrm{D}^s N}{\mathrm{D}t} = \frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} + \frac{D^s N_{(\mathrm{e})}}{Dt}. \tag{10.49}$$

The external entropy production $\frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t}$ is associated with the total flux of nonmechanical power, divided by the absolute temperature θ, thus

$$\frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t} = -\oint_{\partial O} \frac{\mathbf{h} \cdot \mathbf{n}}{\theta} \mathrm{d}a + \int_O \frac{\rho r}{\theta} \mathrm{d}v, \tag{10.50}$$

and the internal entropy production $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t}$ is simply related to a similar specific internal entropy production quantity $\frac{\mathrm{D}^s \sigma}{\mathrm{D}t}$,

$$\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} = \int_O \rho \frac{\mathrm{D}^s \sigma}{\mathrm{D}t} \mathrm{d}v, \tag{10.51}$$

where σ is the specific internal entropy. From experiment and experience it is known that, at constant temperature, the excess of external power over internal power (P_{E} - P_{I}) must be greater than zero; that is to say, at constant temperature, one cannot recover from an object more power than was supplied to the object. It is also a fact of experience and experiment that heat flows from the hotter to the colder parts of an object and not in the reverse way. In equation form we note this last assertion by

$$\mathbf{q} \cdot \nabla\theta \geq 0 \text{ for } P_{\mathrm{E}} - P_{\mathrm{I}} = 0, \tag{10.52}$$

while the former assertion is summarized in the statement (P_{E} - P_{I}) ≥ 0.

Guided by these results it is postulated that the internal entropy production is always greater than or equal to zero $\frac{\mathrm{D}^s \eta}{\mathrm{D}t} \geq 0$. This postulate is a form of the second law of thermodynamics; the postulate of irreversibility. The second law applied to an object occupying a volume O may be stated as $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} = \frac{\mathrm{D}^s N}{\mathrm{D}t} - \frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t} \geq 0$. In terms of $\frac{\mathrm{D}^s N}{\mathrm{D}t}$, $\frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t}$, and $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t}$, the following terminology is customary: an equilibrium state is defined by $\frac{\mathrm{D}^s N}{\mathrm{D}t} = 0$, a reversible process is characterized by $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} = 0$, an irreversible process is characterized by $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} > 0$; $\frac{\mathrm{D}^s N}{\mathrm{D}t} = \frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t}$ characterizes an adiabatic process and only in the case of an adiabatically insulated system does the second law of thermodynamics in the form $\frac{\mathrm{D}^s N}{\mathrm{D}t} \geq 0$ apply. It should be noted that, in general, the various entropies might satisfy the three inequalities $\frac{\mathrm{D}^s N}{\mathrm{D}t} > 0$, $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} > 0$, and $\frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t} < 0$.

In terms of the specific or continuum variables, the second law of thermodynamics $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} = \frac{\mathrm{D}^s N}{\mathrm{D}t} - \frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t} \geq 0$ may be written as

$$\int_O \rho \frac{\mathrm{D}^s \eta}{\mathrm{D}t} \mathrm{d}v + \oint_{\partial O} \frac{\mathbf{h} \cdot \mathbf{n}}{\theta} \mathrm{d}a - \int_O \frac{\rho r}{\theta} \mathrm{d}v \geq 0 \tag{10.53}$$

using (10.48) and (10.50). In order to convert the integral equation (10.53) to a field or point equation the divergence theorem (A183) is employed as well as the argument that was used to convert the integral equation (3.4) to the field equation (3.5). Recall that this was an argument employed four times in Chap. 3. Applying these arguments to (10.53) it follows that

$$\rho \frac{\mathrm{D}^s \eta}{\mathrm{D}t} + \nabla \cdot \left(\frac{\mathbf{h}}{\theta} \right) - \frac{\rho r}{\theta} \geq 0. \tag{10.54}$$

Integral equations such as (3.4) and (10.53) are global statements because they apply to an entire object. However the results (3.5) and (10.54) are local, point wise conditions valid at the typical point (place) in the object. Thus the transitions (3.4) → (3.5) and (10.53) → (10.54) are from the global to the local or from the object to the point (or particle) in the object.

On the molecular level, each macroscopic substate $\{\mathbf{E}, \rho_{(a)}\}$ corresponds to a large set of molecular states. In other words, the relationship between the molecular states and the macroscopic states is injective, they are in one-to-one correspondence. If we attribute an equal probability to each molecular state, the probability that a thermodynamic state $\{\mathbf{E}, \rho_{(a)}\}$ is different from another thermodynamic state $\{\mathbf{E}^*, \rho_{(a)}^*\}$ can be calculated from the number of molecular states to which it corresponds. In statistical physics, the concept of entropy is defined as the logarithm of the number of molecular states that correspond to that particular molecular state (multiplied by Boltzman constant). The entropy thereby provides a measure of the relative probability of different macrostates. The second law, stating that entropy moves towards increasing entropy, simply states that the system has a natural tendency to evolve from less probable states towards more probable states. In this sense one can interpret the second law as being almost a tautology.

10.7 The Entropy Inequality for a Mixture

Returning to the continuum model note that, in terms of the internal energy $\varepsilon(\eta, \mathbf{E}, \rho_{(a)}, \mathbf{X})$, the temperature, stress and electrochemical (or chemical) potential may be defined as the derivatives of $\varepsilon(\eta, \mathbf{E}, \rho_{(a)}, \mathbf{X})$ with respect to entropy, strain, and volume fraction, respectively:

$$\theta = \left(\frac{\partial \varepsilon}{\partial \eta} \right)_{E, \rho_{(a)}}, T = \left(\frac{\partial \varepsilon}{\partial E} \right)_{\eta, \rho_{(a)}}, \mu_{(a)} = \left(\frac{\partial \varepsilon}{\partial \rho_{(a)}} \right)_{E, \eta}. \tag{10.55}$$

The time derivative of the internal energy ε may then be expressed as follows:

$$\frac{\mathrm{D}^s \varepsilon}{\mathrm{D}t} = \theta \frac{\mathrm{D}^s \eta}{\mathrm{D}t} + T : D + \sum_{a=1}^{N} \mu_{(a)} \frac{\mathrm{D}^s \rho_{(a)}}{\mathrm{D}t}. \tag{10.56}$$

The Helmholtz free energy is defined by

$$\Psi(\theta, E, \rho_{(a)}, X) = \varepsilon(\eta, E, \rho_{(a)}, X) - \eta\theta, \tag{10.57}$$

and the derivatives of the free energy Ψ with respect to temperature, strain, and volume fraction yield the entropy, stress, and electrochemical (or chemical) potential, respectively:

$$\eta = -\left(\frac{\partial \Psi}{\partial \theta}\right)_{E,\rho_{(a)}}, T = \left(\frac{\partial \Psi}{\partial E}\right)_{\theta,\rho_{(a)}}, \mu_{(a)} = \left(\frac{\partial \Psi}{\partial \rho_{(a)}}\right)_{E,\theta}. \tag{10.58}$$

The time derivative of the free energy Ψ may then be expressed as follows:

$$\frac{\mathrm{D}^s \Psi}{\mathrm{D}t} = \eta \frac{\mathrm{D}^s \theta}{\mathrm{D}t} + T : D + \sum_{a=1}^{N} \mu_{(a)} \frac{\mathrm{D}^s \rho_{(a)}}{\mathrm{D}t}. \tag{10.59}$$

It is assumed that each constituent of the mixture has the regular properties of a thermodynamic substance, thus the Helmholtz free energy of each constituent $\Psi_{(a)}$ is related to the temperature θ and constituent-specific internal energy $\varepsilon_{(a)}$ and entropy $\eta_{(a)}$ by the component-specific form of (10.57)

$$\Psi_{(a)} = \varepsilon_{(a)} - \theta \eta_{(a)}, \tag{10.60}$$

where

$$\Psi = \frac{1}{\rho} \sum_{a=1}^{N} \rho_{(a)} \Psi_{(a)}. \tag{10.61}$$

Either the condition $\frac{\mathrm{D}^s N_{(\mathrm{i})}}{\mathrm{D}t} = \frac{\mathrm{D}^s N}{\mathrm{D}t} - \frac{\mathrm{D}^s N_{(\mathrm{e})}}{\mathrm{D}t} \geq 0$ or the integral (10.49) or the field equation (10.50) are called the Clausius Duhem inequality for internal entropy production. They are equivalent statements of the second law of thermodynamics. In order to generalize the inequality (10.50) to a mixture, three substitutions into (10.50) are made. First the η in (10.50) is replaced by the density-weighted average of the constituent-specific internal entropy $\eta_{(a)}$, thus

$$\eta = \frac{1}{\rho} \sum_{a=1}^{N} \rho_{(a)} \eta_{(a)}, \tag{10.62}$$

and, second, a similar replacement, the second of (10.37) is made for ρr. Third, the formula (10.26) for the density-weighted sum of all the time derivatives of $\varpi_{(a)}$ following all the constituents is related to the time derivative following the selected constituent is applied to the density-weighted average of the constituent-specific internal entropy $\eta_{(a)}$,

$$\sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^a \eta_{(a)}}{\mathrm{D}t} = \rho \frac{\mathrm{D}^s \eta}{\mathrm{D}t} + \sum_{b=1}^{N-s} \left\{ \nabla \cdot [\eta_{(b)} \rho_{(b)} u_{(b)}] - \eta \nabla \cdot [\rho_{(b)} u_{(b)}] \right\}. \tag{10.63}$$

The entropy inequality for a mixture may now be formulated using the entropy inequality for the single component continuum (10.50) as the guide. The term $\rho \frac{\mathrm{D}^s \eta}{\mathrm{D}t}$ in (10.50) is eliminated using (10.63). The heat supply density r in (10.50) is replaced by that for the mixture given by the second of (10.37), thus entropy inequality for a mixture takes the form

$$\sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^a \eta_{(a)}}{\mathrm{D}t} + \nabla \cdot \left[\sum_{a=1}^{N} \frac{h_{(a)}}{\theta} - \rho_{(a)} \eta_{(a)} u_{(a)} \right] + \sum_{a=1}^{N} \eta \nabla \cdot [\rho_{(b)} u_{(b)}] - \frac{1}{\theta} \sum_{a=1}^{N} \rho_{(a)} r_{(a)} \geq 0. \tag{10.64}$$

It is important to note that, while there were forms of each of the conservation principles for each of the constituents (10.18), (10.20), (10.21) that were summed over to obtain statements of those principles that applied to the mixture as whole, (3.6), (3.29), (3.52), respectively, it was not assumed that there were constituent-specific forms of the entropy inequality (10.64). The literature is somewhat divided on the use of constituent-specific forms of the entropy inequality (Bowen, 1976, Sect. 1.7). The conservative position is to assume only the mixture level inequality. Thus the entropy inequality employed here only makes a statement for the entire mixture, not for any particular constituent.

The remainder of this section presents the development of an alternate form of the entropy inequality (10.64). First, the product $\rho_{(a)} r_{(a)}$ is eliminated between (10.21) and (10.64) and then, second, the result is multiplied by θ, third, it is assumed that the constituent-specific flux vectors, $\mathbf{h}_{(a)}$ and $\mathbf{q}_{(a)}$, are related by

$$h_{(a)} = q_{(a)} + \rho_{(a)} \theta \eta_{(a)} u_{(a)}, \tag{10.65}$$

thus

$$\sum_{a=1}^{N} \rho_{(a)} \left(\theta \frac{\mathrm{D}^a \eta_{(a)}}{\mathrm{D}t} - \frac{\mathrm{D}^a \varepsilon_{(a)}}{\mathrm{D}t} \right) + \theta \nabla \cdot \left[\sum_{a=1}^{N} \left(\frac{q_{(a)}}{\theta} \right) \right] + \sum_{a=1}^{N} \theta \eta \nabla \cdot [\rho_{(b)} u_{(b)}] + \sum_{a=1}^{N} T_{(a)} : D_{(a)} - \nabla \cdot q_{(a)} + \breve{\varepsilon}_{(a)} \geq 0, \tag{10.66}$$

then using (10.38), (10.60) and

$$q = \sum_{a=1}^{N} q_{(a)}, \tag{10.67}$$

(The expression (10.67) for the heat flux is an approximation that neglects several terms associated with diffusion velocities. This point is discussed on page 27 of Bowen (1976)) it follows that

$$-\rho\eta\frac{D^s\theta}{Dt}-\frac{1}{\theta}q\cdot\nabla\theta+\sum_{a=1}^{N}\theta\eta\nabla\cdot[\rho_{(b)}u_{(b)}$$
$$+-\rho\eta\frac{D^s\theta}{Dt}-\frac{1}{\theta}q\cdot\nabla\theta+\sum_{a=1}^{N}\theta\eta\nabla\cdot[\rho_{(b)}u_{(b)}]$$
$$+\sum_{a=1}^{N}\left[T_{(a)}:D_{(a)}-v_{(a)}\cdot\breve{p}_{(a)}-\rho_{(a)}\frac{D^a\Psi_{(a)}}{Dt}\right]\geq 0. \tag{10.68}$$

The expression relating the terms in (10.68) containing the time derivatives of the specific free energy density for the mixture $\Psi_{(a)}$ is replaced by

$$\sum_{a=1}^{N}\rho_{(a)}\frac{\mathrm{D}^a\Psi_{(a)}}{\mathrm{D}t}=\rho\frac{\mathrm{D}^s\Psi}{\mathrm{D}t}+\sum_{b=1}^{N-s}\left\{\nabla\cdot[\Psi_{(b)}\rho_{(b)}u_{(b)}]-\Psi\nabla\cdot[\rho_{(b)}u_{(b)}]\right\}, \tag{10.69}$$

a result that was obtained by substituting $\Psi_{(a)}$ for $\varpi_{(a)}$ in (10.23); thus (10.68) becomes

$$-\rho\frac{\mathrm{D}^s\Psi}{\mathrm{D}t}-\rho\eta\frac{\mathrm{D}^s\theta}{\mathrm{D}t}-\frac{1}{\theta}q\cdot\nabla\theta+\sum_{a=1}^{N}\varepsilon\nabla\cdot[\rho_{(b)}u_{(b)}]$$
$$+\sum_{a=1}^{N}\left[T_{(a)}:D_{(a)}-v_{(a)}\cdot\breve{p}_{(a)}-\nabla\cdot[\Psi_{(b)}\rho_{(b)}u_{(b)}]\right]\geq 0, \tag{10.70}$$

where use of been made of (10.57) in setting $\Psi+\eta\theta=\varepsilon$.

The entropy inequality (10.70) will now be restricted to the case of accelerationless processes. Neglecting both the acceleration and the action-at-a-distance forces, the balance of momentum for the continuum (3.29) reduces to $\nabla\cdot T=0$ and the balance of momentum for each constituent of the continuum (10.20) reduces to $\breve{p}_{(a)}=-\nabla\cdot T_{(a)}$. Thus we now have the following representations for the divergence of the total stress and the divergence of the constituent-specific partial stress

$$\nabla\cdot T=0,\ \nabla\cdot T_{(a)}=-\breve{p}_{(a)}. \tag{10.71}$$

An algebraic development will now be used to obtain an alternate representation for the first two terms of the sum in (10.70). This manipulation begins with the identity that follows easily from (10.71) and the separation of the selected constituent stress-related components from the other stress-related components;

$$\sum_{a=1}^{N}[T_{(a)}:D_{(a)}-v_{(a)}\cdot\breve{p}_{(a)}]=\sum_{a=1}^{N}[T_{(a)}:D_{(a)}+v_{(a)}\cdot\nabla T_{(a)}]$$
$$=T_{(s)}:D_{(s)}+v_{(s)}\cdot\nabla T_{(s)}+\sum_{b=1}^{N-s}[T_{(b)}:D_{(b)}+v_{(b)}\cdot\nabla T_{(b)}]. \tag{10.72}$$

Observing from (10.33) that the selected components of the stress $T_{(s)}$ may be expressed in terms of the remaining components and the total stress.

$$T_{(s)} = T - \sum_{b=1}^{N-s} T_{(b)}, \tag{10.73}$$

then the first two terms of the sum in (10.70) have the alternate representation

$$\sum_{a=1}^{N} [T_{(a)} : D_{(a)} - v_{(a)} \cdot \breve{p}_{(a)}] = T : D_{(s)} + \sum_{b=1}^{N-s} [T_{(b)} : [\nabla \otimes u_{(b)}] + u_{(b)} \cdot \nabla T_{(b)} \tag{10.74}$$

where (10.72) has been employed. Substitution of (10.74) into (10.70) yields the inequality

$$\begin{aligned} &- \rho \frac{\mathrm{D}^s \Psi}{\mathrm{D}t} - \rho\eta \frac{\mathrm{D}^s \theta}{\mathrm{D}t} - \frac{1}{\theta} q \cdot \nabla\theta + \sum_{a=1}^{N} \varepsilon \nabla \cdot [\rho_{(b)} u_{(b)}] + T : D_{(s)} \\ &+ \sum_{b=1}^{N-s} [T_{(b)} : [\nabla \otimes u_{(b)}] + u_{(b)} \cdot \nabla T_{(b)} - \nabla \cdot [\Psi_{(b)} \rho_{(b)} u_{(b)}]] \geq 0. \end{aligned} \tag{10.75}$$

The inequality (10.75) for accelerationless processes in the absence of action-at-a-distance forces will be applied, after further specializing it for a porous medium, in the section after next.

10.8 Constitutive Equations for a Multi-Constituent Porous Medium

In this section the conservation principles of mass, linear momentum and energy, and the entropy inequality are first specialized to the development of a continuum mixture model of a fluid-saturated porous medium containing a porous solid and N-1 other species including a pore fluid and neutral solutes. Then the constitutive equations for the multicomponent porous medium are developed following these same steps outlined in Chap. 4 and employed in Chap. 6 in the development of the four linear continuum theories considered, but here placing a greater emphasis on the use of the entropy inequality to restrict the constitutive assumptions. At the end of the section the reduced entropy inequality is employed to restrict the values of the material coefficients in Fourier' and Darcy' laws.

The porous solid is taken to be the selected constituent in the mixture, hence the "s" subscript or superscript notation now means solid as well as "selected". The porous solid constituent is special because it contains all the other constituents and limits their behavior. From the viewpoint of the modeler it serves as a structure

relative to which the movement of all the other constituents may be referenced. It is therefore singled out for special consideration in models of the porous medium.

Since the true density of each constituent, $\gamma_{(a)}$, is assumed to be constant, the local statement of mass conservation for a single constituent continuum, (10.20), is written for each constituent $a = 1, 2,\ldots,N$ in terms of the local volume fractions, $\phi_{(a)}$, as

$$\frac{\partial \phi_{(a)}}{\partial t} + \nabla \cdot (\phi_{(a)} v_{(a)}) = 0. \tag{10.76}$$

Since the solid constituent is treated as special, the restriction (10.19) is rewritten as

$$\phi_s(x,t) + \sum_{b=1}^{N-s} \phi_{(b)}(x,t) = 1, \tag{10.77}$$

where the summation index b runs over all the constituents except s. Summing all the constituent mass conservation equations (10.76), and employing (10.19), it follows that

$$\nabla \cdot \left(\sum_{a=1}^{N} \phi_{(a)} v_{(a)} \right) = 0. \tag{10.78}$$

Multiplying (10.19) by v_s and subsequently taking the divergence of the result, it follows that

$$\nabla \cdot v_s - \nabla \cdot \left(\sum_{a=1}^{N} \phi_{(a)} v_{(s)} \right) = 0; \tag{10.79}$$

the sum of (10.78) and (10.79) yields

$$\nabla \cdot v_s + \nabla \cdot \left(\sum_{a=1}^{N} \phi_{(a)} u_{(a)} \right) = 0, \tag{10.80a}$$

where the definition of the diffusion velocity, (10.13) or $v_{(a/s)} = v_{(a)} - v_{(s)}$, has been employed. However, since $v_{(s/s)} = 0$, it follows that (10.80a) is equivalent to

$$\nabla \cdot v_s + \nabla \cdot \left(\sum_{b=1}^{N-s} \phi_{(b)} u_{(b)} \right) = 0. \tag{10.80b}$$

It is required that this entropy inequality (10.75) hold for all states of the mixture complying with the balance laws, the incompressibility condition and that it hold

for all states of the mixture. The implementation in the inequality (10.75) of the constraint of incompressibility is accomplished by introducing the Lagrange multipliers p (for this use of Lagrange multipliers, see Example 6.4.1 concerning pressure as a Lagrange multiplier in incompressible fluids; for another method of imposing the incompressibility constraint in poroelasticity, see Chap. 9). The Lagrange multiplier p for incompressibility is introduced by multiplying the following form of (10.76),

$$\nabla \cdot v_s + \left(\sum_{b=1}^{N-s} \phi_{(b)} \nabla \cdot u_{(b)} + u_{(b)} \nabla \cdot \phi_{(b)} \right) = 0, \tag{10.81}$$

by p and adding the result to (10.75), thus (10.75) becomes

$$\begin{aligned} &- \rho \frac{\mathrm{D}^s \Psi}{\mathrm{D}t} - \rho\eta \frac{\mathrm{D}^s \theta}{\mathrm{D}t} - \frac{1}{\theta} q \cdot \nabla\theta + \sum_{a=1}^{N} \varepsilon \nabla \cdot [\rho_{(b)} u_{(b)}] + T^{\mathrm{eff}} : D_{(s)} \\ &+ \sum_{b=1}^{N-s} \{ [T_{(b)} + \phi_{(b)} p 1 + \rho_{(b)} \Psi_{(b)} 1] : [\nabla \otimes u_{(b)}] \} \\ &+ \sum_{b=1}^{N-s} \{ u_{(b)} \cdot [\nabla \cdot T_{(b)} + p \nabla \phi_{(b)} - \nabla \rho_{(b)} \Psi_{(b)}] \} \geq 0, \end{aligned} \tag{10.82}$$

where the effective stress has been introduced:

$$T^{\mathrm{eff}} = T + p1. \tag{10.83}$$

The development of four, relatively simple, constitutive relations was described in Chap. 6. The process of developing constitutive relations was described in Chap. 3. The steps in this process consisted of the constitutive idea and the restrictions associated with the notions of localization, invariance under rigid object motions, determinism, coordinate invariance, and material symmetry. In that development, restrictions on the coefficients representing material properties were developed without recourse to the second law of thermodynamics; ad hoc arguments equivalent to those obtainable from the second law were employed. The present development proceeds by making constitutive assumptions that are consistent with the restrictions of localization, invariance under rigid object motions, determinism and coordinate invariance by assuming a general form for their functional dependence on localized tensorial variables that are invariant under rigid object motions, and that have no dependence upon time except for the present time. Then the entropy inequality is employed to restrict the constitutive assumptions in the manner of Coleman and Noll (1963) based on the philosophy described in the opening quote for this Chapter due to Noll (2009). This is a straightforward process that often appears complex due to the notation for the many factors that must be accounted for in a mixture. To ease the reader into this method a simpler example is first presented.

Example 10.8.1
Restrict the functional dependence of the free energy Ψ, the entropy η and the heat flux vector $\mathbf{q}$ in a rigid isotropic heat conductor using the entropy inequality and arguments concerning the functional dependence of these three functions.

Solution: Democratic but elementary constitutive assumptions are made for the free energy Ψ, the entropy η, and the heat flux $\mathbf{q}$. The independent variables for these three functions are assumed to be same for all three quantities; they are the temperature θ and the temperature gradient $\nabla\theta$, thus $\Psi = \Psi(\theta, \nabla\theta)$, $\eta = \eta(\theta, \nabla\theta)$, $\mathbf{q} = \mathbf{q}(\theta, \nabla\theta)$. The time rate of change of the free energy $\Psi = \Psi(\theta, \nabla\theta)$, determined using the chain rule,

$$\frac{\mathrm{D}^s\Psi}{\mathrm{D}t} = \frac{\partial\Psi}{\partial\theta}\frac{\mathrm{D}^s\theta}{\mathrm{D}t} + \frac{\partial\Psi}{\partial\nabla\theta}\cdot\frac{\mathrm{D}^s\nabla\theta}{\mathrm{D}t},$$

and the constitutive assumptions $\Psi = \Psi(\theta, \nabla\theta)$, $\eta = \eta(\theta, \nabla\theta)$, $q = q(\theta, \nabla\theta)$ are then substituted into the entropy inequality (10.82) and one obtains the inequality

$$-\rho\left(\eta + \frac{\partial\Psi}{\partial\theta}\right)\frac{\mathrm{D}^s\theta}{\mathrm{D}t} - \frac{1}{\theta}q\cdot\nabla\theta - \rho\frac{\partial\Psi}{\partial\nabla\theta}\frac{\mathrm{D}^s\nabla\theta}{\mathrm{D}t} \geq 0.$$

The argument originated by Coleman and Noll (1963) is that this inequality must be true for all possible physical processes and the functional dependence upon the independent variables, in this case $\{\theta, \nabla\theta\}$, must be restricted so that is the case. In view of this set of independent variables, the last summand of the inequality above is linear in the time derivative of the temperature gradient $\mathrm{D}^s\nabla\theta/\mathrm{D}t$. This time derivative is not contained in the set of independent variables $\{\theta, \nabla\theta\}$ and it does not appear elsewhere in the inequality and thus it may be varied when the set of independent variables at a point is held fixed. In order that it not be varied in way that inequality be violated it is necessary that the coefficient $\frac{\partial\Psi}{\partial\nabla\theta}$ of $\frac{\mathrm{D}^s\nabla\theta}{\mathrm{D}t}$ must vanish. However, if $\frac{\partial\Psi}{\partial\nabla\theta} = 0$, it follows that the function dependence of $\Psi = \Psi(\theta, \nabla\theta)$ is reduced to $\Psi = \Psi(\theta)$. When the reduced dependence, $\Psi = \Psi(\theta)$, the inequality above reduces to

$$-\rho\left(\eta + \frac{\partial\Psi}{\partial\theta}\right)\frac{\mathrm{D}^s\theta}{\mathrm{D}t} - \frac{1}{\theta}q\cdot\nabla\theta \geq 0.$$

The argument made above is now repeated to show that the term involving the time derivative of the temperature, $\mathrm{D}^s\theta/\mathrm{D}t$, must vanish, thus one concludes that $\eta = -(\partial\Psi/\partial\theta)$. When the restrictions $\partial\Psi/\partial\nabla\theta = 0$ and $\eta = -(\partial\Psi/\partial\theta)$ are substituted back into the form of the entropy inequality above, it reduces to $-(1/\theta)q\cdot\nabla\theta \geq 0$. In the case when the isotropic form of the Fourier law of heat conduction gives the heat flux, $q = -k\nabla\theta$, the inequality reduces to $(k/\theta)\nabla\theta\cdot\nabla\theta \geq 0$ and it

requires only that k be positive, since the temperature is positive as a consequence of its definition.

Finally, note that the argument involving $\mathrm{D}^s\nabla\theta/\mathrm{D}t$ was performed first, then the argument involving $\mathrm{D}^s\theta/\mathrm{D}t$. This ordering of the arguments is required because $\mathrm{D}^s\theta/\mathrm{D}t$ depends on $\nabla\theta$, $\frac{\mathrm{D}^s\theta}{\mathrm{D}t} = \frac{\partial\theta(x,t)}{\partial t} + v_{(s)} \cdot \nabla[\theta(x,t)]$, a result that is a special case of (10.11).

The method of this example will next be applied to the mixture of interest. Constitutive assumptions will now be made for the free energy Ψ, the entropy η, the effective stress T^{eff}, each constituent-specific free energy $\Psi_{(b)}$, each constituent-specific partial stress $T_{(b)}$ and the heat flux $\mathbf{q}$. The independent variables included in the constitutive assumptions are the localized small strain tensor $\mathbf{E}$, the temperature θ, the temperature gradient $\nabla\theta$, the constituent densities $\rho_{(b)}$ and the diffusion velocities $v_{(b/s)}$. These functional dependencies are expressed as equations in the following forms:

$$\begin{aligned}
\Psi &= \Psi(E,\theta,\nabla\theta,\phi_{(b)},v_{(b/s)}),\ \eta = \eta(E,\theta,\nabla\theta,\phi_{(b)},v_{(b/s)}),\\
T^{\mathrm{eff}} &= T^{\mathrm{eff}}(E,\theta,\nabla\theta,\phi_{(b)},v_{(b/s)}),\ \Psi_{(c)} = \Psi_{(c)}(E,\theta,\nabla\theta,\phi_{(b)},v_{(b/s)})\\
T_{(c)} &= T_{(c)}(E,\theta,\nabla\theta,\phi_{(b)},v_{(b/s)}),\ q = q(E,\theta,\nabla\theta,\phi_{(b)},v_{(b/s)})
\end{aligned} \tag{10.84}$$

These constitutive assumptions are now substituted into the entropy inequality (10.82); in the case of the free energy, Ψ, the chain rule is applied, thus

$$\frac{\mathrm{D}^s\Psi}{\mathrm{D}t} = \frac{\partial\Psi}{\partial\theta}\frac{\mathrm{D}^s\theta}{\mathrm{D}t} + \frac{\partial\Psi}{\partial E} : \frac{\mathrm{D}^sE}{\mathrm{D}t} + \frac{\partial\Psi}{\partial\phi_{(a)}}\frac{\mathrm{D}^s\phi_{(a)}}{\mathrm{D}t} + \frac{\partial\Psi}{\partial v_{(b/s)}}\cdot\frac{\mathrm{D}^sv_{(a/s)}}{\mathrm{D}t} + \frac{\partial\Psi}{\partial\nabla\theta}\cdot\frac{\mathrm{D}^s\nabla\theta}{\mathrm{D}t}, \tag{10.85}$$

thus

$$\begin{aligned}
&-\rho\left(\eta + \frac{\partial\Psi}{\partial\theta}\right)\frac{D^s\theta}{Dt} - \frac{1}{\theta}q\cdot\nabla\theta + \left(T^{eff} - \rho\frac{\partial\Psi}{\partial E}\right) : D_{(s)} - \rho\frac{\partial\Psi}{\partial\nabla\theta}\cdot\frac{D^s\nabla\theta}{Dt}\\
&\quad + -\sum_{a=1}^{N}\rho\frac{\partial\Psi}{\partial v_{(b/s)}}\cdot\frac{D^sv_{(b/s)}}{Dt}\Bigg) + \sum_{b=1}^{N-s}\{[T_{(b)} + \phi_{(b)}\mu_{(b)}1 - \rho_{(b)}\Psi_{(b)}1] : [\nabla\otimes v_{(b/s)}]\}\\
&\quad + \sum_{b=1}^{N-s}\{[\nabla\cdot T_{(b)} + \mu_{(b)}\nabla\phi_{(b)} - \nabla\rho_{(b)}\Psi_{(b)}]\cdot v_{(b/s)}\} \geq 0,
\end{aligned} \tag{10.86}$$

where $\mu_{(b)}$ are the electrochemical potentials of the constituents other than the porous solid,

$$\mu_{(b)} = \rho\frac{\partial\Psi}{\partial\phi_{(b)}} + \gamma_{(b)}\varepsilon + p, \tag{10.87}$$

p is a Lagrange multiplier representing the fluid pressure. We now employ the Coleman and Noll (1963) argument described in the example above, namely that the assumed constitutive dependencies (10.84) must be restricted so that the inequality (10.86) is true for any value of the final set of independent variables. In view of this set of independent variables, the last summand of (10.86) is linear in the time derivative of the diffusion velocity $\frac{\mathrm{D}^s v_{(b/s)}}{\mathrm{D}t}$. This time derivative is not contained in the set of independent variables $(E, \theta, \nabla\theta, \phi_{(b)}, v_{(b/s)})$, $\Psi = \Psi(E, \theta, \nabla\theta, \phi_{(b)}, v_{(b/s)})$ and it does not appear elsewhere in the inequality and thus it may be varied when the set of independent variables at a point is held fixed. It follows that the coefficient $\frac{\partial \Psi}{\partial v_{(b/s)}}$ of $\frac{\mathrm{D}^s v_{(b/s)}}{\mathrm{D}t}$ must vanish, thus the functional dependence of $\Psi = \Psi(E, \theta, \nabla\theta, \phi_{(b)}, v_{(b/s)})$ is reduced to $\Psi = \Psi(E, \theta, \nabla\theta, \phi_{(b)})$. Repeating the same argument, the coefficient $\frac{\partial \Psi}{\partial \nabla\theta}$ of $\frac{\mathrm{D}^s \nabla\theta}{\mathrm{D}t}$ must vanish, thus the functional dependence of $\Psi = \Psi(E, \theta, \nabla\theta, \phi_{(b)})$ is reduced to $\Psi = \Psi(E, \theta, \phi_{(b)})$. In view of the reduced dependence of $\Psi = \Psi(E, \theta, \phi_{(b)})$, a repetition of the previous argument for this reduced set of independent variables is applied to the coefficient of the first summand of (10.86) involving the time derivative of the temperature $\frac{\mathrm{D}^s \theta}{\mathrm{D}t}$, and so the coefficient of the third summand is linear in the solid rate of deformation tensor $D_{(s)}$, thus

$$\eta = -\frac{\partial \Psi}{\partial \theta} \text{ and } T^{\text{eff}} = \rho \frac{\partial \Psi}{\partial E}. \tag{10.88}$$

This result shows that the entropy and effective stress of a porous medium can be derived from a regular strain energy function $\Psi = \Psi(E, \theta, \phi_{(b)})$, which physically has the same meaning as in single phase or multiphasic media, but which can depend on both strain and solute concentrations in the medium. Concerning the summands within the summation signs in (10.86) there is one summand linear in the gradient of the diffusion velocity $\nabla \otimes \mathbf{v}_{(b/s)}$, and this term appears nowhere else in the reduced inequality. In order to satisfy this restriction the following equation for the partial stress is implied:

$$T_{(b)} = \phi_{(b)}(\gamma_{(b)}\Psi_{(b)} - \mu_{(b)})1, \tag{10.89}$$

thus the partial stress of the fluid and the solutes are all seen to be scalars.

When the reduced functional dependence of the free energy to $\Psi = \Psi(E, \theta, \phi_{(b)})$ and the three restrictions (10.88) and (10.89) obtained on the inequality (10.87) are substituted back into (10.87), it reduces to the following:

$$\left.-\frac{1}{\theta} q \cdot \nabla\theta - \sum_{b=1}^{N-s} \phi_{(b)} v_{(b/s)} \cdot \nabla\mu_{(b)}\right\} \geq 0. \tag{10.90}$$

Note that the functional dependence of the terms in (10.90) is not uniform. In particular, note that $\mu_{(b)}$ in (10.90) depends only upon $(E, \theta, \phi_{(b)})$, the independent variables of the free energy, while q depends upon $(E, \theta, \nabla\theta, \phi_{(b)}, v_{(b/s)})$ by assumption (10.84). In the case when the anisotropic form of the Fourier law of heat conduction gives the heat flux, $\mathbf{q} = -\mathbf{K}\cdot\nabla\theta$, the inequality reduces to $(1/\theta)\nabla\theta\cdot\mathbf{K}\cdot\nabla\theta \geq 0$ and it requires only that $\mathbf{K}$ be positive definite, since the temperature is positive as a consequence of its definition. The entropy inequality then reduces to the following:

$$-\sum_{b=1}^{N-s}[v_{(b/s)} \cdot (\phi_{(b)}\nabla\mu_{(b)})] \geq 0. \tag{10.91}$$

Developments beyond this point could include the specification of constitutive equations for the gradients of the chemical potentials, $\phi_{(b)}\nabla\mu_{(b)}$. The inequality (10.91) may then be used to restrict the coefficients in these constitutive relationships.

As an example of the application of this result consider Darcy's law (5.36D) which, in the case of an incompressible fluid, $\rho_f = \rho_o$, is written

$$\mathbf{q} = \phi\mathbf{v} = -\mathbf{H}(p) \cdot \nabla p(\mathbf{x}, t), \mathbf{H}(p) = \mathbf{H}^T(p). \tag{10.92}$$

Considering the chemical potential (10.87) to be a function of the single constituent pressure only, then (10.91) may be written as

$$-\phi v \cdot \nabla p \geq 0. \tag{10.93}$$

Substituting $\phi\mathbf{v}$ from (10.92) into (10.93) it follows that

$$\nabla p \cdot H \cdot \nabla p \geq 0, \tag{10.94}$$

which requires that $\mathbf{H}(p)$ be positive definite. This result is equivalent to (5.29D).

Problem

10.8.1 Derive the statement of mass balance for the entire mixture, (10.76), by summing (10.76) over all constituents and employing (10.77).

10.9 Relevant Literature

Some of the literature relevant to mixture theory is discussed in Sect. 9.14. Bowen (1967) summarized the formative years of mixture theory. A readable history of the subject and its applications in the period 1957–1975 is given by Atkin and Craine (1976a,b). de Boer (1996, 2000) has presented more up-to-date histories. Of key

importance in the development of the mixture theories is the application by Bowen (1967, 1976, 1980, 1982) of a thermodynamically based analytical approach developed by Coleman and Noll (1963) to restrict the form of constitutive equations. At that time (1966–1967) the thermodynamically based Coleman and Noll approach was very new. In his history of the development of theories for porous media, de Boer (2000) credits Goodman and Cowin (1972) with the first use of the decomposition of bulk density into a volume fraction and a true or material density (10.16). The introductory material in this chapter is taken from Bowen (1976, 1980, 1982).

Appendix

The purpose of this appendix is to record the derivation of (30) and some related auxiliary results. Recall that $\varpi_{(a)}$ denotes a generic component-specific property such as $v_{(a)}$ or $\varepsilon_{(a)}$ and we seek a simple formula for $\sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^a \varpi_{(a)}}{\mathrm{D}t}$ to be used in determining the continuum level form of the conservation laws by summing over the single constituent continuum forms of the conservation laws. A formula relating the density-weighted sum of the time derivatives of the selected components to the sum of the density-weighted time derivatives is desired. Recall that the sum of generic constituent-specific quantity per unit mass $\varpi_{(a)}$ is related to its density-weighted sum ϖ by (10.29). The time derivative of (10.29) with respect to the selected component is given by

$$\rho \frac{\mathrm{D}^s \varpi}{\mathrm{D}t} + \varpi \frac{\mathrm{D}^s \rho}{\mathrm{D}t} = \sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^s \varpi_{(a)}}{\mathrm{D}t} + \sum_{a=1}^{N} \varpi_{(a)} \frac{\mathrm{D}^s \rho_{(a)}}{\mathrm{D}t} \tag{10.95}$$

which may be solved for $\sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^s \varpi_{(a)}}{\mathrm{D}t}$, thus

$$\sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^s \varpi_{(a)}}{\mathrm{D}t} = \rho \frac{\mathrm{D}^s \varpi}{\mathrm{D}t} + \varpi \frac{\mathrm{D}^s \rho}{\mathrm{D}t} - \sum_{a=1}^{N} \varpi_{(a)} \frac{\mathrm{D}^s \rho_{(a)}}{\mathrm{D}t}. \tag{10.96}$$

The relationship between the time derivatives with respect to the selected component and with respect to the "a" component is obtained using (10.12)

$$\sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^s \varpi_{(a)}}{\mathrm{D}t} = \sum_{a=1}^{N} \rho_{(a)} \frac{\mathrm{D}^a \varpi_{(a)}}{\mathrm{D}t} + \sum_{a=1}^{N} \rho_{(a)} v_{(s)} \cdot \nabla \varpi_{(a)} - \sum_{a=1}^{N} \rho_{(a)} v_{(a)} \cdot \nabla \varpi_{(a)}; \tag{10.97}$$

this is used to rewrite (10.96) as

$$\sum_{a=1}^{N}\rho_{(a)}\frac{\mathrm{D}^a\boldsymbol{\varpi}_{(a)}}{\mathrm{D}t}=\rho\frac{\mathrm{D}^s\boldsymbol{\varpi}}{\mathrm{D}t}+\boldsymbol{\varpi}\frac{\mathrm{D}^s\rho}{\mathrm{D}t}-\sum_{a=1}^{N}\boldsymbol{\varpi}_{(a)}\frac{\mathrm{D}^s\rho_{(a)}}{\mathrm{D}t}-\sum_{a=1}^{N}\rho_{(a)}\boldsymbol{v}_{(s)}\cdot\nabla\boldsymbol{\varpi}_{(a)}+\sum_{a=1}^{N}\rho_{(a)}\boldsymbol{v}_{(a)}\cdot\nabla\boldsymbol{\varpi}_{(a)}. \tag{10.98}$$

The following relationships, the first obtained from the conservation of mass for the mixture (10.25),

$$\frac{\mathrm{D}^s\rho}{\mathrm{D}t}=\frac{\partial\rho}{\partial t}+\boldsymbol{v}_{(s)}\cdot\nabla\rho=\breve{s}(t)-\nabla\cdot(\rho\boldsymbol{v})+\boldsymbol{v}_{(s)}\cdot\nabla\rho \tag{10.99}$$

and the second obtained from the conservation of mass for the constituent (10.23)

$$\frac{\mathrm{D}^s\rho_{(a)}}{\mathrm{D}t}=\frac{\partial\rho_{(a)}}{\partial t}+\boldsymbol{v}_{(s)}\cdot\nabla\rho_{(a)}=\breve{s}_{(a)}(t)-\nabla\cdot(\rho_{(a)}\boldsymbol{v}_{(a)})+\boldsymbol{v}_{(s)}\cdot\nabla\rho_{(a)} \tag{10.100}$$

will now be used in (10.98). However, before using (10.100) it is multiplied by and summed over all values of "a," thus

$$\begin{aligned}\sum_{a=1}^{N}\rho_{(a)}\frac{D^a\boldsymbol{\varpi}_{(a)}}{Dt}=\rho\frac{D^s\boldsymbol{\varpi}}{Dt}&+\boldsymbol{\varpi}(\breve{s}(t)-\nabla\cdot(\rho\boldsymbol{v})+\boldsymbol{v}_{(s)}\cdot\nabla\rho)\\&-\sum_{a=1}^{N}\boldsymbol{\varpi}_{(a)}\breve{s}_{(a)}(t)+\sum_{a=1}^{N}\boldsymbol{\varpi}_{(a)}\nabla\cdot(\rho_{(a)}\boldsymbol{v}_{(a)})\\&-\sum_{a=1}^{N}\boldsymbol{\varpi}_{(a)}\boldsymbol{v}_{(s)}\cdot\nabla\rho_{(a)}-\sum_{a=1}^{N}\rho_{(a)}\boldsymbol{v}_{(s)}\cdot\nabla\boldsymbol{\varpi}_{(a)}\\&+\sum_{a=1}^{N}\rho_{(a)}\boldsymbol{v}_{(a)}\cdot\nabla\boldsymbol{\varpi}_{(a)}.\end{aligned} \tag{10.101}$$

The second line of the result above is condensed

$$\sum_{a=1}^{N}\rho_{(a)}\frac{\mathrm{D}^a\boldsymbol{\varpi}_{(a)}}{\mathrm{D}t}=\rho\frac{\mathrm{D}^s\boldsymbol{\varpi}}{\mathrm{D}t}+\boldsymbol{\varpi}(\breve{s}(t)-\nabla\cdot(\rho\boldsymbol{v})+\boldsymbol{v}_{(s)}\cdot\nabla\rho)-\sum_{a=1}^{N}\boldsymbol{\varpi}_{(a)}\breve{s}_{(a)}(t)+\sum_{a=1}^{N}\nabla\cdot(\boldsymbol{\varpi}_{(a)}\rho_{(a)}\boldsymbol{v}_{(a)})-\sum_{a=1}^{N}\boldsymbol{v}_{(s)}\cdot\nabla\boldsymbol{\varpi}_{(a)}\rho_{(a)} \tag{10.102}$$

and then the entire equation is algebraically reduced to (10.30).

References

Atkin RJ, Craine RE (1976a) Continuum theories of mixtures: basic theory and historical development. Q J Mech Appl Math XXIX:209–244

Atkin RJ, Craine RE (1976b) Continuum theories of mixtures: applications. J Inst Math Appl 17:153–207

Biot MA (1935) Le problème de la consolidation des matières argileuses sous une charge. Ann Soc Sci Bruxelles B55:110–113

Biot MA (1941) General theory of three-dimensional consolidation. J Appl Phys 12:155–164

Biot MA (1956a) Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low frequency range. J Acoust Soc Am 28:168–178

Biot MA (1956b) Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range. J Acoust Soc Am 28:179–191

Biot MA, Willis DG (1957) The elastic coefficients of the theory of consolidation. J Appl Mech 24:594–601

Biot MA (1962) Generalized theory of acoustic propagation in porous dissipative media. J Acoust Soc Am 34:1254–1264

Biot MA (1972) Theory of finite deformation of porous solids. Indiana Univ Math J 21:597–620

Biot MA (1982) Generalized Lagrangian equations of non-linear reaction-diffusion. Chem Phys 66:11–26

Bowen RM (1967) Toward a thermodynamics and mechanics of mixtures. Arch Rat Mech Anal 24:370–403

Bowen RM (1976) Mixtures and EM Field Theories. In: Eringen A.C. (ed) Continum Physics, Vol 3. Academic Press, New York

Bowen RM (1980) Incompressible porous media models by use of the theory of mixtures. Int J Eng Sci 18:1129–1148

Bowen RM (1982) Compressible porous media models by use of the theory of mixtures. Int J Eng Sci 20:697–735

Coleman BD, Noll W (1963) The thermodynamics of elastic materials with heat conduction. Arch Ration Mech Anal 13:167–178

Cowin SC, Cardoso L. (2012) Mixture theory-based poroelasticity as a model of interstitial tissue growth, Mech Mat 44:47–57

De Boer R (1996) Highlights in the historical development of the porous media theory: toward a consistent macroscopic theory. Appl Mech Rev 49:201–262

De Boer R (2000) Theory of porous media: highlights in the historical development and current state. Springer Verlag, Berlin

Goodman MA, Cowin SC (1972) A continuum theory for granular materials. Arch Ration Mech Anal 44:249–266

Noll W (2009) Thoughts on thermodynamics. 8th International Congress on thermal stresses

Truesdell CA (1957) Sulle basi della termomeccania. Rend Lincei 22(33–38):1158–1166

Chapter 11
Kinematics and Mechanics of Large Elastic Deformations

"The perfectly elastic material is a particular kind of ideal material. It has a single preferred or natural configuration. We think a portion of this material in a certain configuration with material coordinates **X**, which we choose as specifying the reference configuration. However this material is deformed, into whatever configuration it is brought, it always remembers precisely its preferred or natural configuration and attempts to get back to it. When the forces that maintain it its present configuration are released, it will return precisely to initial configuration. It is a material, in other words, which has perfect memory for one state and no memory whatever for any other state. The forces required to maintain it in the configuration $\chi$are completely independent of the matter in which it is brought from its original configuration to the configuration χ, the time it has taken to get there, and all of its intermediate history. This is a highly idealized kind of behavior, but it is one that may be observed in a remarkably good approximation and rubber, for example." Truesdell (1960)

11.1 Large Deformations

Large deformations are more difficult to mathematically model than either small deformations or fluid motions. The difficulty stems from the fact that, for the analysis of large deformations, the knowledge or data associated with at least two different configurations must be maintained. In the case of fluid motions, only the knowledge of the present or instantaneous configuration is necessary and, in the case of small deformations of solids, the difference between the initial reference configurations and the present configuration is a small higher order quantity, and it is neglected. In fact, this neglect of higher order terms between the two configurations is the definition of "small" deformations. These difficulties may be illustrated using the concept of stress in large deformations. For small deformations the only definition of stress employed is force per unit of instantaneous cross-sectional area. This is adequate since there is a negligible difference between the

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2_11,

area in the reference configuration and the area in the instantaneous configuration in the case of small deformations. In practice, in the case of small deformations, the stress is generally calculated approximately using the original cross-sectional area, even though the definition is for the instantaneous cross-sectional area. For large deformations however, the cross-sectional area can change considerably. Consider a rubber band in its unstressed state and measure or visualize the cross-sectional area of the rubber band perpendicular to the long axis of the band, the long axis that forms the closed loop of the band. When the rubber band is stretched, note how the cross-sectional area decreases. When the band is stretched the force on the band is increased and, since stress is defined as force per unit area and the area is decreasing, the increasing stress is due not only to the increasing force, but also to the decreasing cross-sectional area of the rubber band. Each incremental increase in the force increases the stress, but it also reduces the cross-sectional area, thereby further increasing the stress. This feature of the concept of stress at large deformations, the fact that its change is not only due to changing force, but also to changing area, is a characteristic feature of large deformations, namely that the analysis of large deformations requires nonlinear mathematics. Nonlinear mathematics is generally more difficult than linear.

A development of the kinematics of large deformations is presented in this chapter. It begins in the following section with homogenous deformations and continues with the polar decomposition theorem, strain tensors for large deformations, and formulas for the calculation of volume and area change. Using the formulas for the area change, the appropriate definitions of stress for large deformations are then developed. These large deformation stress measures are incorporated in the stress equations of motion. Constitutive equations for both Cauchy elastic and hyperelastic materials capable of large deformations are then considered along with the special cases of isotropic material models and incompressible material models. Some solutions to large deformation anisotropic elastic problems are then described. The chapter closes with a discussion of the literature on this topic.

11.2 Large Homogeneous Deformations

In this section the easily understood and easily illustrated large class of deformations called homogeneous deformations is described. It is most important for the modeler to understand homogeneous deformations because most mechanical testing of materials requires homogeneous deformations and many finite deformation problems for this class of deformations are easily solved. *Homogeneous deformations* are deformations that are exactly the same for all particles, that is to say all particles experience the same deformation, the same strain, and the same rotation. Recalling the representation (2.2) for a motion, $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$, and the fact that the strain and rotation are derivatives of the motion with respect to $\mathbf{X}$, means that the motion must be linear in $\mathbf{X}$ so that the strain and rotation measures such as

the deformation gradients will be constant. Thus a homogeneous deformation is mathematically defined as a deformation that has a representation of the form

$$\mathbf{x} = \mathbf{L}(t) \cdot \mathbf{X}, \tag{11.1}$$

where $\mathbf{L}(t)$ is a tensor independent of $\mathbf{X}$. Recalling the definition (2.13) of the deformation gradient tensor $\mathbf{F}$, $\mathbf{F} = [\nabla_{\mathbf{O}} \otimes \boldsymbol{\chi}(\mathbf{X},t)]^{\mathrm{T}}$, it follows that for a homogeneous deformation,

$$\mathbf{F} = \mathbf{L}. \tag{11.2}$$

Thus, as would be expected for the definition of a homogeneous deformation, the tensor of deformation gradients is independent of $\mathbf{X}$. Lord Kelvin and P. G. Tait in their Treatise on Natural Philosophy extensively developed the geometrically interesting properties of homogeneous deformations. Their study of the effect of homogeneous deformations upon simple geometric figures is briefly reviewed here. Recall that a plane may be defined by its normal and the specification of one point in the plane. Let $\mathbf{a}$ denote the vector normal to a plane and let $\mathbf{X}_{\mathrm{o}}$ denote a point in the plane. Then all the other points in the plane are points $\mathbf{X}$ such that $\mathbf{a}\cdot(\mathbf{X} - \mathbf{X}_{\mathrm{o}}) = 0$. This is so because the condition $\mathbf{a}\cdot(\mathbf{X} - \mathbf{X}_{\mathrm{o}}) = 0$ requires that the vector $(\mathbf{X} - \mathbf{X}_{\mathrm{o}})$ be perpendicular to $\mathbf{a}$, the normal to the plane. If we set $\mathbf{a}\cdot\mathbf{X}_{\mathrm{o}} = c$, a constant, then a material surface that forms a plane may be described in material coordinates by

$$\mathbf{a} \cdot \mathbf{X} = c. \tag{11.3}$$

Substituting the inverse of (11.2), $\mathbf{X} = \mathbf{L}^{-1}\mathbf{x}$, into (11.3) yields

$$\mathbf{a}^* \cdot \mathbf{x} = c, \tag{11.4}$$

where a new constant vector $\mathbf{a}^*$ by $\mathbf{a}^* = \mathbf{a}\cdot\mathbf{L}^{-1}$ has been defined. Eq. (11.3) is also the equation of a plane, a plane in the spatial coordinate system, thus permitting one to conclude that a plane material surface is deformed into a plane spatial surface by a homogeneous deformation. More simply stated, homogeneous deformations map planes into planes. Selecting different values for the constant c in (11.3) and (11.4), it may be concluded that parallel planes will deform into parallel planes since the normals to parallel planes have the same direction. Since the intersection of two planes is a straight line (Fig. 11.1), it follows that parallel straight lines go into parallel straight lines, parallelograms go into parallelograms, and parallelepipeds deform into parallelepipeds. The results are illustrated in Fig. 11.2.

Example Problem 11.2.1

Draw a sketch of the set of parallel lines given by the intersection of the planes $\mathbf{a} = [2\ 3\ 0]^{\mathrm{T}}$ with $c = 0$ and 5, and $X_{III} = 0$. These lines have the representations $2X_I + 3X_{II} = 0$ and $2X_I + 3X_{II} = 5$, respectively. Draw a sketch of the set of parallel lines after subjecting them to the homogeneous deformation

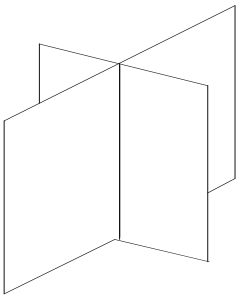


Fig. 11.1 The intersection of two planes is a straight line

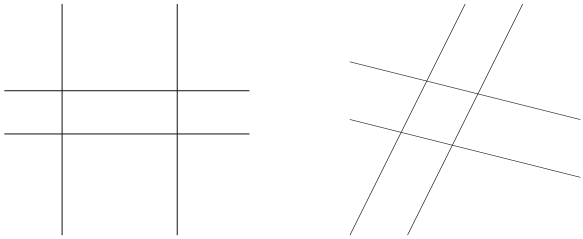


Fig. 11.2 An illustration of sets of parallel lines deforming into sets of parallel lines and a parallelogram deforming into a parallelogram

$$\mathbf{F} = \begin{bmatrix} \sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: A sketch of the set of parallel lines given by $2X_I + 3X_{II} = 0$ and $2X_I + 3X_{II} = 5$ is shown in Fig. 11.3a. The inverse of the homogeneous deformation $\mathbf{F}$ is given by

$$\mathbf{F}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

thus $\mathbf{a}^* = \mathbf{a}\cdot\mathbf{L}^{-1} = [2/\sqrt{3},\ (3/2) - 1/\sqrt{3},\ 0]^{\mathrm{T}}$. The set of deformed parallel lines determined by the intersection of the planes, $\mathbf{a}^* = [2/\sqrt{3},\ (3/2) - 1/\sqrt{3},\ 0]^{\mathrm{T}}$ with $c = 0$ and 5, and $x_3 = 0$. This set of parallel lines are given by $2x_1/\sqrt{3} + ((3/2) - 1/\sqrt{3})x_2 = 0$ and $2x_1/\sqrt{3} + ((3/2) - 1/\sqrt{3})x_2 = 5$ and are sketched in Fig. 11.3b.

The effect of homogeneous deformations on ellipsoids is similar. Recall that the Cartesian equation for an ellipsoid is

$$\frac{X_I^2}{a^2} + \frac{X_{II}^2}{b^2} + \frac{X_{III}^2}{c^2} = 1, \tag{11.5}$$

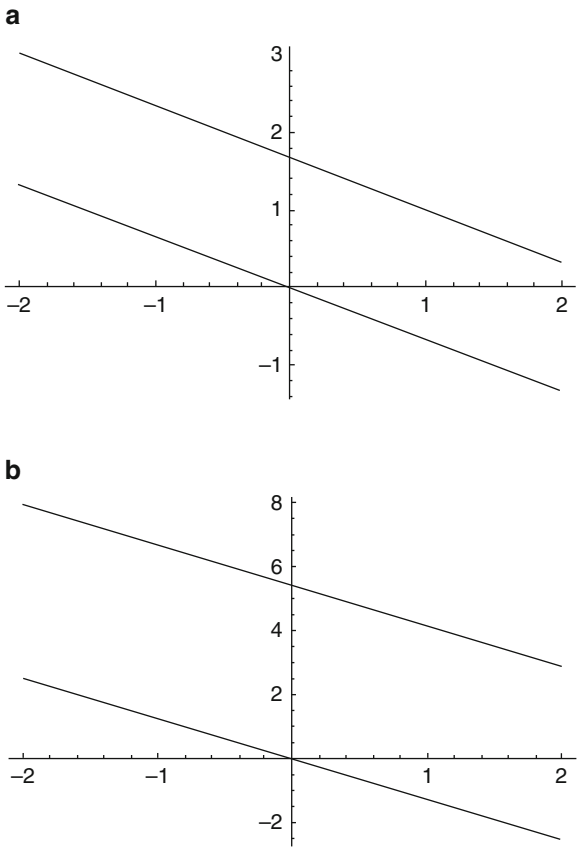


Fig. 11.3 An illustration of Example 11.2.1, a set of parallel lines deforming into a set of parallel lines

where the constants a, b, and c represent the intersection points of the ellipsoid on the Cartesian coordinates axes. In material coordinates an ellipsoid has the representation

$$\mathbf{X} \cdot \mathbf{A} \cdot \mathbf{X} = 1, \tag{11.6}$$

where $\mathbf{A}$ is a constant second rank tensor. To see that (11.5) and (11.6) are equivalent representations of an ellipsoid, let $\mathbf{A}$ be in its principal coordinate system and set $A_{11} = a^{-2}$, $A_{22} = b^{-2}$, $A_{33} = c^{-2}$. Substituting the inverse of (11.2), $\mathbf{X} = \mathbf{L}^{-1}\,\mathbf{x}$, into (11.6) yields

$$\mathbf{x} \cdot \mathbf{A}^* \cdot \mathbf{x} = 1, \tag{11.7}$$

where a new constant second rank tensor $\mathbf{A}^*$ has been defined by the transformation $\mathbf{A}^* = (\mathbf{L}^{-1})^{\mathrm{T}} \cdot \mathbf{A} \cdot \mathbf{L}^{-1}$. Eq. (11.7) is also the equation of an ellipsoid, an ellipsoid in the spatial coordinate system, thus permitting us to conclude that an ellipsoid in the material system is deformed into an ellipsoid in the spatial system by a

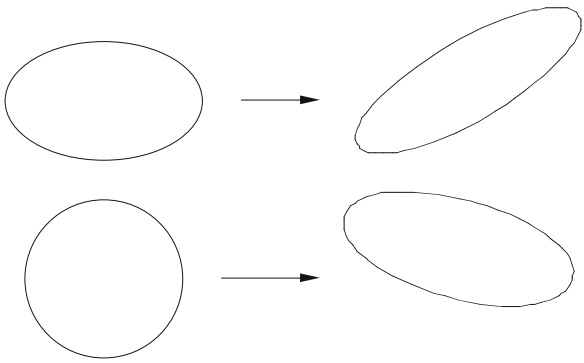


Fig. 11.4 An illustration of an ellipse being deformed into an ellipse and of a circle being deformed into an ellipse

homogeneous deformation. In particular, spheres will deform into ellipsoids and, in planar deformations, ellipses into ellipses (or circles into ellipses). This result is illustrated in the planar case where ellipses deform into ellipses or circles into ellipses in Fig. 11.4.

Example Problem 11.2.2
Draw a sketch of the ellipse given by $(X_I^2/9) + (X_{II}^2/4) = 1$, and then draw a sketch of the same ellipse after it was subjected to the homogeneous deformation of Example 11.2.1.

Solution: A sketch of the ellipse given by $(X_I^2/9) + (X_{II}^2/4) = 1$ is shown in Fig. 11.5a. Using the inverse of the homogeneous deformation **F** determined in Example 11.2.1, the tensor **A** representing the ellipse $(X_I^2/9) + (X_{II}^2/4) = 1$,

$$\mathbf{A} = \begin{bmatrix} \frac{1}{9} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

is transformed into $\mathbf{A}^*$, $\mathbf{A}^* = (\mathbf{L}^{-1})^{\mathrm{T}} \cdot \mathbf{A} \cdot \mathbf{L}^{-1}$,

$$\mathbf{A}^* = \frac{1}{54} \begin{bmatrix} 2 & -1 & 0 \\ -1 & \frac{31}{8} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and the deformed ellipse is given by $(1/27)(x_1^2 + (31/16)x_2^2 - 2x_1x_2) = 1$**.** A sketch of the deformed ellipse is shown in Fig. 11.5b.

The geometric results of this example give an intuitive insight into a number of tissue deformation situations. Consider the case when a circle is inscribed on a tissue and the tissue is then greatly deformed. Fig. 11.6 illustrates the deformation of a circle with a ratio of principal axes of 1:1 through ten steps to a ratio of principal axes of 1:19; this is the deformation of a circle inscribed on a surface subjected to a homogeneous deformation. At a ratio of principal axes of 1:19 the

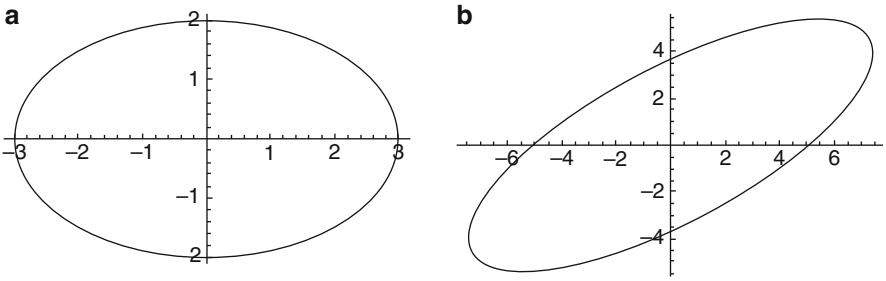


Fig. 11.5 An illustration of Example 11.2.2, an ellipse deforming into an ellipse. (a) left, before deformation, (b) right, after deformation

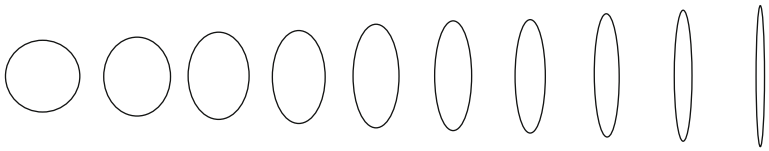


Fig. 11.6 An illustration of a circle deforming into an ellipse by an increasingly unequal biaxial extension. The deformation deforms a principal axes ratio of 1:1 to one of 19:1

deformed circle looks more like a crack than an ellipse or the original circle. It is not an exact analogy but one can, from this result, imagine what happens to the circular hole punched in a cadaver. The skin and epidermis of the cadaver are under a natural tension that quickly deforms the circular hole through the skin into the shape of an elongated ellipse as shown in the right most panel in Fig. 11.6. When these elongated ellipse are connected end to end to form curves they are called split lines, called Langer lines. Langer lines are illustrated in Fig. 11.7 (Danielson 1973; Danielson and Natarajan 1975).

A most significant point concerning homogeneous deformation is that any deformation in a sufficiently small neighborhood of a point is a homogeneous deformation. This may be mathematically verified by expanding the motion $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X},t)$ in a Taylor series in $\mathbf{X}$ about the point $\mathbf{X}_\circ$ and retaining only the first order term, which is the deformation gradient evaluated at $\mathbf{X}_\circ$. Thus, for a sufficiently small domain about the point $\mathbf{X}_\circ$, the deformation gradient evaluated at $\mathbf{X}_\circ$ represents the deformation. Since this deformation gradient is independent of $\mathbf{X}$, the deformation in this small domain is a homogeneous deformation. One easy way to illustrate the general truth of this mathematical result is to draw a small circle or parallelepiped on the skin with removable ink, and then apply a deformation to the skin surface. Alternatively, the small circle can be drawn on a rubber eraser and the eraser may be deformed to visualize the transition. The circle is easily seen to deform into an ellipse and must make an effort to obtain a non-elliptical shape. The same is true of parallelepipeds.

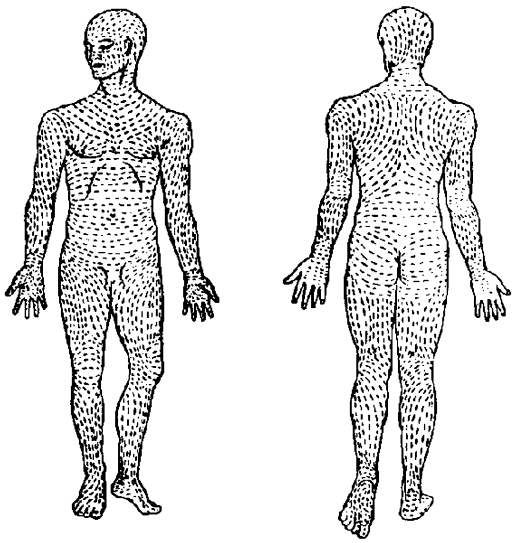


Fig. 11.7 Langer lines on a cadaver. From Danielson (1973)

Problems

11.2.1 Draw a sketch of the set of parallel lines given by $\mathbf{a} = [-2, 1, 0]^{\mathrm{T}}$ and $c = 0$ and 5, $-2X_I + X_{II} = 0$ and $-2X_I + X_{II} = 5$, then draw a sketch of the set of parallel lines after subjecting them to the homogeneous deformation of Example 11.2.1.

11.2.2 Draw a sketch of the ellipse given by $(X_I^2/16) + X_{II}^2 = 1$, and then draw a sketch of the same ellipse after it was subjected to the homogeneous deformation of Example 11.2.1.

11.2.3 Show that the deformation $x_1 = (9/4)X_I$, $x_2 = X_{II}$, $x_1 = X_{II}$ carries the ellipse $(X_I/4)^2 + (X_{II}/9)^2 = 1$ into the circle $(x_1/9)^2 + (x_2/9)^2 = 1$ and that the inverse deformation carries the circle $(X_I/4)^2 + (X_{II}/4)^2 = 1$ into the ellipse $(x_1/9)^2 + (x_2/4)^2 = 1$. Provide a sketch of the undeformed and deformed ellipses and circles.

11.2.4 Why is it not possible for an ellipse to deform into a hyperbola?

11.3 Polar Decomposition of the Deformation Gradients

It can be shown that the deformation gradient $\mathbf{F}$ can be algebraically decomposed in two ways into a pure deformation and a pure rotation. This decomposition is multiplicative and is written

$$\mathbf{F} = \mathbf{R} \cdot \mathbf{U} = \mathbf{V} \cdot \mathbf{R}, \tag{11.8}$$

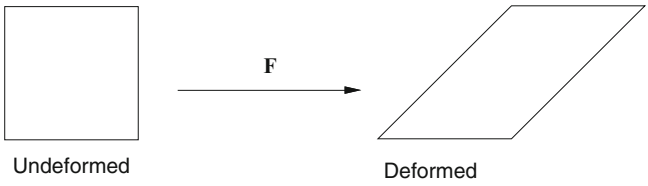


Fig. 11.8 The deformation of a square by **F**

where **R** is an orthogonal tensor ($\mathbf{R}^{\mathrm{T}}\cdot\mathbf{R} = \mathbf{R}\cdot\mathbf{R}^{\mathrm{T}} = \mathbf{1}$) representing the rotation and called the rotation tensor and **U** and **V** are called the right and left stretch tensors, respectively. Both **U** and **V** represent the same pure deformation, but in different ways that will be demonstrated. The right and left stretch tensors, **U** and **V**, are related to **F** by

$$\mathbf{U} = \sqrt{\mathbf{F}^T \cdot \mathbf{F}},\ \mathbf{V} = \sqrt{\mathbf{F} \cdot \mathbf{F}^T}. \tag{11.9}$$

In order to define the square root of a tensor involved in (11.9), the tensor must be symmetric and positive definite. In that case the square root is constructed by transforming the tensor to its principal axes where the eigenvalues are all positive, then the square root of the tensor is the diagonalized tensor coincident with the principal axes but containing the square roots of the eigenvalues along the diagonal. To show that the definitions (11.9) are reasonable it should be shown that the tensors $\mathbf{U}^2$ and $\mathbf{V}^2$ are positive definite,

$$\mathbf{U}^2 = \mathbf{F}^T \cdot \mathbf{F},\ \mathbf{V}^2 = \mathbf{F} \cdot \mathbf{F}^T. \tag{11.10}$$

The positive definite character of $\mathbf{U}^2$ may be seen by letting it operate on the vector **a**, then taking the scalar product with **a**, thus

$$\mathbf{a} \cdot \mathbf{U}^2 \cdot \mathbf{a} = \mathbf{a} \cdot \mathbf{F}^T \cdot \mathbf{F} \cdot \mathbf{a} = (\mathbf{F} \cdot \mathbf{a}) \cdot (\mathbf{F} \cdot \mathbf{a}) \geq 0, \tag{11.11}$$

where the fact that $\mathbf{a}\cdot\mathbf{F}^{\mathrm{T}} = (\mathbf{F}\cdot\mathbf{a})$ has been used. A similar proof will show the positive definite character of $\mathbf{V}^2$. The fact that the tensor **R** is orthogonal follows from the definitions of **U** and/or **V**. From (11.8) **R** is given by $\mathbf{R} = \mathbf{F}\cdot\mathbf{U}^{-1}$ (or $\mathbf{R} = \mathbf{V}^{-1}\cdot\mathbf{F}$) thus the calculation of $\mathbf{R}^{\mathrm{T}}\cdot\mathbf{R}$ (or $\mathbf{R}\cdot\mathbf{R}^{\mathrm{T}}$) yields

$$\mathbf{R}^T \cdot \mathbf{R} = \mathbf{U}^{-1} \cdot \mathbf{F}^T \cdot \mathbf{F} \cdot \mathbf{U}^{-1} = \mathbf{U}^{-1} \cdot \mathbf{U}^2 \cdot \mathbf{U}^{-1} = \mathbf{1},$$

where the fact that **U** and its inverse are symmetric and the definition (11.10) have been employed. As an example of this decomposition consider the **F** associated with a simple shearing deformation illustrated in Fig. 11.8. The decomposition of this simple shearing deformation is shown in Fig. 11.9

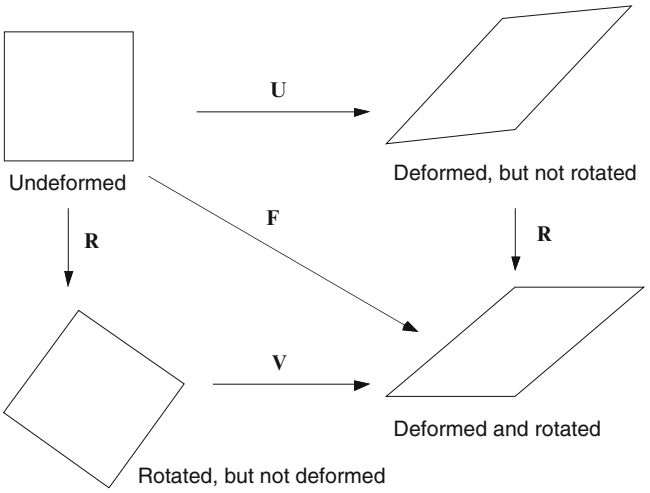


Fig. 11.9 The deformation of a square by **F**, illustrating the polar decomposition of **F**

Example Problem 11.3.1
Determine the polar decomposition of the deformation gradient tensor

$$\mathbf{F} = \begin{bmatrix} \sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: The squares or the right and left stretch tensors are calculated directly from **F**, thus

$$\mathbf{U}^2 = \mathbf{F}^T \cdot \mathbf{F} = \begin{bmatrix} 3 & \sqrt{3} & 0 \\ \sqrt{3} & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathbf{V}^2 = \mathbf{F} \cdot \mathbf{F}^T = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The square roots of these two tensors are constructed by transforming the tensor to its principal axes where the eigenvalues are all positive, then the square root of the tensor is the diagonalized tensor coincident with the principal axes but containing the square roots of the eigenvalues along the diagonal,

$$\mathbf{U} = \begin{bmatrix} \sqrt{3}A & \sqrt{3}B & 0 \\ \sqrt{3}B & 2A + B & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathbf{V} = \begin{bmatrix} 2A & 2B & 0 \\ 2B & 2A & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where

$$A = \cos 15^\circ = \frac{\sqrt{3}+1}{2\sqrt{2}},\ B = \sin 15^\circ = \frac{\sqrt{3}-1}{2\sqrt{2}}.$$

The fact that $\mathbf{U}$ and $\mathbf{V}$ given above are the square roots of $\mathbf{U}^2$ and $\mathbf{V}^2$, respectively, many be verified simply by squaring $\mathbf{U}$ and $\mathbf{V}$. The orthogonal tensor $\mathbf{R}$ may be computed in several ways,

$$\mathbf{R} = \mathbf{V}^{-1}\cdot\mathbf{F} = \mathbf{F}\cdot\mathbf{U}^{-1} = \begin{bmatrix} A & B & 0 \\ -B & A & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The polar right and left decompositions of the given deformation gradient tensor are then given by

$$\mathbf{F} = \mathbf{R}\cdot\mathbf{U} = \begin{bmatrix} A & B & 0 \\ -B & A & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{3}A & \sqrt{3}B & 0 \\ \sqrt{3}B & 2A+B & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and

$$\mathbf{F} = \mathbf{V}\cdot\mathbf{R} = \begin{bmatrix} 2A & 2B & 0 \\ 2B & 2A & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} A & B & 0 \\ -B & A & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

respectively.

Example Problem 11.3.2
Develop a geometric interpretation of the deformation gradient tensor $\mathbf{F}$ of Example Problem 11.3.1 by considering it as representing a homogeneous deformation $\mathbf{x} = \mathbf{F}\cdot\mathbf{X}$ acting on a unit square with vertices (0, 0), (1, 0), (1, 1), (0, 1).

Solution: The scalar equations equivalent to the homogeneous deformation $\mathbf{x} = \mathbf{F}\cdot\mathbf{X}$ where $\mathbf{F}$ is given in the statement of Example 11.3.1 are

$$x_1 = \sqrt{3}X_I + X_{II},\ x_2 = 2X_{II},\ x_3 = X_{III}.$$

This unit square is deformed by $\mathbf{F}$ into a parallelogram with corners at the points (0, 0), ($\sqrt{3}$, 0), (1 + $\sqrt{3}$, 2), and (1, 2) as illustrated in Fig. 11.10. Consider the left decomposition, $\mathbf{F} = \mathbf{V}\cdot\mathbf{R}$, first. In this decomposition the rotation $\mathbf{R}$ is applied first, then the left deformation or left stretch, $\mathbf{V}$. The effect of $\mathbf{R}$ on the unit square is a clockwise rotation of 15°; this is illustrated in Fig. 11.11. Following this rotation of the unit square, there is a left stretch $\mathbf{V}$ that carries the rotated square into the deformed shape illustrated in Fig. 11.9. The other decomposition choice $\mathbf{F} = \mathbf{R}\cdot\mathbf{U}$ reverses the

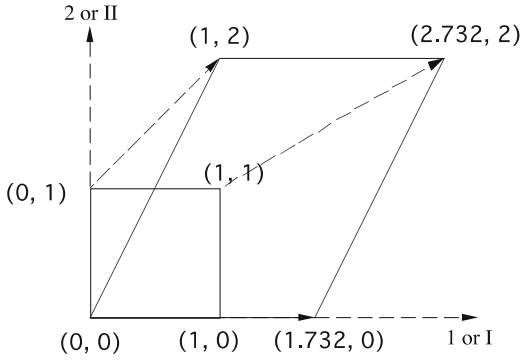


Fig. 11.10 The deformation of a unit square by the **F** of Example 11.3.1

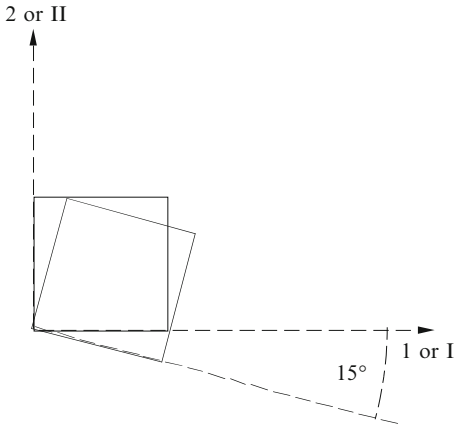


Fig. 11.11 The deformation of a unit square by the **R** of Example 11.3.1

order of the deformation and the rotation. The deformation or right stretch **U** is first applied to the unit square and it deforms the square into the shape illustrated in Fig. 11.12. The clockwise rotation of 15° is then applied and it rotates the deformed shape illustrated in Fig. 11.9 into its final position illustrated in Fig. 11.8

Problems

11.3.1. Prove that $\mathbf{a}\cdot\mathbf{F}^{\mathrm{T}} = \mathbf{F}\cdot\mathbf{a}$.

11.3.2. Using the polar decomposition theorem, $\mathbf{F} = \mathbf{R}\cdot\mathbf{U} = \mathbf{V}\cdot\mathbf{R}$, find the **U** and **R** associated with the **F** given by $\mathbf{F} = \sqrt{2}\begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$.

11.3.3. Find the square root of the matrix

$$\mathbf{V}^2 = \begin{bmatrix} 4 & 1 & -1 \\ 1 & 6 & 0 \\ -1 & 0 & 2 \end{bmatrix}.$$

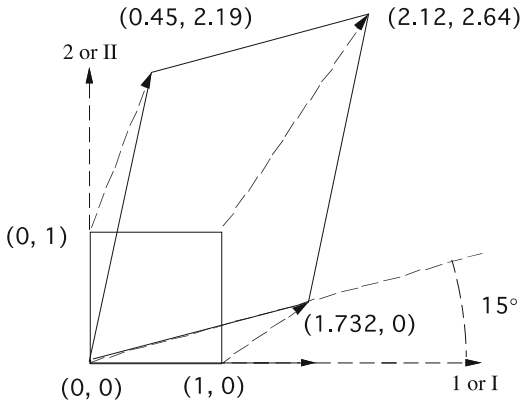


Fig. 11.12 The deformation of a unit square by the **U** of Example 11.3.1

11.3.4. Find the square root of the matrix

$$\mathbf{A} = \begin{bmatrix} 31 & -5\sqrt{3} & 0 \\ -5\sqrt{3} & 21 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

11.3.5. Find the polar decomposition of the tensor **F**,

$$\mathbf{F} = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

11.4 The Strain Measures for Large Deformations

The deformation gradient is the basic measure of local deformational and rotational motion. It maps a small region of the undeformed object into a small region of the deformed object. If the motion is a pure translation with no rotation, then $\mathbf{F} = \mathbf{1}$. If the motion is a rigid object rotation, then $\mathbf{F} = \mathbf{R}$ where **R** is an orthogonal matrix, $\mathbf{R}\cdot\mathbf{R}^{\mathrm{T}} = \mathbf{R}^{\mathrm{T}}\cdot\mathbf{R} = \mathbf{1}$.

The local state of deformation may be investigated by considering the deformation of an infinitesimal material filament denoted by d**X**. In the instantaneous configuration the same material filament has a position represented by d**x**. Recalling the representation (2.2) for a motion, $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X},t)$, and the representation (2.16) for **F**, it is easy to see that d**x** and d**X** are related by $d\mathbf{x} = \mathbf{F}\cdot d\mathbf{X}$, or from the polar decomposition theorem $d\mathbf{x} = \mathbf{R}\cdot\mathbf{U}\cdot d\mathbf{X} = \mathbf{V}\cdot\mathbf{R}\cdot d\mathbf{X}$. Thus a pure rotational motion for which $\mathbf{U} = \mathbf{V} = \mathbf{1}$ the length of d**X** will be preserved but its direction will be rotated. If the motion includes a deformational component, then the length of d**x**

will be different from the length of d$\mathbf{X}$, in general. To calculate this change in length we denote the square of the final length by $ds^2 = d\mathbf{x}\cdot d\mathbf{x}$ and the square of the initial length by $dS^2 = d\mathbf{X}\cdot d\mathbf{X}$. The difference in the squares of these length changes is written $ds^2 - dS^2 = d\mathbf{x}\cdot d\mathbf{x} - d\mathbf{X}\cdot d\mathbf{X}$. This expression for $ds^2 - dS^2$ can be written entirely in terms of the material filament d$\mathbf{X}$ or entirely in terms of its image d$\mathbf{x}$ at time t. To accomplish both of these objectives we observe that, using $d\mathbf{x} = \mathbf{F}\cdot d\mathbf{X}$, $d\mathbf{x}\cdot d\mathbf{x}$ has the representation

$$d\mathbf{x}\cdot d\mathbf{x} = (\mathbf{F}\cdot d\mathbf{X})\cdot(\mathbf{F}\cdot d\mathbf{X}) = d\mathbf{X}\cdot\mathbf{F}^T\cdot\mathbf{F}\cdot d\mathbf{X}, \tag{11.12}$$

and that $d\mathbf{X} = \mathbf{F}^{-1}\cdot d\mathbf{x}$, $d\mathbf{X}\cdot d\mathbf{X}$ has the representation

$$d\mathbf{X}\cdot d\mathbf{X} = (\mathbf{F}^{-1}\cdot d\mathbf{x})\cdot(\mathbf{F}^{-1}\cdot d\mathbf{x}) = d\mathbf{x}\cdot\left(\mathbf{F}^{-1}\right)^T\cdot\mathbf{F}^{-1}d\mathbf{x}. \tag{11.13}$$

The two formulas may then be derived from the expression $ds^2 - dS^2 = d\mathbf{x}\cdot d\mathbf{x} - d\mathbf{X}\cdot d\mathbf{X}$, one by substituting for $d\mathbf{x}\cdot d\mathbf{x}$ from (11.12) and the other by substituting for $d\mathbf{X}\cdot d\mathbf{X}$ from (11.13), thus

$$\begin{aligned} ds^2 - dS^2 &= d\mathbf{x}\cdot d\mathbf{x} - d\mathbf{X}\cdot d\mathbf{X} = \{\mathbf{F}^T\cdot\mathbf{F} - \mathbf{1}\}\, d\mathbf{X}\cdot d\mathbf{X} \\ &= \{\mathbf{1} - \left(\mathbf{F}^{-1}\right)^T\cdot\mathbf{F}^{-1}\}\, d\mathbf{x}\cdot d\mathbf{x}. \end{aligned} \tag{11.14}$$

The Lagrangian or material strain tensor $\mathbf{E}$ and the Eulerian or spatial strain tensor $\mathbf{e}$ are defined by

$$\mathbf{E} = \left(\frac{1}{2}\right)\{\mathbf{F}^T\cdot\mathbf{F} - -\mathbf{1}\},\ \mathbf{e} = \left(\frac{1}{2}\right)\{\mathbf{1} - (\mathbf{F}^{-1})^T\cdot\mathbf{F}^{-1}\}. \tag{11.15}$$

In terms of the strain tensors the change $ds^2 - dS^2$ takes the form

$$ds^2 - dS^2 = 2d\mathbf{X}\cdot\mathbf{E}\cdot d\mathbf{X} = 2d\mathbf{x}\cdot\mathbf{e}\cdot d\mathbf{x}. \tag{11.16}$$

Clearly, if either of these strain tensors is zero, then so is the other and there is no change in length for any material filament, $ds^2 = dS^2$.

The Lagrangian strain tensor $\mathbf{E}$ and Eulerian strain tensor $\mathbf{e}$ are defined for large strains and represented in terms of the deformation gradient and its inverse in (11.15). A component representation of these two tensors in terms of the displacement vector $\mathbf{u}$ will now be obtained. From (2.20) and Fig. 2.4, $\mathbf{u}$ is given by $\mathbf{u} = \mathbf{x} - \mathbf{X} + \mathbf{c}$. If $\mathbf{u}$ is referred to the spatial coordinate system and the spatial gradient taken, or if $\mathbf{u}$ is referred to the material coordinate system and the material gradient taken (Sect. 2.2), then

$$[\nabla_{\mathbf{o}}\otimes\mathbf{u}(\mathbf{X},t)]^T = \mathbf{F}(\mathbf{X},t) - \mathbf{1} \text{ and } [\nabla\otimes\mathbf{u}(\mathbf{x},t)]^T = \mathbf{1} - \mathbf{F}^{--1}(\mathbf{x},t). \tag{2.23}$$

repeated

These equations are then solved for $\mathbf{F}$ and $\mathbf{F}^{-1}$, thus

$$\mathbf{F}(\mathbf{X},t) = \mathbf{1} + [\nabla_{\mathbf{o}} \otimes \mathbf{u}(\mathbf{X},t)]^T \text{ and } \mathbf{F}^{-1}(\mathbf{x},t) = \mathbf{1} - [\nabla \otimes \mathbf{u}(\mathbf{x},t)]^T. \tag{11.17}$$

The Lagrangian strain tensor $\mathbf{E}$ and the Eulerian strain tensor $\mathbf{e}$ are related to the displacement gradients by substituting the two equations (11.17) into the two equations (11.15), thus

$$\mathbf{E} = (1/2)\{[\nabla_{\mathbf{o}} \otimes \mathbf{u}(\mathbf{X},t)]^T + [\nabla_{\mathbf{o}} \otimes \mathbf{u}(\mathbf{X},t)] + [\nabla_{\mathbf{o}} \otimes \mathbf{u}(\mathbf{X},t)][\nabla_{\mathbf{o}} \otimes \mathbf{u}(\mathbf{X},t)]^T\}, \tag{11.18}$$

$$\mathbf{e} = (1/2)\{[\nabla \otimes \mathbf{u}(\mathbf{x},t)]^T + [\nabla \otimes \mathbf{u}(\mathbf{x},t)] - [\nabla \otimes \mathbf{u}(\mathbf{x},t)][\nabla \otimes \mathbf{u}(\mathbf{x},t)]^T\}. \tag{11.19}$$

The expanded component forms of (11.18) and (11.19) are given by

$$E_{I\,I} = \frac{\partial u_I}{\partial x_I} + \frac{1}{2}\left[\left\{\frac{\partial u_I}{\partial x_I}\right\}^2 + \left\{\frac{\partial u_{II}}{\partial x_I}\right\}^2 + \left\{\frac{\partial u_{III}}{\partial x_I}\right\}^2\right],$$

$$E_{II\,II} = \frac{\partial u_{II}}{\partial x_{II}} + \frac{1}{2}\left[\left\{\frac{\partial u_I}{\partial x_{II}}\right\}^2 + \left\{\frac{\partial u_{II}}{\partial x_{II}}\right\}^2 + \left\{\frac{\partial u_{III}}{\partial x_{II}}\right\}^2\right],$$

$$E_{III\,III} = \frac{\partial u_{III}}{\partial x_{III}} + \frac{1}{2}\left[\left\{\frac{\partial u_I}{\partial x_{III}}\right\}^2 + \left\{\frac{\partial u_{II}}{\partial x_{III}}\right\}^2 + \left\{\frac{\partial u_{III}}{\partial x_{III}}\right\}^2\right],$$

$$E_{I\,II} = \frac{1}{2}\left[\frac{\partial u_I}{\partial x_{II}} + \frac{\partial u_{II}}{\partial x_I} + \frac{\partial u_I}{\partial x_I}\frac{\partial u_{II}}{\partial x_I} + \frac{\partial u_{II}}{\partial x_I}\frac{\partial u_{II}}{\partial x_{II}} + \frac{\partial u_{III}}{\partial x_I}\frac{\partial u_{III}}{\partial x_{II}}\right], \tag{11.20}$$

$$E_{I\,III} = \frac{1}{2}\left[\frac{\partial u_I}{\partial x_{III}} + \frac{\partial u_{III}}{\partial x_I} + \frac{\partial u_I}{\partial x_I}\frac{\partial u_{III}}{\partial x_I} + \frac{\partial u_{II}}{\partial x_I}\frac{\partial u_{II}}{\partial x_{III}} + \frac{\partial u_{III}}{\partial x_I}\frac{\partial u_{III}}{\partial x_{III}}\right],$$

$$E_{II\,III} = \frac{1}{2}\left[\frac{\partial u_{II}}{\partial x_{III}} + \frac{\partial u_{III}}{\partial x_{II}} + \frac{\partial u_I}{\partial x_{II}}\frac{\partial u_{II}}{\partial x_{III}} + \frac{\partial u_{II}}{\partial x_{III}}\frac{\partial u_{II}}{\partial x_{III}} + \frac{\partial u_{III}}{\partial x_{II}}\frac{\partial u_{III}}{\partial x_{III}}\right],$$

and

$$e_{11} = \frac{\partial u_1}{\partial x_1} - \frac{1}{2}\left[\left\{\frac{\partial u_1}{\partial x_1}\right\}^2 + \left\{\frac{\partial u_2}{\partial x_1}\right\}^2 + \left\{\frac{\partial u_3}{\partial x_1}\right\}^2\right],$$

$$e_{22} = \frac{\partial u_2}{\partial x_2} - \frac{1}{2}\left[\left\{\frac{\partial u_1}{\partial x_2}\right\}^2 + \left\{\frac{\partial u_2}{\partial x_2}\right\}^2 + \left\{\frac{\partial u_3}{\partial x_2}\right\}^2\right],$$

$$e_{33} = \frac{\partial u_3}{\partial x_3} - \frac{1}{2}\left[\left\{\frac{\partial u_1}{\partial x_3}\right\}^2 + \left\{\frac{\partial u_2}{\partial x_3}\right\}^2 + \left\{\frac{\partial u_3}{\partial x_3}\right\}^2\right],$$

$$e_{12} = \frac{1}{2}\left[\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_1}\frac{\partial u_2}{\partial x_1} - \frac{\partial u_2}{\partial x_1}\frac{\partial u_2}{\partial x_2} - \frac{\partial u_3}{\partial x_1}\frac{\partial u_3}{\partial x_2}\right], \tag{11.21}$$

$$e_{13} = \frac{1}{2}\left[\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} - \frac{\partial u_1}{\partial x_1}\frac{\partial u_3}{\partial x_1} - \frac{\partial u_2}{\partial x_1}\frac{\partial u_2}{\partial x_3} - \frac{\partial u_3}{\partial x_1}\frac{\partial u_3}{\partial x_3}\right],$$

$$e_{23} = \frac{1}{2}\left[\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2} - \frac{\partial u_1}{\partial x_2}\frac{\partial u_2}{\partial x_3} - \frac{\partial u_2}{\partial x_3}\frac{\partial u_2}{\partial x_3} - \frac{\partial u_3}{\partial x_2}\frac{\partial u_3}{\partial x_3}\right],$$

respectively. The products of the displacement gradients appear in Eqs. (11.20) and (11.21) for the strain tensor **E** and **e**. This is called geometrical nonlinearity to distinguish it from the physical or constitutive nonlinearity (e.g. the relation between stress and strain) that will be considered later in this chapter. If the deformation of the object is so small that the square of the displacement gradients can be neglected, and thus the difference between the material and spatial coordinates, then both the Lagrangian strain tensor **E** and the Eulerian strain tensor **e** coincide with the infinitesimal strain tensor (2.44).

The geometrical interpretation of the Lagrangian strain tensor **E** and the Eulerian strain tensor **e** are algebraically straightforward, but not very simple geometrically. If δ_I represents the change in length per unit length in the X_I direction, then the deformation gradient **F** and the Lagrangian strain tensor **E** are given by

$$\begin{aligned}\mathbf{F} &= \mathbf{1} + \delta_I \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \mathbf{E} = \frac{1}{2}\{\mathbf{F}^T \cdot \mathbf{F} - \mathbf{1}\} \\ &= \left(\delta_I + \frac{1}{2}\delta_I^2\right)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};\end{aligned} \tag{11.22}$$

thus all the components of **E** are zero except for,

$$E_{II} = \left(\delta_I + \frac{1}{2}\delta_I^2\right). \tag{11.23}$$

If δ_{II} represents the change in length per unit length in the X_{II} direction, and sin ϕ indicates a shear, then the deformation gradient is given by

$$\mathbf{F} = \begin{bmatrix} 1+\delta_I & (1+\delta_I)\sin\phi & 0 \\ 0 & 1+\delta_{II} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{11.24}$$

and the Lagrangian strain tensor component $E_{I\,II}$ is related to δ_I, δ_{II} and to the change in the right angle ϕ between the filaments lying in the I and II directions by

$$E_{I\,II} = \left(\frac{1}{2}\right)(1+\delta_I)(1+\delta_{II})\sin\phi. \tag{11.25}$$

Thus, unless the extensions δ_I, δ_{II}, and ϕ are small so that the square of each may be neglected, the traditional geometric interpretation for small strains,

$$E_{II} \approx \delta_I, \quad E_{I\,II} \approx \frac{\phi}{2}, \tag{11.26}$$

namely that $E_{I\,I}$ is the extension in the I direction and $E_{I\,II}$ is one-half the change in an angle that was originally a right angle, is not accurate. Eqs. (11.23) and (11.24) representing $E_{I\,I}$ and $E_{I\,II}$ in terms of δ_I, δ_{II}, and ϕ show that the geometric interpretation of the finite strain tensors in terms of extensions and changes in right angles is possible, but is awkward and not very useful due to its nonlinear nature.

Example 11.4.1
Compute the Lagrangian strain tensor $\mathbf{E}$ and the Eulerian strain tensor $\mathbf{e}$ for the motion given by (2.12). Determine the range of values of t for which these two strain measures coincide in this special motion.

Solution: The deformation gradient and inverse deformation gradient for the motion (2.12) are given in Example 2.1.1,

$$\mathbf{F} = \begin{bmatrix} 1+t & t & 0 \\ t & 1+t & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{F}^{-1} = \frac{1}{1+2t}\begin{bmatrix} 1+t & -t & 0 \\ -t & 1+t & 0 \\ 0 & 0 & 1+2t \end{bmatrix}.$$

These values for $\mathbf{F}$ and $\mathbf{F}^{-1}$ can then be substituted into (11.15) for the Lagrangian strain tensor $\mathbf{E}$ and the Eulerian strain tensor $\mathbf{e}$, thus

$$\mathbf{E} = t(1+t)\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{e} = \frac{t(1+t)}{(1+2t)^2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

These expressions are, of course, valid for large strains. If we restrict ourselves to small strains, then the two strain tensors must coincide. Since each component of $\mathbf{E}$ is proportional to $t + t^2$ and each component of $\mathbf{e}$ is proportional to

$$\frac{t(1+t)}{(1+2t)^2} = t - 3t^2 + 8t^3 - \cdots$$

the two strain tensors coincide only if terms of the order t^2 and higher are neglected. Thus the motion (6) is one of small strain only when all terms of order t^2 and higher are neglected.

There are several other strain measures used in the development and analysis of finite deformations. Two of the most widely used are the right and left Cauchy–Green tensors, **C** and **B**, respectively. These two tensors are simply the squares of the right and left stretch tensors, **U** and **V**,

$$\mathbf{C} = \mathbf{U}^2 = \mathbf{F}^T \cdot \mathbf{F}, \ \mathbf{B} = \mathbf{V}^2 = \mathbf{F} \cdot \mathbf{F}^T. \tag{11.27}$$

It is also convenient to introduce the inverse of the left Cauchy–Green tensor denoted by **c**,

$$\mathbf{c} = \mathbf{B}^{-1} = \mathbf{V}^{-2} = \left(\mathbf{F} \times \mathbf{F}^T\right)^{-1}. \tag{11.28}$$

The Lagrangian strain tensor **E** and the Eulerian strain tensor **e** are expressed in terms of **C** and **c** by the following formulas that follow from (11.15), and the definitions of **C** and **c** given as (11.27) and (11.28) above,

$$\mathbf{2E} = \mathbf{C} - \mathbf{1}, \ \mathbf{2e} = \mathbf{1} - \mathbf{c}. \tag{11.29}$$

The eigenvalues of the various strain measures may be interpreted using the concept of *stretch*. The stretch $\lambda_{(\mathbf{N})}$ in the fiber coincident with d**X** is defined by

$$\lambda_{(\mathbf{N})} = \sqrt{\frac{\mathbf{dx} \cdot \mathbf{dx}}{\mathbf{dX} \cdot \mathbf{dX}}}, \tag{11.30}$$

where **N** is a unit vector in the direction of d**X**. The related concept of extension $\delta_{(\mathbf{N})}$ is then defined in terms of stretch by

$$\delta_{(\mathbf{N})} = \lambda_{(\mathbf{N})} - 1. \tag{11.31}$$

As a illustration of homogeneous deformations, and of the relationship of the stretch concept to the various strain measures that have been introduced, the special case of *pure homogeneous deformation*s is considered. A pure homogeneous deformation is a deformation for which the rotation $\mathbf{R} = \mathbf{1}$ and the deformation gradient tensor become symmetric, $\mathbf{F} = \mathbf{U} = \mathbf{V}$. In its principal coordinate system the deformation gradients of a pure homogeneous deformation have the representation

$$\mathbf{F} = \mathbf{U} = \mathbf{V} = \begin{bmatrix} \lambda_I & 0 & 0 \\ 0 & \lambda_{II} & 0 \\ 0 & 0 & \lambda_{III} \end{bmatrix}, \tag{11.32}$$

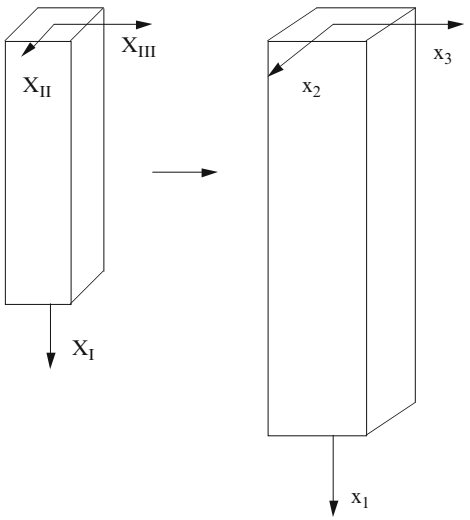


Fig. 11.13 An illustration of a simple extension, a type of pure homogeneous deformation

where λ_I, λ_{II}, λ_{III} are the principal stretches. The principal stretches represent the effect of the deformation upon material filaments in the coordinate directions as they are the ratios of the final length of the filament to the initial length of the filament. In terms of the principal stretches, the deformation tensors **C** and **c** have the representations

$$\mathbf{C} = \begin{bmatrix} \lambda_I^2 & 0 & 0 \\ 0 & \lambda_{II}^2 & 0 \\ 0 & 0 & \lambda_{III}^2 \end{bmatrix}, \ \mathbf{c} = \begin{bmatrix} \lambda_I^{-2} & 0 & 0 \\ 0 & \lambda_{II}^{-2} & 0 \\ 0 & 0 & \lambda_{III}^{-2} \end{bmatrix}, \tag{11.33}$$

and the Lagrangian strain tensor **E** and the Eulerian strain tensor **e**,

$$\begin{aligned} 2\mathbf{E} &= \begin{bmatrix} \lambda_I^2 - 1 & 0 & 0 \\ 0 & \lambda_{II}^2 - 1 & 0 \\ 0 & 0 & \lambda_{III}^2 - 1 \end{bmatrix}, \ 2\mathbf{e} \\ &= \begin{bmatrix} 1 - \lambda_I^{-2} & 0 & 0 \\ 0 & 1 - \lambda_{II}^{-2} & 0 \\ 0 & 0 & 1 - \lambda_{III}^{-2} \end{bmatrix}. \end{aligned} \tag{11.34}$$

Two special cases of pure homogeneous deformation are of particular interest. A *simple extension* is characterized by $\lambda_I \neq \lambda_{II} = \lambda_{III}$ and is illustrated in Fig. 11.13. A *uniform dilation* is characterized by $\lambda = \lambda_I = \lambda_{II} = \lambda_{III}$ and is illustrated in Fig. 11.14. In the case of uniform dilation, $\mathbf{F} = \mathbf{U} = \mathbf{V} = \lambda\mathbf{1}$, $\mathbf{C} = \lambda^2\mathbf{1}$, $\mathbf{c} = \lambda^{-2}\mathbf{1}$, $2\mathbf{E} = (\lambda^2 - 1)\mathbf{1}$, $2\mathbf{e} = (1 - \lambda^{-2})\mathbf{1}$.

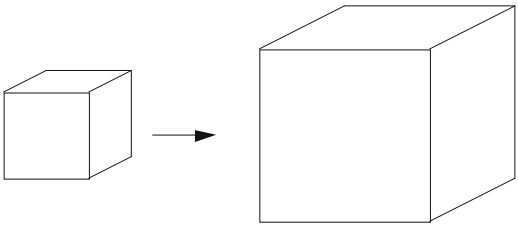


Fig. 11.14 An illustration of a uniform dilation, a type of pure homogeneous deformation

Problems

11.4.1. For the six motions of the form (2.10) given in Problem 2.1.1, namely 2.1.1 (a)–2.1.1(f), compute the Lagrangian strain tensor $\mathbf{E}$ and the Eulerian strain tensor $\mathbf{e}$, the right and left Cauchy–Green tensors, $\mathbf{C}$ and $\mathbf{B}$, respectively, and the inverse of the right Cauchy–Green tensor. Discuss briefly the significance of each of the tensors computed. In particular, explain the form or value of the deformation strain tensors in terms of the motion.

11.4.2. Prove that, in the case of no deformation, the invariants of $\mathbf{C}$ and $\mathbf{c}$ satisfy the following relationships: $I_{\mathbf{c}} = I_{\mathbf{C}} = II_{\mathbf{c}} = II_{\mathbf{C}} = 3$, $III_{\mathbf{c}} = III_{\mathbf{C}} = 1$.

11.4.3. Show that the Jacobian J is related to $III_{\mathbf{C}}$ by $J^2 = III_{\mathbf{C}}$.

11.4.4. A rectangular parallelepiped with a long dimension a_{o} and a square cross-section of dimension b_{o} is deformed by an axial tensile force P into a rectangular parallelepiped with a longer long dimension a and a smaller square cross-section of dimension b.

(a) What are the stretch ratios in the long direction (λ_{L}) and the transverse direction (λ_{T})?

(b) Express the volume of the deformed rectangular parallelepiped, V, as a function of the volume of the undeformed rectangular parallelepiped, V_{o}, and the stretch ratios in the long direction (λ_{L}) and the transverse direction (λ_{T}).

(c) Express the incompressibility condition for the rectangular parallelepiped as a function of the stretch ratios in the long direction (λ_{L}) and the transverse direction (λ_{T}).

(d) Express the cross-sectional area of the deformed rectangular parallelepiped, A, as a function of the cross-sectional area of the undeformed rectangular parallelepiped, A_{o}, and the stretch ratio in the transverse direction (λ_{T}).

11.5 Measures of Volume and Surface Change in Large Deformations

In this section we will consider volume and area measures of deformation. Consider volume deformation first. A material filament denoted by $\mathrm{d}\mathbf{X}$ is mapped into its present position $\mathrm{d}\mathbf{x}$ by the deformation gradient $\mathbf{F}$, $\mathrm{d}\mathbf{x} = \mathbf{F}\cdot\mathrm{d}\mathbf{X}$. By considering the

mapping of three nonplanar d$\mathbf{X}$'s, d$\mathbf{X}(q)$, d$\mathbf{X}(r)$, and d$\mathbf{X}(s)$, into their three-image d$\mathbf{x}$'s, d$\mathbf{x}(q)$, d$\mathbf{x}(r)$, and d$\mathbf{x}(s)$, a representation of the volumetric deformation may be obtained. This algebra uses the triple scalar product (Section A.8, equation (A116)) to calculate the volume associated with the parallelepiped defined by three vectors coincident with the three edges of the parallelepiped that come together at one vertex. The element of volume dV in the undeformed configuration is given by dV = d$\mathbf{X}(q)$·(d$\mathbf{X}(r)$ × d$\mathbf{X}(s)$), and in the deformed configuration by dv = d$\mathbf{x}(q)$·(d$\mathbf{x}(r)$ × d$\mathbf{x}(s)$). Substituting d$\mathbf{x}$ = $\mathbf{F}$·d$\mathbf{X}$ into dv = d$\mathbf{x}(q)$·(d$\mathbf{x}(r)$ × d$\mathbf{x}(s)$) three times, it follows that

$$dv = \{\mathbf{F} \cdot d\mathbf{X}(q)\} \cdot (\{\mathbf{F} \cdot d\mathbf{X}(r)\} \times \{\mathbf{F} \cdot d\mathbf{X}(s)\}), \tag{11.35}$$

which may be expanded and, using the fact that a determinant of a product of matrices is the product of the determinants, rewritten as

$$dv = JdV, \tag{11.36}$$

where J = Det $\mathbf{F}$ and dV = d$\mathbf{X}(q)$·(d$\mathbf{X}(r)$ × d$\mathbf{X}(s)$) are both determinants. Thus the element of volume dV in the undeformed configuration is deformed into a volume dv in the deformed configuration according to the rule dv = JdV where J = Det$\mathbf{F}$ is called the Jacobian of the deformation. In order that no region of positive finite volume be deformed into a region of zero or infinite volume it is required that $0 < J < \infty$.

Consider now the question of the deformation of differential elements of area where similar formulas for area change can be constructed. Let d$\mathbf{A}$ be a differential vector representation of area in the material reference frame obtained by taking the cross product of two different material filaments d$\mathbf{X}$, d$\mathbf{A}$ = d$\mathbf{X}(r)$ × d$\mathbf{X}(s)$. Similarly, let d$\mathbf{a}$ be a differential vector representation of area in the spatial reference frame, representing the deformed shape of the same material area, obtained by taking the cross product of the deformed images d$\mathbf{x}$, d$\mathbf{a}$ = d$\mathbf{x}(r)$ × d$\mathbf{x}(s)$ of the material filaments d$\mathbf{X}(r)$ and d$\mathbf{X}(s)$. The relationship between d$\mathbf{a}$ and d$\mathbf{A}$ is constructed by twice substituting d$\mathbf{x}$ = $\mathbf{F}$·d$\mathbf{X}$ into d$\mathbf{a}$ = d$\mathbf{x}(r)$ × d$\mathbf{x}(s)$,

$$d\mathbf{a} = \{\mathbf{F} \cdot d\mathbf{X}(r)\} \times \{\mathbf{F} \cdot d\mathbf{X}(s)\}. \tag{11.37}$$

The vector d$\mathbf{a}$ is then dotted with the deformation gradient $\mathbf{F}$ from the left, thus

$$d\mathbf{a} \cdot \mathbf{F} = \{\mathbf{F} \cdot d\mathbf{X}(r)\} \times \{\mathbf{F} \cdot d\mathbf{X}(s)\} \cdot \mathbf{F}. \tag{11.38}$$

The right-hand side of this equation may be expanded, as the one for volume was above, and, using the fact that a determinant of a product of matrices is the product of the determinants (Section A.8, page 372), rewritten as

$$d\mathbf{a} \cdot \mathbf{F} = Jd\mathbf{A}. \tag{11.39}$$

On the other hand, the relationship between d$\mathbf{A}$ and d$\mathbf{a}$ is constructed by twice substituting $d\mathbf{X} = \mathbf{F}^{-1}\cdot d\mathbf{x}$ into $d\mathbf{A} = d\mathbf{X}(r) \times d\mathbf{X}(s)$,

$$d\mathbf{A} = \{\mathbf{F}^{-1} \cdot d\mathbf{x}(r)\} \times \{\mathbf{F}^{-1} \cdot d\mathbf{x}(s)\}. \tag{11.40}$$

The vector d$\mathbf{A}$ is then dotted with the inverse deformation gradient $\mathbf{F}^{-1}$ from the left, thus

$$d\mathbf{A} \cdot \mathbf{F}^{-1} = \{\mathbf{F}^{-1} \cdot d\mathbf{x}(r)\} \times \{\mathbf{F}^{-1} \cdot d\mathbf{x}(s)\} \cdot \mathbf{F}^{-1}. \tag{11.41}$$

The right-hand side of this equation may be expanded, as a similar one for d$\mathbf{a}$ was above, and, using the fact that a determinant of a product of matrices is the product of the determinants (Section A.8, page 372), rewritten as

$$d\mathbf{A} \cdot \mathbf{F}^{-1} = J^{-1} d\mathbf{a}. \tag{11.42}$$

These formulas relating d$\mathbf{A}$ and d$\mathbf{a}$ are called the *formulas of Nanson*. These formulas will be employed in relating the various definitions of stress associated with large deformations.

Example 11.5.1
Consider the plane area that forms the right-hand face of the unit cube illustrated in Fig. 11.9. Use the formulas of Nanson to determine the magnitude and orientation of the deformed area as a result of the deformation specified in Example Problem 11.3.1.

Solution: The undeformed area is represented by $d\mathbf{A}^{\mathrm{T}} = (1, 0, 0)$. The value of J is $2\sqrt{3}$ and the tensor of inverse deformation gradients is given by

$$\mathbf{F}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

From (11.42) it follows that $d\mathbf{a} = J d\mathbf{A}\cdot\mathbf{F}^{-1}$, thus

$$d\mathbf{a} = 2\sqrt{3}\left\{ [1 \quad 0 \quad 0] \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}.$$

The deformed representation of the plane area that forms the right-hand face of the unit cube illustrated in Fig. 11.7 is shown as the right sloping right-hand face in Fig. 11.8. From that figure it is seen that the face associated with this area is two units high and one unit wide, and that the unit normal to this face is indeed $(2/\sqrt{3}, -1/\sqrt{3}, 0)$.

Problems

11.5.1. Compute the volume of the parallelepiped whose three edges meeting at one vertex are characterized by the vectors (1, 2, 3), (2, 3, 3), (−3, −2, −1).

11.5.2. Consider the plane area that forms the top face of the deformed parallelogram illustrated in Fig. 11.8. Use the formulas of Nanson to determine the magnitude and orientation of the original area if the deformed area was a consequence of the deformation specified in Example Problem 11.3.1.

11.5.3. Beginning with the formula (11.37) for d**a** use the indicial notation and the formulas for the Det**F** employing the alternator from the Appendix to derive the formula

$$(\mathbf{F}^{-1})^T \cdot d\mathbf{A} = J^{-1} d\mathbf{a}. \text{ (11.42 alternate)}$$

11.6 Stress Measures

The stress equations of motion are the most useful form of Newton's second law in continuum mechanics. In Sect. 3.4 these field equations were shown to have the form

$$\rho\ddot{\mathbf{x}} = \nabla \cdot \mathbf{T} + \rho\mathbf{d}, \ \mathbf{T} = \mathbf{T}^T. \quad (3.37) \text{ repeated}$$

These equations involve the most common stress measure, the Cauchy stress **T** introduced in Chapter 2. It was not called the Cauchy stress in Chapter 4 because there was only one stress tensor under discussion there. The Cauchy stress is referred to the instantaneous or spatial coordinate system; it is measured relative to the instantaneous area. In mechanical testing the phrase "true stress" is used to denote a stress calculated using the instantaneous rather than the original cross-sectional area. Cauchy stress, which is sometimes called Eulerian stress, is therefore a "true stress."

There are a number of different stress measures used in the study of finite deformations of materials; three are considered here. The first is the Cauchy stress. The second stress measure is the first Piola–Kirchhoff stress tensor that is sometimes called the Lagrangian stress tensor. This stress tensor is referred to the reference configuration. Consider an object in both its deformed and undeformed configurations. Since this is the same object in the two configurations, it must have the same total mass M in each configuration, and Eq. (3.1) maybe rewritten to express that fact:

$$M = \int_O \rho_R(\mathbf{X})\, dV = \int_O \rho(\mathbf{x}, t)\, dv, \quad (11.43)$$

where $\rho_R(\mathbf{X})$ is the density in the initial configuration and $\rho(\mathbf{x},t)$ is the density in the instantaneous or deformed configuration; dV is an element of volume in the initial

configuration and dv is an element of volume in the final configuration. Substituting the relationship (11.36) between dv and dV, d$v = J$dV, into (11.43) it follows that

$$M = \int_O \rho_R(\mathbf{X})\, dV = \int_O \rho(\mathbf{x}, t) J\, dV. \tag{11.44}$$

Since this same result must hold for each and every part of the object we conclude that the relationship between the two density fields is given by the simple formula

$$\rho_R = J\rho, \tag{11.45}$$

which is an alternative statement of mass balance. The argument that is used to go from (11.44) to (11.45) is the same argument that was used to go from (3.4) to (3.5), and it was employed three more times in Chapter 3. In order to relate the stresses referred to the two different configurations, a similar procedure to the arguments leading to the result $\rho_R = J\rho$ above is followed. The total force $\mathbf{f}$ acting on an object, or on any particular subpart of the object, is considered. The total force is the same in both configurations; and therefore the product of the stress and a differential area element integrated over the object must be the same in both configurations. Thus

$$\mathbf{f} = \int_{\partial O} \mathbf{T}^{1PK} \cdot d\mathbf{A} = \int_{\partial O} \mathbf{T} \cdot d\mathbf{a}, \tag{11.46}$$

where $\mathbf{T}^{1\mathrm{PK}}$ and d$\mathbf{A}$ are the stress tensor and differential area element in the initial configuration and $\mathbf{T}$ and d$\mathbf{a}$ are the stress and differential area in the instantaneous or deformed configuration. $\mathbf{T}$ is the Cauchy stress, of course. When the relationship of Nanson between d$\mathbf{a}$ and d$\mathbf{A}$ given by (11.42) is substituted into (11.46) we find that

$$\mathbf{f} = \int_{\partial O} \mathbf{T}^{1PK} \cdot d\mathbf{A} = \int_{\partial O} J\mathbf{T} \cdot \left(\mathbf{F}^{-1}\right)^T \cdot d\mathbf{A}. \tag{11.47}$$

Since this must hold for all parts of the object, the same argument as in the transitions from (11.44) to (11.45) and (3.4) to (3.5), it may be concluded that

$$\mathbf{T}^{1PK} = J\mathbf{T} \cdot \left(\mathbf{F}^{-1}\right)^T, \text{~or } \mathbf{T} = J^{-1}\mathbf{T}^{1PK} \cdot \mathbf{F}^T. \tag{11.48}$$

$\mathbf{T}^{1\mathrm{PK}}$ is called the *first Piola–Kirchhoff or* Lagrangian stress tensor. The relation of Cauchy involving the Cauchy stress tensor and the spatial reference frame, namely that the stress vector $\mathbf{t}_{(\mathbf{n})}$ acting on any plane whose normal is $\mathbf{n}$ is given

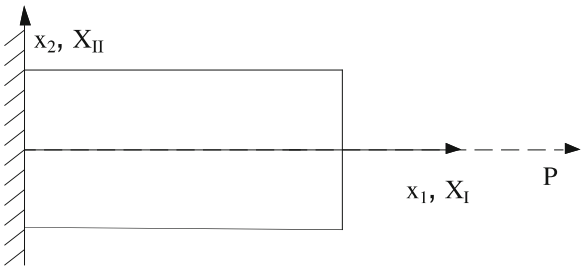


Fig. 11.15 Illustration for a uniaxial stretching deformation due to a load P

by $\mathbf{t}_{(\mathbf{n})} = \mathbf{T}\cdot\mathbf{n}$, may be extended to the first Piola–Kirchhoff stress tensor $\mathbf{T}^{1\mathrm{PK}}$ and the material coordinate system; thus $\mathbf{t}_{(\mathbf{N})} = \mathbf{T}^{1\mathrm{PK}}\cdot\mathbf{N}$, where $\mathbf{N}$ is the unit normal vector to the plane in the material coordinate system. Using this result, the stress equations of motion may be rederived in the material coordinates and in terms of the first Piola–Kirchhoff stress tensor; thus

$$\rho_R\ddot{\mathbf{x}} = \nabla_{\mathbf{o}}\cdot\mathbf{T}^{1PK} + \rho_R\mathbf{d},\ \mathbf{T}^{1PK}\cdot\mathbf{F}^T = \mathbf{F}\cdot\left(\mathbf{T}^{1PK}\right)^T; \tag{11.49}$$

where the density of the initial configuration ρ_R is used as the reference density and the divergence is now with respect to material rather than spatial coordinates. Recall from Chapter 2 (see (2.23)) that the gradient symbol $\nabla_{\mathbf{O}}$ with a subscripted O will indicate a gradient with respect to the material coordinate system $\mathbf{X}$, rather than the usual gradient symbol ∇ used to indicate a gradient with respect to the spatial coordinate system $\mathbf{x}$. The second of (11.49) shows that the first Piola–Kirchhoff stress tensor is not symmetric like the Cauchy stress tensor. To see that this is the case the second of (11.48) may be substituted twice in the second of (3.37) to verify.

In order to have a measure of stress referred to the initial configuration that is symmetric, the *second Piola–Kirchhoff* stress tensor is introduced; this is also called the Kirchhoff stress tensor. This new stress tensor is denoted by $\mathbf{T}^{2\mathrm{PK}}$ and defined as follows:

$$\mathbf{T}^{2PK} \equiv \mathbf{F}^{-1}\cdot\mathbf{T}^{1PK} = J\mathbf{F}^{-1}\cdot\mathbf{T}\cdot\left(\mathbf{F}^{-1}\right)^T. \tag{11.50}$$

Substituting $\mathbf{T}^{1\mathrm{PK}} = \mathbf{F}\cdot\mathbf{T}^{2\mathrm{PK}}$ into (11.49) the equations of motion in terms of the second Piola–Kirchhoff stress tensor are obtained:

$$\rho_R\ddot{\mathbf{x}} = \nabla_{\mathbf{o}}\cdot\left(\mathbf{F}\cdot\mathbf{T}^{2PK}\right) + \rho_R\mathbf{d},\ \mathbf{T}^{2PK} = \left(\mathbf{T}^{2PK}\right)^T. \tag{11.51}$$

This shows that the second Piola–Kirchhoff stress tensor $\mathbf{T}^{2\mathrm{PK}}$ is symmetric.

Example 11.6.1

A solid specimen capable of large deformations is extended by a force of magnitude P in the x_1 or X_I direction (these directions are coincident here). This uniaxial stress situation is illustrated in Fig. 11.15. Determine the Cauchy and the first and second

Piola–Kirchhoff stress tensors in this uniaxial situation. Construct the relationships between each of these tensors in this particular situation.

Solution: The principal axes of extension are obviously the direction of the applied load and the two perpendicular transverse directions. The deformation can be represented in terms of the principal stretches λ_I, λ_{II} and λ_{III} by

$$x_1 = \lambda_I X_I,\ x_2 = \lambda_{II} X_{II},\ x_3 = \lambda_{III} X_{III},$$

and the deformation gradients **F** by (11.32). It follows that the Jacobian J is given by $J = \lambda_I \lambda_{II} \lambda_{III}$. The area perpendicular to the x_1- or X_I-axis will change with the deformation. Using the formula of Nanson (11.42), it follows that

$$d\mathbf{a} = J(\mathbf{F}^{-1})^T \cdot d\mathbf{A} = \lambda_I \lambda_{II} \lambda_{III} \begin{bmatrix} \lambda_I^{-1} & 0 & 0 \\ 0 & \lambda_{II}^{-1} & 0 \\ 0 & 0 & \lambda_{III}^{-1} \end{bmatrix} \begin{bmatrix} dA_I \\ 0 \\ 0 \end{bmatrix},$$

or $\mathrm{d}a_1 = \lambda_{II}\lambda_{III}\mathrm{d}A_I$, $\mathrm{d}a_2 = 0$, and $\mathrm{d}a_3 = 0$; thus we can conclude that the relationship between the instantaneous area A and initial area A_o is $A = \lambda_{II}\lambda_{III}A_\mathrm{o}$. The only nonzero Cauchy stress is $T_{11} = P/A = P/(\lambda_{II}\lambda_{III}A_\mathrm{o})$. From (11.48) the only nonzero component of the first Piola–Kirchhoff stress tensor is given by $T_{1I}^{1PK} = \lambda_{II}\lambda_{III}T_{11} = \lambda_{II}\lambda_{III}P/A = P/A_o$ and, from (11.50) the only nonzero component of the second Piola–Kirchhoff stress tensor is given by

$$T_{1I}^{2PK} = (T_{1I}^{1PK})/\lambda_I = (\lambda_{II}\lambda_{III}T_{11})/\lambda_I = \lambda_{II}\lambda_{III}P/(\lambda_I A) = P/(\lambda_I A_o).$$

In the special case when the material is incompressible, $J = \lambda_I\lambda_{II}\lambda_{III} = 1$, and the cross-section transverse to the extension is symmetric, $\lambda_I = \lambda$, $\lambda_{II} = \lambda_{III} = 1/\sqrt{\lambda}$, then $T_{1I}^{2PK} = (T_{1I}^{1PK})/\lambda = T_{11}/\lambda^2 = P/(\lambda^2 A) = P/(\lambda A_o)$.

Problems

11.6.1. A rectangular parallelepiped with a long dimension a_o and a square cross-section of dimension b_o is deformed by an axial tensile force P into a rectangular parallelepiped with a longer long dimension a and a smaller square cross-section of dimension b. This problem is a continuation of Problem 11.4.4.

(a) Record the expressions for the stresses T_{1I}^{2PK} and T_{11} in the deformed rectangular parallelepiped. Both the x_1 and X_I directions are coincident with the longitudinal axis of the parallelepiped.

(b) If the material of the rectangular parallelepiped is incompressible, what is the relationship between b and b_o?

(c) Record the expression for the stress T_{1I}^{2PK} in terms of the stress T_{11} if the rectangular parallelepiped is incompressible.

11.6.2. Consider a unit cube subjected to a uniaxial extension. A net force P in the x_1 or X_I direction creates this uniaxial extension. The motion is described by $x_1 = (1 + t)X_I, x_2 = X_{II}, x_3 = X_{III}$. Note that each face of the unit cube has an initial area A_o that is unity, $A_o = 1$. Estimate the first Piola–Kirchhoff stress $\mathbf{T}^{1PK}$, then calculate the Cauchy stress $\mathbf{T} = J^{-1}\ \mathbf{T}^{1PK}\cdot\mathbf{F}^T$ and the second Piola–Kirchhoff stress ($\mathbf{T}^{2PK} \equiv \mathbf{F}^{-1}\cdot\mathbf{T}^{1PK} = J\mathbf{F}^{-1}\cdot\mathbf{T}\cdot(\mathbf{F}^{-1})^T$).

11.6.3. Consider a unit cube subjected to a biaxial extension. A net force P_1 in the x_1 or X_I direction and a net force P_2 in the x_2 or X_{II} direction create this biaxial extension. The motion is described by $x_1 = (1 + t)X_I, x_2 = (1 + 2t)X_{II}, x_3 = X_{III}$. Note that each face of the unit cube has an initial area A_o that is unity, $A_o = 1$. Estimate the first Piola–Kirchhoff stress $\mathbf{T}^{1PK}$, then calculate the Cauchy stress $\mathbf{T} = J^{-1}\ \mathbf{T}^{1PK}\cdot\mathbf{F}^T$ and the second Piola–Kirchhoff stress ($\mathbf{T}^{2PK} \equiv \mathbf{F}^{1}\cdot\mathbf{T}^{1PK} = J\mathbf{F}^{-1}\cdot\mathbf{T}\cdot(\mathbf{F}^{-1})^T$).

11.6.4. Consider an object that is the unit cube deformed by deformation gradient tensor $\mathbf{F}$ given in Problem 11.3.1. If the homogeneous *second Piola–Kirchhoff* stress tensor $\mathbf{T}^{2PK}$ is given by

$$T^{2PK} = \begin{bmatrix} 10 & 3 & 0 \\ 3 & 20 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

determine the Cauchy stress tensor $\mathbf{T}$ and the *first Piola–Kirchhoff* stress tensor $\mathbf{T}^{1PK}$.

11.6.5. Consider an object that is the unit cube deformed by deformation gradient tensor $\mathbf{F}$ given in Problem 11.3.1 on page 511. If the homogeneous *second Piola–Kirchhoff* stress tensor $\mathbf{T}^{2PK}$ is given by

$$T^{2PK} = \begin{bmatrix} 10 & 3 & 0 \\ 3 & 20 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

determine the stress vector acting on the sloping face whose normal is $(2/\sqrt{5}, -1/\sqrt{5}, 0)$ in the deformed configuration.

11.7 Finite Deformation Elasticity

An elastic material is a material characterized by a constitutive equation, which specifies that stress is a function of strain only. It is also possible to represent an elastic material by a constitutive equation that specifies stress as a function of the deformation gradients $\mathbf{F}$, provided one keeps in mind that, due to invariance under rigid object rotations, the stress must be independent of the part of $\mathbf{F}$ that represents rotational motion. In terms of the Cauchy stress $\mathbf{T}$ and the deformation gradient $\mathbf{F}$ the constitutive equation for an elastic material can be written $\mathbf{T} = \mathbf{g}(\mathbf{F})$. Invariance under rigid object rotations requires

$$\mathbf{g}(\mathbf{Q}\cdot\mathbf{F}) = \mathbf{Q}\cdot\mathbf{g}(\mathbf{F})\cdot\mathbf{Q}^{T} \tag{11.52}$$

for all orthogonal tensors $\mathbf{Q}$. If we take $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$, where $\mathbf{R}$ is the rigid object rotation and $\mathbf{U}$ the right stretch tensor which are related to $\mathbf{F}$ by $\mathbf{F} = \mathbf{R}\cdot\mathbf{U}$ (Eq. (11.8)), it follows from (11.52) that $\mathbf{g}(\mathbf{R}^{\mathrm{T}}\cdot\mathbf{F}) = \mathbf{g}(\mathbf{U}) = \mathbf{R}^{\mathrm{T}}\cdot\mathbf{g}(\mathbf{F})\cdot\mathbf{R}$; thus $\mathbf{T} = \mathbf{g}(\mathbf{F}) = \mathbf{R}\cdot\mathbf{g}(\mathbf{U})\cdot\mathbf{R}^{\mathrm{T}}$, or

$$\mathbf{T} = \mathbf{R}\cdot\mathbf{f}(\mathbf{C})\cdot\mathbf{R}^{T,} \tag{11.53}$$

where $\mathbf{g}(\mathbf{U}) = \mathbf{f}(\mathbf{C}) = \mathbf{f}(\mathbf{U}^2)$ (since $\mathbf{C} = \mathbf{U}^2$ (Eq. (11.27))). In terms of the first Piola–Kirchhoff stress tensor $\mathbf{T}^{1\mathrm{PK}}$ the constitutive equation for an elastic material is

$$\mathbf{T}^{1PK} = \mathbf{h}(\mathbf{F}), \tag{11.54}$$

or, due to the invariance of constitutive equations under rigid object rotations, $\mathbf{h}(\mathbf{Q}\cdot\mathbf{F}) = \mathbf{Q}\cdot\mathbf{h}(\mathbf{F})$, and taking $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$, $\mathbf{h}(\mathbf{U}) = \mathbf{R}^{\mathrm{T}}\cdot\mathbf{h}(\mathbf{F})$; thus $\mathbf{h}(\mathbf{F}) = \mathbf{R}\cdot\mathbf{h}(\mathbf{U})$, and

$$\mathbf{T}^{1PK} = \mathbf{R}\cdot\mathbf{h}(\mathbf{U}). \tag{11.55}$$

In terms of the second Piola–Kirchhoff stress tensor $\mathbf{T}^{2\mathrm{PK}}$, the constitutive equation for an elastic material has the form

$$\mathbf{T}^{2PK} = \mathbf{t}(\mathbf{C}). \tag{11.56}$$

These constitutive equations are said to describe a material with Cauchy elasticity; that is to say a material in which stress is a function of some measure of the strain or deformation.

11.8 The Isotropic Finite Deformation Stress–Strain Relation

The assumption of isotropic symmetry of a material is an adequate model for many materials. Recall from Chapter 5 that, in the case of a stress strain relation, isotropy means that the response of stress to an applied strain is the same in any direction in the material. The mathematical statement of this notion is that the stress tensor, say $\mathbf{T}^{2\mathrm{PK}}$ in (11.56), is an isotropic function of the right Cauchy–Green tensor $\mathbf{C}$. In order for the tensor valued function $\mathbf{T}^{2\mathrm{PK}} = \mathbf{t}(\mathbf{C})$, given by (11.56) to be isotropic an function, it must satisfy the relation for all orthogonal tensors $\mathbf{Q}$:

$$\mathbf{Q}\cdot\mathbf{T}^{2PK}\cdot\mathbf{Q}^{T} = \mathbf{t}(\mathbf{Q}\cdot\mathbf{C}\cdot\mathbf{Q}^{T}). \tag{11.57}$$

As one can see from the transformation law for a second order tensor (A83) that the definition of an isotropic function (11.57) requires that, when the value of the

argument $\mathbf{C}$ is transformed, the value of the function $\mathbf{T}^{2\mathrm{PK}}$ is transformed in the same manner. The development objective of this section is to use the isotropy requirement (11.57) to restrict the functional form of the relationship (11.56). The first step in this development is to show that the principal axes of $\mathbf{T}^{2\mathrm{PK}}$ must coincide with the principal axes of the right Cauchy–Green tensor $\mathbf{C}$ if the isotropy requirement (11.57) is satisfied.

To show that this is the case let $\mathbf{c}$ denotes an eigenvector of $\mathbf{C}$ corresponding to the eigenvalue, say $\lambda_I^2 C$, thus $\mathbf{C}\cdot\mathbf{c} = \lambda_I^2 c$. This eigenvector $\mathbf{c}$ is used to construct a reflective symmetry transformation $\mathbf{R}^{(\mathbf{c})}$with the properties specified by (5.1) for the vector $\mathbf{a}$. Replacing $\mathbf{a}$ in (4.2) by $\mathbf{c}$, it follows that

$$\mathbf{R}^{(\mathbf{c})} = \mathbf{1} - 2\mathbf{a}\otimes\mathbf{a}, \tilde{}(R_{ij}^{(\mathbf{a})} = \delta_{ij} - 2c_i c_j), \tag{11.58}$$

then, setting $\mathbf{Q} = \mathbf{R}^{(\mathbf{c})}$in (11.57), it follows that

$$\mathbf{R}^{(\mathbf{c})}\cdot\mathbf{T}^{2PK}\cdot\mathbf{R}^{(\mathbf{c})^T} = \mathbf{t}(\mathbf{R}^{(\mathbf{c})}\cdot\mathbf{C}\cdot\mathbf{R}^{(\mathbf{c})^T}). \tag{11.59}$$

For a symmetric second order tensor $\mathbf{A}$ is easy to show, using (11.58), that

$$\mathbf{R}^{(\mathbf{c})}\cdot\mathbf{A}\cdot\mathbf{R}^{(\mathbf{c})} = \mathbf{A} - 2\mathbf{c}\otimes\mathbf{A}\cdot\mathbf{c} - 2\mathbf{A}\cdot\mathbf{c}\otimes\mathbf{c} + 4(\mathbf{c}\cdot\mathbf{A}\cdot\mathbf{c})\mathbf{c}\otimes\mathbf{c} \tag{11.60}$$

and, in particular, when $\mathbf{A} = \mathbf{C}$, it follows from the fact that $\mathbf{c}$ is an eigenvector of $\mathbf{C}$ that

$$\mathbf{R}^{(\mathbf{c})}\cdot\mathbf{C}\cdot\mathbf{R}^{(\mathbf{c})^T} = \mathbf{C}. \tag{11.61}$$

The substitution of (11.61) into (11.59) reduces (11.59) to

$$\mathbf{R}^{(\mathbf{c})}\cdot\mathbf{T}^{2PK}\cdot\mathbf{R}^{(\mathbf{c})^T} = \mathbf{t}(\mathbf{C}). \tag{11.62}$$

Next let $\mathbf{A} = \mathbf{t}(\mathbf{C})$ in the identity (11.60), then it follows with a little algebra, and recalling that $\mathbf{R}^{(\mathbf{c})}\cdot\mathbf{c} = -\mathbf{c}$, that

$$\mathbf{t}(\mathbf{C})\cdot\mathbf{c} = (\mathbf{c}\cdot\mathbf{t}(\mathbf{C})\cdot\mathbf{c})\mathbf{c}. \tag{11.63}$$

Since $\mathbf{c}\cdot\mathbf{t}(\mathbf{C})\cdot\mathbf{c}$ is a scalar, this result shows that $\mathbf{c}$ is an eigenvector of $\mathbf{t}(\mathbf{C})$ as well as $\mathbf{C}$. It then follows that any eigenvector of $\mathbf{C}$ is also an eigenvector of $\mathbf{T}^{2\mathrm{PK}} = \mathbf{t}(\mathbf{C})$. Since $\mathbf{C}$ and $\mathbf{T}^{2\mathrm{PK}} = \mathbf{t}(\mathbf{C})$ have the same set of principal axes, then the eigenvalues $(T_{11}^{2PK}, T_{22}^{2PK}, T_{33}^{2PK})$of $\mathbf{T}^{2\mathrm{PK}}$ are functions of the eigenvalues $(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2)$of $\mathbf{C}$ given by (11.33), thus

$$T_{11}^{2PK} = t_{11}(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2),\ T_{22}^{2PK} = t_{22}(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2),\ T_{33}^{2PK} = t_{33}(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2). \tag{11.64}$$

In order to express more specific function forms of (11.64), recall that each eigenvalue $(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2)$of a matrix like **C** satisfies the same characteristic equation (A54), thus

$$\lambda^6 - I_C\lambda^4 + II_C\lambda^2 - III_C = 0, \tag{11.65}$$

where

$$I_C = tr\mathbf{C},\ II_C = \frac{1}{2}[(tr\mathbf{C})^2 - tr\mathbf{C}^2],\ III_C = Det\mathbf{C}. \tag{11.66}$$

If the expression for $T_{11}^{2PK} = t_{11}(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2)$were expanded in a power series in λ_I^2, (11.65) above could be used to eliminate any term not proportional to $1, \lambda_I^2$or λ_I^4. With this motivation the eigenvalues $(T_{11}^{2PK}, T_{22}^{2PK}, T_{33}^{2PK})$of $\mathbf{T}^{2\mathrm{PK}}$ are expressed as functions of the eigenvalues $(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2)$of **C** as follows:

$$\begin{aligned} T_{11}^{2PK} &= a_o + a_1\lambda_I^2 + a_2\lambda_I^4,\ T_{22}^{2PK} = a_o + a_1\lambda_{II}^2 + a_2\lambda_{III}^4, \\ T_{33}^{2PK} &= a_o + a_1\lambda_{III}^2 + a_2\lambda_{III}^4. \end{aligned} \tag{11.67}$$

This system of equations has a unique solution for the three unknown functions a_o, a_1, and a_2. These functions are elementary symmetric functions of the three eigenvalues $(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2)$or the three (isotropic) invariants of **C**, thus

$$\begin{aligned} a_o &= a_o(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2) = a_o(I_C, II_C, III_C),\ a_1 = a_1(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2) \\ &= a_1(I_C, II_C, III_C),\ a_2 = a_2(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2) = a_2(I_C, II_C, III_C). \end{aligned} \tag{11.68}$$

In the principal coordinate system it then follows that

$$\mathbf{T}^{2PK} = a_o\mathbf{1} + a_1\mathbf{C} + a_2\mathbf{C}^2, \tag{11.69}$$

an expression that is equivalent to (11.67) in the principal coordinate system of **C** (or **t**(**C**)), but that also holds in any arbitrary coordinate system. A necessary and sufficient condition that the constitutive relation (11.56), $\mathbf{T}^{2\mathrm{PK}} = \mathbf{t}(\mathbf{C})$ satisfy the material isotropy requirement (11.57), $\mathbf{Q}\cdot\mathbf{T}^{2\mathrm{PK}}\cdot\mathbf{Q}^{\mathrm{T}} = \mathbf{t}(\mathbf{Q}\cdot\mathbf{C}\cdot\mathbf{Q}^{\mathrm{T}})$, is that $\mathbf{T}^{2\mathrm{PK}} = \mathbf{t}(\mathbf{C})$ have a representation of the form (11.69) with a_o, a_1, and a_2 given by (11.68).

The representation (11.69) of $\mathbf{T}^{2\mathrm{PK}}$ as isotropic function of **C** may also be expressed as an equivalent isotropic relationship between the Cauchy stress **T** and the left Cauchy–Green tensor **B**. The algebraic manipulations that achieve this equivalence begin with recalling from (11.27) that $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\cdot\mathbf{F}$, thus (11.69) may be written in the form

$$\mathbf{T}^{2PK} = a_o\mathbf{1} + a_1\mathbf{F}^T \cdot \mathbf{F} + a_2\mathbf{F}^T \cdot \mathbf{F} \cdot \mathbf{F}^T \cdot \mathbf{F}, \tag{11.70}$$

then premultiplying by $(\mathbf{F}^T)^{-1}$ and post multiplying by $\mathbf{F}^{-1}$, thus

$$(\mathbf{F}^T)^{-1} \cdot \mathbf{T}^{2PK} \cdot \mathbf{F}^{-1} = a_o(\mathbf{F}^T)^{-1} \cdot \mathbf{F}^{-1} + a_1(\mathbf{F}^T)^{-1} \cdot \mathbf{F}^T \cdot \mathbf{F} \cdot \mathbf{F}^{-1} +,$$
$$a_2(\mathbf{F}^T)^{-1} \cdot \mathbf{F}^T \cdot \mathbf{F} \cdot \mathbf{F}^T \cdot \mathbf{F} \cdot \mathbf{F}^{-1}. \tag{11.71}$$

This expression reduces to

$$\mathbf{T} = \frac{a_1}{J}\mathbf{1} + \frac{a_2}{J}\mathbf{B} + \frac{a_o}{J}\mathbf{B}^{-1}, \tag{11.72}$$

when one takes note of (11.27), (11.28), and the second of (11.51). This expression is rewritten in the form

$$\mathbf{T} = h_o\mathbf{1} + h_1\mathbf{B} + h_{-1}\mathbf{B}^{-1}, \tag{11.73}$$

when it is observed that $I_C = I_B$, $II_C = II_B$, and $III_C = III_B = J$ and the following notation is introduced

$$h_o = h_o(I_C, II_C, III_C) = h_o(I_B, II_B, III_B) = \frac{a_1}{J}(I_C, II_C, III_C),$$
$$h_1 = h_1(I_C, II_C, III_C) = h_1(I_B, II_B, III_B) = \frac{a_o}{J}(I_C, II_C, III_C),$$
$$h_{-1} = h_{-1}(I_C, II_C, III_C) = h_{-1}(I_B, II_B, III_B) = \frac{a_2}{J}(I_C, II_C, III_C). \tag{11.74}$$

Problems

11.8.1. Setting $\mathbf{A} = \mathbf{t}(\mathbf{C})$ in the identity (11.60), verify the formula (11.63),$\mathbf{t}(\mathbf{C}) \cdot \mathbf{c} = (\mathbf{c} \cdot \mathbf{t}(\mathbf{C}) \cdot \mathbf{c})\mathbf{c}$.

11.8.2. Verify the result (11.60), then derive the result (11.61) from (11.60).

11.8.3. Determine the Cauchy stress tensor $\mathbf{T}$ and the *second Piola–Kirchhoff* stress tensor $\mathbf{T}^{2PK}$ for an elastic isotropic material subjected to the deformation in Example 11.2.1. Specify the numerical value of the functional dependencies of the functions a_o, a_1, and a_2 as well as those of h_o, h_1, and h_{-1}.

11.9 Finite Deformation Hyperelasticity

A *hyperelastic material* is an elastic material for which the stress is derivable from a scalar potential called a strain energy function. Thus a hyperelastic material is automatically a Cauchy elastic material, but not the reverse. In the case of small deformation elasticity, a strain energy function always exists and therefore the

small deformation theory is hyperelastic. The strain energy per unit volume W is obtained from the specific strain energy, that is to say the strain energy per unit mass by multiplying it by ρ_R, where ρ_R is the density function in the initial configuration. In terms of the Cauchy stress and the first and second Piola–Kirchhoff stress tensors the definition of a hyperelastic material has the following forms:

$$\mathbf{T} = \frac{\rho}{\rho_R}\mathbf{F}\cdot\left(\frac{\partial W}{\partial \mathbf{F}}\right)^T,\ \mathbf{T}^{1PK} = \frac{\partial W}{\partial \mathbf{F}},\ \mathbf{T}^{2PK} = \mathbf{F}^{-1}\cdot\left(\frac{\partial W}{\partial \mathbf{F}}\right). \tag{11.75}$$

On first encounter, the variety of forms for the constitutive relation for hyperelastic materials is bewildering. Not only are there three different stress measures, but there are many different strain measures, **C**, **c**, **F**, **E**, **e**, etc. Thus, for example, if we introduce the right Cauchy–Green deformation tensor **C**, since $\mathbf{C} = \mathbf{U}^2 = \mathbf{F}\mathbf{F}^{\mathrm{T}}$, $\frac{\partial \mathbf{C}}{\partial \mathbf{F}} = 2\mathbf{F}$ then

$$\frac{\partial W}{\partial \mathbf{F}} = 2\mathbf{F}\cdot\frac{\partial W}{\partial \mathbf{C}}, \tag{11.76}$$

and the constitutive relations (11.57) take the form

$$\mathbf{T} = 2\frac{\rho}{\rho_R}\mathbf{F}\cdot\left(\frac{\partial W}{\partial \mathbf{C}}\right)^T\cdot\mathbf{F}^T,\ \mathbf{T}^{1PK} = 2\mathbf{F}\cdot\frac{\partial W}{\partial \mathbf{C}},\ \mathbf{T}^{2PK} = 2\frac{\partial W}{\partial \mathbf{C}}. \tag{11.77}$$

Alternatively, these relations can be expressed in terms of the Lagrangian strain tensor **E**, $2\mathbf{E} = \mathbf{C} - \mathbf{1}$, by (11.29), thus

$$\frac{\partial W}{\partial \mathbf{C}} = \frac{\partial W}{\partial \mathbf{E}} : \frac{\partial \mathbf{E}}{\partial \mathbf{C}} = \frac{1}{2}\frac{\partial W}{\partial \mathbf{E}}, \tag{11.78}$$

and

$$\mathbf{T} = \frac{\rho}{\rho_R}\mathbf{F}\cdot\left(\frac{\partial W}{\partial \mathbf{E}}\right)^T\cdot\mathbf{F}^T,\ \mathbf{T}^{1PK} = \mathbf{F}\cdot\frac{\partial W}{\partial \mathbf{E}},\ \mathbf{T}^{2PK} = \frac{\partial W}{\partial \mathbf{E}}, \tag{11.79}$$

for example.

In the special case of an isotropic hyperelastic material the strain energy function $W = W(\mathbf{C}) = W(\lambda_I^2, \lambda_{II}^2, \lambda_{III}^2)$ depends upon **C** only through the (isotropic) invariants $(I_C, II_C, III_C)$of **C**, thus

$$W = W(I_C, II_C, III_C), \tag{11.80}$$

where $(I_C, II_C, III_C)$are given by (11.66). Substituting this isotropic expression for the strain energy into (11.72) and making use of the following expressions for the derivatives of the invariants I_C, II_C and $III_C$with respect to **C**,

$$\frac{\partial I_C}{\partial \mathbf{C}} = \mathbf{1},\ \frac{\partial II_C}{\partial \mathbf{C}} = I_C\mathbf{1} - \mathbf{C},\ \frac{\partial III_C}{\partial \mathbf{C}} = III_C\mathbf{C}^{-1}. \tag{11.81}$$

it follows that $\mathbf{T}^{2PK}$ has the representation

$$\mathbf{T}^{2PK} = 2\left[\left\{\frac{\partial W}{\partial I_C} + I_C\frac{\partial W}{\partial II_C}\right\}\mathbf{1} - \frac{\partial W}{\partial II_C}\mathbf{C} + \frac{\partial W}{\partial III_C}III_C\mathbf{C}^{-1}\right]. \tag{11.82}$$

This constitutive relation may also be written in a form which contains $\mathbf{C}^2$ rather than $\mathbf{C}^{-1}$,

$$\mathbf{T}^{2PK} = 2\left[\left\{\frac{\partial W}{\partial I_\mathbf{C}} + I_\mathbf{C}\frac{\partial W}{\partial II_\mathbf{C}} + \frac{\partial W}{\partial III_\mathbf{C}}II_\mathbf{C}\right\}\mathbf{1} - \left\{\frac{\partial W}{\partial II_\mathbf{C}} + I_\mathbf{C}\frac{\partial W}{\partial III_\mathbf{C}}\right\}\mathbf{C} + \frac{\partial W}{\partial III_\mathbf{C}}\mathbf{C}^2\right] \tag{11.83}$$

by use of Cayley Hamilton theorem that states that a matrix satisfies its own characteristic equation, thus $\mathbf{C}$ may replace λ^2 in (11.65), $\mathbf{C}^3 - I_C\mathbf{C}^2 + II_C\mathbf{C} - II\ I_C = 0$. A term-by-term comparison of (11.78) with (11.69) shows that the coefficients a_o, a_1, and a_2 in (11.69) are given in the case of hyperelasticity by

$$a_o = 2\left\{\frac{\partial W}{\partial I_C} + I_\mathbf{C}\frac{\partial W}{\partial II_\mathbf{C}} + \frac{\partial W}{\partial III_\mathbf{C}}II_\mathbf{C}\right\},\ a_1 = -2\left\{\frac{\partial W}{\partial II_\mathbf{C}} + I_\mathbf{C}\frac{\partial W}{\partial III_\mathbf{C}}\right\},$$
$$a_2 = 2\frac{\partial W}{\partial III_\mathbf{C}} \tag{11.84}$$

Also, in the case of an isotropic hyperelastic material, the coefficients h_o, h_1, and h_2 in the constitutive relation between the Cauchy stress $\mathbf{T}$ and the left Cauchy–Green tensor $\mathbf{B}$, (11.73), may be expressed in terms of the strain energy function by the following formulas:

$$h_o = \frac{2}{J}\left\{II_\mathbf{C}\frac{\partial W}{\partial II_\mathbf{C}} + \frac{\partial W}{\partial III_\mathbf{C}}III_\mathbf{C}\right\},\ h_1 = \frac{2}{J}\frac{\partial W}{\partial I_\mathbf{C}},\ h_{-1} = -\frac{2}{J}III_\mathbf{C}\frac{\partial W}{\partial II_\mathbf{C}} \tag{11.85}$$

Problems

11.9.1. Derive the result $(\partial W/\partial \mathbf{C}) = (\partial W/\partial \mathbf{E}) : (\partial \mathbf{E}/\partial \mathbf{C}) = (1/2)(\partial W/\partial \mathbf{E})$, (11.78), using the indicial notation and the formula $2\mathbf{E} = \mathbf{C} - \mathbf{1}$. Hint: It is useful to first derive the formula $(\partial E_{\gamma\delta}/\partial C_{\alpha\beta}) = (1/2)\delta_{\beta\delta}\delta_{\alpha\gamma}$ from $2E_{\gamma\delta} = C_{\gamma\delta} - \delta_{\gamma\delta}$.

11.9.2. Derive the result (11.76) using the indicial notation. Hint: It is useful to first derive the formula $(\partial C_{\alpha\beta}/\partial F_{i\gamma}) = (\partial/\partial F_{i\gamma})(F_{k\alpha}F_{k\beta}) = \delta_{ik}\delta_{\alpha\gamma}F_{k\beta} + F_{k\alpha}\delta_{ik}\ \delta_{\beta\gamma}$ beginning from the definition $\mathbf{C} = \mathbf{F}^\mathrm{T}\cdot\mathbf{F}$ in the indicial notation $C_{\alpha\beta} = F_{k\alpha}F_{k\beta}$.

11.9.3. Derive the first result of (11.81), namely $(\partial I_{\mathbf{C}}/\partial \mathbf{C}) = \mathbf{1}$.

11.9.4. Derive the result (11.83), $\mathbf{T}^{2PK} = 2[\{(\partial W/\partial I_{\mathbf{C}}) + I_{\mathbf{C}}(\partial W/\partial II_{\mathbf{C}}) + (\partial W/\partial III_{\mathbf{C}})II_{\mathbf{C}}\}\mathbf{1} - \{(\partial W/\partial II_{\mathbf{C}}) + I_{\mathbf{C}}(\partial W/\partial III_{\mathbf{C}})\}\mathbf{C} + (\partial W \partial III_{\mathbf{C}})\mathbf{C}^2]$, from the last of (11.77), $\mathbf{T}^{2PK} = 2(\partial W/\partial \mathbf{C})$, using (11.80), $W = W(I_C, II_C, III_C)$, (11.81) $\{(\partial I_{\mathbf{C}}/\partial \mathbf{C}) = \mathbf{1}$, $(\partial II_{\mathbf{C}}/\partial \mathbf{C}) = I_{\mathbf{C}}\mathbf{1} - \mathbf{C}$, $(\partial III_{\mathbf{C}}/\partial \mathbf{C}) = III_{\mathbf{C}}\mathbf{C}^{-1}\}$, and the Cayley Hamilton theorem, $\mathbf{C}^3 - I_C\mathbf{C}^2 + II_C\mathbf{C} - III_C = 0$.

11.9.5. Take the derivative of the Cayley Hamilton theorem, $\mathbf{C}^3 - I_C\mathbf{C}^2 + II_C\mathbf{C} - III_C = 0$, with respect to $\mathbf{C}$ and employ the three formulas (11.81) $\{(\partial I_{\mathbf{C}}/\partial \mathbf{C}) = \mathbf{1}, (\partial II_{\mathbf{C}}/\partial \mathbf{C}) = I_{\mathbf{C}}\mathbf{1} - \mathbf{C}, (\partial III_{\mathbf{C}}/\partial \mathbf{C}) = III_{\mathbf{C}}\mathbf{C}^{-1}\}$ to eliminate the expressions that are the derivatives of the invariants with respect to $\mathbf{C}$. Then multiply through this result by $\mathbf{C}$ and simplify. What is the significance of the final result? Is it a correct equation? Would it be a correct equation if the Eq. (11.81) were not correct?

11.10 Incompressible Elasticity

The assumption of incompressibility is an idealization that means that no agency (stress, strain, electric field, temperature, etc.) can change the volume of the model of the material. The Jacobian $J = \text{Det } \mathbf{F}$ relates the element of volume dV in the undeformed configuration to the volume dv in the deformed configuration according to the rule (11.36), $dv = JdV$. The Jacobian J is related to the principal stretches by $J = \lambda_I\lambda_{II}\lambda_{III}$. The requirement of incompressibility may then be expressed in several different algebraic forms related to the deformation, $J = 1$, $\lambda_I\lambda_{II}\lambda_{III} = 1, III_C = III_B = 1$, etc., and to algebraic forms related to the motion such as $\text{tr}\mathbf{D} = \nabla\cdot\mathbf{v} = 0$ (Sect. 6.4). The assumption of incompressibility requires that the density ρ be a constant. The pressure field p in an incompressible material is a Lagrange multiplier (see Example 6.4.1) that serves the function of maintaining the incompressibility constraint, not a thermodynamic variable. Because the volume of the model material cannot change, p does no work; it is a function of $\mathbf{x}$ and t, $p(\mathbf{x}, t)$, to be determined by the solution of the system of differential equations and boundary/initial conditions.

Recall from Sect. 11.7 that the constitutive equation for an elastic material can be written $\mathbf{T} = \mathbf{g}(\mathbf{F})$ in terms of the Cauchy stress $\mathbf{T}$ and the deformation gradient $\mathbf{F}$ or as $\mathbf{T}^{1\text{PK}} = \mathbf{h}(\mathbf{F})$ (11.54) in terms of the first Piola–Kirchhoff stress tensor $\mathbf{T}^{1\text{PK}}$ or as $\mathbf{T}^{2\text{PK}} = \mathbf{t}(\mathbf{C})$ (11.56) in terms of the second Piola–Kirchhoff stress tensor $\mathbf{T}^{2\text{PK}}$ and the right Cauchy–Green tensor $\mathbf{C}$. For incompressible elastic materials the Cauchy stress tensor $\mathbf{T}$ must be replaced by $\mathbf{T} + p\mathbf{1}$ where p is the constitutively indeterminate pressure described above and conveniently interpreted as a Lagrange multiplier. The response functions, say $\mathbf{g}(\mathbf{F})$ above, are defined only for deformations or motions that satisfy the condition $J = \text{Det } \mathbf{F} = 1$. Thus $\mathbf{T} = \mathbf{g}(\mathbf{F})$ is replaced by $\mathbf{T} + p\mathbf{1} = \mathbf{g}(\mathbf{F})$ when the assumption of incompressibility is made, $\mathbf{T}^{1\text{PK}} = \mathbf{h}(\mathbf{F})$, (11.54), becomes

$$\mathbf{T}^{1PK} + p(\mathbf{F}^{-1})^T = \mathbf{h}(\mathbf{F}), \tag{11.86}$$

and $\mathbf{T}^{2\mathrm{PK}} = \mathbf{t}(\mathbf{C})$, (11.56), becomes

$$\mathbf{T}^{2PK} + p\mathbf{C}^{-1} = \mathbf{t}(\mathbf{C}), \tag{11.87}$$

where the functions $\mathbf{h}(\mathbf{F})$ and $\mathbf{t}(\mathbf{C})$ are defined only for deformations or motions that satisfy the incompressibility condition Det $\mathbf{F} = 1$.

When an elastic model material is both isotropic and incompressible there are further simplifications in the constitutive relations. For example (11.69) and (11.73) are now written in the simpler forms

$$\mathbf{T}^{2PK} + p\mathbf{C}^{-1} = a_1\mathbf{C} + a_2\mathbf{C}^2, \tag{11.88}$$

and

$$\mathbf{T} + p\mathbf{1} = h_1\mathbf{B} + h_{-1}\mathbf{B}^{-1}, \tag{11.89}$$

where the functions of the isotropic invariants in these representations also simplify,

$$a_1 = a_1(I_C, II_C),\ a_2 = a_2(I_C, II_C) \tag{11.90}$$

and

$$h_{-1} = h_{-1}(I_C, II_C) = h_{-1}(I_B, II_B),\ h_{-1} = h_{-1}(I_C, II_C) = h_{-1}(I_B, II_B). \tag{11.91}$$

When an elastic model material is isotropic, incompressible and hyperelastic there are even further simplifications in the constitutive relations. In this case the strain energy per unit volume W depends only on $I_C = I_B$ and $II_C = II_B$ and (11.93) reduces to

$$\mathbf{T} = -p\mathbf{1} + h_1\mathbf{B} + h_{-1}\mathbf{B}^{-1}, \tag{11.92}$$

where

$$h_1 = 2\frac{\partial W}{\partial I_C},\ h_{-1} = -2\frac{\partial W}{\partial II_C}. \tag{11.93}$$

Even with all these restrictive assumptions (hyperelasticity, isotropy, and incompressibility) a complete solution of many interesting problems is not possible. Simpler models based on specialized assumptions but which retain the basic characteristics of the nonlinear elastic response have been proposed for polymeric materials and for biological tissues. An example that stems from research on the

constitutive behavior of rubber is the *Mooney–Rivlin* material with the constitutive equation (Mooney 1940; Rivlin and Saunders 1976)

$$\mathbf{T} = -p\mathbf{1} + \mu\left(\frac{1}{2} + \beta\right)\mathbf{B} - \mu\left(\frac{1}{2} - \beta\right)\mathbf{B}^{-1}, \tag{11.94}$$

where μ and β are constants, which has the following strain energy function

$$W = \frac{1}{2}\mu\left[\left(\frac{1}{2} + \beta\right)(I_B - 3) + \left(\frac{1}{2} - \beta\right)(II_B - 3)\right], \tag{11.95}$$

where the inequalities $\mu > 0$ and $(1/2) \geq \beta \geq -(1/2)$ are imposed upon the constants μ and β so that the strain energy W is a positive semidefinite quantity. The special case of the *Mooney–Rivlin* material when $\beta = (1/2)$ is called the *neo-Hookian* material

$$\mathbf{T} = -p\mathbf{1} + \mu\mathbf{B}. \tag{11.96}$$

Employing the assumption of incompressibility is the development of a constitutive model for a soft biological tissue is quite easy to justify because soft tissues contain so much water that their effective bulk compressibility is that of water, 2.3 GPa. When one compares the shear or deviatoric moduli of a soft biological tissue with 2.3 GPa, it is usually orders of magnitude less. Only in the case of hard tissues does the shear or deviatoric moduli approach and exceed (up to an order of magnitude) the effective bulk compressibility of water.

Problems

11.10.1. Derive (11.94) { $\mathbf{T} = -p\mathbf{1} + \mu((1/2) + \beta)\mathbf{B} - \mu((1/2) - \beta)\mathbf{B}^{-1}$ } from (11.95) { $W = (1/2)\mu[((1/2) + \beta)(I_B - 3) + ((1/2) - \beta)(II_B - 3)]$using (11.92) {$\mathbf{T} = -p\mathbf{1} + h_1\mathbf{B} + h_{-1}\mathbf{B}^{-1}$} and (11.93) {$h_1 = 2(\partial W/\partial I_C)$, $h_{-1} = -2(\partial W/\partial II_{\mathbf{C}})$}.

11.10.2. Calculate the components of the Cauchy stress T in a *Mooney–Rivlin* material (11.94) {$\mathbf{T} = -p\mathbf{1} + \mu((1/2) + \beta)\mathbf{B} - \mu((1/2) - \beta)\mathbf{B}^{-1}$} when the material is subjected to a simple shearing deformation given by $x_1 = X_I + kX_{II}, x_2 = X_{II}, x_3 = X_{III}$. Require that the normal stress acting on the surface whose normal is in the x_3 or X_{III} direction be zero.

11.4.2. Calculate the components of the Cauchy stress T in a *neo-Hookian* material (11.96) { $\mathbf{T} = -p\mathbf{1} + \mu\mathbf{B}$ } when the material is subjected to a simple shearing deformation given by $x_1 = X_I + kX_{II}, x_2 = X_{II}, x_3 = X_{III}$. Require that the normal stress acting on the surface whose normal is in the x_3 or X_{III} direction be zero.

11.11 Transversely Isotropic Hyperelasticity

The specialized constitutive equations transversely isotropic hyperelastic materials are developed in this section. Recall from Chapter 4 that transversely isotropic material symmetry is characterized by a unique direction that serves as an axis of rotational symmetry for the material structure. The plane perpendicular or transverse to the unique direction is a plane of isotropy, hence the descriptive term *transverse isotropy*. The particular material symmetry of an object is only constant through infinitesimal deformations; larger deformations will change the type of material symmetry. Thus when the material symmetry of a finitely deformed elastic object is noted, it is the material symmetry of the reference or undeformed configuration, not the material symmetry of the deformed configuration.

The hyperelastic constitutive equation, the first of (11.77), is the starting point of this development,

$$\mathbf{T} = \frac{2}{J}\mathbf{F} \cdot \left(\frac{\partial W}{\partial \mathbf{C}}\right)^T \cdot \mathbf{F}^T, \tag{11.97}$$

where $\rho J = \rho_o$. The selected direction is taken as the $\mathbf{e}_3$ axis and all orthogonal rotations about that axis by an angle ϕ leave the value of $W(\boldsymbol{C})$ unchanged. Let $\mathbf{R}(\phi)$ represent an orthogonal transformation about $\mathbf{e}_3$ by and the angle ϕ,

$$\mathbf{R} \cdot \mathbf{R}^T = \mathbf{1},\ \mathbf{R} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}; \tag{11.98}$$

it follows from

$$\frac{d\mathbf{R}}{d\phi} \cdot \mathbf{R}^T + \mathbf{R} \cdot \frac{d\mathbf{R}^T}{d\phi} = \mathbf{0}, \tag{11.99}$$

that

$$\Omega^T = -\Omega, \quad \textit{where } \Omega \equiv \frac{d\mathbf{R}}{d\phi} \cdot \mathbf{R}^T, \tag{11.100}$$

hence the representation of the components of Ω as an axial vector $\overline{\Omega}$, where

$$\Omega_{ij} = e_{ijk}\bar{\Omega}_k. \tag{11.101}$$

The requirement of the invariance of $W(\mathbf{C})$ under the rotation $\mathbf{R}$ may be written as

$$W(\mathbf{C}) = W(\mathbf{C}'), \quad \textit{where } \mathbf{C}' = \mathbf{R} \cdot \mathbf{C} \cdot \mathbf{R}^T, \tag{11.102}$$

from which it follows

$$\frac{d}{d\phi}W(\mathbf{C}') = 0 = \frac{dW}{d\mathbf{C}'} : \frac{d\mathbf{C}'}{d\phi} = \frac{dW}{d\mathbf{C}'} : [\boldsymbol{\Omega} \cdot \mathbf{C}' + \mathbf{C}' \cdot \boldsymbol{\Omega}^T], \tag{11.103}$$

or

$$\frac{d}{d\phi}W(\mathbf{C}') = 0 = 2\boldsymbol{\Omega} \cdot \mathbf{C}' : \frac{dW}{d\mathbf{C}'}. \tag{11.104}$$

Substitution of (11.101) into (11.104) above yields

$$e_{ij3}\hat{\Omega}_3 C_{im}\frac{dW}{dC_{mj}} = 0, \tag{11.105}$$

which requires that

$$C_{1m}\frac{dW}{dC_{m2}} = C_{2n}\frac{dW}{dC_{n1}}, \tag{11.106}$$

if (11.105) is to be true for all ϕ. Expanding the two sets of summation indexes in (11.106) we obtain

$$C_{12}\left(\frac{dW}{dC_{22}} - \frac{dW}{dC_{11}}\right) + (C_{11} - C_{22})\frac{dW}{dC_{12}} + C_{13}\frac{dW}{dC_{23}} - C_{23}\frac{dW}{dC_{13}} = 0 \tag{11.107}$$

The form of the function $W(\mathbf{C})$ invariant under the rotation (11.98), for all ϕ, is obtained by solving this differential equation (Ericksen and Rivlin 1954; Januzemis 1967). Simplification of this result is obtained if the following notation is introduced,

$$\begin{aligned} &C_{11} = \xi_1 + \xi_2 \cos\xi_3, \quad C_{22} = \xi_1 - \xi_2 \cos\xi_3, \\ &C_{12} = \xi_2 \sin\xi_3, C_{13} = \xi_4 \cos\xi_5, \quad C_{23} = \xi_4 \sin\xi_5, \end{aligned} \tag{11.108}$$

where $\xi_2 \geq 0$, $\xi_4 \geq 0$ in which case (11.107) reduces to

$$2\frac{dW}{d\xi_3} + \frac{dW}{d\xi_5} = 0, \tag{11.109}$$

thus

$$W = W(C_{33}, \xi_1, \xi_2, \xi_3, \xi_3 - 2\xi_5, \xi_4). \tag{11.110}$$

Observe from (11.108) that W cannot depend upon $\xi_3 - 2\xi_5$ when either $\xi_2 = 0$ or $\xi_4 = 0$. To require that W be a single valued function of the components C_{ij} that enter into $\xi_3 - 2\xi_5$, it is necessary for W to be periodic, with period 2π, in $\xi_3 - 2\xi_5$. Therefore it may be assumed that W depends on $\xi_3 - 2\xi_5$ through $\cos(\xi_3 - 2\xi_5)$. From these observations it follows that W can be expressed as a function of the invariants

$$C_{33},\ C_{11} + C_{22} = 2\xi_1,\ (C_{11} - C_{22})^2 + 4C_{12}^2 = 4\xi_2^2,\ C_{13}^2 + C_{23}^2 = \xi_4^2,$$
$$(C_{11} - C_{22})(C_{13}^2 - C_{23}^2) + 4C_{12}C_{23}C_{31} = 2\xi_2\xi_4^2\cos(\xi_3 - 2\xi_5). \tag{11.111}$$

An equivalent set of five invariants is given by the isotropic invariants

$$I = I_{\mathbf{B}} = I_{\mathbf{C}},\ II = II_{\mathbf{B}} = II_{\mathbf{C}},\ III = III_{\mathbf{B}} = III_{\mathbf{C}}, \tag{11.112}$$

and complemented by

$$IV \equiv C_{13}^2 + C_{23}^2 = \xi_4^2,\ V \equiv C_{33}. \tag{11.113}$$

Thus $W(\mathbf{C})$ has the representation $W(\mathbf{C}) = W(I, II, III, IV, V)$ and it follows from (11.97) that the stress has the representation

$$\begin{aligned}\mathbf{T} &= \frac{2}{J}\mathbf{F}\cdot\left(\frac{\partial W}{\partial Y}\right)^T\cdot\left(\frac{\partial Y}{\partial \mathbf{C}}\right)\cdot\mathbf{F}^T,\ \text{or } T_{ij}\\ &= \frac{2}{J}\sum_{Y=I,II,III,IV,V} F_{ik}\left(\frac{\partial W}{\partial Y}\right)^T\left(\frac{\partial Y}{\partial C_{km}}\right)F_{mj}\end{aligned} \tag{11.114}$$

which may be rewritten as

$$\mathbf{T} = \frac{2}{J}\left[\left(II\frac{\partial II}{\partial \mathbf{C}} + III\frac{\partial III}{\partial \mathbf{C}}\right)\mathbf{1} + \frac{\partial I}{\partial \mathbf{C}}\mathbf{B} - III\frac{\partial II}{\partial C}\mathbf{B}^{-1} + \frac{\partial IV}{\partial \mathbf{C}}\mathbf{M} + \frac{\partial V}{\partial \mathbf{C}}\mathbf{N}\right], \tag{11.115}$$

where

$$M_{ij} = \sum_{\alpha=1}^{2} C_{\alpha 3}(F_{i\alpha}F_{j3} + F_{j\alpha}F_{i3}),\ N_{ij} = F_{i3}F_{j3}, \tag{11.116}$$

which are related to IV and V by

$$2M_{ij} = F_{ik}\frac{\partial IV}{\partial F_{jk}},\ 2N_{ij} = F_{ik}\frac{\partial V}{\partial F_{jk}},\ \text{or } 2\mathbf{N} = \mathbf{F}\cdot\left(\frac{\partial V}{\partial \mathbf{F}}\right)^T,\ 2\mathbf{M} = \mathbf{F}\cdot\left(\frac{\partial IV}{\partial \mathbf{F}}\right)^T. \tag{11.117}$$

If this transversely isotropic hyperelastic material is also incompressible, then $III = 1$ and

$$\mathbf{T} = -p\mathbf{1} + 2\left[\frac{\partial I}{\partial \mathbf{C}}\mathbf{B} - \frac{\partial II}{\partial \mathbf{C}}\mathbf{B}^{-1} + \frac{\partial IV}{\partial \mathbf{C}}\mathbf{M} + \frac{\partial V}{\partial \mathbf{C}}\mathbf{N}\right]. \tag{11.118}$$

The remainder of the chapter deals with incompressible transversely isotropic hyperelastic material.

Example 11.11.1
Determine the stress tensor in a rectangular parallelepiped of an incompressible transversely isotropic hyperelastic material in which the unique direction coincides with the long dimension of the parallelepiped and the two transverse dimensions are equal. There is only a force applied to the parallelepiped in the long dimension of the parallelepiped and λ denotes the principal stretch in the long dimension. The stress applied to the parallelepiped is zero in the two transverse dimensions.

Solution: For the situation described above Eqs. (11.28), (11.33), and (11.116) may be used to show that

$$\mathbf{B} = \begin{bmatrix} \lambda^{-1} & 0 & 0 \\ 0 & \lambda^{-1} & 0 \\ 0 & 0 & \lambda^2 \end{bmatrix}, \quad \mathbf{B}^{-1} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda^{-2} \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda^2 \end{bmatrix}, \quad \mathbf{M} = 0,$$

thus from (11.118) it follows that

$$T_{11} = T_{22} = -p + 2\left[\frac{\partial I}{\partial \mathbf{C}}\lambda^{-1} - \frac{\partial II}{\partial \mathbf{C}}\lambda\right],$$

and since these two stress are zero, it follows that $p = 2[(\partial I/\partial \mathbf{C})\lambda^{-1} - (\partial II/\partial \mathbf{C})\lambda]$ and that the axial stress is given by

$$T_{33} = 2\lambda[(\lambda - \lambda^{-2})\left[\frac{\partial I}{\partial \mathbf{C}} + \frac{1}{\lambda}\frac{\partial II}{\partial \mathbf{C}} + \frac{\partial V}{\partial \mathbf{C}}\lambda\right].$$

If the material was isotropic rather than transversely isotropic, then the same result would apply with $\frac{\partial V}{\partial \mathbf{C}} = 0$.

A number of solutions for transversely isotropic hyperelastic materials in cylindrical coordinates were obtained by Ericksen and Rivlin (1954); some of these solutions are contained in the book by Januzemis (1967).

11.12 Relevant Literature

The developments in nonlinear of nonlinear elasticity are described in many books, for example Truesdell (1960), Truesdell and Toupin (1960), Green and Adkins (1960), Januzemis (1967), Treloar (1967), Truesdell and Noll (1965), and Ogden (1984) among many others.

References

Danielson DA (1973) Human skin as an elastic membrane. J Biomech 6:539–546
Danielson DA, Natarajan S (1975) Tension field theory and the stress in stretched skin. J Biomech 8:135–142
Ericksen JL, Rivlin RS (1954) Large elastic deformations of homogeneous anisotropic materials. J Rat Mech Anal 3:281–301
Green AE, Adkins JE (1960) Large Elastic Deformations and Non-Linear Continuum Mechanics. Claredon, Oxford
Januzemis W (1967) Continuum Mechanics. Macmillan, New York
Lockett FJ (1972) Nonlinear Viscoelastic Solids. Academic, London, New York
Mooney M (1940) A theory of large elastic deformation. J Appl Phys 11:582–592
Ogden RW (1984) Non-linear Elastic Deformations. Ellis Horwood Ltd, Chichester
Rivlin RS, Saunders DW (1976) The strain energy function for elastomers. Trans Soc Rheol 20:545–557
Treloar LRG (1967) The Physics of Rubber Elasticity, 2nd edn. Oxford University Press, New York
Truesdell CA (1960) Principles of Continuum Mechanics, Colloquium Lectures in Pure and Figure Applied Science. Socony Mobil Oil Company, Dallas, TX
Truesdell CA, Noll W (1965) The Non-linear Field Theories of Mechanics. In: Flugge S (ed) Handbuch der Physik. Springer Verlag, Berlin
Truesdell CA, Toupin RA (1960) The Classical Field Theories. In: Flugge S (ed) Handbuch der Physik. Springer Verlag, Berlin

Appendix A: Matrices and Tensors

A.1 Introduction and Rationale

The purpose of this appendix is to present the notation and most of the mathematical techniques that will be used in the body of the text. The audience is assumed to have been through several years of college level mathematics that included the differential and integral calculus, differential equations, functions of several variables, partial derivatives, and an introduction to linear algebra. Matrices are reviewed briefly and determinants, vectors, and tensors of order two are described. The application of this linear algebra to material that appears in undergraduate engineering courses on mechanics is illustrated by discussions of concepts like the area and mass moments of inertia, Mohr's circles and the vector cross and triple scalar products. The solutions to ordinary differential equations are reviewed in the last two sections. The notation, as far as possible, will be a matrix notation that is easily entered into existing symbolic computational programs like Maple, Mathematica, Matlab, and Mathcad etc. The desire to represent the components of three-dimensional fourth order tensors that appear in anisotropic elasticity as the components of six-dimensional second order tensors and thus represent these components in matrices of tensor components in six dimensions leads to the nontraditional part of this appendix. This is also one of the nontraditional aspects in the text of the book, but a minor one. This is described in Sect. A.11, along with the rationale for this approach.

A.2 Definition of Square, Column, and Row Matrices

An r by c matrix $\mathbf{M}$ is a rectangular array of numbers consisting of r rows and c columns,

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2,

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & . & . & . & M_{1c} \\ M_{21} & M_{22} & . & . & . & M_{2c} \\ . & . & . & . & . & . \\ M_{r1} & . & . & . & . & M_{rc} \end{bmatrix}. \tag{A.1}$$

The typical element of the array, M_{ij}, is the ith element in the jth column; in this text the elements M_{ij} will be real numbers or functions whose values are real numbers. The transpose of the matrix $\mathbf{M}$ is denoted by $\mathbf{M}^T$ and is obtained from $\mathbf{M}$ by interchanging the rows and columns

$$\mathbf{M}^T = \begin{bmatrix} M_{11} & M_{21} & . & . & . & M_{r1} \\ M_{12} & M_{22} & . & . & . & M_{r2} \\ . & . & . & . & . & . \\ M_{1c} & . & . & . & . & M_{rc} \end{bmatrix}. \tag{A.2}$$

The operation of obtaining $\mathbf{M}^T$ from $\mathbf{M}$ is called transposition. In this text we are interested in special cases of the r by c matrix $\mathbf{M}$. These special cases are those of the square matrix, $r = c = n$, the case of the row matrix, $r = 1, c = n$, and the case of column matrix, $r = n, c = 1$. Further, the special sub-cases of interest are $n = 2$, $n = 3$, and $n = 6$; the sub-case $n = 1$ reduces all three special cases to the trivial situation of a single number or scalar. A square matrix $\mathbf{A}$ has the form

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & . & . & . & A_{1n} \\ A_{21} & A_{22} & . & . & . & A_{2n} \\ . & . & . & . & . & . \\ A_{n1} & . & . & . & . & A_{nn} \end{bmatrix}, \tag{A.3}$$

while row and column matrices, $\mathbf{r}$ and $\mathbf{c}$, have the forms

$$\mathbf{r} = [\, r_1 \quad r_2 \quad . \quad . \quad . \quad r_n \,], \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_n \end{bmatrix}, \tag{A.4}$$

respectively. The transpose of a column matrix is a row matrix, thus

$$\mathbf{c}^T = [\, c_1 \quad c_2 \quad . \quad . \quad . \quad c_n \,]. \tag{A.5}$$

To save space in books and papers the form of $\mathbf{c}$ in (A.5) is used more frequently than the form in the second of (A.4). Wherever possible, square matrices will be denoted by upper case boldface Latin letters, while row and column matrices will be denoted by lower case boldface Latin letters as is the case in equations (A.3) and (A.4).

A.3 The Types and Algebra of Square Matrices

The elements of the square matrix **A** given by (A.3) for which the row and column indices are equal, namely the elements A_{11}, A_{22},..., A_{nn}, are called diagonal elements. The sum of the diagonal elements of a matrix is a scalar called the trace of the matrix and, for a matrix **A**, it is denoted by tr**A**,

$$\mathrm{tr}\mathbf{A} = A_{11} + A_{22} + \cdots + A_{nn}. \tag{A.6}$$

If the trace of a matrix is zero, the matrix is said to be traceless. Note also that $\mathrm{tr}\mathbf{A} = \mathrm{tr}\mathbf{A}^{\mathrm{T}}$. A matrix with only diagonal elements is called a diagonal matrix,

$$\mathbf{A} = \begin{bmatrix} A_{11} & 0 & . & . & . & 0 \\ 0 & A_{22} & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & . & A_{nn} \end{bmatrix}. \tag{A.7}$$

The zero and the unit matrix, **0** and **1**, respectively, constitute the null element, the 0, and the unit element, the 1, in the algebra of square matrices. The zero matrix is a matrix whose every element is zero and the unit matrix is a diagonal matrix whose diagonal elements are all one:

$$\mathbf{0} = \begin{bmatrix} 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & . & 0 \end{bmatrix}, \mathbf{1} = \begin{bmatrix} 1 & 0 & . & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & . & 1 \end{bmatrix}. \tag{A.8}$$

A special symbol, the Kronecker delta δ_{ij}, is introduced to represent the components of the unit matrix. When the indices are equal, $i = j$, the value of the Kronecker delta is one, $\delta_{11} = \delta_{22} = \ldots = \delta_{nn} = 1$ and when they are unequal, $i \neq j$, the value of the Kronecker delta is zero, $\delta_{12} = \delta_{21} = \ldots = \delta_{n1} = \delta_{1n} = 0$. The multiplication of a matrix **A** by a scalar is defined as the multiplication of every element of the matrix **A** by the scalar α, thus

$$\alpha\mathbf{A} = \begin{bmatrix} \alpha A_{11} & \alpha A_{12} & . & . & . & \alpha A_{1n} \\ \alpha A_{21} & \alpha A_{22} & . & . & . & \alpha A_{2n} \\ . & . & . & . & . & . \\ \alpha A_{n1} & . & . & . & . & \alpha A_{nn} \end{bmatrix}. \tag{A.9}$$

It is then easy to show that $1\mathbf{A} = \mathbf{A}$, $-1\mathbf{A} = -\mathbf{A}$, $0\mathbf{A} = \mathbf{0}$, and $\alpha\mathbf{0} = \mathbf{0}$. The addition of matrices is defined only for matrices with the same number of rows and columns. The sum of two matrices, **A** and **B**, is denoted by $\mathbf{A} + \mathbf{B}$, where

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} & . & . & . & A_{1n}+B_{1n} \\ A_{21}+B_{21} & A_{22}+B_{22} & . & . & . & A_{2n}+B_{2n} \\ . & . & . & . & . & . \\ A_{n1}+B_{n1} & . & . & . & . & A_{nn}+B_{nn} \end{bmatrix}. \tag{A.10}$$

Matrix addition is commutative and associative,

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A} \text{ and } \mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}, \tag{A.11}$$

respectively. The following distributive laws connect matrix addition and matrix multiplication by scalars:

$$\alpha(\mathbf{A}+\mathbf{B})=\alpha\mathbf{A}+\alpha\mathbf{B} \text{ and } (\alpha+\beta)\mathbf{A}=\alpha\mathbf{A}+\beta\mathbf{A}, \tag{A.12}$$

where α and β are scalars. Negative square matrices may be created by employing the definition of matrix multiplication by a scalar (A.8) in the special case when $\alpha=-1$. In this case the definition of addition of square matrices (A.10) can be extended to include the subtraction of square matrices, $\mathbf{A}-\mathbf{B}$.

A matrix for which $\mathbf{B}=\mathbf{B}^{\mathrm{T}}$ is said to be a symmetric matrix and a matrix for which $\mathbf{C}=-\mathbf{C}^{\mathrm{T}}$ is said to be a skew-symmetric or anti-symmetric matrix. The symmetric and anti-symmetric parts of a matrix, say $\mathbf{A}^{\mathrm{S}}$ and $\mathbf{A}^{\mathrm{A}}$, are constructed from $\mathbf{A}$ as follows:

$$\text{symmetric part of } \mathbf{A}^{\mathrm{S}}=\frac{1}{2}(\mathbf{A}+\mathbf{A}^{\mathrm{T}}), \text{and} \tag{A.13}$$

$$\text{anti}-\text{symmetric part of } \mathbf{A}^{\mathrm{A}}=\frac{1}{2}(\mathbf{A}-\mathbf{A}^{\mathrm{T}}). \tag{A.14}$$

It is easy to verify that the symmetric part of $\mathbf{A}$ is a symmetric matrix and that the skew-symmetric part of $\mathbf{A}$ is a skew-symmetric matrix. The sum of the symmetric part of $\mathbf{A}$ and the skew-symmetric part of $\mathbf{A}$, $\mathbf{A}^{\mathrm{S}}+\mathbf{A}^{\mathrm{A}}$, is $\mathbf{A}$:

$$\mathbf{A}=\mathbf{A}^{\mathrm{S}}+\mathbf{A}^{\mathrm{A}}=\frac{1}{2}(\mathbf{A}+\mathbf{A}^{\mathrm{T}})+\frac{1}{2}(\mathbf{A}-\mathbf{A}^{\mathrm{T}}). \tag{A.15}$$

This result shows that any square matrix can be decomposed into the sum of a symmetric and a skew-symmetric matrix or anti-symmetric matrix. Using the trace operation introduced above, the representation (A.15) can be extended to three-way decomposition of the matrix $\mathbf{A}$,

$$\mathbf{A}=\frac{\mathrm{tr}\mathbf{A}}{n}\mathbf{1}+\frac{1}{2}\left[\mathbf{A}+\mathbf{A}^{\mathrm{T}}-2\frac{\mathrm{tr}\mathbf{A}}{n}\mathbf{1}\right]+\frac{1}{2}(\mathbf{A}-\mathbf{A}^{\mathrm{T}}). \tag{A.16}$$

The last term in this decomposition is still the skew-symmetric part of the matrix. The second term is the traceless symmetric part of the matrix and the first term is simply the trace of the matrix multiplied by the unit matrix.

Example A.3.1

Construct the three-way decomposition of the matrix $\mathbf{A}$ given by:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Solution: The symmetric and skew-symmetric parts of $\mathbf{A}$, $\mathbf{A}^{\mathrm{S}}$, and $\mathbf{A}^{\mathrm{A}}$, as well as the trace of $\mathbf{A}$ are calculated,

$$\mathbf{A}^{\mathrm{S}} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}}) = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 5 & 7 \\ 5 & 7 & 9 \end{bmatrix}, \mathbf{A}^{\mathrm{A}} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^{\mathrm{T}}) = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}, \mathrm{tr}\mathbf{A} = 15;$$

then, since $n = 3$, it follows from (A.16) that

$$\mathbf{A} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} + \begin{bmatrix} -4 & 3 & 5 \\ 3 & 0 & 7 \\ 5 & 7 & 4 \end{bmatrix} + \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}.$$

Introducing the notation for the deviatoric part of an n by n square matrix $\mathbf{A}$,

$$\mathrm{dev}\mathbf{A} = \mathbf{A} - \frac{\mathrm{tr}\mathbf{A}}{n}\mathbf{1}, \tag{A.17}$$

the representation for the matrix $\mathbf{A}$ given by (A.16) may be rewritten as the sum of three terms

$$\mathbf{A} = \frac{\mathrm{tr}\mathbf{A}}{n}\mathbf{1} + \mathrm{dev}\mathbf{A}^{\mathrm{S}} + \mathbf{A}^{\mathrm{A}}, \tag{A.18}$$

where the first term is called the *isotropic*, or *spherical* (or *hydrostatic* as in hydraulics) part of $\mathbf{A}$. Note that

$$\mathrm{dev}\mathbf{A}^{\mathrm{S}} = \frac{1}{2}(\mathrm{dev}\mathbf{A} + \mathrm{dev}\mathbf{A}^{\mathrm{T}}). \tag{A.19}$$

Example A.3.2

Show that tr (dev$\mathbf{A}$) $= 0$.

Solution: Applying the trace operation to both sides of (A.17) one obtains

$$\mathrm{tr}(\mathrm{dev}\mathbf{A}) = \mathrm{tr}\mathbf{A} - 1/n(\mathrm{tr}\mathbf{A})(\mathrm{tr}\mathbf{1});$$

then, since $\mathrm{tr}\mathbf{1} = n$, it follows that $\mathrm{tr}(\mathrm{dev}\mathbf{A}) = 0$.

The product of two square matrices, **A** and **B**, with equal numbers of rows (columns) is a square matrix with the same number of rows (columns). The matrix product is written as **A**·**B** where **A**·**B** is defined by

$$(\mathbf{A}\cdot\mathbf{B})_{\mathrm{ij}} = \sum_{k=1}^{k=n} A_{\mathrm{ik}}B_{\mathrm{kj}}; \tag{A.20}$$

thus, for example, the element in the *r*th row and *c*th column of the product **A**·**B** is given by

$$(\mathbf{A}\cdot\mathbf{B})_{\mathrm{rc}} = A_{\mathrm{r1}}B_{\mathrm{1c}} + A_{\mathrm{r2}}B_{\mathrm{2c}} + \cdots + A_{\mathrm{rn}}B_{\mathrm{nc}}.$$

The widely used notational convention, called the Einstein summation convention, allows one to simplify the notation by dropping the summation symbol in (A.20) so that

$$(\mathbf{A}\cdot\mathbf{B})_{\mathrm{ij}} = A_{\mathrm{ik}}B_{\mathrm{kj}}, \tag{A.21}$$

where the convention is the understanding that the repeated index, in this case *k*, is to be summed over its *range* of the admissible values from 1 to *n*. This summation convention will be used from this point forward in this Appendix and in the body of the text. For $n = 6$, the range of admissible values is 1–6, including 2, 3, 4, and 5. The two *k* indices are the *summation or dummy* indices; note that the implied summation is unchanged if both of the *k*'*s* are replaced by any other letter of the alphabet. *A summation index* is defined as an index that occurs in a summand twice and only twice. Note that *summands* are terms in equations separated from each other by plus, minus, or equal signs. The existence of summation indices in a summand requires that the summand be summed with respect to those indices over the entire range of admissible values. Note again that the summation index is only a means of stating that a term must be summed, and the letter used for this index is immaterial, thus $A_{\mathrm{im}}B_{\mathrm{mj}}$ has the same meaning as $A_{\mathrm{ik}}B_{\mathrm{kj}}$. The other indices in the formula (A.22), the *i* and *j* indices, are called free indices. A *free index* is free to take on any one of the admissible values in its *range* from 1 to *n*. For example if *n* were 3, the free index could be 1, 2, or 3. A *free* index is formally defined as an index that occurs once and only once in every summand of an equation. The total number of equations that may be represented by an equation with one free index is the range of the admissible values. Thus the equation (A.21) represents n^2 separate equations. For two 2 by 2 matrices *A* and *B*, the product is written as

$$\mathbf{A}\cdot\mathbf{B}=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}$$
$$=\begin{bmatrix}A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22}\\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22}\end{bmatrix}, \tag{A.22}$$

where, in this case, the products (A.20) and (A.21) stand for the $n^2 = 2^2 = 4$ separate equations, the right-hand sides of which are the four elements of the last matrix in (A.22).

The dot between the matrix product **A**·**B** indicates that one index from **A** and one index from **B** is to be summed over. The positioning of the summation index on the two matrices involved in a matrix product is critical and is reflected in the matrix notation by the transpose. In the three equations below, (A.21), study carefully how the positions of the summation indices within the summation sign change in relation to the position of the transpose on the matrices in the associated matrix product:

$$(\mathbf{A}\cdot\mathbf{B}^{\mathrm{T}})_{\mathrm{ij}}=A_{\mathrm{ik}}B_{\mathrm{jk}},\ (\mathbf{A}^{\mathrm{T}}\cdot\mathbf{B})_{\mathrm{ij}}=A_{\mathrm{ki}}B_{\mathrm{kj}},\ (\mathbf{A}^{\mathrm{T}}\cdot\mathbf{B}^{\mathrm{T}})_{\mathrm{ij}}=A_{\mathrm{ki}}B_{\mathrm{jk}}. \tag{A.23}$$

A very significant feature of matrix multiplication is noncommutatively, that is to say **A**·**B** ≠ **B**·**A**. Note, for example, the transposed product **B**·**A** of the multiplication represented in (A.22),

$$\mathbf{B}\cdot\mathbf{A}=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$$
$$=\begin{bmatrix}B_{11}A_{11}+B_{12}A_{21} & B_{11}A_{12}+B_{12}A_{22}\\ B_{21}A_{11}+B_{22}A_{21} & B_{21}A_{12}+B_{22}A_{22}\end{bmatrix}, \tag{A.24}$$

is an illustration of the fact that **A**·**B** ≠ **B**·**A**, in general. If **A**·**B** = **B**·**A**, the matrices **A** and **B** are said to commute. Finally, matrix multiplication is associative,

$$\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})=(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}, \tag{A.25}$$

and matrix multiplication is distributive with respect to addition

$$\mathbf{A}\cdot(\mathbf{B}+\mathbf{C})=\mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C}\ \text{and}\ (\mathbf{B}+\mathbf{C})\cdot\mathbf{A}=\mathbf{B}\cdot\mathbf{A}+\mathbf{C}\cdot\mathbf{A}, \tag{A.26}$$

provided the results of these operations are defined.

Example A.3.3

Construct the products **A**·**B** and **B**·**A** of the matrices **A** and **B** given by:

$$\mathbf{A}=\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix},\mathbf{B}=\begin{bmatrix}10 & 11 & 12\\ 13 & 14 & 15\\ 16 & 17 & 18\end{bmatrix}.$$

Solution: The products $\mathbf{A}\cdot\mathbf{B}$ and $\mathbf{B}\cdot\mathbf{A}$ are given by

$$\mathbf{A}\cdot\mathbf{B}=\begin{bmatrix}84 & 90 & 96\\ 201 & 216 & 231\\ 318 & 342 & 366\end{bmatrix},\mathbf{B}\cdot\mathbf{A}=\begin{bmatrix}138 & 171 & 204\\ 174 & 216 & 258\\ 210 & 261 & 312\end{bmatrix}.$$

Observe that $\mathbf{A}\cdot\mathbf{B}\neq\mathbf{B}\cdot\mathbf{A}$.

The *double dot* notation between the two second order tensors is an extension of the single dot notation between the matrices, $\mathbf{A}\cdot\mathbf{B}$, which indicates that one index from $\mathbf{A}$ and one index from $\mathbf{B}$ are to be summed over; the double dot notation between the matrices, $\mathbf{A}:\mathbf{B}$, indicates that both indices of $\mathbf{A}$ are to be summed with different indices from $\mathbf{B}$, thus

$$\mathbf{A}:\mathbf{B}=A_{\mathrm{ik}}B_{\mathrm{ik}}.$$

This colon notation stands for the same operation as the trace of the product, $\mathbf{A}:\mathbf{B}=\mathrm{tr}(\mathbf{A}\cdot\mathbf{B})$. Although $\mathrm{tr}(\mathbf{A}\cdot\mathbf{B})$ and $\mathbf{A}:\mathbf{B}$ mean the same thing, $\mathbf{A}:\mathbf{B}$ involves fewer characters and it will be the notation of choice. Note that $\mathbf{A}:\mathbf{B}=\mathbf{A}^{\mathrm{T}}:\mathbf{B}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}:\mathbf{B}=\mathbf{A}:\mathbf{B}^{\mathrm{T}}$ but that $\mathbf{A}:\mathbf{B}\neq\mathbf{A}^{\mathrm{T}}:\mathbf{B}$ in general.

In the considerations of mechanics, matrices are often functions of coordinate positions x_1, x_2, x_3, and time t. In this case the matrix is written $\mathbf{A}(x_1, x_2, x_3, t)$ which means that each element of $\mathbf{A}$ is a function of x_1, x_2, x_3, and t,

$$\mathbf{A}(x_1,x_2,x_3,t)=\begin{bmatrix}A_{11}(x_1,x_2,x_3,t) & A_{12}(x_1,x_2,x_3,t) & \dots & A_{1\mathrm{n}}(x_1,x_2,x_3,t)\\ A_{21}(x_1,x_2,x_3,t) & A_{22}(x_1,x_2,x_3,t) & \dots & A_{2\mathrm{n}}(x_1,x_2,x_3,t)\\ . & . & \dots & .\\ A_{\mathrm{n}1}(x_1,x_2,x_3,t) & . & \dots & A_{\mathrm{nn}}(x_1,x_2,x_3,t)\end{bmatrix}. \tag{A.27}$$

Let the operator $\diamondsuit$ stand for a total derivative, or a partial derivative with respect to x_1, x_2, x_3, or t, or a definite or indefinite (single or multiple) integral; then the operation of the operator on the matrix follows the same rule as the multiplication of a matrix by a scalar (A.9), thus

$$\diamondsuit\mathbf{A}(x_1,x_2,x_3,t)=\begin{bmatrix}\diamondsuit A_{11}(x_1,x_2,x_3,t) & \diamondsuit A_{12}(x_1,x_2,x_3,t) & \dots & \diamondsuit A_{1\mathrm{n}}(x_1,x_2,x_3,t)\\ \diamondsuit A_{21}(x_1,x_2,x_3,t) & \diamondsuit A_{22}(x_1,x_2,x_3,t) & \dots & \diamondsuit A_{2\mathrm{n}}(x_1,x_2,x_3,t)\\ . & . & \dots & .\\ \diamondsuit A_{\mathrm{n}1}(x_1,x_2,x_3,t) & . & \dots & \diamondsuit A_{\mathrm{nn}}(x_1,x_2,x_3,t)\end{bmatrix}. \tag{A.28}$$

$$\diamondsuit(\mathbf{A}+\mathbf{B})=\diamondsuit\mathbf{B}+\diamondsuit\mathbf{A}\ \text{and}\ (\diamondsuit_1+\diamondsuit_2)\mathbf{A}=\diamondsuit_1\mathbf{A}+\diamondsuit_2\mathbf{A}, \tag{A.29}$$

where $\diamondsuit_1$ and $\diamondsuit_2$ are two different operators.

Problems

A.3.1. Simplify the following expression by using the Einstein summation index convention for a range of three:

$$
\begin{aligned}
0 &= r_1 w_1 + r_2 w_2 + r_3 w_3, \\
\Psi &= (u_1 v_1 + u_2 v_2 + u_3 v_3)(u_1 v_1 + u_2 v_2 + u_3 v_3), \\
\phi &= A_{11} x_1^2 + A_{22} x_2^2 + A_{33} x_3^2 + A_{12} x_1 x_2 + A_{21} x_1 x_2 + A_{13} x_1 x_3 + A_{31} x_1 x_3 \\
&\quad + A_{23} x_3 x_2 + A_{32} x_3 x_2 .
\end{aligned}
$$

A.3.2. The matrix **M** has the numbers 4, 5, −5 in its first row, −1, 3, −1 in its second row and 7, 1, 1 in its third row. Find the transpose of **M**, the symmetric part of **M**, and the skew-symmetric part of **M**.

A.3.3. Prove that $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$.

A.3.4. Consider the hydrostatic component **H**, the deviatoric component **D** of the symmetric part of **A**, and the skew-symmetric component **S** of the square n by n matrix **A** defined by (A.17) and (A.18),

$$
\mathbf{H} = \frac{\mathrm{tr}\mathbf{A}}{n}\mathbf{1}, \mathbf{D} = \frac{1}{2}\left[\mathbf{A} + \mathbf{A}^{\mathrm{T}} - 2\frac{\mathrm{tr}\mathbf{A}}{n}\mathbf{1}\right] \text{ and } \mathbf{S} = \frac{1}{2}[\mathbf{A} - \mathbf{A}^{\mathrm{T}}].
$$

Evaluate the following: tr**H**, tr**D**, tr**S**, tr(**H**·**D**) = **H**:**D**, tr(**H**·**S**) = **H**:**S**, and tr(**S**·**D**) = **S**:**D**.

A.3.5. For the matrices in example A3.3 show that tr **A**·**B** = tr **B**·**A** = 666. In general, will **A**:**B** = **B**:**A**, or is this a special case?

A.3.6. Prove that **A**:**B** is zero if **A** is symmetric and **B** is skew-symmetric.

A.3.7. Calculate $\mathbf{A}^{\mathrm{T}}\cdot\mathbf{B}$, $\mathbf{A}\cdot\mathbf{B}^{\mathrm{T}}$, and $\mathbf{A}^{\mathrm{T}}\cdot\mathbf{B}^{\mathrm{T}}$ for the matrices **A** and **B** of Example A.3.3.

A.3.8. Find the derivative of the matrix $\mathbf{A}(t)$ with respect to t.

$$
\mathbf{A}(t) = \begin{bmatrix} t & t^2 & \sin\omega t \\ \cosh t & \ln t & 17t \\ 1/t & 1/t^2 & \ln t^2 \end{bmatrix}.
$$

A.3.9. Show that $(\mathbf{A}\cdot\mathbf{B})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\cdot\mathbf{A}^{\mathrm{T}}$.

A.3.10. Show that $(\mathbf{A}\cdot\mathbf{B}\cdot\mathbf{C})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\cdot\mathbf{B}^{\mathrm{T}}\cdot\mathbf{A}^{\mathrm{T}}$.

A.4 The Algebra of n-Tuples

The algebra of column matrices is the same as the algebra of row matrices. The column matrices need only be transposed to be equivalent to row matrices as illustrated in equations (A.4) and (A.5). A phrase that describes both row and

column matrices is n-tuples. This phrase will be used here because it is descriptive and inclusive. A zero n-tuple is an n-tuple whose entries are all zero; it is denoted by $\mathbf{0} = [0, 0, \ldots, 0]$. The multiplication of an n-tuple $\mathbf{r}$ by a scalar α is defined as the multiplication of every element of the n-tuple $\mathbf{r}$ by the scalar α, thus $\alpha\mathbf{r} = [\alpha r_1, \alpha r_2, \ldots, \alpha r_n]$. As with square matrices, it is then easy to show for n-tuples that $1\mathbf{r} = \mathbf{r}$, $-1\mathbf{r} = -\mathbf{r}$, $0\mathbf{r} = \mathbf{0}$, and $\alpha\mathbf{0} = \mathbf{0}$. The addition of n-tuples is only defined for n-tuples with the same n. The sum of two n-tuples, $\mathbf{r}$ and $\mathbf{t}$, is denoted by $\mathbf{r} + \mathbf{t}$, where $\mathbf{r} + \mathbf{t} = [r_1 + t_1, r_2 + t_2, \ldots, r_n + t_n]$. Row-matrix addition is commutative, $\mathbf{r} + \mathbf{t} = \mathbf{t} + \mathbf{r}$, and associative, $\mathbf{r} + (\mathbf{t} + \mathbf{u}) = (\mathbf{r} + \mathbf{t}) + \mathbf{u}$. The following distributive laws connect n-tuple addition and n-tuple multiplication by scalars, thus $\alpha(\mathbf{r} + \mathbf{t}) = \alpha\mathbf{r} + \alpha\mathbf{t}$ and $(\alpha + \beta)\mathbf{r} = \alpha\mathbf{r} + \beta\mathbf{r}$, where α and β are scalars. Negative n-tuples may be created by employing the definition of n-tuple multiplication by a scalar, $\alpha\mathbf{r} = [\alpha r_1, \alpha r_2, \ldots, \alpha r_n]$, in the special case when $\alpha = -1$. In this case the definition of addition of n-tuples, $\mathbf{r} + \mathbf{t} = [r_1 + t_1, r_2 + t_2, \ldots, r_n + t_n]$, can be extended to include the subtraction of n-tuples $\mathbf{r}-\mathbf{t}$, and the difference between n-tuples, $\mathbf{r}-\mathbf{t}$.

Two n-tuples may be employed to create a square matrix. The square matrix formed from $\mathbf{r}$ and $\mathbf{t}$ is called the *open* product of the n-tuples $\mathbf{r}$ and $\mathbf{t}$; it is denoted by $\mathbf{r} \otimes \mathbf{t}$, and defined by

$$\mathbf{r} \otimes \mathbf{t} = \begin{bmatrix} r_1t_1 & r_1t_2 & \ldots & r_1t_n \\ r_2t_1 & r_2t_2 & \ldots & r_2t_n \\ . & . & \ldots & . \\ r_nt_1 & . & \ldots & r_nt_n \end{bmatrix}. \tag{A.30}$$

The American physicist J. Willard Gibbs introduced the concept of the open product of vectors calling the product a *dyad*. This terminology is still used in some books and the notation is spoken of as the *dyadic* notation. The trace of this square matrix, $\mathrm{tr}\{\mathbf{r} \otimes \mathbf{t}\}$ is the scalar product of $\mathbf{r}$ and $\mathbf{t}$,

$$\mathrm{tr}\{\mathbf{r} \otimes \mathbf{t}\} = \mathbf{r} \cdot \mathbf{t} = r_1t_1 + r_2t_2 + \cdots + r_nt_n. \tag{A.31}$$

In the special case of $n = 3$, the skew-symmetric part of the open product $\mathbf{r} \otimes \mathbf{t}$,

$$\frac{1}{2}\begin{bmatrix} 0 & r_1t_2 - r_2t_1 & r_1t_3 - r_3t_1 \\ r_2t_1 - r_1t_2 & 0 & r_2t_3 - r_3t_2 \\ r_3t_1 - r_1t_3 & r_3t_2 - r_2t_3 & 0 \end{bmatrix}, \tag{A.32}$$

provides the components of the cross product of $\mathbf{r}$ and $\mathbf{t}$, denoted by $\mathbf{r} \times \mathbf{t}$, and written as $\mathbf{r} \times \mathbf{t} = [r_2t_3-r_3t_2, r_3t_1-r_1t_3, r_1t_2-r_2t_1]^{\mathrm{T}}$. These points concerning the dot product $\mathbf{r}\cdot\mathbf{t}$ and cross product $\mathbf{r} \times \mathbf{t}$ will be revisited later in this Appendix.

Example A.4.1

Given the n-tuples $\mathbf{a} = [1, 2, 3]$ and $\mathbf{b} = [4, 5, 6]$, construct the open product matrix, $\mathbf{a} \otimes \mathbf{b}$, the skew-symmetric part of the open product matrix, and trace of the open product matrix.

Solution:

$$\mathbf{a}\otimes\mathbf{b}=\begin{bmatrix}4&5&6\\8&10&12\\12&15&18\end{bmatrix},\left(\frac{1}{2}\right)(\mathbf{a}\otimes\mathbf{b}-(\mathbf{a}\otimes\mathbf{b})^{\mathrm{T}})=\left(\frac{3}{2}\right)\begin{bmatrix}0&-1&-2\\1&0&-1\\2&1&0\end{bmatrix}.$$

and $\mathrm{tr}\{\mathbf{a}\otimes\mathbf{b}\}=\mathbf{a}\cdot\mathbf{b}=32$.

Frequently n-tuples are considered as functions of coordinate positions x_1, x_2, x_3, and time t. In this case the n-tuple is written $\mathbf{r}(x_1, x_2, x_3, t)$ that means that each element of $\mathbf{r}$ is a function of x_1, x_2, x_3, and t,

$$\mathbf{r}(x_1,x_2,x_3,t)=[r_1(x_1,x_2,x_3,t),r_2(x_1,x_2,x_3,t),\ldots,r_{\mathrm{n}}(x_1,x_2,x_3,t)]. \tag{A.33}$$

Again letting the operator $\diamondsuit$ stand for a total derivative, or a partial derivative with respect to x_1, x_2, x_3, or t, or a definite or indefinite (single or multiple) integral, then the operation of the operator on the n-tuple follows the same rule as the multiplication of an n-tuple by a scalar (A.9), thus

$$\diamondsuit\mathbf{r}(x_1,x_2,x_3,t)=[\diamondsuit r_1(x_1,x_2,x_3,t),\diamondsuit r_2(x_1,x_2,x_3,t),\ldots,\diamondsuit r_{\mathrm{n}}(x_1,x_2,x_3,t)]. \tag{A.34}$$

The following distributive laws connect matrix addition and operator operations:

$$\diamondsuit(\mathbf{r}+\mathbf{t})=\diamondsuit\mathbf{r}+\diamondsuit\mathbf{t}\text{ and }(\diamondsuit_1+\diamondsuit_2)\mathbf{r}=\diamondsuit_1\mathbf{r}+\diamondsuit_2\mathbf{r}, \tag{A.35}$$

where $\diamondsuit_1$ and $\diamondsuit_2$ are two different operators.

Problems

A.4.1. Find the derivative of the n-tuple $\mathbf{r}(x_1, x_2, x_3, t)=[x_1x_2x_3, 10x_1x_2, \cosh\alpha x_3]^{\mathrm{T}}$ with respect to x_3.

A.4.2. Find the symmetric and skew-symmetric parts of the matrix $\mathbf{r}\otimes\mathbf{s}$ where $\mathbf{r}=[1, 2, 3, 4]$ and $\mathbf{s}=[5, 6, 7, 8]$.

A.5 The Types andLinear Transformations

A system of linear equations

$$\begin{aligned}r_1&=A_{11}t_1+A_{12}t_2+\cdots+A_{1\mathrm{n}}t_{\mathrm{n}},\\r_2&=A_{21}t_1+A_{22}t_2+\cdots+A_{2\mathrm{n}}t_{\mathrm{n}},\\&\cdots\\r_{\mathrm{n}}&=A_{\mathrm{n}1}t_1+A_{\mathrm{n}2}t_2+\ldots+A_{\mathrm{nn}}t_{\mathrm{n}},\end{aligned} \tag{A.36}$$

may be contracted horizontally using the summation symbol, thus

$$\begin{aligned} r_1 &= A_{1k} t_k, \\ r_2 &= A_{2k} t_k, \\ &\cdots \\ r_n &= A_{nk} t_k. \end{aligned} \tag{A.37}$$

Introduction of the free index convention condenses this system of equations vertically,

$$r_i = A_{ik} t_k. \tag{A.38}$$

This result may also be represented in the matrix notation as a combination of n-tuples, **r** and **t**, and a square matrix **A**,

$$\mathbf{r} = \mathbf{A} \cdot \mathbf{t}, \tag{A.39}$$

where the dot between **A** and **t** indicates that the summation is with respect to one index of **A** and one index of **t**, or

$$\begin{bmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & . & . & . & A_{1n} \\ A_{21} & A_{22} & . & . & . & A_{2n} \\ . & . & . & . & . & . \\ A_{n1} & . & . & . & . & A_{nn} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ . \\ . \\ . \\ t_n \end{bmatrix}, \tag{A.40}$$

if the operation of the matrix **A** upon the column matrix **t** is interpreted as the operation of the square matrix upon the n-tuple defined by (A.38). This is an operation very similar to square matrix multiplication. This may be seen easily by rewriting the n-tuple in (A.40) as the first column of a square matrix whose entries are all otherwise zero; thus the operation is one of multiplication of one square matrix by another:

$$\begin{bmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & . & . & . & A_{1n} \\ A_{21} & A_{22} & . & . & . & A_{2n} \\ . & . & . & . & . & . \\ A_{n1} & . & . & . & . & A_{nn} \end{bmatrix} \begin{bmatrix} t_1 & 0 & . & . & . & 0 \\ t_2 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ t_n & . & . & . & . & 0 \end{bmatrix}. \tag{A.41}$$

The operation of the square matrix **A** on the n-tuple **t** is called a *linear transformation* of **t** into the n-tuple **r**. The linearity property is reflected in the

property that $\mathbf{A}$ applied to the sum $(\mathbf{r} + \mathbf{t})$ follows a distributive law $\mathbf{A} \cdot (\mathbf{r} + \mathbf{t}) = \mathbf{A} \cdot \mathbf{r} + \mathbf{A} \cdot \mathbf{t}$ and that multiplication by a scalar α follows the rule $\alpha(\mathbf{A} \cdot \mathbf{r}) = \mathbf{A} \cdot (\alpha\mathbf{r})$. These two properties may be combined into one, $\mathbf{A} \cdot (\alpha\mathbf{r} + \beta\mathbf{t}) = \alpha\mathbf{A} \cdot \mathbf{r} + \beta\mathbf{A} \cdot \mathbf{t}$ where α and β are scalars. The composition of linear transformations is again a linear transformation. Consider the linear transformation $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$, $\mathbf{u} \rightarrow \mathbf{t}$ (meaning $\mathbf{u}$ is transformed into $\mathbf{t}$) which is combined with the linear transformation (A.39) $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}, \mathbf{t} \rightarrow \mathbf{r}$ to transform $\mathbf{u} \rightarrow \mathbf{r}$, thus $\mathbf{r} = \mathbf{A}\cdot\mathbf{B}\cdot\mathbf{u}$, and if we let $\mathbf{C} \equiv \mathbf{A}\cdot\mathbf{B}$, then $\mathbf{r} = \mathbf{C}\cdot\mathbf{u}$. The result of the composition of the two linear transformations, $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ and $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$, is then a new linear transformation $\mathbf{r} = \mathbf{C}\cdot\mathbf{u}$ where the square matrix $\mathbf{C}$ is given by the matrix product $\mathbf{A}\cdot\mathbf{B}$. To verify that it is, in fact, a matrix multiplication, the composition of transformations is done again in the indicial notation. The transformation $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$ in the indicial notation,

$$t_{\mathrm{k}} = B_{\mathrm{km}} u_{\mathrm{m}}, \tag{A.42}$$

is substituted into $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ in the indicial notation (A.38),

$$r_{\mathrm{i}} = A_{\mathrm{ik}} B_{\mathrm{km}} u_{\mathrm{m}}, \tag{A.43}$$

which may be rewritten as

$$r_{\mathrm{i}} = C_{\mathrm{im}} u_{\mathrm{m}}, \tag{A.44}$$

where $\mathbf{C}$ is defined by:

$$C_{\mathrm{im}} = A_{\mathrm{ik}} B_{\mathrm{km}}. \tag{A.45}$$

Comparison of (A.45) with (A.20) shows that $\mathbf{C}$ is the matrix product of $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{C} = \mathbf{A}\cdot\mathbf{B}$. The calculation from (A.42) to (A.45) may be repeated using the Einstein summation convention. The calculation will be similar to the one above with the exception that the summation symbols will appear.

Example A.5.1

Determine the result $\mathbf{r} = \mathbf{C}\cdot\mathbf{u}$ of the composition of the two linear transformations, $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ and $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$, where $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 10 & 11 & 12 \\ 13 & 14 & 15 \\ 16 & 17 & 18 \end{bmatrix}.$$

Solution: The matrix product $\mathbf{A}\cdot\mathbf{B}$ yields the square matrix $\mathbf{C}$ representing the composed linear transformation,

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 84 & 90 & 96 \\ 201 & 216 & 231 \\ 318 & 342 & 366 \end{bmatrix}.$$

It is important to be able to construct the inverse of a linear transformation, $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$, $\mathbf{t} = \mathbf{A}^{-1}\cdot\mathbf{r}$, if it exists. The inverse transformation exists if $\mathbf{A}^{-1}$ can be constructed from $\mathbf{A}$, thus the question is one of the construction of the inverse of a square matrix. The construction of the inverse of a matrix involves the determinant of the matrix and the matrix of the cofactors. DetA denotes the determinant of A. A matrix is said to be *singular* if its determinant is zero, *non-singular* if it is not. The cofactor of the element A_{ij} of $\mathbf{A}$ is denoted by coA_{ij} and is equal to $(-1)^{i+j}$ times the determinant of a matrix constructed from the matrix $\mathbf{A}$ by deleting the row and column in which the element A_{ij} occurs. CoA denotes a matrix formed of the cofactors coA_{ij}.

Example A.5.2

Compute the matrix of cofactors of $\mathbf{A}$, $\mathbf{A} = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$.

Solution: The cofactors of the distinct elements of the matrix $\mathbf{A}$ are co$a = (bc{-}f^2)$, co$b = (ac{-}e^2)$, co$c = (ab{-}d^2)$, co$d = -(dc{-}fe)$, co$e = (df{-}eb)$, and co$f = -(af{-}de)$; thus the matrix of cofactors of $\mathbf{A}$ is

$$\mathbf{coA} = \begin{bmatrix} bc - f^2 & -(dc - fe) & (df - eb) \\ -(dc - fe) & ac - e^2 & -(af - de) \\ (df - eb) & -(af - de) & ab - d^2 \end{bmatrix}.$$

The formula for the inverse of $\mathbf{A}$ is written in terms of $\mathbf{coA}$ as

$$\mathbf{A}^{-1} = \frac{(\mathbf{coA})^{\mathrm{T}}}{\mathrm{Det}\mathbf{A}}, \tag{A.46}$$

where $(\mathbf{coA})^{\mathrm{T}}$ is the matrix of cofactors transposed. The inverse of a matrix is not defined if the matrix is singular. For every non-singular square matrix $\mathbf{A}$ the inverse of $\mathbf{A}$ can be constructed, thus

$$\mathbf{A}\cdot\mathbf{A}^{-1} = \mathbf{A}^{-1}\cdot\mathbf{A} = \mathbf{1}. \tag{A.47}$$

It follows then that the inverse of a linear transformation $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$, $\mathbf{t} = \mathbf{A}^{-1}\cdot\mathbf{r}$, exists if the matrix $\mathbf{A}$ is non-singular, Det$\mathbf{A} \neq 0$.

Example A.5.3

Show that the determinant of a 3-by-3-open product matrix, $\mathbf{a} \otimes \mathbf{b}$, is zero.

Solution:

$$\mathrm{Det}\{\mathbf{a}\otimes\mathbf{b}\} = \mathrm{Det}\begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix} = a_1b_1(a_2b_2a_3b_3 - a_2b_3a_3b_2)$$
$$- a_1b_2(a_2b_1a_3b_3 - a_3b_1a_2b_3) + a_1b_3(a_2b_1a_3b_2 - a_3b_1a_2b_2) = 0.$$

Example A.5.4

Find the inverse of the matrix

$$\mathbf{A} = \begin{bmatrix} 18 & 6 & 6 \\ 6 & 15 & 0 \\ 6 & 0 & 21 \end{bmatrix}$$

Solution: The matrix of cofactors is given by

$$\text{co}\mathbf{A} = \begin{bmatrix} 315 & -126 & -90 \\ -126 & 342 & 36 \\ -90 & 36 & 234 \end{bmatrix},$$

thus the inverse of $\mathbf{A}$ is then given by

$$\mathbf{A}^{-1} = \frac{\text{co}\mathbf{A}^{\text{T}}}{\text{Det}\mathbf{A}} = \frac{1}{243} \begin{bmatrix} 17.5 & -7 & -5 \\ -7 & 19 & 2 \\ -5 & 2 & 13 \end{bmatrix}.$$

The eigenvalue problem for a linear transformation $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ addresses the question of the n-tuple $\mathbf{t}$ being transformed by $\mathbf{A}$ into some scalar multiple of itself, $\lambda\mathbf{t}$. Specifically, for what values of $\mathbf{t}$ and λ does $\lambda\mathbf{t} = \mathbf{A}\cdot\mathbf{t}$? If such values of λ and $\mathbf{t}$ exist, they are called eigenvalues and eigen n-tuples of the matrix $\mathbf{A}$, respectively. The eigenvalue problem is then to find solutions to the equation

$$(\mathbf{A} - \lambda\mathbf{1}) \cdot \mathbf{t} = 0. \tag{A.48}$$

This is a system of linear equations for the elements of the n-tuple $\mathbf{t}$. For the case of $n = 3$ it may be written in the form:

$$\begin{aligned} (A_{11} - \lambda)t_1 + A_{12}t_2 + A_{13}t_3 &= 0, \\ A_{21}t_1 + (A_{22} - \lambda)t_2 + A_{23}t_3 &= 0, \\ A_{31}t_1 + A_{32}t_2 + (A_{33} - \lambda)t_3 &= 0. \end{aligned} \tag{A.49}$$

The standard approach to the solution of a system of linear equations like (A.48) is Cramer's rule. For a system of three equations in three unknowns, (A.36) with $n = 3$,

$$\begin{aligned} r_1 &= A_{11}t_1 + A_{12}t_2 + A_{13}t_3, \\ r_2 &= A_{21}t_1 + A_{22}t_2 + A_{23}t_3, \\ r_3 &= A_{31}t_1 + A_{32}t_2 + A_{33}t_3, \end{aligned} \tag{A.50}$$

Cramer's rule provides the solution for the n-tuple $\mathbf{t} = [t_1, t_2, t_3]$:

$$t_1 = \frac{\begin{vmatrix} r_1 & A_{12} & A_{13} \\ r_2 & A_{22} & A_{23} \\ r_3 & A_{32} & A_{33} \end{vmatrix}}{\mathrm{Det}A}, t_2 = \frac{\begin{vmatrix} A_{11} & r_1 & A_{13} \\ A_{21} & r_2 & A_{23} \\ A_{31} & r_3 & A_{33} \end{vmatrix}}{\mathrm{Det}A}, t_3 = \frac{\begin{vmatrix} A_{11} & A_{12} & r_1 \\ A_{21} & A_{22} & r_2 \\ A_{31} & A_{32} & r_3 \end{vmatrix}}{\mathrm{Det}A}. \tag{A.51}$$

Considering the case $n = 3$ and applying Cramer's rule to the system of equations (A.49) we find that

$$t_1 = \frac{\begin{vmatrix} 0 & A_{12} & A_{13} \\ 0 & A_{22} - \lambda & A_{23} \\ 0 & A_{32} & A_{33} - \lambda \end{vmatrix}}{\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}]}, t_2 = \frac{\begin{vmatrix} A_{11} - \lambda & 0 & A_{13} \\ A_{21} & 0 & A_{23} \\ A_{31} & 0 & A_{33} - \lambda \end{vmatrix}}{\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}]}, t_3$$
$$= \frac{\begin{vmatrix} A_{11} - \lambda & A_{12} & 0 \\ A_{21} & A_{22} - \lambda & 0 \\ A_{31} & A_{32} & 0 \end{vmatrix}}{\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}]}$$

which shows, due to the column of zeros in each numerator determinant, that the only solution is that $\mathbf{t} = [0, 0, 0]$, unless $\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}] = 0$. If $\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}] = 0$, the values of t_1, t_2, and t_3 are all of the form 0/0 and therefore undefined. In this case Cramer's rule provides no information. In order to avoid the trivial solution $\mathbf{t} = [0, 0, 0]$ the value of λ is selected so that $\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}] = 0$. While the argument was specialized to $n = 3$ in order to conserve page space, the result

$$\mathrm{Det}[\mathbf{A} - \lambda \mathbf{1}] = 0 \tag{A.52}$$

holds for all n. This condition forces the matrix $[\mathbf{A} - \lambda \mathbf{1}]$ to be singular and forces the system of equations (A.48) to be linearly dependent. The further solution of (A.52) is explored retaining the assumption of $n = 3$ for convenience, but it should noted that all the manipulations can be accomplished for any n including the values of n of interest here, 2, 3, and 6. In the case of $n = 3$, (A.52) is written in the form

$$\begin{vmatrix} A_{11} - \lambda & A_{12} & A_{13} \\ A_{21} & A_{22} - \lambda & A_{23} \\ A_{31} & A_{32} & A_{33} - \lambda \end{vmatrix} = 0, \tag{A.53}$$

and, when the determinant is expanded, one obtains a cubic equation for λ:

$$\lambda^3 - \mathrm{I}_{\mathbf{A}} \lambda^2 + \mathrm{II}_{\mathbf{A}} \lambda - \mathrm{III}_{\mathbf{A}} = 0 \tag{A.54}$$

where

$$\mathrm{I}_{\mathbf{A}} = \mathrm{tr}\mathbf{A} = \sum_{k=1}^{k=3} A_{\mathrm{kk}} = A_{\mathrm{kk}} = A_{11} + A_{22} + A_{33}, \tag{A.55}$$

$$\mathrm{II}_{\mathbf{A}} = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} + \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix}, \tag{A.56}$$

$$\mathrm{III}_{\mathbf{A}} = \mathrm{Det}\mathbf{A} = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix}. \tag{A.57}$$

This argument then generates a set of three λ's that allow the determinant (A.53) to vanish. We note again that the vanishing of the determinant makes the set of equations (A.49) linearly dependent. Since the system is linearly dependent, all of the components of $\mathbf{t}$ cannot be determined from (A.49). Thus, for each value of λ that is a solution to (A.54), we can find only two ratios of the elements of $\mathbf{t}$, t_1, t_2, and t_3. It follows that, for each eigen n-tuple, there will be one scalar unknown.

In this text we will only be interested in the eigenvalues of symmetric matrices. In Sect. A.7 it is shown that a necessary and sufficient condition for all the eigenvalues to be real is that that the matrix be symmetric.

Example A.5.5

Find the eigenvalues and construct the ratios of the eigen n-tuples of the matrix

$$\mathbf{A} = \begin{bmatrix} 18 & 6 & 6 \\ 6 & 15 & 0 \\ 6 & 0 & 21 \end{bmatrix}. \tag{A.58}$$

Solution: The cubic equation associated with this matrix is, from (A.54), (A.55), (A.56), and (A.57),

$$\lambda^3 - 54\lambda^2 + 891\lambda - 4,374 = 0, \tag{A.59}$$

which has three roots, 27, 18, and 9. The eigen n-tuples are constructed using these eigenvalues. The first eigen n-tuple is obtained by substitution of (A.58) and $\lambda = 27$ into (A.49), thus

$$-9t_1 + 6t_2 + 6t_3 = 0, 6t_1 - 12t_2 = 0, 6t_1 - 6t_3 = 0. \tag{A.60}$$

Note the linear dependence of this system of equations; the first equation is equal to the second multiplied by $(-1/2)$ and added to the third multiplied by (-1). Since there are only two independent equations, the solution to this system of equations is $t_1 = t_3$ and $t_1 = 2t_2$, leaving an undetermined parameter in the eigen n-tuple $\mathbf{t}$. Similar results are obtained by taking $\lambda = 18$ and $\lambda = 9$.

Problems

A.5.1. Show that the eigenvalues of the matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

are 11.345, 0.171, and −0.516.

A.5.2. Construct the inverse of the matrix **A** where $\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$.

A.5.3. Show that the inverse of the matrix **G** of problem A.5.1 is given by

$$\mathbf{G}^{-1} = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 3 & -1 \\ 2 & -1 & 0 \end{bmatrix}.$$

A.5.4. Show that the eigenvalues of the matrix $\mathbf{G}^{-1}$ of problem A.5.3 are the inverse of the eigenvalues of the matrix **G** of problem A.5.1.

A.5.5. Solve the matrix equation $\mathbf{A}^2 = \mathbf{A}\cdot\mathbf{A} = \mathbf{A}$ for **A** assuming that **A** is non-singular.

A.5.6. Why is it not possible to construct the inverse of an open product matrix, $\mathbf{a} \otimes \mathbf{b}$?

A.5.7. Construct a compositional transformation based on the matrix **G** of problem A.5.1 and the open product matrix, $\mathbf{a} \otimes \mathbf{b}$, where the n-tuples are $\mathbf{a} = [1, 2, 3]$ and $\mathbf{b} = [4, 5, 6]$.

A.5.8. If **F** is a square matrix and **a** is an n-tuple, show that $\mathbf{a}^{\mathrm{T}}\cdot\mathbf{F}^{\mathrm{T}} = \mathbf{F}\cdot\mathbf{a}$.

A.6 Vector Spaces

Loosely, vectors are defined as *n-tuples* that follow the *parallelogram law of addition*. More precisely vectors are defined as elements of a vector space called the *arithmetic n-space*. Let A^{n} denote the set of all n-tuples, $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_{\mathrm{N}}]$, $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_{\mathrm{N}}]$, etc., including the zero n-tuple, $\mathbf{0} = [0, 0, 0, \ldots, 0]$, and the negative *n* tuple $-\mathbf{u} = [-u_1, -u_2, -u_3, \ldots, -u_{\mathrm{N}}]$. An *arithmetic n-space* consists of the set A^{n} together with the additive and scalar multiplication operations defined by $\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots u_{\mathrm{N}} + v_{\mathrm{N}}]$ and $\alpha\mathbf{u} = [\alpha u_1, \alpha u_2, \alpha u_3, \ldots, \alpha u_{\mathrm{N}}]$, respectively. The additive operation defined by $\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots u_{\mathrm{N}} + v_{\mathrm{N}}]$ is the *parallelogram law of addition*. The parallelogram law of addition was first introduced and proved experimentally for forces. A vector is defined as an element of a vector space, in our case a particular vector space called the arithmetic *n*-space.

The scalar product of two vectors in *n* dimensions was defined earlier, (A.31). This definition provided a formula for calculating the scalar product $\mathbf{u}\cdot\mathbf{v}$ and the magnitude of the vectors **u** and **v**, $|\mathbf{u}| = \sqrt{\mathbf{u}\cdot\mathbf{u}}$ and $|\mathbf{v}| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$. Thus one can

consider the elementary definition of the scalar product below as the definition of the angle ζ,

$$\mathbf{u} \cdot \mathbf{v} = u_i v_i = |\mathbf{u}||\mathbf{v}| \cos \zeta \tag{A.61}$$

Recalling that there is a geometric interpretation of ζ as the angle between the two vectors **u** and **v** in two or three dimensions, it may seem strange to have the cosζ appear in the formula (A.61), which is valid in n dimensions. However, since **u**·**v** divided by $|\mathbf{u}| \cdot |\mathbf{v}|$ is always less than one and thus the definition (A.61) is reasonable not only for two and three dimensions, but for a space of any finite dimension. It is only in two and three dimensions that the angle ζ may be interpreted as the angle between the two vectors.

Example A.6.1

Show that the magnitude of the sum of the two unit vectors $\mathbf{e}_1 = [1,0]$ and $\mathbf{e}_2 = [\cos\alpha, \sin\alpha]$ can vary in magnitude from 0 to 2, depending on the value of the angle α.

Solution: $\mathbf{e}_1 + \mathbf{e}_2 = [1 + \cos\alpha, \sin\alpha]$, thus $|\mathbf{e}_1 + \mathbf{e}_2| = \sqrt{2}\sqrt{(1 + \cos\alpha)}$. It follows that $|\mathbf{e}_1 + \mathbf{e}_2| = 2$ when $\alpha = 0$, $|\mathbf{e}_1 + \mathbf{e}_2| = 0$ when $\alpha = \pi$, and $|\mathbf{e}_1 + \mathbf{e}_2| = \sqrt{2}$ when $\alpha = \pi/2$. Thus the sum of two unit vectors in two dimensions can point in any direction in the two dimensions and can have a magnitude between 0 and 2.

A set of unit vectors $\mathbf{e}_i$, $i = 1, 2, \ldots, n$, is called an *orthonormal or cartesian basis* of the vector space if all the base vectors are of unit magnitude and are orthogonal to each other, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ for i, j having the range n. From the definition of orthogonality one can see that, when $i \neq j$, the unit vectors $\mathbf{e}_i$ and $\mathbf{e}_j$ are orthogonal. In the case where $i = j$ the restriction reduces to the requirement that the $\mathbf{e}_i$'s be unit vectors. The elements of the n-tuples $v = [v_1, v_2, v_3, \ldots, v_n]$ referred to an orthonormal basis are called cartesian components. An important question concerning vectors is the manner in which their components change as their orthonormal basis is changed. In order to distinguish between the components referred to two different bases of a vector space we introduce two sets of indices. The first set of indices is composed of the lowercase Latin letters i, j, k, m, n, p, etc. which have the admissible values 1, 2, 3, ... n as before; the second set is composed of the lowercase Greek letters α, β, γ, δ, ... etc. whose set of admissible values are the Roman numerals I, II, III, ..., n. The Latin basis refers to the base vectors $\mathbf{e}_i$ while the Greek basis refers to the base vectors $\mathbf{e}_\alpha$. The components of a vector **v** referred to a Latin basis are then v_i, $i = 1, 2, 3, \ldots, n$, while the components of the same vector referred to a Greek basis are v_α, $\alpha = \text{I, II, III}, \ldots, n$. It should be clear that $\mathbf{e}_1$ is not the same as $\mathbf{e}_\text{I}$, v_2 is not the same as v_II, etc., that $\mathbf{e}_1$, v_2 refer to the Latin basis while $\mathbf{e}_\text{I}$, and v_II refers to the Greek basis. The terminology of calling a set of indices "Latin" and the other "Greek" is arbitrary; we could have introduced the second set of indices as i', j', k', m', n', p', etc., which would have had admissible values of $1', 2', 3', \ldots, n$, and subsequently spoken of the unprimed and primed sets of indices.

The range of the indices in the Greek and Latin sets must be the same since both sets of base vectors $\mathbf{e}_i$ and $\mathbf{e}_\alpha$ occupy the same space. It follows then that the two

sets, $\mathbf{e}_i$ and $\mathbf{e}_\alpha$, taken together are linearly dependent and therefore we can write that $\mathbf{e}_i$ is a linear combination of the $\mathbf{e}_\alpha$'s and vice versa. These relationships are expressed as linear transformations,

$$\mathbf{e}_i = Q_{i\alpha}\mathbf{e}_\alpha \text{ and } \mathbf{e}_\alpha = Q^{-1}_{i\alpha}\mathbf{e}_i, \tag{A.62}$$

where $\mathbf{Q} = [Q_{i\alpha}]$ is the matrix characterizing the linear transformation. For unambiguous conversion between the index and matrix notation the Latin index is fixed in the place of the row index (i.e., first) and the Greek index is frozen in the place of the column index (i.e., second) in the notation employed here. In the case of $n = 3$ the first of these equations may be expanded into a system of three equations:

$$\begin{aligned}
\mathbf{e}_1 &= Q_{1\mathrm{I}}\mathbf{e}_\mathrm{I} + Q_{1\mathrm{II}}\mathbf{e}_\mathrm{II} + Q_{1\mathrm{III}}\mathbf{e}_\mathrm{III},\\
\mathbf{e}_2 &= Q_{2\mathrm{I}}\mathbf{e}_\mathrm{I} + Q_{2\mathrm{II}}\mathbf{e}_\mathrm{II} + Q_{2\mathrm{III}}\mathbf{e}_\mathrm{III},\\
\mathbf{e}_3 &= Q_{3\mathrm{I}}\mathbf{e}_\mathrm{I} + Q_{3\mathrm{II}}\mathbf{e}_\mathrm{II} + Q_{3\mathrm{III}}\mathbf{e}_\mathrm{III}.
\end{aligned} \tag{A.63}$$

If one takes the scalar product of $\mathbf{e}_\mathrm{I}$ with each of these equations and notes that since the $\mathbf{e}_\alpha$, α = I, II, III, form an orthonormal basis, then $\mathbf{e}_\mathrm{I}\cdot\mathbf{e}_\mathrm{II} = \mathbf{e}_\mathrm{I}\cdot\mathbf{e}_\mathrm{III} = 0$, and $Q_{1\mathrm{I}} = \mathbf{e}_1\cdot\mathbf{e}_\mathrm{I} = \mathbf{e}_\mathrm{I}\cdot\mathbf{e}_1$, $Q_{2\mathrm{I}} = \mathbf{e}_2\cdot\mathbf{e}_\mathrm{I} = \mathbf{e}_\mathrm{I}\cdot\mathbf{e}_2$, and $Q_{3\mathrm{I}} = \mathbf{e}_3\cdot\mathbf{e}_\mathrm{I} = \mathbf{e}_\mathrm{I}\cdot\mathbf{e}_3$. Repeating the scalar product operation for $\mathbf{e}_{\mathrm{I}\,\mathrm{I}}$ and $\mathbf{e}_{\mathrm{I}\,\mathrm{II}}$ shows that, in general, $Q_{i\alpha} = \mathbf{e}_i\cdot\mathbf{e}_\alpha = \mathbf{e}_\alpha\cdot\mathbf{e}_i$. Recalling that the scalar product of two vectors is the product of magnitudes of each vector and the cosine of the angle between the two vectors (A.61), and that the base vectors are unit vectors, it follows that $Q_{i\alpha} = \mathbf{e}_i\cdot\mathbf{e}_\alpha = \mathbf{e}_\alpha\cdot\mathbf{e}_i$ are just the cosines of angles between the base vectors of the two bases involved. Thus the components of the linear transformation $\mathbf{Q} = [Q_{i\alpha}]$ are the cosines of the angles between the base vectors of the two bases involved. Because the definition of the scalar product (A.61) is valid in n dimensions, all these results are valid in n dimensions even though the two- and three-dimensional geometric interpretation of the components of the linear transformation $\mathbf{Q}$ as the cosines of the angles between coordinate axes is no longer valid.

The geometric analogy is very helpful, so considerations in three dimensions are continued. Three-dimensional Greek and Latin coordinate systems are illustrated on the left-hand side of Fig. A.1. The matrix $\mathbf{Q}$ with components $Q_{i\alpha} = \mathbf{e}_i\cdot\mathbf{e}_\alpha$ relates the components of vectors and base vectors associated with the Greek system to those associated with the Latin system,

$$\mathbf{Q} = [Q_{i\alpha}] = [\mathbf{e}_i\cdot\mathbf{e}_\alpha] = \begin{bmatrix} \mathbf{e}_1\cdot\mathbf{e}_\mathrm{I} & \mathbf{e}_1\cdot\mathbf{e}_\mathrm{II} & \mathbf{e}_1\cdot\mathbf{e}_\mathrm{III} \\ \mathbf{e}_2\cdot\mathbf{e}_\mathrm{I} & \mathbf{e}_2\cdot\mathbf{e}_\mathrm{II} & \mathbf{e}_2\cdot\mathbf{e}_\mathrm{III} \\ \mathbf{e}_3\cdot\mathbf{e}_\mathrm{I} & \mathbf{e}_3\cdot\mathbf{e}_\mathrm{II} & \mathbf{e}_3\cdot\mathbf{e}_\mathrm{III} \end{bmatrix}. \tag{A.64}$$

In the special case when the $\mathbf{e}_1$ and $\mathbf{e}_\mathrm{I}$ are coincident, the relative rotation between the two observers' frames is a rotation about that particular selected and fixed axis, and the matrix $\mathbf{Q}$ has the special form

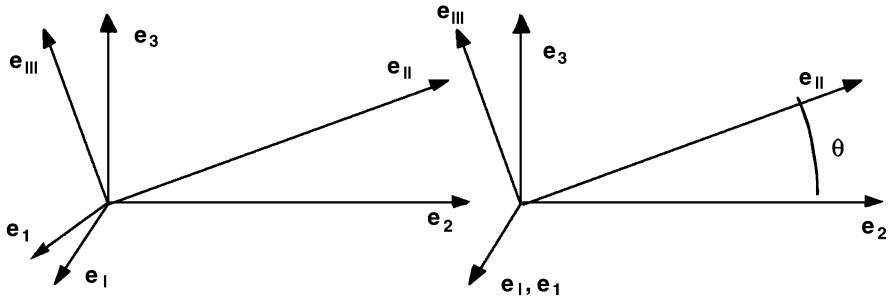


Fig. A.1 The relative rotational orientation between coordinate systems

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}. \tag{A.65}$$

This situation is illustrated on the left in Fig. A.1.

The matrix $\mathbf{Q} = [Q_{i\alpha}]$ characterizing the change from the Latin orthonormal basis $\mathbf{e}_\iota$ in an N-dimensional vector space to the Greek basis $\mathbf{e}_\alpha$ (or vice versa) is a special type of linear transformation called an orthogonal transformation. Taking the scalar product of $\mathbf{e}_i$ with $\mathbf{e}_j$ where $\mathbf{e}_i$ and $\mathbf{e}_j$ both have the representation (A.62),

$$\mathbf{e}_i = Q_{i\alpha}\mathbf{e}_\alpha \text{ and } \mathbf{e}_j = Q_{j\beta}\mathbf{e}_\beta, \tag{A.66}$$

it follows that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = Q_{i\alpha}Q_{j\beta}\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = Q_{i\alpha}Q_{j\beta}\delta_{\alpha\beta} = Q_{i\alpha}Q_{j\alpha}. \tag{A.67}$$

There are a number of steps in the calculation (A.67) that should be considered carefully. First, the condition of orthonormality of the bases has been used twice, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ and $\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = \delta_{\alpha\beta}$. Second, the transition from the term before the last equal sign to the term after that sign is characterized by a change from a double sum to a single sum over n and the loss of the Kronecker delta $\delta_{\alpha\beta}$. This occurs because the sum over β in the double sum is always zero except in the special case when $\alpha = \beta$ due to the presence of the Kronecker delta $\delta_{\alpha\beta}$. Third, a comparison of the last term in (A.67) with the definition of matrix product (A.20) suggests that it is a matrix product of $\mathbf{Q}$ with itself. However, a careful comparison of the last term in (A.67) with the definition of matrix product (A.20) shows that the summation is over a different index in the second element of the product. In order for the last term in (A.67) to represent a matrix product, the α index should appear as the first subscripted index rather than the second. However, this α index may be relocated in the second matrix by using the transposition operation. Thus the last term in equation (A.67) is the matrix product of $\mathbf{Q}$ with $\mathbf{Q}^T$ as may be seen from the first of

equations (A.18). Thus, since the matrix of Kronecker delta components is the unit matrix **1**, it has been shown that

$$\mathbf{1} = \mathbf{Q} \cdot \mathbf{Q}^{\mathrm{T}}. \tag{A.68}$$

If we repeat the calculation of the scalar product, this time using $\mathbf{e}_\alpha$ and $\mathbf{e}_\beta$ rather than $\mathbf{e}_\mathrm{i}$ and $\mathbf{e}_\mathrm{j}$, then it is found that $\mathbf{1} = \mathbf{Q}^{\mathrm{T}}\cdot\mathbf{Q}$ and, combined with the previous result,

$$\mathbf{1} = \mathbf{Q} \cdot \mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{Q}. \tag{A.69}$$

Using the fact that $\mathrm{Det}\mathbf{1} = 1$, and two results that are proved in Sect. A.8, $\mathrm{Det}\mathbf{A}\cdot\mathbf{B} = \mathrm{Det}\mathbf{A}\ \mathrm{Det}\mathbf{B}$, and $\mathrm{Det}\mathbf{A} = \mathrm{Det}\mathbf{A}^{\mathrm{T}}$, it follows from $\mathbf{1} = \mathbf{Q}\cdot\mathbf{Q}^{\mathrm{T}}$ or $\mathbf{1} = \mathbf{Q}^{\mathrm{T}}\cdot\mathbf{Q}$ that $\mathbf{Q}$ is non-singular and $\mathrm{Det}\mathbf{Q} = \pm 1$. Comparing the matrix equations $\mathbf{1} = \mathbf{Q}\cdot\mathbf{Q}^{\mathrm{T}}$ and $\mathbf{1} = \mathbf{Q}^{\mathrm{T}}\cdot\mathbf{Q}$ with the equations defining the inverse of $\mathbf{Q}$, $\mathbf{1} = \mathbf{Q}\cdot\mathbf{Q}^{-1} = \mathbf{Q}^{-1}\cdot\mathbf{Q}$, it follows that

$$\mathbf{Q}^{-1} = \mathbf{Q}^{\mathrm{T}}, \tag{A.70}$$

since the inverse exists ($\mathrm{Det}\mathbf{Q}$ is not singular) and is unique. Any matrix $\mathbf{Q}$ that satisfies equation (A.69) is called an orthogonal matrix. Any change of orthonormal bases is characterized by an orthogonal matrix and is called an orthogonal transformation. Finally, since $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathrm{T}}$ the representations of the transformation of bases (A.62) may be rewritten as

$$\mathbf{e}_\mathrm{i} = Q_{\mathrm{i}\alpha}\mathbf{e}_\alpha \text{ and } \mathbf{e}_\alpha = Q_{\mathrm{i}\alpha}\mathbf{e}_\mathrm{i}. \tag{A.71}$$

Orthogonal matrices are very interesting, useful and easy to handle; their determinant is always plus or minus one and their inverse is obtained simply by computing their transpose. Furthermore, the multiplication of orthogonal matrices has the closure property. To see that the product of two n by n orthogonal matrices is another n by n orthogonal matrix, let $\mathbf{R}$ and $\mathbf{Q}$ be orthogonal matrices and consider their product denoted by $\mathbf{W} = \mathbf{R}\cdot\mathbf{Q}$. The inverse of $\mathbf{W}$ is given by $\mathbf{W}^{-1} = \mathbf{Q}^{-1}\cdot\mathbf{R}^{-1}$ and its transpose by $\mathbf{W}^{\mathrm{T}} = \mathbf{Q}^{\mathrm{T}}\cdot\mathbf{R}^{\mathrm{T}}$. Since $\mathbf{R}$ and $\mathbf{Q}$ are orthogonal matrices, $\mathbf{Q}^{-1}\cdot\mathbf{R}^{-1} = \mathbf{Q}^{\mathrm{T}}\cdot\mathbf{R}^{\mathrm{T}}$, it follows that $\mathbf{W}^{-1} = \mathbf{W}^{\mathrm{T}}$ and therefore $\mathbf{W}$ is orthogonal. It follows then that the set of all orthogonal matrices has the closure property as well as the associative property with respect to the multiplication operation, an identity element (the unit matrix **1** is orthogonal), and an inverse for each member of the set.

Here we shall consider changing the basis to which a given vector is referred. While the vector $\mathbf{v}$ itself is invariant with respect to a change of basis, the components of $\mathbf{v}$ will change when the basis to which they are referred is changed. The components of a vector $\mathbf{v}$ referred to a Latin basis are then v_i, $i = 1, 2, 3, \ldots, n$, while the components of the same vector referred to a Greek basis are v_α, $\alpha = \mathrm{I}, \mathrm{II}, \mathrm{III}, \ldots, n$. Since the vector $\mathbf{v}$ is unique,

$$\mathbf{v} = v_\mathrm{i}\mathbf{e}_\mathrm{i} = v_\alpha\mathbf{e}_\alpha. \tag{A.72}$$

Substituting the second of (A.71) into the second equality of (A.72) one obtains

$$v_{\mathrm{i}}\mathbf{e}_{\mathrm{i}} = Q_{\mathrm{i}\alpha}v_{\alpha}\mathbf{e}_{\mathrm{i}}, \tag{A.73}$$

which may be rewritten as

$$(v_{\mathrm{i}} - Q_{\mathrm{i}\alpha}v_{\alpha})\mathbf{e}_{\mathrm{i}} = 0. \tag{A.74}$$

Taking the dot product of (A.74) with $\mathbf{e}_{\mathrm{j}}$, it follows that the sum over i is only nonzero when $i = j$, thus

$$v_{\mathrm{j}} = Q_{\mathrm{j}\alpha}v_{\alpha}. \tag{A.75}$$

If the first, rather than the second, of (A.71) is substituted into the second equality of (A.72), and similar algebraic manipulations accomplished, one obtains

$$v_{\mathrm{j}} = Q_{\mathrm{j}\alpha}v_{\alpha}. \tag{A.76}$$

The results (A.75) and (A.76) are written in the matrix notation using superscripted (L) and (G) to distinguish between components referred to the Latin or the Greek bases:

$$\mathbf{v}^{(\mathrm{L})} = \mathbf{Q}\cdot\mathbf{v}^{(\mathrm{G})}, \mathbf{v}^{(\mathrm{G})} = \mathbf{Q}^{\mathrm{T}}\cdot\mathbf{v}^{(\mathrm{L})}. \tag{A.77}$$

Problems

A.6.1. Is the matrix

$$\frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -1 & -2 & 2 \\ 2 & -2 & -1 \end{bmatrix}$$

an orthogonal matrix?.

A.6.2. Are the matrices **A**, **B**, **C**, and **Q**, where $\mathbf{Q} = \mathbf{C}\cdot\mathbf{B}\cdot\mathbf{A}$, and where

$$\mathbf{A} = \begin{bmatrix} \cos\Phi & \sin\Phi & 0 \\ -\sin\Phi & \cos\Phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \cos\Psi & 0 & -\sin\Psi \\ 0 & 1 & 0 \\ \sin\Psi & 0 & \cos\Psi \end{bmatrix}$$

all orthogonal matrices?

A.6.3. Does an inverse of the compositional transformation constructed in problem A.5.7 exist?

A.6.4. Is it possible for an open product of vectors to be an orthogonal matrix?

A.6.5. Transform the components of the vector $\mathbf{v}^{(\mathrm{L})} = [1, 2, 3]$ to a new (the Greek) coordinate system using the transformation

$$\mathbf{Q} = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{\frac{3}{2}} & \sqrt{\frac{3}{2}} \\ -\sqrt{3} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\sqrt{2} & \sqrt{2} \end{bmatrix}$$

A.7 Second Order Tensors

Scalars are tensors of order zero; vectors are tensors of order one. Tensors of order two will be defined using vectors. For brevity, we shall refer to "tensors of order two" simply as "tensors" throughout most of this section. For application in physical theories, physicists generated the notion of a tensor, very similar to the notion of a vector, but generalizing the vector concept. In classical dynamics the essential concepts of force, velocity, and acceleration are all vectors; hence the mathematical language of classical dynamics is that of vectors. In the mechanics of deformable media the essential concepts of stress, strain, rate of deformation, etc. are all second order tensors, thus, by analogy, one can expect to deal quite frequently with second order tensors in this branch of mechanics. The reason for this widespread use of tensors is that they enjoy, like vectors, the property of being invariant with respect to the basis, or frame of reference, chosen.

The definition of a tensor is motivated by a consideration of the open or dyadic product of the vectors $\mathbf{r}$ and $\mathbf{t}$. Recall that the square matrix formed from $\mathbf{r}$ and $\mathbf{t}$ is called the open product of the n-tuples $\mathbf{r}$ and $\mathbf{t}$, it is denoted by $\mathbf{r} \otimes \mathbf{t}$, and defined by (A.30) for n-tuples. We employ this same formula to define the open product of the vectors $\mathbf{r}$ and $\mathbf{t}$. Both of these vectors have representations relative to all bases in the vector space, in particular the Latin and the Greek bases, thus from (A.72)

$$\mathbf{r} = r_{\mathrm{i}}\mathbf{e}_{\mathrm{i}} = r_{\alpha}\mathbf{e}_{\alpha},\ \mathbf{t} = t_{\mathrm{j}}\mathbf{e}_{\mathrm{j}} = t_{\beta}\mathbf{e}_{\beta}. \tag{A.78}$$

The open product of the vectors $\mathbf{r}$ and $\mathbf{t}$, $\mathbf{r} \otimes \mathbf{t}$, then has the representation

$$\mathbf{r} \otimes \mathbf{t} = r_{\mathrm{i}}t_{\mathrm{j}}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}} = r_{\alpha}t_{\beta}\mathbf{e}_{\alpha} \otimes \mathbf{e}_{\beta}. \tag{A.79}$$

This is a special type of tensor, but it is referred to the general second order tensor basis, $\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}}$, or $\mathbf{e}_{\alpha} \otimes \mathbf{e}_{\beta}$. A general second order tensor is the quantity $\mathbf{T}$ defined by the formula relative to the bases $\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}}$, $\mathbf{e}_{\alpha} \otimes \mathbf{e}_{\beta}$ and, by implication, any basis in the vector space:

$$\mathbf{T} = T_{ij}\mathbf{e}_i \otimes \mathbf{e}_j = T_{\alpha\beta}\mathbf{e}_\alpha \otimes \mathbf{e}_\beta. \tag{A.80}$$

The formulas (A.78) and (A.80) have similar content in that the vectors **r** and **t** and the tensor **T** are quantities independent of base or coordinate system while the components of **r**, **t**, and **T** may be expressed relative to any basis. In the formulas (A.78) and (A.80), **r**, **t**, and **T** are expressed as components relative to two different bases. The vectors are expressed as components relative to the bases $\mathbf{e}_i$ and $\mathbf{e}_\alpha$, while the tensor **T** is expressed relative to the bases $\mathbf{e}_i \otimes \mathbf{e}_j$ and $\mathbf{e}_\alpha \otimes \mathbf{e}_\beta$. The tensor bases $\mathbf{e}_i \otimes \mathbf{e}_j$ and $\mathbf{e}_\alpha \otimes \mathbf{e}_\beta$ are constructed from the vector bases $\mathbf{e}_i$ and $\mathbf{e}_\alpha$.

Example A.7.1

If the base vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are expressed as $\mathbf{e}_1 = [1, 0, 0]^T$, $\mathbf{e}_2 = [0, 1, 0]^T$, and $\mathbf{e}_3 = [0, 0, 1]^T$, then it follows from (A.77) that $v = v_i e_i$ and we can express **v** in the form

$$v = v_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{A.81}$$

Create a similar representation for **T** given by (A.80) for $n = 3$.

Solution: The representation for **T** given by (A.80), $\mathbf{T} = T_{ij}\ \mathbf{e}_i \otimes \mathbf{e}_j$, involves the base vectors $\mathbf{e}_1 \otimes \mathbf{e}_1$, $\mathbf{e}_1 \otimes \mathbf{e}_2$ etc. These "base vectors" are expressed as matrices of tensor components by

$$e_1 \otimes e_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, e_1 \otimes e_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
$$e_2 \otimes e_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{etc.} \tag{A.82}$$

The representation for $\mathbf{T}$, $\mathbf{T} = T_{ij}\,\mathbf{e}_i \otimes \mathbf{e}_j$, then can be written in analogy to (A.81) as

$$T = T_{11}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + T_{21}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + T_{12}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + T_{13}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$+ T_{31}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + T_{22}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + T_{23}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
$$+ T_{32}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + T_{33}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \square$$

The components of the tensor $\mathbf{T}$ relative to the Latin basis, $\mathbf{T}^{(\mathrm{L})} = [T_{\mathrm{ij}}]$, are related to the components relative to the Greek basis, $\mathbf{T}^{(\mathrm{G})} = [T_{\alpha\beta}]$, by

$$\mathbf{T}^{(\mathrm{L})} = \mathbf{Q} \cdot \mathbf{T}^{(\mathrm{G})} \cdot \mathbf{Q}^{\mathrm{T}} \text{ and } \mathbf{T}^{(\mathrm{G})} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{T}^{(\mathrm{L})} \cdot \mathbf{Q}. \tag{A.83}$$

These formulas relating the components are the tensorial equivalent of vectorial formulas $\mathbf{v}^{(\mathrm{L})} = \mathbf{Q} \cdot \mathbf{v}^{(\mathrm{G})}$ and $\mathbf{v}^{(\mathrm{G})} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{v}^{(\mathrm{L})}$ given by (A.77), and their derivation is similar. First, substitute the second of (A.66) into the (A.80) twice, once for each base vector:

$$\mathbf{T} = T_{\mathrm{ij}}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}} = T_{\alpha\beta}Q_{\mathrm{i}\alpha}Q_{\mathrm{j}\beta}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}}. \tag{A.84}$$

Then gather together the terms referred to the basis $\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}}$, thus

$$(T_{\mathrm{ij}} - T_{\alpha\beta}Q_{\mathrm{i}\alpha}Q_{\mathrm{j}\beta})\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}} = 0. \tag{A.85}$$

Next take the scalar product of (A.85), first with respect to $\mathbf{e}_{\mathrm{k}}$, and then with respect to $\mathbf{e}_{\mathrm{m}}$. One finds that the only nonzero terms that remain are

$$T_{\mathrm{km}} = Q_{\mathrm{k}\alpha}T_{\alpha\beta}Q_{\mathrm{m}\beta}. \tag{A.86}$$

A comparison of the last term in (A.86) with the definition of matrix product (A.20) suggests that it is a triple matrix product involving $\mathbf{Q}$ twice and $\mathbf{T}^{(\mathrm{G})}$ once. Careful comparison of the last term in (A.86) with the definition of matrix product (A.20) shows that the summation is over a different index in the third element of the product. In order for the last term in (A.86) to represent a triple matrix product, the β index should appear as the first subscripted index rather than the second. However, this β index may be relocated in the second matrix by using the transposition operation as shown in the first equation of (A.21). Thus the last term in equation (A.86) is the matrix product of $\mathbf{Q}\cdot\mathbf{T}$ with $\mathbf{Q}^{\mathbf{T}}$. The result is the first equation of (A.83). If the first, rather than the second, of (A.67) is substituted into the second equality of (A.80), and similar algebraic manipulations accomplished, one obtains the second equation of (A.83).

The word tensor is used to refer to the quantity $\mathbf{T}$ defined by (A.80), a quantity independent of any basis. It is also used to refer to the matrix of tensor components relative to a particular basis, for example $\mathbf{T}^{(\mathrm{L})} = [T_{\mathrm{ij}}]$ or $\mathbf{T}^{(\mathrm{G})} = [T_{\alpha\beta}]$. In both cases "tensor" should be "tensor of order two," but the order of the tensor is generally clear from the context. A tensor of order N in a space of n dimensions is defined by

$$\mathbf{B} = B_{\mathrm{ij}\ldots\mathrm{k}}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}} \otimes \ldots \otimes \mathbf{e}_{\mathrm{k}} = B_{\alpha\beta\ldots\gamma}\mathbf{e}_{\alpha} \otimes \mathbf{e}_{\beta} \otimes \ldots \otimes \mathbf{e}_{\gamma}. \tag{A.87}$$

The number of base vectors in the basis is the order N of the tensor. It is easy to see that this definition specializes to that of the second order tensor (A.80). The definition of a vector as a tensor of order one is easy to see, and the definition of a scalar as a tensor of order 0 is trivial.

In the section before last, Sect. A.5 on Linear Transformations, the eigenvalue problem for a linear transformation $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ was considered. Here we extend those results by considering $\mathbf{r}$ and $\mathbf{t}$ to be vectors and $\mathbf{A}$ to be a symmetric second order tensor, $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$. The problem is actually little changed until its conclusion. The eigenvalues are still given by (A.52) or, for $n = 3$ by (A.54). The values of three quantities $I_{\mathbf{A}}$, $II_{\mathbf{A}}$, $III_{\mathbf{A}}$, defined by (A.55), (A.56), (A.57) are the same except that $A_{12} = A_{21}, A_{13} = A_{31}$ and $A_{32} = A_{23}$ due to the assumed symmetry of $\mathbf{A}$, $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$. These quantities may now be called the invariants of the tensor $\mathbf{A}$ since their value is the same independent of the coordinate system chosen for their determination. As an example of the invariance with respect to basis, this property will be derived for $I_{\mathbf{A}} = \mathrm{tr}\,\mathbf{A}$. Let $\mathbf{T} = \mathbf{A}$ in (A.86), then set the indices $k = m$ and sum from one to n over the index k, thus

$$A_{\mathrm{kk}} = T_{\alpha\beta}Q_{\mathrm{k}\alpha}Q_{\mathrm{k}\beta} = A_{\alpha\beta}\delta_{\alpha\beta} = A_{\alpha\alpha} \tag{A.88}$$

The transition across the second equal sign is a simple rearrangement of terms. The transition across the second equal sign is based on the condition

$$Q_{\mathrm{k}\alpha}Q_{\mathrm{k}\beta} = \delta_{\alpha\beta} \tag{A.89}$$

which is an alternate form of (A.67), a form equivalent to $\mathbf{Q}^{\mathrm{T}}\cdot\mathbf{Q} = \mathbf{1}$. The transition across the fourth equal sign employs the definition of the Kronecker delta and the summation over β. The result is that the trace of the matrix of second order tensor components relative to any basis is the same number,

$$A_{\mathrm{kk}} = A_{\alpha\alpha}. \tag{A.90}$$

It may also be shown that $II_{\mathbf{A}}$ and $III_{\mathbf{A}}$ are invariants of the tensor $\mathbf{A}$.

Example A.7.2 (An Extension of Example A.5.5)

Consider the matrix given by (A.58) in Example A.5.5 to be the components of a tensor. Construct the eigenvectors of that tensor and use those eigenvectors to construct an eigenbasis

$$\mathbf{A} = \begin{bmatrix} 18 & 6 & 6 \\ 6 & 15 & 0 \\ 6 & 0 & 21 \end{bmatrix}. \qquad (A.58) \text{ repeated}$$

Solution: The eigenvalues were shown to be 27, 18, and 9. It can be shown that the eigenvalues must always be real numbers if $\mathbf{A}$ is symmetric. Eigen n-tuples were constructed using these eigenvalues. The first eigen n-tuple was obtained by substitution of (A.58) and $\lambda = 27$ into (A.49), thus

$$-9t_1 + 6t_2 + 6t_3 = 0,\ 6t_1 - 12t_2 = 0,\ 6t_1 - 6t_3 = 0. \qquad \text{(A.60) repeated}$$

These three conditions, only two of which are independent, gave $t_1 = t_3$ and $t_1 = 2t_2$, leaving an undetermined parameter in the eigen n-tuple $\mathbf{t}$. Now that $\mathbf{t}$ is a vector, we can specify the length of a vector. Another consequence of the symmetry of $\mathbf{A}$ is that these eigenvectors are orthogonal if the eigenvalues are distinct. Hence, if we set the length of the eigenvectors to be one to remove the undetermined parameter, we will generate an orthonormal basis from the set of three eigenvectors, since the eigenvalues are distinct. If we use the normality condition $t_1^2 + t_2^2 + t_3^2 = 1$ and the results that follow from (A.56), $t_1 = t_3$ and $t_1 = 2t_2$, one finds that

$$\mathbf{t} = \pm\left(\frac{1}{3}\right)(2\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3) \tag{A.91}$$

which shows that both $\mathbf{t}$ and $-\mathbf{t}$ are eigenvectors. This will be true for any eigenvector because they are really eigen-directions. For the second and third eigenvalues, 18 and 9, we find that

$$\mathbf{t} = \pm\left(\frac{1}{3}\right)(-\mathbf{e}_1 - 2\mathbf{e}_2 + 2\mathbf{e}_3) \text{ and } \mathbf{t} = \pm\left(\frac{1}{3}\right)(2\mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3), \tag{A.92}$$

respectively. It is easy to see that these three eigenvectors are mutually orthogonal.

It was noted above that, since the eigenvectors constitute a set of three mutually perpendicular unit vectors in a three-dimensional space, they might be used to form a basis or a coordinate reference frame. Let the three orthogonal eigenvectors be the base vectors $\mathbf{e}_{\mathrm{I}}$, $\mathbf{e}_{\mathrm{II}}$, and $\mathbf{e}_{\mathrm{III}}$ of a Greek reference frame. From (A.91) and (A.92) we form a new reference basis for the example eigenvalue problem, thus

$$\begin{aligned} \mathbf{e}_{\mathrm{I}} &= \left(\frac{1}{3}\right)(2\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3), \mathbf{e}_{\mathrm{II}} = \left(\frac{1}{3}\right)(-\mathbf{e}_1 - 2\mathbf{e}_2 + 2\mathbf{e}_3), \\ \mathbf{e}_{\mathrm{III}} &= \left(\frac{1}{3}\right)(2\mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3). \end{aligned} \tag{A.93}$$

It is easy to verify that both the Greek and Latin base vectors form right-handed orthonormal systems. The orthogonal matrix $\mathbf{Q}$ for transformation from the Latin to the Greek system is given by (A.64) and (A.93) as

$$\mathbf{Q} = [Q_{i\alpha}] = \left(\frac{1}{3}\right)\begin{bmatrix} 2 & -1 & 2 \\ 1 & -2 & -2 \\ 2 & 2 & -1 \end{bmatrix}. \tag{A.94}$$

Substituting the $\mathbf{Q}$ of (A.94) and the $\mathbf{A}$ specified by (A.56) into the second of (A.83) with $\mathbf{T} = \mathbf{A}$,

$$\mathbf{A}^{(\mathrm{G})} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{A}^{(\mathrm{L})} \cdot \mathbf{Q}, \tag{A.95}$$

the following result is determined:

$$\mathbf{A}^{(\mathrm{G})} = \left(\frac{1}{9}\right)\begin{bmatrix} 2 & 1 & 2 \\ -1 & -2 & 2 \\ 2 & -2 & -1 \end{bmatrix}\begin{bmatrix} 18 & 6 & 6 \\ 6 & 15 & 0 \\ 6 & 0 & 21 \end{bmatrix}\begin{bmatrix} 2 & -1 & 2 \\ 1 & -2 & -2 \\ 2 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 27 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 9 \end{bmatrix}. \tag{A.96}$$

Thus, relative to the basis formed of its eigenvectors a symmetric matrix takes on a diagonal form, the diagonal elements being its eigenvalues. This result, which was demonstrated for a particular case, is true in general in a space of any dimension n as long as the matrix is symmetric. □

There are two points in the above example that are always true if the matrix is symmetric. The first is that the eigenvalues are always real numbers and the second is that the eigenvectors are always mutually perpendicular. These points will now be proved in the order stated. To prove that λ is always real we shall assume that it could be complex, and then we show that the imaginary part is zero. This proves that λ is real. If λ is complex, say $\mu + iv$, the associated eigenvector $\mathbf{t}$ must also be complex and we denote it by $\mathbf{t} = \mathbf{p} + i\mathbf{q}$. With these notations (A.48) can be written

$$(\mathbf{A} - \{\mu + iv\}\mathbf{1}) \cdot (\mathbf{p} + i\mathbf{q}) = 0. \tag{A.97}$$

Equating the real and imaginary parts, we obtain two equations,

$$\mathbf{A} \cdot \mathbf{p} = \mu\mathbf{p} - v\mathbf{q}, \mathbf{A} \cdot \mathbf{q} = v\mathbf{p} + \mu\mathbf{q}. \tag{A.98}$$

The symmetry of the matrix $\mathbf{A}$ means that, for the vectors $\mathbf{p}$ and $\mathbf{q}$,

$$\mathbf{p} \cdot \mathbf{A} \cdot \mathbf{q} - \mathbf{q} \cdot \ \mathbf{A} \cdot \mathbf{p} = 0 = -v(\mathbf{p} \cdot \mathbf{p} + \mathbf{q} \cdot \mathbf{q}), \tag{A.99}$$

the last equality following from taking the scalar product of the two equations in (A.98), the first with respect to $\mathbf{q}$ and the second with respect to $\mathbf{p}$. There is only one-way to satisfy $-v(\mathbf{p}\cdot\mathbf{p} + \mathbf{q}\cdot\mathbf{q})$, since the eigenvector cannot be zero and that is to take $v = 0$, hence λ is real. This result also shows that $\mathbf{t}$ must also be real.

We will now show that any two eigenvectors are orthogonal if the two associated eigenvalues are distinct. Let λ_1 and λ_2 be the eigenvalues associated with the eigenvectors $\mathbf{n}$ and $\mathbf{m}$, respectively, then

$$\mathbf{A} \cdot \mathbf{n} = \lambda_1\mathbf{n} \text{ and } \mathbf{A} \cdot \mathbf{m} = \lambda_2\mathbf{m}. \tag{A.100}$$

Substituting the two equations (A.100) into the first and last equalities of (A.99) we find that

$$(\lambda_1 - \lambda_2)\mathbf{n} \cdot \mathbf{m} = 0. \tag{A.101}$$

Thus, if $\lambda_1 \neq \lambda_2$, then **n** and **m** are perpendicular. If the two eigenvalues are not distinct, then any vector in a plane is an eigenvector so that one can always construct a mutually orthogonal set of eigenvectors for a symmetric matrix.

Generalizing Example A.7.2 above from 3 to n it may be concluded that any n by n matrix of symmetric tensor components **A** has a representation in which the eigenvalues lie along the diagonal of the matrix and the off-diagonal elements are all zero. The last expression in (A.96) is a particular example of this when $n = 3$. If the symmetric tensor **A** has n eigenvalues λ_{i}, then a quadratic form ψ may be formed from **A** and a vector n-tuple **x**, thus

$$\psi = \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} = \lambda_{\mathrm{i}} (x_{\mathrm{i}})^2. \tag{A.102}$$

If all the eigenvalues of **A** are positive, this quadratic form is said to be *positive definite* and

$$\psi = \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} = \lambda_{\mathrm{i}} (x_{\mathrm{i}})^2 > 0 \text{ for all } \mathbf{x} \neq \mathbf{0}. \tag{A.103}$$

(If all the eigenvalues of **A** are negative the quadratic form is said to be *negative definite*.) Transforming the tensor **A** to an arbitrary coordinate system the equation (A.102) takes the form

$$\psi = \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} = A_{\mathrm{ij}} x_{\mathrm{i}} x_{\mathrm{i}} > 0 \text{ for all } \mathbf{x} \neq \mathbf{0}. \tag{A.104}$$

A tensor **A** with the property (A.104), when it is used as the coefficients of a quadratic form, is said to be positive definite. In the mechanics of materials there are a number of tensors that are positive definite due to physics they represent. The moment of inertia tensor is an example. Others will be encountered as material coefficients in constitutive equations in Chap. 5.

Problems

A.7.1 Consider two three-dimensional coordinate systems. One coordinate system is a right-handed coordinate system. The second coordinate system is obtained from the first by reversing the direction of the first ordered base vector and leaving the other two base vectors to be identical with those in the first coordinate system. Show that the orthogonal transformation relating these systems is given by

$$\mathbf{Q} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and that its determinant is -1.

A.7.2 Construct the eigenvalues and the eigenvectors of the matrix **T** of tensor components where

$$\mathbf{T} = \frac{1}{2}\begin{bmatrix} 13 & 3\sqrt{3} & \sqrt{3} \\ 3\sqrt{3} & 7 & 1 \\ \sqrt{3} & 1 & 8 \end{bmatrix}.$$

A.7.3 Construct the eigenvalues and the eigenvectors of the matrix **A** of tensor components where

$$\mathbf{A} = \frac{1}{4}\begin{bmatrix} 17 & 3\sqrt{3} \\ 3\sqrt{3} & 11 \end{bmatrix}.$$

A.7.4 Construct the eigenvalues and the eigenvectors of the matrix **A** of tensor components where

$$\mathbf{A} = \frac{1}{3}\begin{bmatrix} 13 & 2 & 4 \\ 2 & 10 & 2 \\ 4 & 2 & 13 \end{bmatrix}.$$

A.7.5 Construct the orthogonal transformation that carries the matrix **B** of Problem A.7.3 into its diagonal form.

A.7.6 Construct the eigenvalues and the eigenvectors of the matrix **A** of tensor components where

$$\mathbf{A} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

A.7.7 Construct the orthogonal transformation that carries the matrix **A** of Problem A.7.5 into its diagonal form.

A.7.8 Construct the eigenvalues and the eigenvectors of the matrix **B** of tensor components where

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

What is the angle between the two eigenvectors?

A.7.9 Show that the eigenvalues of the matrix **H** where

$$\mathbf{H} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 7 & 8 & 9 & 10 & 11 \\ 3 & 8 & 12 & 13 & 14 & 15 \\ 4 & 9 & 13 & 16 & 17 & 18 \\ 5 & 10 & 14 & 17 & 19 & 20 \\ 6 & 11 & 15 & 18 & 20 & 21 \end{bmatrix}$$

are 73.227, 2.018. 1.284, 0.602, 0.162, and −1.294.

A.7.10 Consider the components of the tensor **T** given in problem A.7.2 to be relative to a (Latin) coordinate system and denote them by $\mathbf{T}^{(\mathrm{L})}$. Transform these components to a new coordinate system (the Greek) using the transformation

$$\mathbf{Q} = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{\frac{3}{2}} & \sqrt{\frac{3}{2}} \\ -\sqrt{3} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\sqrt{2} & \sqrt{2} \end{bmatrix}.$$

A.7.11 Show that if a tensor is symmetric (skew-symmetric) in one coordinate system, then it is symmetric (skew-symmetric) in all coordinate systems. Specifically show that if $\mathbf{A}^{(\mathrm{L})} = (\mathbf{A}^{(\mathrm{L})})^{\mathrm{T}}$, then $\mathbf{A}^{(\mathrm{G})} = (\mathbf{A}^{(\mathrm{G})})^{\mathrm{T}}$.

A.8 The Alternator and Vector Cross Products

There is a strong emphasis on the indicial notation in this section. It is advised that the definitions (in Sect. A.3) of free indices and summation indices be reviewed carefully if one is not altogether comfortable with the indicial notation. It would also be beneficial to redo some indicial notation problems.

The alternator in three dimensions is a three index numerical symbol that encodes the permutations that one is taught to use expanding a determinant. Recall the process of evaluating the determinant of the 3 by 3 matrix **A**,

$$\begin{aligned} \mathrm{Det}\mathbf{A} = \mathrm{Det}\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} &= \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \\ = A_{11}A_{22}A_{33} - A_{11}A_{32}A_{23} - A_{12}A_{21}A_{33} + A_{12}A_{31}A_{23} + A_{13}A_{21}A_{32} - A_{13}A_{31}A_{22}. \end{aligned} \tag{A.105}$$

The permutations that one is taught to use expanding a determinant are permutations of a set of three objects. The alternator is denoted by e_{ijk} and defined so that it takes on values +1, 0, or −1 according to the rule:

$$e_{\mathrm{ijk}} \equiv \left\{ \begin{array}{ll} +1 & \text{if } P \text{ is an even permuation} \\ 0 & \text{otherwise} \\ -1 & \text{if } P \text{ is an odd permuation} \end{array} \right\}, P \equiv \left\{ \begin{array}{ccc} 1 & 2 & 3 \\ i & j & k \end{array} \right\} \tag{A.106}$$

where P is the permutation symbol on a set of three objects. The only +1 values of e_{ijk} are e_{123}, e_{231}, and e_{312}. It is easy to verify that 123, 231, and 312 are even

permutations of 123. The only -1 values of e_{ijk} are e_{132}, e_{321}, and e_{213}. It is easy to verify that 132, 321, and 213 are odd permutations of 123. The other 21 components of e_{ijk} are all zero because they are neither even nor odd permutations of 123 due to the fact that one number (either 1, 2, or 3) occurs more than once in the indices (e.g., $e_{122} = 0$ since 122 is not a permutation of 123). One mnemonic device for the even permutations of 123 is to write 123123, then read the first set of three digits 123, the second set 231, and the third set 312. The odd permutations may be read off 123123 also by reading from right to left rather than from left to right; reading from the right (but recording them then from the left, as usual) the first set of three digits 321, the second set 213, and the third set 132.

The alternator may now be employed to shorten the formula (A.105) for calculating the determinant;

$$e_{mnp}\mathrm{Det}\mathbf{A} = e_{ijk}A_{im}A_{jn}A_{kp} = e_{ijk}A_{mi}A_{nj}A_{pk}. \tag{A.107}$$

This result may be verified by selecting the values of *mnp* to be 123, 231, 312, 132, 321, or 213, then performing the summation over the three indices *i*, *j*, and *k* over 1, 2, and 3 as indicated on the right-hand side of (A.107). In each case the result is the right-hand side of (A.105). It should be noted that (A.107) may be used to show $\mathrm{Det}\mathbf{A} = \mathrm{Det}\mathbf{A}^{\mathrm{T}}$.

The alternator may be used to express the fact that interchanging two rows or two columns of a determinant changes the sign of the determinant,

$$e_{mnp}\mathrm{Det}\mathbf{A} = \begin{vmatrix} A_{1m} & A_{1n} & A_{1p} \\ A_{2m} & A_{2n} & A_{2p} \\ A_{3m} & A_{3n} & A_{3p} \end{vmatrix} = \begin{vmatrix} A_{m1} & A_{m2} & A_{m3} \\ A_{n1} & A_{n2} & A_{n3} \\ A_{p1} & A_{p2} & A_{p3} \end{vmatrix}. \tag{A.108}$$

Using the alternator again may combine these two representations:

$$e_{ijk}e_{mnp}\mathrm{Det}\mathbf{A} = \begin{vmatrix} A_{im} & A_{in} & A_{ip} \\ A_{jm} & A_{jn} & A_{jp} \\ A_{km} & A_{kn} & A_{kp} \end{vmatrix}. \tag{A.109}$$

In the special case when $\mathbf{A} = \mathbf{1}$ ($A_{ij} = \delta_{ij}$), an important identity relating the alternator to the Kronecker delta is obtained:

$$e_{ijk}e_{mnp} = \begin{vmatrix} \delta_{im} & \delta_{in} & \delta_{ip} \\ \delta_{jm} & \delta_{jn} & \delta_{jp} \\ \delta_{km} & \delta_{kn} & \delta_{kp} \end{vmatrix}. \tag{A.110}$$

The following special cases of (A.110) provide three more very useful relations between the alternator and the Kronecker delta:

$$e_{mnk}e_{ijk} = \delta_{im}\delta_{jn} - \delta_{jm}\delta_{in}, e_{mjk}e_{ijk} = 2\delta_{im}, e_{ijk}e_{ijk} = 6. \tag{A.111}$$

The first of these relations is obtained by setting the indices p and k equal in (A.110) and then expanding the determinant. The second is obtained from the first by setting the indices n and j equal in first. The third is obtained from the second by setting the indices i and m equal in second.

Example A.8.1

Derive the first of (A.111) from (A.110).
Solution: The first of (A.111) is obtained from (A.110) by setting the indices p and k equal in (A.110) and then expanding the determinant:

$$e_{ijk}e_{mnk} = \begin{vmatrix} \delta_{im} & \delta_{in} & \delta_{ik} \\ \delta_{jm} & \delta_{jn} & \delta_{jk} \\ \delta_{km} & \delta_{kn} & 3 \end{vmatrix},$$

one finds that

$$e_{ijk}e_{mnk} = 3\delta_{im}\delta_{jn} - \delta_{im}\delta_{jk}\delta_{kn} - 3\delta_{in}\delta_{jm} + \delta_{in}\delta_{km}\delta_{jk} + \delta_{ik}\delta_{jm}\delta_{kn} - \delta_{ik}\delta_{km}\delta_{jn}.$$

Carrying out the indicated summation over the index k in the expression above,

$$e_{ijk}e_{mnk} = 3\delta_{im}\delta_{jn} - \delta_{im}\delta_{jn} - 3\delta_{in}\delta_{jm} + \delta_{in}\delta_{jm} + \delta_{in}\delta_{jm} - \delta_{im}\delta_{jn}.$$

This is the desired result, the first of (A.111). □

Example A.8.2

Prove that $\mathrm{Det}(\mathbf{A}\cdot\mathbf{B}) = \mathrm{Det}\mathbf{A}\ \mathrm{Det}\mathbf{B}$.

Solution: Replacing $\mathbf{A}$ in (A.107) by $\mathbf{C}$ and selecting the values of mnp to be 123, then (A.107) becomes

$$\mathrm{Det}\mathbf{C} = e_{ijk}C_{i1}C_{j2}C_{k3} = e_{ijk}C_{1i}C_{2j}C_{3k}.$$

Now $\mathbf{C}$ is replaced by the product $\mathbf{A}\cdot\mathbf{B}$ using

$$C_{i1} = A_{im}B_{m1}, C_{j2} = A_{jn}B_{n2}, C_{k3} = A_{kp}B_{p3},$$

thus

$$\mathrm{Det}\mathbf{A}\cdot\mathbf{B} = e_{ijk}A_{im}B_{m1}A_{jn}B_{n2}A_{kp}B_{p3}, \text{ or } \mathrm{Det}\mathbf{A}\cdot\mathbf{B} = (e_{ijk}A_{im}A_{jn}A_{kp})B_{m1}B_{n2}B_{p3},$$

where the order of the terms in the second sum has been rearranged from the first. Comparison of the first four rearranged terms from the second sum with the right-hand side of (A.107) shows that the first four terms in the sum on the right may be replaced by $e_{mnp}\mathrm{Det}\mathbf{A}$; thus applying the first equation of this solution again with $\mathbf{C}$ replaced by $\mathbf{B}$, the desired result is obtained:

$$\mathrm{Det}\mathbf{A}\cdot\mathbf{B} = \mathrm{Det}\mathbf{A}e_{\mathrm{mnp}}B_{\mathrm{m1}}B_{\mathrm{n2}}B_{\mathrm{p3}} = \mathrm{Det}\mathbf{A}\,\mathrm{Det}\mathbf{B}. \qquad \square$$

Consider now the question of the tensorial character of the alternator. Vectors were shown to be characterized by symbols with one subscript, (second order) tensors were shown to characterized by symbols with two subscripts; what is the tensorial character of a symbol with three subscripts; is it a third order tensor? Almost. Tensors are identified on the basis of their tensor transformation law. Recall the tensor transformations laws (A.75) and (A.76) for a vector, (A.86) for a second order tensor, and (A.87) for a tensor of order n. An equation that contains a transformation law for the alternator is obtained from (A.107) by replacing $\mathbf{A}$ by the orthogonal transformation $\mathbf{Q}$ given by (A.64) and changing the indices as follows: $m \to \alpha$, $n \to \beta$, $p \to \gamma$, thus

$$e_{\alpha\beta\gamma}\mathrm{Det}\mathbf{Q} = e_{\mathrm{ijk}}Q_{\mathrm{i}\alpha}Q_{\mathrm{j}\beta}Q_{\mathrm{k}\gamma}. \tag{A.112}$$

This is an unusual transformation law because the determinant of an orthogonal transformation $\mathbf{Q}$ is either +1 or −1. The expected transformation law, on the basis of the tensor transformations laws (A.75) and (A.76) for a vector, (A.86) for a second order tensor and (A.87) for a tensor of order n, is that $\mathrm{Det}\mathbf{Q} = +1$. $\mathrm{Det}\mathbf{Q} = +1$ occurs when the transformation is between coordinate systems of the same handedness (right-handed to right-handed or left-handed to left-handed). Recall that a right (left) hand coordinate system or orthonormal basis is one that obeys the right (left) hand rule, that is to say if the curl of your fingers in your right (left) hand fist is in the direction of rotation from the first ordered positive base vector into the second ordered positive base vector, your extended thumb will point in the third ordered positive base vector direction. $\mathrm{Det}\mathbf{Q} = -1$ occurs when the transformation is between coordinate systems of the opposite handedness (left to right or right to left). Since handedness does not play a role in the transformation law for even order tensors, this dependence on the sign of $\mathrm{Det}\mathbf{Q}$ and therefore the relative handedness of the coordinate systems for the alternator transformation law is unexpected.

The title to this section mentioned both the alternator and the vector cross product. How are they connected? If you recall the definition of the vector cross product $\mathbf{a}\times\mathbf{b}$ in terms of a determinant, the connection between the two is made,

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2b_3 - b_2a_3)\mathbf{e}_1 + (a_3b_1 - b_3a_1)\mathbf{e}_2 + (a_1b_2 - b_1a_2)\mathbf{e}_3. \tag{A.113}$$

In the indicial notation the vector cross product $\mathbf{a}\times\mathbf{b}$ is written in terms of an alternator as

$$\mathbf{a}\times\mathbf{b} = e_{\mathrm{ijk}}a_{\mathrm{i}}b_{\mathrm{j}}e_{\mathrm{k}}, \tag{A.114}$$

a result that may be verified by expanding it to show that it coincides with (A.113). If $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ denotes the result of the vector cross product, then from (A.114),

$$\mathbf{c} = e_{\mathrm{ijk}} a_{\mathrm{i}} b_{\mathrm{j}} e_{\mathrm{k}}, (c_{\mathrm{k}} = e_{\mathrm{ijk}} a_{\mathrm{i}} b_{\mathrm{j}}). \tag{A.115}$$

Is the vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ a vector or a tensor? It is called a vector, but a second order tensorial character is suggested by the fact that the components of $\mathbf{a} \times \mathbf{b}$ coincide with the components of the skew-symmetric part of $2(\mathbf{a} \otimes \mathbf{b})$, see (A.32). The answer is that the vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ is unusual. Although it is, basically, a second order tensor, it can be treated as a vector as long as the transformations are between coordinate systems of the same handedness. In that case equation (A.113) shows that the alternator transforms as a proper tensor of order three, thus there is no ambiguity in the representation (A.114) for $\mathbf{a} \times \mathbf{b}$. When students first learn about the vector cross product they are admonished (generally without explanation) to always use right handed coordinate systems. This handedness problem is the reason for that admonishment. The "vector" that is the result of the cross product of two vectors has names like "axial vector" or "pseudo vector" to indicate its special character. Typical axial vectors are moments in mechanics and the vector curl operator (Sect. A.11).

Example A.8.3

Prove that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

Solution: In the formula (A.114) let $i \to j$ and $j \to i$, thus

$$\mathbf{a} \times \mathbf{b} = e_{\mathrm{jik}} a_{\mathrm{j}} b_{\mathrm{i}} e_{\mathrm{k}},$$

Next change e_{jik} to $-e_{\mathrm{ijk}}$ and rearrange the order of a_{j} and b_{i}, then the result is proved:

$$\mathbf{a} \times \mathbf{b} = -e_{\mathrm{ijk}} b_{\mathrm{i}} a_{\mathrm{j}} e_{\mathrm{k}} = -\mathbf{b} \times \mathbf{a}. \qquad \square$$

The scalar triple product of three vectors is a scalar formed from three vectors, $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ and the triple vector product is a vector formed from three vectors, $(\mathbf{r} \times (\mathbf{p} \times \mathbf{q}))$. An expression for the scalar triple product is obtained by taking the dot product of the vector $\mathbf{c}$ with the cross product in the representation (A.114) for $\mathbf{a} \times \mathbf{b}$, thus

$$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = e_{\mathrm{jik}} a_{\mathrm{j}} b_{\mathrm{i}} c_{\mathrm{k}}. \tag{A.116}$$

From the properties of the alternator it follows that

$$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = -\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) = -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}). \tag{A.117}$$

If the three vectors **a**, **b**, and **c** coincide with the three nonparallel edges of a parallelepiped, the scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is equal to the volume of the parallelepiped. In the following example a useful vector identity for the triple vector product $(\mathbf{r} \times (\mathbf{p} \times \mathbf{q}))$ is derived.

Example A.8.4

Prove that $(\mathbf{r} \times (\mathbf{p} \times \mathbf{q})) = (\mathbf{r}\cdot\mathbf{q})\mathbf{p} - (\mathbf{r}\cdot\mathbf{p})\mathbf{q}$.

Solution: First rewrite (A.114) with the change $\mathbf{a} \rightarrow \mathbf{r}$, and again with the changes $\mathbf{a} \rightarrow \mathbf{p}$ and $\mathbf{b} \rightarrow \mathbf{q}$, where $\mathbf{b} = (\mathbf{p} \times \mathbf{q})$

$$\mathbf{r} \times \mathbf{b} = e_{\mathrm{ijk}} r_{\mathrm{i}} b_{\mathrm{j}} e_{\mathrm{k}}, \mathbf{b} = \mathbf{p} \times \mathbf{q} = e_{\mathrm{mnj}} p_{\mathrm{m}} q_{\mathrm{n}} e_{\mathrm{j}},$$

Note that the second of these formulas gives the components of **b** as

$$b_{\mathrm{j}} = e_{\mathrm{mnj}} p_{\mathrm{m}} q_{\mathrm{n}}.$$

This formula for the components of **b** is then substituted into the expression for $(\mathbf{r} \times \mathbf{b}) = (\mathbf{r} \times (\mathbf{p} \times \mathbf{q}))$ above, thus

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = e_{\mathrm{ijk}} e_{\mathrm{mnj}} r_{\mathrm{i}} p_{\mathrm{m}} q_{\mathrm{n}} e_{\mathrm{k}}.$$

On the right-hand side of this expression for $\mathbf{r} \times (\mathbf{p} \times \mathbf{q})$, e_{ijk} is now changed to $-e_{\mathrm{ikj}}$ and the first of (A.111) is then employed,

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = -(\delta_{\mathrm{im}}\delta_{\mathrm{kn}} - \delta_{\mathrm{in}}\delta_{\mathrm{km}}) r_{\mathrm{i}} p_{\mathrm{m}} q_{\mathrm{n}} e_{\mathrm{k}}.$$

then summing over k and i

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = r_{\mathrm{i}} p_{\mathrm{k}} q_{\mathrm{i}} e_{\mathrm{k}} - r_{\mathrm{i}} p_{\mathrm{i}} q_{\mathrm{k}} e_{\mathrm{k}} = (\mathbf{r} \cdot \mathbf{q})\mathbf{p} - (\mathbf{r} \cdot \mathbf{p})\mathbf{q}. \qquad \square$$

In the process of calculating area changes on surfaces of objects undergoing large deformations, like rubber or soft tissue, certain identities that involve both vectors and matrices are useful. Two of these identities are derived in the following two examples.

Example A.8.5

Prove that $\mathbf{A}\cdot\mathbf{a}\cdot(\mathbf{A}\cdot\mathbf{b} \times \mathbf{A}\cdot\mathbf{c}) = \mathbf{a}\cdot(\mathbf{b} \times \mathbf{c})\,\mathrm{Det}\mathbf{A}$ where **A** is a 3 by 3 matrix and **a**, **b**, and **c** are vectors.

Solution: Noting the formula for the scalar triple product as a determinant

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

and the representation for the multiplication of **A** times **a**,

$$\mathbf{A}\cdot\mathbf{a}=\begin{bmatrix}A_{11} & A_{12} & A_{13}\\ A_{21} & A_{22} & A_{23}\\ A_{31} & A_{32} & A_{33}\end{bmatrix}\begin{bmatrix}a_1\\ a_2\\ a_3\end{bmatrix}=\begin{bmatrix}A_{11}a_1+A_{12}a_2+A_{13}a_3\\ A_{21}a_1+A_{22}a_2+A_{23}a_3\\ A_{31}a_1+A_{32}a_2+A_{33}a_3\end{bmatrix},$$

then

$$\begin{aligned}&\mathbf{A}\cdot\mathbf{a}\cdot(\mathbf{A}\cdot\mathbf{b}\times\mathbf{A}\cdot\mathbf{c})\\ &=\mathrm{Det}\begin{bmatrix}A_{11}a_1+A_{12}a_2+A_{13}a_3 & A_{21}a_1+A_{22}a_2+A_{23}a_3 & A_{31}a_1+A_{32}a_2+A_{33}a_3\\ A_{11}b_1+A_{12}b_2+A_{13}b_3 & A_{21}b_1+A_{22}b_2+A_{23}b_3 & A_{31}b_1+A_{32}b_2+A_{33}b_3\\ A_{11}c_1+A_{12}c_2+A_{13}c_3 & A_{21}c_1+A_{22}c_2+A_{23}c_3 & A_{31}c_1+A_{32}c_2+A_{33}c_3\end{bmatrix}.\end{aligned}$$

Recalling from Example A.8.1 that $\mathrm{Det}(\mathbf{A}\cdot\mathbf{B})=\mathrm{Det}\mathbf{A}\,\mathrm{Det}\mathbf{B}$, it follows that the previous determinant may be written as a product of determinants,

$$\mathrm{Det}\begin{bmatrix}A_{11} & A_{12} & A_{13}\\ A_{21} & A_{22} & A_{23}\\ A_{31} & A_{32} & A_{33}\end{bmatrix}\mathrm{Det}\begin{bmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{bmatrix}=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})\mathbf{Det}\mathbf{A},$$

which is the desired result. In the last step the fact that the determinant of the transpose of a matrix is equal to the determinant of the matrix, $\mathrm{Det}\mathbf{A}=\mathrm{Det}\mathbf{A}^{\mathrm{T}}$, was employed. □

Example A.8.6

Prove vector identity $(\mathbf{A}\cdot\mathbf{b}\times\mathbf{A}\cdot\mathbf{c})\cdot\mathbf{A}=(\mathbf{b}\times\mathbf{c})\,\mathrm{Det}\mathbf{A}$. where **A** is a 3 by 3 matrix and **b** and **c** are vectors.

Solution: Recall the result of Example A8.5, namely that $\mathbf{A}\cdot\mathbf{a}\cdot(\mathbf{A}\cdot\mathbf{b}\times\mathbf{A}\cdot\mathbf{c})=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})\,\mathrm{Det}\mathbf{A}$, and let $\mathbf{a}=\mathbf{e}_1$, then $\mathbf{e}_2$ and then $\mathbf{e}_3$ to obtain the following three scalar equations:

$$A_{11}w_1+A_{21}w_2+A_{31}w_3=\mathrm{q}_1\mathrm{Det}\mathbf{A}$$

$$A_{12}w_1+A_{22}w_2+A_{32}w_3=\mathrm{q}_2\mathrm{Det}\mathbf{A}$$

$$A_{13}w_1+A_{23}w_2+A_{33}w_3=\mathrm{q}_3\mathrm{Det}\mathbf{A}$$

where $\mathbf{w}=(\mathbf{A}\cdot\mathbf{b}\times\mathbf{A}\cdot\mathbf{c})$, $\mathbf{q}=(\mathbf{b}\times\mathbf{c})$. These three equations many be recast in the matrix notation,

$$\begin{bmatrix}A_{11} & A_{21} & A_{31}\\ A_{12} & A_{22} & A_{32}\\ A_{13} & A_{23} & A_{33}\end{bmatrix}\begin{bmatrix}w_1\\ w_2\\ w_3\end{bmatrix}=\begin{bmatrix}q_1\\ q_2\\ q_3\end{bmatrix}\mathrm{Det}\mathbf{A},$$

or $\mathbf{A}^{\mathrm{T}}\cdot\mathbf{w} = \mathbf{q}\,\mathrm{Det}\mathbf{A}$ and since $\mathbf{w} = (\mathbf{A}\cdot\mathbf{b} \times \mathbf{A}\cdot\mathbf{c})$, $\mathbf{q} = (\mathbf{b} \times \mathbf{c})$,

$$A^{\mathrm{T}}(\mathbf{A}\cdot\mathbf{b} \times \mathbf{A}\cdot\mathbf{c}) = (\mathbf{b} \times \mathbf{c})\mathrm{Det}\mathbf{A}, \text{ or}$$
$$(\mathbf{A}\cdot\mathbf{b} \times \mathbf{A}\cdot\mathbf{c})\cdot\mathbf{A} = (\mathbf{b} \times \mathbf{c})\mathrm{Det}\mathbf{A},$$

□

Problems

A.8.1. Find the cross products $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$ of the two vectors $\mathbf{a} = [1, 2, 3]$ and $\mathbf{b} = [4, 5, 6]$. What is the relationship between $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$?

A.8.2. Show that if $\mathbf{A}$ is a skew-symmetric 3 by 3 matrix, $\mathbf{A} = -\mathbf{A}^{\mathrm{T}}$, then $\mathrm{Det}\mathbf{A} = 0$.

A.8.3. Evaluate $\mathrm{Det}(\mathbf{a} \otimes \mathbf{b})$.

A.8.4. Show $\mathrm{Det}\mathbf{A} = \mathrm{Det}\mathbf{A}^{\mathrm{T}}$.

A.8.5. Show $\mathrm{Det}\mathbf{Q} = \pm 1$ if $\mathbf{Q}^{\mathrm{T}}\cdot\mathbf{Q} = \mathbf{Q}\cdot\mathbf{Q}^{\mathrm{T}} = \mathbf{1}$.

A.8.6. Find the volume of the parallelepiped if the three nonparallel edges of a parallelepiped coincide with the three vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ where $\mathbf{a} = [1, 2, 3]$ meters, $\mathbf{b} = [1, -4, 6]$ meters and $\mathbf{c} = [1, 1, 1]$ meters.

A.8.7. If $\mathbf{v} = \mathbf{a} \times \mathbf{x}$ and $\mathbf{a}$ is a constant vector, using the indicial notation, evaluate the div $\mathbf{v}$ and the curl $\mathbf{v}$.

A.9 The Moment of Inertia Tensor

The mass moment of inertia tensor illustrates many features of the previous sections such as the tensor concept and definition, the open product of vectors, the use of unit vectors and the significance of eigenvalues and eigenvectors. The mass moment of inertia is second moment of mass with respect to an axis. The first and zeroth moment of mass with respect to an axis is associated with the concepts of the center of mass of the object and the mass of the object, respectively. Let $\mathrm{d}v$ represent the differential volume of an object O. The volume of that object V_{O} is then given by

$$V_{\mathrm{O}} = \int_O \mathrm{d}v, \tag{A.118}$$

and, if $\rho(x_1, x_2, x_3, t) = \rho(\mathbf{x}, t)$ is the density of the object O, then the mass M_{O} of O is given by

$$M_{\mathrm{O}} = \int_O \rho(\mathbf{x}, t)\,\mathrm{d}v. \tag{A.119}$$

The centroid $\mathbf{x}_{\mathrm{centroid}}$ and the center of mass $\mathbf{x}_{\mathrm{cm}}$ of the object O are defined by

$$\mathbf{x}_{\mathrm{centroid}} = \frac{1}{V_{\mathrm{O}}}\int_O \mathbf{x}\mathrm{d}v,\ \mathbf{x}_{\mathrm{cm}} = \frac{1}{M_{\mathrm{O}}}\int_O \mathbf{x}\rho(\mathbf{x}, t)\,\mathrm{d}v \tag{A.120}$$

where $\mathbf{x}$ is a position vector locating the differential element of volume or mass with respect to the origin. The power of $\mathbf{x}$ occurring in the integrand indicates the order of the moment of the mass—it is to the zero power in the definition of the mass of the object itself—and it is to the first order in the definition of the mass center. The mass moment of inertia, which is the second moment of mass, arises as a convenient definition that occurs in the development of the conservation of angular momentum for an object as an integral of the moment of the linear momentum of an element of mass, dm, over the object. The expression for the angular momentum $\mathbf{H}$ of an object O as the integral over its volume of the cross product of the position vector $\mathbf{x}$ and the linear momentum $\rho\dot{\mathbf{x}}\mathrm{d}v = \dot{\mathbf{x}}\mathrm{d}m$ of an element of mass dm, is given by

$$\mathbf{H} = \int_O \mathbf{x} \times \rho\dot{\mathbf{x}}\mathrm{d}v. \qquad \text{(4.22, repeated)}$$

If the object is instantaneously rotating about an axis with an angular velocity ω and the rotational velocity $\dot{\mathbf{x}}$ at the mass element, dm is given by

$$\dot{\mathbf{x}} = \omega \times \mathbf{x}.$$

Substitution of this expression for the velocity into (4.22, repeated), an alternate expression for $\mathbf{H}$ is obtained,

$$\mathbf{H} = \int_O \mathbf{x} \times (\omega \times \mathbf{x})\rho\mathrm{d}v. \qquad \text{(A.121)}$$

The integrand in this new representation for $\mathbf{H}$ can be expressed differently using the vector identity $\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = (\mathbf{r} \cdot \mathbf{q})\mathbf{p} - (\mathbf{r} \cdot \mathbf{p})\mathbf{q}$ proved in Example A.8.4, thus

$$\mathbf{x} \times (\omega \times \mathbf{x}) = (\mathbf{x} \cdot \mathbf{x})\omega - (\mathbf{x} \cdot \omega)\mathbf{x}.$$

Incorporating this expression into the integrand in (A.121) it is easily seen that $\mathbf{H}$ has the representation

$$\mathbf{H} = \omega \cdot \int_O [(\mathbf{x} \cdot \mathbf{x})\mathbf{1} - (\mathbf{x} \otimes \mathbf{x})]\rho\mathrm{d}v. \qquad \text{(A.122a)}$$

This result is written more simply as

$$\mathbf{H} = \mathbf{I} \cdot \omega, \qquad \text{(A.122b)}$$

where the definition of the *mass moment of inertia* tensor,

$$\mathbf{I} \equiv \int_O [(\mathbf{x} \cdot \mathbf{x})\mathbf{1} - (\mathbf{x} \otimes \mathbf{x})]\rho\mathrm{d}v, \qquad \text{(A.123)}$$

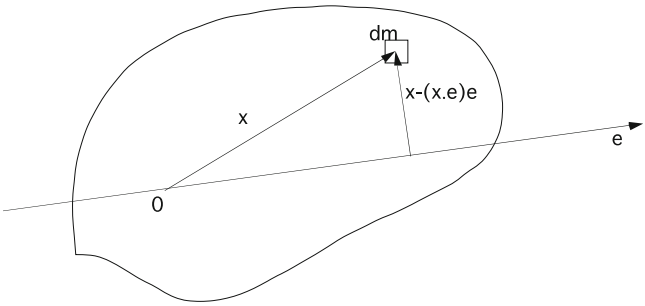


Fig. A.2 A diagram for the calculation of the mass moment of inertia of an object about the axis characterized by the unit vector **e**; **x** is the vector from the origin O of coordinates to the element of mass *dm*, $\mathbf{x}-(\mathbf{x}\cdot\mathbf{e})\mathbf{e}$ is the perpendicular distance from the axis **e** to the element of mass *dm*

has been introduced; note that it is a symmetric tensor. The rotational kinetic energy of the spinning object is then given by

$$K_{\text{rot}} = (1/2)\omega \cdot \mathbf{I} \cdot \omega. \tag{A.124}$$

A second perspective on the mass moment of inertia tensor, without the angular momentum motivation, is the following: Let **e** represent the unit vector passing through the origin of coordinates, then $\mathbf{x}-(\mathbf{x}\ \mathbf{e})\mathbf{e}$ is the perpendicular distance from the **e** axis to the differential element of volume or mass at **x** (Fig. A.2). The second or mass moment of inertia of the object O about the axis **e**, a scalar, is denoted by $I_{\mathbf{ee}}$ and given by

$$I_{\text{ee}} = \int_O (\mathbf{x} - (\mathbf{x}\cdot\mathbf{e})\mathbf{e}) \cdot (\mathbf{x} - (\mathbf{x}\cdot\mathbf{e})\mathbf{e})\rho(\mathbf{x}, t)\, \mathrm{d}v. \tag{A.125a}$$

This expression for $I_{\mathbf{ee}}$ may be changed in algebraic form by noting first that

$$(\mathbf{x} - (\mathbf{x}\cdot\mathbf{e})\mathbf{e}) \cdot (\mathbf{x} - (\mathbf{x}\cdot\mathbf{e})\mathbf{e}) = \mathbf{x}\cdot\mathbf{x} - (\mathbf{x}\cdot\mathbf{e})^2 \text{ and}$$

thus, from A(108),

$$I_{\text{ee}} = \mathbf{e} \cdot \left[\int_O \{(\mathbf{x}\cdot\mathbf{x})\mathbf{1} - (\mathbf{x}\otimes\mathbf{x})\}\rho(\mathbf{x}, t)\, \mathrm{d}v \right] \cdot \mathbf{e}. \tag{A.125b}$$

If the notation (A.123) for the mass moment of inertia tensor **I** is introduced,then the representation for (A.125b) simplifies to

$$I_{\text{ee}} = \mathbf{e} \cdot \mathbf{I} \cdot \mathbf{e}. \tag{A.125c}$$

In this section the mass moment of inertia $\mathbf{I}$ has been referred to as a tensor. A short calculation will demonstrate that the terminology is correct. From (A.123) is easy to see that $\mathbf{I}$ may be written relative to the Latin and Greek coordinate systems as

$$\mathbf{I}^{(\mathrm{L})} = \int_O \{(\mathbf{x}^{(\mathrm{L})} \cdot \mathbf{x}^{(\mathrm{L})})\mathbf{1} - (\mathbf{x}^{(\mathrm{L})} \otimes \mathbf{x}^{(\mathrm{L})})\}\rho(\mathbf{x}^{(\mathrm{L})}, t)\, \mathrm{d}v, \tag{A.126a}$$

and

$$\mathbf{I}^{(\mathrm{G})} = \int_O \{(\mathbf{x}^{(\mathrm{G})} \cdot \mathbf{x}^{(\mathrm{G})})\mathbf{1} - (\mathbf{x}^{(\mathrm{G})} \otimes \mathbf{x}^{(\mathrm{G})})\}\rho(\mathbf{x}^{(\mathrm{G})}, t)\, \mathrm{d}v, \tag{A.126b}$$

respectively. The transformation law for the open product of $\mathbf{x}$ with itself can be calculated by twice using the transformation law for vectors (A.77) applied to $\mathbf{x}$, thus

$$\mathbf{x}^{(\mathrm{L})} \otimes \mathbf{x}^{(\mathrm{L})} = \mathbf{Q} \cdot \mathbf{x}^{(\mathrm{G})} \otimes \mathbf{Q} \cdot \mathbf{x}^{(\mathrm{G})} = \mathbf{Q} \cdot (\mathbf{x}^{(\mathrm{G})} \otimes \mathbf{x}^{(\mathrm{G})}) \cdot \mathbf{Q}^{\mathrm{T}}. \tag{A.127}$$

The occurrence of the transpose in the last equality of the last equation may be more easily perceived by recasting the expression in the indicial notation:

$$\mathbf{x}_{\mathrm{i}}^{(\mathrm{L})}\mathbf{x}_{\mathrm{j}}^{(\mathrm{L})} = Q_{\mathrm{i}\alpha}\mathbf{x}_{\alpha}^{(\mathrm{G})}Q_{\mathrm{j}\beta}\mathbf{x}_{\beta}^{(\mathrm{G})} = Q_{\mathrm{i}\alpha}\mathbf{x}_{\alpha}^{(\mathrm{G})}\mathbf{x}_{\beta}^{(\mathrm{G})}Q_{\mathrm{j}\beta}. \tag{A.128}$$

Now, contracting the open product of vectors in (A.127) above to the scalar product, it follows that since $\mathbf{Q} \cdot \mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{Q} = \mathbf{1}(Q_{\mathrm{i}\alpha}Q_{\mathrm{i}\beta} = \delta_{\alpha\beta})$,

$$\mathbf{x}^{(\mathrm{L})} \cdot \mathbf{x}^{(\mathrm{L})} = \mathbf{x}^{(\mathrm{G})} \cdot \mathbf{x}^{(\mathrm{G})}. \tag{A.129}$$

Combining the results (A.127) and (A.129) it follows that the non-scalar portions of the integrands in (A.126a) and (A.126b) are related by

$$\{(\mathbf{x}^{(\mathrm{L})} \cdot \mathbf{x}^{(\mathrm{L})})\mathbf{1} - (\mathbf{x}^{(\mathrm{L})} \otimes \mathbf{x}^{(\mathrm{L})})\} = \mathbf{Q} \cdot \{(\mathbf{x}^{(\mathrm{G})} \cdot \mathbf{x}^{(\mathrm{G})})\mathbf{1} - (\mathbf{x}^{(\mathrm{G})} \otimes \mathbf{x}^{(\mathrm{G})})\} \cdot \mathbf{Q}^{\mathrm{T}}.$$

Thus from this result and (A.126a) and (A.126b) the transformation law for second order tensors is obtained,

$$\mathbf{I}^{(\mathrm{L})} = \mathbf{Q} \cdot \mathbf{I}^{(\mathrm{G})} \cdot \mathbf{Q}^{\mathrm{T}}, \tag{A.130}$$

and it follows that tensor terminology is correct in describing the mass moment of inertia.

The matrix of tensor components of the moment of inertia tensor $\mathbf{I}$ in a three-dimensional space is given by,

$$\mathbf{I} = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{bmatrix}, \tag{A.131}$$

where the components are given by

$$\begin{aligned} I_{11} &= \int_O (x_2^2 + x_3^2)\rho(\mathbf{x}, t)\,\mathrm{d}v,\ I_{22} = \int_O (x_1^2 + x_3^2)\rho(\mathbf{x}, t)\,\mathrm{d}v, \\ I_{33} &= \int_O (x_2^2 + x_1^2)\rho(\mathbf{x}, t)\,\mathrm{d}v,\ I_{12} = -\int_O (x_1 x_2)\rho(\mathbf{x}, t)\,\mathrm{d}v \\ I_{13} &= -\int_O (x_1 x_2)\rho(\mathbf{x}, t)\,\mathrm{d}v,\ I_{23} = -\int_O (x_2 x_3)\rho(\mathbf{x}, t)\,\mathrm{d}v. \end{aligned} \tag{A.130}$$

Example A.9.1

Determine the mass moment of inertia of a rectangular prism of homogeneous material of density ρ and side lengths a, b, and c about one corner. Select the coordinate systems so that its origin is at one corner and let a, b, c represent the distances along the x_1, x_2, x_3 axes, respectively. Construct the matrix of tensor components referred to this coordinate system.

Solution: The integrations (A.132) yield the following results:

$$\begin{aligned} I_{11} &= \int_O (x_2^2 + x_3^2)\rho(\mathbf{x}, t)\,\mathrm{d}v = \rho \int_O (x_2^2 + x_3^2)\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &= a\rho \int_{0,0}^{b,c} (x_2^2 + x_3^2)\,\mathrm{d}x_2\mathrm{d}x_3 = \frac{\rho abc}{3}(b^2 + c^2), \\ I_{22} &= \frac{\rho abc}{3}(a^2 + c^2),\ I_{33} = \frac{\rho abc}{3}(a^2 + b^2), \\ I_{12} &= -\int_O (x_1 x_2)\rho(\mathbf{x}, t)\,\mathrm{d}v = -\rho c \int_{0,0}^{a.b} (x_1 x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \frac{-\rho abc}{4}(ab) \\ I_{13} &= \frac{-\rho abc}{4}(ac),\ I_{23} = \frac{-\rho abc}{4}(bc), \end{aligned}$$

thus

$$I = \frac{\rho abc}{12}\begin{bmatrix} 4(b^2+c^2) & -3ab & -3ac \\ -3ab & 4(a^2+c^2) & -3bc \\ -3ac & -3bc & 4(a^2+b^2) \end{bmatrix}.$$

□

Example A.9.2

In the special case when the rectangular prism in Example A.9.1 is a cube, that is to say $a = b = c$, find the eigenvalues and eigenvectors of the matrix of tensor components referred to the coordinate system of the example. Then find the matrix of tensor components referred to the principal, or eigenvector, coordinate system.

Solution: The matrix of tensor components referred to this coordinate system is

$$I = \frac{M_O a^2}{12}\begin{bmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{bmatrix}.$$

The eigenvalues of $\mathbf{I}$ are $M_o a^2/6$, $11M_o a^2/12$, and $11M_o a^2/12$. The eigenvector $(1/\sqrt{3})[1, 1, 1]$ is associated with the eigenvalue $M_o a^2/6$. Due the multiplicity of the eigenvalue $11M_o a^2/12$, any vector perpendicular to the first eigenvector $(1/\sqrt{3})$ $[1, 1, 1]$ is an eigenvector associated the multiple eigenvalue $11M_o a^2/12$. Thus any mutually perpendicular unit vectors in the plane perpendicular to the first eigenvector may be selected as the base vectors for the principal coordinate system. The choice is arbitrary. In this example the two perpendicular unit vectors $(1/\sqrt{2})$ $[-1, 0, 1]$ and $(1/\sqrt{6})[1, -2, 1]$ are the eigenvectors associated with the multiple eigenvalue $11M_o a^2/12$, but any perpendicular pair of vectors in the plane may be selected. The orthogonal transformation that will transform the matrix of tensor components referred to this coordinate system to the matrix of tensor components referred to the principal, or eigenvector, coordinate system is then given by

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}.$$

Applying this transformation produced the matrix of tensor components referred to the principal, or eigenvector, coordinate system

$$\mathbf{Q}\cdot\mathbf{I}\cdot\mathbf{Q}^{\mathrm{T}} = \frac{M_O a^2}{12}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 11 \end{bmatrix}.$$

□

Formulas for the mass moment of inertia of a thin plate of thickness t and a homogeneous material of density ρ are obtained by specializing these results. Let the plate be thin in the x_3 direction and consider the plate to be so thin that terms of the order t^2 are negligible relative to the others, then the formulas (A.132) for the components of the mass moment of inertia tensor are given by

$$\begin{aligned} I_{11} &= \rho t \int_O x_2^2 \mathrm{d}x_1 \mathrm{d}x_2,\ I_{22} = \rho t \int_O x_1^2 \mathrm{d}x_1 \mathrm{d}x_2, \\ I_{33} &= \rho t \int_O (x_1^2 + x_2^2)\, \mathrm{d}x_1 \mathrm{d}x_2. \\ I_{12} &= -\rho t \int_O (x_1 x_2)\, \mathrm{d}x_1 \mathrm{d}x_2,\ I_{13} = 0,\ I_{23} = 0 \end{aligned} \tag{A.133}$$

When divided by ρt these components of the mass moment of inertia of a thin plate of thickness t are called the components of the area moment of inertia matrix,

$$\begin{aligned} I_{11}^{\text{Area}} &= \frac{I_{11}}{\rho t} = \int_O x_2^2 \mathrm{d}x_1 \mathrm{d}x_2,\ I_{22}^{\text{Area}} = \frac{I_{22}}{\rho t} = \int_O x_1^2 \mathrm{d}x_1 \mathrm{d}x_2, \\ I_{33}^{\text{Area}} &= \frac{I_{33}}{\rho t} = \int_O (x_1^2 + x_2^2)\, \mathrm{d}x_1 \mathrm{d}x_2, \\ I_{12}^{\text{Area}} &= \frac{I_{12}}{\rho t} = -\int_O (x_1 x_2)\, \mathrm{d}x_1 \mathrm{d}x_2,\ I_{13}^{\text{Area}} = 0, I_{23}^{\text{Area}} = 0. \end{aligned} \tag{A.134}$$

Example A.9.3
Determine the area moment of inertia of a thin rectangular plate of thickness t, height h, and a width of base b, and a homogeneous material of density ρ. Specify precisely where the origin of the coordinate system that you are using is located and how the base vectors of that coordinate system are located relative to the sides of the rectangular plate.

Solution: The coordinate system that makes this problem easy is one that passes through the centroid of the rectangle and has axes that are parallel to the sides of the rectangle. If the base b is parallel to the x_1 axis and height h is parallel to the x_2 axis then the integrations (A.134) yield the following results:

$$I_{11}^{\text{Area}} = \frac{bh^3}{12}, I_{22}^{\text{Area}} = \frac{hb^3}{12}, I_{33}^{\text{Area}} = \frac{bh}{12}(b^2 + h^2), I_{12}^{\text{Area}} = 0, I_{13}^{\text{Area}} = 0, I_{23}^{\text{Area}} = 0. \quad \square$$

Example A.9.4

Determine the area moments and product of inertia of a thin right-triangular plate of thickness t, height h, and a width of base b, and a homogeneous material of density ρ.

Let the base b be along the x_1 axis and the height h be along the x_2 axis and the sloping face of the triangle have end points at $(b, 0)$ and $(0, h)$. Determine the area moments and product of inertia of the right-triangular plate relative to this coordinate system. Construct the matrix of tensor components referred to this coordinate system.

Solution: The integrations (A.134) yield the following results:

$$I_{11}^{\text{Area}} = \int_O x_2^2 \mathrm{d}x_1 \mathrm{d}x_2 = \int_0^h b\Big(1 - \frac{x_2}{h}\Big) x_2^2 \mathrm{d}x_2 = \frac{bh^3}{12},\ I_{22}^{\text{Area}} = \frac{hb^3}{12}$$

$$I_{12}^{\text{Area}} = -\int_O (x_1 x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 = -\left(\frac{1}{2}\right) \int_0^h b^2 \Big(1 - \frac{x_2}{h}\Big)^2 x_2 \mathrm{d}x_2 = -\left(\frac{b^2 h^2}{24}\right),$$

thus the matrix to tensor components referred to this coordinate system is

$$\mathbf{I}^{\text{Area}} = \frac{bh}{24} \begin{bmatrix} 2h^2 & -bh \\ -hb & 2b^2 \end{bmatrix}. \qquad \square$$

Example A.9.5

In the special case when the triangle in Example A.9.4 is an isosceles triangle, that is to say $b = h$, find the eigenvalues and eigenvectors of the matrix of tensor components referred to the coordinate system of the example. Then find the matrix of tensor components referred to the principal, or eigenvector, coordinate system.

Solution: The matrix of tensor components referred to this coordinate system is

$$\mathbf{I}^{\text{Area}} = \frac{bh}{24} \begin{bmatrix} 2h^2 & -bh \\ -hb & 2b^2 \end{bmatrix}.$$

The eigenvalues of $\mathbf{I}$ are $h^4/8$ and $h^4/24$. The eigenvector $(1/\sqrt{2})[1, -1]$ is associated with the eigenvalue $h^4/8$ and the eigenvector $(1/\sqrt{2})[1, 1]$ is associated with the eigenvalue $h^4/24$. The orthogonal transformation that will transform the matrix of tensor components referred to this coordinate system to the matrix of tensor components referred to the principal, or eigenvector, coordinate system is then given by

$$\mathbf{Q} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Applying this transformation produced the matrix of tensor components referred to the principal, or eigenvector, coordinate system

$$\mathbf{Q}\cdot\mathbf{I}^{\text{Area}}\cdot\mathbf{Q}^{\text{T}}=\frac{h^4}{24}\begin{bmatrix}1&0\\0&3\end{bmatrix}. \qquad \square$$

The parallel axis theorem for the moment of inertia matrix $\mathbf{I}$ is derived by considering the mass moment of inertia of the object O about two parallel axes, $I_{\mathbf{ee}}$ about $\mathbf{e}$ and $I_{\mathbf{e}'\mathbf{e}'}$ about $\mathbf{e}'$. $I_{\mathbf{e}'\mathbf{e}'}$ is given by

$$I_{\mathbf{e}'\mathbf{e}'}=\mathbf{e}'\cdot\mathbf{I}'\cdot\mathbf{e}', \tag{A.135}$$

where the moment of inertia matrix $\mathbf{I}'$ is given by

$$\mathbf{I}'=\int_O\{(\mathbf{x}'\cdot\mathbf{x}')\mathbf{1}-(\mathbf{x}'\otimes\mathbf{x}')\}\rho(\mathbf{x}',t)\,\mathrm{d}v'. \tag{A.136}$$

Let $\mathbf{d}$ be a vector perpendicular to both $\mathbf{e}$ and $\mathbf{e}'$ and equal in magnitude to the perpendicular distance between $\mathbf{e}$ and $\mathbf{e}'$, thus $\mathbf{x}'=\mathbf{x}+\mathbf{d}$, $\mathbf{e}\cdot\mathbf{d}=0$, and $\mathbf{e}'\cdot\mathbf{d}=0$. Substituting $\mathbf{x}'=\mathbf{x}+\mathbf{d}$ in $\mathbf{I}'$, it follows that

$$\begin{aligned}\mathbf{I}'=&\int_O\{(\mathbf{x}\cdot\mathbf{x}+\mathbf{d}\cdot\mathbf{d}+2\mathbf{x}\cdot\mathbf{d})\mathbf{1}\}\rho(\mathbf{x},t)\,\mathrm{d}v\\&-\int_O\{(\mathbf{x}\otimes\mathbf{x}+\mathbf{d}\otimes\mathbf{d}+\mathbf{d}\otimes\mathbf{x}+\mathbf{x}\otimes\mathbf{d})\}\rho(\mathbf{x},t)\,\mathrm{d}v\end{aligned} \tag{A.137}$$

or if (A.137) is rewritten so that the constant vector $\mathbf{d}$ is outside the integral signs,

$$\begin{aligned}\mathbf{I}'=&\mathbf{1}\int_O\{(\mathbf{x}\cdot\mathbf{x})\rho(\mathbf{x},t)\,\mathrm{d}v+\mathbf{1}(\mathbf{d}\cdot\mathbf{d})\int_O\rho(\mathbf{x},t)\,\mathrm{d}v+\mathbf{1}\left(2\mathbf{d}\cdot\int_O\mathbf{x}\right)\rho(\mathbf{x},t)\,\mathrm{d}v\\&-\int_O\{(\mathbf{x}\otimes\mathbf{x})\rho(\mathbf{x},t)\,\mathrm{d}v-(\mathbf{d}\otimes\mathbf{d})\int_O\rho(\mathbf{x},t)\,\mathrm{d}v-\mathbf{d}\otimes\int_O\mathbf{x}\rho(\mathbf{x},t)\,\mathrm{d}v-\left(\int_O\mathbf{x}\rho(\mathbf{x},t)\,\mathrm{d}v\right)\otimes\mathbf{d}\end{aligned}$$

then recalling the definitions (A.119) of the mass M_{O} of O and (A.120) of the center of mass $\mathbf{x}_{\text{cm}}$ of the object O, this result simplifies to

$$\mathbf{I}'=\mathbf{I}+\{(\mathbf{d}\cdot\mathbf{d})\mathbf{1}-(\mathbf{d}\otimes\mathbf{d})\}M_{\text{O}}+2M_{\text{O}}\mathbf{1}(\mathbf{x}_{\text{cm}}\cdot\mathbf{d})-M_{\text{O}}(\mathbf{d}\otimes\mathbf{x}_{\text{cm}}+\mathbf{x}_{\text{cm}}\otimes\mathbf{d}). \tag{A.138}$$

Thus, when the origin of coordinates is taken at the center of the mass, it follows that $\mathbf{x}_{\text{cm}}=0$ and

$$\mathbf{I}'=\mathbf{I}_{\text{cm}}+\{(\mathbf{d}\cdot\mathbf{d})\mathbf{1}-(\mathbf{d}\otimes\mathbf{d})\}M_{\text{O}}. \tag{A.139}$$

In the special case of the area moment of inertia this formula becomes

$$\mathbf{I}' = \mathbf{I}_{\text{centroid}} + \{(\mathbf{d}\cdot\mathbf{d})\mathbf{1} - (\mathbf{d}\otimes\mathbf{d})\}A, \tag{A.140}$$

where the $\mathbf{I}$ are now the area moments of inertia and the mass of the object M_{O} has been replaced by the area of the thin plate A.

Example A.9.6

Consider again the rectangular prism of Example A.9.1. Determine the mass moment of inertia tensor of that prism about its centroid or center of mass.

Solution: The desired result, the mass moment of inertia about the centroidal axes is the $\mathbf{I}_{\text{cm}}$ in (A.139) and the moment of inertia about the corner, $\mathbf{I}'$, is the result calculated in Example A.9.1,

$$\mathbf{I}' = \frac{\rho abc}{12}\begin{bmatrix} 4(b^2+c^2) & -3ab & -3ac \\ -3ab & 4(a^2+c^2) & -3bc \\ -3ac & -3bc & 4(a^2+b^2) \end{bmatrix}.$$

The formula (A.139) is then written in the form

$$\mathbf{I}_{\text{cm}} = \mathbf{I}' - \{(\mathbf{d}\cdot\mathbf{d})\mathbf{1} - (\mathbf{d}\otimes\mathbf{d})\}M_{\text{O}},$$

where $M_{\text{o}} = \rho abs$. The vector $\mathbf{d}$ is a vector from the centroid to the corner,

$$\mathbf{d} = -\left(\frac{1}{2}\right)(a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3).$$

Substituting $\mathbf{I}'$ and the formula for $\mathbf{d}$ into the equation for $\mathbf{I}$ above, it follows that the mass moment of inertia of the rectangular prism relative to its centroid is given by

$$\mathbf{I}_{\text{cm}} = \frac{\rho abc}{12}\begin{bmatrix} (b^2+c^2) & 0 & 0 \\ 0 & (a^2+c^2) & 0 \\ 0 & 0 & (a^2+b^2) \end{bmatrix}. \qquad \square$$

Example A.9.7

Consider again the thin right-triangular plate of Example A.9.4. Determine the area moment of inertia tensor of that right-triangular plate about its centroid.

Solution: The desired result, the area moment of inertia about the centroidal axes is the $\mathbf{I}^{\text{Area}}_{\text{centroid}}$ in (A.140) and the moment of inertia, $\mathbf{I}'_{\text{Area}}$, about the corner is the result calculated in Example A.9.4,

$$\mathbf{I}'_{\text{Area}} = \frac{bh}{24}\begin{bmatrix} 2h^2 & -bh \\ -hb & 2b^2 \end{bmatrix}.$$

The formula (A.139) is then written in the form

$$\mathbf{I}^{\text{Area}}_{\text{centroid}} = \mathbf{I}^{'}_{\text{Area}} - \{(\mathbf{d}\cdot\mathbf{d})\mathbf{1} - (\mathbf{d}\otimes\mathbf{d})\}A,$$

where $2A = b$ □.

The vector $\mathbf{d}$ is a vector from the centroid to the corner,

$$\mathbf{d} = -\left(\frac{1}{3}\right)(b\mathbf{e}_1 + h\mathbf{e}_2).$$

Substituting $\mathbf{I}'$ and the formula for $\mathbf{d}$ into the equation for $\mathbf{I}_{\text{centroid}}$ above, it follows that the mass moment of inertia of the rectangular prism relative to its centroid is given by

$$\mathbf{I}^{\text{Area}}_{\text{centroid}} = \frac{bh}{72}\begin{bmatrix} 2h^2 & bh \\ hb & 2b^2 \end{bmatrix}.$$ □

Problems

A.9.1 Find the center of mass of a set of four masses. The form masses and their locations are mass 1 (2 kg) at (3, −1), mass 2 (4 kg) at (4, 4), mass 3 (5 kg) at (−4, 4), mass 4 (1 kg) at (−3, −1).

A.9.2 Under what conditions does the center of mass of an object coincide with the centroid?

A.9.3 Find the centroid of a cylinder of length L with a semicircular cross-section of radius R.

A.9.4 Find the center of mass of a cylinder of length L with a semicircular cross-section of radius R ($R < 2L$) if the density varies in according to the rule $\rho = \rho_\circ(1 + c(x_2)^2)$. The coordinate system for the cylinder has been selected so that x_3 is along its length L, x_2 is across its smallest dimension ($0 \le x_2 \le R$) and x_1 is along its intermediate dimension ($-R \le x_1 \le R$).

A.9.5 Show that the moment of inertia matrix $\mathbf{I}$ is symmetric.

A.9.6 Develop the formulas for the mass moment of inertia of a thin plate of thickness t and a homogeneous material of density ρ. Illustrate these specialized formulas by determining the mass moment of inertia of a thin rectangular plate of thickness t, height h, and a width of base b, and a homogeneous material of density ρ. Specify precisely where the origin of the coordinate system that you are using is located and how the base vectors of that coordinate system are located relative to the sides of the rectangular plate.

A.9.7 In Example A.9.2 the occurrence of a multiple eigenvalue ($7\rho a^3/12$) made any vector perpendicular to the first eigenvector $\mathbf{e}_1 = (1/\sqrt{3})[1, 1, 1]$ an eigenvector associated the multiple eigenvalue $7\rho a^3/12$. In Example A.9.2 the two perpendicular unit vectors $\mathbf{e}_2 = (1/\sqrt{2})[-1, 0, 1]$ and $\mathbf{e}_3 = (1/\sqrt{6})[1, -2, 1]$ were selected as the eigenvectors associated with the multiple

eigenvalue $7\rho a^3/12$, but any two perpendicular vectors in the plane could have been selected. Select two other eigenvectors in the plane and show that these two eigenvectors are given by $\mathbf{e}_{\mathrm{II}} = \cos\gamma\ \mathbf{e}_2 + \sin\gamma\ \mathbf{e}_3$ and $\mathbf{e}_{\mathrm{III}} = -\sin\gamma\ \mathbf{e}_2 + \cos\gamma\ \mathbf{e}_3$. Let $\mathbf{R}$ be the orthogonal transformation between these Latin and Greek systems,

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & \sin\gamma \\ 0 & -\sin\gamma & \cos\gamma \end{bmatrix}.$$

Show that when the Greek coordinate system is used rather than the Latin one, the coordinate transformation that diagonalizes the $\mathbf{I}$ matrix is $\mathbf{Q}\cdot\mathbf{R}$ rather than $\mathbf{Q}$. Show that both $\mathbf{R}\cdot\mathbf{Q}$ and $\mathbf{Q}$ transform the $\mathbf{I}$ matrix into the coordinate system in which it is diagonal,

$$\mathbf{Q}\cdot\mathbf{I}\cdot\mathbf{Q}^{\mathrm{T}} = \mathbf{R}\cdot\mathbf{Q}\cdot\mathbf{I}\cdot\mathbf{Q}^{\mathrm{T}}\cdot\mathbf{R}^{\mathrm{T}} = \frac{\rho a^3}{12}\begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}. \qquad \square$$

A.10 Connection to Mohr's Circles

The material in the section before last, namely the transformation law (A.83) for tensorial components and the eigenvalue problem for linear transformations, is presented in standard textbooks on the mechanics of materials in a more elementary fashion. In those presentations the second order tensor is taken to be the stress tensor and a geometric analog calculator is used for the transformation law (A.83) for tensorial components in two dimensions, and for the solution of the eigenvalue problem in two dimensions. The geometric analog calculator is called the Mohr circle. A discussion of the connection is included to aid in placing the material just presented in perspective.

The special case of the first transformation law from (A.83), $\mathbf{T}^{(\mathrm{L})} = \mathbf{Q}\cdot\mathbf{T}^{(\mathrm{G})}\cdot\mathbf{Q}^{\mathrm{T}}$, is rewritten in two dimensions ($n = 2$) in the form $\boldsymbol{\sigma}' = \mathbf{Q}\cdot\boldsymbol{\sigma}\cdot\mathbf{Q}^{\mathrm{T}}$; thus, $\mathbf{T}^{(\mathrm{L})} = \boldsymbol{\sigma}'$ and $\mathbf{T}^{(\mathrm{G})} = \boldsymbol{\sigma}$, where the matrix of stress tensor components $\boldsymbol{\sigma}$, the matrix of transformed stress tensor components $\boldsymbol{\sigma}'$, and the orthogonal transformation $\mathbf{Q}$ representing the rotation of the Cartesian axes are given by

$$\sigma = \begin{bmatrix} \sigma_x & \tau_{xy} \\ \tau_{xy} & \sigma_y \end{bmatrix}, \sigma' = \begin{bmatrix} \sigma_{x'} & \tau_{x'y'} \\ \tau_{x'y'} & \sigma_{y'} \end{bmatrix}, \mathbf{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{A.141}$$

Expansion of the matrix equation $\boldsymbol{\sigma}' = \mathbf{Q}\cdot\boldsymbol{\sigma}\cdot\mathbf{Q}^{\mathrm{T}}$,

$$\sigma' = \begin{bmatrix} \sigma_{x'} & \tau_{x'y'} \\ \tau_{x'y'} & \sigma_{y'} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \sigma_x & \tau_{xy} \\ \tau_{xy} & \sigma_y \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{A.142}$$

and subsequent use of the double angle trigonometric formulas $\sin 2\theta = 2 \sin \theta \cos \theta$ and $\cos 2\theta = \cos^2 \theta - \sin^2 \theta$ yield the following:

$$\begin{aligned}
\sigma_{x'} &= (1/2)(\sigma_x + \sigma_y) + (1/2)(\sigma_x - \sigma_y)\cos 2\theta + \tau_{xy} \sin 2\theta \\
\sigma_{y'} &= (1/2)(\sigma_x + \sigma_y) - (1/2)(\sigma_x - \sigma_y)\cos 2\theta - \tau_{xy} \sin 2\theta, \\
\tau_{x'y'} &= -(1/2)(\sigma_x - \sigma_y)\sin 2\theta + t_{xy} \cos 2\theta.
\end{aligned} \tag{A.143}$$

These are formulas for the stresses $\sigma_{x'}$, $\sigma_{y'}$, and $\tau_{x'y'}$ as functions of the stresses σ_x, σ_y and τ_{xy}, and the angle 2θ. Note that the sum of the first two equations in (A.143) yields the following expression, which is defined as $2C$,

$$2C \equiv \sigma_{x'} + \sigma_{y'} = \sigma_x + \sigma_y. \tag{A.144}$$

The fact that $\sigma_{x'} + \sigma_{y'} = \sigma_x + \sigma_y$ is a repetition of the result (A.90) concerning the invariance of the trace of a tensor, the first invariant of a tensor, under change of basis. Next consider the following set of equations in which the first is the first of (A.143) incorporating the definition (A.144) and transposing the term involving C to the other side of the equal sign, and the second equation is the third of (A.143):

$$\begin{aligned}
\sigma_{x'} - C &= (1/2)(\sigma_x - \sigma_y)\cos 2\theta + \tau_{xy} \sin 2\theta, \\
\tau_{x'y'} &= -(1/2)(\sigma_x - \sigma_y)\sin 2\theta + \tau_{xy} \cos 2\theta.
\end{aligned}$$

If these equations are now squared and added we find that

$$(\sigma_{x'} - C)^2 + (\tau_{x'y'})^2 = R^2 \tag{A.145}$$

where,

$$R^2 \equiv (1/4)(\sigma_x - \sigma_y)^2 + (\tau_{xy})^2. \tag{A.146}$$

Equation (A.145) is the equation for a circle of radius R centered at the point $\sigma_{x'} = C$, $\tau_{x'y'} = 0$. The circle is illustrated in Fig. A.3.

The points on the circle represent all possible values of $\sigma_{x'}$, $\sigma_{y'}$ and $\tau_{x'y'}$; they are determined by the values of C and R, which are, in turn, determined by σ_x, σ_y, and τ_{xy}. The eigenvalues of the matrix $\boldsymbol{\sigma}$ are the values of the normal stress $\sigma_{x'}$ when the circle crosses the $\sigma_{x'}$ axis. These are given by the numbers $C + R$ and $C - R$, as may be seen from Fig. A.3. Thus Mohr's circle is a graphical analog calculator for the eigenvalues of the two-dimensional second order tensor $\boldsymbol{\sigma}$, as well as a graphical analog calculator for the equation $\boldsymbol{\sigma}' = \mathbf{Q} \cdot \boldsymbol{\sigma} \cdot \mathbf{Q}^{\mathrm{T}}$ representing the transformation of components. The maximum shear stress is simply the radius of the circle R, an important graphical result that is readable from Fig. A.3.

As a graphical calculation device, Mohr's circles may be extended to three dimensions, but the graphical calculation is much more difficult than doing the

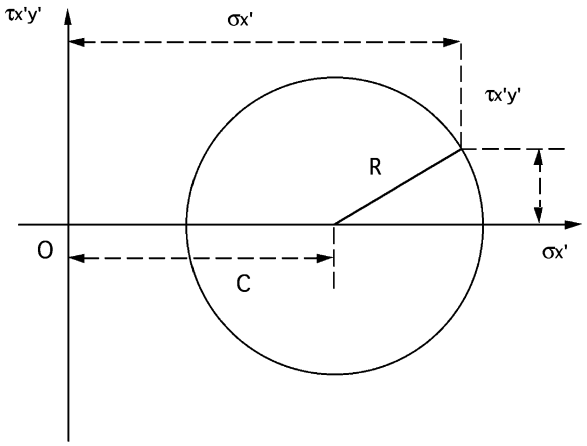


Fig. A.3 An illustration of Mohr's circle for a state of stress

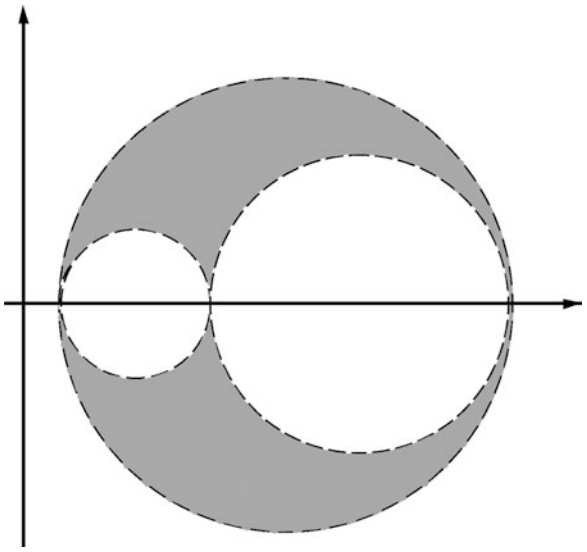


Fig. A.4 Mohr's circles in three dimensions

calculation on a computer so it is no longer done. An illustration of three-dimensional Mohr's circles is shown in Fig. A.4. The shaded region represents the set of points that are possible stress values. The three points where the circles intersect the axis correspond to the three eigenvalues of the three-dimensional stress tensor and the radius of the largest circle is the magnitude of the largest shear stress.

Problems

A.10.1 Construct the two-dimensional Mohr's circle for the matrix **A** given in problem A.7.3.

A.10.2 Construct the three-dimensional Mohr's circles for the matrix **T** given in problem A.7.2.

A.11 Special Vectors and Tensors in Six Dimensions

The fact that the components of a second order tensor in n dimensions can be represented as an n-by-n square matrix allows the powerful algebra of matrices to be used in the analysis of second order tensor components. In general this use of the powerful algebra of matrices is not possible for tensors of other orders. For example in the case of the third order tensor with components A_{ijk} one could imagine a generalization of a matrix from an array with rows and columns to one with rows, columns and a depth dimension to handle the information of the third index. This would be like an n-by-n-by-n cube sub-partitioned into n^3 cells that would each contains an entry similar to the entry at a row/column position in a matrix. Modern symbolic algebra programs might be extended to handle these n-by-n-by-n cubes and to represent them graphically. By extension of this idea, fourth order tensors would require an n-by-n-by-n-by-n hypercube with no possibility of graphical representation. Fortunately for certain fourth order tensors (a case of special interest in continuum mechanics) there is a way to again employ the matrix algebra of n-by-n square matrices in the representation of tensor components. The purposes of this section it to explain how this is done.

The developments in this text will frequently concern the relationship between symmetric second order tensors in three dimensions. The symmetric second order tensors of interest will include stress and stress rate and strain and strain rate, among others. The most general form of a linear relationship between second order tensors in three dimensions involves a three-dimensional fourth order tensor. In general the introduction of tensors of order higher than two involves considerable additional notation. However, since the interest here is only in three-dimensional fourth order tensors that relate three-dimensional symmetric second order tensors, a simple notational scheme can be introduced. The basis of the scheme is to consider a three-dimensional symmetric second order tensor also as a six-dimensional vector, and then three-dimensional fourth order tensors may be associated with second order tensors in a space of six dimensions. When this association is made, all of the algebraic machinery associated with the linear transformations and second order tensors is available for the three-dimensional fourth order tensors.

The first point to be made is that symmetric second order tensors in three dimensions may also be considered as vectors in a six-dimensional space. The one-to-one connection between the components of the symmetric second order tensors **T** and the six-dimensional vector $\hat{\mathbf{T}}$ is described as follows. The definition of

a second order tensor in a space of three dimensions $\mathbf{T}$ is a special case of (A.80) written in the form

$$\mathbf{T} = T_{\mathrm{ij}}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}} = T_{\alpha\beta}\mathbf{e}_{\alpha} \otimes \mathbf{e}_{\beta} \tag{A.147}$$

or, executing the summation in the Latin system,

$$\begin{aligned}\mathbf{T} = {} & T_{11}\mathbf{e}_1 \otimes \mathbf{e}_1 + T_{22}\mathbf{e}_2 \otimes \mathbf{e}_2 + T_{33}\mathbf{e}_3 \otimes \mathbf{e}_3 + T_{23}(\mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2) \\ & + T_{13}(\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1) + T_{12}(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1).\end{aligned} \tag{A.148}$$

If a new set of base vectors defined by

$$\begin{aligned}&\hat{\mathbf{e}}_1 = \mathbf{e}_1 \otimes \mathbf{e}_1,\ \hat{\mathbf{e}}_2 = \mathbf{e}_2 \otimes \mathbf{e}_2,\ \hat{\mathbf{e}}_3 = \mathbf{e}_3 \otimes \mathbf{e}_3,\ \hat{\mathbf{e}}_4 = \frac{1}{\sqrt{2}}(\mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2) \\ &\hat{\mathbf{e}}_5 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1),\ \hat{\mathbf{e}}_6 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1)\end{aligned} \tag{A.149}$$

is introduced as well as a new set of tensor components defined by

$$\hat{T}_1 = T_{11}, \hat{T}_2 = T_{22}, \hat{T}_3 = T_{33}, \hat{T}_4 = \sqrt{2}T_{23}, \hat{T}_5 = \sqrt{2}T_{13}, \hat{T}_6 = \sqrt{2}T_{12}, \tag{A.150}$$

then (A.148) may be rewritten as

$$\hat{\mathbf{T}} = \hat{T}_1\hat{\mathbf{e}}_1 + \hat{T}_2\hat{\mathbf{e}}_2 + \hat{T}_3\hat{\mathbf{e}}_3 + \hat{T}_4\hat{\mathbf{e}}_4 + \hat{T}_5\hat{\mathbf{e}}_5 + \hat{T}_6\hat{\mathbf{e}}_6, \tag{A.151}$$

or

$$\hat{\mathbf{T}} = \hat{T}_{\mathrm{i}}\hat{\mathbf{e}}_{\mathrm{i}} = \hat{T}_{\alpha}\hat{\mathbf{e}}_{\alpha}, \tag{A.152}$$

which is the definition of a vector in six dimensions. This establishes the one-to-one connection between the components of the symmetric second order tensors $\mathbf{T}$ and the six-dimensional vector $\hat{\mathbf{T}}$.

The second point to be made is that fourth order tensors in three dimensions, with certain symmetries, may also be considered as second order tensors in a six-dimensional space. The one-to-one connection between the components of the fourth order tensors in three dimensions $\mathbf{C}$ and the second order tensors in six dimensions vector $\hat{\mathbf{C}}$ is described as follows. Consider next a fourth order tensor c in three dimensions defined by

$$\mathbf{C} = C_{\mathrm{ijkm}}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}} \otimes \mathbf{e}_{\mathrm{k}} \otimes \mathbf{e}_{\mathrm{m}} = C_{\alpha\beta\gamma\delta}\mathbf{e}_{\alpha} \otimes \mathbf{e}_{\beta} \otimes \mathbf{e}_{\gamma} \otimes \mathbf{e}_{\delta}, \tag{A.153}$$

and having symmetry in its first and second pair of indices, $C_{\mathrm{ijkm}} = C_{\mathrm{jikm}}$ and $C_{\mathrm{ijkm}} = C_{\mathrm{ijmk}}$, but not another symmetry in its indices; in particular C_{ijkm} is not

equal to C_{kmij}, in general. The results of interest are for fourth order tensors in three dimensions with these particular symmetries because it is these fourth order tensors that linearly relate two symmetric second order tensors in three dimensions. Due to the special indicial symmetries just described, the change of basis (A.149) may be introduced in (A.153) and may be rewritten as

$$\hat{\mathbf{C}} = \hat{\mathbf{C}}_{\mathrm{ij}}\hat{\mathbf{e}}_{\mathrm{i}} \otimes \hat{\mathbf{e}}_{\mathrm{j}} = \hat{\mathbf{C}}_{\alpha\beta}\hat{\mathbf{e}}_{\alpha} \otimes \hat{\mathbf{e}}_{\beta}, \tag{A.154}$$

where the 36 components of c_{ijkm}, the fourth order tensor in three dimensions (with the symmetries $c_{\mathrm{ijkm}} = c_{\mathrm{jikm}}$ and $c_{\mathrm{ijkm}} = c_{\mathrm{ijmk}}$) are related to the 36 components of $\hat{c}_{\mathrm{ij}}$, the second order tensor in six dimensions by

$$\begin{aligned}
&\hat{c}_{11} = c_{1111}, \hat{c}_{22} = c_{2222}, \hat{c}_{33} = c_{3333}, \hat{c}_{23} = c_{2233}, \hat{c}_{32} = c_{3322},\\
&\hat{c}_{13} = c_{1133}, \hat{c}_{31} = c_{3311}, \hat{c}_{12} = c_{1122}, \hat{c}_{21} = c_{2211},\\
&\hat{c}_{44} = 2c_{2323}, \hat{c}_{55} = 2c_{1313}, \hat{c}_{66} = 2c_{1212}, \hat{c}_{45} = 2c_{2313}, \hat{c}_{54} = 2c_{1323},\\
&\hat{c}_{46} = 2c_{2312}, \hat{c}_{64} = 2c_{1223}, \hat{c}_{56} = 2c_{1312}, \hat{c}_{65} = 2c_{1213},\\
&\hat{c}_{41} = \sqrt{2}c_{2311}, \hat{c}_{14} = \sqrt{2}c_{1123}, \hat{c}_{51} = \sqrt{2}c_{1311}, \hat{c}_{15} = \sqrt{2}c_{1113},\\
&\hat{c}_{61} = \sqrt{2}c_{1211}, \hat{c}_{16} = \sqrt{2}c_{1112}, \hat{c}_{42} = \sqrt{2}c_{2322}, \hat{c}_{24} = \sqrt{2}c_{2223},\\
&\hat{c}_{52} = \sqrt{2}c_{1322}, \hat{c}_{25} = \sqrt{2}c_{2213}, \hat{c}_{62} = \sqrt{2}c_{1222}, \hat{c}_{26} = \sqrt{2}c_{2212},\\
&\hat{c}_{43} = \sqrt{2}c_{2333}, \hat{c}_{34} = \sqrt{2}c_{3323}, \hat{c}_{53} = \sqrt{2}c_{1333}, \hat{c}_{35} = \sqrt{2}c_{3313}, .\\
&\hat{c}_{63} = \sqrt{2}c_{1233}, \hat{c}_{36} = \sqrt{2}c_{3312}.
\end{aligned} \tag{A.155}$$

Using the symmetry of the second order tensors, $\mathbf{T} = \mathbf{T}^{\mathrm{T}}$ and $\mathbf{J} = \mathbf{J}^{\mathrm{T}}$, as well as the two indicial symmetries of c_{ijkm}, the linear relationship between $\mathbf{T}$ and $\mathbf{J}$,

$$T_{\mathrm{ij}} = C_{\mathrm{ijkm}}J_{\mathrm{km}}, \tag{A.156}$$

may be expanded to read

$$\begin{aligned}
T_{11} &= c_{1111}J_{11} + c_{1122}J_{22} + c_{1133}J_{33} + 2c_{1123}J_{23} + 2c_{1113}J_{13} + 2c_{1112}J_{12},\\
T_{22} &= c_{2211}J_{11} + c_{2222}J_{22} + c_{2233}J_{33} + 2c_{2223}J_{23} + 2c_{2213}J_{13} + 2c_{2212}J_{12},\\
T_{33} &= c_{3311}J_{11} + c_{3322}J_{22} + c_{3333}J_{33} + 2c_{3323}J_{23} + 2c_{3313}J_{13} + 2c_{3312}J_{12},\\
T_{23} &= c_{2311}J_{11} + c_{2322}J_{22} + c_{2333}J_{33} + 2c_{2323}J_{23} + 2c_{2313}J_{13} + 2c_{2312}J_{12}\\
T_{13} &= c_{1311}J_{11} + c_{1322}J_{22} + c_{1333}J_{33} + 2c_{1323}J_{23} + 2c_{1313}J_{13} + 2c_{1312}J_{12},\\
T_{12} &= c_{1211}J_{11} + c_{1222}J_{22} + c_{1233}J_{33} + 2c_{1223}J_{23} + 2c_{1213}J_{13} + 2c_{1212}J_{12}.
\end{aligned} \tag{A.157}$$

The corresponding linear relationship between $\hat{\mathbf{T}}$ and $\hat{\mathbf{J}}$,

$$\hat{T}_{\mathrm{i}} = \hat{c}_{\mathrm{ij}}\hat{J}_{\mathrm{j}}, \tag{A.158}$$

may be expanded to read

$$\begin{aligned}
\hat{T}_1 &= \hat{c}_{11}\hat{J}_1 + \hat{c}_{12}\hat{J}_2 + \hat{c}_{13}\hat{J}_3 + \hat{c}_{14}\hat{J}_4 + \hat{c}_{15}\hat{J}_5 + \hat{c}_{16}\hat{J}_6,\\
\hat{T}_2 &= \hat{c}_{21}\hat{J}_1 + \hat{c}_{22}\hat{J}_2 + \hat{c}_{23}\hat{J}_3 + \hat{c}_{24}\hat{J}_4 + \hat{c}_{25}\hat{J}_5 + \hat{c}_{26}\hat{J}_6,\\
\hat{T}_3 &= \hat{c}_{31}\hat{J}_1 + \hat{c}_{32}\hat{J}_2 + \hat{c}_{33}\hat{J}_3 + \hat{c}_{34}\hat{J}_4 + \hat{c}_{35}\hat{J}_5 + \hat{c}_{36}\hat{J}_6,\\
\hat{T}_4 &= \hat{c}_{41}\hat{J}_1 + \hat{c}_{42}\hat{J}_2 + \hat{c}_{43}\hat{J}_3 + \hat{c}_{44}\hat{J}_4 + \hat{c}_{45}\hat{J}_5 + \hat{c}_{46}\hat{J}_6,\\
\hat{T}_5 &= \hat{c}_{51}\hat{J}_1 + \hat{c}_{52}\hat{J}_2 + \hat{c}_{53}\hat{J}_3 + \hat{c}_{54}\hat{J}_4 + \hat{c}_{55}\hat{J}_5 + \hat{c}_{56}\hat{J}_6,\\
\hat{T}_6 &= \hat{c}_{61}\hat{J}_1 + \hat{c}_{62}\hat{J}_2 + \hat{c}_{63}\hat{J}_3 + \hat{c}_{64}\hat{J}_4 + \hat{c}_{65}\hat{J}_5 + \hat{c}_{66}\hat{J}_6.
\end{aligned} \tag{A.159}$$

The advantage to the notation (A.158) or (A.159) as opposed to the notation (A.156) or (A.157) is that there is no matrix representation of (A.156) or (A.157) that retains the tensorial character while there is a simple, direct, and familiar tensorial representation of (A.158) or (A.159). The equations (A.158) or (A.159) may be written in matrix notation as the linear transformation

$$\hat{\mathbf{T}} = \hat{\mathbf{C}} \cdot \hat{\mathbf{J}}. \tag{A.160}$$

Recalling the rule for the transformation for the components of vectors in a coordinate transformation, (A.73), the transformation rule for $\hat{\mathbf{T}}$ or $\hat{\mathbf{J}}$ may be written down by inspection,

$$\hat{\mathbf{J}}^{(\mathrm{L})} = \hat{\mathbf{Q}} \cdot \hat{\mathbf{J}}^{(\mathrm{G})} \text{ and } \hat{\mathbf{J}}^{(\mathrm{G})} = \hat{\mathbf{Q}}^{\mathrm{T}} \cdot \hat{\mathbf{J}}^{(\mathrm{L})}. \tag{A.161}$$

Furthermore, using the result (A.77), the second order tensor $\hat{\mathbf{C}}$ in the space of six dimensions transforms according to the rule

$$\hat{\mathbf{C}}^{(\mathrm{L})} = \hat{\mathbf{Q}} \cdot \hat{\mathbf{C}}^{(\mathrm{G})} \cdot \hat{\mathbf{Q}}^{\mathrm{T}} \text{ and } \hat{\mathbf{C}}^{(\mathrm{G})} = \hat{\mathbf{Q}}^{\mathrm{T}} \cdot \hat{\mathbf{C}}^{(\mathrm{L})} \cdot \hat{\mathbf{Q}}. \tag{A.162}$$

In short, the second order tensor $\hat{\mathbf{C}}$ in the space of six dimensions may be treated exactly like a second order tensor in the space of three dimensions as far as the usual tensorial operations are concerned.

The relationship between components of the second order tensor in three dimensions and the vector in 6 dimensions contained in (A.150) may be written in n-tuple notation for $\mathbf{T}$ and $\mathbf{J}$ ($\hat{\mathbf{T}}$ and $\hat{\mathbf{J}}$),

$$\hat{\mathbf{T}} = [\hat{T}_1, \hat{T}_2, \hat{T}_3, \hat{T}_4, \hat{T}_5, \hat{T}_6]^{\mathrm{T}} = \left[T_{11}, T_{22}, T_{33}, \sqrt{2}T_{23}, \sqrt{2}T_{13}, \sqrt{2}T_{12}\right]^{\mathrm{T}},$$
$$\hat{\mathbf{J}} = [\hat{J}_1, \hat{J}_2, \hat{J}_3, \hat{J}_4, \hat{J}_5, \hat{J}_6]^{\mathrm{T}} = \left[J_{11}, J_{22}, J_{33}, \sqrt{2}J_{23}, \sqrt{2}J_{13}, \sqrt{2}J_{12}\right]^{\mathrm{T}}. \tag{A.163}$$

These formulas permit the conversion of three-dimensional second order tensor components directly to six-dimensional vector components and vice versa. The $\sqrt{2}$ factor that multiplies the last three components of the definition of the six-dimensional vector representation of the three-dimensional second order tensor, (A.150), assures the scalar product of the two six-dimensional vectors is equal to the trace of the product of the corresponding second order tensors,

$$\hat{\mathbf{T}} \cdot \hat{\mathbf{J}} = \mathbf{T} : \mathbf{J}. \tag{A.164}$$

The colon or double dot notation between the two second order tensors illustrated in (A.164) is an extension of the single dot notation between the matrices, $\mathbf{A}\cdot\mathbf{B}$, and indicates that one index from $\mathbf{A}$ and one index from $\mathbf{B}$ are to be summed over; the double dot notation between the matrices, $\mathbf{A}$:$\mathbf{B}$, indicates that both indices of $\mathbf{A}$ are to be summed with different indices from $\mathbf{B}$. As the notation above indicates, the effect is the same as the trace of the product, $\mathbf{A}$:$\mathbf{B} = \mathrm{tr}(\mathbf{A}\cdot\mathbf{B})$. Note that $\mathbf{A}$:$\mathbf{B} = \mathbf{A}^{\mathrm{T}}$:$\mathbf{B}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}$:$\mathbf{B} = \mathbf{A}$:$\mathbf{B}^{\mathrm{T}}$ but that $\mathbf{A}$:$\mathbf{B} \neq \mathbf{A}^{\mathrm{T}}$:$\mathbf{B}$ in general. This notation is applicable for square matrices in any dimensional space.

The vector $\hat{\mathbf{U}} = [1, 1, 1, 0, 0, 0]^{\mathrm{T}}$ is introduced to be the six-dimensional vector representation of the three-dimensional unit tensor $\mathbf{1}$. It is important to note that the symbol $\hat{\mathbf{U}}$ is distinct from the unit tensor in six dimensions that is denoted by $\hat{1}$. Note that $\hat{\mathbf{U}} \cdot \hat{\mathbf{U}} = 3, \hat{\mathbf{U}} \cdot \hat{\mathbf{T}} = \mathrm{tr}\mathbf{T}$ and, using (A.164), it is easy to verify that $\hat{\mathbf{T}} \cdot \hat{\mathbf{U}} = \mathbf{T} : \mathbf{1} = \mathrm{tr}\mathbf{T}$. The matrix $\hat{\mathbf{C}}$ dotted with $\hat{\mathbf{U}}$ yields a vector in six dimensions

$$\hat{\mathbf{C}} \cdot \hat{\mathbf{U}} = \begin{bmatrix} \hat{c}_{11} + \hat{c}_{12} + \hat{c}_{13} \\ \hat{c}_{21} + \hat{c}_{22} + \hat{c}_{23} \\ \hat{c}_{31} + \hat{c}_{32} + \hat{c}_{33} \\ \hat{c}_{41} + \hat{c}_{42} + \hat{c}_{43} \\ \hat{c}_{51} + \hat{c}_{52} + \hat{c}_{53} \\ \hat{c}_{61} + \hat{c}_{62} + \hat{c}_{63} \end{bmatrix}, \tag{A.165}$$

and, dotting again with $\hat{\mathbf{U}}$, a scalar is obtained:

$$\hat{\mathbf{U}} \cdot \hat{\mathbf{C}} \cdot \hat{\mathbf{U}} = \hat{c}_{11} + \hat{c}_{12} + \hat{c}_{13} + \hat{c}_{21} + \hat{c}_{22} + \hat{c}_{23} + \hat{c}_{31} + \hat{c}_{32} + \hat{c}_{33}. \tag{A.166}$$

The transformations rules (A.161) and (A.162) for the vector and second order tensors in six dimensions involve the six-dimensional orthogonal tensor transformation $\hat{\mathbf{Q}}$. The tensor components of $\hat{\mathbf{Q}}$ are given in terms of $\mathbf{Q}$ by

$$\hat{\mathbf{Q}} = \begin{bmatrix} \hat{Q}_{1\mathrm{I}} & \hat{Q}_{1\mathrm{II}} & \hat{Q}_{1\mathrm{III}} & \hat{Q}_{1\mathrm{IV}} & \hat{Q}_{1\mathrm{V}} & \hat{Q}_{1\mathrm{VI}} \\ \hat{Q}_{2\mathrm{I}} & \hat{Q}_{2\mathrm{II}} & \hat{Q}_{2\mathrm{III}} & \hat{Q}_{2\mathrm{IV}} & \hat{Q}_{2\mathrm{V}} & \hat{Q}_{2\mathrm{VI}} \\ \hat{Q}_{3\mathrm{I}} & \hat{Q}_{3\mathrm{II}} & \hat{Q}_{3\mathrm{III}} & \hat{Q}_{3\mathrm{IV}} & \hat{Q}_{3\mathrm{V}} & \hat{Q}_{3\mathrm{VI}} \\ \hat{Q}_{4\mathrm{I}} & \hat{Q}_{4\mathrm{II}} & \hat{Q}_{4\mathrm{III}} & \hat{Q}_{4\mathrm{IV}} & \hat{Q}_{4\mathrm{V}} & \hat{Q}_{4\mathrm{VI}} \\ \hat{Q}_{5\mathrm{I}} & \hat{Q}_{5\mathrm{II}} & \hat{Q}_{5\mathrm{III}} & \hat{Q}_{5\mathrm{IV}} & \hat{Q}_{5\mathrm{V}} & \hat{Q}_{5\mathrm{VI}} \\ \hat{Q}_{6\mathrm{I}} & \hat{Q}_{6\mathrm{II}} & \hat{Q}_{6\mathrm{III}} & \hat{Q}_{6\mathrm{IV}} & \hat{Q}_{6\mathrm{V}} & \hat{Q}_{6\mathrm{VI}} \end{bmatrix}$$

$$= \begin{bmatrix} Q_{1\mathrm{I}}^2 & Q_{1\mathrm{II}}^2 & Q_{1\mathrm{III}}^2 & \sqrt{2}Q_{1\mathrm{II}}Q_{1\mathrm{III}} & \sqrt{2}Q_{1\mathrm{I}}Q_{1\mathrm{III}} & \sqrt{2}Q_{1\mathrm{I}}Q_{1\mathrm{II}} \\ Q_{2\mathrm{I}}^2 & Q_{2\mathrm{II}}^2 & Q_{2\mathrm{III}}^2 & \sqrt{2}Q_{2\mathrm{II}}Q_{2\mathrm{III}} & \sqrt{2}Q_{2\mathrm{I}}Q_{2\mathrm{III}} & \sqrt{2}Q_{2\mathrm{I}}Q_{2\mathrm{II}} \\ Q_{3\mathrm{I}}^2 & Q_{3\mathrm{II}}^2 & Q_{3\mathrm{III}}^2 & \sqrt{2}Q_{3\mathrm{II}}Q_{3\mathrm{III}} & \sqrt{2}Q_{3\mathrm{I}}Q_{3\mathrm{III}} & \sqrt{2}Q_{3\mathrm{I}}Q_{3\mathrm{II}} \\ \sqrt{2}Q_{2\mathrm{I}}Q_{3\mathrm{I}} & \sqrt{2}Q_{2\mathrm{II}}Q_{3\mathrm{II}} & \sqrt{2}Q_{2\mathrm{III}}Q_{3\mathrm{III}} & Q_{2\mathrm{II}}Q_{3\mathrm{III}} + Q_{3\mathrm{II}}Q_{2\mathrm{III}} & Q_{2\mathrm{I}}Q_{3\mathrm{III}} + Q_{3\mathrm{I}}Q_{2\mathrm{III}} & Q_{2\mathrm{I}}Q_{3\mathrm{II}} + Q_{3\mathrm{I}}Q_{2\mathrm{II}} \\ \sqrt{2}Q_{1\mathrm{I}}Q_{3\mathrm{I}} & \sqrt{2}Q_{1\mathrm{II}}Q_{3\mathrm{II}} & \sqrt{2}Q_{1\mathrm{III}}Q_{3\mathrm{III}} & Q_{1\mathrm{II}}Q_{3\mathrm{III}} + Q_{3\mathrm{II}}Q_{1\mathrm{III}} & Q_{1\mathrm{I}}Q_{3\mathrm{III}} + Q_{3\mathrm{I}}Q_{1\mathrm{III}} & Q_{1\mathrm{I}}Q_{3\mathrm{II}} + Q_{3\mathrm{I}}Q_{1\mathrm{II}} \\ \sqrt{2}Q_{1\mathrm{I}}Q_{2\mathrm{I}} & \sqrt{2}Q_{1\mathrm{II}}Q_{2\mathrm{II}} & \sqrt{2}Q_{1\mathrm{III}}Q_{2\mathrm{III}} & Q_{1\mathrm{II}}Q_{2\mathrm{III}} + Q_{2\mathrm{II}}Q_{1\mathrm{III}} & Q_{1\mathrm{I}}Q_{2\mathrm{III}} + Q_{2\mathrm{I}}Q_{1\mathrm{III}} & Q_{1\mathrm{I}}Q_{2\mathrm{II}} + Q_{1\mathrm{II}}Q_{2\mathrm{I}} \end{bmatrix} \tag{A.167}$$

To see that $\hat{\mathbf{Q}}$ is an orthogonal matrix in six dimensions requires some algebraic manipulation. The proof rests on the orthogonality of the three-dimensional $\mathbf{Q}$:

$$\mathbf{Q} \cdot \mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{Q} = \mathbf{1}, \Rightarrow \hat{\mathbf{Q}} \cdot \hat{\mathbf{Q}}^{\mathrm{T}} = \hat{\mathbf{Q}}^{\mathrm{T}} \cdot \hat{\mathbf{Q}} = \hat{\mathbf{1}}. \tag{A.168}$$

In the special case when $\mathbf{Q}$ is given by

$$\mathbf{Q} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{A.169}$$

$\hat{\mathbf{Q}}$ has the representation

$$\hat{\mathbf{Q}} = \begin{bmatrix} \cos^2\alpha & \sin^2\alpha & 0 & 0 & 0 & -\sqrt{2}\cos\alpha\sin\alpha \\ \sin^2\alpha & \cos^2\alpha & 0 & 0 & 0 & \sqrt{2}\cos\alpha\sin\alpha \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & 0 & 0 & -\sin\alpha & \cos\alpha & 0 \\ \sqrt{2}\cos\alpha\sin\alpha & -\sqrt{2}\cos\alpha\sin\alpha & 0 & 0 & 0 & \cos^2\alpha - \sin^2\alpha \end{bmatrix} \tag{A.170}$$

It should be noted that while it is always possible to find $\hat{\mathbf{Q}}$ given $\mathbf{Q}$ by use of (A.167), it is not possible to determine $\mathbf{Q}$ unambiguously given $\hat{\mathbf{Q}}$. Although $\hat{\mathbf{Q}}$ is uniquely determined by $\mathbf{Q}$, the reverse process of finding a $\mathbf{Q}$ given $\hat{\mathbf{Q}}$ is not unique in that there will be a choice of sign necessary in the reverse process. To see this nonuniqueness note that both $\mathbf{Q} = \mathbf{1}$ and $\mathbf{Q} = -\mathbf{1}$ correspond to $\hat{\mathbf{Q}} = \hat{1}$. There are 9 components of $\mathbf{Q}$ that satisfy 6 conditions given by $(\mathrm{A.168})_1$. There are therefore only three independent components of $\mathbf{Q}$. However, there are 36 components of $\hat{\mathbf{Q}}$ that satisfy the 21 conditions given by $(\mathrm{A.168})_2$ and hence 15 independent

components of $\hat{\mathbf{Q}}$. Thus, while (A.167) uniquely determines $\hat{\mathbf{Q}}$ given $\mathbf{Q}$, the components of $\hat{\mathbf{Q}}$ must be considerably restricted in order to determine $\mathbf{Q}$ given $\hat{\mathbf{Q}}$, and selections of signs must be made.

Problems

A.11.1 Construct the six-dimensional vector $\hat{\mathbf{T}}$ that corresponds to the three-dimensional tensor $\mathbf{T}$ given by

$$\mathbf{T} = \frac{1}{2}\begin{bmatrix} 13 & 3\sqrt{3} & \sqrt{3} \\ 3\sqrt{3} & 7 & 1 \\ \sqrt{3} & 1 & 8 \end{bmatrix}.$$

A.11.2 Prove that the relationship $\hat{\mathbf{T}} \cdot \hat{\mathbf{J}} = \mathbf{T} : \mathbf{J}$ (A.164) is correct by substitution of components.

A.11.3 Construct the six-dimensional orthogonal transformation $\hat{\mathbf{Q}}$ that corresponds to the three-dimensional orthogonal transformation $\mathbf{Q}$ where

$$\mathbf{Q} = \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix}.$$

A.11.4 Construct the six-dimensional orthogonal transformation $\hat{\mathbf{Q}}$ that corresponds to the three-dimensional orthogonal transformation $\mathbf{Q}$ where

$$\mathbf{Q} = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{\frac{3}{2}} & \sqrt{\frac{3}{2}} \\ -\sqrt{3} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\sqrt{2} & \sqrt{2} \end{bmatrix}.$$

A.12 The Gradient Operator and the Divergence Theorem

The vectors and tensors introduced are all considered as functions of coordinate positions x_1, x_2, x_3, and time t. In the case of a vector or tensor this dependence is written $\mathbf{r}(x_1, x_2, x_3, t)$ or $\mathbf{T}(x_1, x_2, x_3, t)$, which means that each element of the vector $\mathbf{r}$ or the tensor $\mathbf{T}$ is a function of x_1, x_2, x_3, and t,. The gradient operator is denoted by Δ and defined, in three dimensions, by

$$\nabla = \frac{\partial}{\partial x_1}\mathbf{e}_1 + \frac{\partial}{\partial x_2}\mathbf{e}_2 + \frac{\partial}{\partial x_3}\mathbf{e}_3. \tag{A.171}$$

This operator is called a vector operator because it increases the tensorial order of the quantity operated upon by one. For example, the gradient of a scalar function $f(x_1, x_2, x_3, t)$ is a vector given by

$$\nabla f(x_1, x_2, x_3, t) = \frac{\partial f}{\partial x_1}\mathbf{e}_1 + \frac{\partial f}{\partial x_2}\mathbf{e}_2 + \frac{\partial f}{\partial x_3}\mathbf{e}_3. \tag{A.172}$$

To verify that the gradient operator transforms as a vector consider the operator in both the Latin and Greek coordinate systems,

$$\nabla^{(\mathrm{L})} f(\mathbf{x}^{(\mathrm{L})}, t) = \frac{\partial f}{\partial x_{\mathrm{i}}}\mathbf{e}_{\mathrm{i}} \text{ and } \nabla^{(\mathrm{G})} f(\mathbf{x}^{(\mathrm{G})}, t) = \frac{\partial f}{\partial x_{\alpha}}\mathbf{e}_{\alpha}, \tag{A.173}$$

respectively, and note that, by the chain rule of partial differentiation,

$$\frac{\partial f}{\partial x_{\mathrm{i}}} = \frac{\partial f}{\partial x_{\alpha}}\frac{\partial x_{\alpha}}{\partial x_{\mathrm{i}}}. \tag{A.174}$$

Now since, from (A.77), $\mathbf{x}^{(\mathrm{G})} = \mathbf{Q}^{\mathrm{T}} \cdot \mathbf{x}^{(\mathrm{L})}$, or index notation $x_{\alpha} = Q_{\mathrm{i}\alpha} x_{\mathrm{i}}$ it follows that $Q_{\mathrm{i}\alpha} = \frac{\partial x_{\alpha}}{\partial x_{\mathrm{i}}}$ and, from (A.174), $\dfrac{\partial f}{\partial x_{\mathrm{i}}} = Q_{\mathrm{i}\alpha}\dfrac{\partial f}{\partial x_{\alpha}}$ or

$$\nabla^{(\mathrm{L})} f(\mathbf{x}^{(\mathrm{L})}, t) = \mathbf{Q} \cdot \nabla^{(\mathrm{G})} f(\mathbf{x}^{(\mathrm{G})}, t). \tag{A.175}$$

This shows that the gradient is a vector operator because it transforms like a vector under changes of coordinate systems.

The gradient of a vector function $\mathbf{r}(x_1, x_2, x_3, t)$ is a second order tensor given by

$$\nabla \otimes \mathbf{r}(x_1, x_2, x_3, t) = \frac{\partial r_{\mathrm{j}}}{\partial x_{\mathrm{i}}}\mathbf{e}_{\mathrm{i}} \otimes \mathbf{e}_{\mathrm{j}}, \tag{A.176}$$

where

$$[\nabla \otimes \mathbf{r}]^{\mathrm{T}} = \left[\frac{\partial r_{\mathrm{i}}}{\partial x_{\mathrm{j}}}\right] = \begin{bmatrix} \dfrac{\partial r_1}{\partial x_1} & \dfrac{\partial r_1}{\partial x_2} & \dfrac{\partial r_1}{\partial x_3} \\ \dfrac{\partial r_2}{\partial x_1} & \dfrac{\partial r_2}{\partial x_2} & \dfrac{\partial r_2}{\partial x_3} \\ \dfrac{\partial r_3}{\partial x_1} & \dfrac{\partial r_3}{\partial x_2} & \dfrac{\partial r_3}{\partial x_3} \end{bmatrix}. \tag{A.177}$$

As this example suggests, when the gradient operator is applied, the tensorial order of the quantity operated upon it increases by one. The matrix that is the open product of the gradient and $\mathbf{r}$ is arranged in (A.177) so that the derivative is in the first (or row) position and the vector $\mathbf{r}$ is in the second (or column) position. The divergence operator is a combination of the gradient operator and a contraction operation that results in the reduction of the order of the quantity operated upon to one lower than it

was before the operation. For example the trace of the gradient of a vector function is a scalar called the divergence of the vector, $\mathrm{tr}[\Delta \otimes \mathbf{r}] = \Delta \cdot \mathbf{r} = \mathrm{div}\,\mathbf{r}$,

$$\nabla \cdot \mathbf{r} = \mathrm{div}\,\mathbf{r} = \frac{\partial r_1}{\partial x_1} + \frac{\partial r_2}{\partial x_2} + \frac{\partial r_3}{\partial x_3}. \tag{A.178}$$

The divergence operation is similar to the scalar product of two vectors in that the effect of the operation is to reduce the order of the quantity by two from the sum of the ranks of the combined quantities before the operation. The curl operation is the gradient operator cross product with a vector function $\mathbf{r}(x_1, x_2, x_3, t)$, thus

$$\nabla \times \mathbf{r} = \mathrm{curl}\,\mathbf{r} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ r_1 & r_2 & r_3 \end{vmatrix}. \tag{A.179}$$

A three-dimensional double gradient tensor defined by $\mathbf{O} = \nabla \otimes \nabla$ ($\mathrm{tr}\mathbf{O} = \nabla^2$) and its six-dimensional vector counterpart $\hat{\mathbf{O}}$ ($\hat{\mathbf{O}} \cdot \hat{\mathbf{U}} = \mathrm{tr}\mathbf{O} = \nabla^2$) are often convenient notations to employ. The components of $\hat{\mathbf{O}}$ are

$$\hat{\mathbf{O}} = \left[\frac{\partial^2}{\partial x_1^2}, \frac{\partial^2}{\partial x_2^2}, \frac{\partial^2}{\partial x_3^2}, \sqrt{2}\frac{\partial^2}{\partial x_2 \partial x_3}, \sqrt{2}\frac{\partial^2}{\partial x_1 \partial x_3}, \sqrt{2}\frac{\partial^2}{\partial x_1 \partial x_2}\right]^{\mathrm{T}}, \tag{A.180}$$

and the operation of $\hat{\mathbf{O}}$ on a six-dimensional vector representation of a second order tensor in three dimensions, $\hat{\mathbf{O}} \cdot \hat{\mathbf{T}} = \mathrm{tr}\mathbf{O} \cdot \mathbf{T}$, is given by

$$\hat{\mathbf{O}} \cdot \hat{\mathbf{T}} = \frac{\partial^2 T_{11}}{\partial x_1^2} + \frac{\partial^2 T_{22}}{\partial x_2^2} + \frac{\partial^2 T_{33}}{\partial x_3^2} + 2\frac{\partial^2 T_{23}}{\partial x_2 \partial x_3} + 2\frac{\partial^2 T_{13}}{\partial x_1 \partial x_3} + 2\frac{\partial^2 T_{12}}{\partial x_1 \partial x_2}. \tag{A.181}$$

The divergence of a second order tensor $\mathbf{T}$ is defined in a similar fashion to the divergence of a vector; it is a vector given by

$$\begin{aligned} \nabla \cdot \mathbf{T}(x_1, x_2, x_3, t) = &\left(\frac{\partial T_{11}}{\partial x_1} + \frac{\partial T_{12}}{\partial x_2} + \frac{\partial T_{13}}{\partial x_3}\right)\mathbf{e}_1 + \left(\frac{\partial T_{21}}{\partial x_1} + \frac{\partial T_{22}}{\partial x_2} + \frac{\partial T_{23}}{\partial x_3}\right)\mathbf{e}_2 \\ &+ \left(\frac{\partial T_{31}}{\partial x_1} + \frac{\partial T_{32}}{\partial x_2} + \frac{\partial T_{33}}{\partial x_3}\right)\mathbf{e}_3. \end{aligned} \tag{A.182}$$

The divergence theorem (also called Gauss' theorem, Green's theorem or Ostrogradsky's theorem, depending on the nationality) relates a volume integral to a surface integral over the volume. The divergence of a vector field $\mathbf{r}(x_1, x_2, x_3, t)$ integrated over a volume of space is equal to the integral of the projection of the field $\mathbf{r}(x_1, x_2, x_3, t)$ on the normal to the boundary of the region, evaluated on the boundary of the region, and integrated over the entire boundary

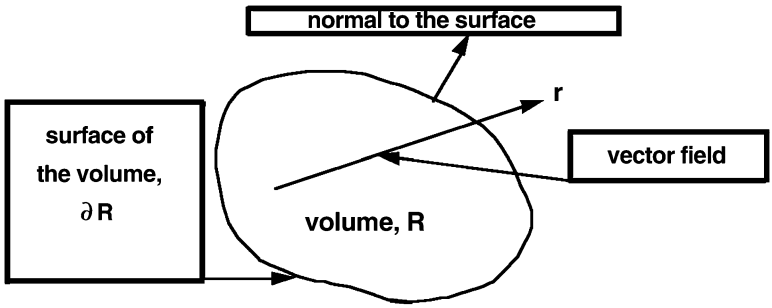


Fig. A.5 An illustration of the geometric elements appearing in the divergence theorem

$$\int_R \nabla \cdot \mathbf{r} \mathrm{d}v = \int_{\partial R} \mathbf{r} \cdot \mathbf{n} \mathrm{d}A, \tag{A.183}$$

where $\mathbf{r}$ represents any vector field, R is a region of three-dimensional space and ∂R is the entire boundary of that region (see Fig. A.5). There are some mathematical restrictions on the validity of (A.183). The vector field $\mathbf{r}(x_1, x_2, x_3, t)$ must be defined and continuously differentiable in the region R. The region R is subject to mathematical restrictions, but any region of interest satisfies these restrictions. For the second order tensor the divergence theorem takes the form

$$\int_R \nabla \cdot \mathbf{T} \mathrm{d}v = \int_{\partial R} \mathbf{T} \cdot \mathbf{n} \mathrm{d}A. \tag{A.184}$$

To show that this version of the theorem is also true if (A.183) is true, the constant vector $\mathbf{c}$ is introduced and used with the tensor field $\mathbf{T}(x_1, x_2, x_3, t)$ to form a vector function field $\mathbf{r}(x_1, x_2, x_3, t)$, thus

$$\mathbf{r} = \mathbf{c} \cdot \mathbf{T}(x_1, x_2, x_3, t). \tag{A.185}$$

Substitution of (A.185) into (A.183) for $\mathbf{r}$ yields

$$\int_R \nabla \cdot (\mathbf{c} \cdot \mathbf{T}) \, \mathrm{d}v = \int_{\partial R} \mathbf{c} \cdot \mathbf{T} \cdot \mathbf{n} \mathrm{d}A, \tag{A.186}$$

and, since $\mathbf{c}$ is a constant vector, (A.186) may be rewritten as

$$\mathbf{c} \cdot \left\{ \int_R \nabla \cdot (\mathbf{T}) \, \mathrm{d}v - \int_{\partial R} \mathbf{T} \cdot \mathbf{n} \mathrm{d}A \right\} = 0. \tag{A.187}$$

This result must hold for all constant vectors **c**, and the divergence theorem for the second order tensor, (A.184), follows.

Stokes theorem relates line integrals to surface integrals,

$$\int_A \nabla \times \mathbf{v} \cdot \mathrm{d}A = \oint_{\partial A} \mathbf{v} \cdot \mathrm{d}\mathbf{x}, \tag{A.188}$$

specifically, it relates the surface integral of the curl of a vector field **v** over a surface A in Euclidean three-space to the line integral of the vector field over its boundary, ∂A. The closed curve of the line integral on ∂A must have positive orientation, such that d**x** points counterclockwise when the surface normal to ∂A points toward the viewer, following the right-hand rule. Note that if **v** is the gradient of a scalar function ϕ, $\mathbf{v} = \nabla\phi$, then (A.188) reduces to

$$0 = \oint_{\partial A} \mathbf{v} \cdot \mathrm{d}\mathbf{x}, \tag{A.189}$$

for all closed paths since $\nabla \times \nabla\phi = 0$, that is to say that the curl of the gradient is zero. It follows that when the vector field **v** satisfies the condition all closed paths, it may be represented as the gradient of a potential, $\mathbf{v} = \nabla\phi$. From (A.179) one can see that $\nabla \times \mathbf{v} = 0$ implies the three conditions

$$\frac{\partial v_1}{\partial x_2} = \frac{\partial v_2}{\partial x_1}, \frac{\partial v_1}{\partial x_3} = \frac{\partial v_3}{\partial x_1}, \frac{\partial v_3}{\partial x_2} = \frac{\partial v_2}{\partial x_3} \text{ or } \frac{\partial v_\mathrm{i}}{\partial x_\mathrm{j}} = \frac{\partial v_\mathrm{j}}{\partial x_\mathrm{i}}, \tag{A.190}$$

which are the conditions that insure that $\mathbf{v} \cdot \mathrm{d}\mathbf{x}$ be an exact differential. Note that these there conditions are equivalent to the requirement that the tensor $\nabla \otimes \mathbf{v}$ be symmetric,

$$\nabla \otimes \mathbf{v} = (\nabla \otimes \mathbf{v})^{\mathrm{T}}. \tag{A.191}$$

We return to this topic in the next section where exact differentials are considered.

Problems

A.12.1 Calculate the gradient of the function $f = (x_1)^2(x_2)^2(x_3)^3$ and evaluate the gradient at the point (1, 2, 3).

A.12.2 Calculate the gradient of a vector function $\mathbf{r}\ (x_1, x_2, x_3) = [(x_1)^2(x_2)^2, x_1x_2, (x_3)^2]$.

A.12.3 Calculate the divergence of a vector function $\mathbf{r}\ (x_1, x_2, x_3) = [(x_1)^2(x_2)^2, x_1x_2, (x_3)^2]$.

A.12.4 Calculate the curl of a vector function $\mathbf{r}\ (x_1, x_2, x_3) = [(x_1)^2(x_2)^2, x_1x_2, (x_3)^2]$.

A.12.5 Calculate the gradient of a vector function $\mathbf{r}(x_1, x_2, x_3) = [(x_2)^2 + (x_3)^2, (x_1)^2 + (x_3)^2, (x_1)^2 + (x_2)^2]$.

A.12.6 Calculate the divergence of a vector function $\mathbf{r}(x_1, x_2, x_3) = [(x_2)^2 + (x_3)^2, (x_1)^2 + (x_3)^2, (x_1)^2 + (x_2)^2]$.

A.12.7 Calculate the curl of a vector function $\mathbf{r}(x_1, x_2, x_3) = [(x_2)^2 + (x_3)^2, (x_1)^2 + (x_3)^2, (x_1)^2 + (x_2)^2]$.

A.12.8 Calculate the divergence of a vector function $\mathbf{r}(x_1, x_2, x_3) = \mathbf{a} \times \mathbf{x}$, where $\mathbf{a}$ is a constant vector.

A.12.9 If $\mathbf{v} = \mathbf{a} \times \mathbf{x}$ and $\mathbf{a}$ is a constant vector, using the indicial notation, evaluate the div $\mathbf{v}$ and the curl $\mathbf{v}$.

A.12.10 Express the integral over a closed surface S, $\int \nabla(\mathbf{x} \cdot \mathbf{x}) \cdot \mathbf{n} dS$, in terms of the total volume V enclosed by the surface S.

A.13 The Index Notation: Review and Summary

The purpose of this section is to gather together most of the results of the previous sections of this appendix relating to the indicial notation and its use. The important concepts and definitions associated with the indicial notation were presented in the earlier sections of this appendix in the places where they were necessary for the logical development of those subjects, a process which left the material relative to the indicial notation distributed amongst many locations. This section contains nothing new other than the gathering together of these results related to the indicial notation as well as the problems related to additional notation culled from the earlier text. The subsection titles below are the important concepts associated with the indicial notation.

The Kronecker Delta

A special symbol, the *Kronecker delta* δ_{ij}, is introduced to represent the components of the unit matrix. When the indices are equal, $i = j$, the value of the Kronecker delta is one, $\delta_{11} = \delta_{22} = \ldots = \delta_{nn} = 1$ and when they are unequal, $i \neq j$, the value of the Kronecker delta is zero, $\delta_{12} = \delta_{21} = \ldots = \delta_{n1} = \delta_{1n} = 0$.

The Einstein Summation Convention

The product of two square matrices, $\mathbf{A}$ and $\mathbf{B}$, with equal numbers of rows (columns) is a square matrix with the same number of rows (columns). The matrix product is written as $\mathbf{A}\cdot\mathbf{B}$ where $\mathbf{A}\cdot\mathbf{B}$ is defined by

$$(\mathbf{A}\cdot\mathbf{B})_{\mathrm{ij}}=\sum_{k=1}^{k=n}A_{\mathrm{ik}}B_{\mathrm{kj}};\tag{A.20}$$

thus, for example, the element in the *r*th row and *c*th column of the product **A**·**B** is given by

$$(\mathbf{A}\cdot\mathbf{B})_{\mathrm{rc}}=A_{\mathrm{r1}}B_{\mathrm{1c}}+A_{\mathrm{r2}}B_{\mathrm{2c}}+\cdots+A_{\mathrm{rn}}B_{\mathrm{nc}}.$$

The widely used notational convention, called the *Einstein summation convention*, allows one to simplify the notation by dropping the summation symbol in (A.20) so that

$$(\mathbf{A}\cdot\mathbf{B})_{\mathrm{ij}}=A_{\mathrm{ik}}B_{\mathrm{kj}},\tag{A.21}$$

where the convention is the understanding that the repeated index, in this case *k*, is to be summed over its *range* of the admissible values from 1 to *n*. For $n = 6$, the range of admissible values is 1–6, including 2, 3, 4, and 5. The two *k* indices are the *summation or dummy* indices; note that the implied summation is unchanged if both of the *k*'*s* are replaced by any other letter of the alphabet.

The Summation Index and Summands

A summation index is defined as an index that occurs in a summand twice and only twice. Note that *summands* are terms in equations separated from each other by plus, minus, or equal signs. The existence of summation indices in a summand requires that the summand be summed with respect to those indices over the entire range of admissible values. Note again that the summation index is only a means of stating that a term must be summed, and the letter used for this index is immaterial, thus $A_{\mathrm{im}}B_{\mathrm{mj}}$ has the same meaning as $A_{\mathrm{ik}}B_{\mathrm{kj}}$ or $A_{\mathrm{k}}B_{\mathrm{k}}$.

The Free Index

The other indices in the formula (A.21), the *i* and *j* indices, are called free indices. A *free index* is free to take on any one of the admissible values in its *range* from 1 to *n*. For example if *n* were 3, the free index could be 1, 2, or 3. A *free index* is formally defined as an index that occurs once and only once in every summand of an equation. The total number of equations that may be represented by an equation with one free index is the range of the admissible values. Thus the equation (A.21) represents n^2 separate equations. For two 2 by 2 matrices *A* and *B*, the product is written as

$$\mathbf{A}\cdot\mathbf{B}=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}=\begin{bmatrix}A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22}\\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22}\end{bmatrix}, \tag{A.22}$$

where, in this case, the products (A.20) and (A.21) stand for the $n^2=2^2=4$ separate equations, the right-hand sides of which are the four elements of the last matrix in (A.22).

The Matrix Transpose Indicial Notation

The dot between the matrix product **A**·**B** indicates that one index from **A** and one index from **B** is to be summed over. The positioning of the summation index on the two matrices involved in a matrix product is critical and is reflected in the matrix notation by the transpose of the matrix. In the three equations below, (A.21), study carefully how the positions of the summation indices within the summation sign change in relation to the position of the transpose on the matrices in the associated matrix product:

$$(\mathbf{A}\cdot\mathbf{B}^{\mathrm{T}})_{\mathrm{ij}}=A_{\mathrm{ik}}B_{\mathrm{jk}},\ (\mathbf{A}^{\mathrm{T}}\cdot\mathbf{B})_{\mathrm{ij}}=A_{\mathrm{ki}}B_{\mathrm{kj}},\ (\mathbf{A}^{\mathrm{T}}\cdot\mathbf{B}^{\mathrm{T}})_{\mathrm{ij}}=A_{\mathrm{ki}}B_{\mathrm{jk}}. \tag{A.23}$$

The Linear Transformation

A system of linear equations representing a *linear transformation*,

$$\begin{aligned} r_1&=A_{11}t_1+A_{12}t_2+\cdots+A_{1\mathrm{n}}t_{\mathrm{n}},\\ r_2&=A_{21}t_1+A_{22}t_2+\cdots+A_{2\mathrm{n}}t_{\mathrm{n}},\\ &\cdots\\ r_{\mathrm{n}}&=A_{\mathrm{n}1}t_1+A_{\mathrm{n}2}t_2+\cdots+A_{\mathrm{nn}}t_{\mathrm{n}} \end{aligned} \tag{A.36}$$

may be contracted horizontally using the summation symbol, thus

$$\begin{aligned} r_1&=A_{1\mathrm{k}}t_{\mathrm{k}},\\ r_2&=A_{2\mathrm{k}}t_{\mathrm{k}},\\ &\cdots\\ r_{\mathrm{n}}&=A_{\mathrm{nk}}t_{\mathrm{k}}. \end{aligned} \tag{A.37}$$

Introduction of the free index convention condenses this system of equations vertically,

$$r_{\mathrm{i}} = A_{\mathrm{ik}} t_{\mathrm{k}}. \tag{A.38}$$

This result may also be represented in the matrix notation as a combination of n-tuples, $\mathbf{r}$ and $\mathbf{t}$, and a square matrix $\mathbf{A}$,

$$\mathbf{r} = \mathbf{A} \cdot \mathbf{t}, \tag{A.39}$$

where the dot between $\mathbf{A}$ and $\mathbf{t}$ indicates that the summation is with respect to one index of $\mathbf{A}$ and one index of $\mathbf{t}$.

The Composition of Linear Transformations

The composition of linear transformations is again a linear transformation. Consider the linear transformation $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$, $\mathbf{u} \rightarrow \mathbf{t}$ (meaning $\mathbf{u}$ is transformed into $\mathbf{t}$) which is combined with the linear transformation (A.39) $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$, $\mathbf{t} \rightarrow \mathbf{r}$ to transform $\mathbf{u} \rightarrow \mathbf{r}$, thus $\mathbf{r} = \mathbf{A}\cdot\mathbf{B}\cdot\mathbf{u}$, and if we let $\mathbf{C} \equiv \mathbf{A}\cdot\mathbf{B}$, then $\mathbf{r} = \mathbf{C}\cdot\mathbf{u}$. The result of the composition of the two linear transformations, $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ and $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$, is then a new linear transformation $\mathbf{r} = \mathbf{C}\cdot\mathbf{u}$ where the square matrix $\mathbf{C}$ is given by the matrix product $\mathbf{A}\cdot\mathbf{B}$. To verify that it is, in fact, a matrix multiplication, the composition of transformations is done again in the indicial notation. The transformation $\mathbf{t} = \mathbf{B}\cdot\mathbf{u}$ in the indicial notation,

$$t_{\mathrm{k}} = B_{\mathrm{km}} u_{\mathrm{m}}, \tag{A.42}$$

is substituted into $\mathbf{r} = \mathbf{A}\cdot\mathbf{t}$ in the indicial notation (A.38),

$$r_{\mathrm{i}} = A_{\mathrm{ik}} B_{\mathrm{km}} u_{\mathrm{m}}, \tag{A.43}$$

which may be rewritten as

$$r_{\mathrm{i}} = C_{\mathrm{im}} u_{\mathrm{m}}, \tag{A.44}$$

where $\mathbf{C}$ is defined by:

$$C_{\mathrm{im}} = A_{\mathrm{ik}} B_{\mathrm{km}}. \tag{A.45}$$

Comparison of (A.45) with (A.20) shows that $\mathbf{C}$ is the matrix product of $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{C} = \mathbf{A}\cdot\mathbf{B}$. The calculation from (A.42) to (A.45) may be repeated using the Einstein summation convention.

Orthogonal Transformations

The matrix $\mathbf{Q} = [Q_{i\alpha}]$ characterizing the change from the Latin orthonormal basis $\mathbf{e}_\iota$ in an N-dimensional vector space to the Greek basis $\mathbf{e}_\alpha$ (or vice versa) is a special type of linear transformation called an orthogonal transformation. Taking the scalar product of $\mathbf{e}_i$ with $\mathbf{e}_j$ where $\mathbf{e}_i$ and $\mathbf{e}_j$ both have the representation (A.62),

$$\mathbf{e}_i = Q_{i\alpha}\mathbf{e}_\alpha \text{ and } \mathbf{e}_j = Q_{j\beta}\mathbf{e}_\beta. \tag{A.66}$$

it follows that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = Q_{i\alpha}Q_{j\beta}\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = Q_{i\alpha}Q_{j\beta}\delta_{\alpha\beta} = Q_{i\alpha}Q_{j\alpha}. \tag{A.67}$$

There are a number of steps in the calculation (A.67) that should be considered carefully. First, the condition of orthonormality of the bases has been used twice, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ and $\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = \delta_{\alpha\beta}$. Second, the transition from the term before the last equal sign to the term after that sign is characterized by a change from a double sum to a single sum over n and the loss of the Kronecker delta $\delta_{\alpha\beta}$. This occurs because the sum over β in the double sum is always zero except in the special case when $\alpha = \beta$ due to the presence of the Kronecker delta $\delta_{\alpha\beta}$. Third, a comparison of the last term in (A.67) with the definition of matrix product (A.20) suggests that it is a matrix product of $\mathbf{Q}$ with itself. However, a careful comparison of the last term in (A.67) with the definition of matrix product (A.20) shows that the summation is over a different index in the second element of the product. In order for the last term in (A.67) to represent a matrix product, the α index should appear as the first subscripted index rather than the second. However, this α index may be relocated in the second matrix by using the transposition operation. Thus the last term in equation (A.67) is the matrix product of $\mathbf{Q}$ with $\mathbf{Q}^T$ as may be seen from the first of equations (A.18). Thus, since the matrix of Kronecker delta components is the unit matrix $\mathbf{1}$, it has been shown that

$$\mathbf{1} = \mathbf{Q} \cdot \mathbf{Q}^T. \tag{A.68}$$

A transformation $\mathbf{Q}$ satisfying (A.68) is to be an *orthogonal transformation*, its inverse is equal to its transpose, $\mathbf{Q}^{-1} = \mathbf{Q}^T$.

Proof of Invariance for the Trace of a Matrix

As an example of the invariance with respect to basis, this property will be derived for $I_\mathbf{A} = \text{tr } \mathbf{A}$. Let $\mathbf{T} = \mathbf{A}$ in (A.86), then set the indices $k = m$ and sum from one to n over the index k, thus

$$A_{kk} = T_{\alpha\beta}Q_{k\alpha}Q_{k\beta} = A_{\alpha\beta}d_{\alpha\beta} = A_{\alpha\alpha}. \tag{A.88}$$

The transition across the second equal sign is a simple rearrangement of terms. The transition across the second equal sign is based on the condition

$$Q_{k\alpha}Q_{k\beta} = \delta_{\alpha\beta} \tag{A.89}$$

which is an alternate form of (A.68), a form equivalent to $\mathbf{Q}^{\mathrm{T}}\cdot\mathbf{Q} = \mathbf{1}$. The transition across the fourth equal sign employs the definition of the Kronecker delta and the summation over β. The result is that the trace of the matrix of second order tensor components relative to any basis is the same number,

$$A_{kk} = A_{\alpha\alpha}. \tag{A.90}$$

The Alternator and the Permutation Symbol

The *alternator* is denoted by e_{ijk} and defined so that it takes on values +1, 0, or −1 according to the rule:

$$e_{ijk} \equiv \left\{ \begin{matrix} +1 & \text{if } P \text{ is an even permuation} \\ 0 & \text{otherwise} \\ -1 & \text{if } P \text{ is an odd permuation} \end{matrix} \right\}, P \equiv \left\{ \begin{matrix} 1 & 2 & 3 \\ i & j & k \end{matrix} \right\}, \tag{A.106}$$

where P is the *permutation symbol* on a set of three objects. The only +1 values of e_{ijk} are e_{123}, e_{231}, and e_{312}. It is easy to verify that 123, 231, and 312 are even permutations of 123. The only −1 values of e_{ijk} are e_{132}, e_{321}, and e_{213}. It is easy to verify that 132, 321, and 213 are odd permutations of 123. The other 21 components of e_{ijk} are all zero because they are neither even nor odd permutations of 123 due to the fact that one number (either 1, 2, or 3) occurs more than once in the indices (e.g., $e_{122} = 0$ since 122 is not a permutation of 123). One mnemonic device for the even permutations of 123 is to write 123123, then read the first set of three digits 123, the second set 231, and the third set 312. The odd permutations may be read off 123123 also by reading from right to left rather than from left to right; reading from the right (but recording them then from the left, as usual) the first set of three digits 321, the second set 213, and the third set 132.

The Alternator and Determinants

The alternator may now be employed to shorten the formula (A.105) for calculating the determinant;

$$e_{mnp}\mathrm{Det}\mathbf{A} = e_{ijk}A_{im}A_{jn}A_{kp} = e_{ijk}A_{mi}A_{nj}A_{pk}. \tag{A.107}$$

This result may be verified by selecting the values of *mnp* to be 123, 231, 312, 132, 321, or 213, then performing the summation over the three indices *i*, *j*, and *k* over 1, 2, and 3 as indicated on the right-hand side of (A.107). In each case the result is the right-hand side of (A.105). It should be noted that (A.107) may be used to show $\mathrm{Det}\mathbf{A} = \mathrm{Det}\mathbf{A}^{\mathrm{T}}$.

The alternator may be used to express the fact that interchanging two rows or two columns of a determinant changes the sign of the determinant,

$$e_{\mathrm{mnp}}\mathrm{Det}A = \begin{vmatrix} A_{1\mathrm{m}} & A_{1\mathrm{n}} & A_{1\mathrm{p}} \\ A_{2\mathrm{m}} & A_{2\mathrm{n}} & A_{2\mathrm{p}} \\ A_{3\mathrm{m}} & A_{3\mathrm{n}} & A_{3\mathrm{p}} \end{vmatrix} = \begin{vmatrix} A_{\mathrm{m}1} & A_{\mathrm{m}2} & A_{\mathrm{m}3} \\ A_{\mathrm{n}1} & A_{\mathrm{n}2} & A_{\mathrm{n}3} \\ A_{\mathrm{p}1} & A_{\mathrm{p}2} & A_{\mathrm{p}3} \end{vmatrix}. \tag{A.108}$$

Using the alternator again may combine these two representations:

$$e_{\mathrm{ijk}}e_{\mathrm{mnp}}\mathrm{Det}\mathbf{A} = \begin{vmatrix} A_{\mathrm{im}} & A_{\mathrm{in}} & A_{\mathrm{ip}} \\ A_{\mathrm{jm}} & A_{\mathrm{jn}} & A_{\mathrm{jp}} \\ A_{\mathrm{km}} & A_{\mathrm{kn}} & A_{\mathrm{kp}} \end{vmatrix}. \tag{A.109}$$

An Important Identity

In the special case when $\mathbf{A} = \mathbf{1}$ ($A_{\mathrm{ij}} = \delta_{\mathrm{ij}}$), an important identity relating the alternator to the Kronecker delta is obtained using (A.109):

$$e_{\mathrm{ijk}}e_{\mathrm{mnp}} = \begin{vmatrix} \delta_{\mathrm{im}} & \delta_{\mathrm{in}} & \delta_{\mathrm{ip}} \\ \delta_{\mathrm{jm}} & \delta_{\mathrm{jn}} & \delta_{\mathrm{jp}} \\ \delta_{\mathrm{kn}} & \delta_{\mathrm{kn}} & \delta_{\mathrm{kp}} \end{vmatrix}. \tag{A.110}$$

The following special cases of (A.110) provide three more very useful relations between the alternator and the Kronecker delta:

$$e_{\mathrm{mnk}}e_{\mathrm{ijk}} = \delta_{\mathrm{im}}\delta_{\mathrm{jn}} - \delta_{\mathrm{jm}}\delta_{\mathrm{in}},\ e_{\mathrm{mjk}}e_{\mathrm{ijk}} = 2\delta_{\mathrm{im}},\ e_{\mathrm{ijk}}e_{\mathrm{ijk}} = 6. \tag{A.111}$$

The first of these relations is obtained by setting the indices *p* and *k* equal in (A.110) and then expanding the determinant. The second is obtained from the first by setting the indices *n* and *j* equal in the first. The third is obtained from the second by setting the indices *i* and *m* equal in the second.

Example A.8.1

Derive the first of (A.111) from (A.110).

Solution: The first of (A.111) is obtained from (A.110) by setting the indices *p* and *k* equal in (A.110) and then expanding the determinant:

$$e_{ijk}e_{mnk} = \begin{vmatrix} \delta_{im} & \delta_{in} & \delta_{ik} \\ \delta_{jm} & \delta_{jn} & \delta_{jk} \\ \delta_{km} & \delta_{kn} & 3 \end{vmatrix},$$

one finds that

$$e_{ijk}e_{mnk} = 3\delta_{im}\delta_{jn} - \delta_{im}\delta_{jk}\delta_{kn} - 3\delta_{in}\delta_{jm} + \delta_{in}\delta_{km}\delta_{jk} + \delta_{ik}\delta_{jm}\delta_{kn} - \delta_{ik}\delta_{km}\delta_{jn}.$$

Carrying out the indicated summation over the index k in the expression above,

$$e_{ijk}e_{mnk} = 3\delta_{im}\delta_{jn} - \delta_{im}\delta_{jn} - 3\delta_{in}\delta_{jm} + \delta_{in}\delta_{jm} + \delta_{in}\delta_{jm} - \delta_{im}\delta_{jn}.$$

This is the desired result, the first of (A.111). □

Example A.8.2

Prove that $\text{Det}(\mathbf{A}\cdot\mathbf{B}) = \text{Det}\mathbf{A}\ \text{Det}\mathbf{B}$.

Solution: Replacing **A** in (A.107) by **C** and selecting the values of *mnp* to be 123, then (A.107) becomes

$$\text{Det}\mathbf{C} = e_{ijk}C_{i1}C_{j2}C_{k3} = e_{ijk}C_{1i}C_{2j}C_{3k}.$$

Now **C** is replaced by the product $\mathbf{A}\cdot\mathbf{B}$ using

$$C_{i1} = A_{im}B_{m1}, C_{j2} = A_{jn}B_{n2}, C_{k3} = A_{kp}B_{p3},$$

thus

$$\text{Det}\mathbf{A}\cdot\mathbf{B} = e_{ijk}A_{im}B_{m1}A_{jn}B_{n2}A_{kp}B_{p3}, \text{ or } \text{Det}\mathbf{A}\cdot\mathbf{B} = (e_{ijk}A_{im}A_{jn}A_{kp})B_{m1}B_{n2}B_{p3},$$

where the order of the terms in the second sum has been rearranged from the first. Comparison of the first four rearranged terms from the second sum with the right-hand side of (A.107) shows that the first four terms in the sum on the right may be replaced by $e_{mnp}\text{Det}\mathbf{A}$; thus applying the first equation of this solution again with **C** replaced by **B**, the desired result is obtained:

$$\text{Det}\mathbf{A}\cdot\mathbf{B} = \text{Det}\mathbf{A}e_{mnp}B_{m1}B_{n2}B_{p3} = \text{Det}\mathbf{A}\ \text{Det}\mathbf{B}.$$ □

The Tensorial Character of the Alternator

Vectors were shown to be characterized by symbols with one subscript, (second rank) tensors were shown to be characterized by symbols with two subscripts; what

is the tensorial character of a symbol with three subscripts; is it a third order tensor? Almost. Tensors are identified on the basis of their tensor transformation law. Recall the tensor transformations laws (A.75) and (A.76) for a vector, (A.86) for a second order tensor and (A.87) for a tensor of order n. An equation that contains a transformation law for the alternator is obtained from (A.107) by replacing $\mathbf{A}$ by the orthogonal transformation $\mathbf{Q}$ given by (A.64) and changing the indices as follows: $m \to \alpha$, $n \to \beta$, $p \to \gamma$, thus

$$e_{\alpha\beta\gamma}\mathrm{Det}\mathbf{Q} = e_{\mathrm{ijk}}Q_{\mathrm{i}\alpha}Q_{\mathrm{j}\alpha}Q_{\mathrm{k}\gamma}. \tag{A.112}$$

This is an unusual transformation law because the determinant of an orthogonal transformation $\mathbf{Q}$ is either +1 or −1. The expected transformation law, on the basis of the tensor transformation laws (A.75) and (A.76) for a vector, (A.86) for a second order tensor and (A.87) for a tensor of order n, is that $\mathrm{Det}\mathbf{Q} = +1$. $\mathrm{Det}\mathbf{Q} = +1$ occurs when the transformation is between coordinate systems of the same handedness (right handed to right handed or left handed to left handed). Recall that a right (left) hand coordinate system or orthonormal basis is one that obeys the right (left) hand rule, that is to say if the curl of your fingers in your right (left) hand fist is in the direction of rotation from the first ordered positive base vector into the second ordered positive base vector, your extended thumb will point in the third ordered positive base vector direction. $\mathrm{Det}\mathbf{Q} = -1$ occurs when the transformation is between coordinate systems of the opposite handedness (left to right or right to left). Since handedness does not play a role in the transformation law for even order tensors, this dependence on the sign of $\mathrm{Det}\mathbf{Q}$ and therefore the relative handedness of the coordinate systems for the alternator transformation law, is unexpected.

The Cross Product of Vectors

In the indicial notation the *vector cross product* $\mathbf{a} \times \mathbf{b}$ is written in terms of an alternator as

$$\mathbf{a} \times \mathbf{b} = e_{\mathrm{ijk}}a_{\mathrm{i}}b_{\mathrm{j}}e_{\mathrm{k}}, \tag{A.114}$$

a result that may be verified by expanding it to show that it coincides with (A.113). If $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ denotes the result of the vector cross product, then from (A.114),

$$\mathbf{c} = e_{\mathrm{ijk}}a_{\mathrm{i}}b_{\mathrm{j}}e_{\mathrm{k}}, (c_{\mathrm{k}} = e_{\mathrm{ijk}}a_{\mathrm{i}}b_{\mathrm{j}}). \tag{A.115}$$

Example A.8.3

Prove that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

Solution: In the formula (A.114) let $i \to j$ and $j \to i$, thus

$$\mathbf{a} \times \mathbf{b} = e_{\mathrm{jik}} a_{\mathrm{j}} b_{\mathrm{i}} e_{\mathrm{k}},$$

Next change e_{jik} to $-e_{\mathrm{ijk}}$ and rearrange the order of a_{j} and b_{i}, then the result is proved:

$$\mathbf{a} \times \mathbf{b} = -e_{\mathrm{ijk}} b_{\mathrm{i}} a_{\mathrm{j}} e_{\mathrm{k}} = -\mathbf{b} \times \mathbf{a}. \qquad \square$$

Scalar Triple Product of Three Vectors

The scalar triple product of three vectors is a scalar formed from three vectors, $\mathbf{a}\cdot(\mathbf{b} \times \mathbf{c})$ and the triple vector product is a vector formed from three vectors, $(\mathbf{r} \times (\mathbf{p} \times \mathbf{q}))$. An expression for the scalar triple product is obtained by taking the dot product of the vector $\mathbf{c}$ with the cross product in the representation (A.114) for $\mathbf{a} \times \mathbf{b}$, thus

$$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = e_{\mathrm{jik}} a_{\mathrm{j}} b_{\mathrm{i}} c_{\mathrm{k}}. \tag{A.116}$$

From the properties of the alternator it follows that

$$\begin{aligned} \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) &= \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = -\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) \\ &= -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}). \end{aligned} \tag{A.117}$$

If the three vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ coincide with the three nonparallel edges of a parallelepiped, the scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is equal to the volume of the parallelepiped. In the following example a useful vector identity for the triple vector product $(\mathbf{r} \times (\mathbf{p} \times \mathbf{q}))$ is derived.

Example A.8.4

Prove that $(\mathbf{r} \times (\mathbf{p} \times \mathbf{q})) = (\mathbf{r}\cdot\mathbf{q})\mathbf{p} - (\mathbf{r}\cdot\mathbf{p})\mathbf{q}$.

Solution: First rewrite (A.114) with the change $\mathbf{a} \rightarrow \mathbf{r}$, and again with the changes $\mathbf{a} \rightarrow \mathbf{p}$ and $\mathbf{b} \rightarrow \mathbf{q}$, where $\mathbf{b} = (\mathbf{p} \times \mathbf{q})$

$$\mathbf{r} \times \mathbf{b} = e_{\mathrm{ijk}} r_{\mathrm{i}} b_{\mathrm{j}} e_{\mathrm{k}}, \;\; \mathbf{b} = \mathbf{p} \times \mathbf{q} = e_{\mathrm{mnj}} p_{\mathrm{m}} q_{\mathrm{n}} e_{\mathrm{j}},$$

Note that the second of these formulas gives the components of $\mathbf{b}$ as

$$b_{\mathrm{j}} = e_{\mathrm{mnj}} p_{\mathrm{m}} q_{\mathrm{n}}.$$

This formula for the components of $\mathbf{b}$ is then substituted into the expression for $(\mathbf{r} \times \mathbf{b}) = (\mathbf{r} \times (\mathbf{p} \times \mathbf{q}))$ above, thus

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = e_{\mathrm{ijk}} e_{\mathrm{mnj}} r_{\mathrm{i}} p_{\mathrm{m}} q_{\mathrm{n}} e_{\mathrm{k}}.$$

On the right-hand side of this expression for $r \times (p \times q)$, e_{ijk} is now changed to $-e_{ikj}$ and the first of (A.111) is then employed,

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = -(\delta_{im}\delta_{kn} - \delta_{in}\delta_{km}) r_i p_m q_n e_k;$$

then summing over k and I,

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = r_i p_k q_i e_k - r_i p_i q_k e_k = (\mathbf{r} \cdot \mathbf{q})\mathbf{p} - (\mathbf{r} \cdot \mathbf{p})\mathbf{q}. \qquad \square$$

Problems

A.3.1 Simplify the following expression by using the Einstein summation index convention for a range of three:

$$\begin{aligned} 0 &= r_1 w_1 + r_2 w_2 + r_3 w_3, \\ \psi &= (u_1 v_1 + u_2 v_2 + u_3 v_3)(u_1 v_1 + u_2 v_2 + u_3 v_3), \\ \phi &= A_{11} x_1^2 + A_{22} x_2^2 + A_{33} x_3^2 + A_{12} x_1 x_2 + A_{21} x_1 x_2 + A_{13} x_1 x_3 + A_{31} x_1 x_3 \\ &\quad + A_{23} x_3 x_2 + A_{32} x_3 x_2. \end{aligned}$$

A.3.3 Prove that $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$.

A.3.9 Show that $(\mathbf{A}\cdot\mathbf{B})^T = \mathbf{B}^T \cdot \mathbf{A}^T$.

A.5.8 If $\mathbf{F}$ is a square matrix and $\mathbf{a}$ is an n-tuple, show that $\mathbf{a}^T \cdot \mathbf{F}^T = \mathbf{F} \cdot \mathbf{a}$.

A.8.2 Show that if $\mathbf{A}$ is a skew-symmetric 3 by 3 matrix, $\mathbf{A} = -\mathbf{A}^T$, then $\mathrm{Det}\mathbf{A} = 0$.

A.8.3 Evaluate $\mathrm{Det}(\mathbf{a} \otimes \mathbf{b})$.

A.8.4 Show $\mathrm{Det}\mathbf{A} = \mathrm{Det}\mathbf{A}^T$.

A.12.9 If $\mathbf{v} = \mathbf{a} \times \mathbf{x}$ and $\mathbf{a}$ is a constant vector, using the indicial notation, evaluate the div $\mathbf{v}$ and the curl $\mathbf{v}$.

A.14 Exact Differentials

In one dimension a differential $\mathrm{d}q = f(x)\,\mathrm{d}x$ is always exact and, in two dimensions, in order that a differential $\mathrm{d}q = \boldsymbol{a}\cdot\mathrm{d}\boldsymbol{x}$ be an exact differential in a simply-connected 2D region R of the x_1, x_2 plane, it is necessary and sufficient that between a_1 and a_2 there exists the relation

$$\frac{\partial a_1}{\partial x_2} = \frac{\partial a_2}{\partial x_1}. \tag{A.192}$$

(Note that the notation often used for this result is $\mathrm{d}q = M(x, y)\,\mathrm{d}x + N(x, y)\,\mathrm{d}y$ leading to the condition $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$). Continuing now with the considerations

associated with Stokes theorem at the end of the section before last, in three dimensions, a differential $\mathrm{d}q = \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{x}$ is an exact differential in a simply-connected 3D region R of the x_1, x_2, x_3 space if between the functions v_1, v_2, *and* v_3 there exist the relations (A.190) or (A.192). If $\mathrm{d}q = \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{x}$ is an exact differential in a simply-connected 3D space, then from (A.189) the integral about any closed path in the space is zero, and furthermore it follows that when the vector field $\mathbf{v}$ satisfies the condition (A.189) all closed paths, it may be represented as the gradient of a potential, $\mathbf{v} = \nabla\phi$.

Having now built the idea of an exact differential from one to three dimensions, we now extend it to six dimensions and the consideration a 6D work differential

$$\mathrm{d}W = \hat{\mathbf{T}} \cdot \mathrm{d}\hat{\mathbf{E}}. \tag{A.193}$$

In Chap. 6 it is shown that, in order that no work can be extracted from an elastic material in a closed path, it is necessary for the work done in all closed paths to be zero, $\oint \hat{\mathbf{T}} \cdot \mathrm{d}\hat{\mathbf{E}} = 0$. (6.14H repeated)

Thus the 6D work differential (A.193) is an exact differential for an elastic material. The conditions parallel to those in 3D, namely (A.190) and (A.191) are that $\hat{\nabla}_{\hat{E}} \otimes \hat{\mathbf{T}}$ be symmetric,

$$\frac{\partial \hat{T}_{\mathrm{i}}}{\partial \hat{E}_{\mathrm{j}}} = \frac{\partial \hat{T}_{\mathrm{j}}}{\partial \hat{E}_{\mathrm{i}}}, \tag{A.194}$$

or

$$\hat{\nabla}_{\hat{E}} \otimes \hat{\mathbf{T}} = \left(\hat{\nabla}_{\hat{E}} \otimes \hat{\mathbf{T}}\right)^{\mathrm{T}}, \tag{A.195}$$

respectively. The parallel to the existence of a potential in 3D is the strain energy U defined by (6.25H) and related to the stress and strain by (6.26H).

Problem

14.1 The conditions for an exact differential are (A.192) in 2D, (A.190) or (A.191) in 3D, and (A.194) or (A.195) in 6D. What are the conditions for an exact differential in 4D if the differential is denoted by $\mathrm{d}q = \boldsymbol{a}\cdot\mathrm{d}\boldsymbol{x}$, $\mathrm{d}q = a_1\mathrm{d}x_1 + a_2\mathrm{d}x_2 + a_3\mathrm{d}x_3 + a_4\mathrm{d}x_4$?

A.15 Tensor Components in Cylindrical Coordinates

In several places use is made of cylindrical coordinates in the solutions to problems in this text. The base vectors in curvilinear coordinates are not generally unit vectors nor do they have the same dimensions. However local Cartesian

coordinates called physical components of the tensors may be constructed if the curvilinear coordinate system is orthogonal. Below, the standard formulas used in this text for cylindrical coordinates are recorded. In the case when cylindrical coordinates are employed, the vectors and tensors introduced are all considered as functions of the coordinate positions r, θ, *and* z in place of the Cartesian coordinates x_1, x_2, and x_3. In the case of a vector or tensor this dependence is written $\mathbf{v}(r, \theta, z, t)$ or $\mathbf{T}(r, \theta, z, t)$, which means that each element of the vector $\mathbf{v}$ or the tensor $\mathbf{T}$ is a function of r, θ, z, and t. The gradient operator is denoted by Δ and defined, in three dimensions, by

$$\nabla = \frac{\partial}{\partial r}\mathbf{e}_{\mathrm{r}} + \frac{1}{r}\frac{\partial}{\partial \theta}\mathbf{e}_{\theta} + \frac{\partial}{\partial z}\mathbf{e}_{\mathrm{z}}, \tag{A.196}$$

where $\mathbf{e}_{\mathrm{r}}$, $\mathbf{e}_{\theta}$, and $\mathbf{e}_{\mathrm{z}}$ are the unit base vectors in the cylindrical coordinate system. The gradient of a scalar function $f(x_1, x_2, x_3, t)$ is a vector given by

$$\nabla f = \frac{\partial f}{\partial r}\mathbf{e}_{\mathrm{r}} + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{e}_{\theta} + \frac{\partial f}{\partial z}\mathbf{e}_{\mathrm{z}}. \tag{A.197}$$

The gradient of a vector function $\mathbf{v}(r, \theta, z, t)$ is given by

$$[\nabla \otimes \mathbf{v}]^{\mathrm{T}} = \begin{bmatrix} \dfrac{\partial v_{\mathrm{r}}}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_{\mathrm{r}}}{\partial \theta} & \dfrac{\partial v_{\mathrm{r}}}{\partial z} \\ \dfrac{\partial v_{\theta}}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_{\theta}}{\partial \theta} & \dfrac{\partial v_{\theta}}{\partial z} \\ \dfrac{\partial v_{\mathrm{z}}}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_{\mathrm{z}}}{\partial \theta} & \dfrac{\partial v_{\mathrm{z}}}{\partial z} \end{bmatrix}. \tag{A.198}$$

The formula for the divergence of the vector $\mathbf{v}$ in cylindrical coordinates is obtained by taking the trace of (A.198), $\mathrm{tr}[\Delta \otimes \mathbf{v}] = \Delta \cdot \mathbf{v} = \mathrm{div}\ \mathbf{v}$. The curl is the gradient operator cross product with a vector function $\mathbf{v}(r, \theta, z, t)$, thus

$$\nabla \times \mathbf{v} = \mathrm{curl}\mathbf{v} = \begin{bmatrix} \mathbf{e}_{\mathrm{r}} & \mathbf{e}_{\theta} & \mathbf{e}_{\mathrm{z}} \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} & \dfrac{\partial}{\partial z} \\ v_{\mathrm{r}} & r v_{\theta} & v_{\mathrm{z}} \end{bmatrix}. \tag{A.199}$$

The form of the double gradient three-dimensional second order tensor defined by $\mathbf{O} = \nabla \otimes \nabla$ ($\mathrm{tr}\mathbf{O} = \nabla^2$) and its six-dimensional vector counterpart $\hat{\mathbf{O}}(\hat{\mathbf{O}} \cdot \hat{\mathbf{U}} = \mathrm{tr}\mathbf{O} = \nabla^2)$ have the following cylindrical coordinate representations:

$$\mathbf{O} = [\nabla \otimes \nabla] = \begin{bmatrix} \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) & \frac{1}{r}\frac{\partial^2}{\partial r \partial \theta} & \frac{\partial^2}{\partial r \partial z} \\ \frac{1}{r}\frac{\partial^2}{\partial r \partial \theta} & \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} & \frac{1}{r}\frac{\partial^2}{\partial z \partial \theta} \\ \frac{\partial^2}{\partial r \partial z} & \frac{1}{r}\frac{\partial^2}{\partial z \partial \theta} & \frac{\partial^2}{\partial z^2} \end{bmatrix}, \tag{A.200}$$

and

$$\hat{\mathbf{O}} = \left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right), \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}, \frac{\partial^2}{\partial z^2}, \sqrt{2}\frac{1}{r}\frac{\partial^2}{\partial z \partial \theta}, \sqrt{2}\frac{\partial^2}{\partial r \partial z}, \sqrt{2}\frac{1}{r}\frac{\partial^2}{\partial r \partial \theta}\right]^{\mathrm{T}}, \tag{A.201}$$

and the operation of $\hat{\mathbf{O}}$ on a six-dimensional vector representation of a second order tensor in three dimensions, $\hat{\mathbf{O}} \cdot \hat{\mathbf{T}} = \mathbf{O} : \mathbf{T} = \mathrm{tr}\mathbf{O} \cdot \mathbf{T}$, is given by

$$\hat{\mathbf{O}} \cdot \hat{\mathbf{T}} = \frac{\partial}{\partial r}\left(r\frac{\partial T_{\mathrm{rr}}}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 T_{\theta\theta}}{\partial \theta^2} + \frac{\partial^2 T_{\mathrm{zz}}}{\partial z^2} + 2\frac{1}{r}\frac{\partial^2 T_{\mathrm{z}\theta}}{\partial z \partial \theta} + 2\frac{\partial^2 T_{\mathrm{rz}}}{\partial r \partial z} + 2\frac{1}{r}\frac{\partial^2 T_{\mathrm{r}\theta}}{\partial r \partial \theta}. \tag{A.202}$$

The divergence of a second order tensor $\mathbf{T}$ is defined in a similar fashion to the divergence of a vector; it is a vector given by

$$\begin{aligned} \nabla \cdot \mathbf{T}(r,\theta,z,t) = &\left(\frac{\partial T_{\mathrm{rr}}}{\partial r} + \frac{1}{r}\frac{\partial T_{\mathrm{r}\theta}}{\partial \theta} + \frac{\partial T_{\mathrm{rz}}}{\partial z}\right)\mathbf{e}_{\mathrm{r}} + \left(\frac{\partial T_{\mathrm{r}\theta}}{\partial r} + \frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial \theta} + \frac{\partial T_{\theta\mathrm{z}}}{\partial z}\right)\mathbf{e}_{\theta} \\ &+ \left(\frac{\partial T_{\mathrm{zr}}}{\partial r} + \frac{1}{r}\frac{\partial T_{\mathrm{z}\theta}}{\partial \theta} + \frac{\partial T_{\mathrm{zz}}}{\partial z}\right)\mathbf{e}_{\mathrm{z}}. \end{aligned} \tag{A.203}$$

The strain–displacement relations (3.52) are written in cylindrical coordinates as

$$\begin{aligned} E_{\mathrm{rr}} &= \frac{\partial u_{\mathrm{r}}}{\partial r}, E_{\theta\theta} = \frac{1}{r}\frac{\partial u_{\theta}}{\partial \theta} + \frac{u_{\mathrm{r}}}{r}, E_{\mathrm{r}\theta} = \frac{1}{2}\left(\frac{1}{r}\frac{\partial u_{\mathrm{r}}}{\partial \theta} + \frac{\partial u_{\theta}}{\partial r} - \frac{u_{\theta}}{r}\right) \\ E_{\mathrm{zz}} &= \frac{\partial u_{\mathrm{z}}}{\partial z}, E_{\mathrm{rz}} = \frac{1}{2}\left(\frac{\partial u_{\mathrm{z}}}{\partial r} + \frac{\partial u_{\mathrm{r}}}{\partial z}\right), E_{\theta\mathrm{z}} = \frac{1}{2}\left(\frac{1}{r}\frac{\partial u_{\mathrm{z}}}{\partial \theta} + \frac{\partial u_{\theta}}{\partial z}\right), \end{aligned} \tag{A.204}$$

and similar formulas apply for the rate of deformation-velocity tensor $\mathbf{D}$, (3.33) if the change of notation from $\mathbf{E}$ to $\mathbf{D}$ and $\mathbf{u}$ to $\mathbf{v}$ is accomplished. The stress equations of motion (4.37) in cylindrical coordinates are

$$\begin{aligned}\frac{\partial T_{\mathrm{rr}}}{\partial r}+\frac{1}{r}\frac{\partial T_{\mathrm{r}\theta}}{\partial \theta}+\frac{\partial T_{\mathrm{rz}}}{\partial z}+\frac{T_{\mathrm{rr}}-T_{\theta\theta}}{r}+\rho d_{\mathrm{r}}&=\rho u\ddot{x}_{\mathrm{r}},\\ \frac{\partial T_{\mathrm{r}\theta}}{\partial r}+\frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial \theta}+\frac{\partial T_{\theta\mathrm{z}}}{\partial z}+\frac{T_{\mathrm{r}\theta}}{r}+\rho d_{\theta}&=\rho\ddot{x}_{\theta},\\ \frac{\partial T_{\mathrm{rz}}}{\partial r}+\frac{1}{r}\frac{\partial T_{\theta\mathrm{z}}}{\partial \theta}+\frac{\partial T_{\mathrm{zz}}}{\partial z}+\frac{T_{\mathrm{rz}}}{r}+\rho d_{\mathrm{z}}&=\rho\ddot{x}_{\mathrm{z}}.\end{aligned} \tag{A.205}$$

Problem

A.15.1 Calculate the components of the rate of deformation-velocity tensor **D**, (3.33) in cylindrical coordinates.

A.16 Laplace Transform Refresher

The solutions to linear differential equations in time are often obtained by the use of Laplace transforms and the Laplace transforms of discontinuous functions. Laplace transforms provide a method for representing and analyzing linear systems using algebraic methods. The Laplace transform variable 's' can directly replace the $d/\mathrm{d}t$ operator in differential equations involving functions whose value at time zero is zero. Most of the readers of this text will have been introduced to Laplace transforms at some time in their past and find it convenient to have the salient points about these transforms refreshed in their minds before solving the differential equations of this type.

The Laplace transform of a function $f(t)$, $0 \leq t \leq \infty$, is defined by

$$L\{f(t)\} = \tilde{f}(s) = \int_0^\infty f(t)\mathrm{e}^{-st}\mathrm{d}t. \tag{A.206}$$

This integral is absolutely convergent if the function $f(t)$ is of exponential order α, that is to say if $f(t)$ is continuous for $0 \leq t \leq \infty$ and $|f(t)|<c\mathrm{e}^{-\alpha t}$ where c and α are constants. The notation for the inverse Laplace transform of the function $\tilde{f}(s)$ is

$$L^{-1}\{\tilde{f}(s)\} = f(t). \tag{A.207}$$

The Laplace transforms of derivatives and integrals are some of the most useful properties of the transform. If the function $f(t)$ is of exponential order and the derivative of $f(t)$ is continuous, then for $s > \alpha$,

$$L\{f'(t)\} = sL\{f(t)\} - f(0) = s\tilde{f}(s) - f(0), \tag{A.208}$$

a result that, by repeated application, can be used to construct similar formulas for higher order derivatives, for example,

$$L\{f''(t)\} = s^2 L\{f(t)\} - sf'(0) - f(0). \tag{A.209}$$

The Laplace transform of an integral is given by

$$L\left\{\int_0^t f(x)\,\mathrm{d}x\right\} = \frac{1}{s}\tilde{f}(s), \tag{A.210}$$

for $s > \alpha$. It follows that, for a function whose value at $t = 0$ is zero, and whose higher order derivatives are zero at $t = 0$, differentiation corresponds to multiplication by s and integration corresponds to division by s. If the function $f(t)$ is of exponential order, then

$$L\{\mathrm{e}^{-at} f(t)\} = \tilde{f}(s + a), \tag{A.211}$$

for $s > \alpha - a$, and if $\tilde{f}(s)$ has derivatives of all orders for $s > \alpha$,

$$\tilde{f}'(s) = L\{-tf(t)\} \text{ or } \tilde{f}^{(\mathrm{n})}(s) = L\{(-t)^{\mathrm{n}} f(t)\}. \tag{A.212}$$

In the solution of differential equations using the Laplace transform it is often necessary to expand the transform into partial fractions before finding the inverse transform. The rational function $P(s)/Q(s)$, where $Q(s)$ is a polynomial with n distinct zeros, $\alpha_1, \alpha_2, \ldots, \alpha_{\mathrm{n}}$, and $P(s)$ is a polynomial of degree less than n, can be written in the form

$$\frac{P(s)}{Q(s)} = \frac{A_1}{s - \alpha_1} + \frac{A_2}{s - \alpha_2} + \cdots + \frac{A_{\mathrm{n}}}{s - \alpha_{\mathrm{n}}} \tag{A.213}$$

where the A_{i} are constants. The constants A_{i} are determined by multiplying both sides of the equation above by $s - \alpha_{\mathrm{i}}$ and letting s approach α_{i}, thus

$$A_{\mathrm{i}} = \lim_{s \to \alpha_{\mathrm{i}}} it \frac{(s - \alpha_{\mathrm{i}})P(s)}{Q(s)} = \lim_{s \to \alpha_{\mathrm{i}}} it \frac{P(s)}{\frac{Q(s) - Q(\alpha_{\mathrm{i}})}{s - \alpha_{\mathrm{i}}}} = \frac{P(\alpha_{\mathrm{i}})}{Q'(\alpha_{\mathrm{i}})}. \tag{A.214}$$

Combining the two previous equations one may write

$$\frac{P(s)}{Q(s)} = \frac{P(\alpha_1)}{Q'(\alpha_1)} \frac{1}{s - \alpha_1} + \cdots + \frac{P(\alpha_{\mathrm{n}})}{Q'(\alpha_{\mathrm{n}})} \frac{1}{s - \alpha_{\mathrm{n}}}, \tag{A.215}$$

whose inverse Laplace transform is given by

Table A.1 A very short table of Laplace transforms

Transform function	Object function	Conditions
$1/s$	1	$s>0$
$1/(s-a)$	e^{at}	$s>0$
$1/(s^2+a^2)$	$(1/a)\sin at$	$a\neq 0,\ s>0$
$s/(s^2+a^2)$	$\cos at$	$s>0$
$1/(s(s+a))$	$(1/a)(1-\mathrm{e}^{-at})$	$s>a$

$$L^{-1}\left\{\frac{P(s)}{Q(s)}\right\}=\frac{P(\alpha_1)}{Q'(\alpha_1)}\mathrm{e}^{\alpha_1 t}+\ldots+\frac{P(\alpha_\mathrm{n})}{Q'(\alpha_\mathrm{n})}\mathrm{e}^{\alpha_\mathrm{n} t}. \tag{A.216}$$

The result above was obtained by using the Table A.1 to verify that

$$L^{-1}\{(s-\alpha_\mathrm{i})^{-1}\}=\mathrm{e}^{\alpha_\mathrm{i} t}.$$

Example A.16.1

Problem: Solve the differential equation $\frac{\mathrm{d}y}{\mathrm{d}t}-5y=\mathrm{e}^{3t}+4$ for $y(0)=0$-*using* Laplace transforms.

Solution: Noting that the Laplace transform of e^{-at} is $\frac{1}{s+a}$ and that the Laplace transform of the derivative of a function is equal to s times the Laplace transform of the function minus the value of the function at $t=0$, the Laplace transform of $\frac{\mathrm{d}y}{\mathrm{d}t}-5y=\mathrm{e}^{3t}+4$ is $s\hat{y}(s)-5\hat{y}(s)=\frac{1}{s-3}+\frac{4}{s}$, thus

$$\hat{y}(s)=\frac{5s-12}{s(s-3)(s-5)}.$$

By partial fractions

$$\begin{aligned}\frac{1}{s(s-3)(s-5)}&=-\frac{4}{5s}-\frac{1}{2(s-3)}+\frac{13}{10(s-5)},\ \text{thus}\\ \hat{y}(s)&=-\frac{4}{5s}-\frac{1}{2(s-3)}+\frac{13}{10(s-5)}.\end{aligned}$$

Using $\mathcal{L}^{-1}\{(s-a)^{-1}\}=\mathrm{e}^{at}$, the inverse Laplace transform is

$$y(t)=-\frac{4}{5}-\frac{1}{2}\mathrm{e}^{3t}+\frac{13}{10}\mathrm{e}^{5t}.$$

The convolution of the functions $f(t)$ and $g(t)$ is indicated by $f(t)*g(t)$ and defined as the integral

$$f(t) * g(t) = \int_0^t f(t-x)g(x)\,\mathrm{d}x. \tag{A.217}$$

It can be shown that the convolution is commutative ($f * g = g * f$), associative ($(f * g) * h = f * (g * h)$), and distributive ($f * (g + h) = f * g + f * h$). The Laplace transform of a convolution is the product of the Laplace transforms, thus

$$L\{f(t) * g(t)\} = \tilde{f}(s)\tilde{g}(s). \tag{A.218}$$

Example A.16.2

Problem: Solve the differential equation $\frac{\mathrm{dy}}{\mathrm{dt}} - 5y = q(t)$ for $y(0) = 0$ in the general case where $q(t)$ is not specified, other than it has a Laplace transform.

Solution: Noting that the Laplace transform of e^{-at} is $\frac{1}{s+a}$ and that the Laplace transform of the derivative of a function is equal to s times the Laplace transform of the function minus the value of the function at $t = 0$, the Laplace transform of $\frac{\mathrm{dy}}{\mathrm{dt}} - 5y = q(t)$ is $s\tilde{y}(s) - 5\tilde{y}(s) = \tilde{q}(s)$, thus $\tilde{y}(s) = \frac{\tilde{q}(s)}{(s-5)}$.

Observe that $\tilde{y}(s)$ is the product of two transformed functions like those that appear on the right-hand side of (A.218) above, thus from (A.218) one can see by partial fractions that

$$L\left\{\int_0^t \mathrm{e}^{5(t-x)}q(x)\,\mathrm{d}x\right\} = \frac{\tilde{q}(s)}{(s-5)} \quad \text{or} \quad \int_0^t \mathrm{e}^{5(t-x)}q(x)\,\mathrm{d}x = L^{-1}\left\{\frac{\tilde{q}(s)}{(s-5)}\right\},$$

thus $y(t) = \int_0^t \mathrm{e}^{5(t-x)}q(x)\,\mathrm{d}x$. This is a general integral to the differential equation for any function $q(t)$.

In the material above it was assumed for simplicity that the functions were continuous in the interval $0 \leq t \leq \infty$. However, one of the most attractive features of using Laplace transforms is their simplicity in dealing with functions that contain discontinuities. The unit or Heaviside step function $h(t-t_\mathrm{o})$ is defined as 0 for $t < t_\mathrm{o}$ and as 1 for $t > t_\mathrm{o}$. Since the function jumps from the value 0 to the value 1 on passage through $t = t_\mathrm{o}$, the approach of $h(t-t_\mathrm{o})$ to the value t_o is therefore different from below than it is from above,

$$h(t_o^-) = \lim_{t->t_o^-} h(t) = 0, h(t_o^+) = \lim_{t->t_o^+} h(t) = 1. \tag{A.219}$$

The function $h(t-t_\mathrm{o})$ is multiple valued at $t = t_\mathrm{o}$ and its value there depends on how the point $t = t_\mathrm{o}$ is approached. The Laplace transform of the discontinuous function, because it integrates the function, removes the discontinuity, thus

$$L\{h(t-t_o)\} = \int_0^\infty h(t-t_o)e^{-st}dt = \int_{t_o}^\infty e^{-st}dt = \frac{e^{-st_o}}{s}. \tag{A.220}$$

The derivative of the unit step function is the delta function $\delta(t)$ which has the value 0 for all values of t except $t = t_o$. The function may be viewed as the limit as ε tends to zero of a function that has the value $1/\varepsilon$ between 0 and ε and is zero everywhere else. These properties of $\delta(t)$ are then

$$\frac{dh(t-t_o)}{dt} = \delta(t-t_o), \delta(t-t_o) = 0 \text{ for } t \neq t_o, \text{and } \int_0^\infty \delta(t-t_o)\,dt = 1. \tag{A.221}$$

If the delta function is multiplied by any other function $f(t)$, the product is zero everywhere except at $t = t_o$ and it follows that

$$\int_0^\infty f(t)\delta(t-t_o)\,dt = f(t_o). \tag{A.222}$$

Example A.16.3

Problem: Use Laplace transforms to solve the ordinary differential equation representing the standard linear solid (1.8) for $\tilde{F}(s)$ as a function of $\tilde{x}(s)$ and vice versa assuming that $F(0^-)$ and $x(0^-)$ are equal to zero. Then determine the creep and relaxation functions from the result.

Solution: The Laplace transform of the differential equation for the standard linear solid (1.8) is given by

$$\frac{\tilde{F}(s)}{k} + \tau_x \frac{s\tilde{F}(s)}{k} = \tilde{x}(s) + \tau_F s\tilde{x}(s). \tag{A.223}$$

The solutions of this transformed equation for $\tilde{F}(s)$, and for $\tilde{x}(s)$, are

$$\frac{\tilde{F}(s)}{k} = \frac{(1+s\tau_F)}{(1+s\tau_x)}\tilde{x}(s) \text{ and } \tilde{x}(s) = \frac{(1+s\tau_x)}{(1+s\tau_F)}\frac{\tilde{F}(s)}{k}, \tag{A.224}$$

respectively. To obtain the creep function one sets $F(t) = h(t)$ in the second of the equations above; thus by partial fractions

$$k\tilde{x}(s) = \frac{(1+s\tau_x)}{(1+s\tau_F)}\frac{1}{s} = \frac{1}{s(1+s\tau_F)} + \frac{\tau_x}{(1+s\tau_F)}, \tag{A.225}$$

and to obtain the relaxation function one sets $x(t) = h(t)$ in the first of the equations above, thus

$$\frac{\tilde{F}(s)}{k} = \frac{(1+s\tau_{\mathrm{F}})}{s(1+s\tau_x)} = \frac{1}{s(1+s\tau_x)} + \frac{\tau_{\mathrm{F}}}{(1+s\tau_x)}. \tag{A.226}$$

Executing the inverse Laplace transforms of $\tilde{x}(s)$ given by (A.225) and $\tilde{F}(s)$ given by (A.226) using inverse transform results in Table A.1, the creep and relaxation functions (1.10) and (1.11) are obtained.

A.17 Direct Integration of First Order Differential Equations

The ordinary differential equation representing the standard linear solid (2.8) may be solved for $F(t)$ given a specified $x(t)$ by direct integration of (2.8), recognizing that it is a first order ordinary differential equation. For the direct integration method note that all solutions to the linear first order differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} + p(t)x = q(t) \tag{A.227}$$

are given by

$$x(t) = \left(\mathrm{e}^{\int -p(t)\,\mathrm{d}t}\right)\left[\int q(t)\mathrm{e}^{\int p(t)\,\mathrm{d}t}\mathrm{d}t + C\right]. \tag{A.228}$$

where C is a constant of integration (Kaplan 1958). If one sets $p(t) = (1/\tau_{\mathrm{F}})$ and $q(t) = (1/k\tau_{\mathrm{F}})F + (\tau_x/k\tau_{\mathrm{F}})(\mathrm{d}F/\mathrm{d}t)$, then (2.8) may be rewritten in the form of the differential equation (A.227) and it follows from (A.228) that

$$x(t) = \left(\mathrm{e}^{-t/\tau_{\mathrm{F}}}\right)\left(\frac{1}{k\tau_{\mathrm{F}}}\right)\left[\int_0^t F(t)\mathrm{e}^{t/\tau_{\mathrm{F}}}\mathrm{d}t + \tau_x\int_0^t\left(\frac{\mathrm{d}F}{\mathrm{d}t}\mathrm{e}^{t/\tau_{\mathrm{F}}}\mathrm{d}t\right]\right] + C\mathrm{e}^{-t/\tau_{\mathrm{F}}}.$$

For both the creep and the relaxation functions the constant C is evaluated using the fact that the dashpot in the standard linear solid cannot extend in the first instant so the deflection at $t = 0$, x_{o}, is due only to the deflection of the two springs in the standard linear solid. The two springs must deflect the same amount, x_{o}, thus the initial force is $F_{\mathrm{o}} = (k + k_{\mathrm{R}})x_{\mathrm{o}}$. This relationship may be rewritten in the form $F_{\mathrm{o}} = (k\tau_{\mathrm{F}}/\tau_x)x_{\mathrm{o}}$, using definitions in (2.9). For the creep function the initial force, F_{o}, is taken to be one unit of force; thus the initial displacement is given by $x(0^+) = x_{\mathrm{o}} = \tau_x/k\tau_{\mathrm{F}}$. For the relaxation function the initial displacement, x_{o}, is taken to be one unit of displacement; thus the initial force F_{o} is given by $F(0^+) = F_{\mathrm{o}} = k\tau_{\mathrm{F}}/\tau_x$. To obtain the creep function one sets $F(t) = h(t)$, thus

$$x(t) = \left(\mathrm{e}^{-t/\tau_\mathrm{F}}\right)\left(\frac{1}{k\tau_\mathrm{F}}\right)\left[\int_0^t h(t)\mathrm{e}^{t/\tau_\mathrm{F}}\mathrm{d}t + \tau_x\int_0^t \delta(t)\mathrm{e}^{t/\tau_\mathrm{F}}\mathrm{d}t\right] + C\mathrm{e}^{-t/\tau_\mathrm{F}}.$$

The integrals in this equation may be evaluated with ease using the definition of the unit step function and the integral formula (A.222), while the initial condition $x(0^+) = x_\mathrm{o} = \tau_x/k\tau_\mathrm{F}$ requires that $C = 0$, thus the creep function $c(t)$ for the standard linear solid (1.10) is again recovered. The relaxation function for the standard linear solid is found by the same methods with some interchange in the roles of $F(t)$ and $x(t)$. In this case the selection of $p(t)$ and $\mathrm{q}(t)$ in (A.227) are that $p(t) = (1/\tau_x)$ and $\mathrm{q}(t) = (k/\tau_x)(x + \tau_\mathrm{F}(\mathrm{d}x/\mathrm{d}t))$, then the solution of (1.8) for $F(t)$ is given by (A.228) as

$$F(t) = \left(\mathrm{e}^{-t/\tau_x}\right)\left(\frac{k}{\tau_x}\right)\left[\int_0^t x(t)\mathrm{e}^{t/\tau_x}\mathrm{d}t + \tau_\mathrm{F}\int_0^t \left(\frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{e}^{t/\tau_x}\mathrm{d}t\right] + C\mathrm{e}^{-t/\tau x}.\right.$$

To obtain the relaxation function one sets $x(t) = h(t)$, thus

$$F(t) = \left(\mathrm{e}^{-t/\tau_x}\right)\left(\frac{k}{\tau_x}\right)\left[\int_0^t h(t)\mathrm{e}^{t/\tau_x}\mathrm{d}t + \tau_\mathrm{F}\int_0^t (\delta(t)\mathrm{e}^{t/\tau_x}\mathrm{d}t\right].$$

The integrals in this equation may again be evaluated using the definition of the unit step function and the integral formula (A.222). Thus setting $F(t) = r(t)$ and using the fact that the initial condition $F(0^+) = F_\mathrm{o} = k\tau_\mathrm{F}/\tau_x$ requires that $C = 0$, the relaxation function for the standard linear solid (1.10) is obtained again.

Index

S.C. Cowin, *Continuum Mechanics of Anisotropic Materials*,
DOI 10.1007/978-1-4614-5025-2,