PAPPALARDO SERIES IN
MECHANICAL ENGINEERING

An Introduction to Finite Elasticity

Volume III of Lecture Notes on
the Mechanics of Elastic Solids

Version 1.0

Rohan Abeyaratne
Quentin Berg Professor of Mechanics
MIT Department of Mechanical Engineering

http://web.mit.edu/abeyaratne/lecture_notes.html

Published online: March 31, 2022

Electronic Publication

Rohan Abeyaratne
Quentin Berg Professor of Mechanics
Department of Mechanical Engineering
77 Massachusetts Institute of Technology
Cambridge, MA 02139-4307, USA

Abeyaratne, Rohan, An Introduction to Finite Elasticity. Volume III of Lecture Notes on the Mechanics of Elastic Solids. / Rohan Abeyaratne – 1st Edition – Cambridge, MA:

ISBN-13: 978-0-9791865-2-3
ISBN-10: 0-9791865-2-8

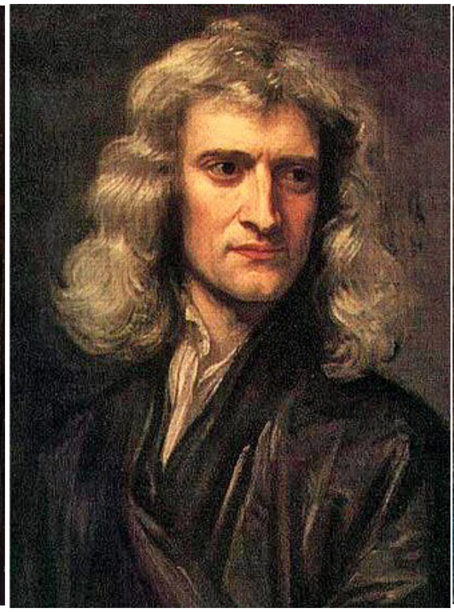
QC

Please send corrections, suggestions and comments to *abeyaratne.vol.3@gmail.com*

Updated March 18, 2022

To the many students I have had the privilege and joy journeying with during their time at MIT: you have taught me so much.

Galileo Galilei
(1564-1642)



Isaac Newton
(1642-1727)

Leonard Euler
(1707-1783)

Augustin-Louis Cauchy
(1789-1857)

Ronald Rivlin
(1915-2005)

From Wikipedia

PREFACE

The MIT Department of Mechanical Engineering offers a series of graduate-level subjects on the Mechanics of Solids and Structures that in recent years has included:

2.071: Mechanics of Solid Materials,
2.072: Mechanics of Continuous Media,
2.074: Solid Mechanics: Elasticity (formerly 2.083),
2.073: Solid Mechanics: Plasticity and Inelastic Deformation,
2.075: Mechanics of Soft Materials,
2.080: Structural Mechanics,
2.094: Finite Element Analysis of Solids and Fluids,
2.095: Molecular Modeling and Simulation for Mechanics, and
2.099: Computational Mechanics of Materials.

I have taught the second and third of these subjects on several occasions and the current four volumes comprise the notes I developed for them. These are *notes, not textbooks*. The earliest rough drafts were written in 1987 and 1988 and they have been expanded and refined on every subsequent occasion when I taught these classes. They are organized as follows:

Volume I: A Brief Review of Some Mathematical Preliminaries

Volume II: Continuum Mechanics

Volume III: An Introduction to Finite Elasticity

Volume IV: Linear Elasticity

This is Volume III.

Until 2018, the subject 2.074 on elasticity treated only the linear(ized) theory. In recent years, several students doing research on “soft materials” and “biomaterials” asked for references to books where they could learn the nonlinear theory on their own and I would direct them to one of the books listed below. In fall 2018 I decided to devote the first part of 2.074 to the nonlinear theory and the rest to the linearized theory. Volume III consists of the notes from the first part (with somewhat more detail than what I actually cover). Due to the limitation of time – less than one semester – the treatment is special in many ways, e.g. inertial effects are not considered. While there is some duplication of material between Volumes II and III, the more narrow focus here should be helpful to the student encountering this material for the first time. An expanded treatment of the underlying theory can be found in the relevant chapters of Volume II.

The content of these notes is entirely classical, in the best sense of the word. While the material covered is not original, some of it is not usually emphasized in textbooks. They include the several boundary-value problems focused on *illustrating nonlinear phenomena* (Chapter 5), strain-energy functions with multiple energy-wells used in the study solid-to-solid phase transitions (Chapter 7), and Cauchy's lattice-based theory of elasticity (Chapter 8).

One of the few positive outcomes of the COVID-19 pandemic was that in the fall of 2020, when 2.074 was taught remotely, I recorded a *few* (amateur) videos on some particular topics. Links to them are provided in the text.

In case you wonder why "an introduction" is about 500 pages long, it is because of the numerous examples and exercises that are included in almost every chapter. They are an essential part of these notes. Many of these problems illustrate general concepts through particular examples. Some provide further details on items touched on in the text. Others generalize previously described special cases. Some concern proofs of results that had simply been quoted before, or they refer to results that will be used in what follows.

The problems are numbered as follows: Problem 2.6 for example can be found at the end of Chapter 2 in the section on Exercises, while Problem 2.6.2 is located within Section 2.6. This distinction between problems identified by two numbers (e.g. 2.6) versus three (e.g. 2.6.2) has been adopted throughout.

My appreciation for mechanics was nucleated by Professors Douglas Amarasekara and Munidasa Ranaweera of the (then) University of Ceylon, and was subsequently shaped and grew substantially under the influence of Professors James K. Knowles and Eli Sternberg of the California Institute of Technology. I have been most fortunate to have had the opportunity to apprentice under these inspiring and distinctive scholars.

I would especially like to acknowledge the innumerable illuminating and stimulating interactions with my mentor, colleague and friend the late Jim Knowles. His influence on me cannot be overstated.

I am also indebted to the many MIT students who have given me enormous fulfillment and joy to be part of their education and for their feedback on these notes.

My understanding of elasticity has benefitted greatly from numerous conversations with many colleagues including Kaushik Bhattacharya, Janet Blume, Eliot Fried, Morton E. Gurtin, Richard D. James, Stelios Kyriakides, David M. Parks, Sensei Phoebus Rosakis,

Stewart Silling and Nicolas Triantafyllidis. My grateful thanks to them all.

I have drawn on a number of sources over the years as I prepared my lectures. I cannot recall every one of them but they certainly include those listed at the end of each chapter. I have found the following articles and books particularly useful:

Volume III: An Introduction to Finite Elasticity

– J. M. Ball, Some recent developments in nonlinear elasticity and its applications to materials science, in *Nonlinear Mathematics and Its Applications*, edited by P.J. Aston, pp. 93–119. Cambridge University Press, 1996.
– P. Chadwick, *Continuum Mechanics: Concise Theory and Problems*, Wiley, 1976. Reprinted by Dover,1999.
– A. Goriely, A. Erlich and C. Goodbrake, C5.1 Solid Mechanics: Online problem sheets, https://courses.maths.ox.ac.uk/node/36846/materials, Oxford University, 2018.
– M.E. Gurtin, E. Fried and L. Anand, *The Mechanics and Thermodynamics of Continua*, Cambridge University Press, 2010.
– J. K. Knowles and E. Sternberg, *(Unpublished) Lecture Notes for AM136: Finite Elasticity*, California Institute of Technology, Pasadena, CA 1978.
– R.W. Ogden, *Nonlinear Elastic Deformations*, Ellis Horwood, 1984. Reprinted by Dover, 1997.
– D.J. Steigmann, *Finite Elasticity Theory*, Oxford, 2017.

For a treatment of the rigorous mathematical underpinnings, the student may refer to:

– S. S. Antman, *Nonlinear Problems of Elasticity*, Springer-Verlag, 1995.
– J.E. Marsden and T.J.R. Hughes, *Mathematical Foundations of Elasticity*, Prentice-Hall, 1983. Reprinted by Dover 1994.

The following notation will be used in Volume III, though there will be a few lapses (for reasons of tradition):

– Greek letters will denote scalars;
– lowercase boldface Latin letters will denote vectors; and
– uppercase boldface Latin letters will denote linear transformations (tensors).

Thus, for example, $\alpha, \beta, \gamma ...$ will denote scalars (real numbers); $\mathbf{a}, \mathbf{b}, \mathbf{c}, ...$ will denote vectors; and $\mathbf{A}, \mathbf{B}, \mathbf{C}, ...$ will denote tensors.

One consequence of this notational convention is that I will *not* use the uppercase boldface letter $\mathbf{X}$ to denote the position vector of a particle in the reference configuration (as many authors do). Being a boldface uppercase letter, my convention would dictate that $\mathbf{X}$ represent some tensor. Instead, I use the lowercase boldface letters $\mathbf{x}$ and $\mathbf{y}$ to denote the respective position *vectors* of a particle in the reference and current configurations. I sometimes lightheartedly refer to this as the "Caltech-Minnesota notation".

I have been frequently asked whether I intend to publish these notes in the form of a traditional textbook, and my answer has aways been "no". These notes are being made available primarily for students who like me, when I was studying in Sri Lanka, could not afford the cost of purchasing a textbook. I therefore intend to make these notes available for free online.

Finally, I would like to express my grateful thanks to Jane and Neil Pappalardo for their friendship and support over many years. The writing of Volume III was supported by the MIT-Pappalardo Series in Mechanical Engineering.

List of symbols

Table 1: Kinematics

Quantity	**Symbol**
Position vector of particle in reference configuration	$\mathbf{x}$
Position vector of particle in deformed configuration	$\mathbf{y}$
Deformation field	$\mathbf{y}(\mathbf{x})$
Displacement field	$\mathbf{u}(\mathbf{x})$
Infinitesimal material fiber in reference configuration	$\mathrm{d}\mathbf{x}$
Infinitesimal material fiber in deformed configuration	$\mathrm{d}\mathbf{y}$
Volume of an infinitesimal part in reference configuration	$\mathrm{d}V_x$
Volume of an infinitesimal part in deformed configuration	$\mathrm{d}V_y$
Infinitesimal (vector) area in reference configuration	$\mathrm{d}A_x\,\mathbf{n}_{\mathrm{R}}$
Infinitesimal (vector) area in deformed configuration	$\mathrm{d}A_y\,\mathbf{n}$
Deformation gradient tensor	$\mathbf{F} = \mathrm{Grad}\,\mathbf{y}$
Jacobian determinant	$\mathrm{J} = \det\mathbf{F}$
Displacement gradient tensor	$\mathbf{H} = \mathrm{Grad}\,\mathbf{u}$
Right (Lagrangian) stretch tensor	$\mathbf{U}$
Left (Eulerian) stretch tensor	$\mathbf{V}$
Rotation tensor	$\mathbf{R}$
Principal stretches	$\lambda_1, \lambda_2, \lambda_3$
Principal directions of Lagrangian stretch	$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$
Principal directions of Eulerian stretch	$\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$
Right Cauchy-Green deformation tensor	$\mathbf{C} = \mathbf{U}^2 = \mathbf{F}^T\mathbf{F}$
Left Cauchy-Green deformation tensor	$\mathbf{B} = \mathbf{V}^2 = \mathbf{F}\mathbf{F}^T$
Principal scalar invariants of $\mathbf{C}$	$\mathrm{I}_1(\mathbf{C}), \mathrm{I}_2(\mathbf{C}), \mathrm{I}_3(\mathbf{C})$
General Lagrangian strain tensor	$\mathbf{E}^{(n)}$
Green Saint-Venant strain tensor	$\mathbf{E}$
Particle velocity	$\mathbf{v}$
Velocity gradient tensor	$\mathbf{L} = \mathrm{grad}\,\mathbf{v}(\mathbf{y}, t)$

Table 2: Mechanics

Quantity	**Symbol**
Cauchy (true) traction vector	$\mathbf{t}(\mathbf{y}, \mathbf{n})$
Normal stress	T_{normal}
Magnitude of resultant shear stress	T_{shear}
Cauchy (true) stress tensor field	$\mathbf{T}(\mathbf{y})$
Principal Cauchy stresses	τ_1, τ_2, τ_3
Principal directions of Cauchy stress	$\mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3$
1st Piola-Kirchhoff traction vector	$\mathbf{s}(\mathbf{x}, \mathbf{n}_{\mathrm{R}})$
1st Piola-Kirchhoff stress tensor field	$\mathbf{S}(\mathbf{x})$
Stress tensor work conjugate to strain tensor $\mathbf{E}^{(n)}$	$\mathbf{S}^{(n)}$
Biot stress tensor	$\mathbf{S}^{(1)}$
2nd Piola-Kirchhoff stress tensor	$\mathbf{S}^{(2)}$
Body force per unit deformed volume	$\mathbf{b}$
Body force per unit reference volume	$\mathbf{b}_{\mathrm{R}}$
Mass density in deformed configuration	ρ
Mass density in reference configuration	ρ_{R}

Table 3: Constitutive Description

Quantity	**Symbol**
Strain energy function	$\widehat{W}(\mathbf{F})$
Strain energy function	$\overline{W}(\mathbf{C})$
Strain energy function	$\widetilde{W}(I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C}))$
Strain energy function	$W^*(\lambda_1, \lambda_2, \lambda_3)$
Reactive stress due to an internal material constraint	$\mathbf{N}$
Reactive pressure due to incompressibility constraint	q
Fiber directions in anisotropic material	$\mathbf{m}_{\mathrm{R}}, \mathbf{m}$
Structural tensor for anisotropic material	$\mathbf{M}$
Additional invariants for anisotropic material	$I_4(\mathbf{C}, \mathbf{M}), I_5(\mathbf{C}, \mathbf{M}), \ldots$

Some Useful Formulae

1. Mathematical background.

$$\text{Orthogonal matrix } [Q]: \qquad Q_{ik}Q_{jk} = Q_{ki}Q_{kj} = \delta_{ij} \tag{1.1}$$

$$\mathbb{T}_{pqr\ldots i\ldots z}\,\delta_{ij} = \mathbb{T}_{pqr\ldots j\ldots z} \tag{1.2}$$

$$\det[A] = e_{ijk}A_{1i}A_{2j}A_{3k} = e_{ijk}A_{i1}A_{j2}A_{k3} = \frac{1}{6}e_{ijk}e_{pqr}A_{ip}A_{jq}A_{kr}. \tag{1.3}$$

$$e_{pqr}\det[A] \;=\; e_{ijk}A_{ip}A_{jq}A_{kr}. \tag{1.4}$$

$$e_{pij}e_{pk\ell} = \delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}. \tag{1.5}$$

$$e_{ijk} = -e_{jik}, \qquad e_{ijk} = -e_{ikj} \tag{1.6}$$

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u} \tag{1.7}$$

$$|\mathbf{u}| = \sqrt{\mathbf{u}\cdot\mathbf{u}} \tag{1.8}$$

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} \tag{1.9}$$

$$|\mathbf{u}| = 0 \qquad \Leftrightarrow \qquad \mathbf{u} = \mathbf{o} \tag{1.10}$$

$$\mathbf{u}\times\mathbf{v} = -\mathbf{v}\times\mathbf{u} \tag{1.11}$$

$$|\mathbf{u}\times\mathbf{v}| = |\mathbf{u}|\,|\mathbf{v}|\,\sin\theta, \qquad (\mathbf{u}\times\mathbf{v})\cdot\mathbf{u} = (\mathbf{u}\times\mathbf{v})\cdot\mathbf{v} = 0 \tag{1.12}$$

$$\mathbf{e}_i\cdot\mathbf{e}_j = \delta_{ij}. \tag{1.13}$$

$$\mathbf{e}_i\times\mathbf{e}_j = e_{ijk}\,\mathbf{e}_k. \tag{1.14}$$

$$\mathbf{v} \;=\; v_i\mathbf{e}_i, \qquad v_i = \mathbf{v}\cdot\mathbf{e}_i \tag{1.15}$$

$$\mathbf{u}\cdot\mathbf{v} = u_iv_i. \tag{1.16}$$

$$|\mathbf{u}| = (\mathbf{u}\cdot\mathbf{u})^{1/2} = (u_ku_k)^{1/2}. \tag{1.17}$$

$$(\mathbf{u}\times\mathbf{v})_i = e_{ijk}u_jv_k. \tag{1.18}$$

$$\mathbf{I}\,\mathbf{u} = \mathbf{u}, \qquad \mathbf{0}\mathbf{u} = \mathbf{o} \qquad \text{for all vectors } \mathbf{u} \tag{1.19}$$

$$\mathbf{A}\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{A}^T\mathbf{v} \qquad \text{for all vectors } \mathbf{u},\mathbf{v} \tag{1.20}$$

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T\mathbf{A}^T \tag{1.21}$$

$$\text{Symmetric tensor } \mathbf{A}: \qquad \mathbf{A} = \mathbf{A}^T \tag{1.22}$$

$$\text{Skew-symmetric tensor } \mathbf{A}: \qquad \mathbf{A} = -\mathbf{A}^T, \tag{1.23}$$

$$\text{Positive definite tensor } \mathbf{A}: \qquad \mathbf{A}\mathbf{u}\cdot\mathbf{u} > 0 \qquad \text{for all vectors } \mathbf{u} \neq \mathbf{0} \tag{1.24}$$

$$\text{Nonsingular tensor } \mathbf{A}: \qquad \mathbf{A}\mathbf{u} = \mathbf{o} \quad \text{if and only if} \quad \mathbf{u} = \mathbf{o} \tag{1.25}$$

$$\text{Nonsingular tensor } \mathbf{A}: \qquad \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{1.26}$$

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \tag{1.27}$$

$$\text{Orthogonal tensor } \mathbf{Q}: \qquad \mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I} \tag{1.28}$$

$$\text{Orthogonal tensor } \mathbf{Q}: \qquad |\mathbf{Q}\mathbf{u}| = |\mathbf{u}| \qquad \text{for all vectors } \mathbf{u} \tag{1.29}$$

$$\mathbf{A} = \mathbf{S} + \mathbf{W}, \qquad \mathbf{S} = \mathbf{S}^T = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T), \quad \mathbf{W} = -\mathbf{W}^T = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) \tag{1.30}$$

$$\text{Nonsingular tensor } \mathbf{F}: \quad \mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}, \quad \mathbf{R} \text{ orthogonal}, \quad \mathbf{U}, \mathbf{V} \text{ symmetric positive definite} \tag{1.31}$$

$$(\mathbf{a}\otimes\mathbf{b})\mathbf{x} = (\mathbf{x}\cdot\mathbf{b})\mathbf{a} \qquad \text{for all vectors } \mathbf{x} \tag{1.32}$$

$$(\mathbf{a}\otimes\mathbf{b})^T = \mathbf{b}\otimes\mathbf{a}, \qquad (\mathbf{a}\otimes\mathbf{b})(\mathbf{c}\otimes\mathbf{d}) \;=\; (\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\otimes\mathbf{d}). \tag{1.33}$$

$$\mathbf{A}(\mathbf{a}\otimes\mathbf{b}) = (\mathbf{A}\mathbf{a})\otimes\mathbf{b}, \qquad (\mathbf{a}\otimes\mathbf{b})\mathbf{A} = \mathbf{a}\otimes(\mathbf{A}^T\mathbf{b}). \tag{1.34}$$

$$\mathbf{e}_1\otimes\mathbf{e}_1 + \mathbf{e}_2\otimes\mathbf{e}_2 + \mathbf{e}_3\otimes\mathbf{e}_3 = \mathbf{I}. \tag{1.35}$$

$$\mathbf{A}\mathbf{e}_j = A_{ij}\,\mathbf{e}_i, \qquad A_{ij} = (\mathbf{A}\mathbf{e}_j)\cdot\mathbf{e}_i. \tag{1.36}$$

$$\mathbf{A} = A_{ij}\mathbf{e}_i\otimes\mathbf{e}_j. \tag{1.37}$$

$$\operatorname{tr}\mathbf{A} = \operatorname{tr}[A] = A_{ii}, \qquad \det\mathbf{A} = \det[A]. \tag{1.38}$$

$$\det(\mathbf{A}\mathbf{B}) = \det\mathbf{A}\;\det\mathbf{B}. \tag{1.39}$$

$$\det(\alpha\mathbf{A}) = \alpha^3\,\det\,\mathbf{A} \tag{1.40}$$

$$\operatorname{tr}(\mathbf{A}\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{A}) \tag{1.41}$$

$$(\mathbf{F}\mathbf{a}\times\mathbf{F}\mathbf{b})\cdot\mathbf{F}\mathbf{c} \;=\; \mathsf{det}\mathbf{F}\;(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} \tag{1.42}$$

$$\mathbf{F}\mathbf{a}\times\mathbf{F}\mathbf{b} \;=\; \det\mathbf{F}\,\mathbf{F}^{-T}(\mathbf{a}\times\mathbf{b}) \tag{1.43}$$

$$\mathbf{A}\cdot\mathbf{B} = \operatorname{tr}(\mathbf{A}\mathbf{B}^T) = A_{ij}B_{ij}. \tag{1.44}$$

$$\mathbf{A}\mathbf{B}\cdot\mathbf{C} = \mathbf{B}\cdot\mathbf{A}^T\mathbf{C} = \mathbf{A}\cdot\mathbf{C}\mathbf{B}^T, \tag{1.45}$$

$$|\mathbf{A}| = (\mathbf{A}\cdot\mathbf{A})^{1/2} = \left[\operatorname{tr}\,(\mathbf{A}\mathbf{A}^T)\right]^{1/2} = (A_{ij}A_{ij})^{1/2}, \tag{1.46}$$

$$(\mathbf{a}\otimes\mathbf{b})_{ij} = a_i b_j. \tag{1.47}$$

$$(\mathbf{A}\mathbf{e}_i)\otimes\mathbf{e}_i = \mathbf{A}. \tag{1.48}$$

$$\operatorname{tr}(\mathbf{a}\otimes\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} \tag{1.49}$$

$$\det(\mathbf{I} + \mathbf{a}\otimes\mathbf{b}) = 1 + \mathbf{a}\cdot\mathbf{b} \tag{1.50}$$

$$(\mathbf{I}+\mathbf{a}\otimes\mathbf{b})^{-1} = 1 - \frac{\mathbf{a}\otimes\mathbf{b}}{1+\mathbf{a}\cdot\mathbf{b}} \qquad (\text{provided } \mathbf{a}\cdot\mathbf{b}\neq -1) \tag{1.51}$$

$$v'_i = Q_{ij}v_j, \qquad \{v'\} = [Q]\{v\} \qquad \text{where } Q_{ij} = \mathbf{e}'_i\cdot\mathbf{e}_j \tag{1.52}$$

$$A'_{ij} = Q_{ip}Q_{jq}A_{pq}, \qquad \{A'\} = [Q][A][Q]^T \qquad \text{where } Q_{ij} = \mathbf{e}'_i\cdot\mathbf{e}_j \tag{1.53}$$

$$\det(\mathbf{A}-\mu\mathbf{I}) \ = \ -\mu^3 + I_1(\mathbf{A})\mu^2 - I_2(\mathbf{A})\mu + I_3(\mathbf{A}) \tag{1.54}$$

where

$$I_1(\mathbf{A}) = \mathrm{tr}(\mathbf{A}), \qquad I_2(\mathbf{A}) = \frac{1}{2}[(\mathrm{tr}(\mathbf{A}))^2 - \mathrm{tr}(\mathbf{A}^2)], \qquad I_3(\mathbf{A}) = \det(\mathbf{A}). \tag{1.55}$$

Eigenvalues and eigenvectors

$$\mathbf{A}\mathbf{a} = \alpha\mathbf{a} \tag{1.56}$$

$$\text{Symmetric tensor } \mathbf{S}: \qquad \mathbf{S} = \sigma_1\mathbf{s}_1\otimes\mathbf{s}_1 + \sigma_2\mathbf{s}_2\otimes\mathbf{s}_2 + \sigma_3\mathbf{s}_3\otimes\mathbf{s}_3 \tag{1.57}$$

$$\frac{dJ}{dt} = J\,\mathbf{F}^{-T}\cdot\frac{d\mathbf{F}}{dt}, \qquad \mathrm{J}(t) = \det\mathbf{F}(t) \tag{1.58}$$

$$\frac{\partial \mathrm{J}}{\partial \mathbf{F}} = \mathrm{J}\mathbf{F}^{-\mathrm{T}}, \qquad \mathrm{J}(\mathbf{F}) = \det\mathbf{F} \tag{1.59}$$

$$\frac{d}{dt}(\mathbf{F}^{-1}) = -\mathbf{F}^{-1}\,\frac{d\mathbf{F}}{dt}\,\mathbf{F}^{-1}, \qquad \mathbf{F} = \mathbf{F}(t) \tag{1.60}$$

If $W(\mathbf{C})$ is defined for all symmetric tensors $\mathbf{C}$ then

$$\left(\frac{\partial W}{\partial \mathbf{C}}\right)_{ij} = \frac{1}{2}\left(\frac{\partial W}{\partial C_{ij}} + \frac{\partial W}{\partial C_{ji}}\right) \tag{1.61}$$

$$\big(\nabla\phi\big)_i = (\mathrm{grad}\,\phi)_i = \frac{\partial\phi}{\partial x_i} \tag{1.62}$$

$$\phi(\mathbf{x}+\delta\mathbf{x}) = \phi(\mathbf{x}) + (\nabla\phi)\cdot\delta\mathbf{x} + o(|\delta\mathbf{x}|) \quad \text{as } |\mathbf{x}|\to 0. \tag{1.63}$$

$$\big(\nabla\mathbf{v}\big)_{ij} = (\mathrm{grad}\,\mathbf{v})_{ij} = \frac{\partial v_i}{\partial x_j} \tag{1.64}$$

$$\mathbf{v}(\mathbf{x}+\delta\mathbf{x}) = \mathbf{v}(\mathbf{x}) + (\nabla\mathbf{v})\delta\mathbf{x} + o(|\delta\mathbf{x}|) \quad \text{as } |\mathbf{x}|\to 0. \tag{1.65}$$

$$\mathrm{div}\,\mathbf{v} = \frac{\partial v_i}{\partial x_i} = \mathrm{tr}\,(\nabla\mathbf{v}) \tag{1.66}$$

$$\big(\mathrm{curl}\ \mathbf{v}\big)_i = e_{ijk}\frac{\partial v_k}{\partial x_j}. \tag{1.67}$$

$$\big(\mathrm{div}\ \mathbf{T}\big)_i = \frac{\partial T_{ij}}{\partial x_j}, \tag{1.68}$$

$$\left(\text{curl } \mathbf{T}\right)_{ij} = e_{ipq} \frac{\partial T_{jq}}{\partial x_p}. \tag{1.69}$$

$$\int_{\partial \mathcal{R}} T_{jk\ldots z} \, n_i \; dA = \int_{\mathcal{R}} \frac{\partial}{\partial x_i} T_{jk\ldots z} \;\; dV \tag{1.70}$$

$$\nabla \phi = \frac{\partial \phi}{\partial R} \, \mathbf{e}_R + \frac{1}{R} \, \frac{\partial \phi}{\partial \Theta} \, \mathbf{e}_\Theta + \frac{\partial \phi}{\partial Z} \, \mathbf{e}_Z. \tag{1.71}$$

$$\begin{aligned} \nabla \mathbf{u} \quad = \quad & \frac{\partial u_R}{\partial R} \left(\mathbf{e}_R \otimes \mathbf{e}_R\right) + \frac{\partial u_\Theta}{\partial R} \left(\mathbf{e}_\Theta \otimes \mathbf{e}_R\right) + \frac{\partial u_Z}{\partial R} \left(\mathbf{e}_Z \otimes \mathbf{e}_R\right) + \\ + \quad & \frac{1}{R} \left(\frac{\partial u_R}{\partial \Theta} - u_\Theta \right) \left(\mathbf{e}_R \otimes \mathbf{e}_\Theta\right) + \frac{1}{R} \left(\frac{\partial u_\Theta}{\partial \Theta} + u_R \right) \left(\mathbf{e}_\Theta \otimes \mathbf{e}_\Theta\right) + \frac{1}{R} \frac{\partial u_Z}{\partial \Theta} \left(\mathbf{e}_Z \otimes \mathbf{e}_\Theta\right) + \\ + \quad & \frac{\partial u_R}{\partial Z} \left(\mathbf{e}_R \otimes \mathbf{e}_Z\right) + \frac{\partial u_\Theta}{\partial Z} \left(\mathbf{e}_\Theta \otimes \mathbf{e}_Z\right) + \frac{\partial u_Z}{\partial Z} \left(\mathbf{e}_Z \otimes \mathbf{e}_Z\right). \end{aligned} \tag{1.72}$$

---

2. Kinematics

$$\text{Deformation :} \qquad \mathbf{y} = \mathbf{y}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x}) \tag{2.1}$$

$$\mathbf{F} = \nabla \mathbf{y} \qquad F_{ij} = \frac{\partial y_i}{\partial x_j} \tag{2.2}$$

$$d\mathbf{y} = \mathbf{F} d\mathbf{x}, \tag{2.3}$$

$$\lambda(\mathbf{m}_\mathrm{R}) = |\mathbf{F}\mathbf{m}_\mathrm{R}| \tag{2.4}$$

$$dV_y = J\ dV_x, \qquad J = \det\ \mathbf{F} \tag{2.5}$$

$$\mathrm{d}A_y\ \mathbf{n} = \mathrm{d}A_x\ J\ \mathbf{F}^{-T}\ \mathbf{n}_\mathrm{R}\ . \tag{2.6}$$

$$\mathbf{F} = \mathbf{R}\,\mathbf{U}, \qquad \mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}, \qquad \mathbf{R} = \mathbf{F}\mathbf{U}^{-1}. \tag{2.7}$$

$$\mathbf{F} = \mathbf{V}\mathbf{R}, \qquad \mathbf{V} = \sqrt{\mathbf{F}\mathbf{F}^T} \qquad \mathbf{R} = \mathbf{V}^{-1}\mathbf{F}. \tag{2.8}$$

$$\mathbf{B} = \mathbf{F}\mathbf{F}^T = \mathbf{V}^2, \qquad \mathbf{C} = \mathbf{F}^T\mathbf{F} = \mathbf{U}^2 \tag{2.9}$$

$$\mathbf{U} = \sum_{i=1}^{3} \lambda_i \mathbf{r}_i \otimes \mathbf{r}_i, \qquad \mathbf{V} = \sum_{i=1}^{3} \lambda_i \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i; \tag{2.10}$$

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \boldsymbol{\ell}_i \otimes \mathbf{r}_i, \qquad \mathbf{R} = \sum_{i=1}^{3} \boldsymbol{\ell}_i \otimes \mathbf{r}_i. \tag{2.11}$$

$$\mathbf{C} = \sum_{i=1}^{3} \lambda_i^2\,(\mathbf{r}_i \otimes \mathbf{r}_i), \qquad \mathbf{B} = \sum_{i=1}^{3} \lambda_i^2\,(\boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i). \tag{2.12}$$

$$I_1(\mathbf{C}) = \mathrm{tr}\,\mathbf{C}, \qquad I_2(\mathbf{C}) = \frac{1}{2}\Big[\big(\mathrm{tr}\,\mathbf{C}\big)^2 - \mathrm{tr}\,\mathbf{C}^2\Big], \qquad I_3(\mathbf{C}) = \det\mathbf{C}, \tag{2.13}$$

$$I_k(\mathbf{C}) = I_k(\mathbf{Q}\mathbf{C}\mathbf{Q}^T) \tag{2.14}$$

$$I_1(\mathbf{C}) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \qquad I_2(\mathbf{C}) = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2, \qquad I_3(\mathbf{C}) = \lambda_1^2\lambda_2^2\lambda_3^2. \tag{2.15}$$

$$\mathbf{E}(\mathbf{U}) = e(\lambda_1)\,\mathbf{r}_1 \otimes \mathbf{r}_1 \ + \ e(\lambda_2)\,\mathbf{r}_2 \otimes \mathbf{r}_2 \ + \ e(\lambda_3)\,\mathbf{r}_3 \otimes \mathbf{r}_3. \tag{2.16}$$

$$\mathbf{E}^{(n)} = \frac{1}{n}\big(\mathbf{U}^n - \mathbf{I}\big) \tag{2.17}$$

$$\text{Biot strain tensor:} \qquad \mathbf{E}^{(1)} = \mathbf{U} - \mathbf{I} \tag{2.18}$$

$$\text{Green-Saint Venant strain tensor:} \qquad \mathbf{E}^{(2)} = \frac{1}{2}(\mathbf{U}^2 - \mathbf{I}) \tag{2.19}$$

$$\mathcal{E}(\mathbf{V}) = e(\lambda_1)\,\boldsymbol{\ell}_1 \otimes \boldsymbol{\ell}_1 \ + \ e(\lambda_2)\,\boldsymbol{\ell}_2 \otimes \boldsymbol{\ell}_2 \ + \ e(\lambda_3)\,\boldsymbol{\ell}_3 \otimes \boldsymbol{\ell}_3. \tag{2.20}$$

$$\text{Piola identity:} \qquad \mathrm{Div}\left(J\,\mathbf{F}^{-\mathrm{T}}\right) = \mathbf{o}, \qquad \frac{\partial}{\partial x_j}\left(J F_{ji}^{-1}\right) = 0 \tag{2.21}$$

$$\text{Piola identity:} \qquad \mathrm{div}\left(J^{-1}\,\mathbf{F}^{\mathrm{T}}\right) = \mathbf{o}, \qquad \frac{\partial}{\partial y_j}\left(J^{-1} F_{ji}\right) = 0 \tag{2.22}$$

$$
\begin{aligned}
\mathbf{F} \quad &= \frac{\partial r}{\partial R}(\mathbf{e}_r \otimes \mathbf{e}_R) + \frac{1}{R}\frac{\partial r}{\partial \Theta}(\mathbf{e}_r \otimes \mathbf{e}_\Theta) + \frac{\partial r}{\partial Z}(\mathbf{e}_r \otimes \mathbf{e}_Z) + \\
&+ r\frac{\partial \theta}{\partial R}(\mathbf{e}_\theta \otimes \mathbf{e}_R) + \frac{r}{R}\frac{\partial \theta}{\partial \Theta}(\mathbf{e}_\theta \otimes \mathbf{e}_\Theta) + r\frac{\partial \theta}{\partial Z}(\mathbf{e}_\theta \otimes \mathbf{e}_Z) + \\
&+ \frac{\partial z}{\partial R}(\mathbf{e}_z \otimes \mathbf{e}_R) \; + \frac{1}{R}\frac{\partial z}{\partial \Theta}(\mathbf{e}_z \otimes \mathbf{e}_\Theta) \; + \frac{\partial z}{\partial Z}(\mathbf{e}_z \otimes \mathbf{e}_Z)\,.
\end{aligned}
\tag{2.23}
$$

$$
\begin{aligned}
\mathbf{B} = \quad & B_{rr}\mathbf{e}_r \otimes \mathbf{e}_r \; + \; B_{\theta\theta}\mathbf{e}_\theta \otimes \mathbf{e}_\theta \; + \; B_{zz}\mathbf{e}_z \otimes \mathbf{e}_z \; + \\
& + B_{r\theta}(\mathbf{e}_r \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_r) \; + \; B_{rz}(\mathbf{e}_r \otimes \mathbf{e}_z + \mathbf{e}_z \otimes \mathbf{e}_r) \; + \\
& + B_{\theta z}(\mathbf{e}_z \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_z),
\end{aligned}
\tag{2.24}
$$

where

$$
\left.
\begin{aligned}
B_{rr} \quad &= \left(\frac{\partial r}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial r}{\partial \Theta}\right)^2 + \left(\frac{\partial r}{\partial Z}\right)^2, \\
B_{\theta\theta} \quad &= r^2\left[\left(\frac{\partial \theta}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial \theta}{\partial \Theta}\right)^2 + \left(\frac{\partial \theta}{\partial Z}\right)^2\right], \\
B_{zz} \quad &= \left(\frac{\partial z}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial z}{\partial \Theta}\right)^2 + \left(\frac{\partial z}{\partial Z}\right)^2, \\
B_{r\theta} \quad &= B_{\theta r} = r\left[\frac{\partial r}{\partial R}\frac{\partial \theta}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial \theta}{\partial \Theta} + \frac{\partial r}{\partial Z}\frac{\partial \theta}{\partial Z}\right], \\
B_{rz} \quad &= B_{zr} = \frac{\partial r}{\partial R}\frac{\partial z}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial z}{\partial \Theta} + \frac{\partial r}{\partial Z}\frac{\partial z}{\partial Z}, \\
B_{\theta z} \quad &= B_{z\theta} = r\left[\frac{\partial \theta}{\partial R}\frac{\partial z}{\partial R} + \frac{1}{R^2}\frac{\partial z}{\partial \Theta}\frac{\partial \theta}{\partial \Theta} + \frac{\partial \theta}{\partial Z}\frac{\partial z}{\partial Z}\right].
\end{aligned}
\right\}
\tag{2.25}
$$

$$
\begin{aligned}
\mathbf{F} \quad &= \frac{\partial r}{\partial R}(\mathbf{e}_r \otimes \mathbf{e}_R) + \frac{1}{R}\frac{\partial r}{\partial \Theta}(\mathbf{e}_r \otimes \mathbf{e}_\Theta) + \frac{1}{R\sin\Theta}\frac{\partial r}{\partial \Phi}(\mathbf{e}_r \otimes \mathbf{e}_\Phi) + \\
&+ r\frac{\partial \theta}{\partial R}(\mathbf{e}_\theta \otimes \mathbf{e}_R) + \frac{r}{R}\frac{\partial \theta}{\partial \Theta}(\mathbf{e}_\theta \otimes \mathbf{e}_\Theta) + \frac{r}{R\sin\Theta}\frac{\partial \theta}{\partial \Phi}(\mathbf{e}_\theta \otimes \mathbf{e}_\Phi) + \\
&+ r\sin\theta\frac{\partial \varphi}{\partial R}(\mathbf{e}_\varphi \otimes \mathbf{e}_R) \; + \frac{r\sin\theta}{R}\frac{\partial \varphi}{\partial \Theta}(\mathbf{e}_\varphi \otimes \mathbf{e}_\Theta) \; + \frac{r\sin\theta}{R\sin\Theta}\frac{\partial \varphi}{\partial \Phi}(\mathbf{e}_\varphi \otimes \mathbf{e}_\Phi)\,.
\end{aligned}
\tag{2.26}
$$

$$
\begin{aligned}
\mathbf{B} = \quad & B_{rr}\mathbf{e}_r \otimes \mathbf{e}_r \; + \; B_{\vartheta\vartheta}\mathbf{e}_\vartheta \otimes \mathbf{e}_\vartheta \; + \; B_{\varphi\varphi}\mathbf{e}_\varphi \otimes \mathbf{e}_\varphi \; + \\
& + B_{r\vartheta}(\mathbf{e}_r \otimes \mathbf{e}_\vartheta + \mathbf{e}_\vartheta \otimes \mathbf{e}_r) \; + \; B_{r\varphi}(\mathbf{e}_r \otimes \mathbf{e}_\varphi + \mathbf{e}_\varphi \otimes \mathbf{e}_r) \; + \\
& + B_{\vartheta\varphi}(\mathbf{e}_\varphi \otimes \mathbf{e}_\vartheta + \mathbf{e}_\vartheta \otimes \mathbf{e}_\varphi),
\end{aligned}
\tag{2.27}
$$

where

$$\left.\begin{aligned}
B_{rr} &= \left(\frac{\partial r}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial r}{\partial \Theta}\right)^2 + \frac{1}{R^2 \sin^2\Theta}\left(\frac{\partial r}{\partial \Phi}\right)^2, \\
B_{\vartheta\vartheta} &= r^2 \left[\left(\frac{\partial \vartheta}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial \vartheta}{\partial \Theta}\right)^2 + \frac{1}{R^2 \sin^2\Theta}\left(\frac{\partial \vartheta}{\partial \Phi}\right)^2\right], \\
B_{\varphi\varphi} &= r^2 \sin^2 \vartheta \left[\left(\frac{\partial \varphi}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial \varphi}{\partial \Theta}\right)^2 + \frac{1}{R^2 \sin^2\Theta}\left(\frac{\partial \varphi}{\partial \Phi}\right)^2\right] \\
B_{r\vartheta} &= B_{\vartheta r} = r \left[\frac{\partial r}{\partial R}\frac{\partial \vartheta}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial \vartheta}{\partial \Theta} + \frac{1}{R^2 \sin^2\Theta}\frac{\partial r}{\partial \Phi}\frac{\partial \vartheta}{\partial \Phi}\right], \\
B_{r\varphi} &= B_{\varphi r} = r \sin \vartheta \left[\frac{\partial r}{\partial R}\frac{\partial \varphi}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial \varphi}{\partial \Theta} + \frac{1}{R^2 \sin^2\Theta}\frac{\partial r}{\partial \Phi}\frac{\partial \varphi}{\partial \Phi}\right], \\
B_{\vartheta\varphi} &= B_{\varphi\vartheta} = r^2 \sin \vartheta \left[\frac{\partial \vartheta}{\partial R}\frac{\partial \varphi}{\partial R} + \frac{1}{R^2}\frac{\partial \varphi}{\partial \Theta}\frac{\partial \vartheta}{\partial \Theta} + \frac{1}{R^2 \sin^2\Theta}\frac{\partial \vartheta}{\partial \Phi}\frac{\partial \varphi}{\partial \Phi}\right].
\end{aligned}\right\} \tag{2.28}$$

---

3. Traction, stress, equilibrium.

$$\mathbf{t}(\mathbf{n}) = \mathbf{T}\mathbf{n} \tag{3.1}$$

$$T_{ij} = t_i(\mathbf{e}_j) = \mathbf{T}\mathbf{e}_j \cdot \mathbf{e}_i, \tag{3.2}$$

$$\mathbf{T} = T_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \qquad \mathbf{T} = \sum_{i=1}^{3} \tau_i\, \mathbf{t}_i \otimes \mathbf{t}_i. \tag{3.3}$$

$$T_{\mathrm{normal}}(\mathbf{n}) = \mathbf{t}(\mathbf{n}) \cdot \mathbf{n} \ = \mathbf{T}\mathbf{n} \cdot \mathbf{n} = T_{ij} n_i n_j = \tau_1\, n_1^2 + \tau_2\, n_2^2 + \tau_3\, n_3^2 \tag{3.4}$$

$$T_{\mathrm{shear}}(\mathbf{n}) = \sqrt{|\mathbf{t}|^2 - T_{\mathrm{normal}}^2} = \sqrt{\big[\mathbf{t}(\mathbf{n}) \cdot \mathbf{t}(\mathbf{n})\big] - \big[\mathbf{t}(\mathbf{n}) \cdot \mathbf{n}\big]^2}. \tag{3.5}$$

$$T_{\mathrm{shear}}^2 = |\mathbf{t}(\mathbf{n})|^2 - T_{\mathrm{normal}}^2 = \tau_1^2 n_1^2 + \tau_2^2 n_2^2 + \tau_3^2 n_3^2 - (\tau_1 n_1^2 + \tau_2 n_2^2 + \tau_3 n_3^2)^2. \tag{3.6}$$

$$\mathrm{div}\, \mathbf{T} + \mathbf{b} = \mathbf{0}, \qquad \big(\mathrm{div}\, \mathbf{T}\big)_i = \frac{\partial T_{ij}}{\partial y_j} \tag{3.7}$$

$$\frac{\partial T_{ij}}{\partial y_j} + b_i \ = \ 0\ , \tag{3.8}$$

$$\mathbf{T} = \mathbf{T}^T \tag{3.9}$$

$$\text{Pure shear stress:} \quad \mathbf{T} = \tau(\mathbf{m} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{m}) \tag{3.10}$$

$$\text{Uniaxial stress:} \quad \mathbf{T} = \sigma\ \mathbf{m} \otimes \mathbf{m} \tag{3.11}$$

$$\mathbf{s}\, dA_x = \mathbf{t}\, dA_y \tag{3.12}$$

$$\mathbf{s} = \mathbf{S}\mathbf{n}_{\mathrm{R}} \tag{3.13}$$

$$\mathbf{S}\mathbf{n}_{\mathrm{R}}\, dA_x = \mathbf{T}\mathbf{n}\, dA_y \tag{3.14}$$

$$1^{\mathrm{st}} \text{ Piola Kirchhoff stress tensor:} \qquad \mathbf{S} = J\mathbf{T}\mathbf{F}^{-T} \tag{3.15}$$

$$\mathbf{S} = S_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \qquad S_{ij} = s_i(\mathbf{e}_j) \tag{3.16}$$

$$\mathrm{Div}\, \mathbf{S} + \mathbf{b}_{\mathrm{R}} = \mathbf{0}, \qquad \big(\mathrm{Div}\, \mathbf{S}\big)_i = \frac{\partial S_{ij}}{\partial x_j} \tag{3.17}$$

$$\frac{\partial S_{ij}}{\partial x_j} + b_i^{\mathrm{R}} \ = \ 0\ , \tag{3.18}$$

$$\mathbf{S}\mathbf{F}^T = \mathbf{F}\mathbf{S}^T \tag{3.19}$$

$$\text{Stress power density} = \mathbf{S} \cdot \dot{\mathbf{F}} = J\ \mathbf{T} \cdot \mathbf{D} \tag{3.20}$$

$$\mathbf{L} = \mathrm{grad}\, \mathbf{v}, \qquad L_{ij} = \frac{\partial v_i}{\partial y_j}. \tag{3.21}$$

$$\mathbf{D} = \frac{1}{2}\left(\mathrm{grad}\, \mathbf{v} + (\mathrm{grad}\, \mathbf{v})^T\right), \qquad D_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial y_j} + \frac{\partial v_j}{\partial y_i}\right) \tag{3.22}$$

$$\dot{\mathbf{F}} = \mathbf{L}\mathbf{F} \tag{3.23}$$

$$\text{Biot stress tensor:} \qquad \mathbf{S}^{(1)} = \frac{1}{2}(\mathbf{S}\mathbf{R} + \mathbf{R}^T\mathbf{S}) \tag{3.24}$$

$$2^{\mathrm{nd}} \text{ Piola Kirchhoff stress tensor:} \qquad \mathbf{S}^{(2)} = J\mathbf{F}^{-1}\mathbf{T}\mathbf{F}^{-T} = \mathbf{F}^{-1}\mathbf{S} \tag{3.25}$$

$$
\begin{aligned}
\frac{\partial T_{rr}}{\partial r} + \frac{1}{r}\frac{\partial T_{r\theta}}{\partial \theta} + \frac{\partial T_{rz}}{\partial z} + \frac{T_{rr} - T_{\theta\theta}}{r} + b_r &= 0, \\
\frac{\partial T_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial \theta} + \frac{\partial T_{\theta z}}{\partial z} + \frac{T_{r\theta} + T_{\theta r}}{r} + b_\theta &= 0, \\
\frac{\partial T_{zr}}{\partial r} + \frac{1}{r}\frac{\partial T_{z\theta}}{\partial \theta} + \frac{\partial T_{zz}}{\partial z} + \frac{T_{zr}}{r} + b_z &= 0\ ,
\end{aligned}
\tag{3.26}
$$

$$
\begin{aligned}
\frac{\partial T_{rr}}{\partial r} + \frac{1}{r}\frac{\partial T_{r\phi}}{\partial \phi} + \frac{1}{r\sin\phi}\frac{\partial T_{r\theta}}{\partial \theta} + \frac{2T_{rr} - T_{\phi\phi} - T_{\theta\theta} + T_{r\phi}\cot\phi}{r} + b_r &= 0, \\
\frac{\partial T_{r\theta}}{\partial r} + \frac{1}{r}\frac{\partial T_{\phi\theta}}{\partial \phi} + \frac{1}{r\sin\phi}\frac{\partial T_{\theta\theta}}{\partial \theta} + \frac{3T_{r\theta} + 2T_{\theta\phi}\cot\phi}{r} + b_\theta &= 0, \\
\frac{\partial T_{r\phi}}{\partial r} + \frac{1}{r}\frac{\partial T_{\phi\phi}}{\partial \phi} + \frac{1}{r\sin\phi}\frac{\partial T_{\phi\theta}}{\partial \theta} + \frac{3T_{r\phi} + (T_{\phi\phi} - T_{\theta\theta})\cot\phi}{r} + b_\phi &= 0.
\end{aligned}
\tag{3.27}
$$

4. Constitutive relation.

$$\mathbf{S}\cdot\dot{\mathbf{F}} = \dot{W} \tag{4.1}$$

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}}, \qquad \mathbf{T} = \mathrm{J}^{-1}\mathbf{S}\mathbf{F}^T = \frac{1}{J}\frac{\partial W}{\partial \mathbf{F}}\,\mathbf{F}^T. \tag{4.2}$$

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{Q}\mathbf{F}) \tag{4.3}$$

$$\widehat{W}(\mathbf{F}) = \overline{W}(\mathbf{C}) \qquad \text{where } \mathbf{C} = \mathbf{F}^T\mathbf{F}. \tag{4.4}$$

$$\mathbf{S} = 2\mathbf{F}\,\frac{\partial \overline{W}}{\partial \mathbf{C}}, \qquad \mathbf{T} = \frac{2}{J}\,\mathbf{F}\,\frac{\partial \overline{W}}{\partial \mathbf{C}}\,\mathbf{F}^T. \tag{4.5}$$

$$\mathcal{G} = \{\mathbf{Q} :\ \mathbf{Q}\mathbf{Q}^T = \mathbf{I},\ \det\mathbf{Q} = 1,\ \widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q}) \text{ for all nonsingular } \mathbf{F}\}.$$

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q}) \quad \text{for all } \mathbf{Q}\in\mathcal{G} \text{ and all nonsingular } \mathbf{F}. \tag{4.6}$$

$$\overline{W}(\mathbf{C}) = \overline{W}\big(\mathbf{Q}\mathbf{C}\mathbf{Q}^T\big) \quad \text{for all } \mathbf{Q}\in\mathcal{G} \text{ and all symmetric positive definite } \mathbf{C}. \tag{4.7}$$

$$W = \widetilde{W}\big(I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C})\big) \tag{4.8}$$

$$\left.\begin{aligned}
\mathbf{T} &= 2J\frac{\partial \widetilde{W}}{\partial I_3}\,\mathbf{I} \;+\; \frac{2}{J}\left[\frac{\partial \widetilde{W}}{\partial I_1} + I_1\frac{\partial \widetilde{W}}{\partial I_2}\right]\mathbf{B} \;-\; \frac{2}{J}\frac{\partial \widetilde{W}}{\partial I_2}\mathbf{B}^2,\\
\mathbf{S} &= 2I_3\frac{\partial \widetilde{W}}{\partial I_3}\,\mathbf{F}^{-T} \;+\; 2\left[\frac{\partial \widetilde{W}}{\partial I_1} + I_1\frac{\partial \widetilde{W}}{\partial I_2}\right]\mathbf{F} \;-\; 2\frac{\partial \widetilde{W}}{\partial I_2}\mathbf{B}\mathbf{F}.
\end{aligned}\right\} \tag{4.9}$$

$$I_1 = \lambda_1^2+\lambda_2^2+\lambda_3^2, \qquad I_2 = \lambda_1^2\lambda_2^2+\lambda_2^2\lambda_3^2+\lambda_3^2\lambda_1^2, \qquad I_3 = \lambda_1^2\lambda_2^2\lambda_3^2. \tag{4.10}$$

$$W = W^*(\lambda_1,\lambda_2,\lambda_3) \tag{4.11}$$

$$W^*(\lambda_1,\lambda_2,\lambda_3) = W^*(\lambda_2,\lambda_1,\lambda_3) = W^*(\lambda_1,\lambda_3,\lambda_2) = \dots. \tag{4.12}$$

$$\mathbf{T} = \tau_1\,\boldsymbol{\ell}_1\otimes\boldsymbol{\ell}_1 + \tau_2\,\boldsymbol{\ell}_2\otimes\boldsymbol{\ell}_2 + \tau_3\,\boldsymbol{\ell}_3\otimes\boldsymbol{\ell}_3 \tag{4.13}$$

$$\tau_1 = \frac{\lambda_1}{\lambda_1\lambda_2\lambda_3}\,\frac{\partial W^*}{\partial\lambda_1}, \qquad \tau_2 = \frac{\lambda_2}{\lambda_1\lambda_2\lambda_3}\,\frac{\partial W^*}{\partial\lambda_2}, \qquad \tau_3 = \frac{\lambda_3}{\lambda_1\lambda_2\lambda_3}\,\frac{\partial W^*}{\partial\lambda_3}. \tag{4.14}$$

$$\mathbf{S} = \sigma_1\,\boldsymbol{\ell}_1\otimes\mathbf{r}_1 + \sigma_2\,\boldsymbol{\ell}_2\otimes\mathbf{r}_2 + \sigma_3\,\boldsymbol{\ell}_3\otimes\mathbf{r}_3 \tag{4.15}$$

$$\sigma_1 = \frac{\partial W^*}{\partial\lambda_1}, \qquad \sigma_2 = \frac{\partial W^*}{\partial\lambda_2}, \qquad \sigma_3 = \frac{\partial W^*}{\partial\lambda_3}. \tag{4.16}$$

$$\dot{\phi} = \frac{\partial\phi}{\partial\mathbf{F}}\cdot\dot{\mathbf{F}} = 0, \tag{4.17}$$

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}} + q\,\frac{\partial\phi}{\partial\mathbf{F}}, \qquad \mathbf{T} = \frac{1}{J}\,\frac{\partial W}{\partial \mathbf{F}}\,\mathbf{F}^T + \frac{q}{J}\,\frac{\partial\phi}{\partial\mathbf{F}}\,\mathbf{F}^T. \tag{4.18}$$

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}} - q\mathbf{F}^{-T}, \qquad \mathbf{T} = \frac{\partial W}{\partial \mathbf{F}}\mathbf{F}^T - q\mathbf{I} \tag{4.19}$$

$$\mathbf{T} = -q\,\mathbf{I} \;+\; 2\left[\frac{\partial \widetilde{W}}{\partial I_1} + I_1\frac{\partial \widetilde{W}}{\partial I_2}\right]\mathbf{B} \;-\; 2\frac{\partial \widetilde{W}}{\partial I_2}\mathbf{B}^2, \tag{4.20}$$

$$\mathbf{S} = -q\,\mathbf{F}^{-T} \;+\; 2\left[\frac{\partial \widetilde{W}}{\partial I_1} + I_1\frac{\partial \widetilde{W}}{\partial I_2}\right]\mathbf{F} \;-\; 2\frac{\partial \widetilde{W}}{\partial I_2}\mathbf{B}\mathbf{F}. \tag{4.21}$$

$$W = W^*(\lambda_1, \lambda_2, \lambda_3) \tag{4.22}$$

$$W^*(\lambda_1, \lambda_2, \lambda_3) = W^*(\lambda_2, \lambda_1, \lambda_3) = W^*(\lambda_1, \lambda_3, \lambda_2) = \dots. \tag{4.23}$$

$$\tau_1 = \lambda_1 \frac{\partial W^*}{\partial \lambda_1} - q, \qquad \tau_2 = \lambda_2 \frac{\partial W^*}{\partial \lambda_2} - q, \qquad \tau_3 = \lambda_3 \frac{\partial W^*}{\partial \lambda_3} - q. \tag{4.24}$$

$$\overline{W}(\mathbf{C}, \mathbf{M}) = \widetilde{W}(I_1, I_2, I_3, I_4, I_5), \qquad W_i = \frac{\partial \widetilde{W}}{\partial I_i},\ i = 1, 2, 3, 4, 5. \tag{4.25}$$

$$I_4(\mathbf{C}, \mathbf{M}) = \mathbf{C} \cdot \mathbf{M}, \qquad I_5(\mathbf{C}, \mathbf{M}) = \mathbf{C}^2 \cdot \mathbf{M}. \qquad \mathbf{M} = \mathbf{m}_{\mathrm{R}} \otimes \mathbf{m}_{\mathrm{R}} \tag{4.26}$$

$$\begin{aligned} \mathbf{T} \quad &= 2JW_3\,\mathbf{I} \ + \ \frac{2}{J}\,[W_1 + I_1 W_2]\ \mathbf{B} \ - \ \frac{2}{J} W_2 \mathbf{B}^2 + \\ &+ \frac{2}{J} W_4 (\mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}_{\mathrm{R}}) \ + \ \frac{2}{J}\, W_5 \Big[(\mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{B}\mathbf{F}\mathbf{m}_{\mathrm{R}}) + (\mathbf{B}\mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}_{\mathrm{R}}) \Big]\ , \end{aligned} \tag{4.27}$$

$$\overline{W}(\mathbf{C}, \mathbf{M}) = \widetilde{W}(I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8), \qquad W_i = \frac{\partial \widetilde{W}}{\partial I_i},\ i = 1, 2, \dots 8. \tag{4.28}$$

$$I_6 = \mathbf{C}\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}'_{\mathrm{R}} \qquad I_7 = \mathbf{C}^2 \mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}'_{\mathrm{R}}, \qquad I_8 = \mathbf{C}\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}. \tag{4.29}$$

$$\begin{aligned} \mathbf{T} \quad &= -q\mathbf{I} + 2W_1\mathbf{B} + 2W_2(I_1\mathbf{B} - \mathbf{B}^2) + \\ &+ 2W_4 \mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}_{\mathrm{R}} + 2W_6 \mathbf{F}\mathbf{m}'_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}'_{\mathrm{R}} + \\ &+ 2W_5(\mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{B}\mathbf{F}\mathbf{m}_{\mathrm{R}} + \mathbf{B}\mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}_{\mathrm{R}}) + 2W_7(\mathbf{F}\mathbf{m}'_{\mathrm{R}} \otimes \mathbf{B}\mathbf{F}\mathbf{m}'_{\mathrm{R}} + \mathbf{B}\mathbf{F}\mathbf{m}'_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}'_{\mathrm{R}}) + \\ &+ W_8(\mathbf{F}\mathbf{m}_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}'_{\mathrm{R}} + \mathbf{F}\mathbf{m}'_{\mathrm{R}} \otimes \mathbf{F}\mathbf{m}_{\mathrm{R}}) \end{aligned} \tag{4.30}$$

Contents

Chapter 1

BRIEF REVIEW OF MATHEMATICAL PRELIMINARIES

When studying the response of a body subjected to some loading, we will encounter entities such as displacement $\mathbf{u}$ and traction $\mathbf{t}$ that are vectors, and quantities such as deformation gradient $\mathbf{F}$ and stress $\mathbf{S}$ that are tensors. We will need to carry out various calculations involving them that require us to use *vector and tensor algebra*. We will sometimes work with the components of these vectors and tensors in a basis, and these are represented as column and square *matrices* respectively, e.g. $\{u\}, \{t\}, [F]$ and $[S]$. Calculations involving matrices can often be carried out expeditiously using *indicial notation*. A typical particle of the body is described by its position vector $\mathbf{x}$ and the displacement vector, for example, varies from particle to particle and is therefore a function of position: $\mathbf{u}(\mathbf{x})$. How displacement varies with position requires us to consider the gradient of the displacement with respect to position, $\nabla\mathbf{u}$, and for such calculations we must rely on *vector and tensor calculus*.

What follows is mostly a list of definitions and properties pertaining to the main mathematical concepts and methods that we will use in these notes. Some proofs are given in the worked examples and exercises. A more detailed treatment of this material can be found in Volume I as well as in the references listed at the end of this chapter. The reader who is familiar with the material in Chapters 1-6 of Volume I can skip this chapter entirely.

The four main topics to be reviewed are (not in this order)

– vector and tensor algebra,

– representation of vectors and tensors in terms of matrices (having chosen a basis),

– the use of indicial notation to simplify calculations, and the

– calculus of vector and tensor fields and of functions of tensors.

Links to some short introductory videos on indicial notation and tensor algebra can be found in Sections 1.2 and 1.3 respectively.

In this chapter we will (*almost*) always use the following notational convention:

		Lowercase Greek letters:
α	.....	scalar
		Lowercase latin letters:
$\{a\}$	.....	3×1 column matrix
$\mathbf{a}$	.....	vector
a_i	.....	i^{th} component of the vector $\mathbf{a}$ in some basis; or
		i^{th} element of the column matrix $\{a\}$
		Uppercase latin letters:
$[A]$	.....	3×3 square matrix
$\mathbf{A}$	.....	second-order tensor (2-tensor) (linear transformation)
A_{ij}	.....	i, j component of the 2-tensor $\mathbf{A}$ in some basis; or
		i, j element of the square matrix $[A]$
		Blackboard bold letters:
$\mathbb{C}$	.....	fourth-order tensor (4-tensor)
$\mathbb{C}_{ijk\ell}$	.....	i, j, k, ℓ component of $\mathbb{C}$ in some basis

While we will closely follow this same notational convention in the *subsequent* chapters as well, there will be a few notable exceptions: for example, we will use the symbol W rather than a lower case Greek letter to denote the (scalar-valued) strain energy function.

1.1 Matrices.

Our discussion here is limited to 3×1 column matrices and 3×3 square matrices.

The element[1] in the ith row and jth column of a 3×3 matrix $[A]$ is denoted by A_{ij} and

[1]We speak of the *elements* of a matrix and the *components* of a vector or tensor.

the element in the ith row of a 3×1 column matrix $\{x\}$ is denoted by x_i:

$$[A] = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}, \qquad \{x\} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

– The magnitude of a column matrix $\{x\}$ is

$$|\{x\}| := (x_1^2 + x_2^2 + x_3^2)^{1/2}, \tag{1.1}$$

and the magnitude of a square matrix $[A]$ is

$$|[A]| := \left[A_{11}^2 + A_{12}^2 + A_{13}^2 + A_{21}^2 + \ldots + A_{33}^2\right]^{1/2}. \tag{1.2}$$

– The product of a 3×3 matrix $[A]$ with a 3×1 matrix $\{x\}$ is a third 3×1 matrix $\{y\}$ whose element in the ith row is the sum of the pairwise products of the elements in the ith row of $[A]$ and the elements of $\{x\}$:

$$\{y\} = [A]\{x\} \qquad \Rightarrow \qquad y_i = \sum_{k=1}^{3} A_{ik} x_k \qquad \text{for each } i = 1, 2, 3. \tag{1.3}$$

– The product of two matrices $[A]$ and $[B]$ is a third matrix $[C]$ whose element in the ith row and jth column is the sum of the pairwise products of the elements in the ith row of $[A]$ and the jth column of $[B]$:

$$[C] = [A][B] \qquad \Rightarrow \qquad C_{ij} = \sum_{k=1}^{3} A_{ik} B_{kj} \qquad \text{for each } i, j = 1, 2, 3. \tag{1.4}$$

While it may be preferable to say "for each $i = 1, 2, 3$ and each $j = 1, 2, 3$" we write it as above for brevity.

– The product of two matrices is not commutative in general: $[A][B] \neq [B][A]$.

– The *transpose* of a matrix $[A]$ is denoted by $[A]^T$. If the element in the ith row and jth column of $[A]$ is A_{ij}, then the element in the ith row and jth column of $[A]^T$ is A_{ji}.

– The transpose of the product of two matrices has the property

$$([A][B])^T = [B]^T[A]^T. \tag{1.5}$$

– A matrix is *symmetric* if

$$[A] = [A]^T, \qquad\qquad A_{ij} = A_{ji} \quad \text{for each } i, j = 1, 2, 3, \tag{1.6}$$

and *skew- (or anti)-symmetric* if

$$[A] = -[A]^T, \qquad\qquad A_{ij} = -A_{ji} \quad \text{for each } i, j = 1, 2, 3. \tag{1.7}$$

If $[A]$ is skew-symmetric, if follows from $(1.7)_2$ that $A_{11} = A_{22} = A_{33} = 0$ and $A_{12} = -A_{21}, A_{23} = -A_{32}, A_{31} = -A_{13}$. Therefore there are only three independent elements in a skew-symmetric matrix and so there is a one-to-one correspondence between skew-symmetric matrices and column matrices.

– Every matrix can be uniquely decomposed into the sum of a symmetric and skew-symmetric matrix:

$$\begin{aligned} &[A] = [S] + [W] \qquad \text{where } [S] = \tfrac{1}{2}([A] + [A]^T), \quad [W] = \tfrac{1}{2}([A] - [A]^T); \\ &A_{ij} = S_{ij} + W_{ij} \qquad \text{where } S_{ij} = \tfrac{1}{2}(A_{ij} + A_{ji}), \quad W_{ij} = \tfrac{1}{2}(A_{ij} - A_{ji}); \end{aligned} \tag{1.8}$$

the second row of (1.8) holds for each $i, j = 1, 2, 3$ and so represents 9 scalar equations.

– The trace and determinant are two scalar-valued functions of a matrix that are encountered frequently. They are defined by

$$\text{tr}\,[A] := A_{11} + A_{22} + A_{33} = \sum_{k=1}^{3} A_{kk}, \tag{1.9}$$

$$\begin{aligned} \det\,[A] \;\; := \;\; & A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + \\ & + A_{13}(A_{21}A_{32} - A_{22}A_{31}). \end{aligned} \tag{1.10}$$

– The determinant of the product of two matrices equals the product of the individual determinants of the two matrices:

$$\det([A][B]) = \det[A]\,\det[B]. \tag{1.11}$$

The determinant of a matrix is unchanged by transposition:

$$\det\left([A]^T\right) = \det[A]. \tag{1.12}$$

– The *identity matrix* $[I]$ has the property

$$[A][I] = [I][A] = [A], \qquad [I]\{x\} = \{x\}$$

for all square matrices $[A]$ and column matrices $\{x\}$. Also, $\det[I] = 1$ and $\text{tr}[I] = 3$.

- A matrix $[A]$ is *nonsingular (or invertible)* if $\det[A] \neq 0$; *singular* if $\det[A] = 0$.

  If $[A]$ is *nonsingular*, then the only column matrix $\{x\}$ for which $[A]\{x\} = \{0\}$ is $\{x\} = \{0\}$.

  If $[A]$ is *nonsingular*, it is invertible in the sense that there is a matrix denoted by $[A]^{-1}$ and called the *inverse* of $[A]$ for which

$$[A][A]^{-1} = [I], \qquad [A]^{-1}[A] = [I]. \tag{1.13}$$

  The inverse of the product of two nonsingular matrices obeys

$$\big([A][B]\big)^{-1} = [B]^{-1}[A]^{-1}. \tag{1.14}$$

- A matrix $[Q]$ is *orthogonal* if it is nonsingular and

$$[Q]^{-1} = [Q]^T. \tag{1.15}$$

  It follows that

$$[Q][Q]^T = [I], \quad [Q]^T[Q] = [I], \tag{1.16}$$

$$\det\,[Q] = \pm 1. \tag{1.17}$$

  An orthogonal matrix whose determinant is $+1$ is said to be *proper orthogonal* and represents a rotation. An orthogonal matrix whose determinant is -1 is said to be *improper orthogonal* and represents a reflection.

- If $\{y\}$ is a 3×1 (column) matrix, then $\{y\}^T$ is the associated 1×3 (row) matrix. The element in the ith column of $\{y\}^T$ equals the element in the ith row of $\{y\}$. If $\{x\}$ is a second column matrix, then

$$\{y\}^T\{x\} := y_1x_1 + y_2x_2 + y_3x_3 = \sum_{i=1}^{3} y_ix_i. \tag{1.18}$$

- The column matrices $\{x\}$ and $\{y\}$ are said to be *orthogonal* if

$$\{y\}^T\{x\} = \sum_{i=1}^{3} y_ix_i = 0. \tag{1.19}$$

- A matrix $[A]$ is *positive definite* if

$$\{x\}^T[A]\{x\} \;=\; \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ij}x_ix_j > 0 \tag{1.20}$$

  for all nonzero column matrices $\{x\}$.

– A scalar α and a column matrix $\{a\}$ $(\neq \{0\})$ for which

$$[A]\{a\} = \alpha\{a\} \tag{1.21}$$

are said to be an *eigenvalue* and *eigen"vector"* of $[A]$.

– If $[A]$ is symmetric, then it has three real eigenvalues $\alpha_1, \alpha_2, \alpha_3$ and three corresponding eigenvectors $\{a^{(1)}\}, \{a^{(2)}\}, \{a^{(3)}\}$. Without loss of generality the eigenvectors can always be chosen so each has unit magnitude and each pair is mutually orthogonal in the sense that

$$\{a^{(i)}\}^T\{a^{(j)}\} = a_1^{(i)}a_1^{(j)} + a_2^{(i)}a_2^{(j)} + a_3^{(i)}a_3^{(j)} = \sum_{k=1}^{3} a_k^{(i)}a_k^{(j)} = \begin{cases} 1 & \text{for} \quad i = j, \\ 0 & \text{for} \quad i \neq j. \end{cases}$$

1.2 Indicial notation.

Three brief videos on indicial notation can be found here.

Indicial notation is convenient when carrying out calculations involving the elements of matrices. When doing so, it is important that one adhere to certain rules/conventions.

The following terminology will be encountered in our discussion below:

– Free index,

– Dummy (or repeated) index,

– Range convention,

– Summation convention, and

– Substitution rule.

– Consider matrices $[A]$, $\{x\}$ and $\{y\}$ satisfying the matrix equation

$$\{y\} = [A]\{x\} \qquad \Leftrightarrow \qquad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{1.22}$$

On carrying out the matrix multiplication, this is equivalent to the system of 3 scalar equations

$$\left.\begin{aligned} y_1 &= A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = \sum_{k=1}^{3} A_{1k}x_k, \\ y_2 &= A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = \sum_{k=1}^{3} A_{2k}x_k, \\ y_3 &= A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = \sum_{k=1}^{3} A_{3k}x_k. \end{aligned}\right\} \tag{1.23}$$

This system of scalar equations can be written more compactly as

$$y_i = \sum_{k=1}^{3} A_{ik}x_k \quad \text{with } i \text{ taking each value in the range } 1,2,3. \tag{1.24}$$

– **Range convention:** We can write (1.24) even more compactly by omitting the phrase "with i taking each value in the range $1, 2, 3$" and simply writing

$$y_i = \sum_{k=1}^{3} A_{ik}x_k \tag{1.25}$$

with the understanding that (1.25) *holds for each value of the index i in the range* $i = 1, 2, 3$. This understanding is referred to as the *range convention.*

Likewise the matrix equation $[C] = [A][B]$ can be written by (1.4) as

$$C_{ij} = \sum_{k=1}^{3} A_{ik}B_{kj}, \tag{1.26}$$

having dropped the phrase "with i and j taking each value in the range $1, 2, 3$".

From here on we shall always use the range convention unless explicitly stated otherwise.

– Observe the distinction between the two types of indices: free indices and repeated (or dummy) indices. The index i in (1.25) is called a *free index* because it is free to take on each value in the range $1, 2, 3$, one at a time. Equation (1.26) involves two free indices i and j, and each, independently, takes each value in the range $1, 2, 3$.

On the other hand the index k (in both equations) is *not* a free index: it is summed over $1, 2, 3$ and is not free to take each value $1, 2, 3$ *one at a time.* Since this index appears twice (in the terms on the right-hand sides), it is called a *repeated index* or (for reasons that will soon become clear) a *dummy index.*

– Now consider the set of equations

$$y_j = \sum_{k=1}^{3} A_{jk} x_k. \tag{1.27}$$

By the range convention, this holds with the free subscript j taking each value in the range $1, 2, 3$. Therefore the set of equations in (1.27) is identical to that in (1.25). This illustrates the fact that *the particular choice of index for the free subscript in an equation is not important provided that the same free subscript appears in every term*[2] *of the equation.*

Likewise the equation

$$C_{pq} = \sum_{k=1}^{3} A_{pk} B_{kq}, \tag{1.28}$$

is equivalent to (1.26) where we have simply used a different pair of free indices p, q.

– If an equation involves n free indices, then it represents 3^n scalar equations. For example (1.25) has 1 free subscript and it represents $3^1 = 3$ scalar equations, while (1.26) has 2 free subscripts and so represents $3^2 = 9$ equations.

– In order to be consistent it is important that *the same free index (or indices) appear once, and only once, in every term of an equation.* For example, the matrix equation $\{y\} = [A]\{x\} + [B]\{x\}$ can be written in scalar form as

$$y_i = \sum_{p=1}^{3} A_{ip} x_p \;+\; \sum_{q=1}^{3} B_{iq} z_q. \tag{1.29}$$

Here we have a free index i on the left-hand side and this same free index i appears in

[2]It is worth clarifying how the word "term" is used in this section. In an equation such as

$$y_i = \sum_{k=1}^{3} A_{ik} x_k + \sum_{q=1}^{3} a_q x_q z_i,$$

when we say "the first term on the right-hand side" we do not mean A_{ik} but rather $\sum A_{ik} x_k$. Two terms are separated by displayed $+, -$ or $=$ signs. If we wrote this out as

$$y_i = A_{i1} x_1 + A_{i2} x_2 + A_{i3} x_3 + \sum_{q=1}^{3} a_q x_q z_i,$$

the first term on the right-hand side would be $A_{i1} x_1$.

each of the terms on the right-hand side. We will never write an equation such as

$$y_i = \sum_{p=1}^{3} A_{jp}x_p \ + \ \sum_{q=1}^{3} B_{jq}z_q$$

which has the free index i on the left-hand side and the free index j on the right-hand side.

Similarly in (1.26), since the free indices i and j appear on the left-hand side they must also appear on the right-hand side. An equation such as $A_{ij} = B_{pq}$ would violate this consistency requirement.

Observe from (1.29) that the same repeated index does not need to appear in every term.

– **Summation convention:** Next, observe that on the right-hand side of equation (1.25) the subscript k is (a) repeated and (b) there is a sum over it. Likewise in the first term on the right-hand side of equation (1.29) the subscript p is repeated and there is a sum over it, and in the second term the subscript q is repeated and there is a sum over it. In each case there is a summation over the repeated index. An example involving 2 repeated indices is

$$A_{11}x_1^2 + A_{12}x_1x_2 + \ldots + A_{33}x_3^2 = \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ij}x_ix_j.$$

On the right-hand side the subscript i is repeated and there is a summation over it, and likewise the subscript j is repeated and there is a summation over it as well.

In view of this observation we can simplify our writing even further *by agreeing to drop the summation sign and instead imposing the rule that summation is implied over a subscript that appears twice in a term.* With this understanding in force, we would write (1.25), (1.26) and (1.29) as

$$y_i = A_{ik}x_k, \qquad C_{ij} = A_{ik}B_{kj}, \qquad y_i = A_{ip}x_p \ + \ B_{qi}z_q, \tag{1.30}$$

respectively with summation on the subscript k in the first and second, and on p and q in the third being implied.

– Note that we can write

$$\mathrm{tr}[A] = A_{ii}.$$

– Since

$$\sum_{k=1}^{3} A_{ik}x_k = \sum_{j=1}^{3} A_{ij}x_j,$$

it follows that

$$y_i = A_{ij}x_j \tag{1.31}$$

is identical to $(1.30)_1$. Thus we see that the particular choice of index for the repeated subscript is not important: it is a dummy index in this sense.

– In order to avoid ambiguity, no subscript is allowed to appear more than twice in any term in general. Thus we shall not write, for example, $A_{ii}x_i = y_i$ since, if we did, the index i would appear 3 times in the term on the left-hand side. We would not know whether the left-hand side denotes $(A_{11} + A_{22} + A_{33})x_i$ or $A_{11}x_1 + A_{22}x_2 + A_{33}x_3$ or perhaps there is no sum over i at all. On a few occasions, usually involving eigenvalues, we will forced to include a term with the same subscript appearing more than twice. In such cases we will make clear that the summation convention is being suspended and the summation is shown explicitly, e.g. see (1.109) where the subscript i appears 3 times.

– **Summary of Rules**:

1. Lower-case latin subscripts take on values in the range $1, 2, 3$.
2. A given index may appear either once or twice in each term of an equation. If it appears once, it is called a free index and it takes on each value in its range, one at a time. If it appears twice, it is called a dummy index and summation is implied over it.
3. The same index may not appear more than twice in the same term.
4. All terms of an equation must have the same free indices.

– *It is important to emphasize that an equation such as $A_{ij} = B_{ik}C_{kj}$ involves scalar quantities and therefore the order in which the scalar factors appear within a term is not significant.* For example

$$A_{ij} = B_{ik}C_{kj}, \tag{1.32}$$

comprises 3^2 scalar equations. Consider the element A_{11}. On setting $i = 1, j = 1$ in (1.32) we get

$$A_{11} = B_{1k}C_{k1} \qquad \Rightarrow \qquad A_{11} = B_{11}C_{11} + B_{12}C_{21} + B_{13}C_{31}.$$

Clearly we can rearrange terms in the right most expression and write

$$A_{11} = C_{11}B_{11} + C_{21}B_{12} + C_{31}B_{13} \qquad \Rightarrow \qquad A_{11} = C_{k1}B_{1k},$$

and so $A_{11} = B_{1k}C_{k1} = C_{k1}B_{1k}$, and more generally

$$A_{ij} = B_{ik}C_{kj} = C_{kj}B_{ik};$$

i.e. we can move B_{ik} to the back of $B_{ik}C_{kj}$ and write $C_{kj}B_{ik}$.

As a matrix equation $A_{ij} = B_{ik}C_{kj}$ corresponds to $[A] = [B][C]$ and so does $A_{ij} = C_{kj}B_{ik}$. The latter does *not* correspond to $[A] = [C][B]$ (as we shall explain below).

– Frequently, in the course of a calculation, we will have to change our choice of indices. For example suppose $y_i = A_{ij}x_j$ and we want to calculate y_iy_i (which is $y_1^2 + y_2^2 + y_3^2$). We *cannot* write $y_iy_i = (A_{ij}x_j)(A_{ij}x_j) = A_{ij}A_{ij}x_jx_j$ because we then have the index j appearing more than twice in the same term. Instead, we would use the equivalent alternative representations $y_i = A_{ij}x_j$ and $y_i = A_{ik}x_k$ to write $y_iy_i = (A_{ij}x_j)(A_{ik}x_k) = A_{ij}A_{ik}x_jx_k$. Observe that no index appears more than twice in each term.

– The indices in an expression can be changed without altering the meaning of an expression *provided that* (a) the positions of the free and repeated indices does not change and (b) one does not violate the preceding rules. Thus, for example, we can change the free index p on both sides of the equation

$$y_p \;=\; A_{pq}x_q$$

to any other index (except q, why not q?), say k, and equivalently write

$$y_k \;=\; A_{kq}x_q.$$

We can also change the repeated subscript q to some other index (except k), say p, and write

$$y_k \;=\; A_{kp}x_p.$$

In fact, we can even write

$$y_q \;=\; A_{qp}x_q.$$

The four preceding sets of equations are identical.

However, the equations $y_p \;=\; A_{qp}x_q$ and $y_p \;=\; A_{pq}x_q$ are *not* identical even though none of the indicial notation rules are violated. This is because the free and repeated indices do not appear in the same *positions*: the free index p on the right-hand side of $y_p = A_{qp}x_q$ is in the second position in A_{qp}, but it is in the first position in A_{pq} in $y_p = A_{pq}x_q$.

Similarly consider $A_{ij} = B_{ik}C_{kj}$. The fact that the second subscript of B_{ik} is the same as the first subscript of C_{kj} is what tells us that in matrix form this arises from $[A] = [B][C]$; see (1.4). In contrast $A_{ij} = B_{ki}C_{kj}$ does *not* represent $[A] = [B][C]$ because the repeated index is in a different position on the right-hand side. .

– We emphasize that while the factors in a term can be moved around, one cannot arbitrarily move indices. The *location of the indices determines the order in which the associated matrices are multiplied.*

When we write $A_{ij}x_j$ it is the second subscript of A_{ij} that also appears in x_j and this is what tells us that $A_{ij}x_j$ is the ith element of $[A]\{x\}$. For this reason $A_{ji}x_j$ is *not* the ith element of $[A]\{x\}$. But $A_{ji}x_j$ is the ith element of some matrix (since it has one free index i). But of what matrix? In order that this represent the product of a 3×3 matrix with a column matrix, the second subscript of the element associated with the 3×3 matrix must also appear in the element associated with the column matrix. To achieve this we can write $A_{ji}x_j = A^T_{ij}x_j$ which is now in the desired form. Therefore $A_{ji}x_j$ is the ith element of $[A^T]\{x\}$. Here $[A^T]$ is the transpose of $[A]$.

Similarly, observe in the matrix multiplication representation (1.32) that the second subscript of B_{ik} is the same as the first subscript of C_{kj}, whence $B_{ik}C_{kj}$ represents the i,j element of the matrix product $[B][C]$. Suppose we have the expression $B_{ki}C_{kj}$. To put it in the preceding form where adjacent subscripts are repeated, we can write $B_{ki}C_{kj} = B^T_{ik}C_{kj}$. The last subscript of B^T_{ik} is now the same as the first subscript of C_{kj} and so $B_{ki}C_{kj} = B^T_{ik}C_{kj}$ represents the i,j element of $[B]^T[C]$.

Alternatively we can write $B_{ki}C_{kj} = C_{kj}B_{ki} = C^T_{jk}B_{ki} = \big([C]^T[B]\big)_{ji}$. Therefore $B_{ki}C_{kj}$ is the j,i (not i,j) element of $[C]^T[B]$. These alternative representations are of course equivalent since $\big([B]^T[C]\big)^T = [C]^T[B]$ by (1.5).

Kronecker delta: The Kronecker delta, δ_{ij}, is defined as

$$\delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \qquad \text{for each } i,j = 1,2,3. \tag{1.33}$$

Observe that δ_{ij} is the element in the ith row and jth column of the identity matrix

$$[I] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\delta_{ii} = 3. \tag{1.34}$$

The Kronecker delta arises, for example, when working with a triplet $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of orthonormal vectors (page 22) and when calculating the derivative $\partial x_i/\partial x_j$ (page 18).

Substitution rule: The following useful property of the Kronecker delta is called the substitution rule. Suppose one wishes to simplify the expression $\delta_{ij}u_j$ for some column matrix $\{u\}$. First note that $\delta_{ij}u_j = \delta_{i1}u_1 + \delta_{i2}u_2 + \delta_{i3}u_3$ where i is a free subscript. Next, consider the choice $i = 1$. Then $\delta_{1j}u_j = \delta_{11}u_1 + \delta_{12}u_2 + \delta_{13}u_3$. Since $\delta_{ij} = 0$ unless $i = j$ and $\delta_{ij} = 1$ if $i = j$ we conclude that the last two terms on the right-hand side of $\delta_{1j}u_j = \delta_{11}u_1 + \cancel{\delta_{12}}u_2 + \cancel{\delta_{13}}u_3$ vanish and so it simplifies to $\delta_{1j}u_j = u_1$.

In a similar manner we see that for any value of the free index i, two terms on the right-hand side of $\delta_{ij}u_j = \delta_{i1}u_1 + \delta_{i2}u_2 + \delta_{i3}u_3$ vanish trivially because the two subscripts of the Kronecker delta in each of those terms will be distinct. The term that remains is the 1$^{\text{st}}$, 2$^{\text{nd}}$ or 3$^{\text{rd}}$ term depending on whether the free index $i = 1, 2$ or 3 respectively. Thus the term that survives on the right-hand side is u_i and so

$$\delta_{ij}u_j = u_i. \tag{1.35}$$

In summary, *(a)* since δ_{ij} is zero unless $j = i$, the expression being simplified has a non-zero value only if $j = i$; and *(b)* when $j = i$, δ_{ij} is unity. Thus we replace the Kronecker delta by unity and simultaneously change the repeated subscript j in the other factor to the non-repeated subscript i of the Kronecker delta. This gives $\delta_{ij}u_j = u_i$.

As a second example suppose that $[A]$ is a square matrix and one wishes to simplify $\delta_{ij}A_{jk}$. Then by the same reasoning[3],

$$\delta_{ij}A_{jk} = \delta_{i1}A_{1k} + \delta_{i2}A_{2k} + \delta_{i3}A_{3k} = A_{ik} \tag{1.36}$$

[3]Observe that these results are immediately apparent by using matrix algebra. In the first example, $\delta_{ij}u_j$ is simply the ith element of the column matrix $[I]\{u\}$. Since $[I]\{u\} = \{u\}$ the result follows at once. Similarly in the second example, $\delta_{ij}A_{jk}$ is simply the i,k-element of the matrix $[I][A]$. Since $[I][A] = [A]$, the result follows.

and so the Kronecker delta has been replaced by unity and the repeated subscript j in A_{jk} has been changed to i.

More generally, if some quantity or expression $\mathbb{T}_{pq\ldots i\ldots z}$ multiplies δ_{ij} with the index i appearing as a subscript in both factors, one simply replaces the Kronecker delta by unity and changes the subscript i in $\mathbb{T}$ to j:

$$\mathbb{T}_{pq\ldots i\ldots z}\, \delta_{ij} = \mathbb{T}_{pq\ldots j\ldots z}. \tag{1.37}$$

Levi-Civita symbol (or alternating symbol or permutation symbol): Consider the following determinant:

$$\det[M] = \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = a_1(b_2c_3 - c_2b_3) - b_1(a_2c_3 - c_2a_3) + c_1(a_2b_3 - b_2a_3)$$

$$= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1.$$

Observe in the final expression that (a) there are no terms in which the subscript $1, 2$ or 3 appears more than once (i.e. the subscripts i, j, k in $a_ib_jc_k$ are distinct); (b) when the subscripts are in cyclic order (i.e. 123, 231, 312) the coefficient is $+1$; and (c) when the subscripts are in anti-cyclic order (i.e. 132, 213, 321) the coefficient is -1. We will encounter this pattern in other settings as well. It is useful therefore to introduce a mathematical entity that captures this. This is the role of the so-called Levi-Civita symbol, also referred to as the alternating or permutation symbol. It is denoted by e_{ijk} and defined as

$$\begin{aligned} e_{ijk} \quad &:= \quad \begin{cases} 0 & \text{if two or more subscripts i, j, k, are equal,} \\ +1 & \text{if the subscripts i, j, k, are in cyclic order,} \\ -1 & \text{if the subscripts i, j, k, are in anticyclic order,} \end{cases} \\ &= \quad \begin{cases} 0 & \text{if two or more subscripts i, j, k, are equal,} \\ +1 & \text{for } (i,j,k) = (1,2,3), (2,3,1), (3,1,2), \\ -1 & \text{for } (i,j,k) = (1,3,2), (2,1,3), (3,2,1). \end{cases} \end{aligned} \tag{1.38}$$

The expression above for the determinant of the matrix $[M]$ can now be written succinctly as

$$\det[M] = e_{ijk} a_i b_j c_k,$$

which can also be written as

$$\det[M] = e_{ijk} M_{1i} M_{2j} M_{3k} = e_{ijk} M_{i1} M_{j2} M_{k3}. \tag{1.39}$$

Other identities involving the determinant include

$$e_{pqr} \det [M] \;=\; e_{ijk} M_{ip} M_{jq} M_{kr}, \tag{1.40}$$

$$\det [M] = \frac{1}{6} e_{ijk} e_{pqr} M_{ip} M_{jq} M_{kr}. \tag{1.41}$$

Another setting in which we shall encounter the Levi-Civita symbol is when working with the vector (cross) product (page 22).

Two useful properties of the Levi-Civita symbol are: (a) the sign of e_{ijk} changes whenever any two adjacent subscripts are switched:

$$e_{ijk} = -e_{jik} = e_{jki} = -e_{kji} = \dots \;, \tag{1.42}$$

(i.e. it is skew-symmetric with respect to every pair of adjacent subscripts) and (b) the Levi-Civita symbol and Kronecker delta are related by

$$e_{pij} e_{pk\ell} = \delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}; \tag{1.43}$$

see Problem 1.3.

1.2.1 Worked examples.

Problem 1.2.1. The matrices $[C], [D]$ and $[E]$ are defined in terms of the two matrices $[A]$ and $[B]$ by

$$[C] = [A][B], \qquad [D] = [B][A], \qquad [E] = [A][B]^T. \tag{i}$$

Express the elements of $[C], [D]$ and $[E]$ in terms of the elements of $[A]$ and $[B]$.

Solution: By the rules of matrix multiplication, the element C_{ij} in the i^{th} row and j^{th} column of $[C]$ is obtained by multiplying the elements of the i^{th} row of $[A]$, pairwise, by the respective elements of the j^{th} column of $[B]$ and summing. So, C_{ij} is obtained by multiplying the elements A_{i1}, A_{i2}, A_{i3} by, respectively, B_{1j}, B_{2j}, B_{3j} and summing. Thus

$$C_{ij} = A_{ik} B_{kj}; \tag{ii}$$

note that i and j are both free indices here and so this represents $3^2 = 9$ scalar equations; moreover summation is carried out over the repeated index k. It follows immediately from (ii) that the equation $[D] = [B][A]$ leads to

$$D_{ij} = B_{ik} A_{kj} \qquad \text{or equivalently} \qquad D_{ij} = A_{kj} B_{ik}, \tag{iii}$$

where the second expression was obtained by simply changing the order in which the factors appear in the first expression (since, as noted previously, the order of the factors within a term is insignificant since these are scalar quantities.) In order to calculate E_{ij}, we first use (ii) to directly obtain $E_{ij} = A_{ik}B^T_{kj}$. However, by definition of transposition, the i, j-element of a matrix $[B]^T$ equals the j, i-element of the matrix $[B]$: $B^T_{ij} = B_{ji}$ and so we can write

$$E_{ij} = A_{ik}B_{jk}. \tag{iv}$$

All four expressions here involve the ik, kj or jk elements of $[A]$ and $[B]$. The precise locations of the subscripts vary and the meaning of the terms depend crucially on these locations. It is worth repeating that the location of the repeated subscript k tells us what term multiplies what term.

Problem 1.2.2. The matrices $[A]$ and $[B]$ are symmetric and skew-symmetric respectively. Show that

$$A_{ij}B_{ij} = 0. \tag{1.44}$$

Remark: This result will be useful in several later calculations.

Solution: We proceed as follows:

$$A_{ij}B_{ij} \overset{(*)}{=} -A_{ji}B_{ji} \overset{(**)}{=} -A_{ij}B_{ij} \qquad \Rightarrow \quad 2A_{ij}B_{ij} = 0 \qquad \Rightarrow \quad A_{ij}B_{ij} = 0 \qquad \square$$

where in step $(*)$ we used the symmetry and skew-symmetry of $[A]$ and $[B]$ while in step $(**)$ we simply changed the dummy subscripts $i \to j, j \to i$.

Problem 1.2.3. Show that

$$e_{ijk}A_{jk} = 0 \tag{1.45}$$

if and only if $[A]$ is a symmetric matrix.

Solution: First suppose that $[A]$ is symmetric. Then, since e_{ijk} is skew-symmetric in j, k while A_{jk} is symmetric in j, k the result follows from (1.44).

Conversely suppose that (1.45) holds. Multiplying it by e_{ipq} gives

$$0 = e_{ipq}e_{ijk}A_{jk} \overset{(1.43)}{=} (\delta_{pj}\delta_{qk} - \delta_{pk}\delta_{qj})A_{jk} = \delta_{pj}\delta_{qk}A_{jk} - \delta_{pk}\delta_{qj}A_{jk} \overset{(**)}{=} A_{pq} - A_{qp} \quad \Rightarrow A_{pq} = A_{qp} \qquad \square$$

and so $[A]$ is symmetric. In step $(**)$ we used the substitution rule (1.37).

Problem 1.2.4. Simplify the expressions

(a) $A_{ij}\delta_{ij}$,

(b) $A_{ij}\delta_{ip}\delta_{jq}$, and

(c) $\delta_{ij}\delta_{ij}$.

Solution:

(a) In the expression $A_{ij}\delta_{ij}$ we have two repeated subscripts and so we can apply the substitution rule to either. Consider for example the repeated subscript j. Then, according to the substitution rule we replace δ_{ij} by unity and change the repeated subscript j in the other factor (i.e. in A_{ij}) to i. This yields $A_{ij}\delta_{ij} = A_{ii}$.

(b) In simplifying $A_{ij}\delta_{ip}\delta_{jq}$ (for clarity) we proceed in two steps. Consider δ_{ip} and the repeated subscript i. Thus we replace δ_{ip} by unity and change the repeated subscript i in the other factor to p. This yields $A_{ij}\delta_{ip}\delta_{jq} = A_{pj}\delta_{jq}$. We can apply the substitution rule again, this time on the index j, which yields $A_{pj}\delta_{jq} = A_{pq}$. Combing the two steps yields $A_{ij}\delta_{ip}\delta_{jq} = A_{pq}$. We could of course have done this in one step.

(c) The expression $\delta_{ij}\delta_{ij}$ involves two repeated subscripts and so we can apply the substitution rule to either one. If we apply the substitution rule on the repeated subscript i we get $\delta_{ij}\delta_{ij} = \delta_{jj}$. Since $\delta_{jj} = 3$ by (1.34) we can simplify further to get $\delta_{ij}\delta_{ij} = 3$.

Problem 1.2.5. Show that (a) $e_{ijp}e_{ijq} = 2\delta_{pq}$ and (b) $e_{ijk}e_{ijk} = 6$.

Solution

(a) We proceed as follows:

$$e_{ijp}e_{ijq} \overset{(1.43)}{=} \delta_{jj}\delta_{pq} - \delta_{jq}\delta_{pj} \overset{(1.34)}{=} 3\delta_{pq} - \delta_{jq}\delta_{pj} \overset{(1.37)}{=} 3\delta_{pq} - \delta_{pq} = 2\delta_{pq}. \qquad \square \qquad (i)$$

(b) Set $p = q = k$ in (i):

$$e_{ijk}e_{ijk} = 2\delta_{kk} \overset{(1.34)}{=} 6. \qquad \square \qquad (ii)$$

Problem 1.2.6. A matrix $[A]$ has the property that the magnitude of the column matrix $[A]\{x\}$ equals the magnitude of the column matrix $\{x\}$ for all column matrices $\{x\}$. Show that $[A]$ is an orthogonal matrix.

(We can view the matrix $[A]$ as mapping the column matrix $\{x\}$ into the column matrix $[A]\{x\}$. The particular matrix $[A]$ in this problem preserves the magnitude of a column matrix under this mapping.)

Solution: Let $\{y\} = [A]\{x\}$. We are told that

$$\{y\}^T\{y\} = \{x\}^T\{x\} \qquad \Leftrightarrow \qquad y_iy_i = x_ix_i. \qquad (i)$$

Since $y_i = A_{ij}x_j$, $(i)_2$ can be written as

$$A_{ij}x_jA_{ik}x_k = x_ix_i \qquad \Leftrightarrow \qquad A_{ij}A_{ik}x_jx_k = x_ix_i. \qquad (ii)$$

We are told that $(i)_1$ holds for all $\{x\}$. Therefore $(ii)_2$ holds for all x_p and so we can differentiate $(ii)_2$ with respect to x_p. This yields

$$\frac{\partial}{\partial x_p}\big(A_{ij}A_{ik}x_jx_k\big) = \frac{\partial}{\partial x_p}\big(x_ix_i\big) \qquad \Rightarrow \qquad A_{ij}A_{ik}\frac{\partial x_j}{\partial x_p}x_k + A_{ij}A_{ik}x_j\frac{\partial x_k}{\partial x_p} = \frac{\partial x_i}{\partial x_p}x_i + x_i\frac{\partial x_i}{\partial x_p}. \qquad (iii)$$

Since x_1, x_2, x_3 are independent variables, it follows that $\partial x_1/\partial x_1 = 1, \partial x_1/\partial x_2 = 0$ etc., i.e. $\partial x_p/\partial x_q = 1$ if $p = q$ and $\partial x_p/\partial x_q = 0$ if $p \neq q$. Thus

$$\frac{\partial x_p}{\partial x_q} = \delta_{pq}. \tag{1.46}$$

Substituting (1.46) into (iii) gives

$$A_{ij}A_{ik}\delta_{jp}x_k + A_{ij}A_{ik}x_j\delta_{kp} = 2\delta_{ip}x_i,$$

which by the substitution rule reduces to

$$A_{ip}A_{ik}x_k + A_{ij}A_{ip}x_j = 2x_p \qquad \Leftrightarrow \qquad A_{ip}A_{ik}x_k + A_{ik}A_{ip}x_k = 2x_p, \qquad (iv)$$

where in getting to the second equation we changed the dummy subscript $j \to k$. Changing the order of the factors in the second term on the left-hand side and simplifying leads to:

$$(iv) \quad \Rightarrow \quad A_{ip}A_{ik}x_k + A_{ip}A_{ik}x_k = 2x_p \quad \Rightarrow \quad 2A_{ip}A_{ik}x_k = 2x_p \quad \Rightarrow \quad A_{ip}A_{ik}x_k = x_p. \qquad (v)$$

Since (v) holds for all x_q we can differentiate it with respect to x_q to get

$$A_{ip}A_{ik}\frac{\partial x_k}{\partial x_q} = \frac{\partial x_p}{\partial x_q} \quad \stackrel{(1.46)}{\Rightarrow} \quad A_{ip}A_{ik}\delta_{kq} = \delta_{pq} \quad \Rightarrow \quad A_{ip}A_{iq} = \delta_{pq}, \qquad (vi)$$

where we again used the substitution rule in getting to the last expression. In matrix form, $(vi)_3$ can be written as

$$[A]^T[A] = [I] \qquad \square$$

which shows that $[A]$ is orthogonal.

1.3 Vector algebra.

- Perhaps the most familiar example of a vector is the geometric entity: an "arrow" which has both length and direction. Figure 2.1 shows the position *vectors* of a particle, $\mathbf{x}$ in the undeformed configuration and $\mathbf{y}$ in the deformed configuration, the displacement *vector* of this particle being $\mathbf{u}$.

  Other commonly encountered vectors in mechanics include velocity, linear momentum, angular velocity, angular momentum, force, traction and torque.

– A collection of vectors (denoted by V), together with certain operations (pertaining to addition, multiplication by a scalar and the null vector), is called a *vector space*.

– The null vector $\mathbf{o}$ has the property that $\mathbf{x} + \mathbf{o} = \mathbf{x}$ for all vectors $\mathbf{x}$ in V.

– A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is said to be **linearly independent** if the only scalars $\alpha_1, \alpha_2, \alpha_3$ for which

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 = \mathbf{o}$$

are $\alpha_1 = \alpha_2 = \alpha_3 = 0$. This implies that no vector in this set can be expressed as a linear combination of the other two. If a vector space V contains 3 linearly independent vectors but does not contain more than 3 linearly independent vectors, we say the **dimension** of V is 3.

– The **scalar product** (or dot product or inner product) of two vectors $\mathbf{u}$ and $\mathbf{v}$ is a scalar that we denote by

$$\mathbf{u} \cdot \mathbf{v}.$$

The scalar product obeys certain rules analogous to those listed in Problem 1.56. A vector space endowed with a scalar product is called a **Euclidean vector space**. Unless stated otherwise, we shall always be concerned with 3-dimensional Euclidean vector spaces which we denote by V (also commonly denoted by $\mathbb{E}^3$).

– The *magnitude* or length of a vector $\mathbf{u}$ is defined as

$$|\mathbf{u}| \coloneqq \sqrt{\mathbf{u} \cdot \mathbf{u}}, \tag{1.47}$$

and the *distance* between two vectors $\mathbf{u}, \mathbf{v}$ is $|\mathbf{u} - \mathbf{v}|$.

– The only vector with zero length is the null vector:

$$|\mathbf{x}| = 0 \qquad \Leftrightarrow \qquad \mathbf{x} = \mathbf{o}, \tag{1.48}$$

where the double arrow (here and throughout these notes) is short-hand for “if and only if”.

– The *angle* θ between two vectors $\mathbf{u} \neq \mathbf{o}$ and $\mathbf{v} \neq \mathbf{o}$ is defined by

$$\cos\ \theta \coloneqq \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|}, \qquad 0 \le \theta \le \pi. \tag{1.49}$$

If $\mathbf{u} \cdot \mathbf{v} = 0$ then $\theta = \pi/2$ and we say the vectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* to each other.

Exercise: In order for the preceding definition of the angle θ to be meaningful, the right-hand side of (1.49) must lie in the interval $[-1, 1]$. Is this true?, i.e. is $-|\mathbf{u}||\mathbf{v}| \le \mathbf{u}\cdot\mathbf{v} \le |\mathbf{u}||\mathbf{v}|$? Why?

Two vectors $\mathbf{a} \neq \mathbf{o}$ and $\mathbf{b} \neq \mathbf{o}$ are said to be parallel if there is a scalar α for which $\mathbf{a} = \alpha\mathbf{b}$.

– The **vector product** (or cross product) of two linearly independent vectors $\mathbf{u}$ and $\mathbf{v}$ is a vector that we denote by

$$\mathbf{u} \times \mathbf{v}.$$

Its magnitude is $|\mathbf{u}|\,|\mathbf{v}|\,\sin\theta$, it is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$, and its sense is given by the right-hand rule. Thus

$$\mathbf{u} \times \mathbf{v} = \big(|\mathbf{u}|\,|\mathbf{v}|\,\sin\theta\big)\,\mathbf{n} \qquad \text{where} \quad \mathbf{n}\cdot\mathbf{u} = \mathbf{n}\cdot\mathbf{v} = 0, \quad \mathbf{n}\cdot(\mathbf{u}\times\mathbf{v}) > 0; \tag{1.50}$$

the unit vector $\mathbf{n}$ is the direction of $\mathbf{u}\times\mathbf{v}$; the inequality in (1.50) tells us that the triplet of vectors $\{\mathbf{u},\mathbf{v},\mathbf{n}\}$ is right-handed.

– The scalar and vector products have the respective properties

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}, \qquad \mathbf{u}\times\mathbf{v} = -\mathbf{v}\times\mathbf{u}. \tag{1.51}$$

– The following useful result is addressed in Problem 1.7: three vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ are linearly independent if and only if

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) \neq 0. \tag{1.52}$$

Geometrically, three vectors are linearly independent if they do not lie in the same plane, i.e. they are non-coplanar.

1.3.1 Worked examples.

Problem 1.3.1. Calculate the area of the triangle OAB defined by the vectors $\overrightarrow{OA} = \mathbf{a}$ and $\overrightarrow{OB} = \mathbf{b}$ shown in Figure 1.1.

Solution: The angle $\theta = \angle AOB$ can be calculated from $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$. Suppose $\theta \in (0, \pi/2)$. By geometry, the height h of the triangle is $h = |\mathbf{b}|\sin\theta$. Thus

$$\text{Area of OAB} = \frac{1}{2}|OA|\,h = \frac{1}{2}|\mathbf{a}|\,h = \frac{1}{2}|\mathbf{a}|\,|\mathbf{b}|\sin\theta \overset{(1.50)}{=} \frac{1}{2}|\mathbf{a}\times\mathbf{b}|.$$

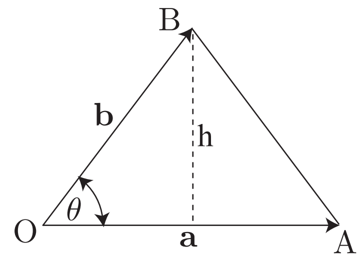


Figure 1.1: Parallelogram OACB.

Problem 1.3.2.

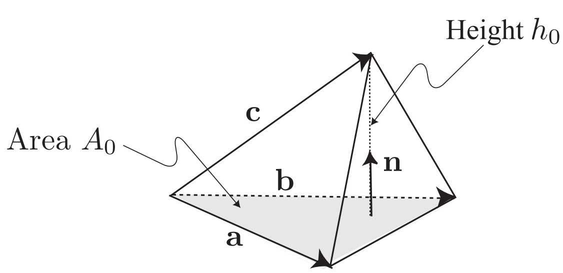
Show that two vectors $\mathbf{u}$ and $\mathbf{v}$ are equal if and only if

$$\mathbf{u}\cdot\mathbf{x} = \mathbf{v}\cdot\mathbf{x} \qquad \text{for all vectors } \mathbf{x}\in\mathsf{V}. \tag{i}$$

Solution: If $\mathbf{u}=\mathbf{v}$ it is clear that (i) holds. It is the converse that needs to be established. Thus suppose that (i) holds, which we can write as $(\mathbf{u}-\mathbf{v})\cdot\mathbf{x}=0$. Since this holds for all vectors $\mathbf{x}\in\mathsf{V}$ it necessarily holds for the particular choice $\mathbf{x}=\mathbf{u}-\mathbf{v}$. This gives $(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})=0$ whence $|\mathbf{u}-\mathbf{v}|=0$. However by (1.48), the only vector with zero length is the null vector and so $\mathbf{u}-\mathbf{v}=\mathbf{o}$. □

Problem 1.3.3. Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, be three linearly independent (i.e. non-coplanar) vectors. Show that the volume V_0 of the tetrahedron formed by them, see Figure 1.2, is

$$V_0 = \frac{1}{6}(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}.$$



$$\text{Volume } V_0 = \tfrac{1}{3}A_0 h_0$$
$$A_0 = \tfrac{1}{2}|\mathbf{a}\times\mathbf{b}|$$
$$h_0 = \mathbf{c}\cdot\mathbf{n}$$
$$\mathbf{n} = \frac{\mathbf{a}\times\mathbf{b}}{|\mathbf{a}\times\mathbf{b}|}$$

Figure 1.2: Volume of the tetrahedron defined by vectors $\mathbf{a},\mathbf{b},\mathbf{c}$.

Solution: Using the symbols introduced in Figure 1.2 we have

$$\mathbf{a}\times\mathbf{b} = |\mathbf{a}\times\mathbf{b}|\,\mathbf{n} = 2A_0\mathbf{n} \quad\Rightarrow\quad (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} = 2A_0(\mathbf{n}\cdot\mathbf{c}) = 2A_0h_0 = 6V_0 \quad\Rightarrow\quad V_0 = \frac{1}{6}(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}.$$

1.3.2 Components of a vector in a basis.

A brief video on the use of indicial notation in vector algebra can be found here.

Basis.

– A set of three linearly independent vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ forms a **basis** for a three-dimensional vector space V. An arbitrary vector $\mathbf{v} \in \mathsf{V}$ can be expressed as a unique linear combination of the three basis vectors.

– The basis is said to be **orthonormal** if each basis vector has unit length and each is orthogonal to the other two:

$$\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_3 \cdot \mathbf{e}_3 = 1, \qquad \mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_2 \cdot \mathbf{e}_3 = \mathbf{e}_3 \cdot \mathbf{e}_1 = 0.$$

This can be written succinctly as

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}, \tag{1.53}$$

where δ_{ij} is the Kronecker delta introduced in (1.33). We shall always restrict attention to orthonormal bases unless explicitly stated otherwise, and so we will drop the adjective "orthonormal" (except when we wish to emphasize it).

The basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is *right-handed* if

$$\mathbf{e}_1 \times \mathbf{e}_2 = +\mathbf{e}_3 \qquad \Leftrightarrow \qquad (\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 = +1.$$

For a right-handed basis it can be readily verified that $\mathbf{e}_1 \times \mathbf{e}_1 = 0$, $\mathbf{e}_1 \times \mathbf{e}_2 = +\mathbf{e}_3$, $\mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2$ etc. and so we again encounter the numbers $0, +1$ and -1 (as we did when looking at the determinant of a matrix). This suggests that we may be able to write $\mathbf{e}_i \times \mathbf{e}_j$ succinctly in terms of the Levi-Civita symbol e_{ijk} introduced in (1.38). Indeed, for a right-handed basis (Exercise),

$$\mathbf{e}_i \times \mathbf{e}_j = e_{ijk}\, \mathbf{e}_k, \qquad\qquad e_{ijk} = (\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_k. \tag{1.54}$$

Components of a vector.

– Given a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for a three-dimensional vector space V, an arbitrary vector $\mathbf{v} \in \mathsf{V}$ can always be expressed as a unique linear combination of the three basis vectors:

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3 \;\;=\;\; v_i\mathbf{e}_i. \tag{1.55}$$

The scalars v_i are called the **components** of $\mathbf{v}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

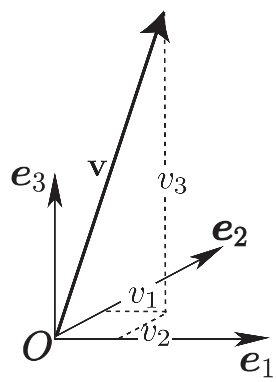


Figure 1.3: Components $\{v_1, v_2, v_3\}$ of a vector $\mathbf{v}$ in an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

– When the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is orthonormal, the components v_1, v_2, v_3 of the vector $\mathbf{v}$ can be calculated from

$$v_i = \mathbf{v} \cdot \mathbf{e}_i. \tag{1.56}$$

This follows by taking the scalar product of (1.55) with $\mathbf{e}_j$. Also, see Figure 1.3.

– Observe from (1.56) that the components v_i of the vector depend on both the vector $\mathbf{v}$ *and* the choice of basis. If we change the basis the components will change even if we don't change the vector.

– The components v_1, v_2, v_3 of $\mathbf{v}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ may be assembled into a column matrix

$$\{v\} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}. \tag{1.57}$$

Thus *a vector, together with a basis, allows the vector to be represented as a column matrix.*

– Once a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is chosen and fixed, there is a unique column matrix $\{v\}$ associated any given vector $\mathbf{v}$ (defined through (1.56), (1.57)); *and conversely,* there is a unique vector $\mathbf{v}$ associated with any given column matrix $\{v\}$ (defined by (1.55), (1.57)) where the components of $\mathbf{v}$ in $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are $\{v\}$. *Thus, once the basis is fixed, there is a one-to-one correspondence between column matrices and vectors.*

For example consider the *vector equation*

$$\mathbf{z} = \mathbf{x} + \mathbf{y}.$$

Taking the scalar product of this equation with each basis vector $\mathbf{e}_i$ gives $\mathbf{z} \cdot \mathbf{e}_i = \mathbf{x} \cdot \mathbf{e}_i + \mathbf{y} \cdot \mathbf{e}_i$ so that by (1.56), we obtain the *system of scalar equations*

$$z_i = x_i + y_i \qquad \text{for each } i = 1, 2, 3,$$

where x_i, y_i and z_i are the components of these vectors in the basis at hand. These components can be assembled into column matrices which allows us to express the preceding equation in *matrix form* as

$$\{z\} = \{x\} + \{y\}.$$

Thus, after choosing the basis, we can express the equation $\mathbf{z} = \mathbf{x} + \mathbf{y}$ in three equivalently forms – vector, components and matrix.

– Keep in mind that the *fundamental notion of a vector stands on its own without the need to refer to its components in a basis.* For example we can speak of the displacement of a particle or the force acting on a particle without needing to say anything about a basis.

– It follows from (1.55) and (1.53) that the *scalar product* of two vectors $\mathbf{u} = u_i\mathbf{e}_i$ and $\mathbf{v} = v_i\mathbf{e}_i$ can be calculated as $\mathbf{u} \cdot \mathbf{v} = (u_i\mathbf{e}_i) \cdot (v_j\mathbf{e}_j) = u_i v_j(\mathbf{e}_i \cdot \mathbf{e}_j) = u_i v_j \delta_{ij} = u_i v_i$ and therefore expressed as

$$\mathbf{u} \cdot \mathbf{v} = u_i v_i = u_1 v_1 + u_2 v_2 + u_3 v_3. \tag{1.58}$$

The *magnitude* of $\mathbf{u}$ can be written as

$$|\mathbf{u}| = (\mathbf{u} \cdot \mathbf{u})^{1/2} = (u_1^2 + u_2^2 + u_3^2)^{1/2} = (u_k u_k)^{1/2}. \tag{1.59}$$

Similarly, it follows from (1.55) and (1.54) that the *vector product* of $\mathbf{u} = u_i\mathbf{e}_i$ and $\mathbf{v} = v_i\mathbf{e}_i$ can be expressed as

$$\mathbf{u} \times \mathbf{v} = (u_j\mathbf{e}_j) \times (v_k\mathbf{e}_k) = u_j v_k\, \mathbf{e}_j \times \mathbf{e}_k \stackrel{(1.54)}{=} u_j v_k\, e_{jki}\mathbf{e}_i = e_{jki} u_j v_k\, \mathbf{e}_i = (e_{ijk} u_j v_k)\, \mathbf{e}_i, \tag{1.60}$$

where e_{ijk} is the Levi-Civita symbol. Equivalently, the ith component of the vector $\mathbf{u} \times \mathbf{v}$ is

$$(\mathbf{u} \times \mathbf{v})_i = e_{ijk} u_j v_k. \tag{1.61}$$

Exercise: Show that $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = e_{ijk} u_i v_j w_k$.

Exercise: Show that $\mathbf{u} \times \mathbf{u} = \mathbf{o}$.

– Finally, we illustrate through an example how one can go back and forth between vectors and their components. For any three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ we want to show that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}. \tag{i}$$

In order to prove this vector identity we shall (implicitly) pick and fix a basis for V and express all vectors in terms of their components in that basis. The left-hand side of (i) is a vector. We start from its ith component:

$$\begin{aligned}
\Big[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})\Big]_i &\overset{(1.61)}{=} e_{ijk}a_j(\mathbf{b} \times \mathbf{c})_k \overset{(1.61)}{=} e_{ijk}a_j(e_{kpq}b_pc_q) = \\
&\overset{(1.42)}{=} e_{kij}e_{kpq}a_jb_pc_q \overset{(1.43)}{=} (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})a_jb_pc_q = \\
&= \delta_{ip}\delta_{jq}a_jb_pc_q - \delta_{iq}\delta_{jp}a_jb_pc_q \overset{(*)}{=} a_qb_ic_q - a_jb_jc_i = \\
&\overset{(1.58)}{=} (\mathbf{a} \cdot \mathbf{c})b_i - (\mathbf{a} \cdot \mathbf{b})c_i
\end{aligned}$$

from which (i) follows. In step $(*)$ we used the substitution rule.

1.3.3 Worked examples.

Problem 1.3.4. For any four vectors $\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}$ show, by expressing the vectors in terms of their components in a basis, that

$$(\mathbf{p} \times \mathbf{q}) \cdot (\mathbf{r} \times \mathbf{s}) = (\mathbf{p} \cdot \mathbf{r})(\mathbf{q} \cdot \mathbf{s}) - (\mathbf{q} \cdot \mathbf{r})(\mathbf{p} \cdot \mathbf{s}). \tag{i}$$

Hence or otherwise, establish Lagrange's identity

$$|\mathbf{a} \times \mathbf{b}|^2 + (\mathbf{a} \cdot \mathbf{b})^2 = |\mathbf{a}|^2|\mathbf{b}|^2 \qquad \text{for all } \mathbf{a} \text{ and } \mathbf{b} \in \mathsf{V}. \tag{ii}$$

Solution: Both $\mathbf{p} \times \mathbf{q}$ and $\mathbf{r} \times \mathbf{s}$ are vectors and the left-hand side of (i) is the scalar product of these vectors. Thus

$$\begin{aligned}
(\mathbf{p} \times \mathbf{q}) \cdot (\mathbf{r} \times \mathbf{s}) &\overset{(1.58)}{=} (\mathbf{p} \times \mathbf{q})_i(\mathbf{r} \times \mathbf{s})_i \overset{(1.61)}{=} (e_{ijk}p_jq_k)(e_{imn}r_ms_n) = e_{ijk}e_{imn}p_jq_kr_ms_n = \\
&\overset{(1.43)}{=} (\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km})p_jq_kr_ms_n = \delta_{jm}\delta_{kn}p_jq_kr_ms_n - \delta_{jn}\delta_{km}p_jq_kr_ms_n = \\
&\overset{(*)}{=} p_mq_nr_ms_n - p_nq_mr_ms_n = (\mathbf{p} \cdot \mathbf{r})(\mathbf{q} \cdot \mathbf{s}) - (\mathbf{p} \cdot \mathbf{s})(\mathbf{q} \cdot \mathbf{r}), \qquad \square
\end{aligned}$$

where in step $(*)$ we used the substitution rule. Pick $\mathbf{p} = \mathbf{r} = \mathbf{a}$ and $\mathbf{q} = \mathbf{s} = \mathbf{b}$ in (i):

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) = (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b}) - (\mathbf{b} \cdot \mathbf{a})(\mathbf{a} \cdot \mathbf{b}),$$

which gives the desired result:

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2. \qquad \square$$

1.4 Tensor algebra.

– In this section we consider **linear transformations** from the vector space $\mathsf{V} \to \mathsf{V}$. Such a transformation $\mathbf{A}$ takes each vector $\mathbf{x} \in \mathsf{V}$ and maps it into another vector in V that we denote by $\mathbf{Ax}$ (where the mapping is subject to certain rules pertaining to addition, multiplication by a scalar and the null linear transformation).

– The most familiar examples of linear transformations are perhaps geometric, say rotation through an angle π about a certain axis which takes each vector into its rotated image, see Figure 1.4.

Figure 2.6 depicts a deformation that carries an infinitesimal material fiber $d\mathbf{x}$ in an undeformed body into its deformed image $d\mathbf{y} = \mathbf{F}d\mathbf{x}$, $\mathbf{F}$ being the deformation gradient tensor. Other examples from mechanics include the inertia "tensor" $\mathbf{J}$ that takes the angular velocity vector $\boldsymbol{\omega}$ into the angular momentum vector $\mathbf{h} = \mathbf{J}\boldsymbol{\omega}$; and the stress tensor $\mathbf{T}$ that takes a unit normal vector $\mathbf{n}$ into the traction vector $\mathbf{t} = \mathbf{Tn}$.

– Let $\mathbf{F}$ be a function (or transformation) that maps each vector $\mathbf{x} \in \mathsf{V}$ into a second vector $\mathbf{F}(\mathbf{x}) \in \mathsf{V}$:

$$\mathbf{x} \to \mathbf{F}(\mathbf{x}), \qquad \mathbf{x} \in \mathsf{V}, \quad \mathbf{F}(\mathbf{x}) \in \mathsf{V}. \tag{1.62}$$

It is said to be a *linear* transformation if

$$\mathbf{F}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{F}(\mathbf{x}) + \beta\mathbf{F}(\mathbf{y}) \tag{1.63}$$

for all scalars α, β and all vectors $\mathbf{x}, \mathbf{y}$ in V. When $\mathbf{F}$ is a linear transformation, we usually omit the parenthesis and write $\mathbf{Fx}$ instead of $\mathbf{F}(\mathbf{x})$. Note that $\mathbf{Fx}$ is a vector, and it is the **image** of $\mathbf{x}$ under the transformation $\mathbf{F}$.

– We shall refer to a linear transformation (from V into V) as a **tensor**.

– Keep in mind that a tensor is defined by the way in which it operates on each vector in V.

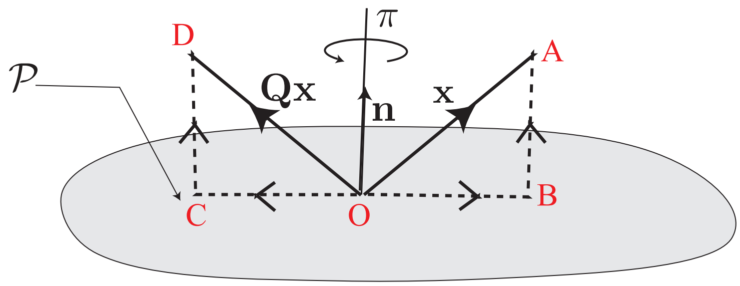


Figure 1.4: The tensor $\mathbf{Q}$ rotates the vector $\mathbf{x} = \overrightarrow{OA}$ through an angle π about the axis $\mathbf{n}$ and takes it to the vector $\mathbf{Qx} = \overrightarrow{OD}$. The plane $\mathcal{P}$ is perpendicular to $\mathbf{n}$

- **Example:** As a geometric example consider the transformation $\mathbf{Q}$ implied by Figure 1.4. It rotates a vector $\mathbf{x}\,(=\overrightarrow{OA})$ through an angle π about an axis $\mathbf{n}$ and carries it into $\mathbf{Qx}\,(=\overrightarrow{OD})$. (Here $\mathbf{x}$ is an arbitrary vector.) The plane $\mathcal{P}$ in the figure is perpendicular to the unit vector $\mathbf{n}$. Observe that $\overrightarrow{OB} = -\overrightarrow{OC}$ and $\overrightarrow{BA} = \overrightarrow{CD}$. Moreover, the magnitude of the vector $\overrightarrow{BA}$ is $\mathbf{x}\cdot\mathbf{n}$ and its direction is $\mathbf{n}$ and therefore $\overrightarrow{BA} = (\mathbf{x}\cdot\mathbf{n})\mathbf{n}$. Thus

$$\overrightarrow{OD} = \overrightarrow{OC} + \overrightarrow{CD} = -\overrightarrow{OB} + \overrightarrow{BA} = -(\overrightarrow{OA} - \overrightarrow{BA}) + \overrightarrow{BA} = -\overrightarrow{OA} + 2\,\overrightarrow{BA},$$

  and so we have

$$\mathbf{Qx} = -\mathbf{x} + 2(\mathbf{x}\cdot\mathbf{n})\mathbf{n} \qquad \text{for all } \mathbf{x} \in \mathsf{V}. \tag{1.64}$$

  Equation (1.64) tells us what $\mathbf{Q}$ does to every vector $\mathbf{x} \in \mathsf{V}$ and therefore it defines $\mathbf{Q}$. Observe that $\mathbf{Q}$ is a linear transformation since $\mathbf{Q}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{Qx} + \beta\mathbf{Qy}$.

- The *identity tensor* $\mathbf{I}$ and the *null tensor* $\mathbf{0}$ have the properties

$$\mathbf{Ix} = \mathbf{x}, \qquad \mathbf{0x} = \mathbf{o} \qquad \text{for all } \mathbf{x} \in \mathsf{V}, \tag{1.65}$$

  where $\mathbf{o}$ is the null vector.

- Continuing to keep in mind that a tensor is defined by the way it operates on vectors, given two tensors $\mathbf{A}$ and $\mathbf{B}$, their product $\mathbf{AB}$ is defined by

$$(\mathbf{AB})\mathbf{x} = \mathbf{A}(\mathbf{Bx}) \qquad \text{for all } \mathbf{x} \in \mathsf{V};$$

  i.e. first the tensor $\mathbf{B}$ operates on the vector $\mathbf{x}$ to produce the vector $\mathbf{Bx}$ and then the tensor $\mathbf{A}$ operates on the vector $\mathbf{Bx}$ to produce the vector $\mathbf{ABx}$. In general, $\mathbf{AB} \neq \mathbf{BA}$.

Tensor product.

– Let $\mathbf{a}$ and $\mathbf{b}$ be two given vectors. Define the associated tensor $\mathbf{T}$ as the transformation that takes an arbitrary vector $\mathbf{x}$ into the vector $\mathbf{Tx}$ defined by

$$\mathbf{Tx} = (\mathbf{x}\cdot\mathbf{b})\mathbf{a} \qquad \text{for all } \mathbf{x} \in \mathsf{V}. \tag{1.66}$$

Observe that corresponding to any vector $\mathbf{x} \in \mathsf{V}$, the right-hand side of (1.66) provides a formula for calculating the vector $\mathbf{Tx}$. Clearly this is a *linear* transformation. This particular tensor $\mathbf{T}$ is called the *tensor product*[4] of the vectors $\mathbf{a}$ and $\mathbf{b}$ and is denoted by

$$\mathbf{T} = \mathbf{a}\otimes\mathbf{b}. \tag{1.67}$$

Thus from (1.66) and (1.67)

$$(\mathbf{a}\otimes\mathbf{b})\mathbf{x} = (\mathbf{x}\cdot\mathbf{b})\mathbf{a} \qquad \text{for all } \mathbf{x} \in \mathsf{V}. \tag{1.68}$$

For example observe that we can now write (1.64) as

$$\mathbf{Qx} \;=\; -\mathbf{x} + 2(\mathbf{x}\cdot\mathbf{n})\mathbf{n} \;=\; -\mathbf{x} + 2(\mathbf{n}\otimes\mathbf{n})\mathbf{x} \;=\; \Big[-\mathbf{I} + 2(\mathbf{n}\otimes\mathbf{n})\Big]\mathbf{x}\,,$$

and so the tensor $\mathbf{Q}$ representing a rotation through an angle π about the axis $\mathbf{n}$ is

$$\mathbf{Q} \;=\; -\mathbf{I} + 2\mathbf{n}\otimes\mathbf{n}. \tag{1.69}$$

Example: Show that

$$(\mathbf{a}\otimes\mathbf{b})(\mathbf{c}\otimes\mathbf{d}) \;=\; (\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\otimes\mathbf{d}) \qquad \text{for all } \mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d} \in \mathsf{V}. \tag{1.70}$$

Solution: Before carrying out any calculation it is useful to observe that on the left-hand side we have a tensor (which is the product of the two tensors $\mathbf{a}\otimes\mathbf{b}$ and $\mathbf{c}\otimes\mathbf{d}$). On the right-hand side we also have a tensor, as we must, (and this is the multiple of the tensor $\mathbf{a}\otimes\mathbf{d}$ by the scalar $\mathbf{b}\cdot\mathbf{c}$). Also, note that if we want to show that two tensors $\mathbf{A}$ and $\mathbf{B}$ are equal, we could show that $\mathbf{Ax} = \mathbf{Bx}$ for all $\mathbf{x} \in \mathsf{V}$. With these in mind, for an arbitrary vector $\mathbf{x}$, we calculate

$$\begin{aligned}
(\mathbf{a}\otimes\mathbf{b})(\mathbf{c}\otimes\mathbf{d})\mathbf{x} \;&=\; (\mathbf{a}\otimes\mathbf{b})\Big((\mathbf{c}\otimes\mathbf{d})\mathbf{x}\Big) \overset{(1.68)}{=} (\mathbf{a}\otimes\mathbf{b})\Big((\mathbf{d}\cdot\mathbf{x})\mathbf{c}\Big) = \\
&\overset{(*)}{=}\; (\mathbf{d}\cdot\mathbf{x})(\mathbf{a}\otimes\mathbf{b})\mathbf{c} \overset{(1.68)}{=} (\mathbf{d}\cdot\mathbf{x})(\mathbf{b}\cdot\mathbf{c})\mathbf{a} \overset{(**)}{=} (\mathbf{b}\cdot\mathbf{c})(\mathbf{d}\cdot\mathbf{x})\mathbf{a} = \\
&=\; \overset{(1.68)}{=} (\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\otimes\mathbf{d})\mathbf{x}
\end{aligned}$$

where in step $(*)$ we used the fact that $\mathbf{d}\cdot\mathbf{x}$ is a scalar and so moved it to the front, and in step $(**)$ we used the fact that $\mathbf{b}\cdot\mathbf{c}$ is a scalar and so moved it to the front. This establishes (1.70).

[4]or outer product or dyadic product

– The set of all tensors (linear transformations from $\mathsf{V} \to \mathsf{V}$) is itself a vector space which we shall denote by Lin (“Lin” standing for linear transformation). It is 9-dimensional. The 9 tensors $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$, are linearly independent and thus form a basis for Lin[5]. Therefore given any tensor $\mathbf{A}$, there is a unique set of nine scalars A_{ij} such that

$$\mathbf{A} = A_{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j; \tag{1.71}$$

the A_{ij}'s are the components of $\mathbf{A}$ in this basis. The components A_{ij} of the tensor depend on the basis. If we change the basis the components will change even if we don't change the tensor.

Our analysis of tensor algebra in this subsection will *not* depend on the choice of basis and the components of the tensor in that basis. Even so, it will sometimes be useful to refer to tensor components, e.g. in (1.73), though we could have postponed all references to components to Section 1.4.3 where we shall say a lot more about them.

The transpose. Symmetric and skew-symmetric tensors.

– Corresponding to any tensor $\mathbf{A}$, there exists a second tensor that we denote by $\mathbf{A}^T$ such that

$$\mathbf{A}\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{A}^T\mathbf{y} \qquad \text{for all vectors } \mathbf{x} \text{ and } \mathbf{y} \in \mathsf{V}; \tag{1.72}$$

$\mathbf{A}^T$ is called the **transpose** of $\mathbf{A}$.

Exercise: Show that

$$\mathbf{A} = A_{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j \qquad \Leftrightarrow \qquad \mathbf{A}^T = A_{ij}\, \mathbf{e}_j \otimes \mathbf{e}_i = A_{ji}\, \mathbf{e}_i \otimes \mathbf{e}_j. \tag{1.73}$$

Exercise: Show that

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T\mathbf{A}^T \qquad \text{for all tensors } \mathbf{A}, \mathbf{B}. \tag{1.74}$$

Example: For all vectors $\mathbf{a}, \mathbf{b}$ and all tensors $\mathbf{A}$ show that

$$\mathbf{a} \otimes \mathbf{b} = (\mathbf{b} \otimes \mathbf{a})^T, \tag{1.75}$$

$$\mathbf{A}(\mathbf{a} \otimes \mathbf{b}) = (\mathbf{A}\mathbf{a}) \otimes \mathbf{b}, \qquad (\mathbf{a} \otimes \mathbf{b})\mathbf{A} = \mathbf{a} \otimes (\mathbf{A}^T\mathbf{b}). \tag{1.76}$$

Solution: In order to establish (1.75) we must show, according to (1.72), that $(\mathbf{a}\otimes\mathbf{b})\mathbf{x}\cdot\mathbf{y} = \mathbf{x}\cdot(\mathbf{b}\otimes\mathbf{a})\mathbf{y}$ for all vectors $\mathbf{x}$ and $\mathbf{y}$. This follows from:

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{x} \cdot \mathbf{y} \overset{(1.68)}{=} [(\mathbf{b} \cdot \mathbf{x})\mathbf{a}] \cdot \mathbf{y} = (\mathbf{b} \cdot \mathbf{x})(\mathbf{a} \cdot \mathbf{y}) = (\mathbf{x} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{y}) = \mathbf{x} \cdot [(\mathbf{a} \cdot \mathbf{y})\mathbf{b}] \overset{(1.68)}{=} \mathbf{x} \cdot (\mathbf{b} \otimes \mathbf{a})\mathbf{y}$$

[5] After introducing the notion of a scalar product between two tensors in (1.117), we can then speak of the magnitude of a tensor and of two tensors being orthogonal. This will allow us to show that these nine tensors in fact form an *orthonormal basis* for Lin; see Problem 1.4.3.

where we have used the fact that $\mathbf{a}\cdot\mathbf{x}$ and $\mathbf{b}\cdot\mathbf{y}$ are scalars.

To show $(1.76)_1$ we proceed as follows: for any vector $\mathbf{x}$,

$$\mathbf{A}(\mathbf{a}\otimes\mathbf{b})\mathbf{x} \overset{(1.68)}{=} \mathbf{A}[(\mathbf{b}\cdot\mathbf{x})\mathbf{a}] = (\mathbf{b}\cdot\mathbf{x})\mathbf{A}\mathbf{a} \overset{(1.68)}{=} (\mathbf{A}\mathbf{a}\otimes\mathbf{b})\mathbf{x},$$

which establishes $(1.76)_1$. The result in $(1.76)_2$ can be obtained similarly.

– A tensor $\mathbf{A}$ is **symmetric** if

$$\mathbf{A} = \mathbf{A}^T, \tag{1.77}$$

and **skew-symmetric** (or anti-symmetric) if

$$\mathbf{A} = -\mathbf{A}^T. \tag{1.78}$$

Several tensors that we will encounter including the Cauchy stress tensor $\mathbf{T}$ and the Lagrangian and Eulerian stretch tensors $\mathbf{U}$ and $\mathbf{V}$ will be symmetric.

Any tensor $\mathbf{A}$ can be uniquely decomposed into the sum of a symmetric tensor $\mathbf{S}$ and a skew-symmetric tensor $\mathbf{W}$:

$$\mathbf{A} = \mathbf{S} + \mathbf{W}, \qquad \mathbf{S} = \mathbf{S}^T, \qquad \mathbf{W} = -\mathbf{W}^T; \tag{1.79}$$

$\mathbf{S}$ is called the *symmetric part* of $\mathbf{A}$ and $\mathbf{W}$ its *skew-symmetric* part. They are given by

$$\mathbf{S} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T), \qquad \mathbf{W} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T). \tag{1.80}$$

Nonsingular tensors.

We are frequently interested in linear transformations that are **one-to-one**. For example suppose two particles of a body are located at $\mathbf{x}_1$ and $\mathbf{x}_2$, and that the tensor $\mathbf{F}$ maps ("deforms" the body and takes) them to the locations $\mathbf{F}\mathbf{x}_1$ and $\mathbf{F}\mathbf{x}_2$. We typically want particles not to coalesce: if the particles $\mathbf{x}_1$ and $\mathbf{x}_2$ are distinct we want their locations $\mathbf{F}\mathbf{x}_1$ and $\mathbf{F}\mathbf{x}_2$ to be distinct, i.e. we want $\mathbf{x}_1 \neq \mathbf{x}_2$ to imply $\mathbf{F}\mathbf{x}_1 \neq \mathbf{F}\mathbf{x}_2$. In addition we usually want particles not to split: if the locations $\mathbf{F}\mathbf{x}_1$ and $\mathbf{F}\mathbf{x}_2$ are distinct we want them to correspond to distinct particles, i.e. we want $\mathbf{F}\mathbf{x}_1 \neq \mathbf{F}\mathbf{x}_2$ to imply $\mathbf{x}_1 \neq \mathbf{x}_2$. Together, they require $\mathbf{F}\mathbf{x}_1 = \mathbf{F}\mathbf{x}_2 \Leftrightarrow \mathbf{x}_1 = \mathbf{x}_2$, i.e. the linear transformation $\mathbf{F}$ must be one-to-one.

The following statements are equivalent: A tensor $\mathbf{F}$ is **nonsingular** (or "one-to-one" or "invertible") if and only if

(a) For all vectors $\mathbf{x}_1, \mathbf{x}_2 \in \mathsf{V}$,

$$\mathbf{F}\mathbf{x}_1 = \mathbf{F}\mathbf{x}_2 \qquad \Leftrightarrow \qquad \mathbf{x}_1 = \mathbf{x}_2. \tag{1.81}$$

(b) The only vector for which $\mathbf{F}\mathbf{x} = \mathbf{o}$ is $\mathbf{x} = \mathbf{o}$:

$$\mathbf{F}\mathbf{x} = \mathbf{o} \qquad \Leftrightarrow \qquad \mathbf{x} = \mathbf{o}. \tag{1.82}$$

(c) The only vector for which $|\mathbf{F}\mathbf{x}| = 0$ is $\mathbf{x} = \mathbf{o}$:

$$|\mathbf{F}\mathbf{x}| = 0 \qquad \Leftrightarrow \qquad \mathbf{x} = \mathbf{o}.$$

(d) There exists a unique tensor that we denote by $\mathbf{F}^{-1}$, and call the **inverse** of $\mathbf{F}$, such that

$$\mathbf{F}\mathbf{F}^{-1} = \mathbf{F}^{-1}\mathbf{F} = \mathbf{I}. \tag{1.83}$$

(e) Also (1.91) below.

Exercise: Show that these statements are equivalent; (e.g. the example on page 32 shows that (1.91) implies (1.82).)

If $[F]$ is the matrix of components of $\mathbf{F}$ is some basis, then $[F]^{-1}$ is the matrix of components of $\mathbf{F}^{-1}$ in that basis.

If $\mathbf{A}$ and $\mathbf{B}$ are both nonsingular, then

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{1.84}$$

We shall denote the transpose of the inverse tensor, which equals the inverse of the transpose, by

$$\mathbf{A}^{-T} := (\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}. \tag{1.85}$$

– A tensor $\mathbf{A}$ is **positive definite** if

$$\mathbf{A}\mathbf{x} \cdot \mathbf{x} > 0 \qquad \text{for all vectors } \mathbf{x} \neq \mathbf{o}. \tag{1.86}$$

Exercise: Show that a positive definite tensor is necessarily nonsingular but a nonsingular tensor need not be positive definite.

Exercise: For any nonsingular tensor $\mathbf{F}$ show that the tensor $\mathbf{F}^T\mathbf{F}$ is symmetric and positive definite.

– The **determinant** of a tensor $\mathbf{A}$ is the scalar defined by

$$\det \mathbf{A} := \frac{\mathbf{A}\mathbf{x} \cdot (\mathbf{A}\mathbf{y} \times \mathbf{A}\mathbf{z})}{\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})} \qquad \text{for all linearly independent vectors } \mathbf{x}, \mathbf{y}, \mathbf{z}. \tag{1.87}$$

We know from Problem 1.3.3 that the volume of the tetrahedron defined by three linearly independent vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ is $\frac{1}{6}\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})$ and therefore the volume of its image under the linear transformation $\mathbf{A}$ is $\frac{1}{6}\mathbf{Ax} \cdot (\mathbf{Ay} \times \mathbf{Az})$. The determinant is therefore the ratio between these two volumes. This *ratio* is independent of the particular choice of $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$.

If $[A]$ is the matrix of components of $\mathbf{A}$ in some basis, one can show from (1.87), (1.71) and (1.39) that (Problem 1.4.6)

$$\det\,\mathbf{A} = \det[A]. \tag{1.88}$$

Observe that we did not use (1.88) as the definition of $\det \mathbf{A}$. This is because the components of a tensor, and therefore the matrix $[A]$, depends on the choice of basis. Therefore $\det[A]$ *may* depend on the choice of basis whereas $\det \mathbf{A}$ should not. See Problem 1.6.1 for further discussion.

The determinant of the product of two tensors has the property

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det \mathbf{A}\ \ \det \mathbf{B}, \tag{1.89}$$

and for any scalar α

$$\det(\alpha\mathbf{A}) = \alpha^3 \det \mathbf{A}. \tag{1.90}$$

– A tensor $\mathbf{A}$ is nonsingular, as defined previously in (1.82), if and only if

$$\det \mathbf{A} \neq 0. \tag{1.91}$$

Example: If $\det \mathbf{A} = 0$ show that there is a nonzero vector $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{o}$; see (1.91) and (1.82).
Solution: Suppose $\det \mathbf{A} = 0$. Then for any three linearly independent vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, equation (1.87) tells us that $\mathbf{Aa} \cdot (\mathbf{Ab} \times \mathbf{Ac}) = 0$. Therefore by (1.52) the vectors $\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}$ are linearly dependent and so there are scalars α, β, γ such that $\alpha\mathbf{Aa} + \beta\mathbf{Ab} + \gamma\mathbf{Ac} = \mathbf{o}$. Since this can be written as $\mathbf{A}(\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}) = \mathbf{o}$ we conclude that $\mathbf{Ax} = \mathbf{o}$ for $\mathbf{x} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}$.

Orthogonal tensors.

Orthogonal tensors will play an important role in both the rigid deformation of a body and the mapping between two different bases.

The following statements are equivalent: A tensor $\mathbf{Q}$ is **orthogonal** if and only if:

(a) It preserves the length of every vector:

$$|\mathbf{Qx}| = |\mathbf{x}| \qquad \text{for all } \mathbf{x} \in \mathsf{V}. \tag{1.92}$$

(b) It preserves the length of every vector and the angle between every pair of vectors:

$$\mathbf{Q}\mathbf{x} \cdot \mathbf{Q}\mathbf{y} = \mathbf{x} \cdot \mathbf{y} \qquad \text{for all } \mathbf{x} \text{ and } \mathbf{y} \in \mathsf{V}. \tag{1.93}$$

(c) It is nonsingular and

$$\mathbf{Q}^{-1} = \mathbf{Q}^T, \tag{1.94}$$

from which it follows that

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}. \tag{1.95}$$

Exercise: Show that these statements are equivalent. (e.g. Problem 1.4.8 shows that (1.92) implies (1.94).)

If $\mathbf{Q}$ is orthogonal, then

$$\det \mathbf{Q} = \pm 1. \tag{1.96}$$

(The converse is not true: $\det \mathbf{Q} = \pm 1$ does not imply that $\mathbf{Q}$ is orthogonal.) If $\det \mathbf{Q} = +1$, $\mathbf{Q}$ is said to be *proper orthogonal*. Otherwise it is *improper orthogonal*. If $\mathbf{Q}$ is improper orthogonal, then $-\mathbf{Q}$ is proper orthogonal and vice versa.

Exercise: Show that the tensor

$$-\mathbf{I} + 2\mathbf{n} \otimes \mathbf{n} \tag{1.97}$$

is proper orthogonal. (We saw earlier in (1.69) that it describes a 180^o rotation about the unit vector $\mathbf{n}$). Show that the tensor

$$\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n} \tag{1.98}$$

is improper orthogonal. (It describes a reflection in the plane perpendicular to $\mathbf{n}$; Problem 1.9.)

1.4.1 Worked examples.

Problem 1.4.1. Consider the tensor $\mathbf{R}$ that maps a vector $\mathbf{x}$ into the vector $\mathbf{R}\mathbf{x}$ according to

$$\mathbf{R}\mathbf{x} = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{n})\mathbf{n} \qquad \text{for all } \mathbf{x} \in \mathsf{V}. \tag{i}$$

Here $\mathbf{n}$ is a given unit vector. Show that

(a) $\mathbf{R}$ is nonsingular,

(b) $\mathbf{R}$ is symmetric,

(c) $\mathbf{R}$ is orthogonal,

(d) $\mathbf{R}$ is in fact improper orthogonal, and

(e) $\mathbf{R}^2 = \mathbf{I}$ and therefore that $\mathbf{R}$ is a square root of the identity tensor $\mathbf{I}$. Note that $\mathbf{R} \neq \mathbf{I}$.

It will be shown in Problem 1.9 that the tensor $\mathbf{R}$ describes reflections in the plane perpendicular to $\mathbf{n}$.

Solution: The solutions below do not use components in a basis. As an exercise, work this problem using components.

Solution 1: By the definition (1.68) of the tensor product we know that $(\mathbf{x} \cdot \mathbf{n})\mathbf{n} = (\mathbf{n} \otimes \mathbf{n})\mathbf{x}$. Therefore we can write (i) equivalently as

$$\mathbf{R}\mathbf{x} = (\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n})\mathbf{x}$$

which holds for all vectors $\mathbf{x}$. Therefore

$$\mathbf{R} = \mathbf{I} - 2\mathbf{n} \otimes \mathbf{n}. \tag{1.99}$$

(a) One way to show that $\mathbf{R}$ is nonsingular is to show that $\det \mathbf{R} \neq 0$. From (1.99) and (1.194),

$$\det \mathbf{R} = \det(\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n}) = 1 + (-2\mathbf{n}) \cdot \mathbf{n} = -1$$

and so $\det \mathbf{R} \neq 0$ and therefore $\mathbf{R}$ is nonsingular. (Alternatively see part (d) of solution-2 below.)

(b) By using (1.75) and (1.99),

$$\mathbf{R}^T = \mathbf{I} - 2\mathbf{n} \otimes \mathbf{n},$$

and so $\mathbf{R} = \mathbf{R}^T$ whence $\mathbf{R}$ is symmetric.

(c) We proceed as follows

$$\begin{aligned} \mathbf{R}\mathbf{R}^T &\overset{(1.99)}{=} (\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n})(\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n}) = \mathbf{I} - 2\mathbf{n} \otimes \mathbf{n} - 2\mathbf{n} \otimes \mathbf{n} + 4(\mathbf{n} \otimes \mathbf{n})(\mathbf{n} \otimes \mathbf{n}) = \\ &\overset{(1.70)}{=} \mathbf{I} - 4\mathbf{n} \otimes \mathbf{n} + 4(\mathbf{n} \cdot \mathbf{n})(\mathbf{n} \otimes \mathbf{n}) = \mathbf{I} - 4\mathbf{n} \otimes \mathbf{n} + 4(\mathbf{n} \otimes \mathbf{n}) = \mathbf{I}. \end{aligned}$$

Thus $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ and so $\mathbf{R}$ is orthogonal.

(d) Consider a right-handed orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ with $\mathbf{e}_3 = \mathbf{n}$. If $\{\mathbf{R}\mathbf{e}_1, \mathbf{R}\mathbf{e}_2, \mathbf{R}\mathbf{e}_3\}$ is left-handed then $\mathbf{R}$ is improper orthogonal. Since $\mathbf{e}_1 \cdot \mathbf{n} = \mathbf{e}_2 \cdot \mathbf{n} = 0$ and $\mathbf{n} \cdot \mathbf{n} = 1$ it follows that

$$\begin{aligned} \mathbf{R}\mathbf{e}_1 &\overset{(1.99)}{=} (\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n})\mathbf{e}_1 = \mathbf{e}_1 - 2(\mathbf{n} \cdot \mathbf{e}_1)\mathbf{n} = \mathbf{e}_1, \\ \mathbf{R}\mathbf{e}_2 &\overset{(1.99)}{=} (\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n})\mathbf{e}_2 = \mathbf{e}_2 - 2(\mathbf{n} \cdot \mathbf{e}_2)\mathbf{n} = \mathbf{e}_2, \\ \mathbf{R}\mathbf{e}_3 &\overset{(1.99)}{=} \mathbf{R}\mathbf{n} = (\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n})\mathbf{n} = \mathbf{n} - 2(\mathbf{n} \cdot \mathbf{n})\mathbf{n} = -\mathbf{n} = -\mathbf{e}_3. \end{aligned}$$

Therefore the orthogonal tensor $\mathbf{R}$ carries the right-hand triplet of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ into the left-handed triplet of vectors $\{\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_3\}$ and therefore is improper orthogonal.

(e) Since $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ and $\mathbf{R} = \mathbf{R}^T$ it follows immediately that $\mathbf{R}^2 = \mathbf{I}$. Therefore $\mathbf{R}$ is a square root of $\mathbf{I}$ (but it is not positive definite. Why?)

Solution 2: Here we will not make use of the representation (1.99).

(a) Another way in which to show that $\mathbf{R}$ is nonsingular is to show that the only vector $\mathbf{x}$ for which $\mathbf{Rx} = \mathbf{o}$ is the null vector $\mathbf{x} = \mathbf{o}$. From (i):

$$\mathbf{Rx} \cdot \mathbf{Rx} = \big[\mathbf{x} - 2(\mathbf{x}\cdot\mathbf{n})\mathbf{n}\big] \cdot \big[\mathbf{x} - 2(\mathbf{x}\cdot\mathbf{n})\mathbf{n}\big] = \mathbf{x}\cdot\mathbf{x} - 2(\mathbf{x}\cdot\mathbf{n})^2 - 2(\mathbf{x}\cdot\mathbf{n})^2 + 4(\mathbf{x}\cdot\mathbf{n})^2 = \mathbf{x}\cdot\mathbf{x}$$

Therefore

$$|\mathbf{Rx}| = |\mathbf{x}| \qquad \text{for all vectors } \mathbf{x} \in \mathsf{V}. \tag{ii}$$

Consequently $|\mathbf{Rx}| = 0$ if and only if $|\mathbf{x}| = 0$ which implies that $\mathbf{Rx} = \mathbf{0}$ if and only if $\mathbf{x} = \mathbf{0}$ (since the only vector whose length vanishes is the null vector). Therefore by (1.82), $\mathbf{R}$ is nonsingular.

(b) According to (1.72) and (1.77) the tensor $\mathbf{R}$ is symmetric if $\mathbf{Rx}\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{Ry}$ for all vectors $\mathbf{x}, \mathbf{y}$. From (i),

$$\mathbf{Rx} = \mathbf{x} - 2(\mathbf{x}\cdot\mathbf{n})\mathbf{n}, \qquad \mathbf{Ry} = \mathbf{y} - 2(\mathbf{y}\cdot\mathbf{n})\mathbf{n},$$

from which it follows that

$$\mathbf{Rx}\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{y} - 2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\mathbf{y}), \qquad \mathbf{x}\cdot\mathbf{Ry} = \mathbf{x}\cdot\mathbf{y} - 2(\mathbf{y}\cdot\mathbf{n})(\mathbf{n}\cdot\mathbf{x}).$$

Thus $\mathbf{Rx}\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{Ry}$ for all vectors $\mathbf{x}, \mathbf{y}$ and so $\mathbf{R}$ is symmetric:

$$\mathbf{R} = \mathbf{R}^T. \tag{iii}$$

(c) By (1.92), an orthogonal tensor is one that preserves the length of every vector, and so it follows immediately from (ii) that $\mathbf{R}$ is orthogonal.

(d) To show that $\mathbf{R}$ is improper orthogonal it is sufficient (since we know $\mathbf{R}$ is orthogonal) to show that $\det\mathbf{R} = -1$. Note from (i) that $\mathbf{Rn} = -\mathbf{n}$. Let $\mathbf{a}$ and $\mathbf{b}\,(\neq \mathbf{a})$ be two vectors orthogonal to $\mathbf{n}$. Since $\mathbf{a}\cdot\mathbf{n} = \mathbf{b}\cdot\mathbf{n} = 0$ it follows from (i) that $\mathbf{Ra} = \mathbf{a}$ and $\mathbf{Rb} = \mathbf{b}$. Therefore taking $\mathbf{x} = \mathbf{n}, \mathbf{y} = \mathbf{a}, \mathbf{z} = \mathbf{b}$ in the definition (1.87) of the determinant we get

$$\det\mathbf{R} := \frac{\mathbf{Rn}\cdot(\mathbf{Ra}\times\mathbf{Rb})}{\mathbf{n}\cdot(\mathbf{a}\times\mathbf{b})} = \frac{-\mathbf{n}\cdot(\mathbf{a}\times\mathbf{b})}{\mathbf{n}\cdot(\mathbf{a}\times\mathbf{b})} = -1.$$

Therefore $\mathbf{R}$ is improper orthogonal. (We could have used this to show that $\mathbf{R}$ is nonsingular in part (a).)

(e) Since $\mathbf{RR}^T = \mathbf{I}$ and $\mathbf{R} = \mathbf{R}^T$ it follows immediately that $\mathbf{R}^2 = \mathbf{I}$. Therefore $\mathbf{R}$ is a square root of $\mathbf{I}$.

Eigenvalues and eigenvectors.

– The **trace** of a tensor $\mathbf{A}$ is the scalar defined by

$$\operatorname{tr}\mathbf{A} := \frac{\mathbf{Ax}\cdot(\mathbf{y}\times\mathbf{z}) + \mathbf{x}\cdot(\mathbf{Ay}\times\mathbf{z}) + \mathbf{x}\cdot(\mathbf{y}\times\mathbf{Az})}{\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z})} \tag{1.100}$$

which is to hold for all linearly independent vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$. It is shown in Problem 1.55 that the right-hand side of (1.100) is in fact independent of the choice of the vectors

$\mathbf{x}, \mathbf{y}, \mathbf{z}$ and therefore depends only on the tensor $\mathbf{A}$, (the dependence being linear). If $[A]$ is the matrix of components of $\mathbf{A}$ in some basis, then

$$\operatorname{tr} \mathbf{A} = \operatorname{tr}[A] = A_{ii}. \tag{1.101}$$

Observe that we did not use (1.101) as the definition of tr$\mathbf{A}$. This is because the components of a tensor depend on the choice of basis and therefore tr$[A]$ *may* depend on the choice of basis whereas tr$\mathbf{A}$ should not. See Problem 1.6.1 for further discussion.

The trace of the product of two tensors has the property

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}). \tag{1.102}$$

Exercise: Show that

$$\operatorname{tr}(\mathbf{a} \otimes \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}. \tag{1.103}$$

– **Characteristic polynomial:** The characteristic polynomial associated with any tensor $\mathbf{A}$ and all scalars μ is (Problem 1.16)

$$\det(\mathbf{A} - \mu\mathbf{I}) \;=\; -\mu^3 + I_1(\mathbf{A})\mu^2 - I_2(\mathbf{A})\mu + I_3(\mathbf{A}), \tag{1.104}$$

where $I_1(\mathbf{A}), I_2(\mathbf{A})$ and $I_3(\mathbf{A})$ are the scalar-valued functions

$$I_1(\mathbf{A}) \coloneqq \operatorname{tr} \mathbf{A}, \qquad I_2(\mathbf{A}) \coloneqq \frac{1}{2}[(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr}(\mathbf{A}^2)], \qquad I_3(\mathbf{A}) \coloneqq \det \mathbf{A}. \tag{1.105}$$

These three functions are called the **principal scalar invariants** of $\mathbf{A}$ (for reasons that will be explained in Section 1.5).

According to the *Cayley-Hamilton theorem*

$$-\mathbf{A}^3 + I_1(\mathbf{A})\mathbf{A}^2 - I_2(\mathbf{A})\mathbf{A} + I_3(\mathbf{A})\mathbf{I} = \mathbf{0}. \tag{1.106}$$

– **Eigenvalues and eigenvectors:** A scalar α and vector $\mathbf{a}\,(\neq \mathbf{o})$ are said to be an *eigenvalue* and *eigenvector* of a tensor $\mathbf{A}$ if

$$\mathbf{Aa} = \alpha\mathbf{a}. \tag{1.107}$$

If $\mathbf{a}$ is an eigenvector of $\mathbf{A}$, so is any scalar multiple of $\mathbf{a}$ and so there is no loss of generality in assuming $\mathbf{a}$ to be a unit vector. The eigenvalues obey the *characteristic equation*

$$\det(\mathbf{A} - \alpha\mathbf{I}) = \; -\alpha^3 + I_1(\mathbf{A})\alpha^2 - I_2(\mathbf{A})\alpha + I_3(\mathbf{A}) = 0 \tag{1.108}$$

which is a cubic equation in α. It has either one or three real roots.

– **Eigenvalues and eigenvectors of a symmetric tensor:** A *symmetric* linear transformation $\mathbf{S}$ has three *real* eigenvalues $\sigma_1, \sigma_2, \sigma_3$. The corresponding eigenvectors $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ can be chosen so they are orthonormal. The eigenvectors are referred to as the *principal directions* of $\mathbf{S}$, and the particular basis $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\}$ is called a *principal basis* for $\mathbf{S}$.

– **Spectral representation of a symmetric tensor:** A *symmetric* linear transformation $\mathbf{S}$ can be expressed as

$$\mathbf{S} = \sigma_1\ (\mathbf{s}_1 \otimes \mathbf{s}_1) + \sigma_2\ (\mathbf{s}_2 \otimes \mathbf{s}_2) + \sigma_3\ (\mathbf{s}_3 \otimes \mathbf{s}_3) = \sum_{i=1}^{3} \sigma_i\ (\mathbf{s}_i \otimes \mathbf{s}_i). \tag{1.109}$$

This is called the *spectral representation* of the symmetric tensor $\mathbf{S}$.

Remark: Observe that the subscript i occurs three times in the rightmost expression in (1.109) and so we have suspended the usual summation convention and explicitly displayed the summation on i; see the discussion below (1.31).

For any positive integer n,

$$\mathbf{S}^n = \sum_{i=1}^{3} \sigma_i^n\ (\mathbf{s}_i \otimes \mathbf{s}_i). \tag{1.110}$$

If $\mathbf{S}$ is symmetric and nonsingular, then none of its eigenvalues vanish and

$$\mathbf{S}^{-1} = \sum_{i=1}^{3} (1/\sigma_i)\ (\mathbf{s}_i \otimes \mathbf{s}_i). \tag{1.111}$$

If $\mathbf{S}$ is symmetric and positive definite, all three eigenvalues are positive, and there is a unique symmetric positive definite tensor $\mathbf{T}$ such that $\mathbf{T}^2 = \mathbf{S}$. The tensor $\mathbf{T}$ is called the *positive definite square root* of $\mathbf{S}$ and denoted by $\mathbf{T} = \sqrt{\mathbf{S}}$: (Problem 1.23)

$$\sqrt{\mathbf{S}} = \sum_{i=i}^{3} \sqrt{\sigma_i}\ (\mathbf{s}_i \otimes \mathbf{s}_i). \tag{1.112}$$

Exercise: In Problem 1.4.1 we showed that both $\mathbf{I} - 2\mathbf{n} \otimes \mathbf{n}$ and $\mathbf{I}$ are square roots of the identity tensor $\mathbf{I}$. Does this contradict the claim of uniqueness associated with (1.112)?

Exercise: Show that the principal scalar invariants of $\mathbf{S}$ can be written as

$$I_1(\mathbf{S}) = \sigma_1 + \sigma_2 + \sigma_3, \qquad I_2(\mathbf{S}) = \sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1, \qquad I_3(\mathbf{S}) = \sigma_1\sigma_2\sigma_3. \tag{1.113}$$

Exercise: If two symmetric tensors $\mathbf{B}$ and $\mathbf{C}$ have the same eigenvalues, we see from (1.113) that they have the same principal scalar invariants. Conversely, if they have the same principal invariants, do they have the same eigenvalues?

Polar decomposition theorem.

A certain nonsingular tensor $\mathbf{F}$ called the deformation gradient tensor will play a pivotal role in describing the deformation of a body. Part of $\mathbf{F}$ will describe a rotation, the rest a stretch/strain. The polar decomposition theorem tells us how to identify these two parts of $\mathbf{F}$.

– The **polar decomposition theorem** states that, corresponding to any nonsingular tensor $\mathbf{F}$, there exist unique symmetric positive definite tensors $\mathbf{U}$ and $\mathbf{V}$ and a unique orthogonal tensor $\mathbf{R}$ such that[6]

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}. \tag{1.114}$$

The eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of $\mathbf{U}$ and $\mathbf{V}$ coincide. Let the corresponding orthonormal eigenvectors of $\mathbf{U}$ and $\mathbf{V}$ be $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Then

$$\mathbf{U} = \sum_{i=1}^{3} \lambda_i \mathbf{u}_i \otimes \mathbf{u}_i, \qquad \mathbf{V} = \sum_{i=1}^{3} \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i, \qquad \mathbf{R} = \sum_{i=1}^{3} \mathbf{v}_i \otimes \mathbf{u}_i. \tag{1.115}$$

The eigenvectors are related by $\mathbf{v}_i = \mathbf{R}\mathbf{u}_i$. Observe from (1.114) and (1.115), together with (1.70), that $\mathbf{F}$ has the representation

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \mathbf{v}_i \otimes \mathbf{u}_i. \tag{1.116}$$

Observe that both bases $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ appear in the right-hand side of (1.116) and so the scalar coefficients of $\mathbf{v}_i \otimes \mathbf{u}_i$ are not the components of $\mathbf{F}$ in either basis (except in the special case where these bases are identical).

Scalar product of two tensors.

As noted previously, the set of all linear transformations on V is a nine-dimensional vector space we denote by Lin. One can define a scalar product on the vector space Lin by

$$\mathbf{A} \cdot \mathbf{B} := \operatorname{tr}(\mathbf{A}\mathbf{B}^T) \qquad \text{for all tensors } \mathbf{A}, \mathbf{B} \in \mathsf{Lin}; \tag{1.117}$$

[6]Given an arbitrary (possibly singular) tensor $\mathbf{F} \in \mathsf{Lin}$, there is a (unique) symmetric, positive semi-definite tensor $\mathbf{U}$, and a (not-necessarily unique) orthogonal tensor $\mathbf{R}$, such that $\mathbf{F} = \mathbf{R}\mathbf{U}$; see Halmos [5], Section 83.

see Problem 1.56 for a justification of this definition and for various properties of the scalar product.

Notation: Since we use lower case boldface letters to denote vectors and upper case boldface letters to denote tensors, it will usually be clear as to whether a dot between two symbols refers to the scalar product between vectors or between tensors. Some authors denote the scalar product between tensors by a colon as in $\mathbf{A} : \mathbf{B}$.

The *magnitude* (norm) of a tensor $\mathbf{A}$ is denoted by $|\mathbf{A}|$ and defined by

$$|\mathbf{A}| = (\mathbf{A} \cdot \mathbf{A})^{1/2} = \left[\text{tr}\,(\mathbf{A}\mathbf{A}^T)\right]^{1/2}. \tag{1.118}$$

According to item (d) of Problem 1.56,

$$|\mathbf{A}| = 0 \qquad \text{if and only if} \quad \mathbf{A} = \mathbf{0}; \tag{1.119}$$

see also (1.132). If $\mathbf{A} \cdot \mathbf{B} = 0$ for two non-null tensors $\mathbf{A}$ and $\mathbf{B}$, we say that $\mathbf{A}$ is orthogonal to $\mathbf{B}$; we shall make use of this in Problem 1.4.3.

A useful identity for all tensors $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathsf{Lin}$ is (Problem 1.13)

$$\mathbf{AB} \cdot \mathbf{C} = \mathbf{B} \cdot \mathbf{A}^T\mathbf{C} = \mathbf{A} \cdot \mathbf{CB}^T. \tag{1.120}$$

Also, for any symmetric tensor $\mathbf{S}$ and skew-symmetric tensor $\mathbf{W}$ (Problem 1.4.7)

$$\mathbf{S} \cdot \mathbf{W} = 0. \tag{1.121}$$

Exercise: For any tensor $\mathbf{A}$ and vectors $\mathbf{x}$ and $\mathbf{y}$ show that

$$\mathbf{A} \cdot (\mathbf{x} \otimes \mathbf{y}) = \mathbf{A}\mathbf{y} \cdot \mathbf{x}, \tag{1.122}$$

where the scalar product on the left is between two tensors while that on the right is between two vectors.

1.4.2 Worked examples.

Problem 1.4.2.

(a) For any skew-symmetric tensor $\mathbf{W}$, show that

$$\mathbf{W}\mathbf{x} \cdot \mathbf{x} = 0 \qquad \text{for all vectors } \mathbf{x} \in \mathsf{V}. \tag{ii}$$

(b) If $\mathbf{S}$ is a symmetric tensor and

$$\mathbf{S}\mathbf{x}\cdot\mathbf{x} = 0 \qquad \text{for all vectors } \mathbf{x}\in\mathsf{V}, \tag{iii}$$

show that $\mathbf{S} = \mathbf{0}$.

Note: In contrast, if $\mathbf{A}$ is an *arbitrary* tensor and $\mathbf{A}\mathbf{x}\cdot\mathbf{x} = 0$ for all vectors $\mathbf{x}$, this does *not* imply that $\mathbf{A} = \mathbf{0}$; only that $\mathbf{A}$ is skew-symmetric.

(c) If the tensor $\mathbf{A}$ obeys

$$\mathbf{A}\mathbf{x}\cdot\mathbf{y} = 0 \qquad \text{for all vectors } \mathbf{x},\mathbf{y}\in\mathsf{V}, \tag{iv}$$

show that $\mathbf{A} = \mathbf{0}$. Note as an immediate consequence that if

$$\mathbf{A}\mathbf{x}\cdot\mathbf{y} = \mathbf{B}\mathbf{x}\cdot\mathbf{y} \qquad \text{for all vectors } \mathbf{x},\mathbf{y}\in\mathsf{V}, \tag{v}$$

then $\mathbf{A} = \mathbf{B}$.

(d) If

$$\mathbf{A}\cdot\mathbf{B} = 0 \qquad \text{for all tensors } \mathbf{B}\in\mathsf{Lin} \tag{vi}$$

show that $\mathbf{A} = \mathbf{0}$.

(e) If

$$\mathbf{A}\cdot\mathbf{B} = 0 \qquad \text{for all \textit{symmetric} tensors } \mathbf{B}\in\mathsf{Lin}, \tag{vii}$$

show that $\mathbf{A}$ must be skew-symmetric. It is important to note that (vii) does *not* imply $\mathbf{A} = \mathbf{0}$.

Solution:

(a) The result follows from the following calculation:

$$\mathbf{W}\mathbf{x}\cdot\mathbf{x} \overset{(1.72)}{=} \mathbf{x}\cdot\mathbf{W}^T\mathbf{x} \overset{(1.78)}{=} -\mathbf{x}\cdot\mathbf{W}\mathbf{x} \overset{(1.51)_1}{=} -\mathbf{W}\mathbf{x}\cdot\mathbf{x}$$

whence (ii) follows.

(b) Since $\mathbf{S}$ is symmetric, it has three real eigenvalues $\sigma_1, \sigma_2, \sigma_3$ and corresponding orthonormal eigenvectors $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$. Since $\mathbf{S}\mathbf{x}\cdot\mathbf{x} = 0$ for all vectors $\mathbf{x}\in\mathsf{V}$ it must necessarily hold for the choice $\mathbf{x} = \mathbf{s}_1$:

$$\mathbf{S}\mathbf{s}_1\cdot\mathbf{s}_1 = 0 \qquad \Rightarrow \qquad \mathbf{S}\mathbf{s}_1\cdot\mathbf{s}_1 = (\sigma_1\mathbf{s}_1)\cdot\mathbf{s}_1 = \sigma_1(\mathbf{s}_1\cdot\mathbf{s}_1) = \sigma_1 = 0.$$

Thus the eigenvalue $\sigma_1 = 0$. Similarly the other eigenvalues also vanish. This implies that $\mathbf{S} = \mathbf{0}$.

(c) Since (iv) holds for all vectors $\mathbf{y}\in\mathsf{V}$ it necessarily holds for the vector $\mathbf{y} = \mathbf{A}\mathbf{x}$. So we have

$$\mathbf{A}\mathbf{x}\cdot\mathbf{A}\mathbf{x} = 0 \qquad \text{for all vectors } \mathbf{x}\in\mathsf{V}.$$

Thus $|\mathbf{A}\mathbf{x}| = 0$ and since the only vector with zero length is the null vector, $\mathbf{A}\mathbf{x} = \mathbf{o}$. Since this holds for all vectors $\mathbf{x}\in\mathsf{V}$, $\mathbf{A} = \mathbf{0}$ by the definition $(1.65)_2$ of the zero tensor.

(d) Since (vi) holds for all tensors $\mathbf{B}\in\mathsf{Lin}$ it necessarily holds for $\mathbf{B} = \mathbf{A}$ and so $\mathbf{A}\cdot\mathbf{A} = 0$. By (1.119), this implies $\mathbf{A} = \mathbf{0}$.

(e) Using the decomposition (1.79) we can write

$$\mathbf{A} = \mathbf{S} + \mathbf{W} \tag{viii}$$

where

$$\mathbf{S} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T), \qquad \mathbf{W} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T), \tag{ix}$$

and so

$$\mathbf{A} \cdot \mathbf{B} = (\mathbf{S} + \mathbf{W}) \cdot \mathbf{B} = \mathbf{S} \cdot \mathbf{B} + \mathbf{W} \cdot \mathbf{B} \overset{(1.121)}{=} \mathbf{S} \cdot \mathbf{B}.$$

Equation (vii) thus tells us that

$$\mathbf{S} \cdot \mathbf{B} = 0 \qquad \text{for all symmetric tensors } \mathbf{B}.$$

Since this is to hold for all symmetric $\mathbf{B}$, and since $\mathbf{S}$ is symmetric, it must necessarily hold for the particular choice $\mathbf{B} = \mathbf{S}$. Therefore

$$\mathbf{S} \cdot \mathbf{S} = 0.$$

By (1.119), this implies $\mathbf{S} = \mathbf{0}$ and so from $(ix)_1$

$$\mathbf{A} = -\mathbf{A}^T$$

which says that $\mathbf{A}$ must be skew-symmetric.

Problem 1.4.3. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis for a Euclidean vector space V and let Lin be the set of all tensors from $\mathsf{V} \to \mathsf{V}$. Show that the nine tensors $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$, are

(a) linearly independent, and

(b) orthonormal.

Therefore if we can show that Lin is 9-dimensional (this is shown in Problem 1.22), then the nine tensors $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$, form an orthonormal basis for Lin.

Solution:

(a) To show that the tensors $\mathbf{e}_i \otimes \mathbf{e}_j$ are linearly independent we must show that the only scalars α_{ij} for which

$$\alpha_{ij}\mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{0} \tag{i}$$

are $\alpha_{ij} = 0$. Suppose (i) holds for some scalars α_{ij}. Operating (i) on the vector $\mathbf{e}_k$ gives

$$\alpha_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)\mathbf{e}_k = \mathbf{o} \quad \Rightarrow \quad \alpha_{ij}(\mathbf{e}_j \cdot \mathbf{e}_k)\mathbf{e}_i = \mathbf{o} \quad \Rightarrow \quad \alpha_{ij}\delta_{jk}\mathbf{e}_i = \mathbf{o} \quad \Rightarrow \quad \alpha_{ik}\mathbf{e}_i = \mathbf{o}. \tag{ii}$$

Taking the scalar product of (ii) with $\mathbf{e}_\ell$ gives $\alpha_{ik}\mathbf{e}_i \cdot \mathbf{e}_\ell = \alpha_{ik}\delta_{i\ell} = \alpha_{\ell k} = 0$. Thus the tensors $\mathbf{e}_i \otimes \mathbf{e}_j$ are linearly independent.

(b) To show that two tensors $\mathbf{A}$ and $\mathbf{B}$ are orthogonal we must show that $\mathbf{A}\cdot\mathbf{B} = 0$ which by (1.117) requires $\mathrm{tr}(\mathbf{A}\mathbf{B}^T) = 0$. Consider the two tensors $\mathbf{e}_i \otimes \mathbf{e}_j$ and $\mathbf{e}_k \otimes \mathbf{e}_\ell$. Their scalar product is

$$(\mathbf{e}_i \otimes \mathbf{e}_j)\cdot(\mathbf{e}_k \otimes \mathbf{e}_\ell) = \mathrm{tr}\Big[(\mathbf{e}_i \otimes \mathbf{e}_j)(\mathbf{e}_k \otimes \mathbf{e}_\ell)^T\Big] \overset{(1.75)}{=} \mathrm{tr}\Big[(\mathbf{e}_i \otimes \mathbf{e}_j)(\mathbf{e}_\ell \otimes \mathbf{e}_k)\Big] \overset{(1.70)}{=} \mathrm{tr}\Big[(\mathbf{e}_j \cdot \mathbf{e}_\ell)(\mathbf{e}_i \otimes \mathbf{e}_k)\Big] =$$
$$= (\mathbf{e}_j \cdot \mathbf{e}_\ell)\,\mathrm{tr}\Big[(\mathbf{e}_i \otimes \mathbf{e}_k)\Big] \overset{(1.103)}{=} (\mathbf{e}_j \cdot \mathbf{e}_\ell)\,(\mathbf{e}_i \cdot \mathbf{e}_k) = \delta_{ik}\delta_{j\ell}$$

Thus

$$(\mathbf{e}_i \otimes \mathbf{e}_j)\cdot(\mathbf{e}_k \otimes \mathbf{e}_\ell) = \delta_{ik}\delta_{j\ell}. \tag{1.123}$$

Therefore if $(\mathbf{e}_i \otimes \mathbf{e}_j) \neq (\mathbf{e}_k \otimes \mathbf{e}_\ell)$, i.e. $i \neq k$ and $j \neq \ell$, we have $(\mathbf{e}_i \otimes \mathbf{e}_j)\cdot(\mathbf{e}_k \otimes \mathbf{e}_\ell) = 0$ and so each of these tensors is orthogonal to the others. On the other hand if $(\mathbf{e}_i \otimes \mathbf{e}_j) = (\mathbf{e}_k \otimes \mathbf{e}_\ell)$, i.e. $i = k$ and $j = \ell$, we have $(\mathbf{e}_i \otimes \mathbf{e}_j)\cdot(\mathbf{e}_k \otimes \mathbf{e}_\ell) = 1$ and so the magnitude of each tensor is unity. Thus the nine tensors $\mathbf{e}_i \otimes \mathbf{e}_j$ are orthonormal.

Problem 1.4.4. Two symmetric tensors $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues. This does not imply that $\mathbf{B} = \mathbf{A}$. However it does imply that there is an orthogonal tensor $\mathbf{Q}$ such that $\mathbf{B} = \mathbf{Q}\mathbf{A}\mathbf{Q}^T$. Prove this.

Solution: Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$, and let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ be the corresponding eigenvectors. Since $\mathbf{A}$ and $\mathbf{B}$ are symmetric, the eigenvectors can be chosen so each set is orthonormal:

$$\mathbf{a}_i \cdot \mathbf{a}_j = \delta_{ij}, \qquad \mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij}. \tag{i}$$

The tensors $\mathbf{A}$ and $\mathbf{B}$ can be represented in spectral form as

$$\mathbf{A} = \sum_{i=1}^{3} \lambda_i \mathbf{a}_i \otimes \mathbf{a}_i, \qquad \mathbf{B} = \sum_{i=1}^{3} \lambda_i \mathbf{b}_i \otimes \mathbf{b}_i. \tag{ii}$$

Whenever an eigenvalue appears in an equation in this problem we will suspend the summation convention (as we have in (ii)) and display the summation explicitly. (We do this because the summation is over an index that appears three times.)

Since the bases $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ are orthonormal, it should not be surprising if the tensor that maps one into the other is orthogonal and perhaps this is the tensor $\mathbf{Q}$ we are after. Define the tensor $\mathbf{Q}$ by

$$\mathbf{Q} = \mathbf{b}_k \otimes \mathbf{a}_k, \tag{iii}$$

and observe that

$$\mathbf{Q}\mathbf{a}_i = (\mathbf{b}_k \otimes \mathbf{a}_k)\mathbf{a}_i = (\mathbf{a}_k \cdot \mathbf{a}_i)\mathbf{b}_k = \delta_{ki}\mathbf{b}_k = \mathbf{b}_i. \tag{iv}$$

Therefore $\mathbf{Q}$ maps each $\mathbf{a}_i \to \mathbf{b}_i$ and thus the basis $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ into the basis $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$. That $\mathbf{Q}$ is orthogonal follows from

$$\mathbf{Q}^T\mathbf{Q} = (\mathbf{b}_i \otimes \mathbf{a}_i)^T(\mathbf{b}_j \otimes \mathbf{a}_j) \overset{(1.75)}{=} (\mathbf{a}_i \otimes \mathbf{b}_i)(\mathbf{b}_j \otimes \mathbf{a}_j) =$$
$$\overset{(1.70)}{=} (\mathbf{b}_i \cdot \mathbf{b}_j)(\mathbf{a}_i \otimes \mathbf{a}_j) \overset{(i)}{=} \delta_{ij}(\mathbf{a}_i \otimes \mathbf{a}_j) \overset{(*)}{=} \mathbf{a}_i \otimes \mathbf{a}_i \overset{(1.133)}{=} \mathbf{I},$$

where in step $(*)$ we used the substitution rule.

It can now be readily shown that $\mathbf{B} = \mathbf{Q}\mathbf{A}\mathbf{Q}^T$:

$$\mathbf{B} \overset{(ii)}{=} \sum_{i=1}^{3} \lambda_i \mathbf{b}_i \otimes \mathbf{b}_i \overset{(iv)}{=} \sum_{i=1}^{3} \lambda_i \mathbf{Q}\mathbf{a}_i \otimes \mathbf{Q}\mathbf{a}_i \overset{(1.76)}{=} \sum_{i=1}^{3} \lambda_i \mathbf{Q}(\mathbf{a}_i \otimes \mathbf{a}_i)\mathbf{Q}^T = \mathbf{Q}\left(\sum_{i=1}^{3} \lambda_i (\mathbf{a}_i \otimes \mathbf{a}_i)\right)\mathbf{Q}^T \overset{(ii)}{=} \mathbf{Q}\mathbf{A}\mathbf{Q}^T.$$

1.4.3 Components of a tensor in a basis.

A few brief videos on the use of indicial notation in tensor algebra can be found here.

– Let Lin be the set of all tensors from the vector space $\mathsf{V} \to \mathsf{V}$. Problem 1.22 shows that the dimension of Lin is 9, and Problem 1.4.3 showed that $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$, are nine orthonormal tensors in Lin (where as usual $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is an orthonormal basis for V). Therefore these 9 tensors form an orthonormal basis for Lin. Consequently, as noted previously, given any tensor $\mathbf{A}$, there is a unique set of nine scalars A_{ij} such that

$$\mathbf{A} = A_{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j. \tag{1.124}$$

The A_{ij}'s are the components of $\mathbf{A}$ in this basis.

– Since the basis of nine tensors is orthonormal, one can derive the following formula for the components A_{ij} (show this):

$$A_{ij} = (\mathbf{A}\mathbf{e}_j) \cdot \mathbf{e}_i. \tag{1.125}$$

This says that the ith component of the vector $\mathbf{A}\mathbf{e}_j$ is A_{ij} which can be equivalently stated as

$$\mathbf{A}\mathbf{e}_j = A_{ij}\mathbf{e}_i. \tag{1.126}$$

Remark: We know that a tensor is characterized by the way it transforms every vector in V. Since any vector can always be expressed in terms of the basis vectors, it follows that in order to define a tensor $\mathbf{A}$ it is sufficient to (only) know how $\mathbf{A}$ transforms the basis vectors, i.e. to know $\mathbf{A}\mathbf{e}_j$ for $j = 1, 2, 3$. According to (1.126) (and also (1.125)) the nine scalars A_{ij} do this.

– Exercise: If $\mathbf{a}$ and $\mathbf{b}$ have components a_i and b_i, show that the components of the tensor $\mathbf{a} \otimes \mathbf{b}$ are

$$(\mathbf{a} \otimes \mathbf{b})_{ij} = a_i b_j. \tag{1.127}$$

If $\mathbf{A}$ and $\mathbf{x}$ have components A_{ij} and x_i, show that the ith component of the vector $\mathbf{Ax}$ is

$$(\mathbf{Ax})_i = A_{ij}x_j. \tag{1.128}$$

If the tensors $\mathbf{A}$ and $\mathbf{B}$ have components A_{ij} and B_{ij} respectively, show that the i,j component of the tensor $\mathbf{AB}$ is

$$(\mathbf{AB})_{ij} = A_{ik}B_{kj}. \tag{1.129}$$

– Exercise: If $\mathbf{Q}$ is an orthogonal tensor, show that

$$Q_{ik}Q_{jk} = Q_{ki}Q_{kj} = \delta_{ij}. \tag{1.130}$$

– Exercise: If the tensors $\mathbf{A}$ and $\mathbf{B}$ have components A_{ij} and B_{ij} in some basis, show that

$$\mathbf{A}\cdot\mathbf{B} = A_{ij}B_{ij}, \tag{1.131}$$

$$|\mathbf{A}| = (\mathbf{A}\cdot\mathbf{A})^{1/2} = (A_{ij}A_{ij})^{1/2} \ = \left(A_{11}^2 + A_{12}^2 + A_{13}^2 + \ldots + A_{33}^3\right)^{1/2}. \tag{1.132}$$

It follows from (1.132) that $|\mathbf{A}| = 0$ if and only if *every* component $A_{ij} = 0$, i.e. if and only if $\mathbf{A} = \mathbf{0}$. Moreover, if $|\mathbf{A}| \to 0$ then each $A_{ij} \to 0$.

– The components of the identity tensor $\mathbf{I}$ in any basis are δ_{ij} and it can be represented as

$$\mathbf{I} = \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 = \mathbf{e}_i \otimes \mathbf{e}_i. \tag{1.133}$$

– The components A_{ij} of a tensor $\mathbf{A}$ in a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ can be assembled into a square matrix:

$$[A] = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}. \tag{1.134}$$

The components A_{ij} depend on *both* the tensor $\mathbf{A}$ *and* the choice of basis.

– Once a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is chosen and fixed, there is a unique matrix $[A]$ associated with any given tensor $\mathbf{A}$; *and conversely* there is a unique tensor $\mathbf{A}$ associated with any given square matrix $[A]$ such that the components of $\mathbf{A}$ in $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are $[A]$. *Thus, once the basis is fixed, there is a one-to-one correspondence between square matrices and tensors.* Just as for vectors, the fundamental notion of a tensor stands on its own, without the need to refer to its components in a basis. For example, the strain (tensor) at a particle does not depend on a basis.

– *Various algebraic operations on vectors and tensors correspond exactly to analogous matrix operations on the associated matrices of components* (once a basis has been chosen). As an example suppose that $\mathbf{y} = \mathbf{A}\mathbf{x}$. Then by (1.128),

$$\mathbf{y} = \mathbf{A}\mathbf{x} \qquad \Leftrightarrow \qquad \{y\} = [A]\{x\} \qquad \Leftrightarrow \qquad y_i = A_{ij}x_j, \tag{1.135}$$

where the column matrices $\{x\}$ and $\{y\}$ are the components of $\mathbf{x}$ and $\mathbf{y}$ respectively. Similarly if $\mathbf{C} = \mathbf{A}\mathbf{B}$, it follows from (1.129) that

$$\mathbf{C} = \mathbf{A}\mathbf{B} \qquad \Leftrightarrow \qquad [C] = [A][B] \qquad \Leftrightarrow \qquad C_{ij} = A_{ik}B_{kj}. \tag{1.136}$$

– Suppose $[A]$ and $[A']$ are the matrices of components of $\mathbf{A}$ in *any two* orthonormal bases. Then one can show that $\mathrm{tr}[A] = \mathrm{tr}[A']$ and $\det[A] = \det[A']$ (Problem 1.6.1). Consequently though the respective statements $\det \mathbf{A} = \det[A]$ and $\mathrm{tr}\mathbf{A} = \mathrm{tr}[A]$ in (1.88) and (1.101) involve components in a basis, they are in fact *independent of the choice of basis* and so could have been used to define the trace and determinant.

– Exercise: Work Problem 1.4.1 using components.

1.4.4 Worked examples.

Problem 1.4.5. It is shown in Problem 1.10 that the tensor $\mathbf{Q}$ describing a rotation through an angle θ about a unit vector $\mathbf{n}$ is defined by

$$\mathbf{Q}\mathbf{x} = \cos\theta\ \mathbf{x} + (1-\cos\theta)(\mathbf{n}\cdot\mathbf{x})\mathbf{n} + \sin\theta\ (\mathbf{n}\times\mathbf{x}) \qquad \text{for all } \mathbf{x} \in \mathsf{V}. \tag{i}$$

Calculate the components of $\mathbf{Q}$ in a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ in the special case where the axis of rotation $\mathbf{n}$ is $\mathbf{e}_3$.

Solution: We shall use (1.126) to calculate the components Q_{ij}. Taking $\mathbf{x} = \mathbf{e}_j$ and $\mathbf{n} = \mathbf{e}_3$ in (i) yields

$$\mathbf{Q}\mathbf{e}_j = \cos\theta\ \mathbf{e}_j + (1-\cos\theta)\delta_{3j}\mathbf{e}_3 + \sin\theta\ e_{3jk}\mathbf{e}_k. \tag{ii}$$

Therefore

$$\left.\begin{aligned}
\mathbf{Q}\mathbf{e}_1 &= \cos\theta\ \mathbf{e}_1 + \sin\theta\ e_{31k}\mathbf{e}_k = \cos\theta\ \mathbf{e}_1 + \sin\theta\ e_{312}\mathbf{e}_2 &&= \cos\theta\ \mathbf{e}_1 + \sin\theta\ \mathbf{e}_2,\\
\mathbf{Q}\mathbf{e}_2 &= \cos\theta\ \mathbf{e}_2 + \sin\theta\ e_{32k}\mathbf{e}_k = \cos\theta\ \mathbf{e}_2 + \sin\theta\ e_{321}\mathbf{e}_1 &&= \cos\theta\ \mathbf{e}_2 - \sin\theta\ \mathbf{e}_1,\\
\mathbf{Q}\mathbf{e}_3 &= \cos\theta\ \mathbf{e}_3 + (1-\cos\theta)\mathbf{e}_3 + \sin\theta\ e_{33k}\mathbf{e}_k &&= \mathbf{e}_3.
\end{aligned}\right\}$$

Therefore we can read off the components of $\mathbf{Q}$ from this using (1.126) to be

$$[Q] = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad \square$$

Problem 1.4.6. The determinant of a tensor $\mathbf{A}$ was defined in (1.87). Show that $\det \mathbf{A} = \det[A]$ where $[A]$ is the matrix of components of $\mathbf{A}$ in a basis.

Solution: Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be an arbitrary set of linearly independent vectors, and let a_i, b_i, c_i and A_{ij} be the components of these vectors and $\mathbf{A}$ in a basis. Then (motivated by the numerator of (1.87)) we evaluate the quantity $(\mathbf{Aa} \times \mathbf{Ab}) \cdot \mathbf{Ac}$:

$$(\mathbf{Aa} \times \mathbf{Ab}) \cdot \mathbf{Ac} \overset{(1.58)}{=} (\mathbf{Aa} \times \mathbf{Ab})_i \; (\mathbf{Ac})_i \overset{(1.61)}{=} e_{ijk} \; (\mathbf{Aa})_j \; (\mathbf{Ab})_k \; (\mathbf{Ac})_i =$$

$$\overset{(1.128)}{=} e_{ijk} \; (A_{jm} \, a_m) \, (A_{kn} \, b_n) \, (A_{is} \, c_s) = e_{ijk} \; A_{is} A_{jm} A_{kn} a_m \, b_n \, c_s.$$

The identity (1.40) for the determinant of a matrix, $e_{smn} \det[A] = e_{ijk} \, A_{is} A_{jm} A_{kn}$, allows us to write this as

$$\begin{aligned} (\mathbf{Aa} \times \mathbf{Ab}) \cdot \mathbf{Ac} &= e_{smn} \det[A] \, a_m \, b_n \, c_s = \det[A] \, e_{smn} \, a_m \, b_n \, c_s \\ &\overset{(1.61)}{=} \det[A] \; (\mathbf{a} \times \mathbf{b})_s \, c_s \overset{(1.58)}{=} \det[A] \; (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}. \end{aligned} \qquad (i)$$

Since $\mathbf{a}, \mathbf{b}, \mathbf{c}$, are linearly independent it follows from (1.52) that $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \neq 0$. Thus, comparing this with (1.87) shows that $\det \mathbf{A} = \det[A]$.

Aside: Observe from (i) that if $\mathbf{A}$ is nonsingular so that $\det \mathbf{A} \neq 0$, then $(\mathbf{Aa} \times \mathbf{Ab}) \cdot \mathbf{Ac} \neq 0$ and so by (1.52) the three vectors $\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}$ are also linearly independent.

Problem 1.4.7. For any symmetric tensor $\mathbf{S}$ and skew-symmetric tensor $\mathbf{W}$ show that

$$\mathbf{S} \cdot \mathbf{W} = 0. \qquad (1.137)$$

Solution: The result follows immediately by writing (1.137) in terms of components and then using the result in Problem 1.2.2.

Problem 1.4.8. An orthogonal tensor was defined in (1.92) as a tensor that preserves length, i.e. $\mathbf{Q}$ is orthogonal if

$$|\mathbf{Qx}| = |\mathbf{x}| \qquad \text{for all vectors } \mathbf{x} \in \mathsf{V}. \qquad (i)$$

Show that an orthogonal tensor is nonsingular and that

$$\mathbf{Q}^T = \mathbf{Q}^{-1}. \qquad (ii)$$

Solution 1: (Using components in a basis.) It follows from (i) that

$$\mathbf{Qx} \cdot \mathbf{Qx} = \mathbf{x} \cdot \mathbf{x} \qquad \Rightarrow \qquad Q_{ij}Q_{ik}x_jx_k = x_kx_k.$$

Since this holds for all x_1, x_2, x_3 we may differentiate it with respect to x_p to get

$$Q_{ij}Q_{ik}\frac{\partial}{\partial x_p}(x_jx_k) = \frac{\partial}{\partial x_p}(x_kx_k) \qquad \Rightarrow \qquad Q_{ij}Q_{ik}(\delta_{jp}x_k + x_j\delta_{pk}) = 2\delta_{pk}x_k,$$

which after using the substitution rule yields $Q_{ip}Q_{ik}x_k + Q_{ij}Q_{ip}x_j = 2x_p$. Changing the dummy subscript $j \to k$ in the second term now leads to

$$Q_{ip}Q_{ik}x_k = x_p.$$

Differentiating this with respect to x_q yields

$$Q_{ip}Q_{ik}\delta_{kq} = \delta_{pq} \qquad \Rightarrow \qquad Q_{ip}Q_{iq} = \delta_{pq} \qquad \Rightarrow \qquad Q^T_{pi}Q_{iq} = \delta_{pq} \qquad \Rightarrow \qquad \mathbf{Q}^T\mathbf{Q} = \mathbf{I}.$$

On taking the determinant of this equation we get $\det \mathbf{Q} = \pm 1$ and so $\det \mathbf{Q} \neq 0$ which implies that $\mathbf{Q}$ is nonsingular. Post-multiplying both sides of the preceding equation by $\mathbf{Q}^{-1}$ now yields (ii).

Solution 2: (Without using components in a basis.) To show that $\mathbf{Q}$ is nonsingular we must show that the only vector $\mathbf{x}$ for which $\mathbf{Qx} = \mathbf{o}$ is $\mathbf{x} = \mathbf{o}$. Suppose $\mathbf{Qx} = \mathbf{o}$. Then $|\mathbf{Qx}| = 0$ and so $|\mathbf{x}| = 0$ by (i). This implies that $\mathbf{x} = \mathbf{o}$ since the only vector with zero length is the null vector; see (1.48). Therefore the only vector $\mathbf{x}$ for which $\mathbf{Qx} = \mathbf{o}$ is the null vector and so by definition (1.82), $\mathbf{Q}$ is nonsingular.

Next we write $|\mathbf{Qx}|^2 = |\mathbf{x}|^2$ as $\mathbf{Qx} \cdot \mathbf{Qx} = \mathbf{x} \cdot \mathbf{x}$ which because of (1.72) implies $\mathbf{Q}^T\mathbf{Qx} \cdot \mathbf{x} = \mathbf{x} \cdot \mathbf{x}$. Thus

$$(\mathbf{Q}^T\mathbf{Q} - \mathbf{I})\mathbf{x} \cdot \mathbf{x} = 0 \qquad \text{for all vectors } \mathbf{x} \in \mathsf{V}. \tag{iii}$$

Recall from Problem 1.4.2(c) that if $\mathbf{Sx} \cdot \mathbf{x} = 0$ for all vectors $\mathbf{x} \in \mathsf{V}$ and $\mathbf{S}$ is a symmetric tensor then $\mathbf{S} = \mathbf{0}$. Since $\mathbf{Q}^T\mathbf{Q} - \mathbf{I}$ is symmetric, it now follows that

$$\mathbf{Q}^T\mathbf{Q} - \mathbf{I} = \mathbf{0}. \tag{iv}$$

Operating on both sides of (iii) with $\mathbf{Q}^{-1}$ (we know that $\mathbf{Q}$ is nonsingular so $\mathbf{Q}^{-1}$ exists) gives the desired result $\mathbf{Q}^{-1} = \mathbf{Q}^T$.

1.5 Invariance. Isotropic functions.

Consider two orthonormal bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ where the first is mapped into the second by the orthogonal tensor $\mathbf{Q}$:

$$\mathbf{e}'_i = \mathbf{Q}\mathbf{e}_i, \quad i = 1, 2, 3, \tag{1.138}$$

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}. \tag{1.139}$$

Vectors:

– For any vector $\mathbf{v}$,

$$\mathbf{Q}\mathbf{v} \cdot \mathbf{e}'_i \overset{(1.138)}{=} \mathbf{Q}\mathbf{v} \cdot \mathbf{Q}\mathbf{e}_i \overset{(1.72)}{=} \mathbf{Q}^T\mathbf{Q}\mathbf{v} \cdot \mathbf{e}_i \overset{(1.139)}{=} \mathbf{v} \cdot \mathbf{e}_i,$$

which, by the definition (1.56) of vector components, says that the components of $\mathbf{Q}\mathbf{v}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ equal the components of $\mathbf{v}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. This is not surprising since we transform the vector $\mathbf{v}$ and the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by the same orthogonal tensor $\mathbf{Q}$ to get $\mathbf{Q}\mathbf{v}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

– If we think of a basis as an "observer" who sees the vector through its components in that basis, then the observer $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ sees the vector $\mathbf{v}$ exactly as the observer $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ sees the vector $\mathbf{Q}\mathbf{v}$.

– Consider a scalar-valued function $\varphi(\mathbf{v})$, $\mathbf{v} \in \mathsf{V}$, that has the property

$$\varphi(\mathbf{Q}\mathbf{v}) = \varphi(\mathbf{v}) \qquad \text{for all orthogonal } \mathbf{Q}. \tag{1.140}$$

Such a function is said to be *isotropic* (or invariant). For example, since $\mathbf{Q}\mathbf{v} \cdot \mathbf{Q}\mathbf{v} \overset{(1.72)}{=} \mathbf{Q}^T\mathbf{Q}\mathbf{v} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{v}$, it follows that

$$\varphi(\mathbf{v}) = \mathbf{v} \cdot \mathbf{v},$$

is isotropic

– Let v_i be the components of $\mathbf{v}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Then there is an associated function $\widehat{\varphi}(v_1, v_2, v_3)$ defined on $\mathbb{R}^3$ such that

$$\varphi(\mathbf{v}) = \widehat{\varphi}(v_1, v_2, v_3).$$

For the example $\varphi(\mathbf{v}) = \mathbf{v} \cdot \mathbf{v}$,

$$\widehat{\varphi}(v_1, v_2, v_3) = v_1^2 + v_2^2 + v_3^2.$$

When $\varphi(\mathbf{v})$ is isotropic, the function $\widehat{\varphi}(v_1, v_2, v_3)$ does *not* depend on the basis, i.e.

$$\widehat{\varphi}(v'_1, v'_2, v'_3) = \widehat{\varphi}(v_1, v_2, v_3),$$

where v'_1, v'_2, v'_3 are the components of $\mathbf{v}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ and the same function $\widehat{\varphi}$ appears on both side of this equation. Every observer $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ sees an isotropic function identically. In the example above,

$$\widehat{\varphi}(v'_1, v'_2, v'_3) = (v'_1)^2 + (v'_2)^2 + (v'_3)^3 = v_1^2 + v_2^2 + v_3^2 = \widehat{\varphi}(v_1, v_2, v_3).$$

– In contrast, the function

$$\varphi(\mathbf{v}) = \mathbf{c} \cdot \mathbf{v}, \qquad \mathbf{c} \neq \mathbf{o},$$

(for example) is not isotropic. In this case

$$\widehat{\varphi}(v_1', v_2', v_3') \neq \widehat{\varphi}(v_1, v_2, v_3),$$

though there is of course a different function $\widetilde{\varphi}(v_1', v_2', v_3')$ such that

$$\varphi(\mathbf{v}) = \widehat{\varphi}(v_1, v_2, v_3) = \widetilde{\varphi}(v_1', v_2', v_3').$$

– *Representation theorem*. Corresponding to any isotropic scalar-valued function $\varphi(\mathbf{v})$ there exists a function $\overline{\varphi}(\cdot)$ such that

$$\varphi(\mathbf{v}) = \overline{\varphi}(|\mathbf{v}|).$$

Tensors:

– Likewise for any tensor $\mathbf{C}$,

$$\mathbf{QCQ}^T\mathbf{e}_j' \cdot \mathbf{e}_i' \overset{(1.138)}{=} \mathbf{QCQ}^T(\mathbf{Qe}_j) \cdot (\mathbf{Qe}_i) \overset{(1.139)}{=} \mathbf{QCe}_j \cdot \mathbf{Qe}_i \overset{(1.72)}{=} \mathbf{Q}^T\mathbf{QCe}_j \cdot \mathbf{e}_i \overset{(1.139)}{=} \mathbf{Ce}_j \cdot \mathbf{e}_i,$$

and thus by the definition (1.125) of tensor components, the components of $\mathbf{QCQ}^T$ in the basis $\{\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\}$ equal the components of $\mathbf{C}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Thus the observer $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ sees the tensor $\mathbf{C}$ exactly as the observer $\{\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\}$ sees the tensor $\mathbf{QCQ}^T$.

– A scalar-valued function $\varphi(\mathbf{C})$ defined for all symmetric tensors $\mathbf{C}$ is said to be *isotropic* (or invariant) if

$$\phi(\mathbf{C}) = \phi(\mathbf{QCQ}^T) \qquad \text{for all orthogonal tensors } \mathbf{Q}. \tag{1.141}$$

An example of such a function is (Problem 1.5.2)

$$\varphi(\mathbf{C}) = \operatorname{tr} \mathbf{C}^2 \qquad \text{for all symmetric } \mathbf{C} \in \mathsf{Lin}.$$

– Let C_{ij} be the components of $\mathbf{C}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Then there is an associated function $\widehat{\varphi}(C_{11}, C_{12}, \ldots, C_{32}, C_{33})$ defined on $\mathbb{R}^9$ such that

$$\varphi(\mathbf{C}) = \widehat{\varphi}(C_{11}, C_{12}, \ldots, C_{32}, C_{33}).$$

For the example $\varphi(\mathbf{C}) = \operatorname{tr} \mathbf{C}^2$,

$$\widehat{\varphi}(C_{11}, C_{12}, \ldots, C_{32}, C_{33}) = C_{11}^2 + C_{12}^2 + \ldots + C_{32}^2 + C_{33}^2.$$

When $\varphi(\mathbf{C})$ is isotropic, the function $\widehat{\varphi}(C_{11}, C_{12}, \ldots, C_{32}, C_{33})$ does *not* depend on the basis, i.e.

$$\widehat{\varphi}(C_{11}, C_{12}, \ldots, C_{32}, C_{33}) = \widehat{\varphi}(C'_{11}, C'_{12}, \ldots, C'_{32}, C'_{33}),$$

where C'_{ij} are the components of $\mathbf{C}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$. Every observer $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ sees an isotropic function identically.

– The three principal scalar invariants $I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C})$ introduced in (1.105) are isotropic functions. (Problem 1.5.2.) It is because of this invariance that they are called "invariants", the reason for the adjective "principal" being that they are the particular invariants that appear in the characteristic polynomial (1.104).

– *Representation theorem.* Corresponding to any isotropic scalar-valued function $\varphi(\mathbf{C})$ of a symmetric tensor $\mathbf{C}$ there exists a function $\overline{\varphi}(\cdot,\cdot,\cdot)$ such that

$$\varphi(\mathbf{C}) = \overline{\varphi}(I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C})). \tag{1.142}$$

– A scalar-valued function $\varphi(\mathbf{C}, \mathbf{M})$ defined for all symmetric tensors $\mathbf{C}$ and $\mathbf{M}$ is said to be jointly isotropic in its arguments if

$$\varphi(\mathbf{C}, \mathbf{M}) = \varphi(\mathbf{QCQ}^T, \mathbf{QMQ}^T) \qquad \text{for all orthogonal } \mathbf{Q}. \tag{1.143}$$

1.5.1 Worked examples.

Problem 1.5.1. Show that the functions

$$(a)\ \phi(\mathbf{C}) = \operatorname{tr}\mathbf{C}^n \text{ where } n \text{ is a positive integer}, \qquad \text{and } (b)\ \phi(\mathbf{C}) = \det\mathbf{C},$$

are isotropic.

Solution: (a) Since

$$\mathbf{C}^n = \underbrace{\mathbf{CC}\ldots\mathbf{CC}}_{n \text{ times}} \tag{i}$$

it follows that

$$(\mathbf{QCQ}^T)^n = \underbrace{\mathbf{QCQ}^T\mathbf{QCQ}^T\ldots\mathbf{QCQ}^T\mathbf{QCQ}^T}_{n \text{ times}} = \mathbf{QC}^n\mathbf{Q}^T. \tag{ii}$$

Therefore

$$\phi(\mathbf{QCQ}^T) = \operatorname{tr}(\mathbf{QCQ}^T)^n \overset{(ii)}{=} \operatorname{tr}(\mathbf{QC}^n\mathbf{Q}^T) \overset{(1.117)}{=} \mathbf{QC}^n \cdot \mathbf{Q} \overset{(1.120)}{=} \mathbf{C}^n \cdot \mathbf{Q}^T\mathbf{Q} \overset{(1.95)}{=} \mathbf{C}^n \cdot \mathbf{I} = \operatorname{tr}\mathbf{C}^n = \phi(\mathbf{C}).$$

Therefore $\phi(\mathbf{C}) = \operatorname{tr}\mathbf{C}^n$ is isotropic.

(b) In view of (1.89) and (1.94),

$$\phi(\mathbf{QCQ}^T) = \det\ \mathbf{QCQ}^T = \det\mathbf{Q}\ \det\mathbf{C}\ \det\mathbf{Q}^T = \det\ \mathbf{C} = \phi(\mathbf{C}),$$

and so $\phi(\mathbf{C}) = \det\ \mathbf{C}$ is isotropic.

Problem 1.5.2. Show that the principal scalar invariant functions,

$$I_1(\mathbf{C}) = \operatorname{tr}\mathbf{C}, \qquad I_2(\mathbf{C}) = \frac{1}{2}\left[\operatorname{tr}\mathbf{C}^2 - (\operatorname{tr}\mathbf{C})^2\right], \qquad I_3(\mathbf{C}) = \det\ \mathbf{C}, \tag{1.144}$$

are isotropic i.e. show that

$$I_i(\mathbf{C}) = I_i(\mathbf{QCQ}^T),\ \ i = 1,2,3, \qquad \text{for all orthogonal } \mathbf{Q}. \tag{1.145}$$

Solution: This follows immediately from the results of Problem 1.5.1.

Problem 1.5.3. Consider the scalar-valued function Φ defined for all symmetric tensors $\mathbf{C}$ by

$$\Phi(\mathbf{C}) = \mathbf{GC}\cdot\mathbf{C} \tag{i}$$

where $\mathbf{G}$ is some constant tensor. If $\Phi(\mathbf{C})$ is isotropic, what does this say about the form of the tensor $\mathbf{G}$?

Solution: Only the symmetric part of $\mathbf{G}$ affects the value of the function Φ and so we might as well assume $\mathbf{G}$ to be symmetric; see Problem (1.14). It follows from (i) that

$$\Phi(\mathbf{QCQ}^T) = \mathbf{GQCQ}^T\cdot\mathbf{QCQ}^T \stackrel{(1.120)}{=} \mathbf{Q}^T\mathbf{GQCQ}^T\cdot\mathbf{CQ}^T \stackrel{(1.120)}{=} \mathbf{Q}^T\mathbf{GQCQ}^T\mathbf{Q}\cdot\mathbf{C} \stackrel{(1.95)}{=} \mathbf{Q}^T\mathbf{GQC}\cdot\mathbf{C}$$

Since Φ is isotropic it follows from (1.143), (i) and this that

$$\mathbf{Q}^T\mathbf{GQC}\cdot\mathbf{C} = \mathbf{GC}\cdot\mathbf{C} \qquad \Rightarrow \qquad (\mathbf{Q}^T\mathbf{GQ} - \mathbf{G})\mathbf{C}\cdot\mathbf{C} = 0.$$

This is to hold for all symmetric tensors $\mathbf{C}$ and so it follows from part (e) of Problem 1.4.2 that $\mathbf{Q}^T\mathbf{GQ} - \mathbf{G}$ must be skew-symmetric:

$$\mathbf{Q}^T\mathbf{GQ} - \mathbf{G} = -\big(\mathbf{Q}^T\mathbf{GQ} - \mathbf{G}\big)^T = -\big(\mathbf{Q}^T\mathbf{GQ}\big)^T - \mathbf{G} \stackrel{(1.74)}{=} -\mathbf{Q}^T\mathbf{GQ} - \mathbf{G}$$

where we have used the symmetry of $\mathbf{G}$ and $\big(\mathbf{Q}^T\big)^T = \mathbf{Q}$. Therefore

$$\mathbf{Q}^T\mathbf{GQ} = \mathbf{G}.$$

Since this is to hold for all orthogonal $\mathbf{Q}$ it follows from Problem 1.35 that $\mathbf{G}$ must be a scalar multiple of the identity:

$$\mathbf{G} = \gamma\mathbf{I} \qquad \text{for some scalar constant } \gamma.$$

Remark: Therefore $\Phi(\mathbf{C}) = \mathbf{GC}\cdot\mathbf{C} = \gamma\mathbf{C}\cdot\mathbf{C} = \gamma\operatorname{tr}(\mathbf{C}^2)$ and so this is consistent with the general representation (1.142).

Problem 1.5.4. It was shown in Problem 1.5.1 that $\operatorname{tr}\mathbf{C}^n$ is isotropic for all positive integers n. Express the principal scalar invariant $I_3(\mathbf{C})$ in terms of $\operatorname{tr}\mathbf{C}$, $\operatorname{tr}\mathbf{C}^2$ and $\operatorname{tr}\mathbf{C}^3$. Do the same for $\operatorname{tr}\mathbf{C}^4$. Hint: Use the Cayley-Hamilton theorem (1.106).

Problem 1.5.5. Show that the functions

$$(a)\ \phi(\mathbf{C},\mathbf{M}) = \mathbf{C}\cdot\mathbf{M}, \qquad \text{and } (b)\ \phi(\mathbf{C},\mathbf{M}) = \mathbf{C}^2\cdot\mathbf{M}.$$

are jointly isotropic in $\mathbf{C}$ and $\mathbf{M}$.

Solution: (a) To show that $\phi(\mathbf{C},\mathbf{M}) = \mathbf{C}\cdot\mathbf{M}$ is isotropic, we proceed as follows:

$$\phi(\mathbf{QCQ}^T,\mathbf{QMQ}^T) = \mathbf{QCQ}^T\cdot\mathbf{QMQ}^T \overset{(1.120)}{=} \mathbf{C}\cdot\mathbf{Q}^T\mathbf{QMQ}^T\mathbf{Q} = \mathbf{C}\cdot\mathbf{M} = \phi(\mathbf{C},\mathbf{M}).$$

(b) That $\phi(\mathbf{C},\mathbf{M}) = \mathbf{C}^2\cdot\mathbf{M}$ is isotropic can be seen from

$$\phi(\mathbf{QCQ}^T,\mathbf{QMQ}^T) = (\mathbf{QCQ}^T)^2\cdot\mathbf{QMQ}^T = (\mathbf{QCQ}^T\mathbf{QCQ}^T)\cdot\mathbf{QMQ}^T = \mathbf{QC}^2\mathbf{Q}^T\cdot\mathbf{QMQ}^T,$$

followed by using (1.120) as in the preceding example.

1.6 Change of basis. Cartesian tensors.

We now look at the components of a vector/tensor in *two* orthonormal bases and examine how these components are related. The vector/tensor stays fixed while the basis changes.

1.6.1 Two orthonormal bases.

We first make some observations on the relation between the two bases. In accordance with the standard way in which this topic is discussed in the literature, let $\mathbf{Q}^T$ be the orthogonal tensor that maps the orthonormal basis $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ into the orthonormal basis $\{\mathbf{e}'_1,\mathbf{e}'_2,\mathbf{e}'_3\}$:

$$\mathbf{e}'_i = \mathbf{Q}^T\mathbf{e}_i, \qquad \mathbf{e}_i = \mathbf{Q}\mathbf{e}'_i. \tag{1.146}$$

One can readily show that the components of $\mathbf{Q}$ in the two bases coincide,

$$\mathbf{Q} = Q_{ij}\mathbf{e}_i\otimes\mathbf{e}_j = Q_{ij}\mathbf{e}'_i\otimes\mathbf{e}'_j,$$

and that

$$Q_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j. \tag{1.147}$$

Given the two bases, equation (1.147) provides a formula for calculating the elements Q_{ij} of the matrix of components $[Q]$. Specifically, Q_{ij} is the cosine of the angle between the basis vectors $\mathbf{e}'_i$ and $\mathbf{e}_j$.

Since $\mathbf{Q}$, and therefore $[Q]$, is orthogonal,

$$[Q][Q]^T = [Q]^T[Q] = [I], \qquad Q_{ik}Q_{jk} = Q_{ki}Q_{kj} = \delta_{ij}. \tag{1.148}$$

If one basis can be rotated into the other as is the case if both bases are right-handed or both are left-handed, then $[Q]$ is proper orthogonal ($\det[Q] = +1$). Otherwise $[Q]$ is improper orthogonal ($\det[Q] = -1$).

Exercise: Show that

$$\mathbf{e}'_i = Q_{ij}\mathbf{e}_j, \qquad \mathbf{e}_i = Q_{ji}\mathbf{e}'_j\ , \tag{1.149}$$

1.6.2 Vectors: 1-tensors.

– Let v_i and v'_i be the components of the same vector $\mathbf{v}$ in the respective bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$:

$$v_i = \mathbf{v} \cdot \mathbf{e}_i, \qquad v'_i = \mathbf{v} \cdot \mathbf{e}'_i.$$

This is illustrated in Figure 1.5.

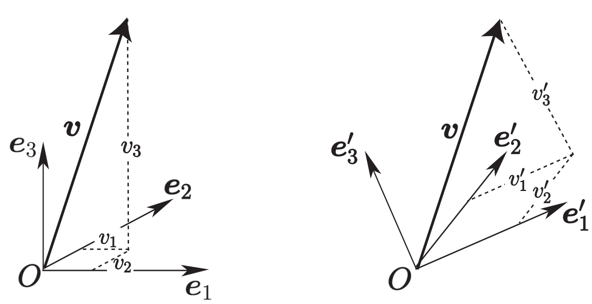


Figure 1.5: Components $\{v_1, v_2, v_3\}$ and $\{v'_1, v'_2, v'_3\}$ of the same vector $\mathbf{v}$ in two different bases.

– These components are related by (Exercise)

$$v'_i = Q_{ij} v_j, \qquad \{v'\} = [Q]\{v\},$$

and equivalently

$$v_i = Q_{ji} v'_j, \qquad \{v\} = [Q]^T\{v'\}.$$

– A quantity whose components v_i and v'_i in (any) two orthonormal bases are related by

$$\left.\begin{aligned} v'_i &= Q_{ij} v_j, & v_i &= Q_{ji} v'_j, \\ \{v'\} &= [Q]\{v\}, & \{v\} &= [Q]^T\{v'\}, \end{aligned}\right\} \tag{1.150}$$

is called a 1st-order cartesian tensor or *1-tensor*.

Observe that if we know the components of a 1-tensor in one basis, its components in any other basis can be calculated using (1.150).

1.6.3 Linear transformations: 2-tensors.

– Let A_{ij} and A'_{ij} be the respective components of the same linear transformation $\mathbf{A}$ in the two bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$:

$$A_{ij} = \mathbf{e}_i \cdot (\mathbf{A}\mathbf{e}_j), \qquad A'_{ij} = \mathbf{e}'_i \cdot (\mathbf{A}\mathbf{e}'_j).$$

– These components are related by (Exercise)

$$A'_{ij} \;=\; Q_{ip} Q_{jq} A_{pq}, \qquad [A'] = [Q][A][Q]^T,$$

the inverse relation being

$$A_{ij} = Q_{pi} Q_{qj} A'_{pq}, \qquad [A] = [Q]^T[A'][Q].$$

– An entity whose components A_{ij} and A'_{ij} in every two orthonormal bases are related by

$$\left.\begin{aligned} A'_{ij} &= Q_{ip} Q_{jq} A_{pq}, & A_{ij} &= Q_{pi} Q_{qj} A'_{pq}, \\ [A'] &= [Q][A][Q]^T, & [A] &= [Q]^T[A'][Q], \end{aligned}\right\} \tag{1.151}$$

is called a 2nd-order cartesian tensor or a *2-tensor*.

Observe that if we know the components of a 2-tensor in one basis, its components in any other basis can be calculated using (1.151).

– Exercise: Show that the components of the identity 2-tensor $\mathbf{I}$ in *every* basis are δ_{ij}.

1.6.4 n-tensors.

The concept of an *n-tensor* can be introduced analogously. Let $\mathbb{T}$ be an entity that is defined in a given basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by a set of 3^n ordered numbers $\mathbb{T}_{i_1 i_2 \ldots . i_n}$. The numbers $\mathbb{T}_{i_1 i_2 \ldots . i_n}$ are called the *components* of $\mathbb{T}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. For example if $\mathbb{T}$ is a scalar, 1-tensor or 2-tensor, it is represented by a set of $3^0, 3^1$ or 3^2 ordered numbers respectively. Let $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ be a second basis related to the first one by an orthogonal matrix [Q] as described by (1.149). Let $\mathbb{T}'_{i_1 i_2 \ldots . i_n}$ be the components of the entity $\mathbb{T}$ in the second basis. If for every choice of bases these two sets of components are related by

$$\mathbb{T}'_{i_1 i_2 \ldots . i_n} \;=\; Q_{i_1 j_1}\, Q_{i_2 j_2}\, \ldots .\, Q_{i_n j_n}\, \mathbb{T}_{j_1 j_2 \ldots . j_n} \tag{1.152}$$

the entity $\mathbb{T}$ is said to be an *n-tensor*. *Thus, the components of a tensor in every basis may be determined if its components in any one basis are known.*

The *quotient rule* states that if $(\mathbb{T}\mathbb{B})_{i_1 i_2 .. i_n}$ is an n-tensor for every m-tensor $\mathbb{B}_{j_1 j_2 \ldots j_m}$, then $\mathbb{T}$ is an ℓ-tensor where

$$(\mathbb{T}\mathbb{B})_{i_1 i_2 .. i_n} = \mathbb{T}_{k_1 k_2 \ldots k_\ell}\, \mathbb{B}_{j_1 j_2 \ldots j_m}. \tag{1.153}$$

Some of the subscripts on the right-hand side of (1.153) maybe repeated. This is called the quotient rule (since it appears to say that the ratio of two tensors is a tensor).

1.6.5 Worked examples.

Problem 1.6.1. Let $[A]$ and $[A']$ be the components of a 2-tensor $\mathbf{A}$ in two orthonormal bases. In general, $[A] \neq [A']$. Show that:

(a) If $[A]$ is symmetric, then so is $[A']$:

$$[A] = [A]^T \qquad \Leftrightarrow \qquad [A'] = [A']^T. \tag{1.154}$$

(b) The trace of the matrices $[A]$ and $[A']$ are equal:

$$\mathrm{tr}\;[A'] = \mathrm{tr}\;[A]. \tag{1.155}$$

(c) The determinants of the matrices $[A]$ and $[A']$ are equal:

$$\det [A'] = \det [A]. \tag{1.156}$$

Remark: Thus, the three characteristics symmetry, trace and determinant are independent of the choice of basis. Therefore there is no ambiguity in defining

(a) a tensor to be symmetric if its matrix of components is symmetric,
(b) the trace of a tensor to be the trace of its matrix of components, and
(c) the determinant of a tensor to be the determinant of its matrix of components.

Solution:

(a) Using the properties $([B][C])^T = [C]^T[B]^T$ and $([B]^T)^T = [B]$ we have

$$[A']^T \overset{(1.151)}{=} \big([Q][A][Q]^T\big)^T = [Q][A]^T[Q]^T = [Q][A][Q]^T = [A']. \qquad \square$$

(b)

$$\mathrm{tr}[A'] = A'_{ii} \overset{(1.151)}{=} \big([Q][A][Q]^T\big)_{ii} = Q_{ij}A_{jk}Q^T_{ki} = Q_{ij}A_{jk}Q_{ik} = Q_{ij}Q_{ik}A_{jk} \overset{(1.130)}{=} \delta_{jk}A_{jk} = A_{jj} = \mathrm{tr}[A]. \quad \square$$

(c) Using the properties $\det([B][C]) = \det[B]\det[C]$ and $\det([B]^T) = \det[B]$ we have

$$\det[A'] \overset{(1.151)}{=} \det\big([Q][A][Q]^T\big) = \det[Q]\ \det[A]\ \det[Q]^T = \big(\det[Q]\big)^2\ \det[A] = \det[A]. \qquad \square$$

Problem 1.6.2. The Kronecker delta obeys

$$\delta_{ij} = Q_{ip}Q_{jq}\delta_{pq} \qquad (i)$$

for all orthogonal matrices $[Q]$ which follows from the substitution rule and (1.130). Show that the Levi-Civita symbol obeys

$$e_{ijk} = \pm Q_{ip}Q_{jq}Q_{kr}e_{pqr} \qquad (ii)$$

for all orthogonal matrices $[Q]$ where the $+$ and $-$ signs hold if $[Q]$ is proper and improper orthogonal respectively.

Remark: If equation (ii) had held with the $\pm$ replaced by $+$ we would have used the term Levi-Civita tensor. Perhaps you have been wondering why we speak of the Levi-Civita *symbol* and not the Levi-Civita tensor!

Solution: Equation (ii) follows immediately upon taking $[A] = [Q]$ in (1.40), multiplying both sides by $Q_{ap}Q_{bq}Q_{cr}$ and using (1.130).

Problem 1.6.3. Consider a scalar-valued function $\phi(\mathbf{A})$. Let $[A]$ be the matrix of components of $\mathbf{A}$ in an arbitrary orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Then, there is a function $\widehat{\phi}(\cdot)$ defined on the set of all 3×3 matrices such that $\phi(\mathbf{A}) = \widehat{\phi}([A])$. Note that the function $\widehat{\phi}$ depends on the choice of basis.

Let $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ be a second orthonormal basis and let $[A']$ be the matrix of components of $\mathbf{A}$ in this basis.

If $\phi(\mathbf{A})$ is an *isotropic function*, show that

$$\widehat{\phi}([A]) = \widehat{\phi}([A']),$$

where the function $\widehat{\phi}$ is the *same* on both sides.

Problem 1.6.4. The triplet of orthonormal vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is carried into the set of vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ by a tensor $\mathbf{R}$:

$$\mathbf{e}'_1 = \mathbf{R}\mathbf{e}_1, \qquad \mathbf{e}'_2 = \mathbf{R}\mathbf{e}_2, \qquad \mathbf{e}'_3 = \mathbf{R}\mathbf{e}_3. \tag{1.157}$$

Show that

(a) the set of vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is orthonormal if and only if $\mathbf{R}$ is orthogonal.

Assume from hereon that $\mathbf{R}$ is orthogonal.

(b) Show that $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \mathbf{e}'_i \otimes \mathbf{e}_i = \mathbf{e}'_1 \otimes \mathbf{e}_1 + \mathbf{e}'_2 \otimes \mathbf{e}_2 + \mathbf{e}'_3 \otimes \mathbf{e}_3. \tag{1.158}$$

Note that (1.133) is a special case of this.

(c) If $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ are both right-handed (or both left-handed), show that $\mathbf{R}$ is proper orthogonal (and therefore represents a rotation). If one is right-handed and the other left-handed, show that $\mathbf{R}$ is improper orthogonal.

(d) If $[R]$ and $[R']$ are the matrices of components of the tensor $\mathbf{R}$ in the respective bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$, show that $[R] = [R']$; and

Solution:

(a) Suppose that $\mathbf{R}$ is orthogonal. To show that the triplet of vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is orthonormal we must show that $\mathbf{e}'_i \cdot \mathbf{e}'_j = \delta_{ij}$. This follows from

$$\mathbf{e}'_i \cdot \mathbf{e}'_j \overset{(1.157)}{=} \mathbf{R}\mathbf{e}_i \cdot \mathbf{R}\mathbf{e}_j \overset{(1.72)}{=} \mathbf{R}^T\mathbf{R}\mathbf{e}_i \cdot \mathbf{e}_j \overset{(1.95)}{=} \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}. \tag{i}$$

where in the last step we used the fact that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is an orthonormal set of vectors.

Conversely suppose that $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is orthonormal. Then it is left as an exercise to show that $\mathbf{R}$ is orthogonal.

(b)

$$\mathbf{e}'_i \otimes \mathbf{e}_i \overset{(1.157)}{=} \mathbf{R}\mathbf{e}_i \otimes \mathbf{e}_i = \mathbf{R}\big(\mathbf{e}_i \otimes \mathbf{e}_i\big) \overset{(1.133)}{=} \mathbf{R}\mathbf{I} = \mathbf{R}.$$

(d) By the definition (1.125) of the components of a tensor in a basis,

$$R_{ij} = \mathbf{e}_i \cdot \mathbf{R}\mathbf{e}_j, \qquad R'_{ij} = \mathbf{e}'_i \cdot \mathbf{R}\mathbf{e}'_j. \tag{ii}$$

Therefore
$$R'_{ij} = \mathbf{e}'_i \cdot \mathbf{R}\mathbf{e}'_j \stackrel{(1.157)}{=} \mathbf{R}\mathbf{e}_i \cdot \mathbf{R}\mathbf{R}\mathbf{e}_j \stackrel{(1.72)}{=} \mathbf{e}_i \cdot \mathbf{R}^T\mathbf{R}\mathbf{R}\mathbf{e}_j \stackrel{(1.95)}{=} \mathbf{e}_i \cdot \mathbf{R}\mathbf{e}_j \stackrel{(ii)_1}{=} R_{ij}.$$

Problem 1.6.5. (See also Problem 1.30.) When we analyze the bending deformation of a block (see page 148), it will be natural to use rectangular cartesian coordinates in the reference configuration and cylindrical polar coordinates in the deformed configuration. In such settings we will work simultaneously with two bases. That is the motivation for this problem.

Given two orthonormal bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$, consider the tensor $\mathbf{F}$ that has the representation
$$\mathbf{F} = \Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j. \tag{i}$$
Note that *both* bases appear on the right-hand side of (i) whence Φ_{ij} are *not* the components of $\mathbf{F}$ in either basis.

(a) If x_i are the components of a vector $\mathbf{x}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, calculate the components of the vector $\mathbf{y} = \mathbf{F}\mathbf{x}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

(b) Derive a representation analogous to (i) for $\mathbf{F}^T$.

(c) Calculate the components of $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and those of $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

(d) Suppose that the two bases are related by $\mathbf{e}'_i = Q_{ij}\mathbf{e}_j$, $\mathbf{e}_i = Q_{ji}\mathbf{e}'_j$ where $[Q]$ is an orthogonal matrix. Calculate the components of $\mathbf{y}$ and $\mathbf{B}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and the components of $\mathbf{C}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

(e) Calculate the components of $\mathbf{F}$ in both bases.

Solution:

(a)

$$\mathbf{y} = \mathbf{F}\mathbf{x} = (\Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j)(x_p\mathbf{e}_p) = \Phi_{ij}x_p(\mathbf{e}'_i \otimes \mathbf{e}_j)\mathbf{e}_p \stackrel{(1.68)}{=} \Phi_{ij}x_p(\mathbf{e}_j \cdot \mathbf{e}_p)\mathbf{e}'_i = \Phi_{ij}x_p\delta_{jp}\mathbf{e}'_i = \Phi_{ip}x_p\ \mathbf{e}'_i, \qquad \square$$

and so we can write
$$\mathbf{y} = y'_i\mathbf{e}'_i \qquad \text{where} \qquad y'_i = \Phi_{ip}x_p, \quad \{y'\} = [\Phi]\{x\};$$
here y'_i are the components of $\mathbf{y}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$. Observe that the vector equation $\mathbf{y} = \mathbf{F}\mathbf{x}$ is equivalent to the matrix equation $\{y'\} = [\Phi]\{x\}$ where $\{y'\}$ are the components of $\mathbf{y}$ in one basis and $\{x\}$ are the components of $\mathbf{x}$ in the other basis.

(b) To determine the transpose we use the definition (1.72), i.e. $\mathbf{F}\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{F}^T\mathbf{y}$. Using (i),
$$\mathbf{F}\mathbf{x} \cdot \mathbf{y} = \Phi_{ij}(\mathbf{e}'_i \otimes \mathbf{e}_j)\mathbf{x} \cdot \mathbf{y} = \Phi_{ij}(\mathbf{e}_j \cdot \mathbf{x})(\mathbf{e}'_i \cdot \mathbf{y}) \stackrel{(*)}{=} \Phi_{ji}(\mathbf{e}_i \cdot \mathbf{x})(\mathbf{e}'_j \cdot \mathbf{y}) = \Phi_{ji}(\mathbf{e}_i \otimes \mathbf{e}'_j)\mathbf{y} \cdot \mathbf{x},$$
where in step $(*)$ we changed the dummy subscripts. Therefore
$$\mathbf{F}^T = \Phi_{ji}(\mathbf{e}_i \otimes \mathbf{e}'_j) \stackrel{(**)}{=} \Phi_{ij}(\mathbf{e}_j \otimes \mathbf{e}'_i), \qquad \square$$

where in step ($**$) we have changed the dummy subscripts again.

(c) First,

$$\begin{aligned}\mathbf{C} &= \mathbf{F}^T\mathbf{F} = (\Phi_{pq}\mathbf{e}'_p \otimes \mathbf{e}_q)^T(\Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j) = (\Phi_{pq}\mathbf{e}_q \otimes \mathbf{e}'_p)(\Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j) = \Phi_{pq}\Phi_{ij}(\mathbf{e}_q \otimes \mathbf{e}'_p)(\mathbf{e}'_i \otimes \mathbf{e}_j) = \\ &\overset{(1.70)}{=} \Phi_{pq}\Phi_{ij}(\mathbf{e}'_p \cdot \mathbf{e}'_i)(\mathbf{e}_q \otimes \mathbf{e}_j) = \Phi_{pq}\Phi_{ij}\delta_{pi}(\mathbf{e}_q \otimes \mathbf{e}_j) = \Phi_{pq}\Phi_{pj}(\mathbf{e}_q \otimes \mathbf{e}_j) = \Phi_{pi}\Phi_{pj}(\mathbf{e}_i \otimes \mathbf{e}_j),\end{aligned}$$

and so we can write

$$\mathbf{C} = C_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j) \qquad \text{where} \qquad C_{ij} = \Phi_{pi}\Phi_{pj}, \quad [C] = [\Phi]^T[\Phi]; \qquad \square$$

here C_{ij} are the components of $\mathbf{C}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

Second,

$$\begin{aligned}\mathbf{B} &= \mathbf{F}\mathbf{F}^T = (\Phi_{pq}\mathbf{e}'_p \otimes \mathbf{e}_q)(\Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j)^T = (\Phi_{pq}\mathbf{e}'_p \otimes \mathbf{e}_q)(\Phi_{ij}\mathbf{e}_j \otimes \mathbf{e}'_i). \overset{(1.70)}{=} \Phi_{pq}\Phi_{ij}(\mathbf{e}_q \cdot \mathbf{e}_j)(\mathbf{e}'_p \otimes \mathbf{e}'_i) = \\ &= \Phi_{pq}\Phi_{ij}\delta_{qj}(\mathbf{e}'_p \otimes \mathbf{e}'_i) = \Phi_{pq}\Phi_{iq}(\mathbf{e}'_p \otimes \mathbf{e}'_i) = \Phi_{iq}\Phi_{jq}(\mathbf{e}'_i \otimes \mathbf{e}'_j),\end{aligned}$$

and so we can write

$$\mathbf{B} = B'_{ij}(\mathbf{e}'_i \otimes \mathbf{e}'_j) \qquad \text{where} \qquad B'_{ij} = \Phi_{iq}\Phi_{jq}, \quad [B'] = [\Phi][\Phi]^T; \qquad \square$$

here B'_{ij} are the components of $\mathbf{B}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

Observe that the components of $\mathbf{C}$ were naturally expressed in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ while those of $\mathbf{B}$ have been expressed in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$. Moreover, even though $[\Phi]$ is *not* the matrix of components of $\mathbf{F}$ in either basis, $[\Phi]^T[\Phi]$ is the matrix of components of $\mathbf{F}^T\mathbf{F}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, and $[\Phi][\Phi]^T$ is the matrix of components of $\mathbf{F}\mathbf{F}^T$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$.

(d) Exercise!

(e) In order to determine the components of $\mathbf{F}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ we must express $\mathbf{F} = \Phi_{ij}\, \mathbf{e}'_i \otimes \mathbf{e}_j$ in the form $\mathbf{F} = F_{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j$ by eliminating $\mathbf{e}'_i$ in favor of $\mathbf{e}_i$. Likewise, to determine the components of $\mathbf{F}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ we must express $\mathbf{F} = \Phi_{ij}\, \mathbf{e}'_i \otimes \mathbf{e}_j$ in the form $\mathbf{F} = F'_{ij}\, \mathbf{e}'_i \otimes \mathbf{e}'_j$ by eliminating $\mathbf{e}_i$ in favor of $\mathbf{e}'_i$. This can be done by using the relations

$$\mathbf{e}'_i = Q_{ij}\mathbf{e}_j, \qquad \mathbf{e}_i = Q_{ji}\mathbf{e}'_j. \qquad (ii)$$

We now find the components of $\mathbf{F}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by using $(ii)_1$ to eliminate $\mathbf{e}'_i$:

$$\mathbf{F} = \Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j \overset{(vi)_1}{=} \Phi_{ij}(Q_{ip}\mathbf{e}_p) \otimes \mathbf{e}_j \overset{(*)}{=} \Phi_{kj}Q_{ki}\mathbf{e}_i \otimes \mathbf{e}_j = Q_{ki}\Phi_{kj}\mathbf{e}_i \otimes \mathbf{e}_j = ([Q]^T[\Phi])_{ij}\mathbf{e}_i \otimes \mathbf{e}_j = F_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$$

(where in step ($*$) we changed the dummy subscripts $i \to k$ and $p \to i$) and so

$$[F] = [Q]^T[\Phi]. \qquad \square$$

Likewise we find the components of $\mathbf{F}$ in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ by using $(ii)_2$ to eliminate $\mathbf{e}_i$:

$$\mathbf{F} = \Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j \overset{(vi)_2}{=} \Phi_{ij}\mathbf{e}'_i \otimes (Q_{pj}\mathbf{e}'_p) = \Phi_{ij}Q_{pj}\mathbf{e}'_i \otimes \mathbf{e}'_p \overset{(*)}{=} \Phi_{ik}Q_{jk}\mathbf{e}'_i \otimes \mathbf{e}'_j = ([\Phi][Q]^T)_{ij}\mathbf{e}'_i \otimes \mathbf{e}'_j = F'_{ij}\mathbf{e}'_i \otimes \mathbf{e}'_j$$

(where in step $(*)$ we changed the dummy subscripts $j \to k$ and $p \to j$) and therefore

$$[F'] = [\Phi][Q]^T. \qquad \square$$

Observe that $[\Phi] = [Q][F] = [F'][Q]$ i.e. $[F'] = [Q][F][Q]^T$.

1.7 Euclidean point space.

It later chapters, when we study the mechanical response of a body, the body will occupy some region of "physical space". The quantities of interest, for example the stress, will vary from point to point within this region. In this section we briefly touch on the connection of physical space – a Euclidean point space – to a Euclidean vector space.

A *Euclidean point space* $\mathcal{E}$ is a collection of elements that we call points. Corresponding to each ordered pair of points $p, q \in \mathcal{E}$, there is a unique vector $\vec{pq}$ in an associated Euclidean vector space V with the properties

$$\vec{pq} = -\vec{qp} \qquad \text{and} \qquad \vec{pq} = \vec{pr} + \vec{rq},$$

for all points $p, q, r \in \mathcal{E}$. Observe that each point is not a vector but each (ordered) pair of points is associated with a vector. Geometric characteristics in $\mathcal{E}$, such as distance and angle, are derived from the vector space V in a natural way, e.g. the distance between two points $p, q \in \mathcal{E}$ is defined as the magnitude of the vector $\vec{pq}$, etc. There is no notion of the addition of two points.

Pick and fix a point $o \in \mathcal{E}$ as depicted in Figure 1.6. Then, corresponding to each point $x \in \mathcal{E}$ there is a unique vector $\mathbf{x} = \vec{ox} \in \mathsf{V}$. We refer to $\mathbf{x}$ as the **position vector** of point x relative to the *origin* o. Observe that if $\mathbf{x}$ and $\mathbf{y}$ are the position vectors of points x and y relative to an origin o, the vector $\mathbf{y} - \mathbf{x}$ is in fact independent of the choice of origin.

An origin $o \in \mathcal{E}$ together with an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for V is referred to as a **frame** which we denote by $\{o; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The components (x_1, x_2, x_3) of the position vector $\mathbf{x}$ in this basis are called the **coordinates** of the point x in this frame. When the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is fixed, we refer to $\{o; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ as a **rectangular cartesian frame**.

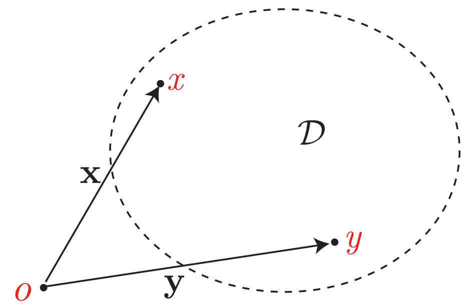


Figure 1.6: Points o, x and y in a Euclidean point space $\mathcal{E}$; ($\mathcal{D}$ is some domain of $\mathcal{E}$). The position vectors of the points x and y relative to the point o – the origin – are the respective vectors $\mathbf{x} = \vec{ox}$ and $\mathbf{y} = \vec{oy}$ of the associated Euclidean vector space.

Since the quantities of interest will vary from point to point in the region of space occupied by the body, they can be treated as functions of the position vector $\mathbf{x}$ (it being implicit that an origin has been chosen).

Let $\mathcal{D}$ be a domain in physical space (a Euclidean point space) and let $\Phi(x)$ be a scalar-valued function defined at each point $x \in \mathcal{D}$. The function $\Phi(x)$ is referred to as a scalar *field* on $\mathcal{D}$. Vector- and tensor fields are defined analogously.

Once an origin $o \in \mathcal{E}$ has been chosen and fixed, there is a one-to-one correspondence between the points x and the associated position vectors $\mathbf{x} = \vec{ox}$. Let D be the set of all position vectors corresponding to the set of all points in $\mathcal{D}$. Then there is a scalar-valued function $\varphi(\mathbf{x})$ defined at each position vector $\mathbf{x} \in \mathsf{D}$ such that

$$\varphi(\mathbf{x}) = \Phi(x), \qquad \mathbf{x} = \vec{ox},$$

for all $x \in \mathcal{D}$. Since the effect of a change of origin from say o to o' is to add the vector $\vec{oo'}$ to all position vectors, the fact that φ depends on the choice of origin is not important. Therefore from hereon we make no distinction between $\varphi, \mathcal{D}$ and Φ, D.

1.8 Calculus.

A brief video on the use of indicial notation in tensor calculus can be found here.

1.8.1 Calculus of scalar, vector and tensor fields.

Throughout this section, $\mathcal{R}$ is a region in three-dimensional space whose boundary is denoted by $\partial\mathcal{R}$. The position vector of a generic point in $\mathcal{R}+\partial\mathcal{R}$ (with respect to some fixed origin) is $\mathbf{x}$. We shall consider a scalar field $\phi(\mathbf{x})$, a vector field $\mathbf{u}(\mathbf{x})$, and a tensor field $\mathbf{T}(\mathbf{x})$, each defined for $\mathbf{x} \in \mathcal{R}+\partial\mathcal{R}$. The region $\mathcal{R}+\partial\mathcal{R}$, and these fields, are assumed to be sufficiently regular so as to permit the calculations carried out below.

In this section we shall work primarily with components in a cartesian (i.e. fixed orthonormal) basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ in which $x_i = \mathbf{x}\cdot\mathbf{e}_i$, $u_i = \mathbf{u}\cdot\mathbf{e}_i$ and $T_{ij} = \mathbf{T}\mathbf{e}_j\cdot\mathbf{e}_i$ denote the respective components of $\mathbf{x}, \mathbf{u}$ and $\mathbf{T}$. In Section 1.8.6 we shall derive corresponding expressions in cylindrical polar coordinates.

Gradient of a scalar field. First consider the scalar field $\phi(\mathbf{x})$ and let $\mathbf{g}(\mathbf{x})$ be the vector field defined by

$$\mathbf{g}(\mathbf{x})\cdot\mathbf{a} := \frac{d}{dt}\phi(\mathbf{x}+t\mathbf{a})\Big|_{t=0} \qquad \text{for all constant vectors } \mathbf{a}, \tag{i}$$

where t is a scalar parameter, or equivalently by[7]

$$\phi(\mathbf{x}+\mathbf{a}) - \phi(\mathbf{x}) = \mathbf{g}(\mathbf{x})\cdot\mathbf{a} + \mathrm{o}(|\mathbf{a}|) \qquad \text{as } |\mathbf{a}| \to 0. \tag{1.159}$$

(Here the remainder term $\mathrm{o}(|\mathbf{a}|)$ is a term that approaches zero faster than $|\mathbf{a}|$.) We call $\mathbf{g}$ the gradient of ϕ and introduce the notation $\mathbf{g} = \operatorname{grad}\phi$ or $\nabla\phi$.

Recall that the partial derivative of ϕ at $\mathbf{x}$ with respect to x_i is defined by

$$\frac{\partial\phi(\mathbf{x})}{\partial x_i} = \frac{d}{dt}\phi(\mathbf{x}+t\mathbf{e}_i)\Big|_{t=0}. \tag{ii}$$

Therefore from (i) and (ii) with $\mathbf{a} = \mathbf{e}_i$ we get

$$\mathbf{g}\cdot\mathbf{e}_i = \frac{\partial\phi}{\partial x_i} \qquad \Rightarrow \qquad \mathbf{g} = \frac{\partial\phi}{\partial x_i}\,\mathbf{e}_i. \tag{iii}$$

Equation $(iii)_2$ is therefore the representation of $\operatorname{grad}\phi$ with respect to the cartesian basis:

$$\operatorname{grad}\,\phi = \nabla\phi = \frac{\partial\phi}{\partial x_i}\,\mathbf{e}_i, \qquad \big(\operatorname{grad}\phi\big)_i = \big(\nabla\phi\big)_i = \frac{\partial\phi}{\partial x_i}. \tag{1.160}$$

[7] We shall make use of this representation when calculating the gradient of a scalar field in cylindrical polar coordinates in Section 1.8.6 .

Gradient of a vector field. Second, consider the vector field $\mathbf{u}(\mathbf{x})$ and let $\mathbf{G}(\mathbf{x})$ be the tensor field defined by

$$\mathbf{G}(\mathbf{x})\mathbf{a} := \frac{d}{dt}\mathbf{u}(\mathbf{x}+t\mathbf{a})\Big|_{t=0} \qquad \text{for all constant vectors } \mathbf{a}, \tag{iv}$$

where t is a scalar parameter, or equivalently by[8]

$$\mathbf{u}(\mathbf{x}+\mathbf{a}) - \mathbf{u}(\mathbf{x}) = \mathbf{G}(\mathbf{x})\mathbf{a} + \mathrm{o}(|\mathbf{a}|) \qquad \text{as } |\mathbf{a}| \to 0. \tag{1.161}$$

We call $\mathbf{G}$ the gradient of $\mathbf{u}$ and introduce the notation $\mathbf{G} = \operatorname{grad}\mathbf{u}$ or $\nabla\mathbf{u}$.

Recall that the partial derivative of $\mathbf{u}$ at $\mathbf{x}$ with respect to x_j is defined by

$$\frac{\partial \mathbf{u}(\mathbf{x})}{\partial x_j} = \frac{d}{dt}\mathbf{u}(\mathbf{x}+t\mathbf{e}_j)\Big|_{t=0}. \tag{v}$$

Then from (iv) and (v) with $\mathbf{a} = \mathbf{e}_j$,

$$\mathbf{G}\mathbf{e}_j = \frac{\partial \mathbf{u}}{\partial x_j}. \tag{vi}$$

If G_{ij} are the components of $\mathbf{G}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$,

$$\mathbf{G} = G_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \tag{vii}$$

then (vi) and (vii) yield

$$(G_{ik}\mathbf{e}_i \otimes \mathbf{e}_k)\mathbf{e}_j = \frac{\partial \mathbf{u}}{\partial x_j} \qquad \Rightarrow \qquad G_{ij}\mathbf{e}_i = \frac{\partial \mathbf{u}}{\partial x_j} \qquad \Rightarrow \qquad G_{ij}\mathbf{e}_i \otimes \mathbf{e}_j = \frac{\partial \mathbf{u}}{\partial x_j} \otimes \mathbf{e}_j. \tag{viii}$$

Therefore from (vii) and $(viii)$

$$\mathbf{G} = \frac{\partial \mathbf{u}}{\partial x_j} \otimes \mathbf{e}_j = \frac{\partial (u_i\mathbf{e}_i)}{\partial x_j} \otimes \mathbf{e}_j = \frac{\partial u_i}{\partial x_j}\mathbf{e}_i \otimes \mathbf{e}_j. \tag{x}$$

Equation (x) is therefore the representation of $\operatorname{grad}\mathbf{u}$ with respect to the cartesian basis:

$$\operatorname{grad}\mathbf{u} = \nabla\mathbf{u} = \frac{\partial u_i}{\partial x_j}\,\mathbf{e}_i \otimes \mathbf{e}_j, \qquad \big(\operatorname{grad}\mathbf{u}\big)_{ij} = \big(\nabla\mathbf{u}\big)_{ij} = \frac{\partial u_i}{\partial x_j}. \tag{1.162}$$

Observe that the gradient of a n-tensor field is a $n+1$-tensor field, while (as we see next) the divergence of a n-tensor field is a $n-1$-tensor field.

[8]We shall make use of this representation when calculating the gradient of a vector field in cylindrical polar coordinates in Section 1.8.6 .

Divergence and curl of a vector field. The divergence of the vector field $\mathbf{u}(\mathbf{x})$ is the scalar field denoted by div $\mathbf{u}$ and defined by

$$\operatorname{div}\mathbf{u} = \operatorname{tr}(\nabla\mathbf{u}). \tag{1.163}$$

By (1.162) and (1.163), in cartesian components,

$$\operatorname{div}\mathbf{u} = \frac{\partial u_i}{\partial x_i}. \tag{1.164}$$

The curl of the vector field $\mathbf{u}(\mathbf{x})$ is the vector field denoted by curl $\mathbf{u}$ which in cartesian components is defined by

$$\operatorname{curl}\mathbf{u} = e_{ijk}\frac{\partial u_k}{\partial x_j}\mathbf{e}_i, \qquad \big(\operatorname{curl}\mathbf{u}\big)_i = e_{ijk}\frac{\partial u_k}{\partial x_j}. \tag{1.165}$$

Divergence and curl of a tensor field. The divergence and curl of the tensor field $\mathbf{T}(\mathbf{x})$ are, respectively, the vector field denoted by div $\mathbf{T}$ and the tensor field denoted by curl $\mathbf{T}$ whose cartesian components are defined by

$$\big(\operatorname{div}\mathbf{T}\big)_i = \frac{\partial T_{ij}}{\partial x_j}, \tag{1.166}$$

$$\big(\operatorname{curl}\mathbf{T}\big)_{ij} = e_{ipq}\frac{\partial T_{jq}}{\partial x_p}. \tag{1.167}$$

Exercise: Show that the divergence of a tensor field $\mathbf{T}(\mathbf{x})$ obeys

$$\big(\operatorname{div}\mathbf{T}\big)\cdot\mathbf{v} = \operatorname{div}\big(\mathbf{T}^T\mathbf{v}\big) \qquad \text{for all constant vectors } \mathbf{v},$$

$$\big(\operatorname{div}\mathbf{T}\big)\cdot\mathbf{v} = \operatorname{div}(\mathbf{T}^T\mathbf{v}) - \mathbf{T}\cdot\operatorname{grad}\mathbf{v} \qquad \text{for all vector fields } \mathbf{v}(\mathbf{x}). \tag{1.168}$$

Note that $\mathbf{T}^T\mathbf{v}$ is a vector and so $\operatorname{div}(\mathbf{T}^T\mathbf{v})$ refers to the divergence of a vector field.

The Laplacian of the scalar field ϕ is denoted by $\nabla^2\phi$ or $\Delta\phi$ and defined as

$$\nabla^2\phi \;=\; \frac{\partial^2\phi}{\partial x_i\partial x_i}\;. \tag{1.169}$$

1.8.2 Divergence theorem.

The *divergence theorem* allows one to relate a surface integral on the boundary $\partial\mathcal{R}$ of a closed region $\mathcal{R}$ to a volume integral over $\mathcal{R}$. In particular, for a scalar field $\phi(\mathbf{x})$, a vector field $\mathbf{u}(\mathbf{x})$ and a tensor field $\mathbf{T}(\mathbf{x})$ one has, with $\mathbf{n}(\mathbf{x})$ being the outward pointing unit vector at points $\mathbf{x}$ on the boundary $\partial\mathcal{R}$,

$$\int_{\partial\mathcal{R}} \phi\mathbf{n}\, dA = \int_{\mathcal{R}} \operatorname{grad}\phi\, dV, \tag{1.170}$$

$$\int_{\partial\mathcal{R}} \mathbf{u}\cdot\mathbf{n}\, dA = \int_{\mathcal{R}} \operatorname{div}\mathbf{u}\, dV, \tag{1.171}$$

$$\int_{\partial\mathcal{R}} \mathbf{T}\mathbf{n}\, dA = \int_{\mathcal{R}} \operatorname{div}\mathbf{T}\, dV, \tag{1.172}$$

or, in terms of components,

$$\int_{\partial\mathcal{R}} \phi n_i\, dA = \int_{\mathcal{R}} \frac{\partial\phi}{\partial x_i}\, dV, \tag{1.173}$$

$$\int_{\partial\mathcal{R}} u_i n_i\, dA = \int_{\mathcal{R}} \frac{\partial u_i}{\partial x_i}\, dV, \tag{1.174}$$

$$\int_{\partial\mathcal{R}} T_{ij} n_j\, dA = \int_{\mathcal{R}} \frac{\partial T_{ij}}{\partial x_j}\, dV. \tag{1.175}$$

For a general field $\mathbb{D}_{i_1 i_2 \ldots i_n}(\mathbf{x})$ (which could be the product of various fields) the divergence theorem gives

$$\int_{\partial\mathcal{R}} \mathbb{T}_{i_1 i_2 \ldots i_n}\, n_k\, dA = \int_{\mathcal{R}} \frac{\partial}{\partial x_k}(\mathbb{T}_{i_1 i_2 \ldots i_n})\, dV. \tag{1.176}$$

1.8.3 Localization.

Let $\mathcal{R}$ be a bounded regular region of three-dimensional space and suppose that the scalar field $\phi(\mathbf{x})$ is defined *and continuous* at all $\mathbf{x} \in \mathcal{R} + \partial\mathcal{R}$. If

$$\int_{\mathcal{D}} \phi(\mathbf{x})\, \mathrm{d}V = 0 \qquad \text{for } \textit{all} \text{ subregions } \mathcal{D} \subset \mathcal{R}, \tag{1.177}$$

then

$$\phi(\mathbf{x}) = 0 \qquad \text{at every point } \mathbf{x} \in \mathcal{R}. \tag{1.178}$$

Since (1.177) holds for all *regions* $\mathcal{D} \subset \mathcal{R}$, it is sometimes said to be a *global* statement, in contrast to (1.178) that holds at each *point* in $\mathcal{R}$ and so is said to be (the associated) *local* statement.

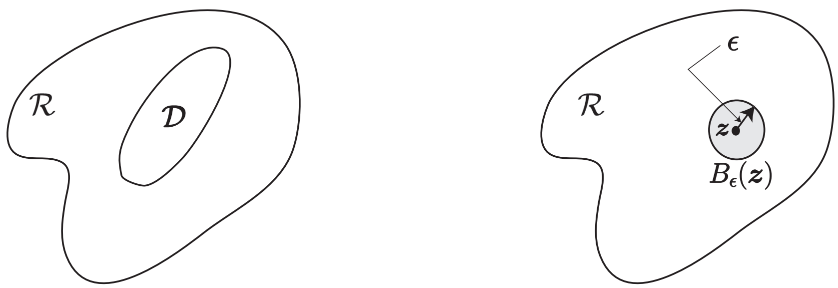


Figure 1.7: The region $\mathcal{R}$, a subregion $\mathcal{D}$, the point $\mathbf{z}$ and a neighborhood $B_\epsilon(\mathbf{z})$ of the point $\mathbf{z}$.

One can prove this by contradiction. Suppose that (1.178) does *not* hold. This implies that there is a point, say $\mathbf{z} \in \mathcal{R}$, at which $\phi(\mathbf{z}) \neq 0$. Suppose that ϕ is positive at this point: $\phi(\mathbf{z}) > 0$. By continuity, ϕ is necessarily (strictly) positive in some neighborhood of $\mathbf{z}$ as well. Let $B_\epsilon(\mathbf{z})$ be a sphere with its center at $\mathbf{z}$ and radius $\epsilon > 0$. We can always choose ϵ sufficiently small (and > 0) so that

$$\phi(\mathbf{x}) > 0 \qquad \text{at all } \mathbf{x} \in B_\epsilon(\mathbf{z});$$

$B_\epsilon(\mathbf{z})$ is a sufficiently small neighborhood of $\mathbf{z}$. Since (1.177) holds for all regions $\mathcal{D}$, we may pick a region $\mathcal{D}$ that is a subset of $B_\epsilon(\mathbf{z})$. Then $\phi(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathcal{D}$. Integrating ϕ over this $\mathcal{D}$ gives

$$\int_{\mathcal{D}} \phi(\mathbf{x}) \, \mathrm{d}V > 0$$

thus contradicting (1.177). An entirely analogous calculation can be carried out in the case $\phi(\mathbf{z}) < 0$. Thus the starting assumption must be false and (1.178) must hold.

1.8.4 Functions of tensors.

– Let $\phi(\mathbf{F})$ be a scalar-valued function defined for all tensors $\mathbf{F} \in \mathsf{Lin}$. Then there is a function $\widehat{\phi}$ such that

$$\phi(\mathbf{F}) = \widehat{\phi}(F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}), \tag{1.179}$$

where the F_{ij}'s are the components of $\mathbf{F}$ in a fixed orthonormal basis. Then $\partial\phi/\partial\mathbf{F}$ denotes the tensor with components

$$\left(\frac{\partial \phi}{\partial \mathbf{F}}\right)_{ij} = \frac{\partial \widehat{\phi}}{\partial F_{ij}}. \tag{1.180}$$

– Let $\phi(\mathbf{C})$ be a scalar-valued function defined for all *symmetric* tensors $\mathbf{C}$. Then $\partial\phi/\partial\mathbf{C}$ must also be symmetric (since ϕ is only defined for symmetric tensors).

Suppose that in terms of components in a fixed basis we have

$$\phi(\mathbf{C}) = \widehat{\phi}(C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{23}, C_{31}, C_{32}, C_{33}). \tag{1.181}$$

In order to ensure that $\partial\phi/\partial\mathbf{C}$ is symmetric when carrying out explicit calculations, one can proceed in either of two ways. The first is to express the function $\widehat{\phi}$ in "symmetric form" such that

$$\widehat{\phi}(C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{23}, C_{31}, C_{32}, C_{33}) = \widehat{\phi}(C_{11}, C_{21}, C_{13}, C_{12}, C_{22}, C_{23}, C_{31}, C_{32}, C_{33})$$

$$\widehat{\phi}(C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{23}, C_{31}, C_{32}, C_{33}) = \widehat{\phi}(C_{11}, C_{12}, C_{31}, C_{21}, C_{22}, C_{23}, C_{13}, C_{32}, C_{33})$$

$$\widehat{\phi}(C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{23}, C_{31}, C_{32}, C_{33}) = \widehat{\phi}(C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{32}, C_{31}, C_{23}, C_{33})$$

One could achieve such a symmetrization of $\widehat{\phi}$ by replacing C_{12} by $\frac{1}{2}(C_{12}+C_{21})$, C_{32} by $\frac{1}{2}(C_{32}+C_{23})$ and so on. Then $\partial\phi/\partial\mathbf{C}$ denotes the *symmetric tensor* with components

$$\left(\frac{\partial\phi}{\partial\mathbf{C}}\right)_{ij} = \frac{\partial\widehat{\phi}}{\partial C_{ij}} = \frac{\partial\widehat{\phi}}{\partial C_{ji}}. \tag{1.182}$$

For example consider the function $\phi(C_{11}, C_{12}, \ldots C_{33}) = C_{12}$. One would not simply differentiate this to get $\partial\phi/\partial C_{12} = 1, \partial\phi/\partial C_{21} = 0$. Instead, one would first symmetrize this by replacing C_{12} by $\frac{1}{2}(C_{12}+C_{21})$ and writing $\phi(C_{11}, C_{12}, \ldots C_{33}) = \frac{1}{2}(C_{12}+C_{21})$, and then differentiating to get $\partial\phi/\partial C_{12} = 1/2, \partial\phi/\partial C_{21} = 1/2$.

The second way is to let $\partial\phi/\partial\mathbf{C}$ be the *symmetric tensor* with components

$$\left(\frac{\partial\phi}{\partial\mathbf{C}}\right)_{ij} = \frac{1}{2}\left(\frac{\partial\phi}{\partial C_{ij}} + \frac{\partial\phi}{\partial C_{ji}}\right). \tag{1.183}$$

In practice it is easier to use (1.183) than to symmetrize the function. See Problem 1.8.6.

1.8.5 Worked examples.

Problem 1.8.1. For any tensor field $\mathbf{A}(\mathbf{x})$ and vector field $\mathbf{u}(\mathbf{x})$ show that

$$\mathbf{u}\cdot \operatorname{div}\mathbf{A} = \operatorname{div}(\mathbf{A}^T\mathbf{u}) - \mathbf{A}\cdot \operatorname{grad}\mathbf{u}. \tag{1.184}$$

Solution:

Since $\mathbf{u}$ and $\operatorname{div}\mathbf{A}$ are vectors, the left-hand side of (1.184) represents their scalar product:

$$\begin{aligned}\mathbf{u}\cdot\operatorname{div}\mathbf{A} = u_i\,(\operatorname{div}\mathbf{A})_i &\overset{(1.166)}{=} u_i\frac{\partial A_{ij}}{\partial x_j} = \\ &= \frac{\partial}{\partial x_j}(A_{ij}u_i) - A_{ij}\frac{\partial u_i}{\partial x_j} \overset{(1.162)}{=} \frac{\partial}{\partial x_j}(\mathbf{A}^T\mathbf{u})_j - A_{ij}\,(\operatorname{grad}\mathbf{u})_{ij} = \\ &\overset{(1.164)}{=} \operatorname{div}(\mathbf{A}^T\mathbf{u}) - \mathbf{A}\cdot\operatorname{grad}\mathbf{u} \qquad \Box\end{aligned}$$

Problem 1.8.2. Consider the vector field $\mathbf{u}(\mathbf{x})$:

$$\mathbf{u}(\mathbf{x}) = \beta\,\frac{\mathbf{x}}{r^3}, \qquad r = |\mathbf{x}| \neq 0, \tag{i}$$

where β is a constant. The tensor field $\mathbf{E}(\mathbf{x})$ is related to $\mathbf{u}(\mathbf{x})$ by

$$\mathbf{E} = \frac{1}{2}\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right), \tag{ii}$$

and the tensor field $\mathbf{S}(\mathbf{x})$ is related to $\mathbf{E}(\mathbf{x})$ by

$$\mathbf{S} = 2\mu\mathbf{E} + \lambda\ \operatorname{tr}(\mathbf{E})\,\mathbf{1}, \tag{iii}$$

where λ and μ are constants. Verify that $\mathbf{S}(\mathbf{x})$ satisfies the differential equation:

$$\operatorname{div}\mathbf{S}(\mathbf{x}) = \mathbf{o}, \qquad |\mathbf{x}| \neq 0. \tag{iv}$$

Solution: We proceed in a straightforward manner by first substituting (i) into (ii) to calculate $\mathbf{E}(\mathbf{x})$; then substituting $\mathbf{E}(\mathbf{x})$ into (iii) to calculate $\mathbf{S}(\mathbf{x})$; and finally checking whether this $\mathbf{S}(\mathbf{x})$ satisfies (iv).

In terms of components, (ii) can be written as

$$E_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{v}$$

and so we need to calculate $\partial u_i/\partial x_j$. Since the expression (i) for u_i involves r, it is convenient to start by calculating $\partial r/\partial x_j$. Observe by differentiating $|\mathbf{x}|^2 = r^2 = x_i\,x_i$ that

$$2r\frac{\partial r}{\partial x_j} = 2\frac{\partial x_i}{\partial x_j}\,x_i = 2\delta_{ij}x_i\ = 2x_j,$$

and therefore

$$\frac{\partial r}{\partial x_j} = \frac{x_j}{r}. \tag{vi}$$

Now differentiating (i), i.e. $u_i = \beta x_i / r^3$, with respect to x_j gives

$$\frac{\partial u_i}{\partial x_j} = \frac{\beta}{r^3}\frac{\partial x_i}{\partial x_j} + \beta x_i\,\frac{\partial(r^{-3})}{\partial x_j} = \beta\,\frac{\delta_{ij}}{r^3} - 3\beta\frac{x_i}{r^4}\,\frac{\partial r}{\partial x_j} \overset{(vi)}{=} \beta\,\frac{\delta_{ij}}{r^3} - 3\beta\,\frac{x_i}{r^4}\,\frac{x_j}{r} = \beta\,\frac{\delta_{ij}}{r^3} - 3\beta\,\frac{x_i\,x_j}{r^5}.$$

Substituting this into (v) gives us E_{ij}:

$$E_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) = \beta\left(\frac{\delta_{ij}}{r^3} - 3\,\frac{x_i\,x_j}{r^5}\right). \qquad (vii)$$

Next, substituting (vii) into (iii) gives S_{ij}:

$$\begin{aligned} S_{ij} &= 2\mu\, E_{ij} + \lambda E_{kk}\delta_{ij} \;=\; 2\mu\beta\left(\frac{\delta_{ij}}{r^3} - \frac{3x_i\,x_j}{r^5}\right) + \lambda\beta\left(\frac{\delta_{kk}}{r^3} - 3\,\frac{x_k\,x_k}{r^5}\right)\delta_{ij} \\ &= 2\mu\beta\left(\frac{\delta_{ij}}{r^3} - \frac{3x_i\,x_j}{r^5}\right) + \lambda\beta\left(\frac{3}{r^3} - 3\,\frac{r^2}{r^5}\right)\delta_{ij} \;=\; 2\mu\beta\left(\frac{\delta_{ij}}{r^3} - \frac{3x_i\,x_j}{r^5}\right). \end{aligned}$$

Finally we use this to calculate $\partial S_{ij}/\partial x_j$

$$\begin{aligned} \frac{1}{2\mu\beta}\,\frac{\partial S_{ij}}{\partial x_j} &= \delta_{ij}\,\frac{\partial}{\partial x_j}\,(r^{-3}) - \frac{3}{r^5}\,\frac{\partial}{\partial x_j}\,(x_i\,x_j) - 3x_i\,x_j\,\frac{\partial}{\partial x_j}(r^{-5}) \\ &= \delta_{ij}\,\left(-\frac{3}{r^4}\frac{\partial r}{\partial x_j}\right) - \frac{3}{r^5}\,(\delta_{ij}\,x_j + x_i\,\delta_{jj}) - 3x_i\,x_j\left(-\frac{5}{r^6}\,\frac{\partial r}{\partial x_j}\right) \\ &\overset{(vi)}{=} -3\,\frac{\delta_{ij}}{r^4}\,\frac{x_j}{r} - \frac{3}{r^5}\,(x_i + 3x_i) + \frac{15x_i\,x_j}{r^6}\,\frac{x_j}{r} \\ &= 0. \end{aligned}$$

Therefore $\partial S_{ij}/\partial x_j = 0$ which establishes the claim.

Problem 1.8.3. Show that

$$\int_{\partial\mathcal{R}} \mathbf{x}\otimes\mathbf{n}\,dA = \mathsf{vol}(R)\,\mathbf{I},$$

where $\mathsf{vol}(\mathcal{R})$ is the volume of the region $\mathcal{R}$.

Solution: In terms of components in a fixed basis, we have to show that

$$\int_{\partial\mathcal{R}} x_i n_j\,dA = \mathsf{vol}(\mathcal{R})\,\delta_{ij}.$$

The result follows immediately by using the divergence theorem:

$$\int_{\partial\mathcal{R}} x_i n_j\;dA \overset{(1.174)}{=} \int_{\mathcal{R}} \frac{\partial x_i}{\partial x_j}\;dV = \int_{\mathcal{R}} \delta_{ij}\;dV \quad = \delta_{ij}\int_{\mathcal{R}} dV \quad = \delta_{ij}\mathsf{vol}(\mathcal{R}).$$

Problem 1.8.4. The scalar-valued functions

$$I_1(\mathbf{C}) = \operatorname{tr}\mathbf{C}, \qquad I_2(\mathbf{C}) = \frac{1}{2}\big[(\operatorname{tr}\mathbf{C})^2 - \operatorname{tr}(\mathbf{C}^2)\big], \qquad I_3(\mathbf{C}) = \det\,\mathbf{C},$$

are defined for all nonsingular symmetric tensors $\mathbf{C}$. Show that

$$\frac{\partial I_1}{\partial \mathbf{C}} = \mathbf{I}, \qquad \frac{\partial I_2}{\partial \mathbf{C}} = I_1\,\mathbf{I} - \mathbf{C} \qquad \frac{\partial I_3}{\partial \mathbf{C}} = I_3\mathbf{C}^{-1}. \tag{1.185}$$

Solution: A direct way in which to establish (1.185) is by working with components and we shall take that approach in Problem 1.8.5. Here we use an alternative (often very convenient) approach.

Consider a one-parameter family of symmetric tensors $\mathbf{C}(t)$ depending smoothly on the parameter t and let $\dot{\mathbf{C}} = \frac{d}{dt}\mathbf{C}$. Since $\operatorname{tr}\mathbf{C} = \mathbf{I}\cdot\mathbf{C}$ we can write $I_1(\mathbf{C}) = \mathbf{I}\cdot\mathbf{C}$. Differentiating both sides of this with respect to t gives

$$\frac{\partial I_1}{\partial \mathbf{C}} \cdot \dot{\mathbf{C}} = \mathbf{I}\cdot\dot{\mathbf{C}} \qquad \Rightarrow \qquad \left(\frac{\partial I_1}{\partial \mathbf{C}} - \mathbf{I}\right)\cdot\dot{\mathbf{C}} = 0.$$

Since this must hold for all $\dot{\mathbf{C}}$ and the term inside the parenthesis is symmetric and does not depend on $\dot{\mathbf{C}}$ it follows[9] that the term in the parenthesis must vanish. Thus

$$\frac{\partial I_1}{\partial \mathbf{C}} = \mathbf{I}. \tag{i}$$

Next consider I_2. Since $\operatorname{tr}\mathbf{C}^2 = \mathbf{C}\cdot\mathbf{C}$ we can write I_2 as

$$I_2(\mathbf{C}) = \frac{1}{2}\big(I_1^2(\mathbf{C}) - \mathbf{C}\cdot\mathbf{C}\big).$$

Differentiating both sides with respect to t gives

$$\frac{\partial I_2}{\partial \mathbf{C}}\cdot\dot{\mathbf{C}} = \frac{1}{2}\left(2I_1\frac{\partial I_1}{\partial \mathbf{C}}\cdot\dot{\mathbf{C}} - \dot{\mathbf{C}}\cdot\mathbf{C} - \mathbf{C}\cdot\dot{\mathbf{C}}\right) = \frac{1}{2}\left(2I_1\frac{\partial I_1}{\partial \mathbf{C}}\cdot\dot{\mathbf{C}} - 2\mathbf{C}\cdot\dot{\mathbf{C}}\right) \overset{(i)}{=} (I_1\,\mathbf{I} - \mathbf{C})\cdot\dot{\mathbf{C}}$$

which leads to

$$\left(\frac{\partial I_2}{\partial \mathbf{C}} - (I_1\,\mathbf{I} - \mathbf{C})\right)\cdot\dot{\mathbf{C}} = 0 \qquad \Rightarrow \qquad \frac{\partial I_2}{\partial \mathbf{C}} = I_1\,\mathbf{I} - \mathbf{C}.$$

Finally consider I_3. Differentiating both sides of $I_3 = \det\mathbf{C}$ with respect to t gives

$$\frac{\partial I_3}{\partial \mathbf{C}}\cdot\dot{\mathbf{C}} = \frac{d}{dt}\det\mathbf{C} \overset{(1.200)}{=} \det\mathbf{C}\;\mathbf{C}^{-T}\cdot\dot{\mathbf{C}} = I_3\;\mathbf{C}^{-1}\cdot\dot{\mathbf{C}} \qquad \Rightarrow \qquad \frac{\partial I_3}{\partial \mathbf{C}} = I_3\;\mathbf{C}^{-1},$$

where we have used the fact that $\mathbf{C}^{-1}$ is a symmetric tensor.

Problem 1.8.5. Given a function $\overline{W}(\mathbf{C})$ defined for all symmetric tensors $\mathbf{C}$, define the function $W(\mathbf{F})$ for all tensors $\mathbf{F} \in \mathsf{Lin}$ by

$$W(\mathbf{F}) = \overline{W}(\mathbf{C}), \qquad \mathbf{C} = \mathbf{F}^T\mathbf{F}. \tag{i}$$

Show that

$$\frac{\partial W}{\partial \mathbf{F}} = 2\mathbf{F}\frac{\partial \overline{W}}{\partial \mathbf{C}}. \tag{ii}$$

[9] This claim needs careful attention, e.g. see Section 3.4 of Gurtin et al. [4]

Solution: In this problem we will work with components. You are encouraged to work this problem using the approach taken in Problem 1.8.4.

Differentiating (i) and using the chain rule,

$$\frac{\partial W}{\partial F_{ij}} = \frac{\partial \overline{W}}{\partial C_{pq}} \frac{\partial C_{pq}}{\partial F_{ij}}. \tag{iii}$$

However in components, $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ reads $C_{pq} = F_{kp}F_{kq}$ and so

$$\frac{\partial C_{pq}}{\partial F_{ij}} = \frac{\partial}{\partial F_{ij}}(F_{kp}F_{kq}) = \frac{\partial F_{kp}}{\partial F_{ij}}F_{kq} + \frac{\partial F_{kq}}{\partial F_{ij}}F_{kp} = \delta_{ki}\delta_{pj}F_{kq} + \delta_{ki}\delta_{qj}F_{kp} = \delta_{pj}F_{iq} + \delta_{qj}F_{ip}. \tag{iv}$$

Substituting (iv) into (iii) and using the substitution rule gives

$$\frac{\partial W}{\partial F_{ij}} = \frac{\partial \overline{W}}{\partial C_{pq}} \frac{\partial C_{pq}}{\partial F_{ij}} = \frac{\partial \overline{W}}{\partial C_{pq}}\Big[\delta_{pj}F_{iq} + \delta_{qj}F_{ip}\Big] = \frac{\partial \overline{W}}{\partial C_{jq}}F_{iq} + \frac{\partial \overline{W}}{\partial C_{pj}}F_{ip} = 2F_{iq}\frac{\partial \overline{W}}{\partial C_{qj}}.$$

and so

$$\frac{\partial W}{\partial \mathbf{F}} = 2\mathbf{F}\frac{\partial \overline{W}}{\partial \mathbf{C}}.$$

Problem 1.8.6. Let $f(\mathbf{C}) = \mathbf{C}^2\mathbf{m}\cdot\mathbf{m}$ be a scalar-valued function defined for all *symmetric* tensors $\mathbf{C}$, $\mathbf{m}$ being a fixed vector. Calculate $\partial f/\partial \mathbf{C}$.

Solution: We will start by showing why one must proceed with care in order to ensure that $\partial f/\partial \mathbf{C}$ is a symmetric tensor.

On writing $f(\mathbf{C})$ in terms of the components C_{ij} and m_i in some fixed cartesian basis:

$$f(\mathbf{C}) = \mathbf{C}^2\mathbf{m}\cdot\mathbf{m} = \mathbf{Cm}\cdot\mathbf{C}^T\mathbf{m} = \mathbf{Cm}\cdot\mathbf{Cm} = (\mathbf{Cm})_i(\mathbf{Cm})_i = C_{ij}m_jC_{ik}m_k = C_{ij}C_{ik}m_jm_k. \tag{i}$$

Therefore

$$\begin{aligned}\frac{\partial f}{\partial C_{pq}} &= \frac{\partial}{\partial C_{pq}}(C_{ij}C_{ik})m_jm_k = \delta_{ip}\delta_{jq}C_{ik}m_jm_k + C_{ij}\delta_{ip}\delta_{kq}m_jm_k = \\ &= C_{pk}m_qm_k + C_{pj}m_jm_q = 2C_{pj}m_qm_j = 2(\mathbf{Cm})_pm_q = 2(\mathbf{Cm}\otimes\mathbf{m})_{pq}\end{aligned}$$

and so we might consider writing

$$\frac{\partial f}{\partial \mathbf{C}} = 2\mathbf{Cm}\otimes\mathbf{m}. \tag{ii}$$

This however would be incorrect since the right-hand side of (ii) is not a symmetric tensor.

To ensure that $\partial f/\partial \mathbf{C}$ is a symmetric tensor we use (1.183) to define $\partial f/\partial \mathbf{C}$ as the symmetric tensor with components

$$\left(\frac{\partial f}{\partial \mathbf{C}}\right)_{pq} = \frac{1}{2}\left(\frac{\partial f}{\partial C_{pq}} + \frac{\partial f}{\partial C_{qp}}\right).$$

Using this on the expression (i) gives

$$\begin{aligned}\frac{\partial f}{\partial C_{pq}} &= \frac{1}{2}\left(\frac{\partial}{\partial C_{pq}}(C_{ij}C_{ik})m_jm_k + \frac{\partial}{\partial C_{qp}}(C_{ij}C_{ik})m_jm_k\right) = \frac{1}{2}(2C_{pj}m_qm_j + 2C_{qj}m_pm_j) = \\ &= C_{pj}m_qm_j + C_{qj}m_pm_j = (\mathbf{Cm})_pm_q + m_p(\mathbf{Cm})_q\end{aligned}$$

which yields

$$\frac{\partial f}{\partial \mathbf{C}} = (\mathbf{C}\mathbf{m}) \otimes \mathbf{m} + \mathbf{m} \otimes (\mathbf{C}\mathbf{m})$$

which is symmetric.

1.8.6 Calculus in orthogonal curvilinear coordinates. An example.

In this section we illustrate working in other orthogonal curvilinear coordinate systems through some examples using cylindrical polar coordinates. A general treatment of orthogonal curvilinear coordinates can be found in Chapter 6 of Volume I.

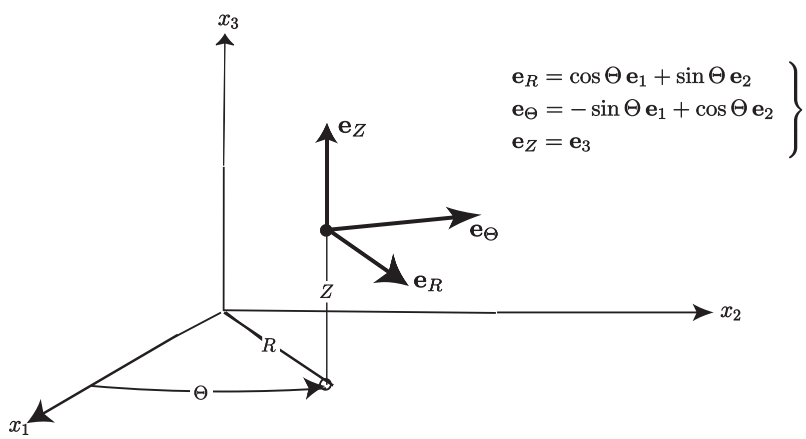


Figure 1.8: Cylindrical polar coordinates (R, Θ, Z) and the associated basis vectors $\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z$.

The rectangular cartesian coordinates (x_1, x_2, x_3) of a point are related to its cylindrical polar coordinates (R, Θ, Z) by

$$x_1 = R\cos\Theta, \quad x_2 = R\sin\Theta, \quad x_3 = Z. \tag{i}$$

As can be seen from Figure 1.8, the basis $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ associated with the cylindrical polar coordinates is related to the fixed rectangular cartesian basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by

$$\mathbf{e}_R(\Theta) = \cos\Theta\ \mathbf{e}_1 + \sin\Theta\ \mathbf{e}_2, \qquad \mathbf{e}_\Theta(\Theta) = -\sin\Theta\ \mathbf{e}_1 + \cos\Theta\ \mathbf{e}_2, \qquad \mathbf{e}_Z = \mathbf{e}_3. \tag{ii}$$

On differentiating the basis vectors $\mathbf{e}_R, \mathbf{e}_\Theta$ and $\mathbf{e}_Z$ with respect to R, Θ and Z we get

$$\begin{gathered}\frac{\partial \mathbf{e}_R}{\partial \Theta} = \mathbf{e}_\Theta, \qquad \frac{\partial \mathbf{e}_\Theta}{\partial \Theta} = -\mathbf{e}_R, \qquad \frac{\partial \mathbf{e}_Z}{\partial \Theta} = 0, \\ \frac{\partial \mathbf{e}_R}{\partial R} = \frac{\partial \mathbf{e}_\Theta}{\partial R} = \frac{\partial \mathbf{e}_Z}{\partial R} = 0, \qquad \frac{\partial \mathbf{e}_R}{\partial Z} = \frac{\partial \mathbf{e}_\Theta}{\partial Z} = \frac{\partial \mathbf{e}_Z}{\partial Z} = 0.\end{gathered} \tag{iii}$$

From Figure 1.8 we see that the position vector $\mathbf{x}$ of a point can be written as

$$\mathbf{x} = \mathbf{x}(R, \Theta, Z) = R\,\mathbf{e}_R + Z\,\mathbf{e}_Z, \tag{iv}$$

and therefore by the chain rule

$$\begin{aligned} d\mathbf{x} &= \frac{\partial \mathbf{x}}{\partial R} dR + \frac{\partial \mathbf{x}}{\partial \Theta} d\Theta + \frac{\partial \mathbf{x}}{\partial Z} dZ = \\ &\overset{(iv)}{=} \frac{\partial}{\partial R}(R\,\mathbf{e}_R + Z\,\mathbf{e}_Z) dR + \frac{\partial}{\partial \Theta}(R\,\mathbf{e}_R + Z\,\mathbf{e}_Z) d\Theta + \frac{\partial}{\partial Z}(R\,\mathbf{e}_R + Z\,\mathbf{e}_Z) dZ = \\ &\overset{(iii)}{=} \mathbf{e}_R\, dR + R\mathbf{e}_\Theta\, d\Theta + \mathbf{e}_Z\, dZ. \end{aligned}$$

It follows by taking the scalar product of this equation with each unit vector $\mathbf{e}_R$, $\mathbf{e}_\Theta$ and $\mathbf{e}_Z$ that

$$dR = \mathbf{e}_R \cdot d\mathbf{x}, \qquad d\Theta = \frac{1}{R}\mathbf{e}_\Theta \cdot d\mathbf{x}, \qquad dZ = \mathbf{e}_Z \cdot d\mathbf{x}. \tag{v}$$

Example: Gradient of a scalar field. Let $\psi(\mathbf{x}) = \psi(R, \Theta, Z)$ be a scalar-valued field. We wish to calculate its gradient, $\nabla\psi$, which we do by relying on the relation $d\psi = \nabla\psi \cdot d\mathbf{x}$ (see (1.159)). Using the chain-rule on $\psi = \psi(R, \Theta, Z)$ gives

$$\begin{aligned} d\psi &= \frac{\partial \psi}{\partial R}\, dR + \frac{\partial \psi}{\partial \Theta}\, d\Theta + \frac{\partial \psi}{\partial Z}\, dZ = \\ &\overset{(v)}{=} \frac{\partial \psi}{\partial R}\, (\mathbf{e}_R \cdot d\mathbf{x}) + \frac{1}{R}\frac{\partial \psi}{\partial \Theta}\, (\mathbf{e}_\Theta \cdot d\mathbf{x}) + \frac{\partial \psi}{\partial Z}\, (\mathbf{e}_Z \cdot d\mathbf{x}) = \\ &= \left(\frac{\partial \psi}{\partial R}\, \mathbf{e}_R + \frac{1}{R}\frac{\partial \psi}{\partial \Theta}\, \mathbf{e}_\Theta + \frac{\partial \psi}{\partial Z}\, \mathbf{e}_Z \right) \cdot d\mathbf{x}, \end{aligned}$$

and therefore, since $d\psi = \nabla\psi \cdot d\mathbf{x}$, we obtain

$$\nabla\psi = \frac{\partial \psi}{\partial R}\, \mathbf{e}_R + \frac{1}{R}\frac{\partial \psi}{\partial \Theta}\, \mathbf{e}_\Theta + \frac{\partial \psi}{\partial Z}\, \mathbf{e}_Z. \tag{1.186}$$

Example: Gradient of a vector field. Let $\mathbf{u}(\mathbf{x})$ be a vector field that can be written in component form as

$$\mathbf{u} = u_R(R, \Theta, Z)\mathbf{e}_R + u_\Theta(R, \Theta, Z)\mathbf{e}_\Theta + u_Z(R, \Theta, Z)\mathbf{e}_Z. \tag{vi}$$

We wish to calculate the gradient of $\mathbf{u}(\mathbf{x})$ in cylindrical polar coordinates which we do by making use of the relation $d\mathbf{u} = \nabla\mathbf{u}\, d\mathbf{x}$; see (1.161). First, from (vi) and the chain rule

$$d\mathbf{u} = \frac{\partial \mathbf{u}}{\partial R} dR + \frac{\partial \mathbf{u}}{\partial \Theta} d\Theta + \frac{\partial \mathbf{u}}{\partial Z} dZ. \tag{vii}$$

Next, we calculate each term on the right-hand side of (vii). For example,

$$
\begin{aligned}
\frac{\partial \mathbf{u}}{\partial \Theta}\, d\Theta \overset{(vi)}{=}& \frac{\partial}{\partial \Theta}(u_R \mathbf{e}_R + u_\Theta \mathbf{e}_\Theta + u_Z \mathbf{e}_Z)\, d\Theta = \\
=& \left(\frac{\partial u_R}{\partial \Theta}\, \mathbf{e}_R + u_R \frac{\partial \mathbf{e}_R}{\partial \Theta} + \frac{\partial u_\Theta}{\partial \Theta}\, \mathbf{e}_\Theta + u_\Theta \frac{\partial \mathbf{e}_\Theta}{\partial \Theta} + \frac{\partial u_Z}{\partial \Theta}\, \mathbf{e}_Z \right) d\Theta = \\
\overset{(iii)}{=}& \left(\frac{\partial u_R}{\partial \Theta}\, \mathbf{e}_R + u_R\, \mathbf{e}_\Theta + \frac{\partial u_\Theta}{\partial \Theta}\, \mathbf{e}_\Theta - u_\Theta\, \mathbf{e}_R + \frac{\partial u_Z}{\partial \Theta}\, \mathbf{e}_Z \right) d\Theta = \\
=& \left(\frac{\partial u_R}{\partial \Theta} - u_\Theta \right) d\Theta\, \mathbf{e}_R + \left(\frac{\partial u_\Theta}{\partial \Theta} + u_R \right) d\Theta\, \mathbf{e}_\Theta + \frac{\partial u_Z}{\partial \Theta}\, d\Theta\, \mathbf{e}_Z = \\
\overset{(v)}{=}& \left(\frac{\partial u_R}{\partial \Theta} - u_\Theta \right) \left(\frac{1}{R}\mathbf{e}_\Theta \cdot d\mathbf{x} \right) \mathbf{e}_R + \left(\frac{\partial u_\Theta}{\partial \Theta} + u_R \right) \left(\frac{1}{R}\mathbf{e}_\Theta \cdot d\mathbf{x} \right) \mathbf{e}_\Theta + \\
& \qquad + \frac{\partial u_Z}{\partial \Theta} \left(\frac{1}{R}\mathbf{e}_\Theta \cdot d\mathbf{x} \right) \mathbf{e}_Z = \\
=& \frac{1}{R} \left(\frac{\partial u_R}{\partial \Theta} - u_\Theta \right) (\mathbf{e}_R \otimes \mathbf{e}_\Theta) d\mathbf{x} + \frac{1}{R} \left(\frac{\partial u_\Theta}{\partial \Theta} + u_R \right) (\mathbf{e}_\Theta \otimes \mathbf{e}_\Theta) d\mathbf{x} + \\
& \qquad + \frac{1}{R}\frac{\partial u_Z}{\partial \Theta}\, (\mathbf{e}_Z \otimes \mathbf{e}_\Theta) d\mathbf{x}.
\end{aligned}
\tag{viii}
$$

Similarly one finds

$$
\frac{\partial \mathbf{u}}{\partial R}\, dR = \frac{\partial u_R}{\partial R}\, (\mathbf{e}_R \otimes \mathbf{e}_R) d\mathbf{x} + \frac{\partial u_\Theta}{\partial R}\, (\mathbf{e}_\Theta \otimes \mathbf{e}_R) d\mathbf{x} + \frac{\partial u_Z}{\partial R}\, (\mathbf{e}_Z \otimes \mathbf{e}_R) d\mathbf{x}. \tag{ix}
$$

$$
\frac{\partial \mathbf{u}}{\partial Z}\, dZ = \frac{\partial u_R}{\partial Z}\, (\mathbf{e}_R \otimes \mathbf{e}_Z) d\mathbf{x} + \frac{\partial u_\Theta}{\partial Z}\, (\mathbf{e}_\Theta \otimes \mathbf{e}_Z) d\mathbf{x} + \frac{\partial u_Z}{\partial Z}\, (\mathbf{e}_Z \otimes \mathbf{e}_Z) d\mathbf{x}. \tag{x}
$$

Therefore combining $(vii), (viii), (ix)$ and (x) yields

$$
\begin{aligned}
d\mathbf{u} =& \frac{\partial u_R}{\partial R}\, (\mathbf{e}_R \otimes \mathbf{e}_R) d\mathbf{x} + \frac{\partial u_\Theta}{\partial R}\, (\mathbf{e}_\Theta \otimes \mathbf{e}_R) d\mathbf{x} + \frac{\partial u_Z}{\partial R}\, (\mathbf{e}_Z \otimes \mathbf{e}_R) d\mathbf{x} + \\
+& \frac{1}{R} \left(\frac{\partial u_R}{\partial \Theta} - u_\Theta \right) (\mathbf{e}_R \otimes \mathbf{e}_\Theta) d\mathbf{x} + \frac{1}{R} \left(\frac{\partial u_\Theta}{\partial \Theta} + u_R \right) (\mathbf{e}_\Theta \otimes \mathbf{e}_\Theta) d\mathbf{x} + \\
& \qquad + \frac{1}{R}\frac{\partial u_Z}{\partial \Theta}\, (\mathbf{e}_Z \otimes \mathbf{e}_\Theta) d\mathbf{x} + \\
+& \frac{\partial u_R}{\partial Z}\, (\mathbf{e}_R \otimes \mathbf{e}_Z) d\mathbf{x} + \frac{\partial u_\Theta}{\partial Z}\, (\mathbf{e}_\Theta \otimes \mathbf{e}_Z) d\mathbf{x} + \frac{\partial u_Z}{\partial Z}\, (\mathbf{e}_Z \otimes \mathbf{e}_Z) d\mathbf{x}.
\end{aligned}
\tag{xi}
$$

Since $d\mathbf{u} = \nabla\mathbf{u}\, d\mathbf{x}$ we can now read off the gradient tensor $\nabla\mathbf{u}$ from (xi) to be

$$\begin{aligned}\nabla\mathbf{u} &= \frac{\partial u_R}{\partial R}(\mathbf{e}_R\otimes\mathbf{e}_R) + \frac{\partial u_\Theta}{\partial R}(\mathbf{e}_\Theta\otimes\mathbf{e}_R) + \frac{\partial u_Z}{\partial R}(\mathbf{e}_Z\otimes\mathbf{e}_R) + \\ &+ \frac{1}{R}\left(\frac{\partial u_R}{\partial\Theta} - u_\Theta\right)(\mathbf{e}_R\otimes\mathbf{e}_\Theta) + \frac{1}{R}\left(\frac{\partial u_\Theta}{\partial\Theta} + u_R\right)(\mathbf{e}_\Theta\otimes\mathbf{e}_\Theta) + \frac{1}{R}\frac{\partial u_Z}{\partial\Theta}(\mathbf{e}_Z\otimes\mathbf{e}_\Theta) + \\ &+ \frac{\partial u_R}{\partial Z}(\mathbf{e}_R\otimes\mathbf{e}_Z) + \frac{\partial u_\Theta}{\partial Z}(\mathbf{e}_\Theta\otimes\mathbf{e}_Z) + \frac{\partial u_Z}{\partial Z}(\mathbf{e}_Z\otimes\mathbf{e}_Z).\end{aligned} \tag{1.187}$$

Example: Divergence of a vector field. The divergence of the vector field $\mathbf{u}(\mathbf{x})$ can be readily found from (1.187) to be

$$\operatorname{div}\mathbf{u} = \operatorname{tr}(\nabla\mathbf{u}) = \frac{\partial u_R}{\partial R} + \frac{u_R}{R} + \frac{1}{R}\frac{\partial u_\Theta}{\partial\Theta} + \frac{\partial u_Z}{\partial Z}. \tag{1.188}$$

Example: Divergence of a tensor field. The divergence of a tensor field, $\operatorname{div}\mathbf{A}(\mathbf{x})$, is a vector field whose three components can be calculated using the identity (1.168) and taking $\mathbf{v} = \mathbf{e}_R, \mathbf{v} = \mathbf{e}_\Theta$ and $\mathbf{v} = \mathbf{e}_Z$ in turn. This calculation is carried out in Section 3.10.1.

1.9 Exercises

1. Matrices and Indicial Notation

Problem 1.1. Evaluate the following expressions:

$$(a)\ e_{ijk}e_{kji}, \qquad (b)\ \delta_{ij}e_{ijk}, \qquad (c)\delta_{ij}\delta_{ik}\delta_{jk}, \qquad (d)\ e_{ik\ell}e_{jk\ell}.$$

Problem 1.2. Derive equations (1.40) and (1.41) involving the determinant of a matrix.

Problem 1.3. From the definitions of δ_{ij} and e_{ijk} show that

$$e_{ijk}e_{pqr} = \det\begin{pmatrix} \delta_{ip} & \delta_{iq} & \delta_{ir} \\ \delta_{jp} & \delta_{jq} & \delta_{jr} \\ \delta_{kp} & \delta_{kq} & \delta_{kr} \end{pmatrix}, \qquad (i)$$

and thus show that

$$e_{ijk}e_{ipq} = \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}. \qquad (ii)$$

Solution:

First, consider the case where any two of the indices i, j, k or p, q, r are the same. Then, on the right-hand side of (i), either two rows or two columns of the matrix coincide and so the determinant vanishes. On the left-hand side of (i), when two of the indices i, j, k or p, q, r are the same, it also vanishes in view of (1.38).

Second, consider the case $(i, j, k) = (p, q, r) = (1, 2, 3)$. Here the right-hand side of (i) equals unity since the matrix is now the identity matrix. The left-hand side also equals unity by (1.38).

Finally, if any two adjacent indices i, j, k or p, q, r are interchanged, the corresponding Levi-Civita symbol on the left-hand side changes sign by (1.42). On the right-hand side, an interchange of two indices results in an interchange of two rows or two columns in the matrix, thus reversing the sign of the determinant.

This establishes (i).

Now set $r = i$ in (i). Then

$$
\begin{aligned}
e_{ijk}e_{pqi} &= \det\begin{pmatrix} \delta_{ip} & \delta_{iq} & \delta_{ii} \\ \delta_{jp} & \delta_{jq} & \delta_{ji} \\ \delta_{kp} & \delta_{kq} & \delta_{ki} \end{pmatrix} \overset{(*)}{=} \det\begin{pmatrix} \delta_{ip} & \delta_{iq} & 3 \\ \delta_{jp} & \delta_{jq} & \delta_{ji} \\ \delta_{kp} & \delta_{kq} & \delta_{ki} \end{pmatrix} = \\
&= \delta_{ip}(\delta_{jq}\delta_{ki} - \delta_{ji}\delta_{kq}) - \delta_{iq}(\delta_{jp}\delta_{ki} - \delta_{ji}\delta_{kp}) + 3(\delta_{jp}\delta_{kq} - \delta_{kp}\delta_{jq}) = \\
&= \delta_{ip}\delta_{jq}\delta_{ki} - \delta_{ip}\delta_{ji}\delta_{kq} - \delta_{iq}\delta_{jp}\delta_{ki} + \delta_{iq}\delta_{ji}\delta_{kp} + 3\delta_{jp}\delta_{kq} - 3\delta_{kp}\delta_{jq} = \\
&\overset{(**)}{=} \delta_{jq}\delta_{kp} - \delta_{jp}\delta_{kq} - \delta_{jp}\delta_{kq} + \delta_{jq}\delta_{kp} + 3\delta_{jp}\delta_{kq} - 3\delta_{kp}\delta_{jq} = \\
&= \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}
\end{aligned}
$$

where step $(*)$ we used $\delta_{ii} = 3$ and in step $(**)$ the substitution rule. (By (1.42) we have $e_{pqi} = -e_{piq} = e_{ipq}$ and so $e_{ijk}e_{pqi} = e_{ijk}e_{ipq}$ and so (ii) is established.)

Problem 1.4. If α_1, α_2 and α_3 are the eigenvalues of a symmetric matrix $[A]$ show that

$$\mathrm{tr}[A] = \alpha_1 + \alpha_2 + \alpha_3, \qquad \det[A] = \alpha_1\alpha_2\alpha_3.$$

Problem 1.5. Let $[F]$ be a nonsingular matrix, $[R]$ an orthogonal matrix and $[U]$ is symmetric matrix such that $[F] = [R][U]$. Show that

(a) $[U]^2 = [F]^T[F]$, and

(b) if $[F]$ is nonsingular, then $[U]^2$ is positive definite.

Problem 1.6. Show that

$$\det\Big([Q]^T[A][Q]\Big) = \det[A], \qquad \mathrm{tr}\Big([Q]^T[A][Q]\Big) = \mathrm{tr}[A],$$

for any orthogonal matrix $[Q]$ and arbitrary matrix $[A]$.

---

2. Vector and tensor algebra.

Problem 1.7. Show that

(a) $\mathbf{u} \times \mathbf{v} = \mathbf{o}$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

(b) $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$ if and only if $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are linearly dependent.

Problem 1.8. Show that

$$(\mathbf{z} \times \mathbf{x}) \times (\mathbf{z} \times \mathbf{y}) = (\mathbf{z} \otimes \mathbf{z})\ (\mathbf{x} \times \mathbf{y}), \qquad (i)$$

$$\mathbf{n} \cdot \big[(\mathbf{n} \times \mathbf{x}) \times (\mathbf{n} \times \mathbf{y})\big] = \mathbf{n} \cdot (\mathbf{x} \times \mathbf{y}), \qquad (ii)$$

where $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are arbitrary vectors and $\mathbf{n}$ is an arbitrary unit vector.

Solution:

(a) The terms $\mathbf{z} \times \mathbf{x}$ and $\mathbf{z} \times \mathbf{y}$ on the left-hand side of (i) each represents a vector and therefore, so does their cross-product $(\mathbf{z} \times \mathbf{x}) \times (\mathbf{z} \times \mathbf{y})$. Its ith component is

$$\begin{aligned}
\Big[(\mathbf{z} \times \mathbf{x}) \times (\mathbf{z} \times \mathbf{y})\Big]_i \ &\overset{(1.61)}{=} e_{ijk}(\mathbf{z} \times \mathbf{x})_j(\mathbf{z} \times \mathbf{y})_k = \\
&\overset{(1.61)}{=} e_{ijk}\left(e_{jpq}z_p x_q\right)\left(e_{krs}z_r y_s\right) \overset{(1.42)}{=} e_{kij}e_{krs}\, e_{jpq}z_p z_r x_q y_s = \\
&\overset{(1.43)}{=} (\delta_{ir}\delta_{js} - \delta_{is}\delta_{jr})\, e_{jpq}z_p z_r x_q y_s = \\
&\overset{(*)}{=} e_{spq}z_p z_i x_q y_s - e_{rpq}z_p z_r x_q y_i \overset{(**)}{=} e_{spq}z_p z_i x_q y_s = \\
&= z_i z_p e_{spq} x_q y_s \overset{(1.42)}{=} z_i z_p e_{pqs} x_q y_s \overset{(1.61)}{=} z_i z_p (\mathbf{x} \times \mathbf{y})_p = \\
&\overset{(1.127)}{=} (\mathbf{z} \otimes \mathbf{z})_{ip}(\mathbf{x} \times \mathbf{y})_p = \Big[(\mathbf{z} \otimes \mathbf{z})(\mathbf{x} \times \mathbf{y})\Big]_i\ .
\end{aligned}$$

In step $(*)$ we used the substitution rule and in step $(**)$ we used (1.45) keeping in mind that $z_p z_r$ is symmetric in p, r and e_{rpq} is skew-symmetric in p, r.

(b) On taking $\mathbf{z} = \mathbf{n}$ in (i) we get

$$(\mathbf{n} \times \mathbf{x}) \times (\mathbf{n} \times \mathbf{y}) = (\mathbf{n} \otimes \mathbf{n})(\mathbf{x} \times \mathbf{y}). \qquad (iii)$$

Each side of this equation is a vector and so we may take its scalar product with $\mathbf{n}$:

$$\mathbf{n} \cdot \Big[(\mathbf{n} \times \mathbf{x}) \times (\mathbf{n} \times \mathbf{y})\Big] = \mathbf{n} \cdot \Big[(\mathbf{n} \otimes \mathbf{n})(\mathbf{x} \times \mathbf{y})\Big]. \qquad (iv)$$

Since the left-hand sides of (iv) and (ii) are identical, it remains to show that the right-hand side of (iv) equals the right-hand side of (ii).

Solution 1: using components in a basis. Simplifying the right-hand side of (iv):

$$\begin{aligned}
\mathbf{n} \cdot \Big[(\mathbf{n} \otimes \mathbf{n})(\mathbf{x} \times \mathbf{y})\Big] \ &\overset{(1.58)}{=} n_i\Big[(\mathbf{n} \otimes \mathbf{n})(\mathbf{x} \times \mathbf{y})\Big]_i \overset{(1.128)}{=} n_i(\mathbf{n} \otimes \mathbf{n})_{ij}(\mathbf{x} \times \mathbf{y})_j \overset{(1.127)}{=} n_i n_i n_j(\mathbf{x} \times \mathbf{y})_j = \\
&= n_j(\mathbf{x} \times \mathbf{y})_j \overset{(1.58)}{=} \mathbf{n} \cdot (\mathbf{x} \times \mathbf{y}),
\end{aligned}$$

which is the right-hand side of (ii). In getting to the second line we used $\mathbf{n}\cdot\mathbf{n} = n_i n_i = 1$.

Solution 2: without using components. Using $(\mathbf{a}\otimes\mathbf{b})\mathbf{c} = (\mathbf{c}\cdot\mathbf{b})\,\mathbf{c}$ with $\mathbf{a}=\mathbf{b}=\mathbf{n}$ and $\mathbf{c}=\mathbf{x}\times\mathbf{y}$,

$$(\mathbf{n}\otimes\mathbf{n})(\mathbf{x}\times\mathbf{y}) = \big[(\mathbf{x}\times\mathbf{y})\cdot\mathbf{n}\big]\,\mathbf{n}, \tag{v}$$

and therefore by taking the scalar product of this equation with $\mathbf{n}$,

$$\mathbf{n}\cdot\Big((\mathbf{n}\otimes\mathbf{n})(\mathbf{x}\times\mathbf{y})\Big) = \big[(\mathbf{x}\times\mathbf{y})\cdot\mathbf{n}\big]\,(\mathbf{n}\cdot\mathbf{n}) = (\mathbf{x}\times\mathbf{y})\cdot\mathbf{n} \tag{vi}$$

since $\mathbf{n}$ is a unit vector. Thus combining (iv) with (vi) establishes (ii).

Problem 1.9. *Reflection in a plane.* Consider a plane $\mathcal{P}$ and let $\mathbf{n}$ be a unit vector normal to it. The operation of reflection in this plane takes a vector $\mathbf{x}$ into the vector $\mathbf{Rx}$. This is illustrated geometrically in Figure 1.9 where $\mathbf{x} = \overrightarrow{OA}$ and $\mathbf{Rx} = \overrightarrow{OC}$.

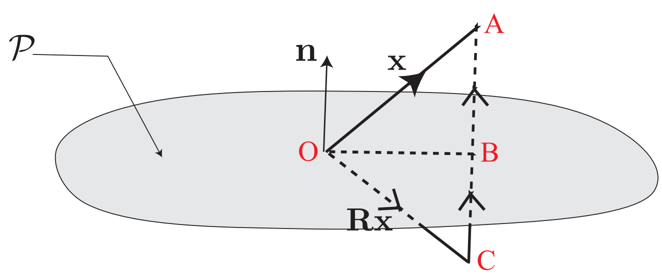


Figure 1.9: The operator $\mathbf{R}$ reflects a vector $\mathbf{x}$ in the plane $\mathcal{P}$.

Determine $\mathbf{R}$ and show that it is precisely the tensor encountered previously in (1.99) of Problem 1.4.1.

Problem 1.10. *Rotation (about an axis) tensor.* (See also Problem 1.4.5.)

(a) The operation of rotation through an angle θ about a unit vector $\mathbf{n}$ takes a vector $\mathbf{x}$ into the vector $\mathbf{Qx}$ as illustrated geometrically in Figure 1.10. Show that

$$\mathbf{Qx} = \cos\theta\;\mathbf{x} + (1-\cos\theta)(\mathbf{n}\cdot\mathbf{x})\mathbf{n} + \sin\theta\;(\mathbf{n}\times\mathbf{x}) \qquad \text{for all } \mathbf{x}\in\mathsf{V}. \tag{1.189}$$

Observe that $\mathbf{Q}$ is a *linear* transformation since $\mathbf{Q}(\alpha\mathbf{x}+\beta\mathbf{y}) = \alpha\mathbf{Qx}+\beta\mathbf{Qy}$.

(b) Show that $\mathbf{Qn} = +\mathbf{n}$. (What geometric transformation does a tensor $\mathbf{Q}$ with the property $\mathbf{Qn} = -\mathbf{n}$ represent? Hint: Consider Problem 1.9.)

(c) Show from (1.189) that the tensor

$$\mathbf{Q} = -\mathbf{I} + 2\mathbf{n}\otimes\mathbf{n}$$

describes a rotation through an angle π about the axis $\mathbf{n}$.

(d) Approximate (1.189) to the case where the angle of rotation is small, $|\theta| \ll 1$.

See also Problem 1.53.

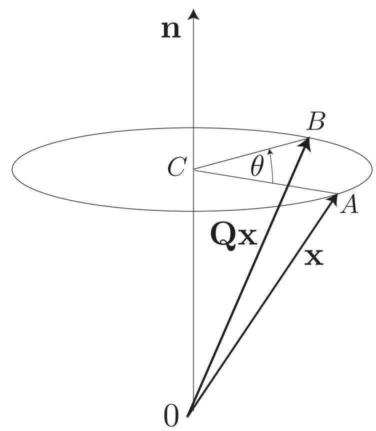


Figure 1.10: The transformation $\mathbf{Q}$ rotates the vector $\mathbf{x} = \vec{OA}$ through an angle θ about the unit vector $\mathbf{n}$ and takes it to $\mathbf{Qx} = \vec{OB}$. (Figure for Problem 1.10.)

Problem 1.11. The non-singular tensor $\mathbf{F}$ maps the three linearly independent vectors $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ into $\{\mathbf{p}, \mathbf{q}, \mathbf{r}\} = \{\mathbf{Fa}, \mathbf{Fb}, \mathbf{Fc}\}$. Show that the volume V_* of the tetrahedron formed by $\mathbf{p}, \mathbf{q}, \mathbf{r}$ is

$$V_* = V_0 \ \det \mathbf{F}, \tag{1.190}$$

where V_0 is the volume of the tetrahedron formed by $\mathbf{a}, \mathbf{b}, \mathbf{c}$. (Hint: see Problem 1.3.3 and equation (1.87)).

Problem 1.12. For a nonsingular tensor $\mathbf{F}$ and arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$ show that

$$\mathbf{Fa} \times \mathbf{Fb} = \det \mathbf{F} \ \mathbf{F}^{-T}(\mathbf{a} \times \mathbf{b}). \tag{1.191}$$

We know from Problem 1.3.1 that the area of the triangle defined by two linearly independent vectors $\mathbf{a}$ and $\mathbf{b}$ is $\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$. The area of its image under the linear transformation $\mathbf{F}$ is therefore $\frac{1}{2}|\mathbf{Fa} \times \mathbf{Fb}|$. Equation 1.191 will be useful when calculating the ratio of these two areas; see Problem 2.43.

Solution:

Without using components in a basis:

$$\det \mathbf{F} \, (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \overset{(1.87)}{=} (\mathbf{Fa} \times \mathbf{Fb}) \cdot \mathbf{Fc} \overset{(1.72)}{=} \mathbf{F}^T(\mathbf{Fa} \times \mathbf{Fb}) \cdot \mathbf{c},$$

and therefore

$$\Big(\det \mathbf{F}\,(\mathbf{a}\times\mathbf{b}) - \mathbf{F}^T(\mathbf{Fa}\times\mathbf{Fb})\Big)\cdot\mathbf{c} = 0 \qquad \text{for all vectors } \mathbf{c}.$$

If $\mathbf{x}\cdot\mathbf{c} = 0$ for all vectors $\mathbf{c}$ it follows that $\mathbf{x} = \mathbf{o}$ and thus here,

$$\mathbf{F}^T(\mathbf{Fa}\times\mathbf{Fb}) = \det \mathbf{F}\,(\mathbf{a}\times\mathbf{b}).$$

When $\mathbf{F}$ is nonsingular this implies

$$\mathbf{Fa}\times\mathbf{Fb} = \det \mathbf{F}\;\mathbf{F}^{-T}(\mathbf{a}\times\mathbf{b}). \qquad \square$$

Problem 1.13. Show that

$$\mathbf{AB}\cdot\mathbf{C} = \mathbf{B}\cdot\mathbf{A}^T\mathbf{C} = \mathbf{A}\cdot\mathbf{CB}^T \tag{1.192}$$

for all tensors $\mathbf{A}, \mathbf{B}$ and $\mathbf{C} \in \mathsf{Lin}$ where the dot between two tensors denotes their scalar product as defined in (1.117).

Solution 1: Without using components:

$$(\mathbf{AB})\cdot\mathbf{C} \overset{(1.117)}{=} \operatorname{tr}\mathbf{ABC}^T \overset{(1.102)}{=} \operatorname{tr}\mathbf{BC}^T\mathbf{A} \overset{(1.74)}{=} \operatorname{tr}\mathbf{B}(\mathbf{A}^T\mathbf{C})^T \overset{(1.117)}{=} \mathbf{B}\cdot(\mathbf{A}^T\mathbf{C}). \qquad \square$$

Similarly

$$(\mathbf{AB})\cdot\mathbf{C} \overset{(1.117)}{=} \operatorname{tr}\mathbf{ABC}^T \overset{(1.74)}{=} \operatorname{tr}\mathbf{A}(\mathbf{CB}^T)^T \overset{(1.117)}{=} \mathbf{A}\cdot(\mathbf{CB}^T) \qquad \square$$

Solution 2: Using components. Since $\mathbf{P}\cdot\mathbf{Q} = P_{ij}Q_{ij}$ for any two tensors $\mathbf{P}$ and $\mathbf{Q}$,

$$\mathbf{AB}\cdot\mathbf{C} = (\mathbf{AB})_{ij}C_{ij} = A_{ik}B_{kj}C_{ij} = B_{kj}A_{ik}C_{ij} = B_{kj}A^T_{ki}C_{ij} = B_{kj}(\mathbf{A}^T\mathbf{C})_{kj} = \mathbf{B}\cdot\mathbf{A}^T\mathbf{C}.$$

Likewise

$$\mathbf{AB}\cdot\mathbf{C} = (\mathbf{AB})_{ij}C_{ij} = A_{ik}B_{kj}C_{ij} = A_{ik}C_{ij}B_{kj} = A_{ik}C_{ij}B^T_{jk} = A_{ik}(\mathbf{CB}^T)_{ik} = \mathbf{A}\cdot\mathbf{CB}^T.$$

Problem 1.14. Consider the scalar-valued function Φ defined for all symmetric tensors $\mathbf{E}$ by

$$\Phi(\mathbf{E}) = \mathbf{CE}\cdot\mathbf{E}; \tag{i}$$

here $\mathbf{C}$ is some constant tensor. Show that "there is no loss of generality in taking $\mathbf{C}$ to be symmetric" (by which we mean that Φ depends only on the symmetric part of $\mathbf{C}$; see (1.80) for what is meant by the symmetric part of a tensor).

Problem 1.15. Show that

$$\frac{\mathbf{x}\cdot(\mathbf{Ay}\times\mathbf{Az}) + \mathbf{Ax}\cdot(\mathbf{y}\times\mathbf{Az}) + \mathbf{Ax}\cdot(\mathbf{Ay}\times\mathbf{z})}{\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z})} = I_2(\mathbf{A}), \tag{1.193}$$

for all linearly independent vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ where $I_2(\mathbf{A})$ is the principal scalar invariant introduced in $(1.105)_2$.

Problem 1.16.

(a) Show for any tensor $\mathbf{A}$ and all scalars λ that

$$\det(\mathbf{A}-\lambda\mathbf{I}) \;=\; -\lambda^3 + I_1(\mathbf{A})\lambda^2 - I_2(\mathbf{A})\lambda + I_3(\mathbf{A}), \tag{i}$$

where

$$I_1(\mathbf{A}) = \mathrm{tr}\,\mathbf{A}, \qquad I_2(\mathbf{A}) = \frac{1}{2}[(\mathrm{tr}\,\mathbf{A})^2 - \mathrm{tr}(\mathbf{A}^2)], \qquad I_3(\mathbf{A}) = \det\,\mathbf{A}. \tag{ii}$$

Note that the identity (i) holds for *all* scalars λ not just the eigenvalues. The eigenvalues are the roots of the cubic equation $\det(\mathbf{A}-\lambda\mathbf{I}) = 0$ – the so-called characteristic equation.

(b) Suppose that $\mathbf{A}$ is symmetric and that $\alpha_1, \alpha_2, \alpha_3$ are its three (necessarily real) eigenvalues. Show that

$$I_1(\mathbf{A}) = \alpha_1+\alpha_2+\alpha_3, \qquad I_2(\mathbf{A}) = \alpha_1\alpha_2+\alpha_2\alpha_3+\alpha_3\alpha_1, \qquad I_3(\mathbf{A}) = \alpha_1\alpha_2\alpha_3. \tag{iii}$$

Solution:

(a) According to (1.87), the determinant of a tensor $\mathbf{B}$ is defined by

$$\det\mathbf{B} := \frac{\mathbf{Bx}\cdot(\mathbf{By}\times\mathbf{Bz})}{\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z})}, \tag{iv}$$

for all linearly independent vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Taking $\mathbf{B} = \mathbf{A}-\lambda\mathbf{I}$,

$$\det(\mathbf{A}-\lambda\mathbf{I}) := \frac{(\mathbf{A}-\lambda\mathbf{I})\mathbf{x}\cdot((\mathbf{A}-\lambda\mathbf{I})\mathbf{y}\times(\mathbf{A}-\lambda\mathbf{I})\mathbf{z})}{\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z})}, \tag{v}$$

on expanding the numerator and using the formulae (1.100) for $I_1(\mathbf{A}) = \mathrm{tr}\,\mathbf{A}$, (1.193) for $I_2(\mathbf{A})$ and (1.87) for $I_3(\mathbf{A}) = \det\,\mathbf{A}$, we get the desired result (i).

(b) Since $\mathbf{A}$ is symmetric it has three real eigenvalues $\alpha_1, \alpha_2, \alpha_3$ and a corresponding set of orthonormal eigenvectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. Consider the orthonormal basis $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ (a principal basis for $\mathbf{A}$). If A_{ij} are the components of $\mathbf{A}$ in this basis then (with the summation convention suspended)

$$A_{ij} \overset{(1.125)}{=} \mathbf{Aa}_j\cdot\mathbf{a}_i = \lambda_j\mathbf{a}_j\cdot\mathbf{a}_i = \lambda_j\delta_{ij}, \tag{vi}$$

having used $\mathbf{Aa}_j = \lambda_j\mathbf{a}_j$. Therefore the off-diagonal terms of $[A]$ vanish and $A_{ii} = \lambda_i$ whence

$$[A] = \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}, \qquad [A]^2 = \begin{pmatrix} \alpha_1^2 & 0 & 0 \\ 0 & \alpha_2^2 & 0 \\ 0 & 0 & \alpha_3^2 \end{pmatrix}. \tag{vii}$$

Thus

$$I_1(\mathbf{A}) = \mathrm{tr}[A] = \alpha_1 + \alpha_2 + \alpha_3, \qquad \mathrm{tr}[A^2] = \alpha_1^2 + \alpha_2^2 + \alpha_3^2, \qquad I_3(\mathbf{A}) = \det[A] = \alpha_1\alpha_2\alpha_3, \qquad \Box$$

and

$$I_2[A] = \frac{1}{2}\left[(\mathrm{tr}\mathbf{A})^2 - \mathrm{tr}(\mathbf{A}^2)\right] = \frac{1}{2}\left[(\alpha_1+\alpha_2+\alpha_3)^2 - (\alpha_1^2+\alpha_2^2+\alpha_3^2)\right] = \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1. \qquad \Box$$

Problem 1.17. *Cayley-Hamilton theorem.* Show that any tensor $\mathbf{A}$ satisfies its own characteristic equation, i.e. show that

$$-\mathbf{A}^3 + I_1(\mathbf{A})\mathbf{A}^2 - I_2(\mathbf{A})\mathbf{A} + I_3(\mathbf{A})\mathbf{I} = 0,$$

where the principal scalar invariants $I_i(\mathbf{A}), i = 1, 2, 3,$ were defined in (1.105).

Problem 1.18. Let

$$\mathbf{A} = \mathbf{u} \otimes \mathbf{v} \qquad (i)$$

for two vectors $\mathbf{u}$ and $\mathbf{v}$.

(a) Show that for all integers $n \geq 2$,
$$\mathbf{A}^n = (\mathbf{u}\cdot\mathbf{v})^{n-1}\mathbf{A}. \qquad (ii)$$

(b) Calculate the principal scalar invariants of $\mathbf{A}$.

(c) Hence or otherwise show that
$$\det(\mathbf{I} + \mathbf{A}) = 1 + \mathbf{u}\cdot\mathbf{v}. \qquad (iii)$$

(d) Determine the eigenvalues of $\mathbf{A}$.

Solution:

(a) We shall establish this by induction. First we show that (ii) holds for $n = 2$. This follows from

$$\mathbf{A}^2 \stackrel{(i)}{=} (\mathbf{u}\otimes\mathbf{v})(\mathbf{u}\otimes\mathbf{v}) \stackrel{(1.70)}{=} (\mathbf{u}\cdot\mathbf{v})(\mathbf{u}\otimes\mathbf{v}) \stackrel{(i)}{=} (\mathbf{u}\cdot\mathbf{v})\mathbf{A}. \qquad (iv)$$

Next, suppose that (ii) holds for some integer $n = N > 2$:

$$\mathbf{A}^N = (\mathbf{u}\cdot\mathbf{v})^{N-1}\mathbf{A}. \qquad (v)$$

Then we show that (ii) holds for $n = N + 1$. This follows from

$$\mathbf{A}^{N+1} = \mathbf{A}^N\mathbf{A} \stackrel{(v)}{=} (\mathbf{u}\cdot\mathbf{v})^{N-1}\mathbf{A}\mathbf{A} = (\mathbf{u}\cdot\mathbf{v})^{N-1}\mathbf{A}^2 \stackrel{(iv)}{=} (\mathbf{u}\cdot\mathbf{v})^N\mathbf{A}.$$

Thus if (ii) holds for $n = N$ it necessarily holds for $n = N + 1$. We know it holds for $n = 2$. It thus follows that it holds for all integers $n \geq 2$.

(b) Observe from (i) that

$$\mathrm{tr}\,\mathbf{A} = A_{ii} = u_i v_i = \mathbf{u}\cdot\mathbf{v}, \tag{vi}$$

and from (ii) that

$$\mathrm{tr}\,\mathbf{A}^n = (\mathbf{u}\cdot\mathbf{v})^{n-1}\mathrm{tr}\,\mathbf{A} \overset{(vi)}{=} (\mathbf{u}\cdot\mathbf{v})^n. \tag{vii}$$

Therefore in particular,

$$\mathrm{tr}\,\mathbf{A}^2 = (\mathbf{u}\cdot\mathbf{v})^2, \qquad \mathrm{tr}\,\mathbf{A}^3 = (\mathbf{u}\cdot\mathbf{v})^3. \tag{$viii$}$$

Thus the first and second principal scalar invariants of $\mathbf{A}$ are

$$I_1(\mathbf{A}) = \mathrm{tr}\,\mathbf{A} = \mathbf{u}\cdot\mathbf{v}, \qquad \Box \tag{ix}$$

$$I_2(\mathbf{A}) = \frac{1}{2}\big[\big(\mathrm{tr}\,\mathbf{A}\big)^2 - \mathrm{tr}\,\mathbf{A}^2\big] = \frac{1}{2}\big[(\mathbf{u}\cdot\mathbf{v})^2 - (\mathbf{u}\cdot\mathbf{v})^2\big] = 0. \qquad \Box \tag{x}$$

There are several ways in which to calculate the third principal scalar invariant $I_3(\mathbf{A}) = \det\,\mathbf{A}$. *Method 1:* According to the Cayley Hamilton theorem, (1.106),

$$\mathbf{A}^3 - I_1(\mathbf{A})\mathbf{A}^2 + I_2(\mathbf{A})\mathbf{A} - I_3(\mathbf{A})\mathbf{I} = \mathbf{O}. \tag{xi}$$

Taking the trace of this equation gives

$$\mathrm{tr}(\mathbf{A}^3) - I_1(\mathbf{A})\mathrm{tr}(\mathbf{A}^2) + I_2(\mathbf{A})\mathrm{tr}(\mathbf{A}) - 3I_3(\mathbf{A}) = 0.$$

Substituting $(viii), (ix)$ and (x) into this gives

$$(\mathbf{u}\cdot\mathbf{v})^3 - (\mathbf{u}\cdot\mathbf{v})(\mathbf{u}\cdot\mathbf{v})^2 + 0 - 3I_3(\mathbf{A}) = 0 \qquad \Rightarrow \qquad I_3(\mathbf{A}) = 0. \qquad \Box \tag{xii}$$

Method 2: Alternatively

$$\begin{aligned} I_3(\mathbf{A}) \quad &\overset{(1.105)_3}{=} \det(\mathbf{A}) \overset{(1.41)}{=} \tfrac{1}{6}e_{ijk}e_{pqr}(\mathbf{A})_{ip}(\mathbf{A})_{jq}(\mathbf{A})_{kr} = \\ &\overset{(i)}{=} \tfrac{1}{6}e_{ijk}e_{pqr}u_iv_pu_jv_qu_kv_r = \tfrac{1}{6}(e_{ijk}u_iu_ju_k)(e_{pqr}v_pv_qv_r). \end{aligned}$$

Next recall from (1.121) that $S_{ij}W_{ij} = 0$ for any symmetric matrix $[S]$ and skew-symmetric matrix $[W]$. Therefore in particular, for any fixed i, since e_{ijk} is skew-symmetric in jk and u_ju_k is symmetric in jk it follows that $e_{ijk}u_ju_k = 0$. Therefore

$$I_3(\mathbf{u}\otimes\mathbf{v}) = 0. \qquad \Box.$$

(c) Setting $\mu = 1$ in the identity (1.104), for any tensor $\mathbf{A}$ one has

$$\det(\mathbf{A}+\mathbf{I}) = 1 + I_1(\mathbf{A}) + I_2(\mathbf{A}) + I_3(\mathbf{A}).$$

Thus for $\mathbf{A} = \mathbf{u}\otimes\mathbf{v}$ we get

$$\det(\mathbf{I}+\mathbf{u}\otimes\mathbf{v}) = 1 + I_1(\mathbf{u}\otimes\mathbf{v}) + I_2(\mathbf{u}\otimes\mathbf{v}) + I_3(\mathbf{u}\otimes\mathbf{v}) \overset{(ix)(x)(xii)}{=} 1 + \mathbf{u}\cdot\mathbf{v}.$$

Thus

$$\det(\mathbf{I}+\mathbf{u}\otimes\mathbf{v}) = 1 + \mathbf{u}\cdot\mathbf{v}. \qquad \Box \tag{1.194}$$

(d) By (1.108) and (1.105), the eigenvalues $\alpha_1, \alpha_2, \alpha_3$ of $\mathbf{A}$ are the roots of the cubic equation

$$\det(\mathbf{A} - \alpha\mathbf{I}) = -\alpha^3 + I_1(\mathbf{A})\alpha^2 - I_2(\mathbf{A})\alpha + I_3(\mathbf{A}) = 0,$$

which, because of (ix), (x) and (xii), simplifies to

$$-\alpha^3 + (\mathbf{u} \cdot \mathbf{v})\alpha^2 = 0.$$

Thus the eigenvalues of $\mathbf{A}$ are

$$\alpha_1 = \alpha_2 = 0, \qquad \alpha_3 = \mathbf{u} \cdot \mathbf{v}.$$

Problem 1.19. Show that an arbitrary tensor $\mathbf{T}$ can be *uniquely* decomposed into the sum,

$$\mathbf{T} = \mathbf{A} + \mathbf{B},$$

of a tensor $\mathbf{A}$ with zero trace and a tensor $\mathbf{B}$ that is a scalar multiple of the identity.

Problem 1.20. You showed in Problem 1.4.7 that $\mathbf{S}\cdot\mathbf{W} = 0$ for all symmetric tensors $\mathbf{S}$ and skew-symmetric tensors $\mathbf{W}$. Hence or otherwise, show that

$$\mathbf{S} \cdot \mathbf{A} = \mathbf{S} \cdot \mathbf{A}^{\text{sym}} \quad \text{for all tensors } \mathbf{A} \text{ and all symmetric tensors } \mathbf{S}, \tag{1.195}$$

where $\mathbf{A}^{\text{sym}}$ is the symmetric part of $\mathbf{A}$, i.e. $\mathbf{A}^{\text{sym}} := \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$.

Problem 1.21. Two symmetric tensors $\mathbf{A}$ and $\mathbf{B}$ are said to be coaxial if they have the same principal directions. Show that $\mathbf{A}$ and $\mathbf{B}$ are coaxial if and only if

$$\mathbf{AB} = \mathbf{BA}. \tag{1.196}$$

Problem 1.22. Show that Lin, the set of all tensors on the vector space V, is nine dimensional.

Solution: Recall from Problem 1.4.3 that $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$, are nine orthonormal tensors in Lin (where as usual $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is an orthonormal basis for V). Since they are orthonormal, they are necessarily linearly independent. In order to show that Lin is nine dimensional, it is therefore sufficient to demonstrate that any tensor $\mathbf{A} \in \mathsf{Lin}$ can be expressed as a linear combination of the nine tensors $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$. To show this, define nine scalars A_{ij} by

$$A_{ij} = (\mathbf{A}\mathbf{e}_j) \cdot \mathbf{e}_i. \tag{i}$$

Observe that this says that the ith component of the vector $\mathbf{A}\mathbf{e}_j$ is A_{ij} which can be equivalently stated as

$$\mathbf{A}\mathbf{e}_j = A_{ij}\mathbf{e}_i. \tag{ii}$$

Then for an arbitrary vector $\mathbf{x} \in \mathsf{V}$ we have

$$\mathbf{A}\mathbf{x} = \mathbf{A}(x_j \mathbf{e}_j) = x_j \mathbf{A}\mathbf{e}_j \stackrel{(ii)}{=} x_j A_{ij} \mathbf{e}_i = A_{ij} x_j \mathbf{e}_i = A_{ij}(\mathbf{x} \cdot \mathbf{e}_j)\mathbf{e}_i = A_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)\mathbf{x}.$$

Since this holds for all vectors $\mathbf{x}$ it follows that

$$\mathbf{A} = A_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j), \tag{1.197}$$

and so an arbitrary tensor $\mathbf{A} \in \mathsf{Lin}$ can be represented as a linear combination of the nine orthonormal tensors $\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, 2, 3$. Thus Lin is nine dimensional (and the nine aforementioned tensors form an orthonormal basis for it.)

Problem 1.23. (*Square root of a symmetric positive definite tensor*) Consider a symmetric positive definite tensor $\mathbf{S}$. Show that it has a *unique* symmetric positive definite square root, i.e. show that there is a unique symmetric positive definite tensor $\mathbf{T}$ for which $\mathbf{T}^2 = \mathbf{S}$.

Solution: In this solution we shall suspend the summation convention on repeated indices and instead show all summations explicitly. Observe that we will have the same subscript appearing 3 times in some equations below.

Since $\mathbf{S}$ is symmetric it has three real eigenvalues $\sigma_1, \sigma_2, \sigma_3$ with corresponding eigenvectors $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ which may be taken to be orthonormal. Furthermore, we know that $\mathbf{S}$ can be represented in its so-called spectral representation

$$\mathbf{S} = \sum_{i=1}^{3} \sigma_i (\mathbf{s}_i \otimes \mathbf{s}_i);$$

see the discussion surrounding (1.109). Since $\mathbf{S}$ is positive definite its eigenvalues are all positive. Hence one can define a linear transformation $\mathbf{T}$ by

$$\mathbf{T} = \sum_{i=1}^{3} \sqrt{\sigma_i} (\mathbf{s}_i \otimes \mathbf{s}_i),$$

and readily verify that $\mathbf{T}$ is symmetric, positive definite and that $\mathbf{T}^2 = \mathbf{S}$. This establishes the existence of a symmetric positive definite square-root of $\mathbf{S}$. What remains is to show uniqueness of this square-root.

Suppose that $\mathbf{S}$ has two symmetric positive definite square roots $\mathbf{T}_1$ and $\mathbf{T}_2 : \mathbf{S} = \mathbf{T}_1^2 = \mathbf{T}_2^2$. Let $\sigma > 0$ and $\mathbf{s}$ be an eigenvalue and corresponding eigenvector of $\mathbf{S}$. Then $\mathbf{S}\mathbf{s} = \sigma\mathbf{s}$ and so $\mathbf{T}_1^2\mathbf{s} = \sigma\mathbf{s}$. Thus we have

$$\mathbf{T}_1^2\mathbf{s} - \sigma\mathbf{s} = (\mathbf{T}_1^2 - \sigma\mathbf{I})\mathbf{s} = (\mathbf{T}_1 + \sqrt{\sigma}\mathbf{I})(\mathbf{T}_1 - \sqrt{\sigma}\mathbf{I})\mathbf{s} = \mathbf{0}\ .$$

If we set $\mathbf{f} = (\mathbf{T}_1 - \sqrt{\sigma}\mathbf{I})\mathbf{s}$ this can be written as

$$\mathbf{T}_1\mathbf{f} = -\sqrt{\sigma}\mathbf{f}\ .$$

Thus either $\mathbf{f}$ is an eigenvector of $\mathbf{T}_1$ corresponding to the eigenvalue $-\sqrt{\sigma}$ (< 0) or $\mathbf{f} = \mathbf{o}$. Since $\mathbf{T}_1$ is positive definite it cannot have a negative eigenvalue. Thus $\mathbf{f} = \mathbf{o}$ and so

$$\mathbf{T}_1\mathbf{s} = \sqrt{\sigma}\mathbf{s}\ .$$

A similar calculation shows that $\mathbf{T}_2\mathbf{s} = \sqrt{\sigma}\mathbf{s}$. Thus

$$\mathbf{T}_1\mathbf{s} = \mathbf{T}_2\mathbf{s}.$$

This holds for *every* eigenvector $\mathbf{s}$ of $\mathbf{S}$: i.e. $\mathbf{T}_1\mathbf{s}_i = \mathbf{T}_2\mathbf{s}_i, \quad i = 1, 2, 3$. Since the triplet of eigenvectors form a basis for the underlying vector space this implies that $\mathbf{T}_1\mathbf{x} = \mathbf{T}_2\mathbf{x}$ for any vector $\mathbf{x}$. Thus $\mathbf{T}_1 = \mathbf{T}_2$.

Problem 1.24. *The summation convention is suspended in this problem and all summations are shown explicitly.* The polar decomposition theorem (1.114) states that any nonsingular linear transformation $\mathbf{F}$ can be represented uniquely in the forms $\mathbf{F} = \mathbf{RU} = \mathbf{VR}$ where $\mathbf{R}$ is orthogonal and $\mathbf{U}$ and $\mathbf{V}$ are symmetric and positive definite. Let $\lambda_i, \mathbf{u}_i, \ i = 1, 2, 3$ be the eigenvalues and eigenvectors of $\mathbf{U}$. As stated just below (1.114), the eigenvalues of $\mathbf{V}$ are the same as those of $\mathbf{U}$ and the corresponding eigenvectors $\mathbf{v}_i$ of $\mathbf{V}$ are given by $\mathbf{v}_i = \mathbf{R}\mathbf{u}_i$. [Exercise: show this.] Thus $\mathbf{U}$ and $\mathbf{V}$ have the spectral decompositions

$$\mathbf{U} = \sum_{i=1}^{3} \lambda_i \mathbf{u}_i \otimes \mathbf{u}_i \ , \qquad \mathbf{V} = \sum_{i=1}^{3} \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i \ .$$

Show that

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \mathbf{v}_i \otimes \mathbf{u}_i \ , \qquad \mathbf{R} = \sum_{i=1}^{3} \mathbf{v}_i \otimes \mathbf{u}_i \ .$$

Problem 1.25. (The summation convention is suspended in this problem.) Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be two orthonormal sets of vectors. Suppose that a tensor $\mathbf{F}$ has the representation

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \mathbf{v}_i \otimes \mathbf{u}_i.$$

If all three λ_i's are non-zero, show that

$$\mathbf{F}^{-1} = \sum_{i=1}^{3} \lambda_i^{-1} \mathbf{u}_i \otimes \mathbf{v}_i.$$

---

3. Change of basis.

Problem 1.26. (See also Problem 1.27.) Consider two bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ and define the scalars Q_{ij} in the usual way by

$$Q_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j. \qquad (i)$$

Show that

$$\mathbf{e}_i = Q_{ji}\mathbf{e}'_j, \qquad \mathbf{e}'_i = Q_{ij}\mathbf{e}_j. \qquad (ii)$$

Let $\mathbf{Q}$ be the tensor whose components in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are Q_{ij}. Show that

$$\mathbf{e}'_i = \mathbf{Q}^T\mathbf{e}_i. \qquad (iii)$$

Therefore $\mathbf{Q}^T$ is the orthogonal transformation that carries the first basis into the second (not $\mathbf{Q}$).

Problem 1.27. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ be two orthonormal bases. Define the tensor $\mathbf{R}$ by

$$\mathbf{R} = \mathbf{e}'_i \otimes \mathbf{e}_i = \mathbf{e}'_1 \otimes \mathbf{e}_1 + \mathbf{e}'_2 \otimes \mathbf{e}_2 + \mathbf{e}'_3 \otimes \mathbf{e}_3. \tag{i}$$

(a) Show that $\mathbf{R}$ is an orthogonal tensor.

(b) Show that $\mathbf{R}$ takes the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ into the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$, i.e. show that

$$\mathbf{e}'_i = \mathbf{R}\mathbf{e}_i. \tag{ii}$$

(c) For any tensor $\mathbf{A}$, let $[A]$ and $[A']$ be its respective matrices of components in the bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$. Show that

$$[A'] = [R]^T[A][R], \tag{iii}$$

where $[R]$ is the matrix of components of $\mathbf{R}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

Remark: Note that $\mathbf{R}$ is the transpose of the tensor $\mathbf{Q}$ in Problem 1.26.

Problem 1.28. Suppose that the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is obtained by rotating the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ through an angle θ about the unit vector $\mathbf{e}_3$; see Figure 1.11. Write out the transformation rule (1.151) explicitly for a 2-tensor $\mathbf{A}$ whose matrix of components in $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is

$$[A] = \begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Problem 1.29. Let $[A]$ and $[A']$ be the components of a 2-tensor $\mathbf{A}$ in two bases. Show that the two matrices $[A]$ and $[A']$ have the same eigenvalues.

Hint: The eigenvalues of a matrix are the roots of the characteristic equation. For matrices $[A]$ and $[A']$ the respective characteristic equations are

$$\det([A] - \lambda[I]) = 0 \qquad \text{and} \qquad \det([A'] - \lambda[I]) = 0.$$

To show that these two matrices have the same eigenvalues it is sufficient to show that the two characteristic equations are identical. When the characteristic equations are written out each has the form given in (1.108). Therefore we should aim to show that

$$I_1([A]) = I_1([A']), \qquad I_2([A]) = I_2([A']), \qquad I_3([A]) = I_3([A']).$$

For this we must show that $\mathsf{trace}[A] = \mathsf{trace}[A']$, $\mathsf{trace}([A]^2) = \mathsf{trace}([A']^2)$ and $\mathsf{det}[A] = \mathsf{det}[A']$.

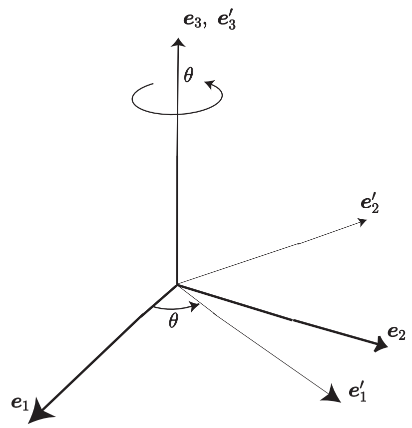


Figure 1.11: A basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ obtained by rotating the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ through an angle θ about the unit vector $\mathbf{e}_3$.

Problem 1.30. (See also Problem 1.6.5.) A tensor $\mathbf{F}$ has the representation

$$\mathbf{F} = \Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j, \tag{i}$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ are two right-handed orthonormal bases.

(a) Show that $\det \mathbf{F} = \det[\Phi]$ where $[\Phi]$ is the matrix whose i,j-element is Φ_{ij}.

(b) Suppose $\mathbf{F}$ is nonsingular. Since $(\mathbf{e}_i \otimes \mathbf{e}'_j)(\mathbf{e}'_i \otimes \mathbf{e}_j) = \mathbf{e}_i \otimes \mathbf{e}_i = \mathbf{I}$, the tensor $\mathbf{F}^{-1}$ has the representation

$$\mathbf{F}^{-1} = \Psi_{ij}\mathbf{e}_i \otimes \mathbf{e}'_j. \tag{ii}$$

Show that $[\Psi] = [\Phi]^{-1}$ where $[\Psi]$ is the matrix whose i,j-element is Ψ_{ij}.

Solution:

(a) From (i) we obtain $\mathbf{F}\mathbf{e}_k = \Phi_{ij}(\mathbf{e}'_i \otimes \mathbf{e}_j)\mathbf{e}_k = \Phi_{ij}(\mathbf{e}_j \cdot \mathbf{e}_k)\mathbf{e}'_i = \Phi_{ij}\delta_{jk}\mathbf{e}'_i$ and so

$$\mathbf{F}\mathbf{e}_k = \Phi_{ik}\mathbf{e}'_i. \tag{iii}$$

By taking $\mathbf{x} = \mathbf{e}_1, \mathbf{y} = \mathbf{e}_2, \mathbf{z} = \mathbf{e}_3$ in (1.87) we get

$$\begin{aligned}
\det \mathbf{F} &= \mathbf{F}\mathbf{e}_1 \cdot (\mathbf{F}\mathbf{e}_2 \times \mathbf{F}\mathbf{e}_3) = \\
&\overset{(iii)}{=} (\Phi_{i1}\mathbf{e}'_i) \cdot \Big[(\Phi_{j2}\mathbf{e}'_j) \times (\Phi_{k3}\mathbf{e}'_k)\Big] = \Phi_{i1}\Phi_{j2}\Phi_{k3}\,\mathbf{e}'_i \cdot (\mathbf{e}'_j \times \mathbf{e}'_k) = \\
&\overset{(1.54)}{=} \Phi_{i1}\Phi_{j2}\Phi_{k3}e_{jki} = e_{ijk}\Phi_{i1}\Phi_{j2}\Phi_{k3} \overset{(1.39)}{=} \det[\Phi]
\end{aligned}$$

□

Alternatively, from Problem 1.6.5, we have $[F] = [Q]^T[\Phi]$ whence $\det \mathbf{F} = \det[F] = \det[\Phi]$ where we have used the fact that $\det[Q] = +1$ which is a consequence of $[Q]$ being proper orthogonal since the two bases are both right-handed and orthonormal.

(b) We can assume $\mathbf{F}^{-1}$ to have the form (ii) and need to show that $[\Psi] = [\Phi]^{-1}$. From (i) and (ii)

$$\mathbf{F}\mathbf{F}^{-1} = (\Phi_{ij}\mathbf{e}'_i \otimes \mathbf{e}_j)(\Psi_{k\ell}\mathbf{e}_k \otimes \mathbf{e}'_\ell) = \Phi_{ij}\Psi_{k\ell}(\mathbf{e}_j \cdot \mathbf{e}_k)(\mathbf{e}'_i \otimes \mathbf{e}'_\ell) = \Phi_{ij}\Psi_{k\ell}\delta_{jk}(\mathbf{e}'_i \otimes \mathbf{e}'_\ell) = \Phi_{ik}\Psi_{k\ell}(\mathbf{e}'_i \otimes \mathbf{e}'_\ell)$$

Since this equals the identity, which we can write as $\mathbf{I} = \mathbf{e}'_i \otimes \mathbf{e}'_i = \delta_{i\ell}\,\mathbf{e}'_i \otimes \mathbf{e}'_\ell$, it follows that

$$\Phi_{ik}\Psi_{k\ell} = \delta_{i\ell}, \qquad [\Phi][\Psi] = [I]. \qquad \square$$

---

4. Invariant (isotropic) functions.

Problem 1.31.

(a) A scalar-valued function $\phi(\mathbf{x}) : \mathsf{V} \to \mathbb{R}$ is said to be isotropic if $\phi(\mathbf{x}) = \phi(\mathbf{Q}\mathbf{x})$ for all orthogonal $\mathbf{Q}$. Show that $\phi(\mathbf{x})$ is isotropic if and only if there is a function $\widehat{\phi}$ such that

$$\phi(\mathbf{x}) = \widehat{\phi}(|\mathbf{x}|) \qquad \text{for all } \mathbf{x} \in \mathsf{V}.$$

(Here $\mathbb{R}$ is the set of all real numbers.)

(b) A vector-valued function $\mathbf{u}(\mathbf{x}) : \mathsf{V} \to \mathsf{V}$ is said to be isotropic if $\mathbf{Q}\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{Q}\mathbf{x})$ for all orthogonal $\mathbf{Q}$. Show that $\mathbf{u}(\mathbf{x})$ is isotropic if and only if there is a function $\widehat{\phi}$ such that

$$\mathbf{u}(\mathbf{x}) = \widehat{\phi}(|\mathbf{x}|)\,\mathbf{x} \qquad \text{for all } \mathbf{x} \in \mathsf{V}.$$

Problem 1.32. Suppose that the principal scalar invariants of two symmetric tensors $\mathbf{B}$ and $\mathbf{C}$ have the same values:

$$I_1(\mathbf{B}) = I_1(\mathbf{C}), \qquad I_2(\mathbf{B}) = I_2(\mathbf{C}), \qquad I_3(\mathbf{B}) = I_3(\mathbf{C}). \qquad (i)$$

Show that there necessarily exists an orthogonal tensor $\mathbf{Q}$ such that

$$\mathbf{B} = \mathbf{Q}\mathbf{C}\mathbf{Q}^T. \qquad (ii)$$

Problem 1.33. Let $\phi(\mathbf{A})$ be an isotropic scalar-valued function defined for all symmetric tensors $\mathbf{A}$. Show that there exists a function $\widehat{\phi}$ such that

$$\phi(\mathbf{A}) = \widehat{\phi}(I_1(\mathbf{A}), I_2(\mathbf{A}), I_3(\mathbf{A}))$$

where the I_i's are the principal scalar invariant functions defined in (1.105).

Solution:

Let $\phi(\mathbf{A})$ be an isotropic scalar-valued function:

$$\phi(\mathbf{A}) = \phi(\mathbf{QAQ}^T) \qquad (i)$$

for all symmetric tensors $\mathbf{A}$ and orthogonal tensors $\mathbf{Q}$. It is sufficient for us to show that

$$\phi(\mathbf{B}) = \phi(\mathbf{C}) \qquad (ii)$$

whenever

$$I_1(\mathbf{B}) = I_1(\mathbf{C}), \qquad I_2(\mathbf{B}) = I_2(\mathbf{C}), \qquad I_3(\mathbf{B}) = I_3(\mathbf{C}). \qquad (iii)$$

From the result in Problem 1.32, whenever (iii) holds there is an orthogonal tensor $\mathbf{Q}$ such that

$$\mathbf{B} = \mathbf{QCQ}^T. \qquad (iv)$$

Therefore $\phi(\mathbf{B}) \overset{(iv)}{=} \phi(\mathbf{QCQ}^T) \overset{(i)}{=} \phi(\mathbf{C})$ and thus (ii) holds.

Problem 1.34. Let $\phi(\mathbf{A})$ be the function defined for all symmetric tensors $\mathbf{A}$ by

$$\phi(\mathbf{A}) = \frac{1}{2}\mathbb{C}_{ijk\ell}A_{ij}A_{k\ell}. \qquad (i)$$

The components here have been taken with respect to a fixed basis and $\mathbb{C}$ is a constant 4-tensor. If ϕ is isotropic, find the most general form of ϕ and also of $\mathbb{C}$.

Solution:

We know from the general representation (1.142) of an isotropic function that $\phi(\mathbf{A})$ can be written as a function of the three principal invariants: $\phi(\mathbf{A}) = \widehat{\phi}(I_1(\mathbf{a}), I_2(\mathbf{A}), I_3(\mathbf{a}))$. Observe that (i) is the most general *quadratic* function of $\mathbf{A}$, and note from (1.105) that $I_1(\mathbf{A})$ is a linear function of $\mathbf{A}$, $I_2(\mathbf{A})$ is a quadratic function of $\mathbf{A}$, and $I_3(\mathbf{A})$ is a cubic function of $\mathbf{A}$. It therefore follows that the most general isotropic quadratic function can be written as

$$\phi(\mathbf{A}) = c_1\big(I_1(\mathbf{A})\big)^2 + c_2 I_2(\mathbf{A}) \qquad (ii)$$

for two constants c_1 and c_2. Since

$$I_1(\mathbf{A}) = \mathrm{tr}(\mathbf{A}) = A_{ii}, \qquad I_2(\mathbf{A}) = \frac{1}{2}[(\mathrm{tr}(\mathbf{A}))^2 - \mathrm{tr}(\mathbf{A}^2)] = \frac{1}{2}\Big(A_{ii}A_{jj} - A_{ik}A_{ki}\Big), \qquad (iii)$$

(ii) can be written as

$$\phi(\mathbf{A}) = c_1 A_{ii}A_{jj} + \frac{1}{2}c_2\Big(A_{ii}A_{jj} - A_{ik}A_{ki}\Big),$$

which leads to

$$\phi(\mathbf{A}) = c_3 A_{ii}A_{jj} + c_4 A_{ik}A_{ki} = c_3\big(\mathrm{tr}\mathbf{A}\big)^2 + c_4\mathrm{tr}\big(\mathbf{A}^2\big) \qquad \square \qquad (iv)$$

for two other constants c_3 and c_4. This is the most general isotropic function of the form (i).

Observe that

$$A_{ii}A_{jj} = \delta_{pq}\delta_{rs}A_{pq}A_{rs}, \qquad A_{ij}A_{ij} = \frac{1}{2}(\delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr})A_{pq}A_{rs}$$

and so we can write (iv) as

$$\phi(\mathbf{A}) = c_3\delta_{pq}\delta_{rs}A_{pq}A_{rs} + \frac{1}{2}c_4(\delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr})A_{pq}A_{rs} = \frac{1}{2}\Big[2c_3\delta_{pq}\delta_{rs} + c_4(\delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr})\Big]A_{pq}A_{rs}$$

which is of the form (i) with

$$\mathbb{C}_{pqrs} = \frac{1}{2}\Big[2c_3\delta_{pq}\delta_{rs} + c_4(\delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr})\Big]. \qquad \square$$

We will encounter this 4-tensor when studying isotropic linear elastic materials.

Problem 1.35. A symmetric tensor $\mathbf{B}$ that has the property

$$\mathbf{QBQ}^T = \mathbf{B} \qquad \text{for all orthogonal tensors } \mathbf{Q} \tag{1.198}$$

is said to be *isotropic*. Show that $\mathbf{B}$ is isotropic if and only if it has the representation $\mathbf{B} = \beta\mathbf{I}$ for some scalar β.

Problem 1.36. Let $\mathbf{F}(\mathbf{B})$ be a symmetric tensor-valued function that is defined for all symmetric tensors $\mathbf{B}$. Such a function is said to be *isotropic* if $\mathbf{F}(\mathbf{QBQ}^T) = \mathbf{QF}(\mathbf{B})\mathbf{Q}^T$ for all orthogonal tensors $\mathbf{Q}$. Show that $\mathbf{F}(\mathbf{B})$ is isotropic if and only if it has the representation

$$\widehat{\mathbf{F}}(\mathbf{B}) = \beta_2\mathbf{B}^2 + \beta_1\mathbf{B} + \beta_0\mathbf{I}, \tag{1.199}$$

where the β_j's are functions of the principal scalar invariants of $\mathbf{B}$.

---

5. Calculus.

Problem 1.37. Let $\mathbf{F}(t)$ be a one-parameter family of nonsingular tensors that depend smoothly on the parameter t. Show that

(a)
$$\frac{d}{dt}(\det\mathbf{F}) = \det\mathbf{F}\,\mathrm{tr}(\dot{\mathbf{F}}\mathbf{F}^{-1}) = \mathrm{J}\,\mathbf{F}^{-T}\cdot\dot{\mathbf{F}}, \tag{1.200}$$

where $J = \det\mathbf{F}$ and $\dot{\mathbf{F}} = d\mathbf{F}/dt$.

(b) Show also that
$$\frac{d}{dt}\big(\mathbf{F}^{-1}\big) = -\mathbf{F}^{-1}\dot{\mathbf{F}}\mathbf{F}^{-1}. \tag{1.201}$$

See also Problem 1.45.

Problem 1.38. Let ϕ, $\mathbf{v}$, and $\mathbf{w}$, be a scalar field and two vector fields, respectively. Show that:

(a) $\mathsf{div}(\phi\mathbf{v}) = \phi\,\mathsf{div}\mathbf{v} + \mathbf{v}\cdot\mathsf{grad}\phi$,

(b) $\mathsf{grad}(\phi\mathbf{v}) = \phi\,\mathsf{grad}\mathbf{v} + \mathbf{v}\otimes\mathsf{grad}\phi$,

(c) $\mathsf{grad}(\mathbf{v}\cdot\mathbf{w}) = (\boldsymbol{\nabla}\mathbf{w})^T\mathbf{v} + (\boldsymbol{\nabla}\mathbf{v})^T\mathbf{w}$,

(d) $\mathsf{div}(\mathbf{v}\otimes\mathbf{w}) = (\mathsf{div}\,\mathbf{w})\mathbf{v} + (\mathsf{div}\,\mathbf{v})\mathbf{w}$.

Problem 1.39. Calculate the gradient of a scalar field and the gradient and divergence of a vector field in spherical polar coordinates (R, Θ, Φ) with associated basis vectors $\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi$ as defined by equations (2.80) and (2.81) on page 149.

Problem 1.40. (*Localization*) "Localization" refers to deriving local equations (i.e. equations that hold at each point in a region $\mathcal{R}$) from a global statement in integral form. In Section 1.8.3 we encountered one circumstance in which localization is possible. Here we look at a second.

Let $\phi(\mathbf{x})$ be a continuous scalar-valued function defined for all $\mathbf{x}\in\mathcal{R}$. Suppose that

$$\int_{\mathcal{R}} \phi(\mathbf{x})\psi(\mathbf{x})\;dV = 0$$

for *all* continuous functions $\psi(\mathbf{x})$ defined on $\mathcal{R}$. Show that this implies $\phi(\mathbf{x}) = 0$ at all points in $\mathcal{R}$.

Remark: In Section 1.8.3 the global statement held for *all* subregions of $\mathcal{R}$; in contrast here, we have a single integral statement that hold on $\mathcal{R}$. On the other hand here, the integrand involves an arbitrary function ψ while such a function was absent in Section 1.8.3.

Problem 1.41. Let $\mathbf{A}(\mathbf{x})$ and $\mathbf{b}(\mathbf{x})$ be smooth tensor and vector fields respectively defined on some regular region $\mathcal{R}$. Suppose that

$$\int_{\partial\mathcal{R}} \mathbf{A}\mathbf{n}\cdot\mathbf{w}\,dA + \int_{\mathcal{R}} \mathbf{b}\cdot\mathbf{w}\,dV = \int_{\mathcal{R}} \mathbf{A}\cdot\nabla\mathbf{w}\,dV, \qquad (i)$$

for *all* smooth vector fields $\mathbf{w}(\mathbf{x})$. Here $\partial\mathcal{R}$ is the boundary of $\mathcal{R}$. Show by localization that

$$\mathrm{div}\,\mathbf{A} + \mathbf{b} = \mathbf{o} \qquad \text{at each } \mathbf{x}\in\mathcal{R}. \qquad (ii)$$

One often speaks of (i) as being the *weak form* of the differential statement (ii), and (ii) as being the *strong form* of the integral statement (i). Note that (i) does not require $\mathbf{A}(\mathbf{x})$ to be differentiable while (ii) does.

Problem 1.42. Let $\mathbf{u}(\mathbf{x})$ be a smooth vector field defined on some region $\mathcal{R}$. Suppose that

$$\int_{\partial\mathcal{D}} \mathbf{u}(\mathbf{x}) \otimes \mathbf{n}(\mathbf{x})\ \mathrm{d}A \ = \ \mathbf{0} \qquad \text{for all subregions } \mathcal{D} \subset \mathcal{R}, \tag{i}$$

where $\mathbf{n}(\mathbf{x})$ is the unit outward normal vector at a point $\mathbf{x}$ on the boundary $\partial\mathcal{D}$. Show by using the divergence theorem and localization that (i) holds if and only if

$$\boldsymbol{\nabla}\mathbf{u} = \mathbf{0} \qquad \text{at each } \mathbf{x} \in \mathcal{R}. \tag{ii}$$

Problem 1.43.

(a) Let $\mathbf{S}(\mathbf{x})$ be a continuously differentiable tensor field on $\mathcal{R}$ with the property

$$\int_{\partial\mathcal{D}} \mathbf{S}(\mathbf{x})\mathbf{n}(\mathbf{x})\ \mathrm{d}A \ = \ \mathbf{o} \qquad \text{for all subregions } \mathcal{D} \subset \mathcal{R}, \tag{i}$$

where $\mathbf{n}(\mathbf{x})$ is the unit outward normal vector at a point $\mathbf{x}$ on the boundary $\partial\mathcal{D}$. Show by using the divergence theorem and localization that (i) implies

$$\mathsf{div}\ \mathbf{S} = \mathbf{o} \qquad \text{at each } \mathbf{x} \in \mathcal{R}. \tag{ii}$$

(b) Conversely, if $\mathbf{S}(\mathbf{x})$ is a tensor field such that (ii) holds, show that then (i) holds.

(c) Suppose that $\mathbf{S}(\mathbf{x})$ is a smooth tensor field that obeys (ii). Show that

$$\int_{\mathcal{D}} \mathbf{Sn} \cdot \mathbf{w}\, dA = \int_{\mathcal{D}} \mathbf{S} \cdot \nabla\mathbf{w}\, dV$$

for any smooth vector field $\mathbf{w}$.

Problem 1.44. Reconsider the tensor field $\mathbf{S}(\mathbf{x})$ introduced in part (a) of Problem 1.43. Suppose that in addition to equation (i) there, $\mathbf{S}(\mathbf{x})$ also has the property that

$$\int_{\partial\mathcal{D}} \mathbf{x} \times \mathbf{S}(\mathbf{x})\mathbf{n}(\mathbf{x})\ \mathrm{d}A \ = \ \mathbf{o} \qquad \text{for all subregions } \mathcal{D} \subset \mathcal{R}. \tag{iii}$$

Show that if (i) and (iii) hold, then

$$\mathbf{S} = \mathbf{S}^T \qquad \text{at each } \mathbf{x} \in \mathcal{R}. \tag{iv}$$

Hint: Use the divergence theorem, localization and equation (ii) of Problem 1.43.

---

6. Functions of a tensor.

Problem 1.45. The function $\mathrm{J}(\mathbf{F}) = \det \mathbf{F}$ is defined for all nonsingular tensors $\mathbf{F}$. Show that

$$\frac{\partial \mathrm{J}}{\partial \mathbf{F}} = \mathrm{J}\mathbf{F}^{-\mathrm{T}}. \tag{1.202}$$

Problem 1.46. Consider the function $W(\mathbf{F})$ defined for all nonsingular tensors $\mathbf{F}$ by $W(\mathbf{F}) = \widehat{W}(I_1(\mathbf{C}))$ where $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ and $I_1(\mathbf{C}) = \operatorname{tr}\mathbf{C}$. Calculate the components of the 4-tensor

$$\frac{\partial^2 W(\mathbf{F})}{\partial F_{ij}\partial F_{k\ell}}.$$

Solution:

First note that

$$\frac{\partial F_{ij}}{\partial F_{k\ell}} = \delta_{ik}\delta_{j\ell}, \tag{i}$$

$$\mathbf{C} = \mathbf{F}^T\mathbf{F} \qquad \Rightarrow \qquad C_{ij} = F_{ki}F_{kj}, \tag{ii}$$

$$I_1 = \operatorname{tr}\mathbf{C} = C_{ii} = F_{ki}F_{ki}. \tag{iii}$$

Therefore

$$\frac{\partial I_1}{\partial F_{pq}} \overset{(iii)}{=} \frac{\partial}{\partial F_{pq}}(F_{ki}F_{ki}) \overset{(i)}{=} \delta_{kp}\delta_{iq}F_{ki} + F_{ki}\delta_{kp}\delta_{iq} = F_{pq} + F_{pq} = 2F_{pq}, \tag{iv}$$

and so

$$\frac{\partial W(I_1)}{\partial F_{k\ell}} = W'(I_1)\frac{\partial I_1}{\partial F_{k\ell}} \overset{(iv)}{=} 2W'(I_1)F_{k\ell}. \tag{v}$$

Thus

$$\begin{aligned}\frac{\partial^2 W(I_1)}{\partial F_{ij}\partial F_{k\ell}} = \frac{\partial}{\partial F_{ij}}\left(\frac{\partial W(I_1)}{\partial F_{k\ell}}\right) \quad &\overset{(v)}{=} \frac{\partial}{\partial F_{ij}}(2W'(I_1)F_{k\ell}) \overset{(i)}{=} 2W''(I_1)\frac{\partial I_1}{\partial F_{ij}}F_{k\ell} + 2W'(I_1)\delta_{ki}\delta_{\ell j} = \\ &\overset{(iv)}{=} 4W''(I_1)F_{ij}F_{k\ell} + 2W'(I_1)\delta_{ki}\delta_{\ell j}, \end{aligned} \qquad \square$$

where a prime denotes differentiation with respect to the argument.

Problem 1.47. The scalar valued function $\widehat{W}(\mathbf{F})$ is defined for all nonsingular tensors $\mathbf{F}$, and the scalar valued function $\overline{W}(\mathbf{C})$ is defined for all symmetric positive definite tensors $\mathbf{C}$. Suppose that these two functions are related by

$$\widehat{W}(\mathbf{F}) = \overline{W}(\mathbf{C}) \qquad \text{where} \quad \mathbf{C} = \mathbf{F}^T\mathbf{F}. \tag{i}$$

Furthermore, suppose the tensors $\mathbf{S}$, $\mathbf{T}$ and $\mathbf{F}$ are related by

$$\mathbf{S} = \frac{\partial \widehat{W}}{\partial \mathbf{F}}(\mathbf{F}), \qquad \mathbf{T} = \frac{1}{J}\mathbf{S}\mathbf{F}^T \quad \text{where} \quad J = \det\,\mathbf{F}. \tag{ii}$$

Derive an expression for $\mathbf{T}$ in terms of $\partial\overline{W}(\mathbf{C})/\partial\mathbf{C}$ and $\mathbf{F}$. Specialize it to the case where

$$\overline{W}(\mathbf{C}) = \operatorname{tr}\mathbf{C}. \tag{iii}$$

Solution:

From the chain rule and $(i)_1$,

$$\frac{\partial \widehat{W}}{\partial F_{ij}} = \frac{\partial \overline{W}}{\partial C_{pq}} \frac{\partial C_{pq}}{\partial F_{ij}}. \tag{iv}$$

We now calculate the last term in (iv) by differentiating $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ with respect to $\mathbf{F}$:

$$\frac{\partial C_{pq}}{\partial F_{ij}} = \frac{\partial}{\partial F_{ij}}(F^T_{pk}F_{kq}) = \frac{\partial}{\partial F_{ij}}(F_{kp}F_{kq}) = \delta_{ki}\delta_{pj}F_{kq} + F_{kp}\delta_{ki}\delta_{qj} = \delta_{pj}F_{iq} + F_{ip}\delta_{qj}.$$

Substituting this back into (iv) gives

$$\frac{\partial \widehat{W}}{\partial F_{ij}} = \delta_{pj}F_{iq}\frac{\partial \overline{W}}{\partial C_{pq}} + F_{ip}\delta_{qj}\frac{\partial \overline{W}}{\partial C_{pq}} = F_{iq}\frac{\partial \overline{W}}{\partial C_{jq}} + F_{ip}\frac{\partial \overline{W}}{\partial C_{pj}} \overset{(*)}{=} F_{iq}\frac{\partial \overline{W}}{\partial C_{qj}} + F_{iq}\frac{\partial \overline{W}}{\partial C_{qj}} = 2F_{iq}\frac{\partial \overline{W}}{\partial C_{qj}}$$

where in step $(*)$, we used the symmetry of $\mathbf{C}$ to write $C_{qj} = C_{jq}$ in the first term, and changed the dummy subscript p to q in the second term. Thus we have

$$\frac{\partial \widehat{W}}{\partial \mathbf{F}} = 2\mathbf{F}\frac{\partial \overline{W}}{\partial \mathbf{C}}. \tag{v}$$

Substituting this into $(ii)_1$ and the result into $(ii)_2$ gives

$$\mathbf{T} = \frac{2}{J}\mathbf{F}\frac{\partial \overline{W}}{\partial \mathbf{C}}\mathbf{F}^T. \qquad \Box \tag{vi}$$

Now suppose that $\overline{W}(\mathbf{C}) = \operatorname{tr}\mathbf{C}$. Then

$$\frac{\partial}{\partial C_{ij}}(\operatorname{tr}\mathbf{C}) = \frac{\partial}{\partial C_{ij}}(C_{kk}) = \delta_{ki}\delta_{kj} = \delta_{ij} \qquad \Rightarrow \qquad \frac{\partial}{\partial \mathbf{C}}(\operatorname{tr}\mathbf{C}) = \mathbf{I}.$$

Substituting this into (vi) gives

$$\mathbf{T} = \frac{2}{J}\mathbf{F}\mathbf{I}\mathbf{F}^T = \frac{2}{J}\mathbf{F}\mathbf{F}^T. \tag{vii}$$

Problem 1.48. Let $\mathbb{C}$ be a constant 4-tensor and suppose that the function $\widehat{W}(\mathbf{E})$ is defined for all symmetric 2-tensors $\mathbf{E}$ by

$$\widehat{W}(\mathbf{E}) = W(E_{11}, E_{12}, ...E_{33}) = \frac{1}{2}\,\mathbb{C}_{ijk\ell}E_{ij}E_{k\ell},$$

where $\mathbb{C}_{ijkl}$ and E_{ij} are the components of $\mathbb{C}$ and $\mathbf{E}$ in some fixed basis. Calculate

$$\frac{\partial W}{\partial E_{ij}} \qquad \text{and} \qquad \frac{\partial^2 W}{\partial E_{ij}\partial E_{kl}}.$$

Following the discussion in Section 1.8.4, ensure that these derivative have the proper symmetries.

Remark: We will encounter this function W later in the linear theory of elasticity where it will correspond to the strain energy density at a point in the body with $\mathbf{E}$ being the infinitesimal strain tensor and $\mathbb{C}$ a tensor of elastic moduli.

7. Additional problems.

Problem 1.49. Consider an N-dimensional Euclidean vector space. It is useful to consider an N-dimensional vector space (rather than a 3-dimensional one) since we will use this result later for the 9-dimensional vector space of all linear transformations. Let $\mathbf{a}_1$ and $\mathbf{a}_2$ be two non-null vectors. If $\mathbf{a}_2$ is perpendicular to all vectors perpendicular to $\mathbf{a}_1$, show that $\mathbf{a}_2$ is parallel to $\mathbf{a}_1$.

Solution:

Consider an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}$ with $\mathbf{a}_1$ parallel to, say, the basis vector $\mathbf{e}_N$. Then $\mathbf{a}_1 = \alpha_1 \mathbf{e}_N$ for some nonzero scalar α_1. The set of all vectors perpendicular to $\mathbf{a}_1$ is now the set spanned by $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{N-1}\}$. Since the vector $\mathbf{a}_2$ is perpendicular to all vectors perpendicular to $\mathbf{a}_1$, it must be perpendicular to each of these $N-1$ vectors and so it must be of the form $\mathbf{a}_2 = \alpha_2 \mathbf{e}_N$. This shows that $\mathbf{a}_2$ is parallel to $\mathbf{a}_1$.

An alternative proof that doesn't rely on a basis is the following where we proceed in three steps:

(a) First we show that corresponding to *any* two non-null vectors $\mathbf{a}_1, \mathbf{a}_2$ that there is a scalar α and a vector $\mathbf{n}$ perpendicular to $\mathbf{a}_1$ such that

$$\mathbf{a}_2 = \alpha \mathbf{a}_1 + \mathbf{n}, \qquad \mathbf{n} \cdot \mathbf{a}_1 = 0. \tag{i}$$

(b) Then we show that the representation (i) is unique.

(c) Finally we use (a) and (b) to establish the desired result.

(a) Given any $\mathbf{a}_1 \neq \mathbf{o}$ and $\mathbf{a}_2$, define α and $\mathbf{n}$ by

$$\alpha = \frac{\mathbf{a}_2 \cdot \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_1}, \qquad \mathbf{n} = \mathbf{a}_2 - \alpha \mathbf{a}_1. \tag{ii}$$

Then $\mathbf{n} \cdot \mathbf{a}_1 = \mathbf{a}_2 \cdot \mathbf{a}_1 - \alpha \mathbf{a}_1 \cdot \mathbf{a}_1 = \mathbf{o}$ which establishes (i).

(b) Suppose this representation is not unique. Then there exist a scalar β and vector $\mathbf{m}$ such that

$$\mathbf{a}_2 = \beta \mathbf{a}_1 + \mathbf{m}, \qquad \mathbf{m} \cdot \mathbf{a}_1 = 0. \tag{iii}$$

Subtracting (iii) from (i) gives

$$(\alpha - \beta)\mathbf{a}_1 + (\mathbf{n} - \mathbf{m}) = \mathbf{o}. \tag{iv}$$

Taking the scalar product of this with $\mathbf{a}_1$, and then using $(i)_2$ and $(iii)_2$ shows that $\beta = \alpha$. Then (iv) reduces to $\mathbf{m} = \mathbf{n}$. Thus if both (i) and (iii) hold then necessarily $\beta = \alpha$ and $\mathbf{m} = \mathbf{n}$. Thus the representation (i) is unique.

(c) By the result above

$$\mathbf{a}_2 = \alpha \mathbf{a}_1 + \mathbf{n}, \qquad \mathbf{a}_1 \cdot \mathbf{n} = 0. \tag{v}$$

Now suppose that $\mathbf{a}_2 \cdot \mathbf{x} = 0$ for all $\mathbf{x}$ for which $\mathbf{a}_1 \cdot \mathbf{x} = 0$. By $(v)_2$, one such $\mathbf{x}$ is $\mathbf{x} = \mathbf{n}$. Therefore it follows that $\mathbf{a}_2 \cdot \mathbf{n} = 0$. Thus taking the scalar product of $(v)_1$ with $\mathbf{n}$ yields $\mathbf{n} \cdot \mathbf{n} = 0$. Consequently $\mathbf{n} = \mathbf{o}$ and hence $\mathbf{a}_2 = \alpha \mathbf{a}_1$.

Problem 1.50. *Projection tensor.* The "**projection tensor**" $\mathbf{P}$ projects vectors onto a given plane $\mathcal{P}$. It takes any vector $\mathbf{v} \in \mathsf{V}$ into the vector $\mathbf{Pv} \in \mathcal{P}$ as illustrated geometrically in Figure 1.12. Determine $\mathbf{P}$. Show that $\mathbf{P}$ is singular.

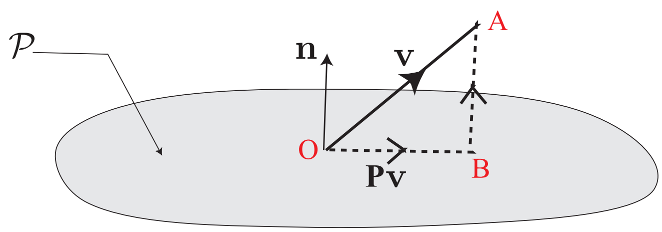


Figure 1.12: The projection $\mathbf{Pv}$ of a vector $\mathbf{v}$ onto the plane $\mathcal{P}$.

Solution:

Let $\mathbf{v}$ be an arbitrary vector: $\overrightarrow{OA} = \mathbf{v}$. Its image after projections is $\overrightarrow{OB} = \mathbf{Pv}$. Let $\mathbf{n}$ be a unit vector normal to the plane $\mathcal{P}$. Observe from Figure 1.12 that the vector $\overrightarrow{BA}$ has magnitude $\mathbf{v} \cdot \mathbf{n}$ and is in the direction $\mathbf{n}$. Thus

$$\overrightarrow{BA} = (\mathbf{v} \cdot \mathbf{n})\mathbf{n}.$$

Consequently

$$\mathbf{Pv} \ = \overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OA} - \overrightarrow{BA} = \mathbf{v} - (\mathbf{v} \cdot \mathbf{n})\mathbf{n}.$$

Thus

$$\mathbf{Pv} \ = \ \mathbf{v} - (\mathbf{v} \cdot \mathbf{n})\mathbf{n} \qquad \text{for all vectors} \quad \mathbf{v} \in \mathsf{V}. \tag{1.203}$$

This completely defines $\mathbf{P}$ since it tells us how it operates on an arbitrary vector $\mathbf{v}$.

Remark: Note that $\mathbf{P}$ is a *linear* operator since $\mathbf{P}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{Px} + \beta\mathbf{Py}$. Thus it is a linear transformation (tensor).

Remark: Recall that the tensor product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is the tensor, denoted by $\mathbf{a} \otimes \mathbf{b}$, that has the property $(\mathbf{a} \otimes \mathbf{b})\mathbf{v} = (\mathbf{b} \cdot \mathbf{v})\mathbf{a}$ for all vectors $\mathbf{v} \in \mathsf{V}$. Therefore we can express the projection tensor defined by (1.203) equivalently as

$$\mathbf{P} = \mathbf{I} - \mathbf{n} \otimes \mathbf{n}. \tag{1.204}$$

Remark: Observe from (1.203) or (1.204) or geometrically from Figure 1.12 that $\mathbf{Pn} = \mathbf{o}$. Since $\mathbf{n} \neq \mathbf{o}$ it follows from (1.82) that $\mathbf{P}$ is singular.

Remark: The components P_{ij} of $\mathbf{P}$ in a basis will be calculated in Problem 1.51 .

Problem 1.51. Determine the components in an arbitrary basis of the projection tensor $\mathbf{P}$ introduced in Problem 1.50.

Problem 1.52. Let $\mathbf{W}$ be a skew-symmetric tensor, i.e.

$$\mathbf{W} = -\mathbf{W}^T. \qquad (i)$$

Show that there is a vector $\mathbf{w}$ such that

$$\mathbf{W}\mathbf{x} = \mathbf{w} \times \mathbf{x} \qquad \text{for all vectors } \mathbf{x} \in \mathsf{V}. \qquad (ii)$$

In terms of components in a basis, show that

$$w_i = -\frac{1}{2} e_{ijk} W_{jk}. \qquad (iii)$$

Remark: Recall from the comment below (1.7) that any skew-symmetric matrix $[W]$ has *only three independent elements*, e.g. W_{12}, W_{23}, W_{31}. This is because the elements on the diagonal of a skew-symmetric matrix $[W]$ are zero and $W_{12} = -W_{21}, W_{23} = -W_{32}$ and $W_{31} = -W_{13}$. It is not surprising therefore that one can associate a vector $\mathbf{w}$ (which has three independent components) with each skew-symmetric tensor $\mathbf{W}$.

Problem 1.53. Consider the rotation tensor $\mathbf{Q}$ introduced in (1.189).

(a) Determine the components of $\mathbf{Q}$ in an arbitrary basis.

(b) Verify that $\mathbf{Q}$ is proper orthogonal.

(c) Show that

$$\mathsf{trace}\, \mathbf{Q} = 1 + 2\cos\theta \qquad (i)$$

and

$$\sin\theta\, n_i = -\frac{1}{2} e_{ipq} Q_{pq}. \qquad (ii)$$

Therefore, given a proper orthogonal $\mathbf{Q}$, one can find the rotation angle θ from (i) and thereafter the rotation axis $\mathbf{n}$ from (ii).

(d) Suppose that the axis of rotation coincides with one of the basis vectors, say $\mathbf{n} = \mathbf{e}_3$. Specialize your answer to part (a) and display the matrix $[Q]$.

Problem 1.54. The tensor $\mathbf{Q}$ is proper orthogonal.

(a) Show that $\det(\mathbf{Q} - \mathbf{I}) = 0$ and hence deduce that unity is an eigenvalue of $\mathbf{Q}$.

(b) Let $\mathbf{a}$ be the eigenvector of $\mathbf{Q}$ corresponding to the eigenvalue $\lambda = 1$. Consider an orthonormal basis $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. Show that $\mathbf{Qb}$ and $\mathbf{Qc}$ are perpendicular to $\mathbf{a}$ and therefore lie in the plane spanned by $\mathbf{b}$ and $\mathbf{c}$. Hence deduce that for some angle θ,

$$\mathbf{Qb} = \cos\theta\mathbf{b} + \sin\theta\mathbf{c}, \qquad \mathbf{Qc} = -\sin\theta\mathbf{b} + \cos\theta\mathbf{c}.$$

(c) Show that $\mathbf{Q}$ has the representation

$$\mathbf{Q} = \mathbf{a}\otimes\mathbf{a} + \cos\theta(\mathbf{b}\otimes\mathbf{b} + \mathbf{c}\otimes\mathbf{c}) + \sin\theta(\mathbf{c}\otimes\mathbf{b} - \mathbf{b}\otimes\mathbf{c}). \tag{i}$$

Hence show that (same as (1.189))

$$\mathbf{Qx} = \cos\theta\,\mathbf{x} + (1-\cos\theta)(\mathbf{a}\cdot\mathbf{x})\mathbf{a} + \sin\theta\,(\mathbf{a}\times\mathbf{x}),$$

for any vector $\mathbf{x}$.

(d) Calculate the principal scalar invariants of $\mathbf{Q}$

(e) Show that $\mathbf{Q}$ has no real eigenvalues other than unity.

Problem 1.55. The trace of a tensor was defined in (1.100). Show that the right-hand side of that formula, i.e.

$$\frac{\mathbf{Ax}\cdot(\mathbf{y}\times\mathbf{z}) + \mathbf{x}\cdot(\mathbf{Ay}\times\mathbf{z}) + \mathbf{x}\cdot(\mathbf{y}\times\mathbf{Az})}{\mathbf{x}\cdot(\mathbf{y}\times\mathbf{z})}, \tag{i}$$

is independent of the choice of the linearly independent vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$.

Problem 1.56. *Scalar (dot) product of two tensors.* Consider the scalar-valued function

$$f(\mathbf{A}, \mathbf{B}) \;=\; \mathsf{trace}(\mathbf{AB}^T)$$

defined for all tensors $\mathbf{A}$ and $\mathbf{B} \in \mathsf{Lin}$. Show that this function f has the following properties for all tensors $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathsf{Lin}$ and all scalars α:

(a) $f(\mathbf{A}, \mathbf{B}) = f(\mathbf{B}, \mathbf{A})$,

(b) $f(\alpha\mathbf{A}, \mathbf{B}) = \alpha f(\mathbf{A}, \mathbf{B})$,

(c) $f(\mathbf{A}+\mathbf{C}, \mathbf{B}) = f(\mathbf{A}, \mathbf{B}) + f(\mathbf{C}, \mathbf{B})$ and

(d) $f(\mathbf{A}, \mathbf{A}) > 0$ provided $\mathbf{A} \neq \mathbf{0}$ and $f(\mathbf{0}, \mathbf{0}) = 0$.

Problem 1.57. Show that

$$(\mathbf{I}+\mathbf{a}\otimes\mathbf{b})^{-1}=\mathbf{I}-\frac{\mathbf{a}\otimes\mathbf{b}}{1+\mathbf{a}\cdot\mathbf{b}} \qquad (\text{provided } \mathbf{a}\cdot\mathbf{b}\neq -1)$$

Remark: It is shown in Problem 1.18 that

$$\det(\mathbf{I}+\mathbf{a}\otimes\mathbf{b})=1+\mathbf{a}\cdot\mathbf{b}.$$

Problem 1.58. Given any tensor $\mathbf{A}$, there is a tensor called its cofactor and denoted by $\mathbf{A}^*$ with the property

$$\mathbf{A}^*(\mathbf{a}\times\mathbf{b})=\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b}$$

for all vectors $\mathbf{a}$ and $\mathbf{b}$; see Chadwick [2]. If $\mathbf{A}$ is nonsingular show that

$$\mathbf{A}^*=(\det\mathbf{A})\,\mathbf{A}^{-T}.$$

Solution:

We can obtain the desired result if we can show that $\det\mathbf{A}\,(\mathbf{a}\times\mathbf{b})=\mathbf{A}^T(\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b})$. Thus we evaluate the ith component of $\mathbf{A}^T(\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b})$:

$$\begin{aligned}\Big(\mathbf{A}^T(\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b})\Big)_i &= A_{ji}(\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b})_j=A_{ji}e_{jk\ell}(\mathbf{A}\mathbf{a})_k(\mathbf{A}\mathbf{b})_\ell=A_{ji}e_{jk\ell}(A_{kp}a_p)(A_{\ell q}b_q)=\\ &= e_{jk\ell}A_{ji}A_{kp}A_{\ell q}a_pb_q\overset{(1.40)}{=}\det\mathbf{A}\,e_{ipq}a_pb_q=\det\mathbf{A}\,(\mathbf{a}\times\mathbf{b})_i\end{aligned}$$

and so

$$\mathbf{A}^T(\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b})=\det\mathbf{A}\,(\mathbf{a}\times\mathbf{b}).$$

Thus when $\mathbf{A}$ is nonsingular.

$$\mathbf{A}\mathbf{a}\times\mathbf{A}\mathbf{b}=(\det\mathbf{A})\,\mathbf{A}^{-T}(\mathbf{a}\times\mathbf{b}),$$

whence $\mathbf{A}^*=(\det\mathbf{A})\,\mathbf{A}^{-T}$.

To solve without using components, consider $\mathbf{A}^T\mathbf{A}^*$ acting on the set basis vectors, and dot the results with the set of basis vectors.

Problem 1.59. Let $\mathbf{S}$ be a symmetric tensor and $\mathbf{u}$ a vector. Show that $\mathbf{u}$ is an eigenvector of $\mathbf{S}$ if and only if $\mathbf{S}\mathbf{u}\otimes\mathbf{u}=\mathbf{u}\otimes\mathbf{S}\mathbf{u}$.

Problem 1.60. *4-tensors.* As before: vectors in the 3-dimensional Euclidean vector space V are denoted by lowercase boldface latin letters; 2-tensors are denoted by uppercase boldface latin letters and a 2-tensor $\mathbf{A}$ is a linear transformation that takes a vector $\mathbf{x}\in\mathsf{V}$ into another vector in V that we denote by $\mathbf{A}\mathbf{x}$ (subject to certain rules); the collection of all 2-tensors is itself a vector space that we denote by Lin; 4-tensors are denoted with uppercase blackboard letters and a 4-tensor $\mathbb{L}$ is a linear transformation that takes a 2-tensor $\mathbf{A} \in \mathsf{Lin}$ into another 2-tensor in Lin that we denote by $\mathbb{L}\mathbf{A}$ (subject to certain rules).

We shall denote the set of all 4-tensors by LinLin.

Reference: G. Del Piero [3].

– The identity 4-tensor $\mathbb{I}$ and null 4-tensor $\mathbb{O}$ obey

$$\mathbb{I}\mathbf{A} = \mathbf{A}, \qquad \mathbb{O}\mathbf{A} = \mathbf{0} \qquad \text{for all 2-tensors } \mathbf{A} \in \mathsf{Lin}.$$

– The product of two 4-tensors $\mathbb{C}$ and $\mathbb{D}$ is defined by

$$(\mathbb{C}\mathbb{D})\mathbf{A} = \mathbb{C}(\mathbb{D}\mathbf{A}) \qquad \text{for all 2-tensors } \mathbf{A} \in \mathsf{Lin}$$

(a) Exercise: Let $\mathbb{T}$ be the particular 4-tensor that takes any 2-tensor $\mathbf{A} \in \mathsf{Lin}$ into the 2-tensor $\mathbf{A}^T \in \mathsf{Lin}$:

$$\mathbb{T}\mathbf{A} = \mathbf{A}^T \qquad \text{for all 2-tensors } \mathbf{A} \in \mathsf{Lin}.$$

It is called the **transposition 4-tensor**.

(a1) Show that $\mathbb{T}$ is invertible and its inverse is $\mathbb{T}$:

$$\mathbb{T}\mathbb{T} = \mathbb{I};$$

(a2) Define the 4-tensor $\mathbb{S}$ by

$$\mathbb{S} = \frac{1}{2}(\mathbb{I} + \mathbb{T}).$$

Show that $\mathbb{S}\mathbb{T} = \mathbb{T}\mathbb{S} = \mathbb{S}$, $\mathbb{S}\mathbb{S} = \mathbb{S}$ and

$$\mathbb{S}\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) \qquad \text{for all 2-tensors } \mathbf{A} \in \mathsf{Lin}.$$

– Define the **transpose** $\mathbb{L}^T$ of the tensor $\mathbb{L}$ by

$$\mathbb{L}^T\mathbf{A} \cdot \mathbf{B} = \mathbf{A} \cdot \mathbb{L}\mathbf{B} \qquad \text{for all 2-tensors } \mathbf{A}, \mathbf{B} \in \mathsf{Lin}.$$

(b) Exercise: Show that $(\mathbb{C}\mathbb{D})^T = \mathbb{D}^T\mathbb{C}^T$.

– Define the **tensor product of two 2-tensors A** and **B** to be the 4-tensor, denoted by $\mathbf{A} \boxtimes \mathbf{B}$, for which

$$(\mathbf{A} \boxtimes \mathbf{B})\mathbf{X} = \mathbf{A}\mathbf{X}\mathbf{B}^T \qquad \text{for all 2-tensors } \mathbf{X} \in \mathsf{Lin}.$$

(c) Exercise: Show that

(c1) $(\mathbf{A} \boxtimes \mathbf{I})\mathbf{X} = \mathbf{A}\mathbf{X}$

(c2) $(\mathbf{I} \boxtimes \mathbf{A}^T)\mathbf{X} = \mathbf{X}\mathbf{A}$

(c3) $(\mathbf{A} \boxtimes \mathbf{B})^T = \mathbf{A}^T \boxtimes \mathbf{B}^T$.

(c4) $\mathbf{A}\mathbb{L}(\mathbf{X}\mathbf{B}) = (\mathbf{A} \boxtimes \mathbf{I})\mathbb{L}(\mathbf{I} \boxtimes \mathbf{B}^T)\mathbf{X}$

(c5) $(\mathbf{A} \boxtimes \mathbf{B})(\mathbf{C} \boxtimes \mathbf{D}) = \mathbf{AC} \boxtimes \mathbf{BD}$

(c6) $\mathbb{T}(\mathbf{A} \boxtimes \mathbf{B}) = (\mathbf{B} \boxtimes \mathbf{A})\mathbb{T}$

(c7) $(\mathbf{A} \boxtimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \boxtimes \mathbf{B}^{-1}$.

– We say that $\mathbb{L}$ has the first minor symmetry if

$$\mathbb{L} = \mathbb{T}\mathbb{L},$$

where $\mathbb{T}$ is the transposition tensor introduced earlier. We say that $\mathbb{L}$ has the second minor symmetry if

$$\mathbb{L} = \mathbb{L}\mathbb{T}$$

Verify that $\mathbb{L}$ has both minor symmetries if

$$\mathbb{L} = \mathbb{S}\mathbb{L}\mathbb{S} \qquad \text{where} \quad \mathbb{S} = \frac{1}{2}(\mathbb{I} + \mathbb{T})$$

We say that $\mathbb{L}$ has the major symmetry (or simply we say that $\mathbb{L}$ is symmetric) if

$$\mathbb{A} = \mathbb{A}^T.$$

(d) Exercise: Show that if $\mathbb{L}$ has the major symmetry and one of the minor symmetries, it necessarily has the other minor symmetry.

– Components: Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be an orthonormal basis for the Euclidean vector space V. Define the 3^4 numbers $\mathbb{L}_{ijk\ell}$ by

$$\mathbb{L}_{ijk\ell} = \big(\mathbb{L}(\mathbf{e}_k \otimes \mathbf{e}_\ell)\big) \cdot (\mathbf{e}_i \otimes \mathbf{e}_j)$$

Show that

$$\mathbb{L} = \mathbb{L}_{ijk\ell}(\mathbf{e}_i \otimes \mathbf{e}_k) \boxtimes (\mathbf{e}_j \otimes \mathbf{e}_\ell)$$

$$\big(\mathbb{L}\mathbf{A}\big)_{ij} = \mathbb{L}_{ijk\ell}A_{k\ell}, \qquad (\mathbf{A} \boxtimes \mathbf{B})_{ijk\ell} = A_{ik}B_{j\ell}.$$

If the components of $\mathbb{L}$ are $\mathbb{L}_{ijk\ell}$ what are the components of $\mathbb{L}^T$? If a tensor $\mathbb{L}$ has in turn (a) the first minor symmetry, (b) the second minor symmetry, and (c) the major symmetry, what does this imply about the components $\mathbb{L}_{ijk\ell}$?

– Let $\mathbf{F}(\mathbf{X})$ be a 2-tensor valued function of all 2-tensors $\mathbf{X}$. Assuming $\mathbf{F}(\mathbf{X})$ is differentiable at $\mathbf{X}$, its gradient is the 4-tensor denoted by $\nabla\mathbf{F}$ for which

$$\mathbf{F}(\mathbf{X} + \mathbf{H}) = \mathbf{F}(\mathbf{X}) + (\nabla\mathbf{F}(\mathbf{X}))\mathbf{H} + o(|\mathbf{H}|) \qquad \text{with} \quad \lim_{|\mathbf{H}| \to 0} \frac{o(|\mathbf{H}|}{|\mathbf{H}|} \to 0.$$

(e) Exercise: Show that the components of this 4-tensor are

$$(\nabla F)_{ijk\ell} = \frac{\partial F_{ij}}{\partial X_{k\ell}}.$$

References:

1. R. Abeyaratne, Chapters 1, 2, 3 and 5 of *A Brief Review of Some Mathematical Preliminaries*, Volume I in the series *Lecture Notes on The Mechanics of Solids*, http://web.mit.edu/abeyaratne/lecture_notes.html.

2. P. Chadwick, Chapter 1 of *Continuum Mechanics: Concise Theory and Problems*, Dover, 1999.

3. G. Del Piero, Some properties of the set of fourth-order tensors, with application to elasticity, *Journal of Elasticity*, 9(1979), pp. 245-261.

4. M.E. Gurtin, E. Fried and L. Anand, Chapters 1-4 and 17 of *The Mechanics and Thermodynamics of Continua*, Cambridge University Press, 2010.

5. P. R. Halmos, *Finite-Dimensional Vector Spaces*, Springer-Verlag, Berlin 1974.

6. C.S. Jog, Chapter 1 of Continuum Mechanics: Foundations and Applications of Mechanics, Cambridge, 2015.

7. J.K. Knowles, *Linear Vector Spaces and Cartesian Tensors*, Oxford University Press, New York, 1997.

8. R. W. Ogden, Chapter 1 of *Non-linear Elastic Deformations*, Dover, 1997.

Chapter 2

Kinematics: Finite Deformation

Several short videos on the material in Sections 2.3 - 2.6 can be found here.

In this chapter we shall consider *purely geometric issues* ("kinematics") associated with the deformation of a body. At this stage we will not address the *causes* of the deformation, such as what the applied loading is, nor will we discuss the characteristics of the material of which the body is composed, assuming only that it can be described as a continuum. Our focus will be entirely on kinematic considerations[1].

Problem 2.2 shows that the familiar notion of strain as defined in linear theories of solid mechanics is deficient when considering finite (i.e. large) deformations. This is why it is necessary that we devote some time to a careful analysis of the kinematics of large deformations. An even more detailed discussion, especially of time dependent entities such as strain-rate, can be found in the references [1, 3, 5, 7] listed at the end of this chapter.

A roadmap of this chapter is as follows: in Section 2.1 we introduce the notion of a deformation: $\mathbf{y} = \mathbf{y}(\mathbf{x})$. Some homogeneous deformations $\mathbf{y}(\mathbf{x}) = \mathbf{F}\mathbf{x} + \mathbf{b}$ such as a pure stretch, simple shear and a rigid deformation are discussed in Section 2.2. In Section 2.3 we introduce the deformation gradient tensor $\mathbf{F}(\mathbf{x})$, the central ingredient needed to describe the deformation in the neighborhood of a particle $\mathbf{x}$. We then consider in Section 2.4 an

[1] It is worth mentioning that in developing a continuum theory for a material, the appropriate kinematic description of the body is not entirely independent of the nature of the forces. For example, in describing the interaction between particles in a dielectric material subjected to an electric field, one might allow for internal forces *and internal couples* between every pair of points in the body. This in turn requires that the kinematics allow for *independent* displacement *and rotation* fields in the body. In general, the kinematics and the forces must be *conjugate* to each other in order to construct a self-consistent theory.

infinitesimal material curve, material surface and material region in the reference configuration and examine the geometric characteristics of their images in the deformed configuration. The decomposition of a general deformation gradient tensor $\mathbf{F}$ into the product of a rigid rotation $\mathbf{R}$ and pure stretches $\mathbf{U}, \mathbf{V}$ is described in Section 2.5. Section 2.6 introduces the notion of strain. In Section 2.7.1 we calculate the deformation gradient tensor $\mathbf{F}$ and the left Cauchy-Green deformation tensor $\mathbf{B}$ in cylindrical and spherical polar coordinates. We discuss material and spatial descriptions of a field in Section 2.8. Finally we linearize the preceding results in Section 2.9. In the appendix, Section 2.11, we touch on the material time derivative and a transport formula.

2.1 Deformation

In this chapter we examine how the geometric characteristics of one configuration of the body (the "deformed" configuration) *are related* to those of a second configuration (the "undeformed" or "reference" configuration). Thus we are necessarily concerned with two configurations of the body[2].

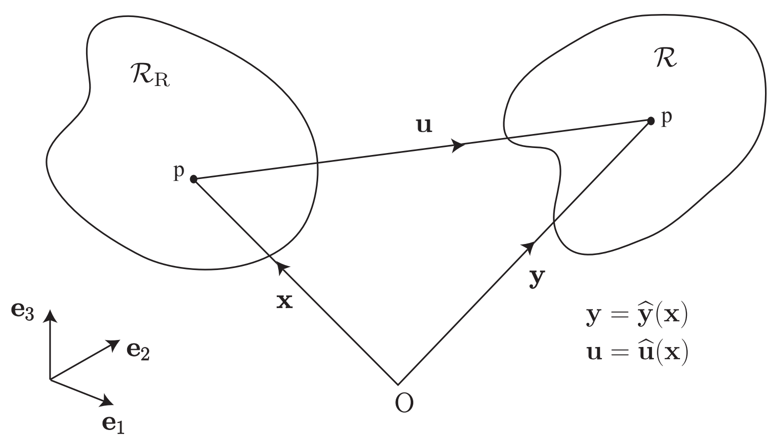


Figure 2.1: The respective regions $\mathcal{R}_R$ and $\mathcal{R}$ occupied by a body in the reference and deformed configurations; the position vectors of a generic particle p in these two configurations are denoted by $\mathbf{x}$ and $\mathbf{y}$. The displacement of this particle is $\mathbf{u}$.

In the deformed configuration the body occupies a region $\mathcal{R}$ of physical space while the corresponding region in a reference configuration is $\mathcal{R}_R$. The position vector of a generic

[2]This is in contrast to the study of many fluids where only the current configuration needs to be considered.

particle p in the reference configuration is denoted by $\mathbf{x}$ and the deformation takes this particle to the position $\widehat{\mathbf{y}}(\mathbf{x})$ in the deformed configuration. This is illustrated in Figure 2.1. We write

$$\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}), \qquad \mathbf{x} \in \mathcal{R}_{\mathrm{R}}, \quad \mathbf{y} \in \mathcal{R}. \tag{2.1}$$

We refer to $\widehat{\mathbf{y}}(\mathbf{x})$ as the **deformation**. We use the "hat" over $\mathbf{y}$ in order to distinguish the *function* $\widehat{\mathbf{y}}(\cdot)$ from its value $\mathbf{y}$. As we progress through these notes, we will usually omit the "hat" unless the context does not make clear whether we are referring to $\widehat{\mathbf{y}}$ or $\mathbf{y}$, or when we wish to emphasize the distinction.

The reference configuration serves two main purposes. One, geometric *changes*, e.g. the change in length of a fiber, are measured with respect to this configuration. Two, it provides a convenient way in which to "label" particles of the body: since there is a one-to-one correspondence between a particle p and its position $\mathbf{x}$ in the reference configuration[3], we can uniquely identify a particle by $\mathbf{x}$. Whenever there is no confusion in doing so, *we shall speak of "the particle* $\mathbf{x}$*" rather than "the particle located at* $\mathbf{x}$ *in the reference configuration".*

The reference configuration is an *arbitrary* conveniently chosen configuration, the only requirement being that it be a "possible" configuration that the body *can* occupy. It need not, for example, be the initial configuration of a body undergoing a motion. Unless explicitly stated otherwise, we shall always consider one fixed reference configuration.

The **displacement** $\widehat{\mathbf{u}}(\mathbf{x})$ of the particle $\mathbf{x}$ is

$$\widehat{\mathbf{u}}(\mathbf{x}) = \widehat{\mathbf{y}}(\mathbf{x}) - \mathbf{x}, \tag{2.2}$$

as shown in Figure 2.1. The functions $\widehat{\mathbf{y}}$ (and $\widehat{\mathbf{u}}$) are defined on $\mathcal{R}_{\mathrm{R}}$, i.e. at every $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$.

For physical reasons we require that (a) a single particle $\mathbf{x}$ not split into two particles and occupy two locations $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ in the deformed configuration, and (b) that two particles $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ both not occupy the same location $\mathbf{y}$ in the deformed configuration. Therefore we take the deformation $\mathbf{x} \to \widehat{\mathbf{y}}(\mathbf{x})$ to be one-to-one.

Unless explicitly stated otherwise, we will assume $\widehat{\mathbf{y}}(\mathbf{x})$ to be "smooth", i.e. that it may be differentiated as many times as needed, and that these derivatives are continuous on $\mathcal{R}_{\mathrm{R}}$. There are situations in which this must be relaxed: for example, if we consider a "dislocation" it will be necessary to allow the displacement field to be discontinuous across a surface in the body; or if we consider the deformation of a "two-phase composite material"

[3]See Chapter 1 of Volume II for a more careful discussion of what we mean by "a particle" and "a configuration" in the continuum theory.

we must allow the gradient of the displacement field to be discontinuous across the interface between the two phases.

Finally, we pick and fix (an arbitrary) right-handed orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. When referring to components of vector and tensor quantities, it will always be with respect to this basis (unless explicitly stated otherwise). In particular, denoting the components of $\mathbf{x}$ and $\mathbf{y}$ in this basis by $x_i = \mathbf{x} \cdot \mathbf{e}_i$ and $y_i = \mathbf{y} \cdot \mathbf{e}_i$, we write the deformation (2.1) in component form as

$$y_i = \widehat{y}_i(x_1, x_2, x_3) = x_i + \widehat{u}_i(x_1, x_2, x_3). \tag{2.3}$$

The rectangular cartesian coordinates of a particle in the reference and deformed configurations are (x_1, x_2, x_3) and (y_1, y_2, y_3) respectively.

2.2 Some homogeneous deformations.

In a **homogeneous deformation**, the position $\mathbf{y}(\mathbf{x})$ of a particle in the deformed configuration depends linearly on its position $\mathbf{x}$ in the reference configuration and so the deformation has the form

$$\mathbf{y} = \mathbf{y}(\mathbf{x}) = \mathbf{F}\mathbf{x} + \mathbf{b}, \tag{2.4}$$

where $\mathbf{F}$ is a *constant* tensor with positive determinant and $\mathbf{b}$ is a *constant* vector representing a rigid translation. In component form,

$$y_i = F_{ik} x_k + b_i.$$

Note that

$$\frac{\partial y_i}{\partial x_j} = F_{ik} \frac{\partial x_k}{\partial x_j} = F_{ik} \delta_{kj} = F_{ij}.$$

Exercise: Show that the set of points lying on a straight line/plane/ellipsoid in the reference configuration are mapped by a homogeneous deformation into a straight line/plane/ellipsoid in the deformed configuration.

2.2.1 Pure stretch.

The constant tensor $\mathbf{F}$ in a pure stretch $\mathbf{y}(\mathbf{x}) = \mathbf{F}\mathbf{x}$ is symmetric and positive definite. We will see in Section 2.5 that this implies that such a deformation does not involve a rigid rotational part.

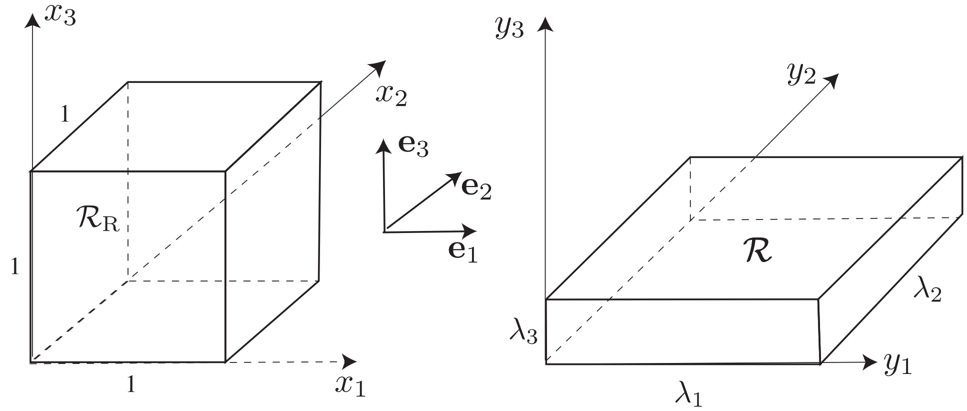


Figure 2.2: Pure homogeneous stretch of a cube. A unit cube in the reference configuration is carried into an orthorhombic region of dimensions $\lambda_1 \times \lambda_2 \times \lambda_3$.

Consider a body that occupies a unit cube in a reference configuration with its edges aligned with the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ as shown in Figure 2.2. The body is subjected to the deformation

$$y_1 = \lambda_1 x_1, \qquad y_2 = \lambda_2 x_2, \qquad y_3 = \lambda_3 x_3, \tag{2.5}$$

where the three $\lambda_i' s$ are positive constants. This deformation maps the $1 \times 1 \times 1$ undeformed cube $\mathcal{R}_R$ into a $\lambda_1 \times \lambda_2 \times \lambda_3$ orthorhombic region $\mathcal{R}$ as shown in Figure 2.2. The positive constants λ_1, λ_2 and λ_3 represent the ratios by which the three edges of the cube *stretch* in the respective directions $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. This deformation is called a **pure stretch**.

Observe that a material fiber parallel to an edge of the cube in the reference configuration simply undergoes a stretch and no rotation under this deformation. However, this is not true of all material fibers, e.g. a fiber oriented along a diagonal of a face of the cube will undergo both a length change and a rotation.

The deformation (2.5) can be written in matrix form as

$$\{y\} = [F]\{x\} \qquad \text{where} \qquad [F] = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \tag{2.6}$$

and in tensor form as

$$\mathbf{y} = \mathbf{F}\mathbf{x} \qquad \text{where} \qquad \mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 \; + \; \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 \; + \; \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{2.7}$$

The 3×3 matrix $[F]$ in $(2.6)_2$ is the matrix of components of the tensor $\mathbf{F}$ in (2.7). In the special case where the deformed and reference configurations coincide, i.e. the body is undeformed, then $\mathbf{y}(\mathbf{x}) = \mathbf{x}$ and so $\lambda_1 = \lambda_2 = \lambda_3 = 1$ whence $[F] = [I], F_{ij} = \delta_{ij}$ and $\mathbf{F} = \mathbf{I}$.

It will be useful for future purposes to note that the deformation gradient tensor (2.7) in this example is symmetric and positive definite, with principal directions $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and corresponding principal stretches $\lambda_1, \lambda_2, \lambda_3$.

We now consider some particular pure stretches:

– **Pure dilatation:** The special case $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$ of (2.5) describes a pure dilatation of the body. In this case

$$\mathbf{F} = \lambda(\mathbf{e}_1 \otimes \mathbf{e}_1 \ + \ \mathbf{e}_2 \otimes \mathbf{e}_2 \ + \ \mathbf{e}_3 \otimes \mathbf{e}_3) = \lambda \mathbf{I},$$

and so the deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ specializes to $\mathbf{y} = \lambda\, \mathbf{x}$. This shows that *all* dimensions of the body are uniformly scaled by the factor λ.

– **Isochoric pure stretch:** A deformation is said to be isochoric if it is volume preserving at each point of the body. In the case of the pure stretch (2.7) this requires

$$\lambda_1 \lambda_2 \lambda_3 = 1. \tag{2.8}$$

Note that this is a constraint on the three stretches in that their values cannot be prescribed independently.

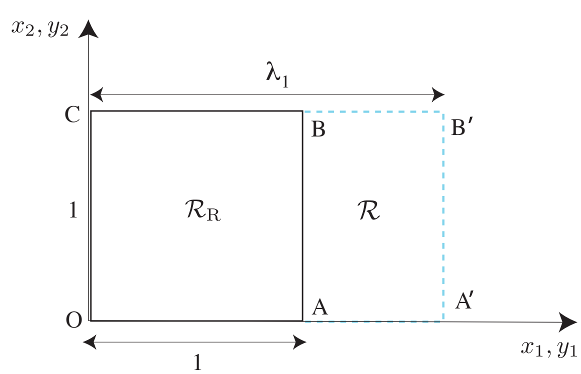


Figure 2.3: Uniaxial stretch in the $\mathbf{e}_1$-direction. A unit cube $\mathcal{R}_R$ in the reference configuration is carried into a $\lambda_1 \times 1 \times 1$ tetragonal region $\mathcal{R}$ in the deformed configuration.

– **Uniaxial stretch:** The deformation

$$y_1 = \lambda x_1, \qquad y_2 = x_2, \qquad y_3 = x_3, \tag{2.9}$$

is illustrated in Figure 2.3. It describes a uniaxial stretch in the $\mathbf{e}_1$-direction. If $\lambda > 1$ the stretch is an elongation, if $\lambda < 1$ a contraction. (The terms "tensile" and "compressive" refer to stress not deformation.) This deformation can be written in tensor form as $\mathbf{y} = \mathbf{F}\mathbf{x}$ by taking

$$\mathbf{F} = \lambda \mathbf{e}_1 \otimes \mathbf{e}_1 \,+\, \mathbf{e}_2 \otimes \mathbf{e}_2 \,+\, \mathbf{e}_3 \otimes \mathbf{e}_3 = \mathbf{I} + (\lambda - 1)\mathbf{e}_1 \otimes \mathbf{e}_1.$$

More generally, the deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ with

$$\mathbf{F} = \mathbf{I} + (\lambda - 1)\mathbf{m}_{\mathrm{R}} \otimes \mathbf{m}_{\mathrm{R}} \tag{2.10}$$

represents a uniaxial stretch in the direction of the unit vector $\mathbf{m}_{\mathrm{R}}$. If the body is composed of an incompressible material, its volume cannot change and so it cannot undergo a uniaxial stretch of the form (2.9) (except for the trivial one where $\lambda = 1$).

– **Isochoric uniaxial stretch with equal lateral stretch:** If the body undergoes a stretch λ in the x_1-direction and equal lateral stretches in the x_2- and x_3-directions, and the deformation is isochoric, then the deformation is described by

$$y_1 = \lambda x_1, \qquad y_2 = \lambda^{-1/2} x_2, \qquad y_3 = \lambda^{-1/2} x_3. \tag{2.11}$$

Since $\lambda_1 = \lambda, \lambda_2 = \lambda_3 = \lambda^{-1/2}$ we have $\lambda_1\lambda_2\lambda_3 = 1$.

2.2.2 Simple shear.

The displacement field associated with a **simple shearing deformation** has components

$$u_1 = kx_2, \qquad u_2 = 0, \qquad u_3 = 0,$$

where k is a constant. Observe that the displacement (vector) of every particle has only an $\mathbf{e}_1$-component and the magnitude of this displacement increases linearly with x_2. This carries the cube $\mathcal{R}_{\mathrm{R}}$ into the sheared region $\mathcal{R}$ as shown in Figure 2.4. One refers to a plane $x_2 =$ constant as a *shearing (or glide) plane*, the x_1-direction as the *shearing direction* and the scalar k as the *amount of shear*.

The deformation $\mathbf{y}(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x})$ associated with a simple shear has components

$$y_1 = x_1 + kx_2, \qquad y_2 = x_2, \qquad y_3 = x_3. \tag{2.12}$$

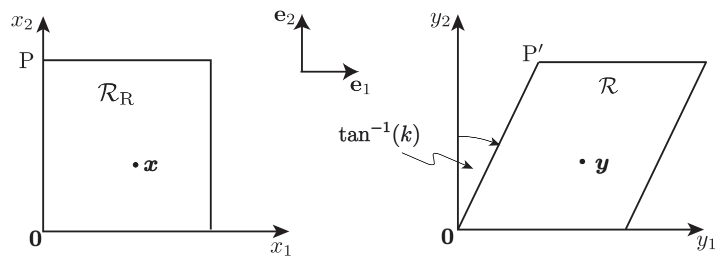


Figure 2.4: Simple shear of a cube. Each plane x_2 = constant undergoes a displacement in the x_1-direction by the amount kx_2.

The deformation (2.12) can be written in matrix form as

$$\{y\} = [F]\{x\} \qquad \text{where} \qquad [F] = \begin{pmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.13}$$

and tensor form as

$$\mathbf{y} = \mathbf{F}\mathbf{x} \qquad \text{where} \qquad \mathbf{F} = \mathbf{I} + k\ \mathbf{e}_1 \otimes \mathbf{e}_2.$$

Note that $\det \mathbf{F} = 1$ and therefore a simple shear preserves volume. (That $\det \mathbf{F}$ is a measure of volume change is discussed in Section 2.4.3.)

More generally, if $\mathbf{n}_R$ and $\mathbf{m}_R$ are arbitrary unit vectors that are orthogonal, $|\mathbf{m}_R| = |\mathbf{n}_R| = 1, \mathbf{m}_R \cdot \mathbf{n}_R = 0$, a simple shear whose glide plane normal is $\mathbf{n}_R$ and shear direction is $\mathbf{m}_R$ is described by the deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ where

$$\mathbf{F} = \mathbf{I} + k\mathbf{m}_R \otimes \mathbf{n}_R. \tag{2.14}$$

One can of course consider combinations of the preceding homogeneous deformations. For example consider

$$\mathbf{y} = \mathbf{F}_1\mathbf{F}_2\mathbf{x} \qquad \text{where} \quad \mathbf{F}_1 = \mathbf{I} + \alpha\mathbf{a} \otimes \mathbf{a}, \quad \mathbf{F}_2 = \mathbf{I} + k\mathbf{m} \otimes \mathbf{n},$$

where the vectors $\mathbf{a}, \mathbf{m}, \mathbf{n}$ have unit length and $\mathbf{m} \cdot \mathbf{n} = 0$. We can consider this deformation in two steps as $\mathbf{y} = \mathbf{F}_1(\mathbf{F}_2\mathbf{x})$: in the first step a particle goes from $\mathbf{x} \to \mathbf{F}_2\mathbf{x}$ corresponding to a simple shearing of the body. In the second step it goes from $\mathbf{F}_2\mathbf{x} \to \mathbf{F}_1(\mathbf{F}_2\mathbf{x})$, and the body undergoes a uniaxial stretching. Figure 2.5 illustrates such a deformation in the particular case $\mathbf{a} = \mathbf{n} = \mathbf{e}_2, \mathbf{m} = \mathbf{e}_1$. Perhaps it is worth pointing out that the individual tensors $\mathbf{F}_1$, $\mathbf{F}_2$ enter the tensor $\mathbf{F}$ *multiplicatively* (not additively), i.e. as $\mathbf{F} = \mathbf{F}_1\mathbf{F}_2$ not $\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2$.

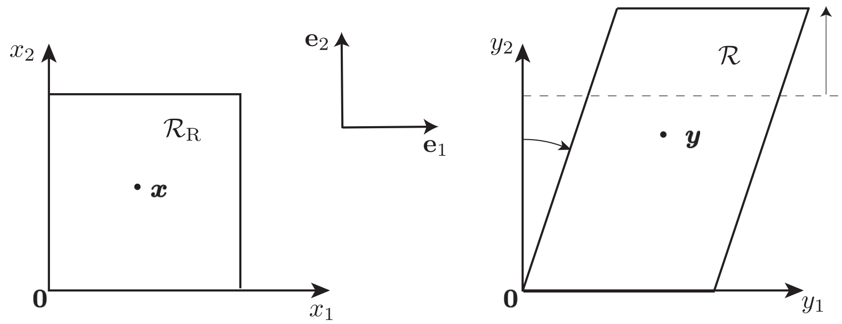


Figure 2.5: A unit cube subjected to a simple shear (with glide plane normal $\mathbf{e}_2$) and a uniaxial stretch in the direction $\mathbf{e}_2$.

2.2.3 Rigid deformation.

A deformation is said to be rigid if the distance between all pairs of particles remains unchanged, i.e. if the distance $|\mathbf{x}_1 - \mathbf{x}_2|$ between any two particles $\mathbf{x}_1$ and $\mathbf{x}_2$ in the reference configuration equals the distance $|\mathbf{y}(\mathbf{x}_2) - \mathbf{y}(\mathbf{x}_1)|$ between them in the deformed configuration:

$$|\mathbf{y}(\mathbf{x}_2) - \mathbf{y}(\mathbf{x}_1)|^2 = |\mathbf{x}_2 - \mathbf{x}_1|^2 \qquad \text{for all } \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{R}_{\mathrm{R}}. \tag{2.15}$$

It can be shown (Problem 2.41) that a deformation is rigid if and only if it has the form

$$\mathbf{y}(\mathbf{x}) = \mathbf{Q}\mathbf{x} + \mathbf{b}, \tag{2.16}$$

where $\mathbf{Q}$ is a constant *orthogonal* tensor and $\mathbf{b}$ is a constant vector[4]. When the deformation preserves orientation, it follows because of (2.24) below that $\det \mathbf{Q} = +1$ and therefore $\mathbf{Q}$ is proper orthogonal and represents a rigid rotation. The vector $\mathbf{b}$ represents a rigid translation.

A rigid *material* (or rigid body) is a material that can *only* undergo rigid deformations.

All of the deformations considered in this section were homogeneous in the sense that they were of the form (2.4) with $\mathbf{F}$ being a *constant* tensor. Most deformations are *not* of this form, a simple example being

$$\left.\begin{aligned} y_1 &= x_1 \cos \alpha x_3 - x_2 \sin \alpha x_3, \\ y_2 &= x_1 \sin \alpha x_3 + x_2 \cos \alpha x_3, \\ y_3 &= x_3, \end{aligned}\right\}$$

[4]Recall from Problem 1.4.8 that an orthogonal tensor $\mathbf{Q}$ preserves length, i.e. $|\mathbf{Q}\mathbf{x}| = |\mathbf{x}|$ for all vectors $\mathbf{x}$. Therefore it is immediately clear that the deformation (2.16) obeys (2.15). What requires proof is the converse, that (2.15) implies (2.16).

where α is a constant. This can be shown to represent a torsional deformation about the x_3-axis in which each plane x_3 = constant rotates by an angle αx_3. [Exercise.]

2.3 Deformation in the neighborhood of a particle. Deformation gradient tensor.

We now return to a general deformation $\mathbf{y} = \mathbf{y}(\mathbf{x})$. Questions such as "what is the state of stress at a particle $\mathbf{x}$?" depend not only on the deformation at $\mathbf{x}$ but also on the deformation of all particles in a neighborhood of $\mathbf{x}$. Thus we now turn to characterizing the deformation in the *neighborhood* of a generic particle. Intuitively, we expect the deformation of a small ball of material centered at $\mathbf{x}$ to consist of a combination of a rigid translation, a rigid rotation and a "straining", notions that we shall make precise in what follows.

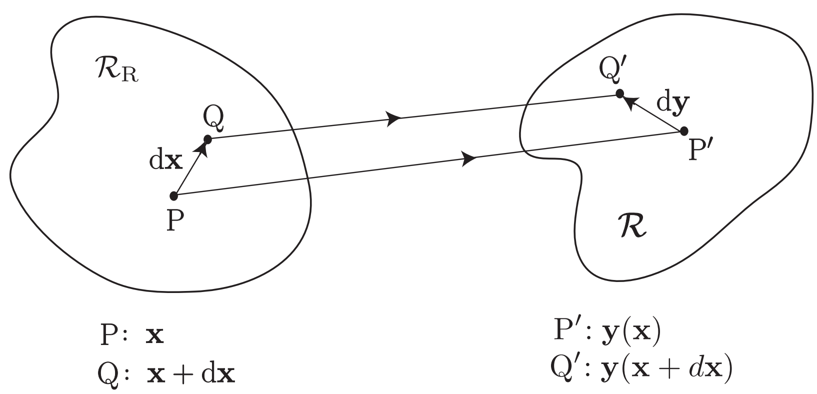


Figure 2.6: An infinitesimal material fiber in the reference and deformed configurations.

Consider two particles $\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$ located in the reference configuration at P and Q as depicted in Figure 2.6. The material fiber joining them is $\overrightarrow{PQ} = d\mathbf{x}$. In the deformed configuration these particles are located at P' and Q' with respective position vectors $\mathbf{y}(\mathbf{x})$ and $\mathbf{y}(\mathbf{x} + d\mathbf{x})$. The deformed image of this *material* fiber[5] is

$$\overrightarrow{P'Q'} = d\mathbf{y} = \mathbf{y}(\mathbf{x} + d\mathbf{x}) - \mathbf{y}(\mathbf{x}). \tag{2.17}$$

From (1.161), we have $\mathbf{y}(\mathbf{x} + d\mathbf{x}) = \mathbf{y}(\mathbf{x}) + \nabla\mathbf{y}(\mathbf{x})\, d\mathbf{x} + \mathrm{o}(|d\mathbf{x}|)$ where the tensor $\nabla\mathbf{y}$ is the gradient of the deformation $\mathbf{y}(\mathbf{x})$. Thus when the two particles are *close to each other* we

[5]The fibers PQ and $P'Q'$ being *material* fibers refers to the fact that they are comprised of the same set of particles.

can write

$$d\mathbf{y} = \nabla\mathbf{y}\ d\mathbf{x} + \mathrm{o}(|\mathrm{d}\mathbf{x}|). \tag{2.18}$$

We denote the **deformation gradient tensor** $\nabla\mathbf{y}$ at particle $\mathbf{x}$ by

$$\boxed{\mathbf{F}(\mathbf{x}) := \nabla\mathbf{y}(\mathbf{x})}, \tag{2.19}$$

and write (2.18) formally as

$$\boxed{d\mathbf{y} = \mathbf{F}\, d\mathbf{x}.} \tag{2.20}$$

Note that (2.20) does *not* assume the deformation or deformation gradient to be "small"; only the two particles to be close to each other.

The deformation gradient tensor $\mathbf{F}$ carries an infinitesimal material fiber $d\mathbf{x}$ in the undeformed configuration into $d\mathbf{y} = \mathbf{F}\, d\mathbf{x}$ in the deformed configuration. It describes the deformation of *every* infinitesimal material fiber through $\mathbf{x}$, and therefore it describes the deformation of the entire neighborhood of $\mathbf{x}$. Thus one can calculate all local changes in geometry at $\mathbf{x}$ in terms of $\mathbf{F}(\mathbf{x})$, e.g. the change in length of a material fiber, the change in angle between two material fibers etc. We will carry out these calculations in Section 2.4.

The deformation gradient tensor is the principal entity used to study the deformation in the neighborhood of a particle. It characterizes *both* the rigid rotation and the "strain" at $\mathbf{x}$.

The equation $d\mathbf{y} = \mathbf{F}\, d\mathbf{x}$ is the local version in the vicinity of the particle $\mathbf{x}$ of the equation $\mathbf{y} = \mathbf{F}\mathbf{x}$ that we had previously when examining homogeneous deformations in Section 2.2.

Given $\mathbf{F}(\mathbf{x})$, we can calculate the deformation of every material fiber (through $\mathbf{x}$). Conversely, given the deformation of any *three* linearly independent material fibers, one can calculate $\mathbf{F}$, and therefore determine the deformation of all other material fibers (Problem 2.5).

The deformation gradient tensor $\mathbf{F}(\mathbf{x})$ is a 2-tensor field whose cartesian components[6],

$$F_{ij}(\mathbf{x}) = \frac{\partial y_i}{\partial x_j}(\mathbf{x}), \tag{2.21}$$

[6]Given a function $\phi(\mathbf{x})$, the partial derivative $\partial\phi/\partial x_k$ is often denoted by $\phi_{,k}$. If we were to adopt this notation we would write $F_{ij} = y_{i,j}$. However, since we have both referential coordinates x_1, x_2, x_3 and spatial coordinates y_1, y_2, y_3, we will encounter both $\partial/\partial x_i$ and $\partial/\partial y_i$. In order not to confuse one partial derivative with the other, we shall write-out the partial derivatives explicitly and *not* adopt the subscript comma notation.

correspond to the elements of a 3×3 matrix field $[F(\mathbf{x})]$. In terms of components, (2.20) reads

$$dy_i = F_{ij}\, dx_j. \tag{2.22}$$

In physically realizable deformations we expect (a) a single fiber $d\mathbf{x}$ to not split into two fibers $d\mathbf{y}^{(1)}$ and $d\mathbf{y}^{(2)}$, and (b) two fibers $d\mathbf{x}^{(1)}$ and $d\mathbf{x}^{(2)}$ not to coalesce into a single fiber $d\mathbf{y}$. This requires $d\mathbf{y} = \mathbf{F}d\mathbf{x}$ to be a one-to-one relation between $d\mathbf{x}$ and $d\mathbf{y}$ whence $\mathbf{F}$ must be *nonsingular*. The **Jacobian** determinant, J, therefore cannot vanish:

$$J \; := \; \det \mathbf{F} \neq 0. \tag{2.23}$$

Without any further restrictions, a deformation might map a right-handed triplet of vectors into a left-handed triplet of vectors (which would imply that the body has been turned "inside out" like a sock). We say that the deformation *preserves "orientation"* if every right-handed (linearly independent) triplet of material fibers $\{d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}, d\mathbf{x}^{(3)}\}$ is carried into a right-handed triplet of vectors $\{d\mathbf{y}^{(1)}, d\mathbf{y}^{(2)}, d\mathbf{y}^{(3)}\}$. According to Problem 2.42, orientation is preserved if and only if

$$J \; = \; \det \mathbf{F} > 0. \tag{2.24}$$

In these notes we will mostly be concerned with orientation-preserving deformations[7] and therefore, unless explicitly stated otherwise, assume (2.24) to hold.

<u>Caution</u>: In these notes we use the term "orientation-preserving" in two different ways. In the sense of the preceding paragraph, it refers to the preservation of the right-handedness (or left-handedness) of a triplet of vectors. When concerned with a particular material fiber, if its direction (orientation) in the reference and deformed configurations is the same, we shall say its orientation is preserved. The context should make clear the sense in which the term is being used.

In the special case where the deformed and reference configurations coincide, $\mathbf{y}(\mathbf{x}) = \mathbf{x}$ and so

$$\mathbf{F}(\mathbf{x}) = \mathbf{I} \qquad \text{in the reference configuration.} \tag{2.25}$$

[7]Problem 2.17 concerns the eversion of a hollow cylinder where the body is turned "inside out". Such deformations do not preserve orientation. See also Problem 5.13.

2.4 Change of length, orientation, angle, volume and area.

The deformation gradient tensor $\mathbf{F}(\mathbf{x})$ characterizes the deformation of *all* (infinitesimal) material fibers $d\mathbf{x}$ at the particle $\mathbf{x}$. We can therefore calculate various geometric quantities of interest (near $\mathbf{x}$) in terms of $\mathbf{F}$. In particular we now calculate the *local*[8] change in length of a fiber, change in angle between two fibers, change in volume of an infinitesimal material region and the change in area of an infinitesimal material surface, all in terms of $\mathbf{F}$.

The change in length is related to the notion of fiber stretch (or normal strain), the change in angle to the notion of shear strain and the change in volume to the notion of volumetric (or dilatational) strain. The change in area is indispensable when calculating the traction (force per unit area) on a surface.

2.4.1 Change of length and direction.

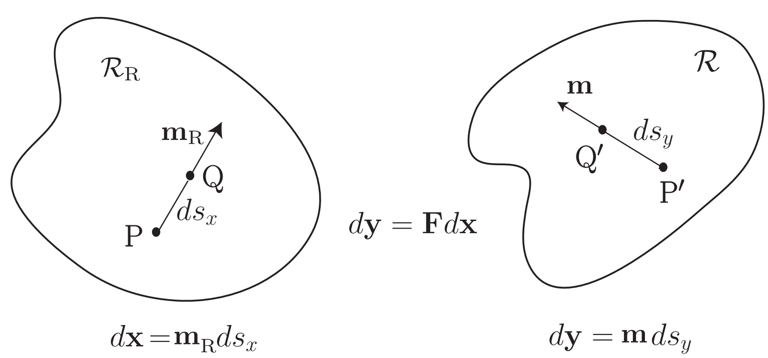


Figure 2.7: The infinitesimal material fiber $\overrightarrow{PQ}$ has length ds_x and direction $\mathbf{m}_{\rm R}$ in the reference configuration and length ds_y and direction $\mathbf{m}$ in the deformed configuration.

Consider a material fiber that has length ds_x and direction $\mathbf{m}_{\rm R}$ in the reference configuration. Then $d\mathbf{x} = (ds_x)\mathbf{m}_{\rm R}$. If ds_y and $\mathbf{m}$ denote its length and direction in the deformed configuration, then $d\mathbf{y} = (ds_y)\mathbf{m}$. Given ds_x and $\mathbf{m}_{\rm R}$ we want to calculate ds_y and $\mathbf{m}$.

[8] These are *local* changes in the sense that they refer to changes of *infinitesimally small* line, area and volume elements at $\mathbf{x}$.

Since $d\mathbf{y}$ and $d\mathbf{x}$ are related by $d\mathbf{y} = \mathbf{F}d\mathbf{x}$, it follows that

$$(ds_y)\mathbf{m} = (ds_x)\mathbf{F}\mathbf{m}_{\mathrm{R}}. \tag{2.26}$$

On taking the magnitude of both sides of this vector equation we get $ds_y|\mathbf{m}| = ds_x|\mathbf{F}\mathbf{m}_{\mathrm{R}}|$ and so the deformed length of the fiber is

$$ds_y = ds_x|\mathbf{F}\mathbf{m}_{\mathrm{R}}|. \tag{2.27}$$

The **stretch** λ at the particle $\mathbf{x}$ in the direction $\mathbf{m}_{\mathrm{R}}$ is defined as the ratio

$$\lambda := \lim_{ds_x \to 0} \frac{ds_y}{ds_x}, \tag{2.28}$$

and so

$$\boxed{\lambda = \lambda(\mathbf{m}_{\mathrm{R}}) = |\mathbf{F}\mathbf{m}_{\mathrm{R}}|.} \tag{2.29}$$

This is the stretch of the fiber with referential direction $\mathbf{m}_{\mathrm{R}}$. Exercise: among all fibers of all orientations at $\mathbf{x}$, which has the maximum stretch?

The stretch λ is related to the relative change in length by

$$\frac{ds_y - ds_x}{ds_x} = \lambda - 1.$$

We will return to this later in (2.68).

The direction $\mathbf{m}$ of this fiber in the deformed configuration is found from (2.26) and (2.27) to be

$$\mathbf{m} = \frac{\mathbf{F}\mathbf{m}_{\mathrm{R}}}{|\mathbf{F}\mathbf{m}_{\mathrm{R}}|}. \tag{2.30}$$

It is worth noting from (2.29) and (2.30) that

$$\lambda\mathbf{m} = \mathbf{F}\mathbf{m}_{\mathrm{R}}. \tag{2.31}$$

2.4.2 Change of angle.

In order to calculate the change in angle between two fibers we make use of the fact that the angle between two vectors appears in the expression for the scalar product.

Consider two fibers $d\mathbf{x}^{(1)}$ and $d\mathbf{x}^{(2)}$ in the reference configuration as shown in Figure 2.8. They have the same length ds_x and are oriented in the respective directions $\mathbf{m}_{\mathrm{R}}^{(1)}$ and $\mathbf{m}_{\mathrm{R}}^{(2)}$. Thus

$$d\mathbf{x}^{(1)} = ds_x\mathbf{m}_{\mathrm{R}}^{(1)}, \qquad d\mathbf{x}^{(2)} = ds_x\mathbf{m}_{\mathrm{R}}^{(2)}. \tag{2.32}$$

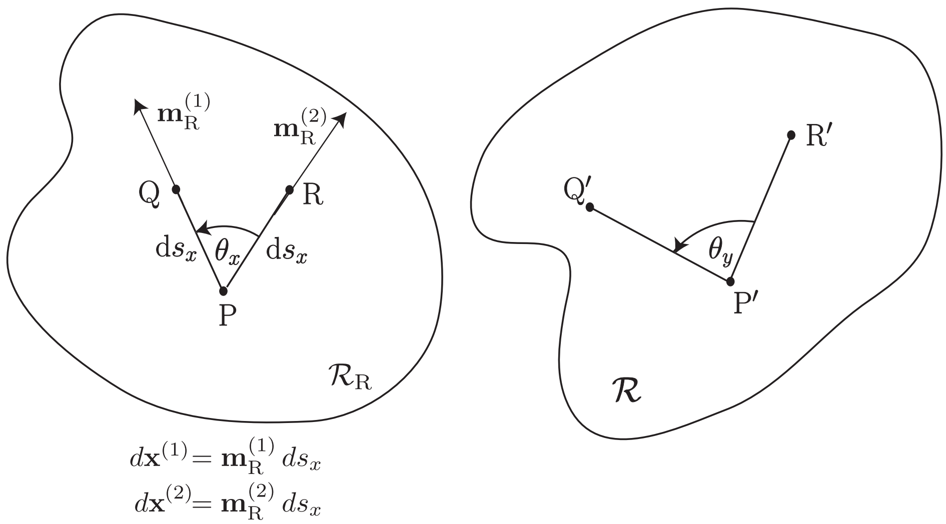


Figure 2.8: In the reference configuration two infinitesimal material fibers have equal length ds_x and directions $\mathbf{m}_{\mathrm{R}}^{(1)}$ and $\mathbf{m}_{\mathrm{R}}^{(2)}$. The angle between them in the reference and deformed configurations are θ_x and θ_y respectively.

If θ_x denotes the angle between them, by the definition of the scalar product,

$$d\mathbf{x}^{(1)} \cdot d\mathbf{x}^{(2)} = |d\mathbf{x}^{(1)}||d\mathbf{x}^{(2)}| \cos\theta_x \quad \Rightarrow \quad \cos\theta_x = \frac{d\mathbf{x}^{(1)}}{|d\mathbf{x}^{(1)}|} \cdot \frac{d\mathbf{x}^{(2)}}{|d\mathbf{x}^{(2)}|} \overset{(2.32)}{=} \mathbf{m}_{\mathrm{R}}^{(1)} \cdot \mathbf{m}_{\mathrm{R}}^{(2)}. \tag{2.33}$$

We want to calculate the angle between this pair of fibers in the deformed configuration.

In the deformed configuration these fibers are characterized by

$$d\mathbf{y}^{(1)} = \mathbf{F} d\mathbf{x}^{(1)} \overset{(2.32)}{=} ds_x\, \mathbf{F}\mathbf{m}_{\mathrm{R}}^{(1)}, \qquad d\mathbf{y}^{(2)} = \mathbf{F} d\mathbf{x}^{(2)} \overset{(2.32)}{=} ds_x\, \mathbf{F}\mathbf{m}_{\mathrm{R}}^{(2)}. \tag{2.34}$$

Letting θ_y denote the angle between them, again by the definition of the scalar product of two vectors we have

$$\cos\theta_y = \frac{d\mathbf{y}^{(1)}}{|d\mathbf{y}^{(1)}|} \cdot \frac{d\mathbf{y}^{(2)}}{|d\mathbf{y}^{(2)}|} \overset{(2.34)}{=} \frac{\mathbf{F}\mathbf{m}_{\mathrm{R}}^{(1)} \cdot \mathbf{F}\mathbf{m}_{\mathrm{R}}^{(2)}}{|\mathbf{F}\mathbf{m}_{\mathrm{R}}^{(1)}||\mathbf{F}\mathbf{m}_{\mathrm{R}}^{(2)}|}. \tag{2.35}$$

Given the deformation gradient tensor $\mathbf{F}$ and the directions $\mathbf{m}_{\mathrm{R}}^{(1)}$ and $\mathbf{m}_{\mathrm{R}}^{(2)}$ of two fibers in the reference configuration, (2.35) gives the angle between them in the deformed configuration.

The decrease in angle, $\gamma := \theta_x - \theta_y$, is the *shear* associated with the directions $\mathbf{m}_{\mathrm{R}}^{(1)}, \mathbf{m}_{\mathrm{R}}^{(2)}$: $\gamma = \gamma(\mathbf{m}_{\mathrm{R}}^{(1)}, \mathbf{m}_{\mathrm{R}}^{(2)})$. Exercise: among all pairs of fibers at $\mathbf{x}$, which pair undergoes the maximum change in angle, i.e. maximum shear?

2.4.3 Change of volume.

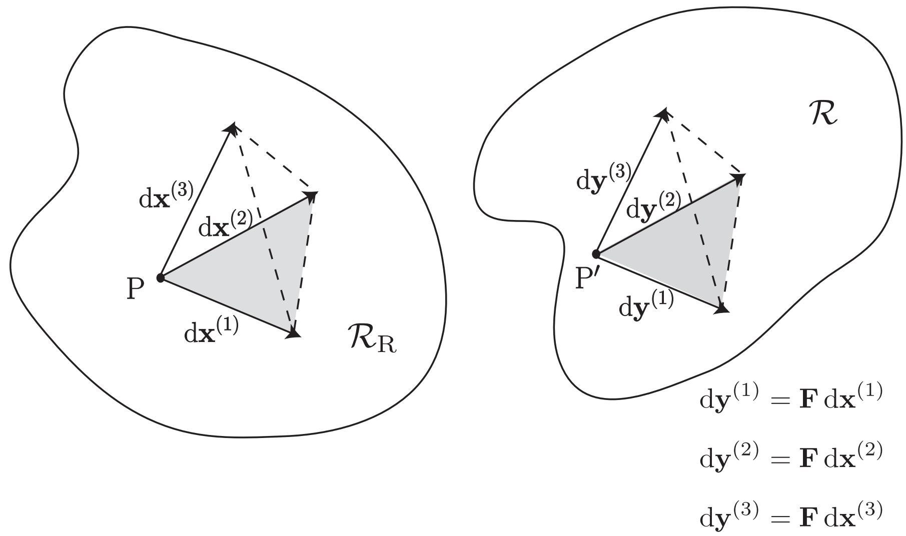


Figure 2.9: Three infinitesimal material fibers defining a tetrahedral region. The volumes of the tetrahedrons in the reference and deformed configurations are dV_x and dV_y respectively.

Consider three linearly independent material fibers $d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}, d\mathbf{x}^{(3)}$ in the reference configuration that form a tetrahedron of volume dV_x as shown in Figure 2.9. The deformation carries them into $d\mathbf{y}^{(1)} = \mathbf{F}\mathbf{x}^{(1)}, d\mathbf{y}^{(2)} = \mathbf{F}d\mathbf{x}^{(2)}, d\mathbf{y}^{(3)} = \mathbf{F}d\mathbf{x}^{(3)}$. If dV_y denotes the volume of the tetrahedron formed by the deformed fibers, according to Problem 1.11,

$$\boxed{dV_y = J\ dV_x} \qquad \text{where} \quad \mathrm{J} = \det \mathbf{F}, \tag{2.36}$$

having used $\det \mathbf{F} > 0$. This relates the volumes of an infinitesimal part of the body in the reference and deformed configurations.

Observe from (2.36) that a deformation preserves the volume of *every* infinitesimal part of the body if and only if

$$J(\mathbf{x}) = 1 \qquad \text{for all} \quad \mathbf{x} \in \mathcal{R}_\mathrm{R}. \tag{2.37}$$

Such a deformation is said to be **isochoric** or locally volume preserving.

An incompressible *material* is a material that can *only* undergo isochoric deformations. Keep in mind that a material that is not incompressible can undergo an isochoric deformation, e.g. a simple shear.

2.4.4 Change of area.

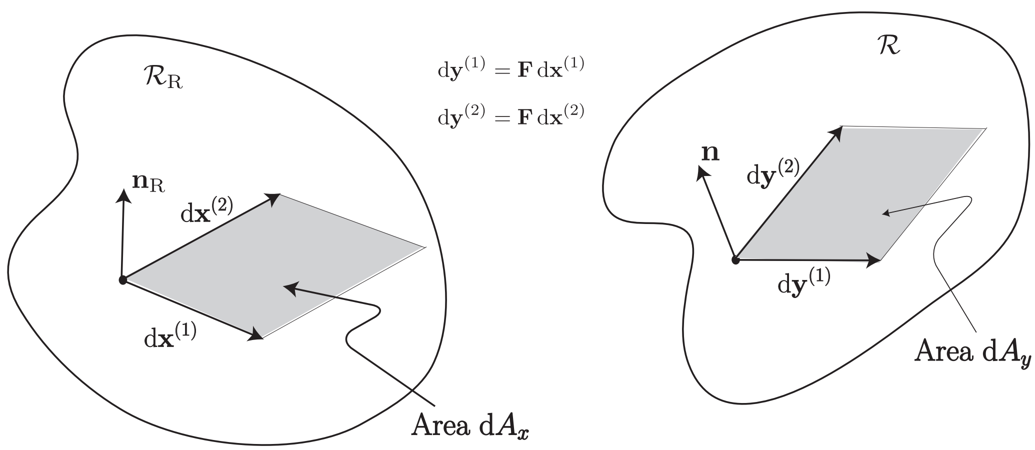


Figure 2.10: The two infinitesimal material fibers $d\mathbf{x}^{(1)}$ and $d\mathbf{x}^{(2)}$ define a parallelogram of area dA_x and unit normal vector $\mathbf{n}_\mathrm{R}$. The corresponding quantities in the deformed configuration are $d\mathbf{y}^{(1)}$, $d\mathbf{y}^{(2)}$, dA_y and $\mathbf{n}$.

Finally we turn to the relationship between two area elements in the reference and deformed configurations. Consider two linearly independent material fibers $d\mathbf{x}^{(1)}$ and $d\mathbf{x}^{(2)}$ in the reference configuration that form a parallelogram as shown in Figure 2.10. Let $\mathrm{d}A_x$ denote its area and let $\mathbf{n}_\mathrm{R}$ be a unit vector normal to the plane of the parallelogram. The deformation carries these fibers into $d\mathbf{y}^{(1)} = \mathbf{F}d\mathbf{x}^{(1)}$ and $d\mathbf{y}^{(2)} = \mathbf{F}d\mathbf{x}^{(2)}$. Let $\mathrm{d}A_y$ and $\mathbf{n}$ be the area and unit normal vector respectively of the parallelogram defined by $d\mathbf{y}^{(1)}$ and $d\mathbf{y}^{(2)}$. According to Problem 2.43 these two vector areas are related by

$$\boxed{\mathrm{d}A_y\,\mathbf{n} = \mathrm{d}A_x\,J\,\mathbf{F}^{-T}\,\mathbf{n}_\mathrm{R}\ .} \tag{2.38}$$

Equation (2.38) is known as *Nanson's formula*. The relation between the scalar areas dA_y and dA_x is found by taking the magnitude of both sides of this vector equation which leads to

$$\mathrm{d}A_y = \mathrm{d}A_x\,J\,|\mathbf{F}^{-T}\,\mathbf{n}_\mathrm{R}|. \tag{2.39}$$

The relation between the unit normal vectors $\mathbf{n}_\mathrm{R}$ and $\mathbf{n}$ is obtained by using (2.39) in (2.38):

$$\mathbf{n} = \frac{\mathbf{F}^{-T}\mathbf{n}_\mathrm{R}}{|\mathbf{F}^{-T}\mathbf{n}_\mathrm{R}|}. \tag{2.40}$$

It is worth noting that a material fiber in the direction $\mathbf{n}_\mathrm{R}$ in the reference configuration maps into $\mathbf{F}\mathbf{n}_\mathrm{R}$, which in general is *not* the direction $\mathbf{n}$ given by (2.40). The vectors $\mathbf{n}_\mathrm{R}$ and

$\mathbf{n}$ are defined by the fact that they are normal to the particular material surface elements being considered. They are not attached to a material fiber. To see this clearly, consider a simple shear as illustrated in Figure 2.11. Here $P'Q'$ is the image of PQ and the unit vectors $\mathbf{n}_\mathrm{R}$ and $\mathbf{n}$ are defined as being normal to PQ and $P'Q'$ respectively. A material fiber in the direction $\mathbf{n}_\mathrm{R}$ in the reference configuration (i.e. on the dashed green line) remains on the dashed line. Its direction in the deformed configuration is $\mathbf{e}_1$, not $\mathbf{n}$.

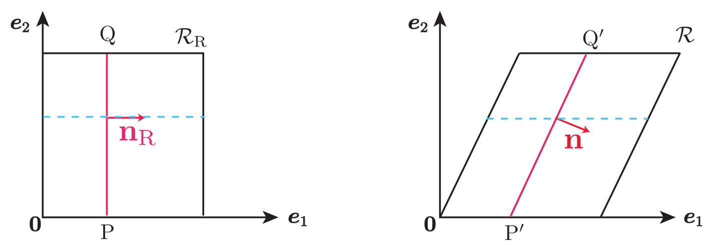


Figure 2.11: $P'Q'$ is the image of PQ. The unit vectors $\mathbf{n}_\mathrm{R}$ and $\mathbf{n}$ are normal to PQ and $P'Q'$ respectively. A material fiber that is in the direction $\mathbf{n}_\mathrm{R}$ in the reference configuration (i.e. on the dashed green line) remains on the dashed green line in the deformed configuration – its direction in the deformed configuration is $\mathbf{e}_1$, not $\mathbf{n}$.

2.4.5 Worked examples.

Problem 2.4.1. The region $\mathcal{R}_\mathrm{R}$ occupied by a body in a reference configuration is a unit cube.

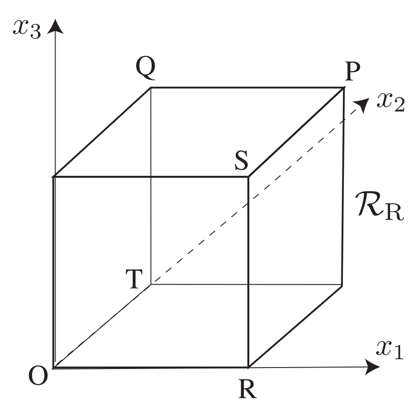


Figure 2.12: Unit cube $\mathcal{R}_\mathrm{R}$ occupied by a body in its reference configuration. (Problem 2.4.1)

The body undergoes the pure stretch $\mathbf{y} = \mathbf{F}\mathbf{x}$ described by

$$y_1 = \lambda_1 x_1, \qquad y_2 = \lambda_2 x_2, \qquad y_3 = \lambda_3 x_3, \tag{i}$$

where the components have been taken with respect to an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ aligned with the edges of the cube, see Figure 2.12. Derive relationships between the λ's in each of the following cases:

(a) The body is composed of an incompressible material.

(b) The length of the fiber $\overrightarrow{OP}$ remains unchanged by the deformation.

(c) The angle between the fibers $\overrightarrow{OP}$ and $\overrightarrow{QR}$ remains unchanged by the deformation.

(d) The area of the plane $RSQT$ remains unchanged by the deformation.

(e) The orientation of the plane $RSQT$ remains unchanged by the deformation.

Solution: The deformation gradient tensor and its inverse are

$$\mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \mathbf{F}^{-1} = \lambda_1^{-1} \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2^{-1} \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda 1_3 \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{ii}$$

(a) In general, the volumes of an infinitesimal part of the body in the reference and deformed configurations are related by $dV_y = J dV_x$, $J = \det \mathbf{F}$. If the material is incompressible then $dV_y = dV_x$ and so $J = 1$:

$$J = \det \mathbf{F} = \lambda_1 \lambda_2 \lambda_3 = 1. \qquad \square$$

(b) In general, the length ds_y of a deformed material fiber is given by (2.27) where ds_x and $\mathbf{m}_\mathrm{R}$ are the length and direction of the fiber in the reference configuration. Thus if the fiber does not change length, then $ds_x = ds_y$ and so the stretch $\lambda = 1$:

$$|\mathbf{F}\mathbf{m}_\mathrm{R}| = 1. \tag{iii}$$

The fiber of interest $\overrightarrow{\mathrm{OP}}$ can be expressed as $\overrightarrow{\mathrm{OP}} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$ and therefore the unit vector $\mathbf{m}_\mathrm{R}$ in the direction of $\overrightarrow{\mathrm{OP}}$ is

$$\mathbf{m}_\mathrm{R} = \frac{\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3}{\sqrt{3}}. \tag{iv}$$

Substituting $(ii)_1$ and (iv) into (iii) and simplifying leads to

$$\frac{\lambda_1^2}{3} + \frac{\lambda_2^2}{3} + \frac{\lambda_3^2}{3} = 1. \qquad \square$$

(c) In general, the angle θ_x between two material fibers that are in the directions of the unit vectors $\mathbf{m}_\mathrm{R}^{(1)}$ and $\mathbf{m}_\mathrm{R}^{(2)}$ in the reference configuration is given by (2.33) and the corresponding angle θ_y in the deformed configuration between these same two fibers is given by (2.35). Thus if the angle remains unchanged by the deformation we must have

$$\frac{\mathbf{F}\mathbf{m}_\mathrm{R}^{(1)}}{|\mathbf{F}\mathbf{m}^{(1)}|} \cdot \frac{\mathbf{F}\mathbf{m}^{(2)}}{|\mathbf{F}\mathbf{m}_\mathrm{R}^{(2)}|} = \mathbf{m}_\mathrm{R}^{(1)} \cdot \mathbf{m}_\mathrm{R}^{(2)} . \tag{v}$$

The unit vectors $\mathbf{m}_{\mathrm{R}}^{(1)}$ and $\mathbf{m}_{\mathrm{R}}^{(2)}$ in the directions of the material fibers $\overrightarrow{\mathrm{OP}}$ and $\overrightarrow{\mathrm{QR}}$ are

$$\mathbf{m}_{\mathrm{R}}^{(1)} = (\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)/\sqrt{3}, \qquad \mathbf{m}_{\mathrm{R}}^{(2)} = (\mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3)/\sqrt{3}. \tag{vi}$$

Thus substituting $(ii)_1$ and (vi) into (v) and simplifying leads to

$$2\lambda_1^2 = \lambda_2^2 + \lambda_3^2. \qquad \square$$

(d) In general, infinitesimal elements of area in the reference and deformed configurations are related by $dA_y = J|\mathbf{F}^{-T}\mathbf{n}_{\mathrm{R}}|\, dA_x$. If a particular area element remains unchanged, $dA_y = dA_x$, the deformation must be much that

$$J|\mathbf{F}^{-T}\mathbf{n}_{\mathrm{R}}| = 1, \tag{vii}$$

where $J = \det \mathbf{F}$ and $\mathbf{n}_{\mathrm{R}}$ is a unit vector normal to the surface of interest in the reference configuration. The unit vector normal to the plane RSQT is

$$\mathbf{n}_{\mathrm{R}} = \frac{1}{\sqrt{2}}(\mathbf{e}_1 + \mathbf{e}_2). \tag{$viii$}$$

Substituting $(ii)_2$ and $(viii)$ into (vii) and simplifying leads to

$$\lambda_1\lambda_2\lambda_3\left(\frac{1}{2\lambda_1^2} + \frac{1}{2\lambda_2^2}\right)^{1/2} = 1. \qquad \square$$

(e) In general, the unit vectors $\mathbf{n}_{\mathrm{R}}$ and $\mathbf{n}$ normal to a surface in the reference and deformed configurations are related by (2.40). If the orientation of this surface does not change, then $\mathbf{n}_{\mathrm{R}} = \mathbf{n}$ in which case (2.40) yields

$$\mathbf{n}_{\mathrm{R}} = \mathbf{F}^{-T}\mathbf{n}_{\mathrm{R}}/|\mathbf{F}^{-T}\mathbf{n}_{\mathrm{R}}|. \tag{ix}$$

Substituting $(ii)_2$ and $(viii)$ into (ix) and simplifying yields

$$\lambda_1 = \lambda_2, \qquad \square$$

(which is precisely what one would expect intuitively).

Problem 2.4.2. (Spencer) A body undergoes an arbitrary homogeneous deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$. Consider the set of particles that lie on a sphere of radius b in the deformed configuration. Show that in the undeformed configuration these particles lie on the surface of an ellipsoid and determine the lengths of the three major axes of the ellipsoid. Under what condition on $\mathbf{F}$ is this ellipsoid a sphere of radius a?

Solution: Let $\mathcal{S}$ denote the spherical surface of interest in the deformed configuration, and let $\mathcal{S}_{\mathrm{R}}$ be its image in the undeformed configuration. We pick the origin to be at the center of $\mathcal{S}$. Then the position vector $\mathbf{y}$ of a point on $\mathcal{S}$ and the radius b of $\mathcal{S}$ are related by

$$\mathcal{S}: \qquad \mathbf{y}\cdot\mathbf{y} = b^2. \tag{i}$$

Let $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ be a particle on $\mathcal{S}_{\rm R}$ and consider the deformation gradient tensor[9]

$$\mathbf{F} = \lambda_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2\mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3\mathbf{e}_3 \otimes \mathbf{e}_3.$$

The deformation takes $\mathbf{x} \to \mathbf{y}$ according to

$$\mathbf{y} = \mathbf{F}\mathbf{x} = (\lambda_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2\mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3\mathbf{e}_3 \otimes \mathbf{e}_3)(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3) = x_1\lambda_1\mathbf{e}_1 + x_2\lambda_2\mathbf{e}_2 + x_3\lambda_3\mathbf{e}_3. \qquad (ii)$$

Substituting (ii) into (i) gives

$$(x_1\lambda_1\mathbf{e}_1 + x_2\lambda_2\mathbf{e}_2 + x_3\lambda_3\mathbf{e}_3) \cdot (x_1\lambda_1\mathbf{e}_1 + x_2\lambda_2\mathbf{e}_2 + x_3\lambda_3\mathbf{e}_3) = b^2$$

which when simplified leads to

$$\mathcal{S}_{\rm R}: \qquad \frac{x_1^2}{b^2/\lambda_1^2} + \frac{x_2^2}{b^2/\lambda_2^2} + \frac{x_3^2}{b^2/\lambda_3^2} = 1.$$

Therefore the surface $\mathcal{S}_{\rm R}$ is an ellipsoid and the lengths of its (semi)-major axes are $b/\lambda_1, b/\lambda_2$ and b/λ_3.

The surface $\mathcal{S}_{\rm R}$ is a sphere of radius a if $\lambda_1 = \lambda_2 = \lambda_3 = b/a$.

Problem 2.4.3. The region $\mathcal{R}_R$ occupied by a body in a reference configuration is a hollow sphere. In spherical polar coordinates, a general deformation takes the particle located at (R, Θ, Φ) in the reference configuration into the location (r, θ, φ) in the deformed configuration. In a spherically *symmetric* deformation one has

$$r = r(R), \qquad \theta = \Theta, \qquad \varphi = \Phi. \qquad (o)$$

(a) In terms of $r(R)$, calculate the stretch of a material fiber in the radial direction? What is the stretch of a material fiber perpendicular to the radial direction?

(b) Determine the function $r(R)$ (to the extent possible) in each of the following cases:

  (b1) the material is incompressible,

  (b2) the material is inextensible in the radial direction (perhaps there are very stiff fibers in the radial direction),

  (b3) the material is inextensible in circumferential directions (perhaps there are very stiff fibers in the circumferential directions).

(c) Suppose that the inner and outer radii of the body are A and B in the reference configuration, and the inner radius is a in the deformed configuration. Calculate the outer radius of the body in each of the preceding cases?

Solution:

[9] Why did we take $\mathbf{F}$ to be a pure stretch?

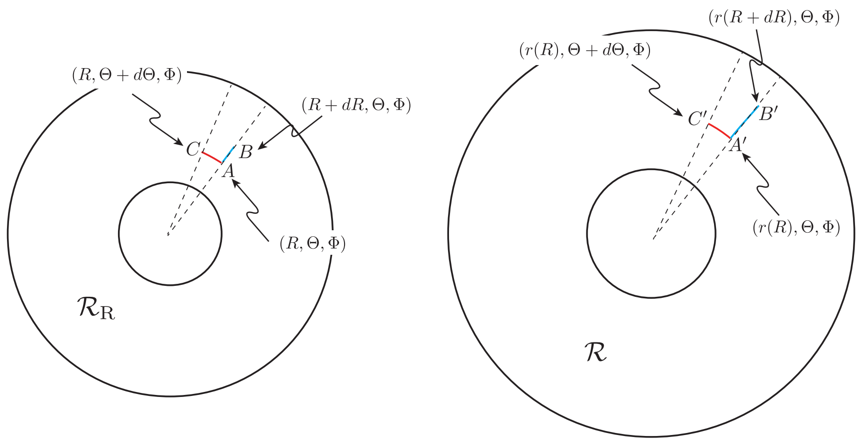


Figure 2.13: The radial and circumferential material fibers $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are mapped by the spherically symmetric deformation into $\overrightarrow{A'B'}$ and $\overrightarrow{A'C'}$ respectively. The spherical polar coordinates of the points A, B, C, A', B' and C' are shown in the figure.

(a) In order to calculate the stretch of a radial fiber, consider two particles located in the reference configuration at $A : (R, \Theta, \Phi)$ and $B : (R + dR, \Theta, \Phi)$ as shown in Figure 2.13. The spherically symmetric deformation (o) maps them to $A' : (r(R), \Theta, \Phi)$ and $B' : (r(R + dR), \Theta, \Phi)$ in the deformed configuration. Thus the deformation takes the infinitesimal radial material fiber

$$\overrightarrow{AB} = d\mathbf{x} = dR\,\mathbf{e}_R \quad \to \quad \overrightarrow{A'B'} = d\mathbf{y} = (r(R + dR) - r(R))\,\mathbf{e}_r,$$

and therefore the stretch of this fiber is

$$\lambda_R = \lim_{|AB|\to 0} \frac{|A'B'|}{|AB|} = \lim_{|d\mathbf{x}|\to 0} \frac{|d\mathbf{y}|}{|d\mathbf{x}|} = \lim_{dR\to 0} \frac{r(R + dR) - r(R)}{dR} = r\,'(R). \qquad \square \qquad (i)$$

Next, in order to calculate the stretch of a fiber perpendicular to the radial direction consider two particles located in the reference configuration at $A : (R, \Theta, \Phi)$ and $C : (R, \Theta + d\Theta, \Phi)$ as shown in Figure 2.13. The spherically symmetric deformation (o) maps them to $A' : (r(R), \Theta, \Phi)$ and $C' : (r(R), \Theta + d\Theta, \Phi)$ in the deformed configuration. The deformation takes the infinitesimal material fiber in the circumferential Θ direction

$$\overrightarrow{AC} = d\mathbf{x} = Rd\Theta\,\mathbf{e}_\Theta \quad \to \quad \overrightarrow{A'C'} = d\mathbf{y} = (r(R)d\Theta)\,\mathbf{e}_\theta,$$

and so the stretch of this fiber is

$$\lambda_\Theta = \lim_{|AC|\to 0} \frac{|A'C'|}{|AC|} = \lim_{|d\mathbf{x}|\to 0} \frac{|d\mathbf{y}|}{|d\mathbf{x}|} = \lim_{d\Theta\to 0} \frac{r(R)d\Theta}{Rd\Theta} = \frac{r(R)}{R}. \qquad \square \qquad (ii)$$

By symmetry, the stretch of a fiber in the circumferential Φ direction is

$$\lambda_\Phi = \lambda_\Theta = \frac{r(R)}{R}.$$

Note that the angle between any pair of the three material fibers considered above (the third being in the circumferential Φ direction) is $\pi/2$ in both the reference and deformed configurations. Therefore by Problem

2.5.1 the three stretches $\lambda_R, \lambda_\Theta$ and λ_Φ are in fact the principal stretches and the radial and circumferential directions are the principal directions (for both the Lagrangian and Eulerian stretch tensors).

($b1$) Incompressibility requires $\det \mathbf{F} = 1$:

$$\det \mathbf{F} = \lambda_R \lambda_\Theta \lambda_\Phi = \frac{r^2(R)}{R^2} r'(R) = 1,$$

which leads to the differential equation

$$r^2(R)\, r'(R) = R^2 \qquad \Rightarrow \qquad \frac{1}{3}\frac{d}{dR}\big(r(R)\big)^3 = R^2.$$

Integrating this yields

$$r(R) = \left[R^3 + c_1\right]^{1/3}, \qquad \square \tag{iv}$$

where c_1 is an arbitrary constant.

($b2$) Inextensibility in the radial direction requires $\lambda_R = 1$:

$$\lambda_R = r'(R) = 1 \qquad \Rightarrow \qquad r(R) = R + c_2, \qquad \square \tag{v}$$

where c_2 is an arbitrary constant.

($b3$) Inextensibility in the circumferential direction requires $\lambda_\Theta = 1$:

$$\lambda_\Theta = \frac{r(R)}{R} = 1 \qquad \Rightarrow \qquad r(R) = R. \qquad \square \tag{vi}$$

(c) Let b denote the (unknown) outer radius of the body in the deformed configuration. Then

$$r(A) = a, \qquad r(B) = b. \tag{vii}$$

In the incompressible case (iv) and (vii) give

$$b = \left[B^3 - A^3 + a^3\right]^{1/3}. \qquad \square$$

In the radially inextensible case (v) and (vii) give

$$b = B - A + a. \qquad \square$$

In the circumferentially inextensible case (vi) and (vii) give

$$b = B \qquad \text{(and in fact } a = A). \qquad \square$$

Remark: You could have deduced these values of b directly by physical considerations.

2.5 Stretch and rotation.

As mentioned previously, the deformation gradient tensor $\mathbf{F}(\mathbf{x})$ completely characterizes the deformation in the vicinity of the particle $\mathbf{x}$. Part of this deformation is a rigid rotation, the rest a "distortion", i.e. a "stretch/strain". We now explore this decomposition and examine various features of the rotation and stretch.

2.5.1 Right (or Lagrangian) Stretch Tensor U.

– According to the polar decomposition theorem stated in Section 1.4, every nonsingular tensor $\mathbf{F}$ can be written uniquely as the product of an orthogonal tensor $\mathbf{R}$ and a symmetric positive definite tensor $\mathbf{U}$ as

$$\boxed{\mathbf{F} = \mathbf{R}\,\mathbf{U}.} \tag{2.41}$$

When $\det \mathbf{F} > 0$, $\mathbf{R}$ is proper orthogonal and therefore represents a *rotation*. Since the tensor $\mathbf{U}$ is symmetric and positive definite, it describes a pure *stretch*; see the remarks below (2.7).

– Since $\mathbf{U}$ is symmetric and positive definite tensor we can express it in terms of its three real positive eigenvalues λ_1, λ_2 and λ_3 and corresponding orthonormal eigenvectors $\mathbf{r}_1, \mathbf{r}_2$ and $\mathbf{r}_3$ as

$$\mathbf{U} = \sum_{i=1}^{3} \lambda_i \mathbf{r}_i \otimes \mathbf{r}_i; \tag{2.42}$$

see (1.109). Equivalently, the matrix of components of $\mathbf{U}$ in its principal basis $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ is

$$[U] = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \tag{2.43}$$

which may be compared with the matrix[10] $[F]$ in (2.6) representing a pure stretch.

The tensor $\mathbf{U}$ is called the **right stretch tensor**[11] and the λ_i's are the **principal stretches**, the $\mathbf{r}_i$'s the corresponding **principal directions**. As we shall see shortly, $\mathbf{U}$ can be viewed as a **Lagrangian stretch tensor**.

– Let λ and $\mathbf{r}$ be one of the eigenvalues and corresponding eigenvectors of $\mathbf{U}$ and consider a referential material fiber that is in the direction $\mathbf{r}$: $d\mathbf{x} = dx\,\mathbf{r}$. When $\mathbf{U}$ operates on this fiber it is carried into $\mathbf{U}d\mathbf{x} = dx\,\mathbf{U}\mathbf{r} = \lambda dx\,\mathbf{r} = \lambda\,\mathrm{d}\mathbf{x}$ and so the stretch tensor $\mathbf{U}$ simply stretches this fiber by λ without rotation.

Now consider an infinitesimal rectangular parallelepiped of dimensions $dx_1 \times dx_2 \times dx_3$ with its edges aligned with the principal directions $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$. The tensor $\mathbf{U}$ carries it

[10] In equation (2.6), the deformation was homogeneous and so the matrix of components of $\mathbf{F}$ had this form at every point in the body. Here, (2.43) holds locally, at the point $\mathbf{x}$ under consideration.

[11] "Right" because $\mathbf{U}$ appears on the right-hand side of the expression $\mathbf{RU}$.

into a rectangular parallelepiped of dimensions $\lambda_1 dx_1 \times \lambda_2 dx_2 \times \lambda_3 dx_3$ (without rotating it). When $\mathbf{R}$ acts on it, this infinitesimal part of the body will rotate rigidly, and so in particular, the angle between the edges remains $\pi/2$.

– The deformation of a generic material fiber can also be viewed in two-steps:

$$d\mathbf{x} \quad \xrightarrow{\text{stretch}} \quad \mathbf{U}d\mathbf{x} \quad \xrightarrow{\text{rigid rotation}} \quad \mathbf{R}(\mathbf{U}d\mathbf{x}) = d\mathbf{y} \tag{2.44}$$

In the first part of the deformation where d$\mathbf{x} \to \mathbf{U}d\mathbf{x}$, the fiber $d\mathbf{x}$ is subjected to a pure stretch in the directions of the eigenvectors of $\mathbf{U}$, the amounts of stretch being equal to the eigenvalues of $\mathbf{U}$. Since $d\mathbf{x}$ will not be parallel to $\mathbf{U}d\mathbf{x}$ in general, the fiber will also rotate when it undergoes the stretching deformation $d\mathbf{x} \to \mathbf{U}d\mathbf{x}$. However, this is not a rigid rotation since the length of the fiber changes.

In the second step $\mathbf{U}d\mathbf{x} \to \mathbf{R}(\mathbf{U}d\mathbf{x})$, the stretched fiber is rigidly rotated by $\mathbf{R}$.

The stretch $\lambda(\mathbf{m}_{\mathrm{R}})$ of a fiber in an arbitrary direction $\mathbf{m}_{\mathrm{R}}$ can be written in terms of the principal stretches using (2.29) as

$$\lambda^2(\mathbf{m}_{\mathrm{R}}) = \lambda_1^2 m_1^2 + \lambda_2^2 m_2^2 + \lambda_3^2 m_3^2, \tag{2.45}$$

where the m_i's are the components of $\mathbf{m}_{\mathrm{R}}$ in a principal basis for $\mathbf{U}$, (Problem 2.19).

– Since the determinant of a proper orthogonal tensor is 1, it follows from $\det \mathbf{F} = \det(\mathbf{RU}) = \det \mathbf{R} \, \det \mathbf{U} = \det \mathbf{U}$ that

$$J = \det \mathbf{F} = \lambda_1 \lambda_2 \lambda_3. \tag{2.46}$$

– The right stretch tensor $\mathbf{U}$ is often referred to as the **Lagrangian stretch tensor**. This is because expressions for the length of a fiber, the angle between two fibers, area of a surface element etc. in the deformed configuration can be calculated in terms of just $\mathbf{U}$ and the referential geometry. For example, given a referential material fiber $d\mathbf{x}$, its length in the deformed configuration is $|d\mathbf{y}| = |\mathbf{F}\, d\mathbf{x}| = |\mathbf{RU}\, d\mathbf{x}| = |\mathbf{U}d\mathbf{x}|$ which shows that this length depends only on $\mathbf{U}$ and $d\mathbf{x}$ and not $\mathbf{R}$.

Exercise: By using $\mathbf{F} = \mathbf{RU}$ in (2.35), (2.36) and (2.39), show that changes in angle, volume and area can be expressed in terms of $\mathbf{U}$ and the referential geometry without involving $\mathbf{R}$.

On the other hand if we are given the fiber $d\mathbf{y}$ in the deformed configuration, we cannot calculate its undeformed length without knowing both $\mathbf{U}$ and $\mathbf{R}$. This follows by taking the magnitude of the vector equation $d\mathbf{x} = \mathbf{U}^{-1}\mathbf{R}^T d\mathbf{y}$. In this sense, $\mathbf{U}$ is not an Eulerian stretch tensor.

The principal directions $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ of $\mathbf{U}$ are known as the **Lagrangian principal directions** of stretch.

– *Principal scalar invariants.* Looking ahead, in Chapter 4.3 we will encounter (scalar-valued) functions[12] of the Lagrangian stretch tensor, $\varphi = \varphi(\mathbf{U})$, that have the property $\varphi(\mathbf{U}) = \varphi(\mathbf{Q}\mathbf{U}\mathbf{Q}^T)$ for all orthogonal tensors $\mathbf{Q}$. Such functions are called scalar-valued invariants (or isotropic functions). In Section 1.5 we saw that the **principal scalar invariants** of $\mathbf{U}$,

$$I_1(\mathbf{U}) = \operatorname{tr}\mathbf{U}, \qquad I_2(\mathbf{U}) = \frac{1}{2}\Big[\big(\operatorname{tr}\mathbf{U}\big)^2 - \operatorname{tr}\mathbf{U}^2\Big], \qquad I_3(\mathbf{U}) = \det\mathbf{U}, \tag{2.47}$$

have this invariance $I_i(\mathbf{U}) = I_i(\mathbf{Q}\mathbf{U}\mathbf{Q}^T)$.

One can show using (2.42) and (2.47) that the principal scalar invariants of $\mathbf{U}$ can be written in terms of the principal stretches as (Problem 1.16)

$$I_1(\mathbf{U}) = \lambda_1 + \lambda_2 + \lambda_3, \qquad I_2(\mathbf{U}) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1, \qquad I_3(\mathbf{U}) = \lambda_1\lambda_2\lambda_3. \tag{2.48}$$

– Finally we turn to the question of how, given $\mathbf{F}$, one might go about calculating $\mathbf{U}$ and $\mathbf{R}$. From (2.41) we have $\mathbf{F}^T\mathbf{F} = (\mathbf{R}\mathbf{U})^T\mathbf{R}\mathbf{U} = \mathbf{U}\mathbf{R}^T\mathbf{R}\mathbf{U} = \mathbf{U}^2$. The tensor $\mathbf{F}^T\mathbf{F}$ is symmetric and positive definite and therefore $\mathbf{U}$ is its unique, symmetric, positive definite square root (Problem 1.23):

$$\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}. \tag{2.49}$$

After determining the stretch $\mathbf{U}$ from (2.49) the rotation $\mathbf{R}$ can be found from

$$\mathbf{R} = \mathbf{F}\mathbf{U}^{-1}. \tag{2.50}$$

In Problem 2.5.2 we shall work out the details of this calculation for a simple shear deformation.

2.5.2 Left (or Eulerian) Stretch Tensor V.

– The alternative version of the polar decomposition theorem (Section 1.4) provides a second representation for $\mathbf{F}$. According to this part of the theorem, every tensor $\mathbf{F}$

[12]representing the energy stored per unit volume in an isotropic elastic body.

with det $\mathbf{F} > 0$ can be written uniquely as the product of a symmetric positive definite tensor $\mathbf{V}$ and a proper orthogonal tensor $\mathbf{R}$ as

$$\boxed{\mathbf{F} = \mathbf{V}\mathbf{R}.} \tag{2.51}$$

The rotation tensor $\mathbf{R}$ here is identical to that in the preceding representation and $\mathbf{V}$ is the unique, symmetric, positive definite square root

$$\mathbf{V} = \sqrt{\mathbf{F}\mathbf{F}^T}. \tag{2.52}$$

The tensor $\mathbf{V}$ is called the **left stretch tensor**[13] and as we shall see shortly, can be thought of as an **Eulerian stretch tensor**. The principal values of $\mathbf{V}$ are the same as those of $\mathbf{U}$. The stretch tensor $\mathbf{V}$ can be expressed as

$$\mathbf{V} = \sum_{i=1}^{3} \lambda_i \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i, \tag{2.53}$$

where $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ are the principal directions of $\mathbf{V}$. The principal directions of $\mathbf{U}$ and $\mathbf{V}$ are related by

$$\boldsymbol{\ell}_i = \mathbf{R}\mathbf{r}_i. \tag{2.54}$$

The deformation of a generic fiber can now be written as

$$d\mathbf{y} = \mathbf{V}\,(\mathbf{R}\,d\mathbf{x}), \tag{2.55}$$

and so interpreted as a rigid rotation, $d\mathbf{x} \to \mathbf{R}\,d\mathbf{x}$, followed by stretching in the directions $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ by $\lambda_1, \lambda_2, \lambda_3$.

– The tensors $\mathbf{F}$ and $\mathbf{R}$ have the following representations (Problem 1.24):

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \boldsymbol{\ell}_i \otimes \mathbf{r}_i, \qquad \mathbf{R} = \sum_{i=1}^{3} \boldsymbol{\ell}_i \otimes \mathbf{r}_i. \tag{2.56}$$

Since $\mathbf{F}\mathbf{r}_i = \lambda_i \boldsymbol{\ell}_i$ (no sum on i), it follows that a fiber in the direction $\mathbf{r}_i$ in the reference configuration is carried by $\mathbf{F}$ into the direction $\boldsymbol{\ell}_i$.

Exercise: Show that

$$\mathbf{F}^{-1} = \sum_{i=1}^{3} \lambda_i^{-1} \mathbf{r}_i \otimes \boldsymbol{\ell}_i. \tag{2.57}$$

[13] "Left" because $\mathbf{V}$ appears on the left-hand side of the expression $\mathbf{VR}$.

The left stretch tensor $\mathbf{V}$ is also called the **Eulerian stretch tensor**. Given a material fiber $d\mathbf{y}$ in the deformed configuration, its length in the reference configuration can be expressed as $|\mathbf{F}^{-1}\,d\mathbf{y}| = |(\mathbf{VR})^{-1}\,d\mathbf{y}| = |\mathbf{R}^T\mathbf{V}^{-1}\,d\mathbf{y}| = |\mathbf{V}^{-1}d\mathbf{y}|$ showing that it can be calculated in terms of just $\mathbf{V}$ and $d\mathbf{y}$ without involving $\mathbf{R}$.

The principal directions $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ of $\mathbf{V}$ are known as the **Eulerian principal directions** of stretch. Since

$$\mathbf{F}\mathbf{r}_i = \mathbf{R}\mathbf{U}\mathbf{r}_i = \lambda_i\mathbf{R}\mathbf{r}_i = \lambda_i\boldsymbol{\ell}_i \qquad \text{(no sum on } i\text{)}$$

it follows that a referential material fiber in a Lagrangian principal direction is mapped by the deformation into a fiber in an Eulerian principal direction.

– The principal scalar invariants of $\mathbf{U}$ and $\mathbf{V}$ coincide: $I_i(\mathbf{U}) = I_i(\mathbf{V})$.

2.5.3 Cauchy–Green deformation tensors.

– In Problem 2.5.2 where we calculate the Lagrangian stretch tensor $\mathbf{U}$ associated with a simple shear deformation, we will see that this calculation is quite tedious, mainly because of having to find the square root of $\mathbf{F}^T\mathbf{F}$. However, since there is a one-to-one relation between $\mathbf{U}$ and $\mathbf{U}^2$, and similarly between $\mathbf{V}$ and $\mathbf{V}^2$, we can just as well use $\mathbf{U}^2$ and $\mathbf{V}^2$ as our measures of stretch and these are much easier to calculate. These two tensors, usually denoted by $\mathbf{C}$ and $\mathbf{B}$,

$$\boxed{\mathbf{C} := \mathbf{F}^T\mathbf{F} = \mathbf{U}^2, \qquad \mathbf{B} := \mathbf{F}\mathbf{F}^T = \mathbf{V}^2,} \tag{2.58}$$

are referred to as the **right** and **left Cauchy-Green deformation tensors** respectively. They represent Lagrangian and Eulerian measures of the deformation. Observe from (2.29) that

$$\lambda^2 = \mathbf{F}\mathbf{m}_{\mathrm{R}} \cdot \mathbf{F}\mathbf{m}_{\mathrm{R}} \overset{(1.72)}{=} \mathbf{F}^T\mathbf{F}\mathbf{m}_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}} = \mathbf{C}\mathbf{m}_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}, \tag{2.59}$$

where $\mathbf{m}_{\mathrm{R}}$ is the direction of a material fiber in the reference configuration and λ is its stretch. Likewise, from $\lambda\mathbf{F}^{-1}\mathbf{m} = \mathbf{m}_{\mathrm{R}}$ it follows that

$$\frac{1}{\lambda^2} = |\mathbf{F}^{-1}\mathbf{m}|^2 = (\mathbf{F}^{-1}\mathbf{m}) \cdot (\mathbf{F}^{-1}\mathbf{m}) \overset{(1.72)}{=} (\mathbf{F}^{-T}\mathbf{F}^{-1}\mathbf{m}) \cdot \mathbf{m} = \mathbf{B}^{-1}\mathbf{m} \cdot \mathbf{m}, \tag{2.60}$$

where $\mathbf{m}$ is the direction of a material fiber in the deformed configuration and λ is its stretch. Exercise: Equivalently, show that

$$|d\mathbf{y}|^2 - |d\mathbf{x}|^2 = (\mathbf{C} - \mathbf{I})d\mathbf{x} \cdot d\mathbf{x}, \qquad |d\mathbf{y}|^2 - |d\mathbf{x}|^2 = (\mathbf{I} - \mathbf{B}^{-1})d\mathbf{y} \cdot d\mathbf{y}. \tag{2.61}$$

Note that the eigenvalues of $\mathbf{C}$ and $\mathbf{B}$ are λ_1^2, λ_2^2 and λ_3^2, where the λ_i's are the principal stretches, and the eigenvectors of $\mathbf{C}$ and $\mathbf{B}$ are the same as those of $\mathbf{U}$ and $\mathbf{V}$ respectively. Thus the two Cauchy-Green tensors admit the spectral representations

$$\mathbf{C} = \sum_{i=1}^{3} \lambda_i^2 \, (\mathbf{r}_i \otimes \mathbf{r}_i), \qquad \mathbf{B} = \sum_{i=1}^{3} \lambda_i^2 \, (\boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i). \tag{2.62}$$

– The principal scalar invariants of $\mathbf{C} = \mathbf{U}^2$ can be written in terms of the principal stretches as

$$I_1(\mathbf{C}) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \qquad I_2(\mathbf{C}) = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2, \qquad I_3(\mathbf{C}) = \lambda_1^2\lambda_2^2\lambda_3^2. \tag{2.63}$$

The principal scalar invariants of $\mathbf{B}$ and $\mathbf{C}$ coincide: $I_i(\mathbf{C}) = I_i(\mathbf{B})$.

2.5.4 Worked examples.

Problem 2.5.1. Consider three referential material fibers oriented in mutually orthogonal directions $\mathbf{m}_{\mathrm{R}}^{(1)}, \mathbf{m}_{\mathrm{R}}^{(2)}, \mathbf{m}_{\mathrm{R}}^{(3)}$. If the angle between each pair of these fibers remains $\pi/2$ in the deformed configuration, show that these directions are principal directions of the Lagrangian stretch tensor $\mathbf{U}$.

Solution: Since a fiber in the $\mathbf{m}_{\mathrm{R}}^{(1)}$ direction remains orthogonal to fibers in the $\mathbf{m}_{\mathrm{R}}^{(2)}$ and $\mathbf{m}_{\mathrm{R}}^{(3)}$ directions it follows from (2.35) that

$$\mathbf{F}\mathbf{m}_{\mathrm{R}}^{(1)} \cdot \mathbf{F}\mathbf{m}_{\mathrm{R}}^{(2)} = 0, \qquad \mathbf{F}\mathbf{m}_{\mathrm{R}}^{(1)} \cdot \mathbf{F}\mathbf{m}_{\mathrm{R}}^{(3)} = 0.$$

On using (1.72) and (2.58) this tells us that

$$\mathbf{C}\mathbf{m}_{\mathrm{R}}^{(1)} \cdot \mathbf{m}_{\mathrm{R}}^{(2)} = 0, \qquad \mathbf{C}\mathbf{m}_{\mathrm{R}}^{(1)} \cdot \mathbf{m}_{\mathrm{R}}^{(3)} = 0,$$

where $\mathbf{C} = \mathbf{F}^T\mathbf{F}$. Therefore the vector $\mathbf{C}\mathbf{m}_{\mathrm{R}}^{(1)}$ is perpendicular to both $\mathbf{m}_{\mathrm{R}}^{(2)}$ and $\mathbf{m}_{\mathrm{R}}^{(3)}$ and so must be parallel to $\mathbf{m}_{\mathrm{R}}^{(1)}$. Thus $\mathbf{C}\mathbf{m}_{\mathrm{R}}^{(1)} = \gamma\mathbf{m}_{\mathrm{R}}^{(1)}$ for some scalar γ from which it follows that $\mathbf{m}_{\mathrm{R}}^{(1)}$ is an eigenvector of the right Cauchy-Green tensor $\mathbf{C} = \mathbf{U}^2$ and therefore of $\mathbf{U}$.

That $\mathbf{m}_{\mathrm{R}}^{(2)}$ and $\mathbf{m}_{\mathrm{R}}^{(3)}$ are principal directions of $\mathbf{U}$ follows similarly.

Problem 2.5.2. (See also Problems 2.35 and 2.36.) Calculate the principal stretches $\lambda_1, \lambda_2, \lambda_3$ and principal Lagrangian stretch directions $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ associated with the simple shear deformation

$$y_1 = x_1 + kx_2, \quad y_2 = x_2, \quad y_3 = x_3; \qquad k > 0. \tag{i}$$

Here the components have been taken with respect to a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

Graphically illustrate the simple shear deformation in the form $\mathbf{y} = \mathbf{F}\mathbf{x} = \mathbf{R}(\mathbf{U}\mathbf{x})$.

Solution: Since[14]

$$\mathbf{U} = \lambda_1 \mathbf{r}_1 \otimes \mathbf{r}_2 + \lambda_2 \mathbf{r}_2 \otimes \mathbf{r}_2 + \lambda_3 \mathbf{r}_3 \otimes \mathbf{r}_3, \qquad \mathbf{C} = \lambda_1^2 \mathbf{r}_1 \otimes \mathbf{r}_2 + \lambda_2^2 \mathbf{r}_2 \otimes \mathbf{r}_2 + \lambda_3^2 \mathbf{r}_3 \otimes \mathbf{r}_3, \tag{ii}$$

to find the λ_i 's and $\mathbf{r}_i$'s we solve the eigenvalue problem for $\mathbf{C} = \mathbf{F}^T\mathbf{F} = \mathbf{U}^2$.

It follows from (i) and $F_{ij} = \partial y_i/\partial x_j$ that the deformation gradient tensor is

$$\mathbf{F} = \mathbf{I} + k\mathbf{e}_1 \otimes \mathbf{e}_2, \tag{iii}$$

and therefore from $\mathbf{C} = \mathbf{F}^T\mathbf{F}$,

$$\mathbf{C} = (\mathbf{I} + k\mathbf{e}_2 \otimes \mathbf{e}_1)(\mathbf{I} + k\mathbf{e}_1 \otimes \mathbf{e}_2) = \mathbf{e}_1 \otimes \mathbf{e}_1 + k(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + (1+k^2)\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{vi}$$

The eigenvalues of $\mathbf{C}$ are the roots λ of the equation $\det[\mathbf{C} - \lambda^2\mathbf{I}] = 0$:

$$\det[\mathbf{C} - \lambda^2\mathbf{I}] \;=\; \det\begin{pmatrix} 1-\lambda^2 & k & 0 \\ k & 1+k^2-\lambda^2 & 0 \\ 0 & 0 & 1-\lambda^2 \end{pmatrix} = (1-\lambda^2)(\lambda^4 - (2+k^2)\lambda^2 + 1) = 0. \tag{vii}$$

The roots of this equation are

$$\lambda_1^2 = \frac{2+k^2+k\sqrt{k^2+4}}{2}\ (\geq 1), \qquad \lambda_2^2 = \frac{2+k^2-k\sqrt{k^2+4}}{2}\ (\leq 1), \qquad \lambda_3^2 = 1. \tag{viii}$$

The eigenvector $\mathbf{r} = r_1\mathbf{e}_1 + r_2\mathbf{e}_2 + r_3\mathbf{e}_3$ corresponding to the eigenvalue λ^2 is given by $(\mathbf{C} - \lambda^2\mathbf{I})\mathbf{r} = \mathbf{o}$:

$$\begin{pmatrix} 1-\lambda^2 & k & 0 \\ k & 1+k^2-\lambda^2 & 0 \\ 0 & 0 & 1-\lambda^2 \end{pmatrix}\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{ix}$$

For each $\lambda = \lambda_i$ this can be solved for r_1, r_2, r_3 thus leading to the eigenvectors

$$\mathbf{r}_1 = \cos\theta_r\ \mathbf{e}_1 + \sin\theta_r\ \mathbf{e}_2, \qquad \mathbf{r}_2 = -\sin\theta_r\ \mathbf{e}_1 + \cos\theta_r\ \mathbf{e}_2, \qquad \mathbf{r}_3 = \mathbf{e}_3, \tag{x}$$

where we have set

$$\tan 2\theta_r = -\frac{2}{k}, \qquad \frac{\pi}{4} \leq \theta_r < \frac{\pi}{2}. \tag{xii}$$

The angle θ_r is depicted in Figure 2.14. For the reasons given in the footnote on page 134, we anticipated $\lambda_3 = 1$ and $\mathbf{r}_3 = \mathbf{e}_3$. To find the principal stretches we simply take the square roots of $(viii)$:

$$\lambda_1 = \frac{\sqrt{k^2+4}+k}{2}\ (\geq 1), \qquad \lambda_2 = \frac{\sqrt{k^2+4}-k}{2}\ (\leq 1), \qquad \lambda_3 = 1. \qquad \square \tag{xiii}$$

We may now visualize the simple shear deformation $\mathbf{y} = \mathbf{F}\mathbf{x} = \mathbf{R}(\mathbf{U}\mathbf{x})$ in two steps as follows: First, the deformation $\mathbf{x} \to \mathbf{U}\mathbf{x}$ stretches the square $OABC$ in Figure 2.15 by the amounts λ_1, λ_2 in the principal directions $\mathbf{r}_1, \mathbf{r}_2$ leading to the region $OA'B'C'$. This is then followed by the deformation $\mathbf{U}\mathbf{x} \to \mathbf{R}(\mathbf{U}\mathbf{x})$ which rigidly rotates $OA'B'C'$ into the region OA_*B_*C which is the region occupied by the deformed body.

[14]Since this simple shear is a planar deformation in the x_1, x_2-plane with no stretch in the x_3-direction, one of the principal directions, say $\mathbf{r}_3$, will be $\mathbf{e}_3$ and the corresponding principal stretch $\lambda_3 = 1$. Moreover, since a simple shear is isochoric, $\det\mathbf{U} = \lambda_1\lambda_2\lambda_3 = \lambda_1\lambda_2 = 1$. Therefore one knows a priori that $\mathbf{U} = \lambda_1\mathbf{r}_1 \otimes \mathbf{r}_1 + \lambda_1^{-1}\mathbf{r}_2 \otimes \mathbf{r}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3$.

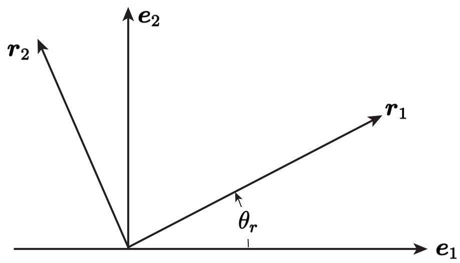


Figure 2.14: Principal directions $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ of the (right) Lagrangian stretch tensor $\mathbf{U}$. When $k \to 0$, the angle $\theta_r \to \pi/4$.

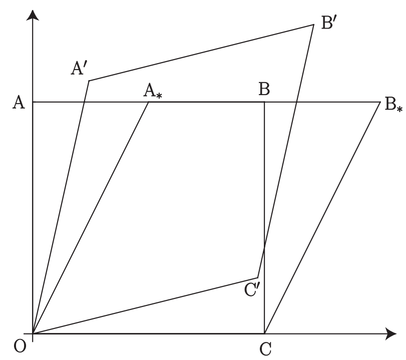


Figure 2.15: Simple shear deformation viewed in two steps: $\mathbf{y} = \mathbf{F}\mathbf{x} = \mathbf{R}(\mathbf{U}\mathbf{x})$. The pure stretch $\mathbf{x} \to \mathbf{U}\mathbf{x}$ takes the region $OABC \to O'A'B'C'$ and the subsequent rotation $\mathbf{U}\mathbf{x} \to \mathbf{R}(\mathbf{U}\mathbf{x})$ takes $OA'B'C' \to OA_*B_*C$.

Remark: The tensor $\mathbf{U}$ can be expressed with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by substituting (x) into $(ii)_1$ which leads to

$$\mathbf{U} = \frac{1}{\sqrt{4+k^2}}\Big(2\mathbf{e}_1 \otimes \mathbf{e}_1 + k(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + (2+k^2)\,\mathbf{e}_2 \otimes \mathbf{e}_2\Big) + \mathbf{e}_3 \otimes \mathbf{e}_3. \qquad (xvi)$$

The rotation tensor $\mathbf{R}$ can now be calculated using $\mathbf{R} = \mathbf{F}\mathbf{U}^{-1}$ which leads to

$$\mathbf{R} = \frac{1}{\sqrt{4+k^2}}\Big(2\mathbf{e}_1 \otimes \mathbf{e}_1 + k\mathbf{e}_1 \otimes \mathbf{e}_2 - k\mathbf{e}_2 \otimes \mathbf{e}_1 + 2\,\mathbf{e}_2 \otimes \mathbf{e}_2\Big) + \mathbf{e}_3 \otimes \mathbf{e}_3. \qquad (xix)$$

Problem 2.5.3. The region $\mathcal{R}_\mathrm{R}$ occupied by a body in a reference configuration is the unit cube

$$\mathcal{R}_\mathrm{R} = \{(x_1, x_2, x_3)\ :\ -1/2 \le x_1 \le 1/2,\ -1/2 \le x_2 \le 1/2,\ -1/2 \le x_3 \le 1/2\}.$$

The basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are parallel to the edges of the cube and all components are taken with respect to this basis. The body is subjected to the homogeneous deformation

$$\{y\} = [F]\{x\} \qquad \text{where} \qquad [F] = \begin{pmatrix} -a & 0 & 0 \\ 0 & -b & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad a > 0,\ b > 0, \qquad (i)$$

that takes the particle at $(x_1, x_2, x_3) \in \mathcal{R}_{\rm R}$ to $(y_1, y_2, y_3) \in \mathcal{R}$. The negative signs are not typos.

(a) Calculate the components of the stretch tensors $\mathbf{U}, \mathbf{V}$ and the rotation tensor $\mathbf{R}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

(b) Determine the principal stretches and principal directions of $\mathbf{U}$.

(c) Sketch the region $\mathcal{R}$ occupied by the body in the deformed configuration.

Solution: Observe that $[F] = [F]^T$ and therefore $[V]^2 = [B] = [F][F]^T = [F]^2$ and likewise $[U]^2 = [C] = [F]^T[F] = [F]^2$. Thus

$$[U]^2 = [V]^2 = \begin{pmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \Rightarrow \qquad [U] = [V] = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \square \tag{ii}$$

having made use of $a > 0$ and $b > 0$. Observe that $[U]$ and $[V]$ are symmetric and positive definite. The principal stretches are $a > 0, b > 0$ and 1 and the corresponding principal directions (of both $\mathbf{U}$ and $\mathbf{V}$) are $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$. $\square$

The components of the rotation tensor are

$$[R] = [F][U]^{-1} = \begin{pmatrix} -a & 0 & 0 \\ 0 & -b & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad \square \tag{iii}$$

The matrix $[R]$ is proper orthogonal ($\det[R] = +1$). It represents a rotation through an angle π with the axis of rotation being $\mathbf{e}_3$ as shown in Figure 2.16.

In Figure 2.16, the points A, B, C and D in the reference configuration have the respective coordinates $(-1/2, -1/2, 0), (1/2, -1/2, 0), (1/2, 1/2, 0)$ and $(-1/2, 1/2, 0)$. The deformation maps them into the points A′, B′, C′ and D′ in the deformed configuration with coordinates $(a/2, b/2, 0), (-a/2, b/2, 0), (-a/2, -b/2, 0)$ and $(a/2, -b/2, 0)$. Observe the stretches by a and b in the $\mathbf{e}_1$- and $\mathbf{e}_2$-directions and the rotation through π about $\mathbf{e}_3$.

Problem 2.5.4. Bending of a block.

A body occupies the rectangular parallelepiped region $\mathcal{R}_{\rm R} = \{(x_1, x_2, x_3) \mid -A \le x_1 \le A, -B \le x_2 \le B, -C \le x_3 \le C\}$ in a reference configuration. The left-hand figure in Figure 2.17 shows a side view of this block looking down the x_3-axis. The block is subjected to a bending deformation in the x_1, x_2-plane that carries a generic particle from (x_1, x_2, x_3) to (y_1, y_2, y_3) and the region $\mathcal{R}_{\rm R} \to \mathcal{R}$ as shown in Figure 2.17. Specifically, the deformation has the following characteristics:

– The body stretches uniformly in the x_3-direction in the sense that $y_3 = \Lambda x_3$ for some positive constant Λ.

– Every plane $x_3 = \mathsf{constant}$ in $\mathcal{R}_{\rm R}$ deforms identically.

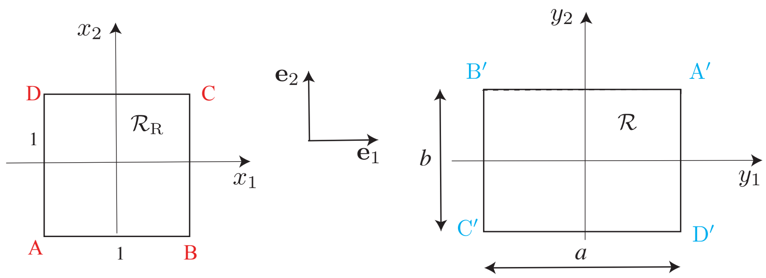


Figure 2.16: Figure for Problem 2.5.3: Mid-plane of the body looking down the x_3-axis. The body has been stretched by a, b and 1 in the $\mathbf{e}_1$-, $\mathbf{e}_2$- and $\mathbf{e}_3$-directions and rotated through an angle π about the $\mathbf{e}_3$-direction.

- Each horizontal straight line $x_2 = \mathsf{constant}$ is carried into a straight line in the deformed configuration, e.g. MN $\to$ M′N′. Moreover (on each plane $x_3 = \mathsf{constant}$) the family of such straight lines corresponding to the various values of x_2 all pass through the same point $(y_1, y_2) = (0, 0)$ as depicted in Figure 2.17.
- Each vertical straight line $x_1 = \mathsf{constant}$ is deformed into a circular arc centered at $(0, 0)$ as shown in Figure 2.17, e.g. PQ $\to$ P′Q′.
- The deformation is symmetric with respect to the x_1, x_3-plane.

(a) Given the shapes of the regions $\mathcal{R}$ and $\mathcal{R}_{\rm R}$, and the nature of the bending deformation described, it is natural to use rectangular cartesian coordinates (x_1, x_2, x_3) and cylindrical polar coordinates (r, θ, z) to describe the undeformed and deformed configurations respectively, with associated basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. Thus express the deformation as

$$\left.\begin{aligned} y_1 &= r(x_1, x_2, x_3)\cos\theta(x_1, x_2, x_3), \\ y_2 &= r(x_1, x_2, x_3)\sin\theta(x_1, x_2, x_3), \\ y_3 &= \Lambda x_3, \end{aligned}\right\} \qquad (i)$$

and determine the form of the functions $r(x_1, x_2, x_3)$ and $\theta(x_1, x_2, x_3)$ to the extent possible.

(b) Calculate the deformation gradient tensor $\mathbf{F}$. (Since $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ are the natural bases to use when characterizing the reference and deformed configurations, you will find that the natural representation for $\mathbf{F}$ is with respect to the tensor basis $\mathbf{e}_r \otimes \mathbf{e}_1, \mathbf{e}_r \otimes \mathbf{e}_2, \ldots$.)

(c) By factoring $\mathbf{F}$ appropriately determine the stretch tensors $\mathbf{U}$ and $\mathbf{V}$ and the rotation tensor $\mathbf{R}$.

(d) What are the principal stretches and the principal Eulerian and Lagrangian directions of stretch?

(e) Specialize your answer from part (a) to the case where the deformation is isochoric.

Solution:

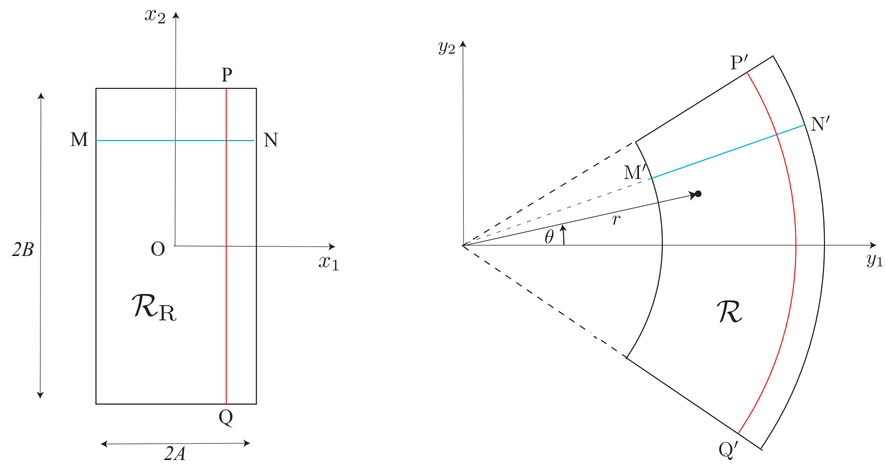


Figure 2.17: Left: In a reference configuration the body occupies the rectangular parallelepiped region $\mathcal{R}_\mathrm{R} = \{(x_1, x_2, x_3) : -A \le x_1 \le A, -B \le x_2 \le B, -C \le x_3 \le C\}$. Right: The region $\mathcal{R}$ in the deformed configuration. A vertical straight line in the reference configuration, e.g. PQ, is mapped into a circular arc in the deformed configuration, e.g. P′Q′. A horizontal straight line in the reference configuration, e.g. MN, is carried into a radial straight line, e.g. M′N′.

(a) Since every plane $x_3 =$ constant deforms identically, it follows that $r(x_1, x_2, x_3)$ and $\theta(x_1, x_2, x_3)$ must be independent of x_3 and so we can write the deformation as

$$y_1 = r(x_1, x_2)\cos\theta(x_1, x_2), \quad y_2 = r(x_1, x_2)\sin\theta(x_1, x_2), \quad y_3 = \Lambda x_3. \tag{ii}$$

– As one moves from M towards N along the horizontal straight line MN, the coordinate x_2 remains constant while x_1 increases. Thus considering its image M′N′, (the particle label) x_1 increases as one moves from M′ towards N′. However the orientation of M′N′, i.e. the angle $\theta(x_1, x_2)$, does not change. It follows that $\theta(x_1, x_2)$ cannot depend on x_1 and so

$$\theta = \theta(x_2). \tag{iii}$$

In view of symmetry,

$$\theta(x_2) = -\theta(-x_2). \tag{iv}$$

– Similarly, as one moves along the vertical straight line PQ, the coordinate x_1 remains constant while x_2 varies. Thus considering its image P′Q′, (the particle label) x_2 varies along it while its radius $r(x_1, x_2)$ does not change. It follows that $r(x_1, x_2)$ must be independent of x_2 and so

$$r = r(x_1). \tag{v}$$

Thus the bending deformation described in the problem is characterized by

$$y_1 = r(x_1)\cos\theta(x_2), \qquad y_2 = r(x_1)\sin\theta(x_2), \qquad y_3 = \Lambda x_3, \qquad \square \tag{vi}$$

where $r(x_1) > 0$ and $\theta(x_2) \in [-\pi, \pi]$ are arbitrary functions (with θ being an odd function).

(*b*) *Method 1:* The components of $\mathbf{F}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ can be calculated by differentiating (vi) with respect to the x_1, x_2 and x_3 and using $F_{ij} = \partial y_i/\partial x_j$. This yields

$$\mathbf{F} = r' \cos\theta\, \mathbf{e}_1 \otimes \mathbf{e}_1 - r\theta' \sin\theta\, \mathbf{e}_1 \otimes \mathbf{e}_2 + r' \sin\theta\, \mathbf{e}_2 \otimes \mathbf{e}_1 + r\theta' \cos\theta\, \mathbf{e}_2 \otimes \mathbf{e}_2 + \Lambda\, \mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \square \qquad (vii)$$

This representation of $\mathbf{F}$ does not provide much insight.

Method 2: It is more natural to calculate the components of the deformation gradient tensor with respect to the mixed basis $\mathbf{e}_r \otimes \mathbf{e}_1, \mathbf{e}_r \otimes \mathbf{e}_2, \ldots, \mathbf{e}_z \otimes \mathbf{e}_3$. We shall do this using the basic relation $d\mathbf{y} = \mathbf{F}d\mathbf{x}$. First note that

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \qquad \Rightarrow \qquad d\mathbf{x} = dx_1\, \mathbf{e}_1 + dx_2\, \mathbf{e}_2 + dx_3\, \mathbf{e}_3, \qquad (viii)$$

$$\mathbf{y} = r\mathbf{e}_r + z\mathbf{e}_z \qquad \Rightarrow \qquad d\mathbf{y} = dr\, \mathbf{e}_r + r\, d\mathbf{e}_r + dz\, \mathbf{e}_z, \qquad (ix)$$

where in calculating $d\mathbf{y}$ we had to keep in mind that the basis vector $\mathbf{e}_r$ is not a constant but rather $\mathbf{e}_r = \mathbf{e}_r(\theta)$. The basis vectors $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ are related to the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by

$$\mathbf{e}_r(\theta) = \cos\theta\, \mathbf{e}_1 + \sin\theta\, \mathbf{e}_2, \qquad \mathbf{e}_\theta(\theta) = -\sin\theta\, \mathbf{e}_1 + \cos\theta\, \mathbf{e}_2, \qquad \mathbf{e}_z = \mathbf{e}_3, \qquad (x)$$

and therefore

$$d\mathbf{e}_r = -\sin\theta\, d\theta\, \mathbf{e}_1 + \cos\theta\, d\theta\, \mathbf{e}_2 = (-\sin\theta\, \mathbf{e}_1 + \cos\theta\, \mathbf{e}_2)d\theta = \mathbf{e}_\theta\, d\theta. \qquad (xi)$$

Substituting (xi) into (ix) gives

$$d\mathbf{y} = dr\, \mathbf{e}_r + r d\theta\, \mathbf{e}_\theta + dz\, \mathbf{e}_z.$$

Equation (vi) tells us that $r = r(x_1)$, $\theta = \theta(x_2)$ and $z = \Lambda x_3$ and so this can be written as

$$d\mathbf{y} = r' dx_1\, \mathbf{e}_r + r\theta' dx_2\, \mathbf{e}_\theta + \Lambda dx_3\, \mathbf{e}_z,$$

where a prime denotes differentiation with respect to the argument. Finally using $(viii)_2$ in the preceding equation gives

$$\begin{aligned} d\mathbf{y} &= r'(d\mathbf{x} \cdot \mathbf{e}_1)\, \mathbf{e}_r + r\theta'(d\mathbf{x} \cdot \mathbf{e}_2)\, \mathbf{e}_\theta + \Lambda(d\mathbf{x} \cdot \mathbf{e}_3)\, \mathbf{e}_z = \\ &= r'(\mathbf{e}_r \otimes \mathbf{e}_1)\, d\mathbf{x} + r\theta'(\mathbf{e}_\theta \otimes \mathbf{e}_2)\, d\mathbf{x} + \Lambda(\mathbf{e}_z \otimes \mathbf{e}_3)d\mathbf{x} \\ &= \Big[r'(\mathbf{e}_r \otimes \mathbf{e}_1) + r\theta'(\mathbf{e}_\theta \otimes \mathbf{e}_2) + \Lambda(\mathbf{e}_z \otimes \mathbf{e}_3)\Big]d\mathbf{x} \\ &= \mathbf{F}\, d\mathbf{x}, \end{aligned}$$

where

$$\mathbf{F} = r'(x_1)\, (\mathbf{e}_r \otimes \mathbf{e}_1) + r\theta'(x_2)\, (\mathbf{e}_\theta \otimes \mathbf{e}_1) + \Lambda\, (\mathbf{e}_z \otimes \mathbf{e}_1). \qquad \square \qquad (xii)$$

Remark: Observe that we can write (vii) as

$$\mathbf{F} = r'(\mathbf{e}_1 \cos\theta + \mathbf{e}_2 \sin\theta) \otimes \mathbf{e}_1 + r\theta'(-\mathbf{e}_1 \sin\theta + \mathbf{e}_2 \cos\theta) \otimes \mathbf{e}_2 + \Lambda\, \mathbf{e}_3 \otimes \mathbf{e}_3.$$

Substituting (x) into this gives (xii).

(*c*) The expression (xii) for $\mathbf{F}$ can be factored in two ways[15]:

$$\mathbf{F} = (\mathbf{e}_r \otimes \mathbf{e}_1 + \mathbf{e}_\theta \otimes \mathbf{e}_2 + \mathbf{e}_z \otimes \mathbf{e}_3)(r'\mathbf{e}_1 \otimes \mathbf{e}_1 + r\theta'\mathbf{e}_2 \otimes \mathbf{e}_2 + \Lambda\mathbf{e}_3 \otimes \mathbf{e}_3), \qquad (xiii)$$

[15] Alternatively one can calculate $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ and $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ from (xii) and then calculate $\mathbf{U}$ and $\mathbf{V}$ from the results.

$$\mathbf{F} = (r'\mathbf{e}_r \otimes \mathbf{e}_r + r\theta'\mathbf{e}_\theta \otimes \mathbf{e}_\theta + \Lambda\mathbf{e}_z \otimes \mathbf{e}_z)(\mathbf{e}_r \otimes \mathbf{e}_1 + \mathbf{e}_\theta \otimes \mathbf{e}_2 + \mathbf{e}_z \otimes \mathbf{e}_3). \qquad (xiv)$$

Therefore the rotation tensor $\mathbf{R}$ in the polar decomposition is

$$\mathbf{R} = \mathbf{e}_r \otimes \mathbf{e}_1 + \mathbf{e}_\theta \otimes \mathbf{e}_2 + \mathbf{e}_z \otimes \mathbf{e}_3, \qquad \square \qquad (xv)$$

which is in fact the rotation that takes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ into $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. Then, from $\mathbf{F} = \mathbf{RU}$ and $(xiii)$ we find

$$\mathbf{U} = r'\mathbf{e}_1 \otimes \mathbf{e}_1 + r\theta'\mathbf{e}_2 \otimes \mathbf{e}_2 + \Lambda\mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \square \qquad (xvi)$$

and similarly from $\mathbf{F} = \mathbf{VR}$ and (xiv) we obtain

$$\mathbf{V} = r'\mathbf{e}_r \otimes \mathbf{e}_r + r\theta'\mathbf{e}_\theta \otimes \mathbf{e}_\theta + \Lambda\mathbf{e}_z \otimes \mathbf{e}_z. \qquad \square \qquad (xvii)$$

(d) It follows from (xvi) and $(xvii)$ that the principal stretches are

$$\lambda_1 = r'(x_1), \qquad \lambda_2 = r(x_1)\theta'(x_2), \qquad \lambda_3 = \Lambda. \qquad \square \qquad (xviii)$$

Moreover (xvi) shows that the principal directions of $\mathbf{U}$ are $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, while $(xvii)$ gives the principal directions of $\mathbf{V}$ to be $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. $\square$

(e) Incompressibility requires

$$\lambda_1\lambda_2\lambda_3 = 1 \qquad \Rightarrow \qquad r(x_1)r'(x_1)\theta'(x_2)\Lambda = 1.$$

By separating variables we conclude that

$$\Lambda r(x_1)r'(x_1) = \frac{1}{\theta'(x_2)} = c_1 \text{ (constant)},$$

which when integrated yields

$$r(x_1) = \sqrt{\frac{2c_1}{\Lambda}x_1 + \beta}, \qquad \theta(x_2) = \frac{x_2}{c_1} + c_2,$$

where c_1, c_2 and β are constants. Since $\theta(x_2)$ is an odd function $c_2 = 0$. Finally it is convenient to let $c_1 = 1/\alpha$ and write this as

$$r(x_1) = \left[\frac{2x_1}{\alpha\Lambda} + \beta\right]^{1/2}, \qquad \theta(x_2) = \alpha x_2, \qquad \square \qquad (xix)$$

where α and β are constants. If the total angle subtended by the block in the deformed configuration is 2ϕ, then $\alpha = \phi/B$. Observe that in this case the principal stretches are functions only of x_1:

$$\lambda_1 = r'(x_1), \qquad \lambda_2 = \alpha r(x_1), \qquad \lambda_3 = \Lambda.$$

2.6 Strain.

It is clear that $\mathbf{U}$ (or $\mathbf{V}$) is the essential ingredient that characterizes the non-rigid part of the deformation gradient. "Strain" is simply an alternative measure of this part of the deformation, the only essential distinction between strain and stretch being that (by convention) the strain vanishes in a rigid deformation whereas the stretch equals the identity $\mathbf{I}$. Thus for *example* we might say that $\mathbf{U}-\mathbf{I}$ is the strain where $\mathbf{U}$ is the stretch.

Various measures of strain are used in the literature, examples of which we shall describe below. It should be pointed out that the continuum theory does not prefer[16] one strain measure over another; each is a one-to-one function of the stretch tensor and so all strain measures are equivalent. In fact, one does not even have to introduce the notion of strain and the theory could be based entirely on the stretch tensors $\mathbf{U}$ and $\mathbf{V}$.

The various measures of **Lagrangian strain** used in the literature are all related to the Lagrangian stretch $\mathbf{U}$ in a one-to-one manner. Examples include the Green Saint-Venant strain tensor, the Biot strain tensor, the generalized Green Saint-Venant strain tensor and the Hencky (or logarithmic) strain tensor, defined by the respective expressions

$$\begin{aligned}
&\text{Green Saint-Venant:} && \frac{1}{2}(\mathbf{U}^2-\mathbf{I})=\sum_{i=1}^{3}\frac{1}{2}(\lambda_i^2-1)\,\mathbf{r}_i\otimes\mathbf{r}_i,\\
&\text{Biot :} && \mathbf{U}-\mathbf{I}=\sum_{i=1}^{3}(\lambda_i-1)\,\mathbf{r}_i\otimes\mathbf{r}_i,\\
& && \frac{1}{m}(\mathbf{U}^m-\mathbf{I})=\sum_{i=1}^{3}\frac{1}{m}(\lambda_i^m-1)\,\mathbf{r}_i\otimes\mathbf{r}_i,\\
&\text{Hencky :} && \ln\mathbf{U}=\sum_{i=1}^{3}\ln\,\lambda_i\,\mathbf{r}_i\otimes\mathbf{r}_i,
\end{aligned}\tag{2.64}$$

where m is a non-zero integer. Observe that all of these strain measures vanish in a rigid deformation, i.e. when $\mathbf{U}=\mathbf{I}$. They are all symmetric and their principal directions are the Lagrangian principal directions $\{\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3\}$; the associated **principal strains** are

$$\tfrac{1}{2}(\lambda_i^2-1),\qquad \lambda_i-1,\qquad \tfrac{1}{m}(\lambda_i^m-1),\qquad \text{and}\qquad \ln\lambda_i, \tag{2.65}$$

[16]It may happen that the constitutive relation for a particular material takes an especially simple form when one particular strain measure is used in its characterization, while a different strain measure might lead to a simple constitutive description for some other material. This might then lead to a preference for one strain measure over another for a particular material.

respectively. In Problem 2.38 you are asked to calculate the components of the Green Saint-Venant strain tensor in simple shear.

Similarly, various measures of **Eulerian strain** are used in the literature, all of them being related to the Eulerian stretch $\mathbf{V}$ in a one-to-one manner. Examples include the Almansi strain, the generalized Almansi strain and the logarithmic strain, defined by the respective expressions

$$\begin{aligned} &\frac{1}{2}(\mathbf{I}-\mathbf{V}^{-2}) = \sum_{i=1}^{3} \frac{1}{2}(1-\lambda_i^{-2})\, \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i \\ &\frac{1}{m}(\mathbf{V}^m - \mathbf{I}) = \sum_{i=1}^{3} \frac{1}{m}(\lambda_i^m - 1)\, \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i, \\ &\ln \mathbf{V} = \sum_{i=1}^{3} \ln\, \lambda_i\, \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i, \end{aligned} \tag{2.66}$$

where m is a non-zero integer. The principal directions of all of these symmetric strain tensors are the Eulerian principal directions $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$.

The preceding examples may be *unified and generalized* as follows: Let $e(\lambda)$ be an arbitrary (for the moment) scalar-valued function defined for $0 < \lambda < \infty$, and consider defining the Lagrangian strain tensor $\mathbf{E}(\mathbf{U})$ to be the tensor with eigenvectors $\mathbf{r}_i$ and corresponding eigenvalues $e(\lambda_i)$, i.e.

$$\mathbf{E}(\mathbf{U}) = e(\lambda_1)\, \mathbf{r}_1 \otimes \mathbf{r}_1 \ + \ e(\lambda_2)\, \mathbf{r}_2 \otimes \mathbf{r}_2 \ + \ e(\lambda_3)\, \mathbf{r}_3 \otimes \mathbf{r}_3. \tag{2.67}$$

Note that the **principal strains** associated with this tensor are $e(\lambda_1), e(\lambda_2)$ and $e(\lambda_3)$ and the corresponding principal directions are $\mathbf{r}_1, \mathbf{r}_2$ and $\mathbf{r}_3$.

In the undeformed configuration the principal stretches are unity and we would like the strain to vanish there. Thus we require $e(1) = 0$. Next consider a "small deformation" of the body in which λ is close to unity. Then Taylor expansion of the function $e(\lambda)$ about $\lambda = 1$ leads to

$$e(\lambda) = e(1) + e'(1)(\lambda - 1) + \ldots = e'(1)\left(\frac{ds_y}{ds_x} - 1\right) + \ldots = e'(1)\ \frac{ds_y - ds_x}{ds_x} + \ldots. \tag{2.68}$$

In order that this coincide with the familiar definition of normal strain in an infinitesimal deformation, i.e. in order that $e \approx (ds_y - ds_x)/ds_x$, we must have $e'(1) = 1$. Finally we require the normal strain $e(\lambda)$ to increase monotonically as the stretch λ increases and so we impose $e'(\lambda) > 0$. Note then that the principal strain $e(\lambda)$ is positive for extensions ($\lambda > 1$) and negative for contractions ($\lambda < 1$).

Thus we define a Lagrangian strain $\mathbf{E}(\mathbf{U})$ by (2.67) where the function $e(\lambda)$ is required to have the properties

$$\begin{aligned} \text{a)} \quad e(1) &= 0, \\ \text{b)} \quad e'(1) &= 1, \\ \text{c)} \quad e'(\lambda) &> 0 \quad \text{for all} \quad \lambda > 0. \end{aligned} \tag{2.69}$$

Observe that all Lagrangian strain tensors defined by (2.67) are symmetric. Their diagonal components E_{11}, E_{22} and E_{33} are known as the *normal* components of strain, while the off-diagonal components E_{12}, E_{23} and E_{31} are the *shear* components of strain. In the *principal basis* $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$, the matrix $[E]$ of strain components is diagonal and so the shear strains vanish in this basis, the normal components being the *principal strains*.

A generalized Eulerian strain tensor can be defined analogously by

$$\mathcal{E}(\mathbf{V}) = e(\lambda_1)\,\boldsymbol{\ell}_1 \otimes \boldsymbol{\ell}_1 \;+\; e(\lambda_2)\,\boldsymbol{\ell}_2 \otimes \boldsymbol{\ell}_2 \;+\; e(\lambda_3)\,\boldsymbol{\ell}_3 \otimes \boldsymbol{\ell}_3, \tag{2.70}$$

where $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ are the principal directions of the Eulerian stretch tensor $\mathbf{V}$ and $e(\lambda)$ obeys (2.69).

2.6.1 Remarks on the Green Saint-Venant strain tensor.

While, as noted already, the theory does not prefer one strain measure over another, the Green Saint-Venant strain tensor has been used frequently in the (especially older) literature. It is therefore worth devoting some attention to it. From (2.58) and (2.64) the Green Saint-Venant strain tensor is

$$\mathbf{E} = \frac{1}{2}(\mathbf{U}^2 - \mathbf{I}) = \frac{1}{2}(\mathbf{C} - \mathbf{I}) = \frac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I}). \tag{2.71}$$

– First we wish to express $\mathbf{E}$ in terms of the displacement gradient tensor $\nabla\mathbf{u}$. Since $\mathbf{y} = \mathbf{x} + \mathbf{u}$ it follows that

  $$\mathbf{F} = \nabla\mathbf{y} = \nabla(\mathbf{x} + \mathbf{u}) = \mathbf{I} + \nabla\mathbf{u}.$$

  Substituting this into (2.71) yields

  $$\mathbf{E} = \frac{1}{2}\Big((\mathbf{I} + \nabla\mathbf{u})^T(\mathbf{I} + \nabla\mathbf{u}) - \mathbf{I}\Big) = \frac{1}{2}\Big(\nabla\mathbf{u} + (\nabla\mathbf{u})^T + (\nabla\mathbf{u})^T\nabla\mathbf{u}\Big). \tag{2.72}$$

  Since $\nabla\mathbf{u}$ is the tensor with cartesian components $\partial u_i/\partial x_j$, this can be written in terms of cartesian components as

  $$E_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{\partial u_k}{\partial x_i}\frac{\partial u_k}{\partial x_j}\right), \tag{2.73}$$

  where the summation convention is being used on the repeated index k.

– Second we wish to interpret the components of the Green Saint-Venant strain tensor in terms of changes in length and changes in angle. Recall that the stretch of a material fiber in the direction $\mathbf{m}_\mathrm{R}$ in the reference configuration is

$$\lambda(\mathbf{m}_\mathrm{R}) \overset{(2.29)}{=} |\mathbf{F}\mathbf{m}_\mathrm{R}| = \sqrt{\mathbf{F}\mathbf{m}_\mathrm{R}\cdot\mathbf{F}\mathbf{m}_\mathrm{R}} = \sqrt{\mathbf{F}^T\mathbf{F}\mathbf{m}_\mathrm{R}\cdot\mathbf{m}_\mathrm{R}} \overset{(2.71)}{=} \sqrt{(2\mathbf{E}\mathbf{m}_\mathrm{R}\cdot\mathbf{m}_\mathrm{R}+1)}. \qquad (i)$$

Now consider a fiber oriented in the direction $\mathbf{m}_\mathrm{R} = \mathbf{e}_1$. Its stretch is

$$\lambda(\mathbf{e}_1) = |\mathbf{F}\mathbf{e}_1| = \sqrt{(2\mathbf{E}\mathbf{e}_1\cdot\mathbf{e}_1+1)} \overset{(1.125)}{=} \sqrt{(2E_{11}+1)}. \qquad (ii)$$

Since the stretch λ represents the ratio of the deformed and undeformed lengths of the fiber, i.e. $\lambda = ds_y/ds_x$, it now follows that

$$\frac{ds_y - ds_x}{ds_x} = \lambda(\mathbf{e}_1) - 1 = \sqrt{(2E_{11}+1)} - 1. \qquad (2.74)$$

Thus we conclude that the change in length relative to the original length of a fiber in the direction $\mathbf{e}_1$ depends only on the normal strain E_{11}. It is not equal to E_{11} but is fully determined by E_{11}. If the normal strain $|E_{11}| \ll 1$, this approximates to the familiar expression

$$\frac{ds_y - ds_x}{ds_x} \approx E_{11}. \qquad (iii)$$

Analogous calculations may be carried out for fibers in the directions $\mathbf{e}_2$ and $\mathbf{e}_3$.

Now consider the change in angle between two fibers in the directions $\mathbf{e}_1$ and $\mathbf{e}_2$. From (2.35),

$$\cos\theta_y = \frac{\mathbf{F}\mathbf{e}_1\cdot\mathbf{F}\mathbf{e}_2}{|\mathbf{F}\mathbf{e}_1|\,|\mathbf{F}\mathbf{e}_2|} = \frac{\mathbf{F}^T\mathbf{F}\mathbf{e}_1\cdot\mathbf{e}_2}{|\mathbf{F}\mathbf{e}_1|\,|\mathbf{F}\mathbf{e}_2|} \overset{(ii)}{=} \frac{\mathbf{F}^T\mathbf{F}\mathbf{e}_1\cdot\mathbf{e}_2}{\sqrt{(2E_{11}+1)}\,\sqrt{(2E_{22}+1)}} \overset{(2.71)}{=} \frac{2\mathbf{E}\mathbf{e}_1\cdot\mathbf{e}_2}{\sqrt{(2E_{11}+1)}\,\sqrt{(2E_{22}+1)}}, \qquad (2.75)$$

where θ_y is the angle between the two fibers in the deformed configuration. In view of (1.125) this yields

$$\cos\theta_y = \frac{2\,E_{12}}{\sqrt{(2E_{11}+1)}\,\sqrt{(2E_{22}+1)}}, \qquad (2.76)$$

which shows that the angle between these two fibers in the deformed configuration depends on the shear strain E_{12} *and* the normal strains E_{11} and E_{22}. If the strains are small, i.e. $|E_{11}| \ll 1, |E_{22}| \ll 1$ and $|E_{12}| \ll 1$, this approximates to leading order to the familiar expression

$$\theta_x - \theta_y \approx 2E_{12},$$

where $\theta_x = \pi/2$.

– It is illuminating to calculate the components of $\mathbf{E}$ in a simple shear deformation

$$y_1 = x_1 + kx_2, \qquad y_2 = x_2, \qquad y_3 = x_3.$$

Differentiating this with respect to x_j and using $F_{ij} = \partial y_i/\partial x_j$ leads to

$$\mathbf{F} = \mathbf{I} + k\mathbf{e}_1\otimes\mathbf{e}_2, \qquad \mathbf{C} = \mathbf{I} + k(\mathbf{e}_1\otimes\mathbf{e}_2 + \mathbf{e}_2\otimes\mathbf{e}_1) + k^2\mathbf{e}_2\otimes\mathbf{e}_2,$$

and so the components of the Green Saint-Venant strain tensor are

$$[E] = \frac{1}{2}\big([C]-[I]\big) = \begin{pmatrix} 0 & k/2 & 0 \\ k/2 & k^2/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Observe that the normal strain $E_{22} \neq 0$. This is related to the fact that the line OP in Figure 2.4 increases in length when it deforms into OP'. Note that $E_{22} = O(k^2)$ so that in a linearized theory with a small amount of shear, $|k| \ll 1$, this term would be neglected, leading to a strain tensor whose only nonzero components are the shear strains E_{12} and E_{21}.

2.7 Some other coordinate systems.

2.7.1 Cylindrical polar coordinates.

In this section we illustrate working in other coordinate systems by calculating expressions for the deformation gradient tensor and the left Cauchy-Green deformation tensor in cylindrical polar coordinates.

In rectangular cartesian coordinates, a deformation is characterized by the mapping

$$y_1 = \widehat{y}_1(x_1, x_2, x_3), \quad y_2 = \widehat{y}_2(x_1, x_2, x_3) \quad y_3 = \widehat{y}_3(x_1, x_2, x_3), \tag{i}$$

that takes the particle with coordinates (x_1, x_2, x_3) in the reference configuration to the point (y_1, y_2, y_3) in the deformed configurations. Let (R, Θ, Z) be the cylindrical polar coordinates of this particle in the undeformed configuration so that (see Figure 2.18)

$$x_1 = R\cos\Theta, \quad x_2 = R\sin\Theta, \quad x_3 = Z, \tag{ii}$$

and let (r, θ, z) be its cylindrical polar coordinates in the deformed configuration so that

$$y_1 = r\cos\theta, \quad y_2 = r\sin\theta, \quad y_3 = z. \tag{iii}$$

By combining (i) with (ii) and (iii) we can characterize the deformation in the form,

$$r = \widehat{r}(R, \Theta, Z), \qquad \theta = \widehat{\theta}(R, \Theta, Z), \qquad z = \widehat{\theta}(R, \Theta, Z). \tag{iv}$$

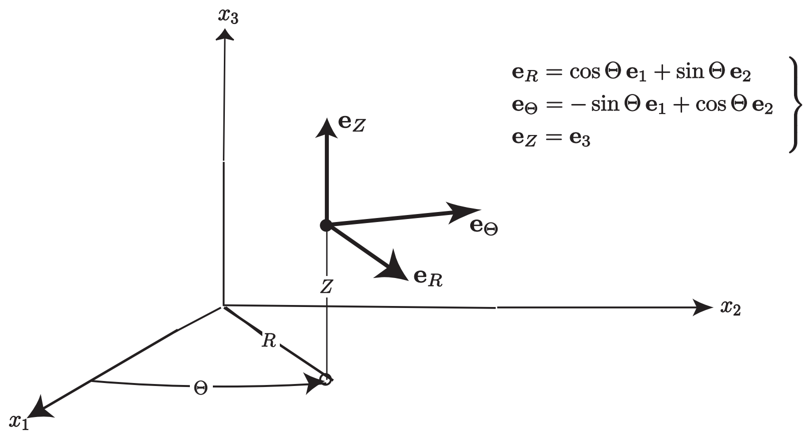


Figure 2.18: Cylindrical polar coordinates (R, Θ, Z) and the associated basis vectors $\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z$.

The basis $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ associated with the cylindrical polar coordinates (R, Θ, Z) is shown in Figure 2.18 and is related to the fixed cartesian basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by

$$\left.\begin{aligned} \mathbf{e}_R(\Theta) &= \cos\Theta\ \mathbf{e}_1 + \sin\Theta\ \mathbf{e}_2, \\ \mathbf{e}_\Theta(\Theta) &= -\sin\Theta\ \mathbf{e}_1 + \cos\Theta\ \mathbf{e}_2, \\ \mathbf{e}_Z &= \mathbf{e}_3, \end{aligned}\right\} \qquad (v)$$

while the corresponding relation for the basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ associated with the coordinates in the deformed configuration is

$$\left.\begin{aligned} \mathbf{e}_r(\theta) &= \cos\theta\ \mathbf{e}_1 + \sin\theta\ \mathbf{e}_2, \\ \mathbf{e}_\theta(\theta) &= -\sin\theta\ \mathbf{e}_1 + \cos\theta\ \mathbf{e}_2, \\ \mathbf{e}_z &= \mathbf{e}_3. \end{aligned}\right\} \qquad (vi)$$

Observe that in general, the basis $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ differs from the basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$.

The position vectors of a particle in the undeformed and deformed configurations are, see Figure 2.18, $\mathbf{x} = R\,\mathbf{e}_R + Z\,\mathbf{e}_Z$ and $\mathbf{y} = r\mathbf{e}_r + z\mathbf{e}_z$, or in a little more detail,

$$\mathbf{x} = \widehat{\mathbf{x}}(R, \Theta, Z) = R\,\mathbf{e}_R + Z\,\mathbf{e}_Z, \qquad (vii)$$

$$\mathbf{y} = \widehat{\mathbf{y}}(R, \Theta, Z) = \Big(r\mathbf{e}_r + z\mathbf{e}_z\Big)\Big|_{r=\widehat{r},\ \theta=\widehat{\theta},\ z=\widehat{z}}\ , \qquad (viii)$$

respectively. We wish to calculate the deformation gradient tensor using $(iv), (vii)$ and $(viii)$. The general approach involves calculating the vectors $d\mathbf{y}$ and $d\mathbf{x}$ independently and then recognizing that they are related by $d\mathbf{y} = \mathbf{F}\, d\mathbf{x}$. A general treatment of orthogonal curvilinear coordinates can be found in Chapter 6 of Volume I.

We now proceed to calculate $d\mathbf{y}$. First, from $(viii)$ and the chain rule

$$d\mathbf{y} = \frac{\partial\widehat{\mathbf{y}}}{\partial R}dR + \frac{\partial\widehat{\mathbf{y}}}{\partial\Theta}d\Theta + \frac{\partial\widehat{\mathbf{y}}}{\partial Z}dZ. \tag{ix}$$

Next, we calculate each term on the right-hand side of (ix):

$$\frac{\partial\widehat{\mathbf{y}}}{\partial R} \stackrel{(viii)}{=} \frac{\partial}{\partial R}(r\mathbf{e}_r + z\mathbf{e}_z) = \frac{\partial r}{\partial R}\mathbf{e}_r + r\frac{\partial \mathbf{e}_r}{\partial R} + \frac{\partial z}{\partial R}\mathbf{e}_z = \frac{\partial r}{\partial R}\mathbf{e}_r + r\frac{\partial \mathbf{e}_r}{\partial \theta}\frac{\partial\theta}{\partial R} + \frac{\partial z}{\partial R}\mathbf{e}_z,$$

where in getting to the second equality we used the fact that $\mathbf{e}_z$ does not depend on R, and in getting to the last equality we used the fact that according to (vi) the unit vector $\mathbf{e}_r$ depends on θ but not on r or z. Observe from (vi) that $\partial\mathbf{e}_r/\partial\theta = \mathbf{e}_\theta$. Thus we can write the preceding equation as

$$\frac{\partial\widehat{\mathbf{y}}}{\partial R} = \frac{\partial r}{\partial R}\mathbf{e}_r + r\frac{\partial\theta}{\partial R}\mathbf{e}_\theta + \frac{\partial z}{\partial R}\mathbf{e}_z. \tag{x}$$

The other terms in (ix) can be calculated similarly leading to

$$\frac{\partial\widehat{\mathbf{y}}}{\partial \Theta} = \frac{\partial r}{\partial \Theta}\mathbf{e}_r + r\frac{\partial\theta}{\partial \Theta}\mathbf{e}_\theta + \frac{\partial z}{\partial \Theta}\mathbf{e}_z, \tag{xi}$$

$$\frac{\partial\widehat{\mathbf{y}}}{\partial Z} = \frac{\partial r}{\partial Z}\mathbf{e}_r + r\frac{\partial\theta}{\partial Z}\mathbf{e}_\theta + \frac{\partial z}{\partial Z}\mathbf{e}_z. \tag{xii}$$

Substituting $(x), (xi), (xii)$ in (ix) now leads to

$$\begin{aligned} d\mathbf{y} \;\; &= \frac{\partial r}{\partial R}dR\,\mathbf{e}_r + r\frac{\partial\theta}{\partial R}dR\,\mathbf{e}_\theta + \frac{\partial z}{\partial R}dR\,\mathbf{e}_z + \\ &\quad + \frac{\partial r}{\partial \Theta}d\Theta\,\mathbf{e}_r + r\frac{\partial\theta}{\partial \Theta}d\Theta\,\mathbf{e}_\theta + \frac{\partial z}{\partial \Theta}d\Theta\,\mathbf{e}_z + \\ &\quad + \frac{\partial r}{\partial Z}dZ\,\mathbf{e}_r + r\frac{\partial\theta}{\partial Z}dZ\,\mathbf{e}_\theta + \frac{\partial z}{\partial Z}dZ\,\mathbf{e}_z. \end{aligned} \tag{$xiii$}$$

The next step is to express $dR, d\Theta$ and dZ in $(xiii)$ in terms of $d\mathbf{x}$. From (vii)

$$d\mathbf{x} = \frac{\partial\widehat{\mathbf{x}}}{\partial R}dR + \frac{\partial\widehat{\mathbf{x}}}{\partial\Theta}d\Theta + \frac{\partial\widehat{\mathbf{x}}}{\partial Z}dZ, \tag{xiv}$$

where

$$\frac{\partial\widehat{\mathbf{x}}}{\partial R} = \frac{\partial}{\partial R}(R\mathbf{e}_R + Z\mathbf{e}_Z) \stackrel{(v)}{=} \frac{\partial R}{\partial R}\,\mathbf{e}_R = \mathbf{e}_R,$$

$$\frac{\partial\widehat{\mathbf{x}}}{\partial \Theta} = \frac{\partial}{\partial \Theta}(R\mathbf{e}_R + Z\mathbf{e}_Z) \stackrel{(v)}{=} R\frac{\partial\mathbf{e}_R}{\partial\Theta} \stackrel{(v)}{=} R\mathbf{e}_\Theta,$$

$$\frac{\partial \widehat{\mathbf{x}}}{\partial Z} = \frac{\partial}{\partial Z}(R\mathbf{e}_R + Z\mathbf{e}_Z) \overset{(v)}{=} \mathbf{e}_Z.$$

Therefore (xiv) can be written as

$$d\mathbf{x} = dR\,\mathbf{e}_R + Rd\Theta\,\mathbf{e}_\Theta + dZ\,\mathbf{e}_Z, \qquad (xv)$$

from which we obtain

$$dR = d\mathbf{x}\cdot\mathbf{e}_R, \qquad d\Theta = R^{-1}\,d\mathbf{x}\cdot\mathbf{e}_\Theta, \qquad dZ = d\mathbf{x}\cdot\mathbf{e}_Z. \qquad (xvi)$$

Substituting (xvi) into $(xiii)$ gives

$$\begin{aligned} d\mathbf{y} \;&= \frac{\partial r}{\partial R}(d\mathbf{x}\cdot\mathbf{e}_R)\,\mathbf{e}_r + r\frac{\partial \theta}{\partial R}(d\mathbf{x}\cdot\mathbf{e}_R)\,\mathbf{e}_\theta + \frac{\partial z}{\partial R}(d\mathbf{x}\cdot\mathbf{e}_R)\,\mathbf{e}_z + \\ &+\frac{\partial r}{\partial \Theta}(R^{-1}\,d\mathbf{x}\cdot\mathbf{e}_\Theta)\,\mathbf{e}_r + r\frac{\partial \theta}{\partial \Theta}(R^{-1}\,d\mathbf{x}\cdot\mathbf{e}_\Theta)\,\mathbf{e}_\theta + \frac{\partial z}{\partial \Theta}(R^{-1}\,d\mathbf{x}\cdot\mathbf{e}_\Theta)\,\mathbf{e}_z \\ &+\frac{\partial r}{\partial Z}(d\mathbf{x}\cdot\mathbf{e}_Z)\,\mathbf{e}_r + r\frac{\partial \theta}{\partial Z}(d\mathbf{x}\cdot\mathbf{e}_Z)\,\mathbf{e}_\theta + \frac{\partial z}{\partial Z}(d\mathbf{x}\cdot\mathbf{e}_Z)\,\mathbf{e}_z, \end{aligned}$$

which can be written as

$$\begin{aligned} d\mathbf{y} \;&= \frac{\partial r}{\partial R}(\mathbf{e}_r\otimes\mathbf{e}_R)\,d\mathbf{x} + r\frac{\partial \theta}{\partial R}(\mathbf{e}_\theta\otimes\mathbf{e}_R)\,d\mathbf{x} + \frac{\partial z}{\partial R}(\mathbf{e}_z\otimes\mathbf{e}_R)\,d\mathbf{x} + \\ &+\frac{1}{R}\frac{\partial r}{\partial \Theta}(\mathbf{e}_r\otimes\mathbf{e}_\Theta)\,d\mathbf{x} + \frac{r}{R}\frac{\partial \theta}{\partial \Theta}(\mathbf{e}_\theta\otimes\mathbf{e}_\Theta)\,d\mathbf{x} + \frac{1}{R}\frac{\partial z}{\partial \Theta}(\mathbf{e}_z\otimes\mathbf{e}_\Theta)\,d\mathbf{x} + \\ &+\frac{\partial r}{\partial Z}(\mathbf{e}_r\otimes\mathbf{e}_Z)\,d\mathbf{x} + r\frac{\partial \theta}{\partial Z}(\mathbf{e}_\theta\otimes\mathbf{e}_Z)\,d\mathbf{x} + \frac{\partial z}{\partial Z}(\mathbf{e}_z\otimes\mathbf{e}_Z)\,d\mathbf{x}. \end{aligned}$$

Since $d\mathbf{y} = \mathbf{F}d\mathbf{x}$ we can now read off the deformation gradient tensor $\mathbf{F}$ to be

$$\begin{aligned} \mathbf{F} \;&= \frac{\partial r}{\partial R}(\mathbf{e}_r\otimes\mathbf{e}_R) + \frac{1}{R}\frac{\partial r}{\partial \Theta}(\mathbf{e}_r\otimes\mathbf{e}_\Theta) + \frac{\partial r}{\partial Z}(\mathbf{e}_r\otimes\mathbf{e}_Z) + \\ &+r\frac{\partial \theta}{\partial R}(\mathbf{e}_\theta\otimes\mathbf{e}_R) + \frac{r}{R}\frac{\partial \theta}{\partial \Theta}(\mathbf{e}_\theta\otimes\mathbf{e}_\Theta) + r\frac{\partial \theta}{\partial Z}(\mathbf{e}_\theta\otimes\mathbf{e}_Z) + \\ &+\frac{\partial z}{\partial R}(\mathbf{e}_z\otimes\mathbf{e}_R) \;+ \frac{1}{R}\frac{\partial z}{\partial \Theta}(\mathbf{e}_z\otimes\mathbf{e}_\Theta) \;+ \frac{\partial z}{\partial Z}(\mathbf{e}_z\otimes\mathbf{e}_Z)\,. \end{aligned} \qquad (2.77)$$

Observe that the representation (2.77) involves both bases $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. The scalar coefficients are not the components of $\mathbf{F}$ in either basis. See Problem 1.6.5.

The left and right Cauchy Green deformation tensors $\mathbf{B}$ and $\mathbf{C}$ can now be readily calculated. We will see below that when (2.77) is used to calculate $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ the result comes out in terms of the components of $\mathbf{B}$ in the basis $\mathbf{e}_r\otimes\mathbf{e}_r, \mathbf{e}_r\otimes\mathbf{e}_\theta, \ldots \mathbf{e}_z\otimes\mathbf{e}_z$. On the other hand when $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ is calculated the result emerges in terms of the components of $\mathbf{C}$ in the basis $\mathbf{e}_R \otimes \mathbf{e}_R, \mathbf{e}_R \otimes \mathbf{e}_\Theta, \ldots \mathbf{e}_Z \otimes \mathbf{e}_Z$. This is consistent with our previous reference to $\mathbf{B}$ as an Eulerian deformation tensor and $\mathbf{C}$ as a Lagrangian deformation tensor. Of course one can express either tensor in any basis of one's choice.

Turning to the left Cauchy-Green tensor $\mathbf{B} = \mathbf{F}\mathbf{F}^T$, it is straightforward to use (2.77) to show that as

$$\begin{aligned} \mathbf{B} = \; & B_{rr}\mathbf{e}_r \otimes \mathbf{e}_r \; + \; B_{\theta\theta}\mathbf{e}_\theta \otimes \mathbf{e}_\theta \; + \; B_{zz}\mathbf{e}_z \otimes \mathbf{e}_z \; + \\ & + B_{r\theta}(\mathbf{e}_r \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_r) \; + \; B_{rz}(\mathbf{e}_r \otimes \mathbf{e}_z + \mathbf{e}_z \otimes \mathbf{e}_r) \; + \\ & + B_{\theta z}(\mathbf{e}_z \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_z), \end{aligned} \tag{2.78}$$

where

$$\left.\begin{aligned} B_{rr} &= \left(\frac{\partial r}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial r}{\partial \Theta}\right)^2 + \left(\frac{\partial r}{\partial Z}\right)^2, \\ B_{\theta\theta} &= r^2\left[\left(\frac{\partial \theta}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial \theta}{\partial \Theta}\right)^2 + \left(\frac{\partial \theta}{\partial Z}\right)^2\right], \\ B_{zz} &= \left(\frac{\partial z}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial z}{\partial \Theta}\right)^2 + \left(\frac{\partial z}{\partial Z}\right)^2, \\ B_{r\theta} &= B_{\theta r} = r\left[\frac{\partial r}{\partial R}\frac{\partial \theta}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial \theta}{\partial \Theta} + \frac{\partial r}{\partial Z}\frac{\partial \theta}{\partial Z}\right], \\ B_{rz} &= B_{zr} = \frac{\partial r}{\partial R}\frac{\partial z}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial z}{\partial \Theta} + \frac{\partial r}{\partial Z}\frac{\partial z}{\partial Z}, \\ B_{\theta z} &= B_{z\theta} = r\left[\frac{\partial \theta}{\partial R}\frac{\partial z}{\partial R} + \frac{1}{R^2}\frac{\partial z}{\partial \Theta}\frac{\partial \theta}{\partial \Theta} + \frac{\partial \theta}{\partial Z}\frac{\partial z}{\partial Z}\right]. \end{aligned}\right\} \tag{2.79}$$

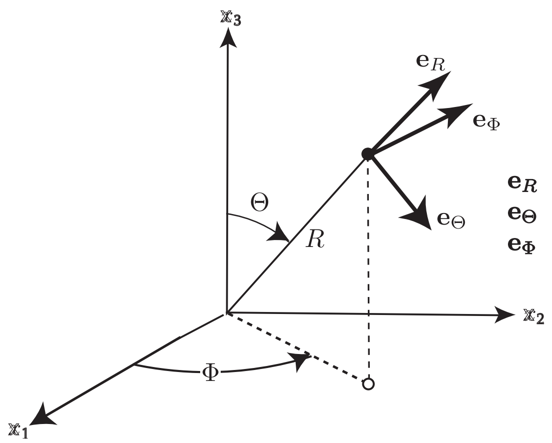
2.7.2 Spherical polar coordinates.

Let (x_1, x_2, x_3) denote the rectangular cartesian coordinates of a particle in the reference configuration, and let (R, Θ, Φ) be its spherical polar coordinates; (see Figure 2.19) Then

$$x_1 = R\sin\Theta\cos\Phi, \qquad x_2 = R\sin\Theta\sin\Phi, \qquad x_3 = R\cos\Theta. \tag{2.80}$$

The associated basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi\}$ are related by

$$\left.\begin{aligned} \mathbf{e}_R &= (\sin\Theta\cos\Phi)\;\mathbf{e}_1 + (\sin\Theta\sin\Phi)\;\mathbf{e}_2 + \cos\Theta\;\mathbf{e}_3, \\ \mathbf{e}_\Theta &= (\cos\Theta\cos\Phi)\;\mathbf{e}_1 + (\cos\Theta\sin\Phi)\;\mathbf{e}_2 - \sin\Theta\;\mathbf{e}_3, \\ \mathbf{e}_\Phi &= -\sin\Phi\;\mathbf{e}_1 + \cos\Phi\;\mathbf{e}_2. \end{aligned}\right\} \tag{2.81}$$



$$\left.\begin{array}{lcrcrcr}
\mathbf{e}_R & = & (\sin\Theta\cos\Phi)\ \mathbf{e}_1 & + & (\sin\Theta\sin\Phi)\ \mathbf{e}_2 & + & \cos\Theta\ \mathbf{e}_3 \\
\mathbf{e}_\Theta & = & (\cos\Theta\cos\Phi)\ \mathbf{e}_1 & + & (\cos\Theta\sin\Phi)\ \mathbf{e}_2 & - & \sin\Theta\ \mathbf{e}_3 \\
\mathbf{e}_\Phi & = & -\sin\Phi\ \mathbf{e}_1 & + & \cos\Phi\ \mathbf{e}_2 & &
\end{array}\right\}$$

Figure 2.19: Spherical polar coordinates (R, Θ, Φ) and the associated basis vectors $\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi$.

Let (y_1, y_2, y_3) denote the rectangular cartesian coordinates of a particle in the deformed configuration, and let (r, ϑ, φ) be its spherical polar coordinates. Then

$$y_1 = r\sin\vartheta\cos\varphi, \qquad y_2 = r\sin\vartheta\sin\varphi, \qquad y_3 = r\cos\vartheta. \tag{2.82}$$

The associated basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_r, \mathbf{e}_\vartheta, \mathbf{e}_\varphi\}$ are related by

$$\left.\begin{array}{lcrcrcr}
\mathbf{e}_r & = & (\sin\vartheta\cos\varphi)\ \mathbf{e}_1 & + & (\sin\vartheta\sin\varphi)\ \mathbf{e}_2 & + & \cos\vartheta\ \mathbf{e}_3, \\
\mathbf{e}_\vartheta & = & (\cos\vartheta\cos\varphi)\ \mathbf{e}_1 & + & (\cos\vartheta\sin\varphi)\ \mathbf{e}_2 & - & \sin\vartheta\ \mathbf{e}_3, \\
\mathbf{e}_\varphi & = & -\sin\varphi\ \mathbf{e}_1 & + & \cos\varphi\ \mathbf{e}_2. & &
\end{array}\right\} \tag{2.83}$$

The deformation can be characterized by

$$r = r(R, \Theta, \Phi), \qquad \vartheta = \vartheta(R, \Theta, \Phi), \qquad \varphi = \varphi(R, \Theta, \Phi), \tag{2.84}$$

The deformation gradient tensor can be expressed in spherical polar coordinates with respect to the mixed bases $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$ as

$$\begin{aligned}
\mathbf{F} &= \frac{\partial r}{\partial R}(\mathbf{e}_r \otimes \mathbf{e}_R) + \frac{1}{R}\frac{\partial r}{\partial \Theta}(\mathbf{e}_r \otimes \mathbf{e}_\Theta) + \frac{1}{R\sin\Theta}\frac{\partial r}{\partial \Phi}(\mathbf{e}_r \otimes \mathbf{e}_\Phi) + \\
&+ r\frac{\partial \theta}{\partial R}(\mathbf{e}_\theta \otimes \mathbf{e}_R) + \frac{r}{R}\frac{\partial \theta}{\partial \Theta}(\mathbf{e}_\theta \otimes \mathbf{e}_\Theta) + \frac{r}{R\sin\Theta}\frac{\partial \theta}{\partial \Phi}(\mathbf{e}_\theta \otimes \mathbf{e}_\Phi) + \\
&+ r\sin\theta\frac{\partial \varphi}{\partial R}(\mathbf{e}_\varphi \otimes \mathbf{e}_R) + \frac{r\sin\theta}{R}\frac{\partial \varphi}{\partial \Theta}(\mathbf{e}_\varphi \otimes \mathbf{e}_\Theta) + \frac{r\sin\theta}{R\sin\Theta}\frac{\partial \varphi}{\partial \Phi}(\mathbf{e}_\varphi \otimes \mathbf{e}_\Phi)\,.
\end{aligned} \tag{2.85}$$

The corresponding representation for the left Cauchy-Green tensor $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ is

$$\begin{aligned}\mathbf{B} = \; & B_{rr}\mathbf{e}_r \otimes \mathbf{e}_r \; + \; B_{\vartheta\vartheta}\mathbf{e}_\vartheta \otimes \mathbf{e}_\vartheta \; + \; B_{\varphi\varphi}\mathbf{e}_\varphi \otimes \mathbf{e}_\varphi \; + \\ & + B_{r\vartheta}(\mathbf{e}_r \otimes \mathbf{e}_\vartheta + \mathbf{e}_\vartheta \otimes \mathbf{e}_r) \; + \; B_{r\varphi}(\mathbf{e}_r \otimes \mathbf{e}_\varphi + \mathbf{e}_\varphi \otimes \mathbf{e}_r) \; + \\ & + B_{\vartheta\varphi}(\mathbf{e}_\varphi \otimes \mathbf{e}_\vartheta + \mathbf{e}_\vartheta \otimes \mathbf{e}_\varphi),\end{aligned} \tag{2.86}$$

where

$$\left.\begin{aligned}
B_{rr} \;&= \left(\frac{\partial r}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial r}{\partial \Theta}\right)^2 + \frac{1}{R^2 \sin^2\Theta}\left(\frac{\partial r}{\partial \Phi}\right)^2, \\
B_{\vartheta\vartheta} \;&= r^2\left[\left(\frac{\partial \vartheta}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial \vartheta}{\partial \Theta}\right)^2 + \frac{1}{R^2 \sin^2\Theta}\left(\frac{\partial \vartheta}{\partial \Phi}\right)^2\right], \\
B_{\varphi\varphi} \;&= r^2 \, \sin^2 \vartheta\left[\left(\frac{\partial \varphi}{\partial R}\right)^2 + \frac{1}{R^2}\left(\frac{\partial \varphi}{\partial \Theta}\right)^2 + \frac{1}{R^2 \sin^2\Theta}\left(\frac{\partial \varphi}{\partial \Phi}\right)^2\right] \\
B_{r\vartheta} \;&= B_{\vartheta r} = r\left[\frac{\partial r}{\partial R}\frac{\partial \vartheta}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial \vartheta}{\partial \Theta} + \frac{1}{R^2 \sin^2\Theta}\frac{\partial r}{\partial \Phi}\frac{\partial \vartheta}{\partial \Phi}\right], \\
B_{r\varphi} \;&= B_{\varphi r} = r \, \sin \vartheta\left[\frac{\partial r}{\partial R}\frac{\partial \varphi}{\partial R} + \frac{1}{R^2}\frac{\partial r}{\partial \Theta}\frac{\partial \varphi}{\partial \Theta} + \frac{1}{R^2 \sin^2\Theta}\frac{\partial r}{\partial \Phi}\frac{\partial \varphi}{\partial \Phi}\right], \\
B_{\vartheta\varphi} \;&= B_{\varphi\vartheta} = r^2 \, \sin \vartheta\left[\frac{\partial \vartheta}{\partial R}\frac{\partial \varphi}{\partial R} + \frac{1}{R^2}\frac{\partial \varphi}{\partial \Theta}\frac{\partial \vartheta}{\partial \Theta} + \frac{1}{R^2 \sin^2\Theta}\frac{\partial \vartheta}{\partial \Phi}\frac{\partial \varphi}{\partial \Phi}\right].
\end{aligned}\right\} \tag{2.87}$$

2.7.3 Worked examples.

Problem 2.8.1. (Ogden) *(Extension and torsion of a solid circular cylinder.)* Let (R, Θ, Z) and (r, θ, z) be cylindrical polar coordinates of a particle in the reference and deformed configurations respectively with associated bases $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. The region $\mathcal{R}_R$ occupied by the body in a reference configuration is a circular cylinder of radius A and length L.

The body is uniformly stretched axially to a length $\ell = \Lambda L$ (with an accompanying transverse contraction), and the stretched cylinder in then subjected to a torsional deformation. Thus the cross section at Z in the reference configuration, displaces to ΛZ and then rotates through an angle $\alpha\Lambda Z$. In particular, one end of the cylinder is held fixed while the other rotates through an angle $\alpha\ell$. This deformation has the form

$$r = r(R), \qquad \theta = \Theta + \alpha\Lambda Z, \qquad z = \Lambda Z. \tag{i}$$

(a) Calculate the deformation gradient tensor.

(b) Determine $r(R)$ assuming the material to be incompressible,

(c) By factoring the deformation gradient tensor, show that locally, at each point of the body, the deformation is comprised of a rigid rotation, followed by a pure stretch, followed be a simple shear with shearing direction $\mathbf{e}_\theta$ and glide plane normal $\mathbf{e}_z$.

(d) Calculate the right Cauchy Green deformation tensor $\mathbf{C}$.

e) Calculate the principal stretches and the principal Lagrangian stretch directions.

Solution:

(a) The deformation gradient tensor is found by specializing (2.77) to the deformation (i) which leads to

$$\mathbf{F} = r'(R)(\mathbf{e}_r \otimes \mathbf{e}_R) + \frac{r}{R}(\mathbf{e}_\theta \otimes \mathbf{e}_\Theta) + \alpha\Lambda r(\mathbf{e}_\theta \otimes \mathbf{e}_Z) + \Lambda(\mathbf{e}_z \otimes \mathbf{e}_Z)\,. \qquad \square$$

(b) When the material is incompressible

$$\det \mathbf{F} = 1 \qquad \Rightarrow \quad \Lambda \frac{r}{R}\frac{dr}{dR} = 1 \qquad \Rightarrow \quad r(R) = \sqrt{c + R^2/\Lambda},$$

where c is a constant of integration. However, since particles on the axis of the cylinder undergo no radial displacement it is necessary that $r(0) = 0$. This implies that $c = 0$ and therefore

$$r(R) = \Lambda^{-1/2} R. \qquad \square \qquad (ii)$$

It is convenient to set

$$\lambda := \Lambda^{-1/2}$$

so that $r'(R) = r(R)/R = \lambda$ and so

$$\mathbf{F} = \lambda(\mathbf{e}_r \otimes \mathbf{e}_R + \mathbf{e}_\theta \otimes \mathbf{e}_\Theta) + \alpha\Lambda r(\mathbf{e}_\theta \otimes \mathbf{e}_Z) + \Lambda(\mathbf{e}_z \otimes \mathbf{e}_Z)\,. \qquad (iii)$$

(c) It can be readily verified that we can factor the deformation gradient tensor (iii) and write it as

$$\mathbf{F} = \mathbf{F}_1\mathbf{F}_2\mathbf{F}_3, \qquad (iv)$$

where

$$\mathbf{F}_3 = \mathbf{e}_r \otimes \mathbf{e}_R + \mathbf{e}_\theta \otimes \mathbf{e}_\Theta + \mathbf{e}_z \otimes \mathbf{e}_Z, \qquad (v)$$

$$\mathbf{F}_2 = \lambda\mathbf{e}_r \otimes \mathbf{e}_r + \lambda\mathbf{e}_\theta \otimes \mathbf{e}_\theta + \Lambda\mathbf{e}_z \otimes \mathbf{e}_z, \qquad (vi)$$

$$\mathbf{F}_1 = \mathbf{I} + \alpha r\, \mathbf{e}_\theta \otimes \mathbf{e}_z. \qquad (vii)$$

Comparing (v) with (1.158) shows that $\mathbf{F}_3$ describes the rigid rotation that takes $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ into $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. Comparing (vi) with (2.7) shows that $\mathbf{F}_2$ describes a pure stretch with extension Λ in the z-direction and lateral contraction λ in the plane normal to it. Comparing (vii) with (2.14) shows that $\mathbf{F}_1$ describes a simple shear of amount αr in the shearing direction $\mathbf{e}_\theta$ with glide plane normal $\mathbf{e}_z$.

(d) The left Cauchy-Green tensor is

$$\mathbf{C} = \mathbf{F}^T\mathbf{F} \overset{(ii)}{=} \lambda^2(\mathbf{e}_R \otimes \mathbf{e}_R) + \lambda^2(\mathbf{e}_\Theta \otimes \mathbf{e}_\Theta) + \alpha\lambda\Lambda r(\mathbf{e}_\Theta \otimes \mathbf{e}_Z + \mathbf{e}_Z \otimes \mathbf{e}_\Theta) + \Lambda^2(1+\alpha^2 r^2)(\mathbf{e}_Z \otimes \mathbf{e}_Z). \quad (viii)$$

(e) To find the principal stretches and principal directions of Lagrangian stretch, we can solve the eigenvalue problem for $\mathbf{C}$. However because of the special nature of the deformation under consideration, one can find the eigenvalues more easily. From the form of $\mathbf{C}$ (specifically since there are no shear components associated with $\mathbf{e}_R$) we conclude immediately that one principal value and principal direction are

$$\lambda_1 = \lambda = \Lambda^{-1/2}, \qquad \mathbf{r}_1 = \mathbf{e}_R. \qquad (ix)$$

In a principal basis,

$$\mathbf{C} = \lambda_1^2(\mathbf{r}_1 \otimes \mathbf{r}_1) + \lambda_2^2(\mathbf{r}_2 \otimes \mathbf{r}_2) + \lambda_3^2(\mathbf{r}_3 \otimes \mathbf{r}_3). \qquad (x)$$

Calculating $\operatorname{tr}\mathbf{C}$ from $(viii)$ and (x) and equating the results gives

$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 2\lambda^2 + \Lambda^2(1+\alpha^2 r^2) \qquad \Rightarrow \qquad \lambda_2^2 + \lambda_3^2 = \lambda^2 + \Lambda^2(1+\alpha^2 r^2). \qquad (xi)$$

Next calculating $\det\mathbf{F}$ from (iii) and equating the result to $\lambda_1\lambda_2\lambda_3$ yields

$$\lambda_1\lambda_2\lambda_3 = \lambda^2\Lambda \qquad \Rightarrow \qquad \lambda_2\lambda_3 = \lambda\Lambda. \qquad (xii)$$

Equations (xi) and (xii) can be solved for the principal stretches λ_2 and λ_3. (Note that in general $\lambda_2 \neq \lambda$ and $\lambda_3 \neq \Lambda$.)

Let the corresponding principal directions be

$$\mathbf{r}_2 = \cos\psi\,\mathbf{e}_\Theta + \sin\psi\,\mathbf{e}_Z, \qquad \mathbf{r}_3 = -\sin\psi\,\mathbf{e}_\Theta + \cos\psi\,\mathbf{e}_Z, \qquad (xiii)$$

for some to-be-determined angle ψ. Substituting $(ix)_2$ and $(xiii)$ into (x) and comparing the resulting expression with $(viii)$ leads to

$$\lambda_2^2\cos^2\psi + \lambda_3^2\sin^2\psi = \lambda^2, \qquad \lambda_2^2\sin^2\psi + \lambda_3^2\cos^2\psi = \Lambda^2(1+\alpha^2 r^2), \qquad (\lambda_2^2 - \lambda_3^2)\sin\psi\cos\psi = \alpha\lambda\Lambda r.$$

These are three equations for λ_2, λ_3 and ψ. We do not need to solve them for the two principal stretches since we found them above in (xi) and (xii). By eliminating λ_2^2 and λ_3^2 from the three preceding equations we find

$$\tan 2\psi = \frac{2\alpha\lambda\Lambda r}{\lambda^2 - \Lambda^2(1+\alpha^2 r^2)}. \qquad (xiv)$$

Thus the other two principal directions of $\mathbf{C}$ are given by $(xiii), (xiv)$.

2.8 Spatial and referential descriptions of a field.

Consider a scalar field $\phi(\mathbf{y})$ defined on the region $\mathcal{R}$ occupied by the body in the deformed configuration. For example it might represent the temperature field in the deformed body, with $\phi(\mathbf{y})$ being the temperature at the particle whose position in the deformed configuration is $\mathbf{y}$. This function $\phi(\mathbf{y})$, defined on $\mathcal{R}$, can be expressed as a function of $\mathbf{x}$ defined on $\mathcal{R}_{\mathrm{R}}$ by "changing variables" from $\mathbf{y} \to \mathbf{x}$:

$$\widehat{\phi}(\mathbf{x}) := \phi(\mathbf{y})\Big|_{\mathbf{y}=\widehat{\mathbf{y}}(\mathbf{x})} \qquad \text{for all } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}, \tag{2.88}$$

where $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x})$ is the deformation. *Note that $\widehat{\phi}(\mathbf{x})$ is not the temperature of the undeformed body.* The functions $\phi(\mathbf{y})$ and $\widehat{\phi}(\mathbf{x})$ both give the temperature at the same particle p in the deformed body, the particle having been identified in two different ways – in the former by its location $\mathbf{y}$ in the deformed configuration and in the latter by its position $\mathbf{x}$ in the reference configuration. One refers to the representation $\widehat{\phi}(\mathbf{x})$ as the **referential or material description** of the field under consideration, and $\phi(\mathbf{y})$ as its **spatial description**.

Since the deformation $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x})$ is one-to-one, there exists an inverse deformation

$$\mathbf{x} = \overline{\mathbf{x}}(\mathbf{y}) \tag{2.89}$$

that carries $\mathcal{R} \to \mathcal{R}_{\mathrm{R}}$. Therefore any function $\widehat{\phi}(\mathbf{x})$ defined on $\mathcal{R}_{\mathrm{R}}$ can be written as a function $\phi(\mathbf{y})$ defined on $\mathcal{R}$ by "changing variables" from $\mathbf{x} \to \mathbf{y}$:

$$\phi(\mathbf{y}) = \widehat{\phi}(\mathbf{x})\Big|_{\mathbf{x}=\overline{\mathbf{x}}(\mathbf{y})} \qquad \text{for all } \mathbf{y} \in \mathcal{R}. \tag{2.90}$$

If one takes the gradient of $\widehat{\phi}(\mathbf{x})$ with respect to $\mathbf{x}$ one gets a certain gradient vector that we will denote by $\mathrm{Grad}\,\phi$. It has cartesian components $\partial\widehat{\phi}/\partial x_i$. On the other hand if one takes the gradient of $\phi(\mathbf{y})$ with respect to $\mathbf{y}$ one gets a different gradient vector denoted by $\mathrm{grad}\,\phi$ with cartesian components $\partial\phi/\partial y_i$. Since we use two different symbols "Grad" and "grad" to denote these two gradients, it is not necessary to write, for example, $\mathrm{Grad}\,\widehat{\phi}$, it being understood that "Grad" applies on the referential field (and "grad" on the spatial field). The relation between $\mathrm{Grad}\,\phi$ and $\mathrm{grad}\,\phi$ can be determined by differentiating both sides of (2.88) with respect to $\mathbf{x}$, or both sides of (2.90) with respect to $\mathbf{y}$, and using the chain rule. (Problem 2.8.2)

In general, for arbitrary scalar fields $\phi(\mathbf{y})$ and $\psi(\mathbf{x})$, $\mathrm{grad}\,\phi$ and $\mathrm{Grad}\,\psi$ denote the respective vector fields with Cartesian components $\partial\phi/\partial y_i$ and $\partial\psi/\partial x_i$:

$$\mathrm{grad}\,\phi = \frac{\partial\phi}{\partial y_i}\,\mathbf{e}_i, \qquad \mathrm{Grad}\,\psi = \frac{\partial\psi}{\partial x_i}\,\mathbf{e}_i. \tag{2.91}$$

For arbitrary vector fields $\mathbf{a}(\mathbf{y})$ and $\mathbf{b}(\mathbf{x})$, $\mathrm{grad}\,\mathbf{a}$ and $\mathrm{Grad}\,\mathbf{b}$ denote the respective tensor fields with Cartesian components $\partial a_i/\partial y_j$ and $\partial b_i/\partial x_j$:

$$\mathrm{grad}\,\mathbf{a} = \frac{\partial a_i}{\partial y_j}\,\mathbf{e}_i \otimes \mathbf{e}_j, \qquad \mathrm{Grad}\,\mathbf{b} = \frac{\partial b_i}{\partial x_j}\,\mathbf{e}_i \otimes \mathbf{e}_j, \tag{2.92}$$

and

$$\operatorname{div} \mathbf{a} = \frac{\partial a_i}{\partial y_i}, \qquad \operatorname{Div} \mathbf{b} = \frac{\partial b_i}{\partial x_i}. \tag{2.93}$$

For arbitrary tensor fields $\mathbf{A}(\mathbf{y})$ and $\mathbf{B}(\mathbf{x})$, div $\mathbf{A}$ and Div $\mathbf{B}$ denote the respective vector fields with Cartesian components $\partial A_{ij}/\partial y_j$ and $\partial B_{ij}/\partial x_j$:

$$\operatorname{div} \mathbf{A} = \frac{\partial A_{ij}}{\partial y_j}\,\mathbf{e}_i, \qquad \operatorname{Div} \mathbf{B} = \frac{\partial B_{ij}}{\partial x_j}\,\mathbf{e}_i. \tag{2.94}$$

Let $\mathbf{u}(\mathbf{y})$ and $\widehat{\mathbf{u}}(\mathbf{x})$ be spatial and material descriptions of a vector-valued field:

$$\widehat{\mathbf{u}}(\mathbf{x}) = \mathbf{u}(\mathbf{y})\Big|_{\mathbf{y}=\widehat{\mathbf{y}}(\mathbf{x})}, \qquad \mathbf{u}(\mathbf{y}) = \widehat{\mathbf{u}}(\mathbf{x})\Big|_{\mathbf{x}=\overline{\mathbf{x}}(\mathbf{y})}. \tag{i}$$

Suppose that $\mathbf{u}(\mathbf{y})$ satisfies the differential equation

$$\operatorname{div} \mathbf{u}(\mathbf{y}) = 0 \qquad \text{at each } \mathbf{y} \in \mathcal{R}. \tag{ii}$$

Then by mapping $\mathbf{y} \to \mathbf{x}$ ("changing variables from $\mathbf{y}$ to $\mathbf{x}$") we have

$$\operatorname{div} \mathbf{u} = \frac{\partial u_i}{\partial y_i} = \frac{\partial \widehat{u}_i}{\partial x_j}\frac{\partial x_j}{\partial y_i} = \frac{\partial \widehat{u}_i}{\partial x_j} F^{-1}_{ji} = \frac{\partial \widehat{u}_i}{\partial x_j} F^{-T}_{ij} = \operatorname{Grad} \widehat{\mathbf{u}} \cdot \mathbf{F}^{-T}, \tag{iii}$$

where the dot in the last expression denotes the scalar product between the tensors Grad $\widehat{\mathbf{u}}$ and $\mathbf{F}^{-T}$. Thus the referential version of the differential equation (ii) is

$$\operatorname{Grad} \widehat{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{F}^{-T}(\mathbf{x}) = 0 \quad \text{at each } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}. \tag{iv}$$

Observe that the differential equation (iv) *holds on* $\mathcal{R}_{\mathrm{R}}$.

2.8.1 Worked examples.

Problem 2.8.1. Let $\widehat{\phi}(\mathbf{y})$ and $\widehat{\mathbf{v}}(\mathbf{y})$ be a scalar and vector field defined on $\mathcal{R}$ and let $\phi(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$ be the corresponding scalar and vector fields defined on $\mathcal{R}_{\mathrm{R}}$ through the deformation $\mathbf{y} = \mathbf{y}(\mathbf{x})$:

$$\phi(\mathbf{x}) = \widehat{\phi}(\mathbf{y}(\mathbf{x})), \qquad \mathbf{v}(\mathbf{x}) = \widehat{\mathbf{v}}(\mathbf{y}(\mathbf{x})).$$

Show that

$$\operatorname{Grad} \phi = \mathbf{F}^T \operatorname{grad} \widehat{\phi}, \qquad \operatorname{Grad} \mathbf{v} = \operatorname{grad} \widehat{\mathbf{v}}\ \mathbf{F}, \tag{2.95}$$

Here, in cartesian coordinates,

$$\Big(\operatorname{Grad} \phi\Big)_i = \frac{\partial \phi}{\partial x_i}, \qquad \Big(\operatorname{grad} \widehat{\phi}\Big)_i = \frac{\partial \widehat{\phi}}{\partial y_i}, \qquad \Big(\operatorname{Grad} \mathbf{v}\Big)_{ij} = \frac{\partial v_i}{\partial x_j}, \qquad \Big(\operatorname{grad} \widehat{\mathbf{v}}\Big)_{ij} = \frac{\partial \widehat{v}_i}{\partial y_j}.$$

Problem 2.8.2. Let $\phi(\mathbf{y})$ and $\mathbf{v}(\mathbf{y})$ be a scalar and vector field defined on $\mathcal{R}$ and let $\widehat{\phi}(\mathbf{x})$ and $\widehat{\mathbf{v}}(\mathbf{x})$ be the corresponding scalar and vector fields defined on $\mathcal{R}_{\rm R}$ through the deformation $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x})$:

$$\widehat{\phi}(\mathbf{x}) = \phi(\widehat{\mathbf{y}}(\mathbf{x})), \qquad \widehat{\mathbf{v}}(\mathbf{x}) = \mathbf{v}(\widehat{\mathbf{y}}(\mathbf{x})).$$

Show that

$$\operatorname{Grad}\widehat{\phi} = \mathbf{F}^T \operatorname{grad}\phi, \qquad \operatorname{Grad}\widehat{\mathbf{v}} = \operatorname{grad}\mathbf{v}\ \mathbf{F}, \tag{2.96}$$

Here, in cartesian coordinates,

$$\Big(\operatorname{Grad}\widehat{\phi}\Big)_i = \frac{\partial\widehat{\phi}}{\partial x_i}, \qquad \Big(\operatorname{grad}\phi\Big)_i = \frac{\partial\phi}{\partial y_i}, \qquad \Big(\operatorname{Grad}\widehat{\mathbf{v}}\Big)_{ij} = \frac{\partial\widehat{v}_i}{\partial x_j}, \qquad \Big(\operatorname{grad}\mathbf{v}\Big)_{ij} = \frac{\partial v_i}{\partial y_j}.$$

Remark: It is not necessary to include the "hats" in (2.96) and the line below it since we use two different symbols, Grad and grad, for the two gradients. It is understood that ϕ is expressed referentially in $\operatorname{Grad}\phi$ and spatially in $\operatorname{grad}\phi$.

Problem 2.8.3. Let $\mathbf{a}(\mathbf{y})$ and $\mathbf{b}(\mathbf{x})$ be spatial and referential representations of two vector fields. Suppose they are related by

$$\mathbf{b} = J\mathbf{F}^{-1}\mathbf{a}, \tag{i}$$

where $\mathbf{F} = \nabla\mathbf{y}$ is the deformation gradient tensor and $J = \det\mathbf{F}$ is the Jacobian determinant. Show that

$$\operatorname{Div}\mathbf{b} = J\operatorname{div}\mathbf{a}. \tag{ii}$$

Solution: The result follows from

$$\begin{aligned}\operatorname{Div}\mathbf{b} &= \frac{\partial b_i}{\partial x_i} \overset{(i)}{=} \frac{\partial}{\partial x_i}(JF^{-1}_{ij}a_j) = \frac{\partial}{\partial x_i}(JF^{-T}_{ji}a_j) = \frac{\partial}{\partial x_i}(JF^{-T}_{ji})a_j + JF^{-T}_{ji}\frac{\partial a_j}{\partial x_i} = \\ &\overset{(2.120)}{=} JF^{-T}_{ji}\frac{\partial a_j}{\partial x_i} = JF^{-1}_{ij}\frac{\partial a_j}{\partial y_k}\frac{\partial y_k}{\partial x_i} = JF^{-1}_{ij}\frac{\partial a_j}{\partial y_k}F_{ki} = J\delta_{kj}\frac{\partial a_j}{\partial y_k} = J\frac{\partial a_j}{\partial y_j} = J\operatorname{div}\mathbf{a}\end{aligned}$$

where we used the Piola identity (2.120) in getting to the second line.

Problem 2.8.4. Let $\mathbf{A}(\mathbf{y})$ and $\mathbf{B}(\mathbf{x})$ be spatial and referential representations of two tensor fields. Suppose they are related by

$$\mathbf{B} = J\mathbf{A}\mathbf{F}^{-T}. \tag{i}$$

Show that

$$\operatorname{Div}\mathbf{B} = J\operatorname{div}\mathbf{A}, \tag{ii}$$

where the divergence of a tensor field is given in (2.94).

Solution: The result follows from

$$
\begin{aligned}
(\operatorname{Div}\mathbf{B})_k \quad &= \frac{\partial B_{ki}}{\partial x_i} \overset{(i)}{=} \frac{\partial}{\partial x_i}(JA_{kj}F^{-T}_{ji}) = \frac{\partial}{\partial x_i}(JF^{-T}_{ji}A_{kj}) = \frac{\partial}{\partial x_i}(JF^{-T}_{ji})A_{kj} + JF^{-T}_{ji}\frac{\partial A_{kj}}{\partial x_i} = \\
&\overset{(2.120)}{=} JF^{-T}_{ji}\frac{\partial A_{kj}}{\partial x_i} = JF^{-1}_{ij}\frac{\partial A_{kj}}{\partial y_p}\frac{\partial y_p}{\partial x_i} = JF^{-1}_{ij}\frac{\partial A_{kj}}{\partial y_p}F_{pi} = J\delta_{pj}\frac{\partial A_{kj}}{\partial y_p} = J\frac{\partial A_{kj}}{\partial y_j} = J\operatorname{div}\mathbf{A}
\end{aligned}
$$

where we used the Piola identity (2.120) in getting to the second line.

2.9 Linearization.

The *displacement field* $\mathbf{u}(\mathbf{x})$ is related to the deformation by

$$\mathbf{u}(\mathbf{x}) = \mathbf{y}(\mathbf{x}) - \mathbf{x}. \tag{2.97}$$

The associated *displacement gradient tensor*

$$\mathbf{H} \coloneqq \nabla\mathbf{u}, \tag{2.98}$$

has cartesian components

$$H_{ij} = \frac{\partial u_i}{\partial x_j}. \tag{2.99}$$

Since $\mathbf{F} = \nabla\mathbf{y}$, it follows from (2.97) and (2.98) that the displacement gradient tensor $\mathbf{H}$ and the deformation gradient tensor $\mathbf{F}$ are related by

$$\mathbf{F} = \mathbf{I} + \mathbf{H}. \tag{2.100}$$

The various kinematic quantities encountered previously, such as the stretches $\mathbf{U}, \mathbf{V}$, the rotation $\mathbf{R}$ and the strain $\mathbf{E}$, were all expressed in terms of the deformation gradient tensor $\mathbf{F}$, and so they can all be represented instead in terms of the displacement gradient tensor $\mathbf{H}$.

In the reference configuration we have $\mathbf{F} = \mathbf{I}$ and so $\mathbf{H} = \mathbf{0}$ by (2.100). If, in some sense (to be made precise), the body is deformed by a "small" amount, then $\mathbf{F}$ is close to $\mathbf{I}$ and $\mathbf{H}$ is close to $\mathbf{0}$. This is indeed the case in many physical circumstances and our goal in this section is to derive approximations for $\mathbf{U}, \mathbf{V}, \mathbf{R}, \mathbf{E}$ etc. in this particular setting. Note that when $\mathbf{F}$ is close to $\mathbf{I}$, both the stretch $\mathbf{U}$ *and* rotation $\mathbf{R}$ are close to $\mathbf{I}$, and so the setting we are considering involves small amounts of both stretching and rotation.

Before proceeding to derive suitable approximations, we first recall what we mean by the magnitude of a tensor and what it means when that magnitude is small. We note three preliminary algebraic results:

– First, recall from (1.118) that the norm (or magnitude) of a tensor $\mathbf{A}$ is defined as

$$|\mathbf{A}| = \sqrt{\mathbf{A}\cdot\mathbf{A}} \; = \; \sqrt{\operatorname{tr}(\mathbf{A}^T\mathbf{A})}. \tag{2.101}$$

In terms of the components A_{ij} this reads

$$|\mathbf{A}| = (A_{11}^2 + A_{12}^2 + A_{13}^2 + A_{21}^2 + \cdots + A_{33}^2)^{1/2}. \tag{2.102}$$

Observe that $|\mathbf{A}| > 0$ for all $\mathbf{A} \neq \mathbf{0}$. Moreover if $|\mathbf{A}| \to 0$ then *each* component $A_{ij} \to 0$ as well.

– Second, let $\mathbf{Z}(\mathbf{H})$ be a function that is defined for all 2-tensors $\mathbf{H}$ and whose values are also 2-tensors. We say that $\mathbf{Z}(\mathbf{H}) = \mathrm{O}(|\mathbf{H}|^n)$ as $|\mathbf{H}| \to 0$ if there exists a number $\alpha > 0$ such that $|\mathbf{Z}(\mathbf{H})| < \alpha|\mathbf{H}|^n$ as $|\mathbf{H}| \to 0$.

– And third, for a symmetric tensor $\mathbf{A}$ and real number m,

$$(\mathbf{I} + \mathbf{A})^m = \mathbf{I} + m\mathbf{A} + \mathrm{O}(|\mathbf{A}|^2) \quad \text{as} \quad |\mathbf{A}| \to 0 \tag{2.103}$$

which can be readily established in a principal basis of $\mathbf{A}$.

We are now in a position to linearize our preceding kinematic quantities in the special case of an **infinitesimal deformation**, defined as a deformation in which $\mathbf{H} = \nabla\mathbf{u} = \mathbf{F} - \mathbf{I}$ is small. To this end we set

$$\epsilon = |\mathbf{H}| \tag{2.104}$$

and conclude that as $\epsilon \to 0$,

$$\begin{aligned}
\mathbf{C} = \mathbf{U}^2 \;&=\; \mathbf{F}^T\mathbf{F} &&\overset{(2.100)}{=} \quad \mathbf{I} + \mathbf{H} + \mathbf{H}^T + \mathrm{O}(\epsilon^2),\\
\mathbf{B} = \mathbf{V}^2 \;&=\; \mathbf{F}\mathbf{F}^T &&\overset{(2.100)}{=} \quad \mathbf{I} + \mathbf{H} + \mathbf{H}^T + \mathrm{O}(\epsilon^2),\\
\mathbf{U} = \sqrt{\mathbf{U}^2} \;&=\; \sqrt{\mathbf{C}} &&\overset{(2.103)}{=} \quad \mathbf{I} + \tfrac{1}{2}(\boldsymbol{H} + \mathbf{H}^T) + \mathrm{O}(\epsilon^2),\\
\mathbf{V} = \sqrt{\mathbf{V}^2} \;&=\; \sqrt{\mathbf{B}} &&\overset{(2.103)}{=} \quad \mathbf{I} + \tfrac{1}{2}(\boldsymbol{H} + \mathbf{H}^T) + \mathrm{O}(\epsilon^2),\\
\mathbf{R} \;&=\; \mathbf{F}\mathbf{U}^{-1} &&\overset{(2.100),(2.103)}{=} \quad \mathbf{I} + \tfrac{1}{2}(\mathbf{H} - \mathbf{H}^T) + \mathrm{O}(\epsilon^2),
\end{aligned} \tag{2.105}$$

where the first two equations follow immediately by using $\mathbf{F} = \mathbf{I} + \mathbf{H}$, and we have used (2.103) in deriving the last three equations. By $(2.105)_1$ the Green Saint-Venant strain tensor can be written as

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I}) = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T) + \mathrm{O}(\epsilon^2). \tag{2.106}$$

More generally the Lagrangian strain tensor (2.67) can be linearized by Taylor expanding $e(\lambda)$ about $\lambda = 1$ giving

$$\mathbf{E}(\mathbf{U}) = \sum_{i=1}^{3} e(\lambda_i)\mathbf{r}_i \otimes \mathbf{r}_i = \sum_{i=1}^{3}(\lambda_i - 1)\mathbf{r}_i \otimes \boldsymbol{r}_i + \mathrm{O}\left(\epsilon^2\right) = \mathbf{U} - \mathbf{I} + \mathrm{O}\left(\epsilon^2\right) = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T) + \mathrm{O}(\epsilon^2), \tag{2.107}$$

where we have used the properties $e(1) = 0, e'(1) = 1$.

One can also show that

$$J = \det \mathbf{F} = 1 + \operatorname{tr}\mathbf{H} + \mathrm{O}\left(\epsilon^2\right) = 1 + \operatorname{Div}\mathbf{u} + \mathrm{O}\left(\epsilon^2\right), \tag{2.108}$$

where in terms of its cartesian components, $\operatorname{Div}\mathbf{u} = \operatorname{tr}(\nabla\mathbf{u}) = \partial u_i/\partial x_i$.

Observe from $(2.105)_{3,4}$ and (2.107) that the stretch and strain tensors depend on $\mathbf{H}$ through only its symmetric part. Therefore we define a 2-tensor $\boldsymbol{\varepsilon}$ by

$$\boldsymbol{\varepsilon} \;:=\; \tfrac{1}{2}(\mathbf{H} + \mathbf{H}^T), \tag{2.109}$$

with cartesian components

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{2.110}$$

Note that the stretch tensors $\mathbf{U}$ and $\mathbf{V}$ and a general Lagrangian strain tensor $\mathbf{E}$ can be approximated as

$$\mathbf{U} = \mathbf{I} + \boldsymbol{\varepsilon} + \mathrm{O}\left(\epsilon^2\right), \qquad \mathbf{V} = \mathbf{I} + \boldsymbol{\varepsilon} + \mathrm{O}\left(\epsilon^2\right), \qquad \mathbf{E} = \boldsymbol{\varepsilon} + \mathrm{O}\left(\epsilon^2\right). \tag{2.111}$$

The symmetric tensor $\boldsymbol{\varepsilon}$ is known as the **infinitesimal strain tensor** and plays a central role in the theory of solids undergoing infinitesimal deformations. It is worth comparing the expression (2.110) for the components of the infinitesimal strain tensor with the corresponding expression (2.73) for those of the Green Saint-Venant strain tensor.

Remark: In Problem 2.2 it is shown that $\boldsymbol{\varepsilon}$ does *not* vanish in a rigid rotation $\mathbf{y} = \mathbf{Q}\mathbf{x}$ and therefore is *not* a suitable measure of strain in a finite deformation. However it is shown there (and below) that $\boldsymbol{\varepsilon}$ does vanish to leading order in an infinitesimal rigid rotation and therefore *is* appropriate in the study of infinitesimal deformations.

Remark: An eigenvalue ε_i of $\boldsymbol{\varepsilon}$ is related to the corresponding principal stretch λ_i by

$$\lambda_i = 1 + \varepsilon_i + \mathrm{O}(\epsilon^2). \tag{2.112}$$

Remark: Upon linearization, equations (2.74) and (2.76) read

$$\varepsilon_{11} \approx \frac{ds_y - ds_x}{ds_x}, \qquad \varepsilon_{12} \approx \frac{1}{2}\cos\theta_y \approx \frac{1}{2}(\pi/2 - \theta_y). \tag{2.113}$$

It follows from this that when the deformation is infinitesimal, the normal strain component ε_{11} represents the change in length per reference length of a fiber that was in the x_1-direction in the reference configuration; and that the shear strain component ε_{12} represents one half the decrease in angle between two fibers that were in the x_1- and x_2-directions in the reference configuration.

Next, observe from $(2.105)_5$ that the rotation tensor $\mathbf{R}$ depends only on the skew-symmetric part of $\mathbf{H}$. Therefore we define a 2-tensor $\boldsymbol{\omega}$ by

$$\boldsymbol{\omega} \;:=\; \tfrac{1}{2}(\mathbf{H} - \mathbf{H}^T), \tag{2.114}$$

whose cartesian components are

$$\omega_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right). \tag{2.115}$$

Observe that the rotation tensor $\mathbf{R}$ can be approximated as

$$\mathbf{R} = \mathbf{I} + \boldsymbol{\omega} + \mathrm{O}\left(\epsilon^2\right). \tag{2.116}$$

The tensor $\boldsymbol{\omega}$ is known as the **infinitesimal rotation tensor** and it is skew-symmetric (not orthogonal!)

Consider the particular displacement field

$$\mathbf{u}(\mathbf{x}) = \mathbf{W}\mathbf{x} + \mathbf{b} \tag{2.117}$$

where $\mathbf{W}$ is a constant skew-symmetric tensor and $\mathbf{b}$ is a constant vector. It is readily seen by substituting this into (2.109) that the associated infinitesimal strain field vanishes. Thus (2.117) describes an **infinitesimal rigid displacement**.

Remark: It is useful to observe from (2.20), (2.100) and (2.114) that a fiber $d\mathbf{x}$ in the reference configuration and its image $d\mathbf{y}$ in the deformed configuration are related by

$$d\mathbf{y} = d\mathbf{x} + \boldsymbol{\varepsilon}\, d\mathbf{x} + \boldsymbol{\omega}\, d\mathbf{x} + \mathrm{O}(\epsilon^2), \tag{2.118}$$

which shows that in the linearized theory the local deformation can be *additively* decomposed into a strain and a rotation. This is in contrast to the multiplicative decomposition $\mathrm{d}\mathbf{y} = \mathbf{R}\mathbf{U}\mathrm{d}\mathbf{x}$ for a finite deformation.

Remark: Note from (2.108), (2.114) and (2.36) that

$$\frac{dV_y - dV_x}{dV_x} = J - 1 = \mathrm{Div}\, \mathbf{u} + \mathrm{O}(\epsilon^2) = \mathrm{tr}\,(\nabla \mathbf{u}) + \mathrm{O}(\epsilon^2) = \frac{\partial u_i}{\partial x_i} + \mathrm{O}(\epsilon^2) = \mathrm{tr}\, \boldsymbol{\varepsilon} + \mathrm{O}(\epsilon^2). \quad (2.119)$$

Thus the volumetric strain is measured by $\mathrm{tr}\, \boldsymbol{\varepsilon}$ in the infinitesimal deformation theory.

Remark: As noted previously, when $|\mathbf{H}|$ is small, *both* the strain and rotation are small. There are certain physical circumstances in which one wants to carry out a different linearization, i.e., linearization based on the smallness of some quantity other than $\nabla \mathbf{u}$. For example, consider rolling up a sheet of paper. If the rolled-up configuration is the deformed configuration and the flat one the reference configuration, in this situation one has large rotations $\mathbf{R}$ but small strains $\mathbf{U} - \mathbf{I}$. Thus one might wish to linearize based on the assumption that $|\mathbf{U} - \mathbf{I}|$ is small (but leave $\mathbf{R}$ arbitrary). Note that under these conditions $|\mathbf{H}|$ will not be small.

2.10 Exercises.

Problem 2.1. Bending of a block.

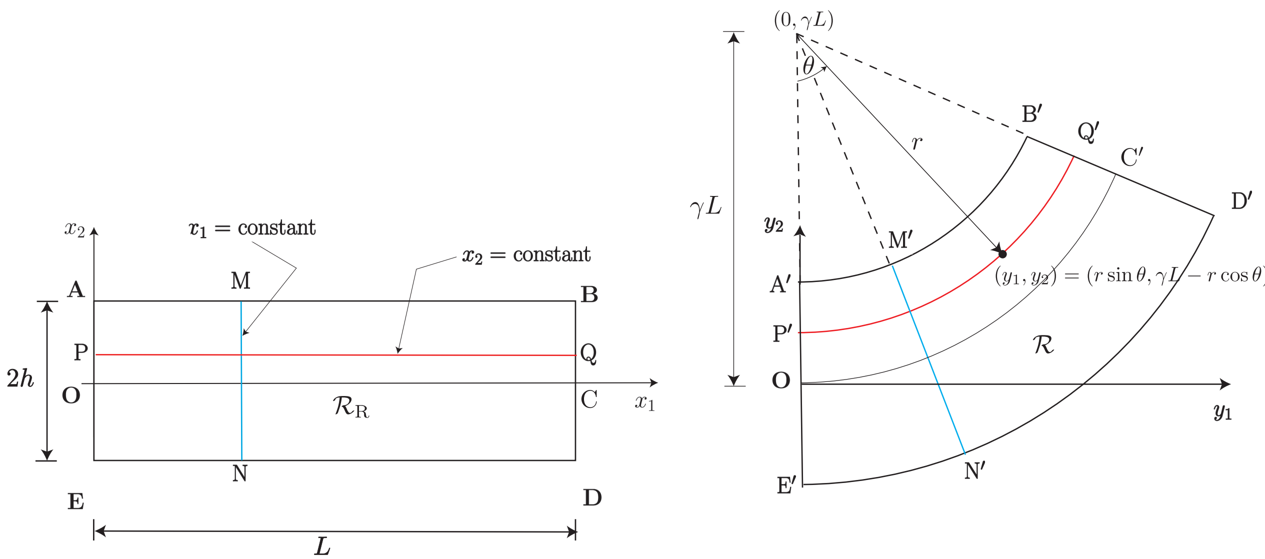


Figure 2.20: Left: Region $\mathcal{R}_{\rm R} = \{(x_1, x_2, x+_3) \, : \, 0 \le x_1 \le L, -h \le x_2 \le h, -b \le x_3 \le b\}$ occupied by a body in a reference configuration. Right: The region $\mathcal{R}$ occupied by the body in the deformed configuration. The points P′, Q′, M′, N′, etc. are the images in the deformed configuration of the points P, Q, M, N, etc. in the reference configuration. Vertical straight lines, e.g. MN, in the reference configuration are mapped into straight lines, e.g. M′N′, that pass through the point $(0, \gamma L)$. Horizontal straight lines in the reference configuration, e.g. PQ, are carried into circular arcs, e.g. P′Q′. (Figure for Problem 2.1)

A body occupies the region $\mathcal{R}_{\rm R} = \{(x_1, x_2, x_3) \, : \, 0 \le x_1 \le L, -h \le x_2 \le h, -b \le x_3 \le b\}$ in a reference configuration. Figure 2.20 shows a side view of this block looking down the x_3-axis. The body is subjected to a bending deformation that carries $\mathcal{R}_{\rm R}$ into the region $\mathcal{R}$ shown. The deformation has the features described in the figure caption and is therefore analogous to that in Problem 2.5.4. An analysis like that in Problem 2.5.4 shows that the deformation must have the form (Exercise)

$$y_1 = r(x_2)\sin\theta(x_1), \qquad y_2 = \gamma L - r(x_2)\cos\theta(x_1), \qquad y_3 = x_3, \tag{i}$$

where the functions $r(x_2)$ and $\theta(x_1)$ are subject to

$$r(0) = \gamma L, \qquad \theta(0) = 0, \tag{ii}$$

and

$$r(x_2) > 0, \qquad r'(x_2) < 0, \qquad \theta'(x_1) > 0. \tag{iii}$$

(a) Consider the (undeformed) material fiber $d\mathbf{x} = ds_x\,\mathbf{e}_1$ at an arbitrary point $(x_1, x_2, 0)$ in the body. Calculate its stretch $\lambda_1(x_1, x_2)$. Similarly consider the undeformed material fiber $d\mathbf{x} = ds_x\,\mathbf{e}_2$ and calculate its stretch $\lambda_2(x_1, x_2)$. Caution: we do not know that these are the *principal* stretches so the use of the symbols λ_1, λ_2 is probably not ideal.

(b) Suppose that a material fiber that lies on the x_2-axis in the undeformed configuration remains unstretched by the deformation. What additional information can you infer about $r(x_2)$ and $\theta(x_1)$?

(c) Specialize the stretch $\lambda_1(x_1, x_2)$ to the case where the material is incompressible. Show that $\lambda_1(x_1, x_2)$ varies nonlinearly with x_2. In addition, show that $\lambda_1(x_1, 0) = 1$, $\lambda_1(x_1, x_2) < 1$ for $x_2 > 0$ and $\lambda_1(x_1, x_2) > 1$ for $x_2 < 0$. Do **not** assume the material to be incompressible from here on.

(d) Calculate the components of the Green Saint-Venant and infinitesimal strain tensors $\mathbf{E}$ and $\boldsymbol{\varepsilon}$. Comment on the distinction.

(e) Under what conditions is the deformation infinitesimal? Specialize your preceding expressions for $\boldsymbol{\varepsilon}(\mathbf{x})$ and $\mathbf{u}(\mathbf{x})$ to this case. Make use of the fact that $\lambda_1 = \lambda_2 = 1$ in the reference configuration and therefore that in an infinitesimal deformation $|\lambda_1 - 1| \ll 1$ and $|\lambda_2 - 1| \ll 1$.

Problem 2.2. A rigid rotation of a body is described by the deformation $\mathbf{y} = \mathbf{Q}\mathbf{x}$ where $\mathbf{Q}$ is proper orthogonal. Consider the particular rigid rotation

$$\mathbf{Q} = \cos\theta(\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2) + \sin\theta(\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1) + \mathbf{e}_3 \otimes \mathbf{e}_3 \qquad (i)$$

that describes a rotation through an angle θ about the $\mathbf{e}_3$-axis; see Problem 1.4.5. The deformation $\mathbf{y} = \mathbf{Q}\mathbf{x}$ when written out explicitly in component form reads

$$\left.\begin{aligned} y_1 &= x_1\,\cos\theta + x_2\,\sin\theta, \\ y_2 &= -x_1\,\sin\theta + x_2\,\cos\theta, \\ y_3 &= x_3. \end{aligned}\right\} \qquad (ii)$$

Use (2.110) to calculate the components of the *infinitesimal strain tensor* $\boldsymbol{\varepsilon}$ associated with the deformation (ii). Explain why this strain tensor does not vanish even though the deformation is rigid.

Suppose that the deformation is infinitesimal in the sense that $|\theta| \ll 1$. Show that $\boldsymbol{\varepsilon}$ does vanish to leading order in this case. Moreover, show that the deformation (ii) reduces to the form (2.117) of an infinitesimal rigid rotation. *This shows that the infinitesimal strain tensor $\boldsymbol{\varepsilon}$ is not a good measure of strain for finite deformations but is appropriate for the study of infinitesimal deformations.*

Homogeneous deformations

Problem 2.3. (Based on Chadwick) (This problem will be revisited in Chapter 6 when we discuss the constitutive relation of an anisotropic material.) An *incompressible* body is reinforced by embedding two families of straight *inextensible* fibers in it as depicted in Figure 2.21. The fiber directions $\mathbf{m}_{\mathrm{R}}^{\pm}$ in the reference configuration are

$$\mathbf{m}_{\mathrm{R}}^{\pm} = \cos\Theta\,\mathbf{e}_1 \pm \sin\Theta\,\mathbf{e}_2, \qquad 0 < \Theta < \pi/2.$$

The body is subjected to the homogeneous deformation

$$y_1 = \lambda_1 x_1, \qquad y_2 = \lambda_2 x_2, \qquad y_3 = \lambda_3 x_3, \tag{o}$$

where the stretches λ_i's are > 0.

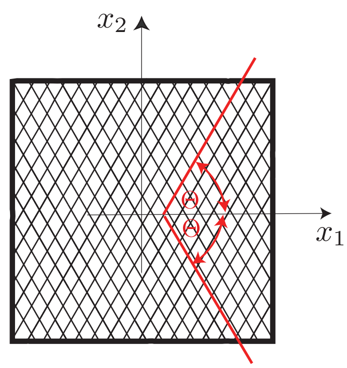


Figure 2.21: A cubic block reinforced with two families of inextensible fibers making angles $\pm\Theta$ with the x_1-axis in the reference configuration. (Figure for Problem 2.3)

(a) Show that in view of the kinematic constraints used in modeling the material, i.e. inextensibility and incompressibility, the only deformations (of the above form) that this body can sustain are those that obey

$$\lambda_1^2\cos^2\Theta + \lambda_2^2\sin^2\Theta = 1, \qquad \lambda_1\lambda_2\lambda_3 = 1.$$

(b) Show that the value of the stretch λ_1 is restricted to the range $0 < \lambda_1 < 1/\cos\Theta$. Why do you think λ_1 cannot be increased beyond a certain value?

(c) Let $\pm\theta$ be the angles that the fibers make with the y_1-axis in the deformed configuration. Analyze how θ varies as a function of λ_1. What value does θ approach when $\lambda_1 \to 1/\cos\Theta$? Now explain why the value of λ_1 cannot be increased beyond $1/\cos\Theta$.

(d) Analyze the variation of λ_2 as a function of λ_1. In particular, show that λ_2 decreases monotonically as λ_1 increases and that $\lambda_2 \to 0$ when $\lambda_1 \to 1/\cos\Theta$.

(e) Analyze the variation of λ_3 as a function of λ_1. In particular, show that as λ_1 increases, the body first contracts in the x_3-direction until λ_3 reaches the value $\sin 2\Theta$ and expands thereafter with $\lambda_3 \to \infty$ when $\lambda_1 \to 1/\cos\Theta$.

(f) Calculate the value of the angle between the two families of fibers in the deformed configuration when the body has contracted to its minimum value $\lambda_3 = \sin 2\Theta$ (corresponding to $\lambda_1 = 1/(\sqrt{2}\cos\Theta)$).

(g) Calculate the "Poisson's ratios" $-d\lambda_2/d\lambda_1$ at $\lambda_1 = 1$ and $-d\lambda_3/d\lambda_1$ at $\lambda_1 = 1$. Under what conditions (if any) do they take the value 1/2?

Problem 2.4. A body occupies a unit cube in a reference configuration:

$$\mathcal{R}_{\mathrm{R}} = \{(x_1, x_2, x_3) \,:\, -1/2 < x_1 < 1/2,\ -1/2 < x_2 < 1/2,\ -1/2 < x_3 < 1/2\}.$$

It is subjected to the deformation

$$y_1 = ax_1 + bx_2, \qquad y_2 = bx_1 + ax_2, \qquad y_3 = cx_3, \tag{i}$$

where a, b and c are constants. We are told that $c > 0$.

(a) Under what conditions on a, b, c does this deformation preserve orientation (in the sense that every right-handed linearly independent triplet of vectors is mapped into a right-handed triplet of vectors)?

(b) Consider a plane $\mathcal{S}_{\mathrm{R}}$ defined by $x_1 + x_2 = \mathsf{constant}$ in the reference configuration. Under what conditions on a, b, c does the area of a patch on this surface not change due to the deformation (i)?

(c) Under what conditions on a, b, c does *every* material fiber on the plane $\mathcal{S}_{\mathrm{R}}$ remain *unstretched* by the deformation?

Problem 2.5. (P. Rosakis) In this problem you are to show that, given the deformation of any *three* linearly independent material fibers, you can calculate $\mathbf{F}$, and therefore determine the deformation of *all* material fibers.

Consider three distinct non-coplanar material fibers identified by the three linearly independent vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. (These fibers need not be perpendicular to each other and need not have the same lengths.) The body is subjected to a homogeneous deformation that carries these fibers into $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ respectively. You are given $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{b}_1, \mathbf{b}_2$ and $\mathbf{b}_3$ and in order to establish the desired result, it suffices to derive a formula for the deformation gradient tensor in terms of these six vectors (alone).

Problem 2.6. (Based on Goriely et al.) A body undergoes a simple shear deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$, $\mathbf{F} = \mathbf{I} + k\,\mathbf{a}\otimes\mathbf{b}$, where $\mathbf{a}$ is the shearing direction and $\mathbf{b}$ is glide plane normal. Consider a plane $\mathcal{S}_{\mathrm{R}}$ in the reference configuration that is perpendicular to the unit vector $\mathbf{n}_{\mathrm{R}} = \cos\theta\,\mathbf{a} + \sin\theta\,\mathbf{b}$. Calculate the ratio $(\Delta A_y/\Delta A_x)^2$ where ΔA_x is the area of a surface element on $\mathcal{S}_{\mathrm{R}}$ and ΔA_y is the area of its image in the deformed configuration. Considering all such planes $\mathcal{S}_{\mathrm{R}}$, on which do the maximum and minimum values of $(\Delta A_y/\Delta A_x)^2$ occur? Calculate those values.

Problem 2.7. Let $\mathbf{y} = \mathbf{F}_1\mathbf{x}$ and $\mathbf{y} = \mathbf{F}_2\mathbf{x}$ be two arbitrary homogeneous deformations. Suppose that the deformation $\mathbf{y} = \mathbf{F}_1\mathbf{F}_2\mathbf{x}$ is a simple shear. Is the deformation $\mathbf{y} = \mathbf{F}_2\mathbf{F}_1\mathbf{x}$ also a simple shear? If it is, (either in general or under special circumstances), what is the associated amount of shear, glide plane normal and direction of shear?

Problem 2.8. (Ogden) Show that a simple shear with amount of shear k_1, shear direction $\mathbf{m}_1$ and glide plane normal $\mathbf{n}_1$ is commutative with a simple shear with amount of shear k_2, shear direction $\mathbf{m}_2$ and glide plane normal $\mathbf{n}_2$ if and only if

$$\text{either} \quad (a)\ \mathbf{m}_1 = \pm\mathbf{m}_2 \quad \text{or} \quad (b)\ \mathbf{n}_1 = \pm\mathbf{n}_2. \tag{i}$$

In case (a) show that the composite deformation is a simple shear with shear direction $\mathbf{m}_1$, glide plane normal $k_1\mathbf{n}_1 \pm k_2\mathbf{n}_2$ and amount of shear $(k_1^2 + k_2^2 \pm 2k_1k_2\mathbf{n}_1 \cdot \mathbf{n}_2)^{1/2}$. In case (b) show that the composite deformation is a simple shear with shear direction $k_1\mathbf{m}_1 \pm k_2\mathbf{m}_2$, glide plane normal $\mathbf{n}$ and amount of shear $(k_1^2 + k_2^2 \pm 2k_1k_2\mathbf{m}_1 \cdot \mathbf{m}_2)^{1/2}$.

Problem 2.9.

(a) Under what conditions does the direction of a material fiber remain unchanged (invariant) in a given deformation?

(b) The region $\mathcal{R}_\mathrm{R}$ occupied by a body in a reference configuration is a unit cube. The orthonormal basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are aligned with the edges of the cube. Consider the following *isochoric* deformation:

$$\mathbf{y} = \big(\lambda_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3\big)\big(\mathbf{I} + k\,\mathbf{e}_1 \otimes \mathbf{e}_2\big)\mathbf{x}, \qquad \lambda_1 \neq 1, \lambda_2 \neq 1, k \neq 0. \tag{i}$$

Describe the physical nature of this deformation and list as many invariant directions as you can (based on your intuition).

(c) Now show mathematically that there are exactly three directions that remain invariant in this deformation and determine them.

Problem 2.10. (Ogden) Consider a planar pure stretch

$$\mathbf{F} = \lambda_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{i}$$

Let

$$\mathbf{m}_\mathrm{R}^{(1)} = \cos\Theta\,\mathbf{e}_1 + \sin\Theta\,\mathbf{e}_2, \qquad \mathbf{m}_\mathrm{R}^{(2)} = -\sin\Theta\,\mathbf{e}_1 + \cos\Theta\,\mathbf{e}_2, \tag{ii}$$

be the (mutually orthogonal) directions of two material fibers in the reference configuration.

(a) Calculate the change in angle between them due to the deformation (i).

(b) Show that the maximum angle change among all such pairs of material fibers is

$$\sin^{-1}\left(\frac{\lambda_1^2-\lambda_2^2}{\lambda_1^2+\lambda_2^2}\right).$$

Problem 2.11. Calculate the components of the Lagrangian logarithmic strain tensor $\mathbf{E} = \ln \mathbf{U}$ associated with a simple shear deformation

$$y_1 = x_1 + kx_2, \qquad y_2 = x_2, \qquad y_3 = x_3.$$

Cylindrical and spherical bodies.

Problem 2.12. *(Deformation of a hollow circular tube)* A body occupies a hollow circular cylindrical region $\mathcal{R}_{\rm R}$ in a reference configuration with inner radius A, outer radius B and length L:

$$\mathcal{R}_{\rm R} = \{(x_1,x_2,x_3) : A < (x_1^2+x_2^2)^{1/2} < B,\ 0 < x_3 < L\}.$$

All components of vectors and tensors are taken with respect to the right-handed orthonormal basis $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ shown in Figure 2.22. A particle located at (x_1,x_2,x_3) in the reference configuration is carried to (y_1,y_2,y_3) by the deformation

$$\left.\begin{aligned} y_1 &= f(R)\Big[x_1\cos\phi(x_3) - x_2\sin\phi(x_3)\Big],\\ y_2 &= f(R)\Big[x_2\cos\phi(x_3) + x_1\sin\phi(x_3)\Big],\\ y_3 &= \Lambda x_3, \end{aligned}\right\}, \qquad R = (x_1^2+x_2^2)^{1/2}. \qquad (i)$$

Here $f(R)$ and $\phi(x_3)$ are smooth functions defined for $A \le R \le B$ and $0 \le x_3 \le L$ respectively and $\Lambda > 0$ is a constant. Describe the physical nature of this deformation: in particular, consider the particles that, in the undeformed configuration, lie on a circle $x_1^2+x_2^2 = c^2$ (on the cross section at some fixed x_3). Determine (and describe) the curve on which these particles lie in the deformed configuration. Do the same for the particles on a radial straight line $x_2 = cx_1$. Determine the region $\mathcal{R}$ occupied by the body in the deformed configuration.

Problem 2.13. (Spencer) *(Deformation of a solid circular cylinder)* The region occupied by a body in a reference configuration is a solid circular cylinder of radius A. Coordinate axes are chosen such that the axis of the cylinder coincides with the x_3-axis. The body undergoes the deformation

$$\left.\begin{aligned} y_1 &= \lambda[x_1\cos(\alpha x_3) + x_2\sin(\alpha x_3)],\\ y_2 &= \lambda[-x_1\sin(\alpha x_3) + x_2\cos(\alpha x_3)],\\ y_3 &= \Lambda x_3, \end{aligned}\right\} \qquad (i)$$

where α, λ and Λ are positive constants.

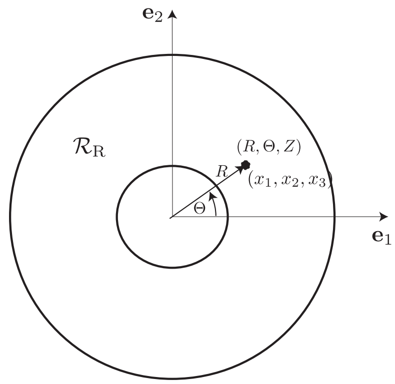


Figure 2.22: Cross-section of the region $\mathcal{R}_{\mathrm{R}}$ occupied by the body in a reference configuration: a hollow circular cylinder of inner radius A, outer radius B (and length L). (Figure for Problem 2.12)

(a) Describe this deformation.

(b) Consider a material fiber on the outer surface of the cylinder that, in the *reference* configuration is parallel to the axis of the cylinder. Calculate the stretch of this material fiber due to the deformation. Observe from your answer that though this referential fiber lies in the x_3-direction, it is not Λ alone that contributes to its stretch. Can you derive your answer by "physical" (elementary geometric) arguments alone?

(c) Consider a material fiber on the outer surface of the cylinder that, in the *deformed* configuration, is parallel to the axis of the cylinder. Calculate the stretch of this fiber. You may find useful the result in Problem 1.30, regarding the inverse of a tensor expressed in a mixed basis. An alternative way in which to calculate $\mathbf{F}^{-1}$ is to realize that it is the deformation gradient tensor for the inverse deformation $R = R(r, \theta, z), \Theta = \Theta(r, \theta, z), Z = Z(r, \theta, z)$.

Problem 2.14. *(Inflation and extension of a hollow circular tube)* The region $\mathcal{R}_{\mathrm{R}}$ occupied by a body in a reference configuration is a hollow circular cylinder of inner radius A, outer radius B and length L. It is subjected to the radially symmetric deformation

$$r = r(R), \qquad \theta = \Theta, \qquad z = \Lambda\, Z, \qquad (i)$$

where (R, Θ, Z) and (r, θ, z) are the respective cylindrical polar coordinates of a particle in the reference and deformed configurations, where $\Lambda > 0$ is a constant and $r(R) > 0$.

(a) Calculate the principal stretches.

(b) Determine $r(R)$ (to the extent possible) if the material is incompressible. Assume that $r'(R) > 0$.

(c) Denote the stretch in the circumferential direction by $\lambda(R)$:

$$\lambda(R) = r(R)/R. \qquad (ii)$$

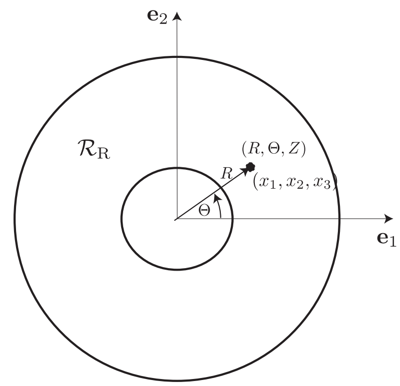


Figure 2.23: The region $\mathcal{R}_\mathrm{R}$ occupied by a body in a reference configuration is a hollow circular cylinder of inner radius A and outer radius B. (Figure for Problem 2.14)

Show that

$$A^2(\lambda_a^2\Lambda - 1) = B^2(\lambda_b^2\Lambda - 1) = R^2(\lambda^2(R)\Lambda - 1), \qquad (iii)$$

where $\lambda_a = \lambda(A), \lambda_b = \Lambda(B)$.

Problem 2.15. (Ogden) *(Combined axial and azimuthal shear of a tube)* Let (R, Θ, Z) and (r, θ, z) be cylindrical polar coordinates of a particle in the reference and deformed configurations respectively with associated bases $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. The region $\mathcal{R}_R$ occupied by the body in a reference configuration is a hollow circular cylinder of inner radius A, outer radius B and length L.

Consider a deformation of the form

$$r = R, \qquad \theta = \Theta + \phi(R), \qquad z = Z + w(R). \qquad (i)$$

In the special case $\phi \equiv 0$ this describes an "axial (telescopic) shearing" of the tube while the special case $w \equiv 0$ describes an "azimuthal shearing". Note that neither of these particular deformations, nor (i), is a torsional deformation.

We will study the stress in this tube in Problem 3.10.1 of Chapter 3.

(a) Calculate the deformation gradient tensor $\mathbf{F}$.

(b) By factoring $\mathbf{F}$ into the product of three tensors, show that locally, at each point of the body, the deformation is comprised of a rigid rotation, followed by a simple shear with glide plane normal $\mathbf{e}_r$ and shear direction $\mathbf{e}_\theta$, followed be a simple shear with glide plane normal $\mathbf{e}_r$ and shear direction $\mathbf{e}_z$. Determine the associated amounts of shear.

(c) If the material is incompressible, what does this tell you (if anything) about $\phi(R)$ and $w(R)$?

(d) Show that the composition of the two simple shears in part (*b*) corresponds to a simple shear with shearing direction $\mathbf{a}$, glide plane normal $\mathbf{e}_r$ and amount of shear k where

$$\mathbf{a} = \sin\beta\,\mathbf{e}_\theta + \cos\beta\,\mathbf{e}_z, \quad \tan\beta = \frac{k_1}{k_2}, \qquad k = k = \left[k_1^2 + k_2^2\right]^{1/2}, \quad k_1 = R\phi'(R), \quad k_2 = w'(R). \quad (ii)$$

Remark: Note that you now have the deformation gradient tensor factored as $\mathbf{F} = \mathbf{KQ}$ where $\mathbf{Q}$ is a rotation and $\mathbf{K}$ a simple shear. Keep in mind that this is *not* the polar decomposition of $\mathbf{F}$ and that $\mathbf{K} \neq \mathbf{V}$.

(e) Calculate the matrix of components of the left Cauchy Green deformation tensor $\mathbf{B}$ in the basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. (Express your answer in terms of k_1 and k_2.)

(f) Calculate the principal stretches and principal Eulerian stretch directions. Hint: You can make use the results of Problem 2.36. (Express your answer in terms of $k, \mathbf{a}, \mathbf{e}_r, \mathbf{e}_\theta$ and $\mathbf{e}_z$.)

Problem 2.16. (Ogden) *(Inflation of a hollow spherical shell)* Let (R, Θ, Φ) and (r, θ, φ) be spherical polar coordinates in the reference and deformed configurations respectively with associated bases $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$. The hollow spherical region $\mathcal{R}_R$ occupied by an incompressible body in a reference configuration is described by

$$A \leq R \leq B, \qquad 0 \leq \Theta \leq 2\pi, \qquad 0 \leq \Phi \leq 2\pi. \qquad (i)$$

The body is subjected to a spherically symmetric deformation

$$r = r(R), \qquad \theta = \Theta, \qquad \varphi = \Phi, \qquad (ii)$$

and its inner and outer radii in the deformed configuration are a and b respectively.

(a) Calculate the principal stretches.

(b) Let

$$\lambda(R) = r(R)/R, \qquad \lambda_a = r(A)/A, \qquad \lambda_b = r(B)/B; \qquad (iii)$$

these are the stretches in the circumferential direction at $R, R = A$ and $R = B$ respectively. Show that

$$A^3(\lambda_a^3 - 1) = R^3(\lambda^3 - 1) = B^3(\lambda_b^3 - 1), \qquad (iv)$$

and hence show that either $\lambda_a \geq \lambda_b \geq 1$ or $\lambda_a \leq \lambda_b \leq 1$.

(c) Now suppose that the sphere is thin-walled in the sense that

$$\varepsilon := T/R \ll 1, \qquad (v)$$

where $T = B - A$ and $R = (A+B)/2$ are the wall-thickness and mean radius of the sphere respectively in the reference configuration. Let λ_a and λ_b be the stretch at the inner and outer wall as in (*iii*) above, and let $\lambda = r/R$ be the *mean* stretch where r is the *mean* radius of the deformed sphere. Derive approximate expressions for the stretches λ_a and λ_b keeping terms of order ε. Your results will be of the form

$$\lambda_a = \lambda + \Delta\lambda_a\,\varepsilon + \mathrm{O}(\varepsilon^2), \qquad \lambda_b = \lambda + \Delta\lambda_b\,\varepsilon + \mathrm{O}(\varepsilon^2). \qquad (vi)$$

Problem 2.17. *(Eversion of a circular cylindrical tube)* A hollow circular cylindrical tube has inner and outer radii A and B and length L in a reference configuration. Choose rectangular cartesian coordinates such that the region occupied by the body in this configuration is

$$\mathcal{R}_{\mathrm{R}} = \{(x_1, x_2, x_3) \ : A^2 < x_1^2 + x_2^2 < B^2, \ -L/2 < x_3 < L/2\}.$$

Consider "everting" the tube by turning it inside out – imagine a sock being turned inside out. In particular, this deformation maps the inner surface in the reference configuration into the outer surface in the deformed configuration and the outer surface in the reference configuration into the inner surface in the deformed configuration. Assume that the everted shape of the body is a hollow circular cylinder of inner and outer radii a and b and length ℓ. (In order to maintain the body in this particular deformed configuration one may have to apply a suitable loading on the tube. Otherwise the everted body may not be a hollow circular cylinder with flat ends.) If (r, θ, z) and (R, Θ, Z) denote the cylindrical polar coordinates of a particle in the deformed and reference configurations, take the deformation to have the form

$$r = r(R), \qquad \theta = \Theta, \qquad z = z(Z), \tag{i}$$

where

$$r(A) = b, \qquad r(B) = a, \qquad z(L/2) = -\ell/2, \qquad z(-L/2) = \ell/2. \tag{ii}$$

(a) Determine the deformation, i.e. $r(R)$ and $z(Z)$, assuming

  (a1) that radial and axial fibers do not stretch, and alternatively

  (a2) that the deformation is isochoric. Calculate $\mathbf{C}, \mathbf{U}$ and $\mathbf{R}$.

(b) Verify that if you repeat this deformation, i.e. you evert the deformed configuration, you recover the reference configuration.

Problem 2.18. Calculate explicit expressions for the deformation gradient tensor $\mathbf{F}$ and the left and right Cauchy-Green tensors $\mathbf{B}$ and $\mathbf{C}$ using spherical polar coordinates (R, Θ, Φ) in the undeformed configuration and (r, θ, φ) in the deformed configuration.

Some general considerations.

Problem 2.19. The stretch $\lambda(\mathbf{m}_{\mathrm{R}})$ of a fiber oriented (in the reference configuration) in the direction $\mathbf{m}_{\mathrm{R}}$ is

$$\lambda(\mathbf{m}_{\mathrm{R}}) = |\mathbf{F}\mathbf{m}_{\mathrm{R}}|. \tag{i}$$

(a) Maximize $\lambda(\mathbf{m}_{\mathrm{R}})$ over all fiber directions $\mathbf{m}_{\mathrm{R}}$.

(b) Show that $\lambda(\mathbf{m}_{\mathrm{R}})$ can be written in terms of the principal stretches as

$$\lambda(\mathbf{m}_{\mathrm{R}}) = \sqrt{\lambda_1^2 m_1^2 + \lambda_2^2 m_2^2 + \lambda_3^2 m_3^2}, \qquad (ii)$$

where the m_k's are the components of $\mathbf{m}_{\mathrm{R}}$ in the principal basis of the right stretch tensor $\mathbf{U}$.

Problem 2.20. Show for any isochoric deformation that

$$I_1(\mathbf{C}) \geq 3, \qquad (i)$$

where $I_1(\mathbf{C}) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2$ is the first principal invariant of $\mathbf{C}$. Moreover, show that $I_1 = 3$ if and only if $\lambda_1 = \lambda_2 = \lambda_3 = 1$.

Likewise show for isochoric deformations that

$$I_2(\mathbf{C}) \geq 3, \qquad (ii)$$

where $I_2(\mathbf{C}) = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2$ is the second principal invariant of $\mathbf{C}$, and that $I_2 = 3$ if and only if $\lambda_1 = \lambda_2 = \lambda_3 = 1$.

Problem 2.21. (Piola identity) Show that

$$\int_{\mathcal{D}} \mathbf{n}\, dA_y = \mathbf{0}$$

where $\mathcal{D}$ is an arbitrary subregion of $\mathcal{R}$ and $\mathbf{n}$ is the outward pointing unit vector normal to its boundary $\partial\mathcal{D}$.

By using the relation $\mathbf{n} dA_y = J\mathbf{F}^{-T}\mathbf{n}_{\mathrm{R}} dA_x$ or otherwise, show that

$$\mathrm{Div}(J\mathbf{F}^{-T}) = \mathbf{0}, \qquad (2.120)$$

where for any tensor field $\mathbf{A}(\mathbf{x})$, $\mathrm{Div}\,\mathbf{A}$ denotes the vector field with cartesian components $\partial A_{ij}/\partial x_j$.

Similarly show that

$$\mathrm{div}(J^{-1}\mathbf{F}^T) = \mathbf{0}, \qquad (2.121)$$

where for any tensor field $\mathbf{A}(\mathbf{y})$, $\mathrm{div}\,\mathbf{A}$ denotes the vector field with cartesian components $\partial A_{ij}/\partial y_j$.

Problem 2.22. Show that

$$\frac{\partial \lambda_i}{\partial \mathbf{C}} = \frac{1}{2\lambda_i}\mathbf{r}_i \otimes \mathbf{r}_i \qquad \text{and} \qquad \frac{\partial \lambda_i}{\partial \mathbf{F}} = \boldsymbol{\ell}_i \otimes \mathbf{r}_i, \qquad (i)$$

where *the summation convention for repeated subscript is suspended.*

Problem 2.23. Let $i_1(\mathbf{E}), i_2(\mathbf{E}), i_3(\mathbf{E})$ be the principal scalar invariants of the Green Saint-Venant strain tensor $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$:

$$i_1 = \operatorname{tr}\mathbf{E}, \qquad i_2 = \frac{1}{2}(i_1^2 - \operatorname{tr}\mathbf{E}^2), \qquad i_3 = \det \mathbf{E}. \tag{i}$$

Show that a deformation is isochoric if

$$i_1 + 2i_2 + 4i_3 = 0. \tag{ii}$$

Problem 2.24. Show that the cartesian components $F_{ij}(\mathbf{x})$ of the deformation gradient tensor field must necessarily obey the following system of partial differential equations (**compatibility equations**)

$$\frac{\partial F_{ij}}{\partial x_k} = \frac{\partial F_{ik}}{\partial x_j}. \tag{2.122}$$

Problem 2.25. Suppose that the right Cauchy-Green tensor field equals the identity at every point in the body: $\mathbf{C}(\mathbf{x}) = \mathbf{I}$ for all $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$. Show that the deformation gradient tensor field $\mathbf{F}(\mathbf{x})$ is orthogonal at each $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$. *Moreover*, show that $\mathbf{F}(\mathbf{x})$ is independent of $\mathbf{x}$ and therefore is a *constant* orthogonal tensor.

Problem 2.26. Suppose that the deformation gradient tensor field has the form $\mathbf{F}(\mathbf{x}) = \nabla\mathbf{y}(\mathbf{x}) = \phi(\mathbf{x})\mathbf{A}$ where ϕ is a positive scalar-valued function of $\mathbf{x}$ and $\mathbf{A}$ is a constant tensor with $\det \mathbf{A} > 0$. Show that $\phi(\mathbf{x})$ must be independent of $\mathbf{x}$ and therefore a constant.

Problem 2.27. Consider two deformations $\mathbf{y} = \mathbf{y}_1(\mathbf{x})$ and $\mathbf{y} = \mathbf{y}_2(\mathbf{x})$ related by a rigid deformation, i.e. the deformations are such that $\mathbf{y}_2(\mathbf{x}) = \mathbf{Q}\mathbf{y}_1(\mathbf{x}) + \mathbf{b}$ where $\mathbf{Q}$ is a constant rotation tensor and $\mathbf{b}$ is a constant vector. Show that the right Cauchy-Green tensors $\mathbf{C}_1$ and $\mathbf{C}_2$ associated with these two deformations coincide: $\mathbf{C}_1 = \mathbf{C}_2$.

What is the corresponding relation between the left Cauchy-Green tensors $\mathbf{B}_1$ and $\mathbf{B}_2$?

Conversely, is it true that if $\mathbf{C}_1(\mathbf{x}) = \mathbf{C}_2(\mathbf{x})$ at each $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$ then the two deformations differ by a rigid deformation?

Problem 2.28. Among the various experiments on rubber that Rivlin and Saunders [6] carried out are some where they subjected a thin incompressible rubber sheet to a pure stretch

$$y_1 = \lambda_1 x_1, \qquad y_2 = \lambda_2 x_2, \qquad y_3 = \lambda_3 x_3. \tag{i}$$

They varied the stretches λ_1, λ_2 keeping the value of the invariant $I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2$ fixed[17]; see Section 4.6.1. Show that

$$\lambda_2^2 = \frac{1}{2}\Big[I_1 - \lambda_1^2 \pm \sqrt{(I_1 - \lambda_1^2)^2 - 4/\lambda_1^2}\Big], \qquad (ii)$$

and sketch the contours of the closed curves defined by (ii) in the λ_1, λ_2-plane corresponding to different fixed values of I_1.

Problem 2.29. (See also Problem 2.30.) Consider a planar surface $\mathcal{S}_{\mathrm{R}}$ that passes through the region $\mathcal{R}_{\mathrm{R}}$ occupied by a body in the reference configuration. Let $\mathbf{n}_{\mathrm{R}}$ be a unit vector normal to $\mathcal{S}_{\mathrm{R}}$ and let $\mathcal{R}_{\mathrm{R}}^+$ denote the side into which $\mathbf{n}_{\mathrm{R}}$ points, $\mathcal{R}_{\mathrm{R}}^-$ the other side. Thus $\mathcal{S}_{\mathrm{R}}$ is a planar interface between two parts of the body. Consider the piecewise homogeneous deformation

$$\mathbf{y} = \begin{cases} \mathbf{F}^+\mathbf{x} & \text{for } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}^+, \\ \mathbf{F}^-\mathbf{x} & \text{for } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}^-, \end{cases} \qquad (i)$$

where $\mathbf{F}^\pm$ are constant non-singular tensors. Show that this deformation is continuous across $\mathcal{S}_{\mathrm{R}}$ if and only if there is a constant vector $\mathbf{a}$ for which

$$\mathbf{F}^+ - \mathbf{F}^- = \mathbf{a} \otimes \mathbf{n}_{\mathrm{R}}. \qquad (2.123)$$

This is known as the *Hadamard jump (compatibility) condition.* It plays an important role in studying interfaces between two material phases.

Interpret (2.123) geometrically. Specifically, with $\mathcal{S}$ being the image of $\mathcal{S}_{\mathrm{R}}$ in the deformed configuration, show that $\mathbf{F}^+$ differs from $\mathbf{F}^-$ by a simple shear with shearing direction $\mathbf{e}$ and glide plane normal $\mathbf{n}$ and a uniaxial extension in the direction $\mathbf{n}$. (Here the unit vectors $\mathbf{e}$ and $\mathbf{n}$ are in the plane $\mathcal{S}$ and normal to $\mathcal{S}$ respectively).

Problem 2.30. (See also Problem 2.29.) Now consider the time-dependent version of Problem 2.29. Consider a planar surface $\mathcal{S}_t$ that passes through the region $\mathcal{R}_{\mathrm{R}}$ occupied by a body in the reference configuration. The surface propagates through the reference configuration with velocity $V_n \mathbf{n}_R$ where V_n is the constant propagation speed and $\mathbf{n}_R$ is the constant unit vector that is normal to $\mathcal{S}_t$ at all times. Note that since this surface propagates through the reference configuration, different particles lie on it at different times, and therefore it is not a material surface. (This is in contrast to the interface between two materials in a composite material.) Let $\mathcal{R}_{\mathrm{R}t}^+$ denote the side into which $\mathbf{n}_{\mathrm{R}}$ points, $\mathcal{R}_{\mathrm{R}t}^-$ the other side. Consider the piecewise homogeneous motion

$$\mathbf{y}(\mathbf{x}, t) = \begin{cases} \mathbf{F}^+\mathbf{x} + \mathbf{v}^+ t & \text{for } \mathbf{x} \in \mathcal{R}_{\mathrm{R}t}^+,\ t > t_0, \\ \mathbf{F}^-\mathbf{x} + \mathbf{v}^- t & \text{for } \mathbf{x} \in \mathcal{R}_{\mathrm{R}t}^-,\ t > t_0, \end{cases}$$

[17] They also did experiments keeping I_2 fixed.

where $\mathbf{F}^{\pm}$ are constant non-singular tensors and $\mathbf{v}^{\pm}$ are constant vectors. Show that this motion is continuous across $\mathcal{S}_t$ at all times if and only if there is a constant vector $\mathbf{a}$ for which

$$\mathbf{F}^{+} - \mathbf{F}^{-} = \mathbf{a} \otimes \mathbf{n}_{\mathrm{R}}, \tag{2.124}$$

and

$$\mathbf{v}^{+} - \mathbf{v}^{-} = -V_n(\mathbf{F}^{+} - \mathbf{F}^{-})\mathbf{n}_R. \tag{2.125}$$

These *Hadamard jump (compatibility) conditions* generalize the special one in Problem 2.29. They play an important role in studying interfaces between two material phases.

Problem 2.31. This problem arises when studying the microstructure of a certain two-phase material. In one phase, the crystallographic lattice underlying the material is cubic and this is called the austenite phase. In the other, the lattice is tetragonal and this phase is called the martensite phase. There are three variants of the martensite phase. Suppose that the reference configuration $\mathbf{F} = \mathbf{I}$ corresponds to the austenite phase. The three stretch tensors $\mathbf{U}_k$ given below in $(i), (ii)$ describe the deformation from the austenite phase into the three martensite variants.

The answer to question (a) below is yes and therefore an interface, oriented in a specific way, separating one martensite variant from another can exist. The answer to question (b) is no and therefore an interface separating one martensite variant from austenite cannot exist. *Reference:* K. Bhattacharya, *Microstructure of Martensite*, Oxford, 2003.

Let $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ be an orthonormal basis. Consider the 3 symmetric positive definite tensors $\mathbf{U}_1, \mathbf{U}_2$ and $\mathbf{U}_3$ defined by

$$\mathbf{U}_k = \alpha\mathbf{I} + (\beta - \alpha)\mathbf{r}_k \otimes \mathbf{r}_k, \qquad \alpha \neq 1,\ \beta \neq 1,\ \alpha \neq \beta,\ \alpha > 0,\ \beta > 0, \qquad k = 1, 2, 3. \tag{i}$$

Here α and β are constant (lattice) parameters. The components of these three tensors in the basis $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ are

$$[U_1] = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}, \qquad [U_2] = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \alpha \end{pmatrix}, \qquad [U_3] = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}. \tag{ii}$$

Therefore the deformation $\mathbf{y} = \mathbf{U}_1\mathbf{x}$ takes a $1 \times 1 \times 1$ cube and stretches it in the $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ directions by stretches β, α, α and maps the cube into a $\beta \times \alpha \times \alpha$ tetragon. The deformations $\mathbf{y} = \mathbf{U}_2\mathbf{x}$ and $\mathbf{y} = \mathbf{U}_3\mathbf{x}$ are similar with the stretch by β being in the $\mathbf{r}_2$ and $\mathbf{r}_3$ directions respectively. Consider the Hadamard compatibility condition (2.124):

$$\mathbf{F}^{+} = \mathbf{F}^{-} = \mathbf{a} \otimes \mathbf{n}_{\mathrm{R}}. \tag{iii}$$

(a) Here you want to study the possibility of a two-phase deformation that involves two variants of martensite, one on each side of the interface. Accordingly take $\mathbf{F}^{+} = \mathbf{R}^{+}\mathrm{U}_2,\ \mathbf{F}^{-} = \mathbf{U}_1$ in (iii). Do there exist a proper orthogonal tensor $\mathbf{R}^{+}$, a unit vector $\mathbf{n}_{\mathrm{R}}$ and a vector $\mathbf{a}$ such that (iii) holds? If yes, find $\mathbf{R}^{+}, \mathbf{n}_{\mathrm{R}}$ and $\mathbf{a}$.

(b) Here you want to study the possibility of a two-phase deformation that involves austenite on one side of the interface and one variant of martensite on the other. Accordingly take $\mathbf{F}^+ = \mathbf{R}^+\mathbf{U}_2,\ \mathbf{F}^- = \mathbf{I}$. Do there exist a proper orthogonal tensor $\mathbf{R}^+$, a unit vector $\mathbf{n}_\mathrm{R}$ and a vector $\mathbf{a}$ such that (*iii*) holds? If yes, find $\mathbf{R}^+, \mathbf{n}_\mathrm{R}$ and $\mathbf{a}$.

Problem 2.32. *(Average deformation gradient tensor.)* Define the average value of the deformation gradient tensor field in a body to be

$$\overline{\mathbf{F}} := \frac{1}{\mathsf{vol}} \int_{\mathcal{R}_\mathrm{R}} \mathbf{F}(\mathbf{x})\, dV_x,$$

where vol is the volume of the region $\mathcal{R}_\mathrm{R}$. Show that

$$\overline{\mathbf{F}} = \frac{1}{\mathsf{vol}} \int_{\partial\mathcal{R}_\mathrm{R}} \mathbf{y}(\mathbf{x}) \otimes \mathbf{n}_\mathrm{R}\, dA_x$$

and therefore that the average value of the deformation gradient tensor field in a body depends only on the deformation of the boundary $\partial\mathcal{R}_\mathrm{R}$. Show this

- first in the case where $\mathbf{F}(\mathbf{x})$ is continuous on $\mathcal{R}_\mathrm{R}$, and
- second in the case where $\mathbf{F}(\mathbf{x})$ is piecewise continuous on $\mathcal{R}_\mathrm{R}$. Specifically, suppose there is a surface $\mathcal{S}_\mathrm{R} \subset \mathcal{R}_\mathrm{R}$ with $\mathbf{F}(\mathbf{x})$ being continuous on either side of $\mathcal{S}_\mathrm{R}$ but discontinuous at $\mathcal{S}_\mathrm{R}$ but with the deformation $\mathbf{y}(\mathbf{x})$ being continuous on $\mathcal{R}_\mathrm{R}$ including on $\mathcal{S}_\mathrm{R}$.

Problem 2.33. *Decomposition of an arbitrary isochoric planar deformation gradient tensor:* Show that any planar isochoric deformation deformation gradient tensor $\mathbf{F}$ is equivalent to a suitable simple shear followed by a rotation, i.e. show that one can express such a tensor $\mathbf{F}$ as

$$\mathbf{F} = \mathbf{OK},$$

where $\mathbf{O}$ is proper orthogonal and $\mathbf{K} = \mathbf{I} + k\mathbf{a} \otimes \mathbf{b}$ for some scalar k and mutually orthogonal unit vectors $\mathbf{a}$ and $\mathbf{b}$. Note: if the deformation is planar in the plane spanned by $\mathbf{r}_1$ and $\mathbf{r}_2$ then $\mathbf{a}$ and $\mathbf{b}$ are in that same plane and the rotation $\mathbf{O}$ is about $\mathbf{r}_3$.

Problem 2.34. *Decomposition of an arbitrary isochoric deformation gradient tensor:* Show that an arbitrary isochoric homogeneous deformation can be viewed as a uniaxial extension with accompanying lateral contraction, a simple shear in the plane normal to the direction of extension, and a rigid rotation.

Problem 2.35. (See also Problem 2.5.2 .) Calculate the principal stretches associated with the simple shear

$$y_1 = x_1 + kx_2, \qquad y_2 = x_2, \qquad y_3 = x_3, \qquad k > 0. \tag{i}$$

In Problem 2.5.2 we used a direct (but tedious) way by calculating $\mathbf{F}$, then $\mathbf{C} = \mathbf{F}^T\mathbf{F}$, determining its eigenvalues, and taking their square roots. Instead, carry out your calculations by getting two different expressions for the first invariant $I_1(\mathbf{C})$ and equating them, keeping in mind that the deformation is planar and isochoric.

Problem 2.36. In Problem 2.5.2 we determined the rotation $\mathbf{R}$ and Lagrangian stretch tensor $\mathbf{U}$ associated with a simple shear $\mathbf{y} = \mathbf{F}\mathbf{x}$, $\mathbf{F} = \mathbf{I} + k\mathbf{e}_1 \otimes \mathbf{e}_2$, and then graphically interpreted that deformation viewed as $\mathbf{y} = \mathbf{R}(\mathbf{U}\mathbf{x})$. Carry out a corresponding graphical interpretation of a simple shear deformation represented as $\mathbf{y} = \mathbf{V}(\mathbf{R}\mathbf{x})$.

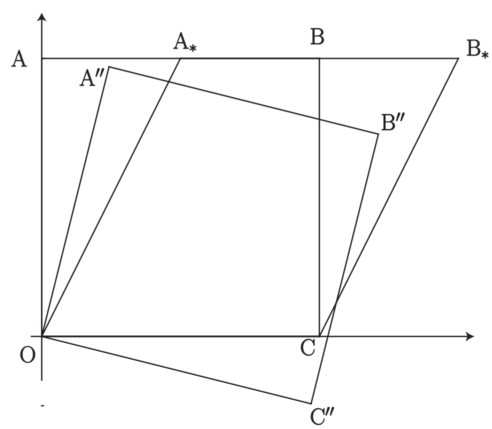


Figure 2.24: Problem 2.36: Simple shear deformation $\mathbf{y} = (\mathbf{I} + k\mathbf{e}_1 \otimes \mathbf{e}_2)\mathbf{x} = \mathbf{V}(\mathbf{R}\mathbf{x})$ viewed in two steps: The rotation $\mathbf{x} \to \mathbf{R}\mathbf{x}$ takes the region $OABC \to OA''B''C''$ and the pure stretch $\mathbf{R}\mathbf{x} \to \mathbf{V}(\mathbf{R}\mathbf{x})$ takes $OA''B''C'' \to OA_*B_*C$.

Remark: The tensor $\mathbf{V}$ can be readily expressed with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by using $\mathbf{V} = \mathbf{F}\mathbf{R}^T$ together with the results of Problem 2.5.2. This leads to

$$\mathbf{V} = \frac{1}{\sqrt{4+k^2}}\Big((2+k^2)\mathbf{e}_1 \otimes \mathbf{e}_1 + k(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + 2\,\mathbf{e}_2 \otimes \mathbf{e}_2\Big) + \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{ii}$$

Problem 2.37. Two material fibers AB and AC in the reference configuration have equal length s_0 and are oriented in the respective directions $\mathbf{e}_1$ and $\mathbf{e}_2$. A homogeneous deformation

$$\mathbf{y} = \mathbf{F}\mathbf{x}, \qquad \mathbf{F} = \ \text{constant},$$

maps these fibers into $A'B'$ and $A'C'$ that have lengths s_1, s_2 with the angle between them being $\pi/2 - \phi$. The quantities s_0, s_1, s_2 and ϕ have been measured.

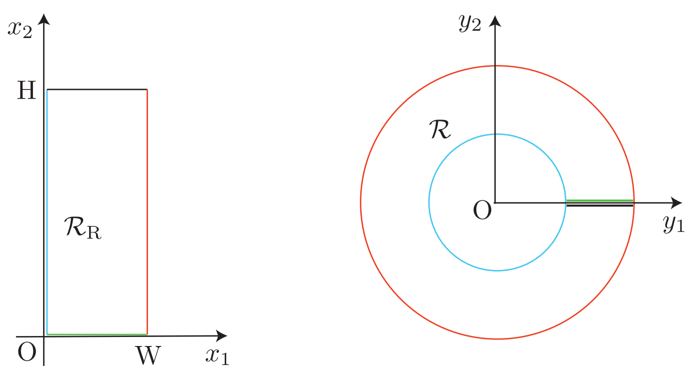
Calculate the strain components E_{11}, E_{22} and E_{12} in terms of s_0, s_1, s_2 and ϕ where $\mathbf{E}$ is the Green Saint-Venant strain tensor. Linearize your answer to the case of an infinitesimal deformation.

Problem 2.38. In Problem 2.5.2 we calculated the stretch tensor $\mathbf{U}$ and rotation tensor $\mathbf{R}$ associated with a simple shear. Linearize those results for small amounts of shear k and thus derive the specializations of $(2.111)_1$ and (2.116) to simple shear. Calculate also the infinitesimal strain tensor $\boldsymbol{\varepsilon}$ and compare and contrast your result with the expression you derived for the Green Saint-Venant strain tensor in Section 2.6.1.

Problem 2.39. This problem involves a planar deformation and for convenience we shall display only the in-plane equations. As shown in Figure 2.25, the body occupies a rectangular strip of width W and height H in a reference configuration. Coordinate axes are chosen such that

$$\mathcal{R}_{\mathrm{R}} = \{(x_1, x_2) :\ 0 \le x_1 \le W,\ 0 \le x_2 \le H\}.$$

The deformation takes the point $(x_1, x_2) \to (y_1, y_2)$ and the region $\mathcal{R}_{\mathrm{R}} \to \mathcal{R}$. Let (r, θ) be the polar



(a) Reference configuration (b) Deformed configuration

Figure 2.25: The body occupying a rectangular strip in the reference configuration, is rolled up into its deformed configuration. The deformation takes $(x_1, x_2) \to (y_1, y_2)$ and $\mathcal{R}_{\mathrm{R}} \to \mathcal{R}$. Despite the figure on the right, in part (a) of this problem do not assume the region $\mathcal{R}$ to be a circular annulus. Figure for Problem 2.39.

coordinates in the deformed configuration,

$$y_1 = r\cos\theta,\ \ y_2 = r\sin\theta, \tag{i}$$

with associated basis vectors

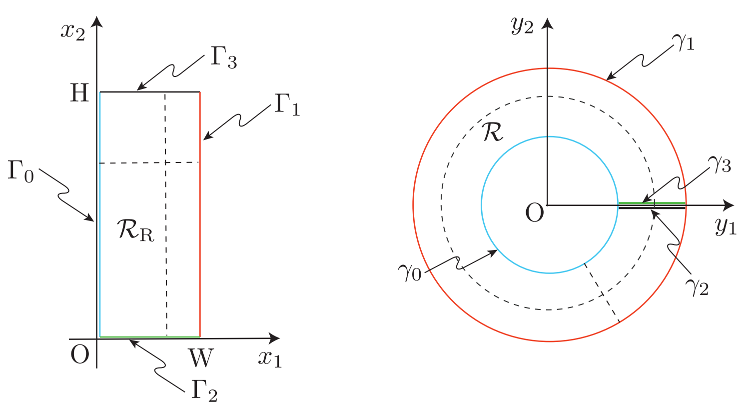
$$\mathbf{e}_r(\theta) = \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2, \qquad \mathbf{e}_\theta(\theta) = -\sin\theta\,\mathbf{e}_1 + \cos\theta\,\mathbf{e}_2. \tag{ii}$$

The deformation can be characterized by

$$r = r(x_1, x_2), \quad \theta = \theta(x_1, x_2), \qquad (x_1, x_2) \in \mathcal{R}_{\mathrm{R}}. \tag{iii}$$

(a) Though the figure shows the region in the deformed configuration to be a circular annulus, in this part of the problem do *not* assume the deformation (*iii*) to possess any form of symmetry. (You would have to consider such non-symmetric deformations if, for example, your goal was to study the stability of the cylindrically symmetric one.) Calculate the deformation gradient tensor $\mathbf{F}$ and the left Cauchy-Green tensor $\mathbf{B}$ using the bases $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta\}$ in the reference and deformed configurations respectively. Derive a condition on $r(x_1, x_2), \theta(x_1, x_2)$ and their partial derivatives if the material is incompressible.

(b) Now consider the special case where the deformation carries each vertical line $x_1 =$ constant in $\mathcal{R}_{\mathrm{R}}$ into a circle $r =$ constant in $\mathcal{R}$, and each horizontal line $x_2 =$ constant in $\mathcal{R}_{\mathrm{R}}$ into a radial line $\theta =$ constant in $\mathcal{R}$. This is illustrated by the dashed curves in Figure 2.26. The left- and right-hand boundaries $x_1 = 0$ and $x_1 = W$ map into circles of radii r_0 and r_1 respectively. The bottom edge of the strip $x_2 = 0$ and the top edge of the strip $x_2 = H$ map into the respective radial lines $\theta = 0$ and $\theta = 2\pi$. What form do the



(a) Reference configuration (b) Deformed configuration

Figure 2.26: The body, occupying a rectangular strip in the reference configuration, is rolled up into a circular annulus in its deformed configuration. As shown by the dashed curves, every vertical line maps into a circle and every horizontal line maps into a radial line. The boundaries $\Gamma_i \to \gamma_i, i = 0, 1, 2, 3$.

functions $r(x_1, x_2)$ and $\theta(x_1, x_2)$ have in this case? Specialize your expressions for $\mathbf{F}$ and $\mathbf{B}$ from part (a) to this case. Moreover, use the incompressibility condition to find $r(x_1, x_2)$ and $\theta(x_1, x_2)$. What are the principal stretches λ_r and λ_θ?

(c) Finally, suppose that the boundary Γ_0 (see Figure 2.26) is not stretched by this deformation. Specialize the deformation and the principal stretches from part (b) to this case.

Problem 2.40. (*Measures of volumetric and shape change.*)

(a) Multiplicatively decompose an arbitrary deformation gradient tensor $\mathbf{F}$ into the product of a tensor $\alpha\mathbf{I}$ that captures the entire volume change associated with $\mathbf{F}$ and a tensor $\overline{\mathbf{F}}$ that involves no volume change, i.e. given $\mathbf{F}$, find α and $\overline{\mathbf{F}}$ such that

$$\mathbf{F} = (\alpha\mathbf{I})\overline{\mathbf{F}} = \alpha\overline{\mathbf{F}} \qquad \text{where} \quad \det\overline{\mathbf{F}} = 1. \tag{i}$$

One speaks of the part $\alpha\mathbf{I}$ as the volumetric part of the deformation gradient tensor $\mathbf{F}$ while the part $\overline{\mathbf{F}}$ is the "shape change" (plus rotation) part of $\mathbf{F}$.

(b) Define the "modified left Cauchy-Green deformation tensor" $\overline{\mathbf{B}}$ and its principal scalar invariants $\overline{I}_1, \overline{I}_2, \overline{I}_3$ by

$$\overline{\mathbf{B}} := \overline{\mathbf{F}}\,\overline{\mathbf{F}}^T, \qquad \overline{I}_1 := \operatorname{tr}\overline{\mathbf{B}}, \qquad \overline{I}_2 := \frac{1}{2}\big[(\operatorname{tr}\overline{\mathbf{B}})^2 - \operatorname{tr}\overline{\mathbf{B}}^2\big], \qquad \overline{I}_3 := \det\overline{\mathbf{B}}. \tag{2.126}$$

Derive expressions for $\mathbf{B}$, I_1, I_2, J in terms of $\overline{\mathbf{B}}$, $\overline{I}_1, \overline{I}_2, \overline{I}_3$ where, as usual

$$\mathbf{B} = \mathbf{F}\,\mathbf{F}^T, \qquad I_1 = \operatorname{tr}\mathbf{B}, \qquad I_2 = \frac{1}{2}\big[(\operatorname{tr}\mathbf{B})^2 - \operatorname{tr}\mathbf{B}^2\big], \qquad J = \sqrt{I_3} = \sqrt{\det\mathbf{B}}, \tag{ii}$$

and show that there is a one-to-one relation between $\{I_1, I_2, J\}$ and $\{\overline{I}_1, \overline{I}_2, J\}$.

(c) Derive linearized expressions for the volumetric and shape change measures

$$\alpha\mathbf{I} - \mathbf{I} \qquad \text{and} \qquad \overline{\mathbf{E}} := \frac{1}{2}\left[\overline{\mathbf{F}}\,\overline{\mathbf{F}}^T - \mathbf{I}\right]$$

when the displacement gradient is small. Express your answers in terms of the infinitesimal strain and rotation tensors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\omega}$.

Problem 2.41. *Rigid deformation.* A deformation $\mathbf{y}(\mathbf{x})$ is said to be rigid if it preserves the distance between all pairs of particles, i.e. if (2.15) holds. Show that a deformation is rigid if and only if it has the form

$$\mathbf{y}(\mathbf{x}) = \mathbf{Q}\mathbf{x} + \boldsymbol{b} \tag{i}$$

where $\mathbf{Q}$ is a constant orthogonal tensor and $\boldsymbol{b}$ is a constant vector.

Problem 2.42. *"Orientation" preserving deformation.* A triplet of vectors $\{d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}, d\mathbf{x}^{(3)}\}$ is right-handed if

$$(d\mathbf{x}^{(1)} \times d\mathbf{x}^{(2)}) \cdot d\mathbf{x}^{(3)} > 0.$$

A deformation is said to preserve orientation if every right-handed linearly-independent triplet of material fibers $\{d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}, d\mathbf{x}^{(3)}\}$ is carried into a right-handed triplet of fibers $\{d\mathbf{y}^{(1)}, d\mathbf{y}^{(2)}, d\mathbf{y}^{(3)}\}$. Show that a deformation $\mathbf{y}(\mathbf{x})$ is orientation preserving if and only if

$$\det\mathbf{F} > 0 \qquad \text{where } \mathbf{F} = \nabla\mathbf{y}. \tag{i}$$

Problem 2.43. *(Change of Area).* Derive *Nanson's formula*, i.e. calculate the relationship between two material area elements $dA_x\mathbf{n}_\mathrm{R}$ and $dA_y\mathbf{n}$ in the reference and deformed configurations respectively; see Figure 2.10.

References:

1. R. Abeyaratne, Chapters 2 and 3 of *Continuum Mechanics*, Volume II in the series *Lecture Notes on The Mechanics of Solids*, http://web.mit.edu/abeyaratne/lecture_notes.html.

2. R. J. Atkin and N. Fox, Chapter 1 of *An Introduction to the Theory of Elasticity*, Longman, 1980.

3. P. Chadwick, Chapter 2 of *Continuum Mechanics: Concise Theory and Problems*, Dover, 1999.

4. A. Goriely, A. Erlich and C. Goodbrake, C5.1 Solid Mechanics: Online problem sheets, https://courses.maths.ox.ac.uk/node/36846/materials, Oxford University, 2018.

5. M.E. Gurtin, E. Fried and L. Anand, Part III of *The Mechanics and Thermodynamics of Continua*, Cambridge University Press, 2010.

6. R. S. Rivlin and D. W. Saunders, Large elastic deformations of isotropic materials. VII Experiments on the deformation of rubber, *Philosophical Transactions of the Royal Society of London. Series A.*, Volume 243, 1951, pp. 34 - 288.

7. R. W. Ogden, Chapter 2 of *Non-linear Elastic Deformations*, Dover, 1997.

8. A.J.M. Spencer, *Continuum Mechanics*, Chapters 4 and 6, Longman, 1980.

2.11 Appendix

2.11.1 The material time derivative.

Next consider a time-dependent **motion** of the body on some time interval $[t_0, t_1]$. The motion takes the particle located at $\mathbf{x}$ in the reference configuration to the location $\widehat{\mathbf{y}}(\mathbf{x}, t)$ at time t:

$$\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}, t), \qquad \mathbf{x} \in \mathcal{R}_{\mathrm{R}}, \quad \mathbf{y} \in \mathcal{R}_t, \quad t \in [t_0, t_1], \tag{2.127}$$

$\mathcal{R}_t$ being the region of space occupied by the body at time t. Note that $\mathcal{R}_t$ evolves with time. Since there is a one-to-one relation between $\mathbf{x}$ and $\mathbf{y}$ at each time, there is an inverse mapping

$$\mathbf{x} = \overline{\mathbf{x}}(\mathbf{y}, t)$$

that takes $\mathcal{R}_t \to \mathcal{R}_{\mathrm{R}}$.

Keep in mind that the location of a particle in the reference configuration serves as a convenient tag by which to identify the particle and so $\mathbf{x}$ serves as a proxy for a particle label. Thus when we want to consider the rate of change of some field at a fixed particle, we consider its rate of change at fixed $\mathbf{x}$.

Now consider a field $\phi(\mathbf{y}, t)$ defined on $\mathcal{R}_t$. Though this represents ϕ spatially, suppose we want to calculate its time rate of change at a fixed particle – the so-called **material time derivative** of ϕ. We shall use a superior dot to denote this time rate of change of ϕ by writing $\dot{\phi}$. In order to calculate $\dot{\phi}$ we first use the motion $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}, t)$ to map ϕ into the reference configuration thus obtaining its material representation $\widehat{\phi}(\mathbf{x}, t)$ where

$$\widehat{\phi}(\mathbf{x}, t) = \phi(\widehat{\mathbf{y}}(\mathbf{x}, t), t). \tag{2.128}$$

Then by $\dot{\phi}$ we mean

$$\dot{\phi} = \frac{\partial \widehat{\phi}}{\partial t}(\mathbf{x}, t).$$

In particular, the velocity of a particle at time t is the time rate of change of position at a fixed particle:

$$\widehat{\mathbf{v}}(\mathbf{x}, t) := \dot{\mathbf{y}} = \frac{\partial \widehat{\mathbf{y}}}{\partial t}(\mathbf{x}, t). \tag{2.129}$$

Following the discussion in Section 2.8, we can express the velocity field spatially in the form $\mathbf{v}(\mathbf{y}, t)$ or referentially in the form $\widehat{\mathbf{v}}(\mathbf{x}, t)$ where these two representations are related by

$$\mathbf{v}(\mathbf{y}, t) = \widehat{\mathbf{v}}(\overline{\mathbf{x}}(\mathbf{y}, t), t), \qquad \widehat{\mathbf{v}}(\mathbf{x}, t) = \mathbf{v}(\widehat{\mathbf{y}}(\mathbf{x}, t), t).$$

Returning to a generic function $\phi(\mathbf{y},t)$, we can calculate its material time derivative by differentiating (2.128) with respect to time (keeping $\mathbf{x}$ fixed) and using the chain rule:

$$\dot{\phi} = \frac{\partial \widehat{\phi}}{\partial t}(\mathbf{x},t) = \frac{\partial \phi}{\partial y_i}(\mathbf{y},t)\frac{\partial \widehat{y}_i}{\partial t}(\mathbf{x},t) + \frac{\partial \phi}{\partial t}(\mathbf{y},t) = \frac{\partial \phi}{\partial y_i}v_i + \frac{\partial \phi}{\partial t}$$

which we can write as

$$\dot{\phi} = \mathbf{v}\cdot \operatorname{grad}\phi + \frac{\partial \phi}{\partial t}, \tag{2.130}$$

where $\operatorname{grad}\phi$ is the vector field with cartesian components $\partial\phi/\partial y_i$.

Exercise: Show that

$$\dot{J} = J\,\mathsf{div}\,\mathbf{v} \tag{2.131}$$

where $J = \det\mathbf{F}$. In cartesian components, $\mathsf{div}\,\mathbf{v} = \partial v_i/\partial y_i$.

Exercise: The *velocity gradient tensor* $\mathbf{L}$ is the tensor with cartesian components $\partial v_i/\partial y_j$:

$$\mathbf{L} := \operatorname{grad}\mathbf{v}, \qquad L_{ij} = \frac{\partial v_i}{\partial y_j}(\mathbf{y},t). \tag{2.132}$$

Show that

$$\dot{\mathbf{F}} = \mathbf{L}\mathbf{F} \qquad \text{and} \qquad \operatorname{div}\mathbf{v} = \operatorname{tr}\mathbf{L}. \tag{2.133}$$

2.11.2 A transport theorem.

In subsequent chapters we will need to calculate the rate of change of energy associated with some part of the body. Suppose that this part occupies a subregion $\mathcal{D}_t \subset \mathcal{R}_t$ at time t, keeping in mind that even though $\mathcal{D}_t$ moves through space, the same material particles are associated with it at all times. We will therefore have to evaluate a term of the form

$$\frac{d}{dt}\int_{\mathcal{D}_t}\phi(\mathbf{y},t)\,dV_y.$$

If $\mathcal{D}_t$ did not depend on t we would simply take the derivative inside the integral but here we must pay attention to the fact that $\mathcal{D}_t$ is time-dependent. In order to get around the time dependency of $\mathcal{D}_t$, we map $\mathcal{D}_t$ into the (time-independent) region $\mathcal{D}_{\mathrm{R}}$ that it occupies in the reference configuration using the motion $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x},t)$. Under this mapping $\mathcal{D}_t \to \mathcal{D}_{\mathrm{R}}$, $\phi(\mathbf{y},t) \to \widehat{\phi}(\mathbf{x},t)$ and $dV_y \to J dV_x$. Accordingly

$$\frac{d}{dt}\int_{\mathcal{D}_t}\phi(\mathbf{y},t)\,dV_y = \frac{d}{dt}\int_{\mathcal{D}_{\mathrm{R}}}\widehat{\phi}(\mathbf{x},t)\,J\,dV_x.$$

We can now take the derivative inside the integral since $\mathcal{D}_{\mathrm{R}}$ is time independent:

$$\begin{aligned}
\frac{d}{dt}\int_{\mathcal{D}_t}\phi(\mathbf{y},t)\,dV_y \quad &= \frac{d}{dt}\int_{\mathcal{D}_{\mathrm{R}}}\widehat{\phi}(\mathbf{x},t)\,J(\mathbf{x},t)\,dV_x = \int_{\mathcal{D}_{\mathrm{R}}}\frac{\partial}{\partial t}\Big(\widehat{\phi}(\mathbf{x},t)\,J(\mathbf{x},t)\Big)\,dV_x = \\
&= \int_{\mathcal{D}_{\mathrm{R}}}(\dot{\phi}J + \phi\dot{J})\,dV_x \overset{(2.131)}{=} \int_{\mathcal{D}_{\mathrm{R}}}(\dot{\phi}J + \phi J\,\mathsf{div}\,\mathbf{v})\,dV_x = \\
&= \int_{\mathcal{D}_{\mathrm{R}}}(\dot{\phi} + \phi\,\mathsf{div}\,\mathbf{v})\,J\,dV_x = \int_{\mathcal{D}_t}(\dot{\phi} + \phi\,\mathsf{div}\,\mathbf{v})\,dV_y,
\end{aligned}$$

where in getting to the very last expression we reverted from $\mathcal{D}_{\mathrm{R}} \to \mathcal{D}_t$ and $J dV_x \to dV_y$. Therefore we have the **transport formula**

$$\frac{d}{dt}\int_{\mathcal{D}_t} \phi\, dV_y = \int_{\mathcal{D}_t} (\dot{\phi} + \phi\, \mathsf{div}\, \mathbf{v})\, dV_y \tag{2.134}$$

for the function $\phi(\mathbf{y}, t)$. Similar transport formulae can be written for vector and tensor fields, as well as for fields defined on a moving surface $\mathcal{S}_t$ or a moving curve $\mathcal{L}_t$; see Volume II.

Finally we note an illuminating alternative form of (2.134). First, we can rewrite (2.130) as

$$\dot{\phi} = \operatorname{grad}\phi \cdot \mathbf{v} + \frac{\partial \phi}{\partial t}(\mathbf{y}, t) = \operatorname{div}(\phi \mathbf{v}) - \phi \operatorname{div}\mathbf{v} + \frac{\partial \phi}{\partial t}(\mathbf{y}, t),$$

where $\operatorname{div}\mathbf{v}$ is the scalar field $\partial v_i/\partial y_i$. Substituting this into (2.134) yields

$$\frac{d}{dt}\int_{\mathcal{D}_t} \phi\, dV_y = \int_{\mathcal{D}_t} \left(\frac{\partial \phi}{\partial t} + \operatorname{div}(\phi \mathbf{v})\right) dV_y. \tag{2.135}$$

Finally we use the divergence theorem (1.171) to rewrite the last term thus obtaining the following alternate form of the transport formula:

$$\frac{d}{dt}\int_{\mathcal{D}_t} \phi\, dV_y = \int_{\mathcal{D}_t} \frac{\partial \phi}{\partial t}\, dV_y + \int_{\partial\mathcal{D}_t} \phi\, \mathbf{v}\cdot\mathbf{n}\, dA_y. \tag{2.136}$$

In this representation, the last term characterizes the flux of ϕ across the boundary $\partial\mathcal{D}_t$.

2.11.3 Exercises.

Problem 2.11.1. Consider the particular (time-dependent) *motion* $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}, t)$:

$$y_1 = a(t)x_1 + b(t)x_2, \qquad y_2 = c(t)x_2, \qquad y_3 = d(t)x_3. \tag{i}$$

Calculate the particle velocity field and express it in both referential (material) form and spatial form.

Calculate the particle acceleration field and express it in spatial form.

Calculate the components of $\operatorname{Grad}\mathbf{v}$, the tensor with cartesian components $\partial v_i(\mathbf{x}, t)/\partial x_j$.

Calculate the components of the velocity gradient tensor $\mathbf{L} = \operatorname{grad}\mathbf{v}$ where $\operatorname{grad}\mathbf{v}$ is the tensor with cartesian components $\partial v_i(\mathbf{y}, t)/\partial y_j$.

Calculate also the *stretching tensor* field $\mathbf{D}(\mathbf{y}, t)$:

$$\mathbf{D} \coloneqq \frac{1}{2}(\mathbf{L} + \mathbf{L}^T). \tag{2.137}$$

Problem 2.11.2. A body undergoes a motion $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}, t)$ and occupies a region $\mathcal{R}_t$ at time t. Calculate the rate of change of the surface area of the outer boundary of the body:

$$\frac{d}{dt}\int_{\partial\mathcal{R}_t} dA_y. \tag{i}$$

Problem 2.11.3. *(A transport theorem.)* A body undergoes a motion $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}, t)$ and occupies a region $\mathcal{R}_t$ at time t. Let $\mathcal{S}_t$ be an evolving *material surface* in the interior of $\mathcal{R}_t$ – by a material surface we mean that the same material particles lie on $\mathcal{S}_t$ at all times even though $\mathcal{S}_t$ moves through space. Let $\mathbf{g}(\mathbf{y}, t)$ be a smooth vector field defined for all $\mathbf{y} \in \mathcal{R}_t$ at each t. Show that

$$\frac{d}{dt}\int_{\mathcal{S}_t} \mathbf{g} \cdot \mathbf{n}\, dA_y = \int_{\mathcal{S}_t} (\dot{\mathbf{g}} + (\operatorname{tr}\mathbf{L})\,\mathbf{g} - \mathbf{L}\mathbf{g}) \cdot \mathbf{n}\, dA_y, \tag{2.138}$$

where $\dot{\mathbf{g}}$ is the material time derivative of $\mathbf{g}$ as defined just above (2.131) and $\mathbf{L} = \operatorname{grad}\mathbf{v}$ is the velocity gradient tensor defined in (2.132).

Chapter 3

Force, Equilibrium Principles and Stress

In this chapter we consider the equilibrium principles of force and moment balance and their consequences. The analysis holds no matter what the constitutive characteristics of the material, provided only that it can be modeled as a continuum. Our focus will be on purely mechanical issues. A more complete discussion (including inertial effects) can be found in the references listed at the end of this chapter

A roadmap of this chapter is as follows: in Section 3.1 we introduce the notion of force, more specifically body force and traction, and discuss their various attributes. The global balance laws for force and moment equilibrium are stated in Section 3.2, and from them we deduce the notion of stress and discuss it in Section 3.3. Section 3.4 is devoted to deriving the field equations associated with the balance laws. Principal stresses and principal directions are discussed in Section 3.5. The analysis and discussion up to this point are carried out entirely using the geometric characteristics of the deformed configuration without any mention of a reference configuration or the deformation. It is often useful however to work with an (equivalent) formulation with respect to a reference configuration. Accordingly in Section 3.7 we reformulate the *geometric aspects* of the preceding analysis to be those associated with a reference configuration and the first Piola Kirchhoff stress tensor is introduced. Section 3.8 considers the rate at which stress does work – the stress power – , and the notion of work-conjugate stress-strain pairs is discussed in Section 3.8.1. The preceding results are linearized in Section 3.9. Finally in Section 3.10 we examine the equilibrium equations in cylindrical and spherical polar coordinates.

All fields encountered in this chapter will be assumed to be smooth. That is, we assume them to be differentiable as many times as needed, and that these derivatives are continuous. This must be relaxed when, for example, we consider a two-phase material where the stress field will be discontinuous at the interfaces between the phases (Problem 3.24).

3.1 Force.

We are concerned with the *deformed configuration* of the body. In this configuration the body occupies a region $\mathcal{R}$, and an arbitrary part of the body[1] occupies a region $\mathcal{D}$ that is a subregion of $\mathcal{R}$. It is convenient to refer to $\mathcal{D}$ as a part of the body (rather than to use the more cumbersome but precise phrase "the region occupied by a part of the body"). In this configuration, a generic particle is located at $\mathbf{y} \in \mathcal{R}$. Inertial effects are not considered and when we refer to time t, we only use it to discuss a one-parameter family of configurations.

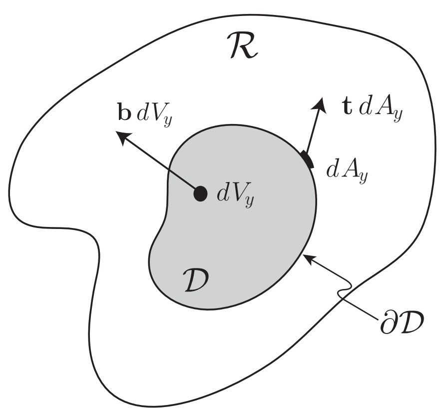


Figure 3.1: Forces acting on a part $\mathcal{D}$ of the body: the traction $\mathbf{t}$ is a force per unit area acting at points on the boundary $\partial\mathcal{D}$ due to contact between $\mathcal{D}$ and the rest of the body across the surface $\partial\mathcal{D}$. The body force density $\mathbf{b}$ is a force per unit volume acting at points in the interior of $\mathcal{D}$ applied by agents outside the body.

We now turn our attention to the forces acting on an arbitrary part $\mathcal{D}$ of the deformed body. As depicted in Figure 3.1 we assume there are two types of forces: *body forces* that act at each point in the interior of $\mathcal{D}$ and are applied by agents outside of the body, and *contact forces* or **tractions** that act at points on the boundary $\partial\mathcal{D}$ of $\mathcal{D}$ and represent forces due to

[1]A *part* of a body involves the same set of particles in all configurations. For a more careful discussion, see Volume II.

contact between $\mathcal{D}$ and the rest of the body[2] across the surface $\partial\mathcal{D}$. The body force density $\mathbf{b}$ is a force *per unit (deformed) volume*[3], while the contact force density $\mathbf{t}$ is a force *per unit (deformed) surface area*; see Figure 3.1.

In order to characterize a force, we must specify how it contributes to (a) the resultant force, (b) the resultant moment about an arbitrary fixed point, and (c) how it does work.

Since $\mathbf{b}$ is a force per unit volume distributed over $\mathcal{D}$, its resultant is its volume integral over $\mathcal{D}$. Similarly since $\mathbf{t}$ is a force per unit area distributed over the boundary $\partial\mathcal{D}$, its resultant is its surface integral over $\partial\mathcal{D}$. The *resultant external force* on the part $\mathcal{D}$ under consideration is then taken to be

$$\int_{\mathcal{D}} \mathbf{b}\,\mathrm{d}V_y + \int_{\partial\mathcal{D}} \mathbf{t}\,\mathrm{d}A_y\,; \tag{3.1}$$

the *resultant moment* of the external forces acting on $\mathcal{D}$ about an arbitrary fixed point O is taken to be

$$\int_{\mathcal{D}} \mathbf{y}\times\mathbf{b}\,\mathrm{d}V_y + \int_{\partial\mathcal{D}} \mathbf{y},\times\mathbf{t}\,\mathrm{d}A_y \tag{3.2}$$

where $\mathbf{y}$ is position with respect to O; and the *rate of working* of the external forces acting on $\mathcal{D}$ is taken to be

$$\int_{\mathcal{D}} \mathbf{b}\cdot\mathbf{v}\,\mathrm{d}V_y + \int_{\partial\mathcal{D}} \mathbf{t}\cdot\mathbf{v}\,\mathrm{d}A_y\,, \tag{3.3}$$

where $\mathbf{v}$ is particle velocity. Note that $\mathbf{t}$ represents a force per unit *deformed* area and $\mathbf{b}$ a force per unit *deformed* volume.

In order for the formulae (3.1) - (3.3) to be useful, we must specify the variables on which $\mathbf{b}$ and $\mathbf{t}$ depend. We expect that the body force density may depend on position $\mathbf{y}$ and so we assume that

$$\mathbf{b} = \mathbf{b}(\mathbf{y}). \tag{3.4}$$

We now turn to the traction $\mathbf{t}$. It too will depend on the position $\mathbf{y}$ but it cannot depend *only* on $\mathbf{y}$. To see this consider Figure 3.2. The two figures there both show the same region $\mathcal{D}$, and the point A in both is the same. Its position vector is $\mathbf{y}_A$. In the left-hand figure, $\mathcal{D}_1$ and $\mathcal{D}_2$ are two parts of the body and A lies on the interface between them. The right-hand figure shows two different parts of the body, $\mathcal{D}_3$ and $\mathcal{D}_4$, and A lies on the interface between these two parts as well. The interface between $\mathcal{D}_3$ and $\mathcal{D}_4$ is different to that between $\mathcal{D}_1$

[2]If part of $\partial\mathcal{D}$ coincides with a part of $\partial\mathcal{R}$, the contact force on that part of the surface is applied by an outside agent.

[3]One can alternatively characterize the body force as a force per unit mass.

and $\mathcal{D}_2$ though A lies on both interfaces. If the traction $\mathbf{t}$ depended *only* on $\mathbf{y}$, then the traction at A would be $\mathbf{t}(\mathbf{y}_A)$ and the force per unit area applied by[4] $\mathcal{D}_1$ on $\mathcal{D}_2$ at A and the force per unit area applied by $\mathcal{D}_3$ on $\mathcal{D}_4$ at A would both be $\mathbf{t}(\mathbf{y}_A)$. However we do not expect the force applied by $\mathcal{D}_1$ on $\mathcal{D}_2$ to be the same as that applied by $\mathcal{D}_3$ on $\mathcal{D}_4$.

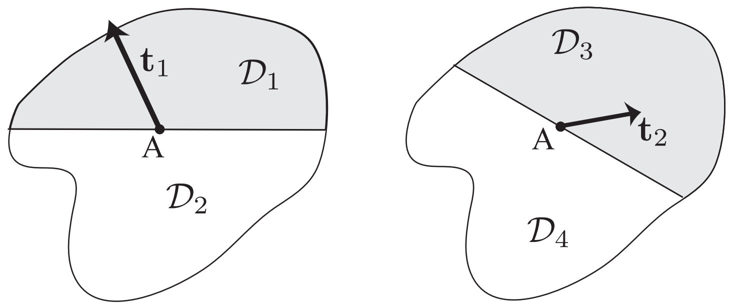


Figure 3.2: Traction depends on the surface on which it acts: Regions $\mathcal{D}_1$ and $\mathcal{D}_2$ are occupied by two parts of a body, while $\mathcal{D}_3$ and $\mathcal{D}_4$ are occupied by a different pair of parts. Planar interfaces separate these parts while the point A is common to both interfaces. The traction $\mathbf{t}_1$ in the figure on the left is applied at A by the material in $\mathcal{D}_1$ on that in $\mathcal{D}_2$. The traction $\mathbf{t}_2$ in the figure on the right is applied at A by the material in $\mathcal{D}_3$ on that in $\mathcal{D}_4$. Even though both tractions are associated with the same point $\mathbf{y}_A$ there is no reason to expect that $\mathbf{t}_1 = \mathbf{t}_2$.

Therefore the traction must depend on the specific surface at $\mathbf{y}$ on which it acts. To first order, a surface is defined by its unit normal vector $\mathbf{n}$, and so we shall assume that

$$\mathbf{t} = \mathbf{t}(\mathbf{y}, \mathbf{n}). \tag{3.5}$$

According to this, the dependence of the traction on the surface is only through the normal vector and not, for example, the curvature of the surface. The ansatz (3.5) is known as *Cauchy's hypothesis*.

It is worth emphasizing that according to (3.3) the traction $\mathbf{t}(\mathbf{y}, \mathbf{n})$ denotes the force per unit area *applied by* the part outside $\mathcal{D}$ *on* the material inside $\mathcal{D}$. Now consider a (not-necessarily closed) surface $\mathcal{S}$ in the body and let $\mathbf{y}$ be a point on this surface and let $\mathbf{n}$ be a unit vector that is normal to $\mathcal{S}$ as shown in Figure 3.3. The side of $\mathcal{S}$ into which $\mathbf{n}$ points is referred to as the *positive side* of $\mathcal{S}$ and the other is the *negative side*. By convention, the traction vector $\mathbf{t}(\mathbf{y}, \mathbf{n})$ denotes the force per unit area applied *by the material on the positive side on the material on the negative side*.

[4]Question: Is this the force applied by $\mathcal{D}_2$ on $\mathcal{D}_1$ or by $\mathcal{D}_1$ on $\mathcal{D}_2$?

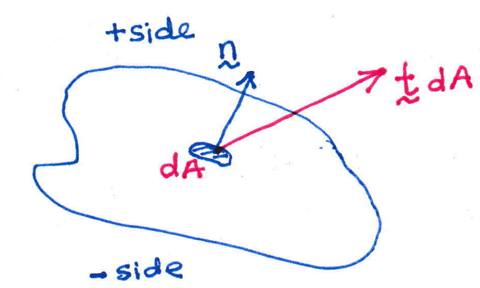


Figure 3.3: The force vector $\mathbf{t}(\mathbf{n})\, dA_y$ acting on an infinitesimal surface element of surface $\mathcal{S}$ at a point $\mathbf{y}$. This force is applied by the material on the positive side on the material on the negative side. The positive side is the one into which the unit normal vector $\mathbf{n}$ points.

Is this consistent with our earlier discussion of the traction on the closed surface $\partial\mathcal{D}$? If the unit normal vector $\mathbf{n}$ on $\partial\mathcal{D}$ is taken so *it points out* of $\mathcal{D}$, then the positive side of the surface is the outside of $\mathcal{D}$ and so $\mathbf{t}(\mathbf{y}, \mathbf{n})$ is the traction applied by the part outside of $\mathcal{D}$ on $\mathcal{D}$. This is exactly what we had earlier, the point being that the unit normal vector should be pointing outwards.

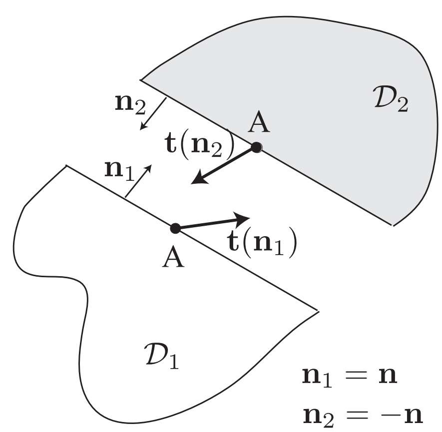


Figure 3.4: The unit *outward* normal vector to $\mathcal{D}_1$ is $\mathbf{n}_1$ and so the traction that is applied by $\mathcal{D}_2$ on $\mathcal{D}_1$ is $\mathbf{t}(\mathbf{n}_1)$. The unit *outward* normal vector to $\mathcal{D}_2$ is $\mathbf{n}_2$ and so the traction applied by $\mathcal{D}_1$ on $\mathcal{D}_2$ is $\mathbf{t}(\mathbf{n}_2)$. While $\mathbf{n}_1 = -\mathbf{n}_2$ we do not know (yet) whether $\mathbf{t}(\mathbf{n}_1) = -\mathbf{t}(\mathbf{n}_2)$.

Continuing to focus on the dependence of the traction on the outward unit normal vector, consider the body shown in Figure 3.4. In order to calculate the traction acting *on* $\mathcal{D}_1$ at A, we draw the unit normal vector to the interface that points outward from $\mathcal{D}_1$. This is denoted by $\mathbf{n}_1$ in the figure. Thus the traction acting on $\mathcal{D}_1$ is $\mathbf{t}(\mathbf{y}_A, \mathbf{n}_1)$; this is applied by $\mathcal{D}_2$. On the other hand if we want to calculate the traction acting *on* $\mathcal{D}_2$ at A, we draw the unit normal vector $\mathbf{n}_2$ that points outward from $\mathcal{D}_2$. Thus the traction acting *on* $\mathcal{D}_2$ is $\mathbf{t}(\mathbf{y}_A, \mathbf{n}_2)$; this is applied by $\mathcal{D}_1$. We do not (yet) know how $\mathbf{t}(\mathbf{y}_A, \mathbf{n}_1)$ relates to $\mathbf{t}(\mathbf{y}_A, \mathbf{n}_2)$ though $\mathbf{n}_1 = -\mathbf{n}_2$, i.e. how $\mathbf{t}(\mathbf{y}, \mathbf{n})$ relates to $\mathbf{t}(\mathbf{y}, -\mathbf{n})$.

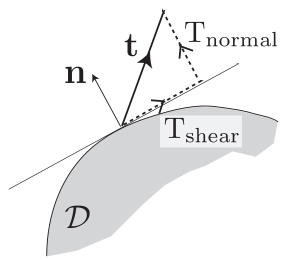


Figure 3.5: Components of the traction $\mathbf{t}$: normal stress T_{normal} and resultant shear stress T_{shear} .

Finally we note that the traction acts in a direction that *need not be normal to the surface*, it being determined by the internal forces within the body, i.e., as depicted in Figure 3.5, $\mathbf{t}(\mathbf{n})$ need not be in the direction $\mathbf{n}$. The component of traction that is normal to the surface is called the *normal stress* and we denote it by T_{normal}:

$$T_{\text{normal}}(\mathbf{n}) = \mathbf{t}(\mathbf{n}) \cdot \mathbf{n}\,. \tag{3.6}$$

When $T_{\text{normal}} > 0$ we say it is *tensile*, *compressive* when $T_{\text{normal}} < 0$. The tangential component of $\mathbf{t}$ is called the *resultant shear stress* and we denote its magnitude by T_{shear}. By the Pythagorean theorem,

$$T_{\text{shear}}(\mathbf{n}) = \sqrt{|\mathbf{t}|^2 - T_{\text{normal}}^2} = \sqrt{\big[\mathbf{t}(\mathbf{n}) \cdot \mathbf{t}(\mathbf{n})\big] - \big[\mathbf{t}(\mathbf{n}) \cdot \mathbf{n}\big]^2}. \tag{3.7}$$

In the two preceding equations we have suppressed the dependency on $\mathbf{y}$.

A natural (and important) question to ask is: "from among all planes through a given point, on which is $T_{\text{normal}} = \mathbf{t}(\mathbf{n}) \cdot \mathbf{n}$ largest? And on which is it smallest?" This requires one to consider $T_{\text{normal}}(\mathbf{n})$ as a function of the unit vector $\mathbf{n}$ and to find the specific vector(s) $\mathbf{n}$ at which it has its extrema. One can ask a similar question for the shear stress $T_{\text{shear}}(\mathbf{n})$. We shall revisit these questions once we have more information on how $\mathbf{t}(\mathbf{n})$ depends on $\mathbf{n}$.

3.2 Force and moment equilibrium.

The *equilibrium principle for force balance* postulates that the resultant force on every part of the body vanishes:

$$\int_{\mathcal{D}} \mathbf{b}\, \mathrm{d}V_y + \int_{\partial\mathcal{D}} \mathbf{t}\, \mathrm{d}A_y = \mathbf{o} \qquad \text{for all } \mathcal{D} \subset \mathcal{R}. \tag{3.8}$$

Similarly, the *equilibrium principle of moment balance* postulates that the resultant moment (about a fixed point O) on every part of the body vanishes:

$$\int_{\mathcal{D}} \mathbf{y} \times \mathbf{b}\, \mathrm{d}V_y + \int_{\partial\mathcal{D}} \mathbf{y} \times \mathbf{t}\, \mathrm{d}A_y = \mathbf{o} \qquad \text{for all } \mathcal{D} \subset \mathcal{R}. \tag{3.9}$$

Both (3.8) and (3.9) must hold for *every part* of the body.

An equation that holds at each point $\mathbf{y}$ is said to be "local" while one that holds for each part $\mathcal{D}$ is said to be "global". Global statements such as (3.8) and (3.9) are convenient when formulating the basic balance principles. When solving a specific boundary-value problem however it is more convenient to have a local version of that principle. The local statement corresponding to a balance law is said to be the associated **field equation**.

From the discussion in Section 3.1 we know that the integrand of the surface integral term in (3.8) depends on the unit normal vector $\mathbf{n}$. If this dependence is linear, and we do not yet know if this is true, then the integrand would have the form $\mathsf{A}\mathbf{n}$ where A is some 2-tensor. In this event we can use the divergence theorem to rewrite the surface integral as a volume integral, and the equation would have the form of a single volume integral over $\mathcal{D}$ that is to vanish. Since this balance law is to hold for all parts $\mathcal{D}$ of the body, then provided the integrand is continuous, we conclude by localization (Section 1.8.3) that the integrand itself must vanish at each point $\mathbf{y} \in \mathcal{R}$. This leads to the field equation associated with (3.8). We could simplify (3.9) similarly. This is what we shall carry out in Section 3.4 below, but before we do that we must show that the traction $\mathbf{t}(\mathbf{y}, \mathbf{n})$ depends linearly on $\mathbf{n}$.

3.3 Consequences of force balance. Stress.

We now explore several implications of force balance. The focus in this section is on how the traction vector $\mathbf{t}(\mathbf{y}, \mathbf{n})$ depends on the unit vector $\mathbf{n}$. The position $\mathbf{y}$ will play no central role in our discussion and so it will be convenient to suppress $\mathbf{y}$ and write $\mathbf{t}(\mathbf{n})$ instead of $\mathbf{t}(\mathbf{y}, \mathbf{n})$.

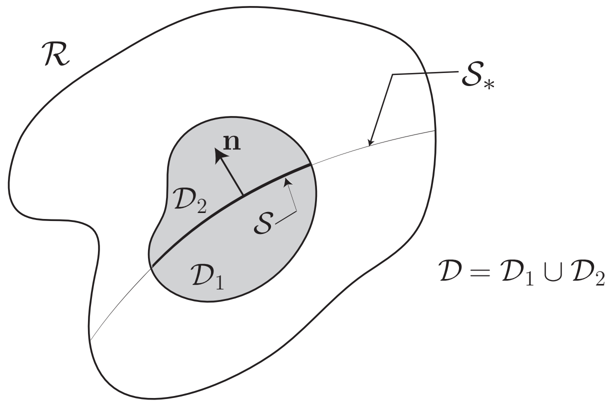


Figure 3.6: A surface $\mathcal{S}_*$ contained within $\mathcal{R}$ intersects the sub-region $\mathcal{D}$ and separates it into two parts $\mathcal{D}_1$ and $\mathcal{D}_2$.

Consequence (1): The "equal and opposite" property $\mathbf{t}(\mathbf{n}) = -\mathbf{t}(-\mathbf{n})$.

Let $\mathcal{S}_*$ be an arbitrary surface contained within the region $\mathcal{R}$ occupied by the deformed body as shown in Figure 3.6. Pick a sub-region $\mathcal{D}$ that is intersected by $\mathcal{S}_*$ and is thus separated into regions $\mathcal{D}_1$ and $\mathcal{D}_2$: $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$. $\mathcal{S}$ is the portion of $\mathcal{S}_*$ that is contained within $\mathcal{D}$ and is therefore the interface between $\mathcal{D}_1$ and $\mathcal{D}_2$. Note that the unit normal vector $\mathbf{n}$ on $\mathcal{S}$ shown in the figure is outward to $\mathcal{D}_1$ whereas $-\mathbf{n}$ is outward to $\mathcal{D}_2$. Thus when force balance (3.8) is applied to $\mathcal{D}_1$, the traction term will involve the integral of $\mathbf{t}(\mathbf{n})$ over $\mathcal{S}$, whereas when it is applied to $\mathcal{D}_2$, it will involve the integral of $\mathbf{t}(-\mathbf{n})$ over $\mathcal{S}$. We now apply (3.8) to each of the regions $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}$ individually, and then subtract the first two of the resulting equations from the third. This leads to (Exercise)

$$\int_{\mathcal{S}} [\mathbf{t}(\mathbf{n}) + \mathbf{t}(-\mathbf{n})] \; dA_y \;=\; 0. \tag{3.10}$$

Since this must hold for arbitrary choices of $\mathcal{D}$, and therefore for arbitrary choices of $\mathcal{S}$, it follows by localization[5] that the integrand must vanish at each point on $\mathcal{S}$. Thus we conclude that

$$\mathbf{t}(-\mathbf{n}) = -\mathbf{t}(\mathbf{n}) \tag{3.11}$$

for all unit vectors $\mathbf{n}$.

Observe that this is the analog for a continuum of Newton's third law for particles. It says that the traction exerted on the positive side of a surface by the negative side, is equal

[5]See Section 1.8.3 for the volume integral version of localization.

in magnitude and opposite in direction to the traction exerted on the negative side by the positive side. While this appears to be a consequence of force balance and not a separate postulate, it is in fact implicitly buried within the assumption that the force on $\mathcal{D}$ is given by (3.1).

Consequence (2): The traction $\mathbf{t}(\mathbf{n})$ is a linear function of $\mathbf{n}$.

We now derive an expression for the traction on a plane oriented in an arbitrary direction $\mathbf{n}$ in terms of the tractions on three mutually orthogonal planes, e.g. planes normal to the basis vectors $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$. This leads to a second, critically important, consequence of force balance, namely that the traction vector $\mathbf{t}(\mathbf{n})$ depends *linearly* on the normal vector $\mathbf{n}$. This is called Cauchy's Theorem.

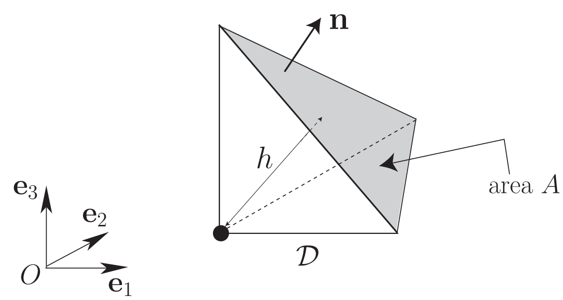


Figure 3.7: Tetrahedral subregion $\mathcal{D}$ of the body.

In order to establish this, consider the tetrahedral subregion $\mathcal{D}$ shown in Figure 3.7 with three of its faces parallel to the coordinate planes. Observe that the unit outward normal vectors to the four faces of $\mathcal{D}$ are $\mathbf{n}$, $-\mathbf{e}_1$, $-\mathbf{e}_2$ and $-\mathbf{e}_3$. Moreover, if the area of the shaded face is A, one can readily show from geometry that the area, A_k, of the face normal to $\mathbf{e}_k$ is $n_k A$. Next we apply force balance to this tetrahedron, and take the limit of the resulting equation as the height $h \to 0$ keeping the orientations of all faces fixed. In this limit the volumetric term (which involves the body force) approaches zero like h^3 whereas the area terms (which involve the traction) approach zero like h^2. Therefore the volumetric term goes to zero faster than the area terms and so only the area terms survive in this limit leading to

$$\lim_{h\to 0} \; A\mathbf{t}(\mathbf{n}) + A_1\mathbf{t}(-\mathbf{e}_1) + A_2\mathbf{t}(-\mathbf{e}_2) + A_3\mathbf{t}(-\mathbf{e}_3) = \mathbf{0}. \tag{3.12}$$

Because of (3.11) and $A_k = n_k A$, this leads to

$$\mathbf{t}(\mathbf{n}) = n_1\mathbf{t}(\mathbf{e}_1) + n_2\mathbf{t}(\mathbf{e}_2) + n_3\mathbf{t}(\mathbf{e}_3) = \mathbf{t}(\mathbf{e}_k)n_k. \tag{3.13}$$

Equation (3.13) tells us that *if we know the tractions on three mutually orthogonal planes*, for example the planes normal to the basis vectors $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$, *that information alone* *is enough to calculate the traction on an arbitrary plane* (whose normal is $\mathbf{n}$). Observe by writing (3.13) as

$$\mathbf{t}(n_1\mathbf{e}_1+n_2\mathbf{e}_2+n_3\mathbf{e}_3) = n_1\mathbf{t}(\mathbf{e}_1)+n_2\mathbf{t}(\mathbf{e}_2)+n_3\mathbf{t}(\mathbf{e}_3) \qquad \text{for all } n_1, n_2, n_3 \text{ with } n_1^2+n_2^2+n_3^2 = 1,$$

that $\mathbf{t}(\mathbf{n})$ is a linear function of $\mathbf{n}$ on the set of all unit vectors.

Consequence (3): The stress tensor $\mathbf{T}$.

As observed above, in order to calculate the traction on an arbitrary plane, we only need know the tractions $\mathbf{t}(\mathbf{e}_1), \mathbf{t}(\mathbf{e}_2), \mathbf{t}(\mathbf{e}_3)$ on the three coordinate planes. It is natural therefore to "label" the components of these three traction vectors. Since each traction vector has three components, we have a total of nine components to label.

Accordingly let $T_{ij}, i, j = 1, 2, 3$, be the *ith* component of the traction $\mathbf{t}(\mathbf{e}_j)$:

$$T_{ij} \coloneqq t_i(\mathbf{e}_j) = \mathbf{t}(\mathbf{e}_j) \cdot \mathbf{e}_i.$$

This is illustrated in Figure 3.8(a). Note that the second subscript of T_{ij} identifies the surface on which the traction acts and the first identifies the direction of that traction component. Thus each T_{ij} represents a force per unit deformed area acting on a particular coordinate plane in a particular direction. An equivalent way in which to write the preceding equation is $\mathbf{t}(\mathbf{e}_j) = T_{ij}\mathbf{e}_i$. Thus we have

$$\boxed{T_{ij} \coloneqq t_i(\mathbf{e}_j) = \mathbf{t}(\mathbf{e}_j) \cdot \mathbf{e}_i, \qquad\qquad \mathbf{t}(\mathbf{e}_j) = T_{ij}\mathbf{e}_i.} \tag{3.14}$$

In order to determine the traction components on a face whose outward normal is in the *negative* j^{th}-direction we observe from (3.11), (3.14) that

$$\mathbf{t}(-\mathbf{e}_j) = -\mathbf{t}(\mathbf{e}_j) = -T_{ij}\mathbf{e}_i = T_{ij}(-\mathbf{e}_i). \tag{3.15}$$

Therefore the force/area acting on a surface with unit normal $-\mathbf{e}_j$, in the direction $-\mathbf{e}_i$, is T_{ij}. This is illustrated in Figure 3.8(b).

The 9 elements T_{ij} may be assembled into a matrix $[T]$. The elements T_{11}, T_{22} and T_{33} on the diagonal of $[T]$ are known as the **normal stress** components; the off-diagonal terms $T_{ij}, i \neq j$, are the **shear stress** components.

We now return to the expression (3.13) for the traction on an arbitrary surface and substitute (3.14) and $n_j = \mathbf{e}_j \cdot \mathbf{n}$ into it:

$$\mathbf{t}(\mathbf{n}) \overset{(3.13)}{=} \mathbf{t}(\mathbf{e}_j)n_j \overset{(3.14)}{=} (T_{ij}\mathbf{e}_i)n_j = T_{ij}n_j\mathbf{e}_i = T_{ij}(\mathbf{e}_j \cdot \mathbf{n})\mathbf{e}_i = T_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)\mathbf{n} = \Big(T_{ij}\,\mathbf{e}_i \otimes \mathbf{e}_j\Big)\mathbf{n}\,. \tag{3.16}$$

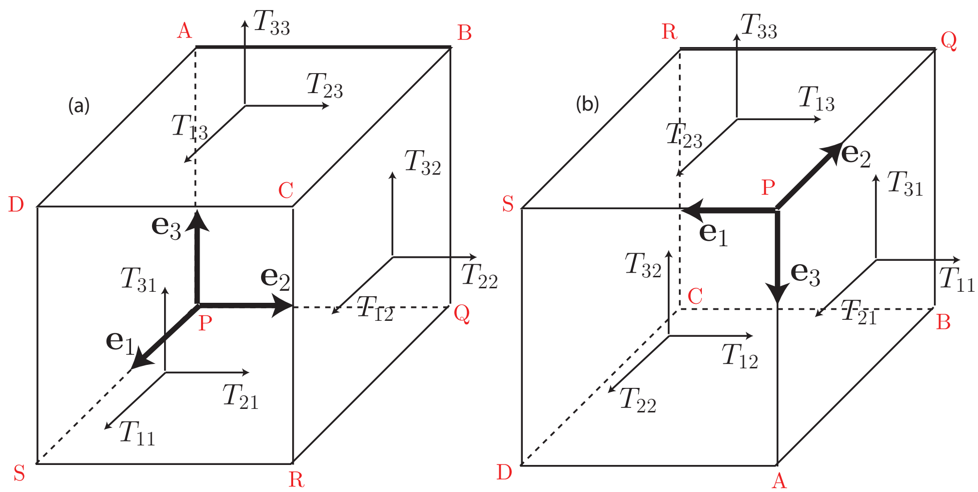


Figure 3.8: The figure shows two views of the same cubic region and the same basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Figure (a) shows the stress components T_{ij} acting on the three faces with outward normals $+\mathbf{e}_1$, $+\mathbf{e}_2$ and $+\mathbf{e}_3$. Observe how the figure is consistent with $T_{ij} = \mathbf{t}(\mathbf{e}_j) \cdot (\mathbf{e}_i)$. Figure (b) shows the stress components T_{ij} acting on the three faces with outward normals $-\mathbf{e}_1$, $-\mathbf{e}_2$ and $-\mathbf{e}_3$. Note in this case the consistency with $T_{ij} = \mathbf{t}(-\mathbf{e}_j) \cdot (-\mathbf{e}_i)$.

Let $\mathbf{T}$ be the second-order tensor whose components in the basis $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3\}$ are T_{ij}, i.e[6].

$$\boxed{\mathbf{T} = T_{ij}\mathbf{e}_i \otimes \mathbf{e}_j.} \tag{3.17}$$

It now follows that (3.16) can be written as $\mathbf{t}(\mathbf{n}) = \mathbf{T}\mathbf{n}$, or by writing out all the arguments:

$$\boxed{\mathbf{t}(\mathbf{y}, \mathbf{n}) = \mathbf{T}(\mathbf{y})\mathbf{n}.} \tag{3.18}$$

In terms of components,

$$t_i(\mathbf{n}) = T_{ij}n_j, \qquad\qquad \{t\} = [T]\{n\}. \tag{3.19}$$

The tensor $\mathbf{T}(\mathbf{y})$ is known as the **Cauchy stress tensor**. Observe that $\mathbf{T}$ *does not depend on* the normal vector $\mathbf{n}$. Therefore we may speak of the stress *at a point*. In contrast, when speaking of traction, we must speak of the traction *on a surface through a point*. When $\mathbf{T}(\mathbf{y})$ is known, equation (3.18) can be used to calculate the traction $\mathbf{t}(\mathbf{y}, \mathbf{n})$ on *any* plane through $\mathbf{y}$. The equilibrium principle of moment balance will show that $\mathbf{T}$ is symmetric.

As noted earlier, the component T_{ij} of the stress tensor represents the i^{th} component of the force per unit area acting on a surface whose normal is in the j^{th} direction. It is

[6]See Section 1.4.3, in particular (1.197).

worth reiterating that we have been concerned with the region occupied by the *deformed* body and therefore (a) the surface referenced above must be normal to the j^{th} direction in the *deformed* configuration, and (b) the area referenced above refers to area in the *deformed* configuration. The middle figure in Figure 3.16 illustrates this in a special case.

The normal stress and the magnitude of the resultant shear stress introduced in (3.6) and (3.7) can now be written in the following respective forms with the dependence on $\mathbf{n}$ made explicit:

$$T_{\text{normal}}(\mathbf{n}) = \mathbf{Tn} \cdot \mathbf{n}\,, \tag{3.20}$$

$$T_{\text{shear}}(\mathbf{n}) = \sqrt{\big(\mathbf{Tn} \cdot \mathbf{Tn}\big) - \big(\mathbf{Tn} \cdot \mathbf{n}\big)^2}. \tag{3.21}$$

In Section 3.5 we shall discuss the maximum values of these two quantities (over all unit vectors $\mathbf{n}$).

In cylindrical polar coordinates (for example) we have

$$\begin{aligned}\mathbf{T} \;\; &= T_{rr}\mathbf{e}_r \otimes \mathbf{e}_r + T_{r\theta}\mathbf{e}_r \otimes \mathbf{e}_\theta + T_{rz}\mathbf{e}_r \otimes \mathbf{e}_z + \\ &\qquad + T_{\theta r}\,\mathbf{e}_\theta \otimes \mathbf{e}_r + T_{\theta\theta}\,\mathbf{e}_\theta \otimes \mathbf{e}_\theta + T_{\theta z}\,\mathbf{e}_\theta \otimes \mathbf{e}_z + \\ &\qquad + T_{zr}\,\mathbf{e}_z \otimes \mathbf{e}_r + T_{z\theta}\,\mathbf{e}_z \otimes \mathbf{e}_\theta + T_{zz}\,\mathbf{e}_z \otimes \mathbf{e}_z.\end{aligned} \tag{3.22}$$

Consider, for example, a surface with outward unit normal vector $\mathbf{e}_r$. The traction on this surface is

$$\mathbf{t}(\mathbf{e}_r) = \mathbf{T}\mathbf{e}_r = T_{rr}\mathbf{e}_r + T_{\theta r}\,\mathbf{e}_\theta + T_{zr}\,\mathbf{e}_z,$$

and we see (again) that the first subscript tells us the direction of a traction component and the second indicates the surface on which it acts.

3.3.1 Some particular stress tensors.

Consider the stress tensor $\mathbf{T}(\mathbf{y})$ at a particular point $\mathbf{y}$. Let T_{ij} be the components of $\mathbf{T}$ in an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

– *Uniaxial stress.* The particular case where the only nonzero component of stress is $T_{11} = T$, i.e.

$$\mathbf{T} = \mathrm{T}\,\mathbf{e}_1 \otimes \mathbf{e}_1, \qquad [T] = \begin{pmatrix} T & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

describes a uniaxial stress in the $\mathbf{e}_1$-direction. Observe that the traction on an arbitrary plane is $\mathbf{t}(\mathbf{n}) = \mathbf{T}\mathbf{n} = T(\mathbf{e}_1 \otimes \mathbf{e}_1)\mathbf{n} = T\,(\mathbf{n}\cdot\mathbf{e}_1)\,\mathbf{e}_1 = Tn_1\,\mathbf{e}_1$. Thus the traction on *every* plane acts in the $\mathbf{e}_1$-direction, though its value depends on the plane (through n_1). A uniaxial stress of magnitude T in some direction $\mathbf{m}$ is described by

$$\mathbf{T} = \mathrm{T}\,\mathbf{m} \otimes \mathbf{m}.$$

– *Hydrostatic stress.* The special case where $\mathbf{T}$ has the form

$$\mathbf{T} = \mathrm{T}\,\mathbf{I}, \qquad [T] = \begin{pmatrix} T & 0 & 0 \\ 0 & T & 0 \\ 0 & 0 & T \end{pmatrix},$$

describes a hydrostatic stress (a pure pressure $-T$). Observe that the traction on an arbitrary plane is $\mathbf{t}(\mathbf{n}) = \mathbf{T}\mathbf{n} = T\,\mathbf{n}$. Thus the traction on any plane acts in the direction normal to that plane and has magnitude T.

– *Pure shear.* Finally,

$$\mathbf{T} = \tau(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1), \qquad [T] = \begin{pmatrix} 0 & \tau & 0 \\ \tau & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

describes a *pure shear* stress state with respect to the $\mathbf{e}_1$, $\mathbf{e}_2$ directions. A pure shear with respect to an arbitrary pair of orthogonal directions $\mathbf{a}$ and $\mathbf{b}$ is described by

$$\mathbf{T} = \tau(\mathbf{a} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{a}).$$

3.3.2 Worked examples.

Problem 3.3.1. The surface $\mathcal{S}$ in Figure 3.9 with unit outward normal vector $\mathbf{e}_2$ is traction-free. Therefore

$$\mathbf{t}(\mathbf{e}_2) = \mathbf{o} \quad \overset{(3.14)_1}{\Rightarrow} \quad T_{k2} = 0 \quad \Rightarrow \quad T_{12} = T_{22} = T_{32} = 0 \qquad \text{on } \mathcal{S}.$$

Therefore the three components of stress T_{12}, T_{22}, T_{32} must vanish on $\mathcal{S}$. Note that *it is not necessary* that the remaining stress components vanish. The surface $\mathcal{S}$ is traction-free but that does not mean a point on the surface has to be stress-free. That is, $\mathbf{t}(\mathbf{n}) = \mathbf{o}$ for some $\mathbf{n}$ does not imply $\mathbf{T} = \mathbf{0}$.

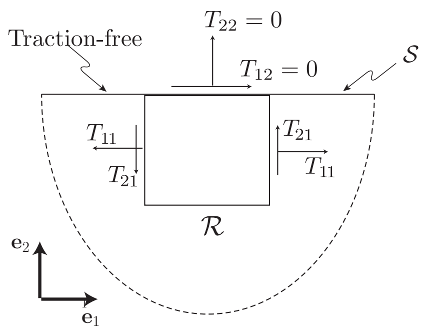


Figure 3.9: The surface $\mathcal{S}$, perpendicular to $\mathbf{e}_2$, is traction-free, and so the stress components $T_{12} = T_{22} = T_{32} = 0$ on $\mathcal{S}$. However the stress components T_{11}, T_{33} and T_{13} need not vanish on $\mathcal{S}$: the surface is traction-free but a point on the surface might not be stress-free.

Problem 3.3.2. The region $\mathcal{R}$ occupied by a body in the deformed configuration is a prismatic cylinder whose cross section is an equilateral triangle as shown in Figure 3.10. Determine the normal and shear traction components that must be applied (as shown in the left-hand figure) such that the Cauchy stress tensor is a pure shear

$$\mathbf{T} = \tau(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1). \qquad (i)$$

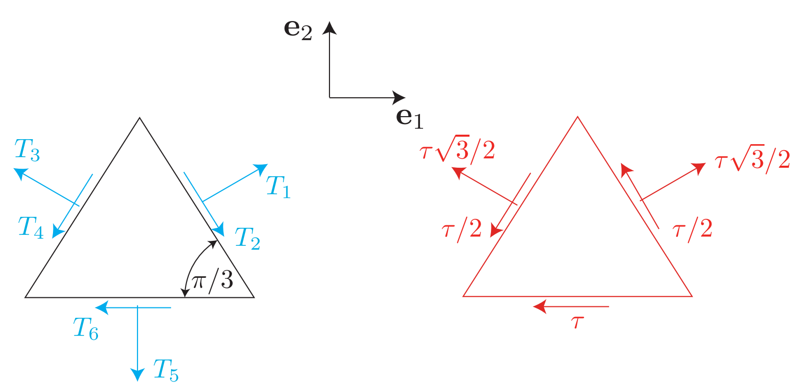


Figure 3.10: Left: Traction components to be determined.

Solution: First consider the bottom surface. The unit outward normal is $\mathbf{n} = -\mathbf{e}_2$ and so the traction vector on this surface is

$$\mathbf{t} = \mathbf{Tn} = \Big[\tau(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1)\Big](-\mathbf{e}_2) = -\tau\mathbf{e}_1. \qquad (ii)$$

The traction shown in the figure is $\mathbf{t} = -T_6\mathbf{e}_1 - T_5\mathbf{e}_2$ and therefore on comparing this with (ii): $T_5 = 0, T_6 = \tau$.

Next consider the upper right-hand surface. The unit outward normal vector is

$$\mathbf{n} = (\sqrt{3}/2)\mathbf{e}_1 + (1/2)\mathbf{e}_2, \qquad (iii)$$

and so the traction on this surface is

$$\mathbf{t} = \mathbf{T}\mathbf{n} \overset{(i),(iii)}{=} \Big[\tau(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1)\Big] \left(\frac{\sqrt{3}}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2 \right) = \frac{\tau}{2}\mathbf{e}_1 + \frac{\tau\sqrt{3}}{2}\mathbf{e}_2. \qquad (iv)$$

The traction component T_1 in the figure is in the direction $\mathbf{n}$ and so

$$T_1 = \mathbf{t}\cdot\mathbf{n} \overset{(iv),(iii)}{=} \left(\frac{\tau}{2}\mathbf{e}_1 + \frac{\tau\sqrt{3}}{2}\mathbf{e}_2 \right) \Big((\sqrt{3}/2)\mathbf{e}_1 + (1/2)\mathbf{e}_2 \Big) = \frac{\tau\sqrt{3}}{2}. \qquad (v)$$

The traction component T_2 in the figure is in the direction

$$(1/2)\mathbf{e}_1 - (\sqrt{3}/2)\mathbf{e}_2, \qquad (vi)$$

(that is perpendicular to $\mathbf{n}$) and so

$$T_2 = \mathbf{t}\cdot\left(\frac{1}{2}\mathbf{e}_1 - \frac{\sqrt{3}}{2}\mathbf{e}_2 \right) \overset{(iv),(vi)}{=} \left(\frac{\tau}{2}\mathbf{e}_1 + \frac{\tau\sqrt{3}}{2}\mathbf{e}_2 \right)\cdot\left(\frac{1}{2}\mathbf{e}_1 - \frac{\sqrt{3}}{2}\mathbf{e}_2 \right) = -\frac{\tau}{2}.$$

A similar calculation gives

$$T_3 = -\tau\sqrt{3}/2, \qquad T_4 = \tau/2.$$

The right-hand figure in Figure 3.10 displays these results. As an exercise you may wish to confirm force and moment equilibrium.

Problem 3.3.3. The stress tensor $\mathbf{T}$ at a particular point in a certain body corresponds to a state of pure shear of magnitude τ with respect to the directions $\mathbf{e}'_1, \mathbf{e}'_2$. As shown in Figure 3.11, the vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ are obtained by rotating the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ through an angle $\pi/4$ about $\mathbf{e}_3$. Calculate the components of $\mathbf{T}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

Solution: Since $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is obtained by rotating $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ through an angle $\pi/4$ about $\mathbf{e}_3$,

$$\mathbf{e}'_1 = \frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2, \qquad \mathbf{e}'_2 = -\frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2, \qquad \mathbf{e}'_3 = \mathbf{e}_3. \qquad (i)$$

We are told that

$$\mathbf{T} = \tau(\mathbf{e}'_1 \otimes \mathbf{e}'_2 + \mathbf{e}'_2 \otimes \mathbf{e}'_1). \qquad (ii)$$

Substituting (i) into (ii) and simplifying, for example

$$\mathbf{e}'_1 \otimes \mathbf{e}'_2 = \left(\frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2 \right) \otimes \left(-\frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2 \right) = -\frac{1}{2}\mathbf{e}_1 \otimes \mathbf{e}_1 - \frac{1}{2}\mathbf{e}_2 \otimes \mathbf{e}_1 + \frac{1}{2}\mathbf{e}_1 \otimes \mathbf{e}_2 + \frac{1}{2}\mathbf{e}_2 \otimes \mathbf{e}_2,$$

leads to

$$\mathbf{T} = -\tau\mathbf{e}_1 \otimes \mathbf{e}_1 + \tau\mathbf{e}_1 \otimes \mathbf{e}_1. \qquad (iii)$$

Thus the matrix of components of $\mathbf{T}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is

$$[T] = \begin{pmatrix} -\tau & 0 & 0 \\ 0 & \tau & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad (iv)$$

Observe from (iv) that $\mathbf{T}$ can be viewed as the superposition of a uniaxial compressive stress τ in the $\mathbf{e}_1$-direction and a uniaxial tensile stress τ in the $\mathbf{e}_2$-direction (when $\tau > 0$). We also know that (this same stress tensor) $\mathbf{T}$ can be viewed as a pure shear with respect to $\mathbf{e}'_1, \mathbf{e}'_2$. This is depicted in Figure 3.11. This example illustrates how the components of $\mathbf{T}$ depend on the basis.

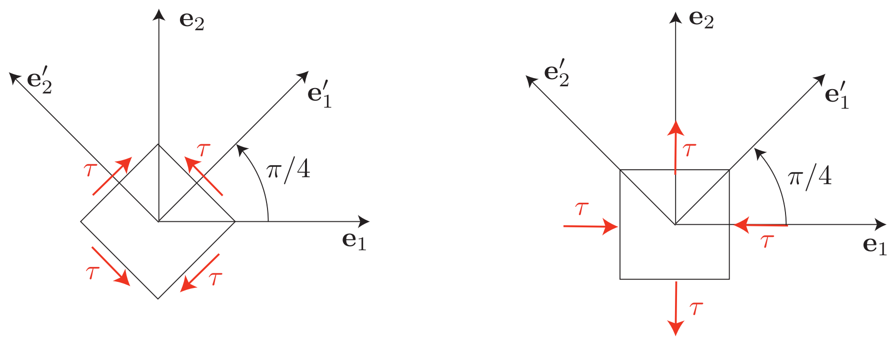


Figure 3.11: The stress tensor $\mathbf{T}$ is a simple shear of magnitude τ with respect to the directions $\mathbf{e}'_1, \mathbf{e}'_2$, and equivalently the superposition of a uniaxial compressive stress τ in the $\mathbf{e}_1$-direction and a uniaxial tensile stress τ in the $\mathbf{e}_2$-direction. (Problem 3.3.3)

Problem 3.3.4. (Continued in Problem 3.21.) Consider a body that in the deformed configuration occupies the annular sector $a \le r \le b, -\beta \le \theta \le \beta, -1/2 \le z \le 1/2$ shown in Figure 3.12. We are using cylindrical polar coordinates (r, θ, z) in the deformed configuration with associated basis vectors $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. Assume that the Cauchy stress components $T_{zr} = T_{z\theta} = T_{zz} = 0$, so that

$$\mathbf{T} = T_{rr}\mathbf{e}_r \otimes \mathbf{e}_r + T_{r\theta}(\mathbf{e}_r \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_r) + T_{\theta\theta}\, \mathbf{e}_\theta \otimes \mathbf{e}_\theta. \qquad (i)$$

Assume further that the remaining stress components depend on r and θ (but not z).

Determine the restrictions on the stress components $T_{rr}(r,\theta), T_{r\theta}(r,\theta)$ and $T_{\theta\theta}(r,\theta)$ arising from the following requirements: (a) the outer curved boundary $r = b$ is traction-free; (b) the inner curved boundary $r = a$ is also traction-free; (c) the resultant force on the top inclined surface $\mathcal{S}$ vanishes; and (d) the resultant moment on $\mathcal{S}$ about O is $m\mathbf{e}_z$.

Later, in Problem 3.21, after we have developed the equilibrium equations, we will explore the consequences of equilibrium.

The kinematics of the bending of a rectangular block into a shape like the one shown in Figure 3.12 was examined previously in Problem 2.5.4. Here we are not told what the undeformed configuration of the body is.

Solution: First consider the outer curved surface $r = b$. Since the outward pointing unit normal vector to it is $\mathbf{e}_r$, the traction $\mathbf{T}\mathbf{e}_r$ acting on this surface has radial and circumferential components T_{rr} and $T_{\theta r}$

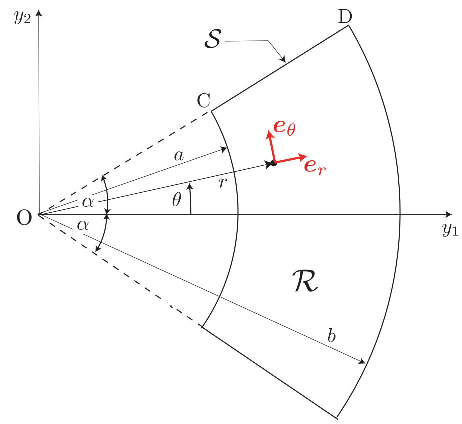


Figure 3.12: In the deformed configuration the body occupies the annular sector $\mathcal{R} = \{(r,\theta,z) : a \le r \le b, -\beta \le \theta \le \beta, -1/2 \le z \le 1/2\}$. (Figure for Problem 3.3.4 .)

respectively. Since (each point on) this surface is traction-free, it follows that

$$T_{rr}(b,\theta) = T_{\theta r}(b,\theta) = 0 \qquad \text{for all} \quad -\beta \le \theta \le \beta. \tag{ii}$$

Likewise, at the traction-free inner curved surface $r = a$ we have

$$T_{rr}(a,\theta) = T_{\theta r}(a,\theta) = 0 \qquad \text{for all} \quad -\beta \le \theta \le \beta. \tag{iii}$$

Next consider the flat inclined surface $\mathcal{S}$ on which $\theta = \beta$. Keep in mind that, in general, $\mathbf{e}_r = \mathbf{e}_r(\theta), \mathbf{e}_\theta = \mathbf{e}_\theta(\theta)$. The outward pointing unit vector on it is $\mathbf{e}_\theta$, or more precisely $\mathbf{e}_\theta(\beta)$, and so from (i), the traction on $\mathcal{S}$ is

$$\mathbf{t} = \mathbf{T}\mathbf{e}_\theta\Big|_{\theta=\beta} = \big(T_{r\theta}\,\mathbf{e}_r + T_{\theta\theta}\,\mathbf{e}_\theta\big)\Big|_{\theta=\beta} = T_{r\theta}(r,\beta)\,\mathbf{e}_r(\beta) + T_{\theta\theta}(r,\beta)\,\mathbf{e}_\theta(\beta). \tag{iv}$$

The resultant force on this surface is therefore

$$\begin{aligned}\int_{\mathcal{S}} \mathbf{t}\,dA_y = \int_a^b\int_{-1/2}^{1/2} \mathbf{t}\,dzdr = \int_a^b \mathbf{t}\,dr \ &= \int_a^b \Big[T_{r\theta}(r,\beta)\,\mathbf{e}_r(\beta) + T_{\theta\theta}(r,\beta)\,\mathbf{e}_\theta(\beta)\Big]\,dr = \\ &= \left[\int_a^b T_{r\theta}(r,\beta)\,dr\right]\mathbf{e}_r(\beta) + \left[\int_a^b T_{\theta\theta}(r,\beta)\,dr\right]\mathbf{e}_\theta(\beta).\end{aligned}$$

Since the resultant force on this surface vanishes it follows that

$$\int_a^b T_{r\theta}(r,\beta)\,dr = 0, \qquad \int_a^b T_{\theta\theta}(r,\beta)\,dr = 0. \tag{v}$$

The resultant moment on $\mathcal{S}$ about O is

$$\begin{aligned}\int_{\mathcal{S}} \mathbf{y}\times\mathbf{t}\,dA_y = \int_a^b\int_{-1/2}^{1/2} \mathbf{y}\times\mathbf{t}\,dzdr = \int_a^b \mathbf{y}\times\mathbf{t}\,dr \ &= \int_a^b r\,\mathbf{e}_r(\beta)\times\Big[T_{r\theta}(r,\beta)\,\mathbf{e}_r(\beta) + T_{\theta\theta}(r,\beta)\,\mathbf{e}_\theta(\beta)\Big]\,dr = \\ &= \left[\int_a^b rT_{\theta\theta}(r,\beta)\,dr\right]\mathbf{e}_z, \qquad (vi)\end{aligned}$$

where we have used $\mathbf{e}_r \times \mathbf{e}_r = \mathbf{o}$ and $\mathbf{e}_r \times \mathbf{e}_\theta = \mathbf{e}_z$. (Please derive ($vi$) using physical arguments, without using the vector cross-product.) We are told that the resultant moment on this surface is $m\mathbf{e}_z$ and so

$$\left[\int_a^b rT_{\theta\theta}(r,\beta)\, dr\right] = m. \qquad (vii)$$

3.4 Field equations.

We are now in a position to derive the local versions – the field equations – of the global equilibrium principles for force and moment balance (3.8) and (3.9).

Consequence (4): Equilibrium equations.

Consider *force balance* (3.8), which in component form reads

$$\int_{\partial\mathcal{D}} t_i \,\mathrm{d}A_y + \int_{\mathcal{D}} b_i \,\mathrm{d}V_y = 0 \ . \qquad (3.23)$$

Considering the first term, we first trade traction for stress using (3.18), and then convert the surface integral into a volume integral by using the divergence theorem:

$$\int_{\partial\mathcal{D}} t_i \,\mathrm{d}A_y = \int_{\partial\mathcal{D}} T_{ij}n_j \,\mathrm{d}A_y = \int_{\mathcal{D}} \frac{\partial T_{ij}}{\partial y_j}\, dV_y, \qquad (3.24)$$

and so (3.23) yields

$$\int_{\mathcal{D}} \left(\frac{\partial T_{ij}}{\partial y_j} + b_i\right) \mathrm{d}V_y = 0 \ . \qquad (3.25)$$

Since this must hold for all parts $\mathcal{D}$ of the body, and assuming the integrand to be continuous, it follows by localization (Section 1.8.3) that the integrand itself must vanish at each point in $\mathcal{R}$. We thus conclude that

$$\frac{\partial T_{ij}}{\partial y_j} + b_i \ = \ 0 \qquad \text{at each } \mathbf{y} \in \mathcal{R}, \qquad (3.26)$$

which can be written in basis-free form by using (1.166) (page 64) as

$$\mathrm{div}\, \mathbf{T} + \mathbf{b} = \mathbf{0} \qquad \text{at each } \mathbf{y} \in \mathcal{R}. \qquad (3.27)$$

The *equilibrium equation* (3.27) is the field equations corresponding to force balance. It must hold at each point in the body.

Conversely, when the equilibrium equation (3.27) and the traction-stress relation (3.18) hold, then the global force balance law (3.8) holds. (Show this.)

Consequence (5): Symmetry of the stress tensor.

We turn next to *moment balance* (3.9). Recall that for any two vectors $\mathbf{a}$ and $\mathbf{b}$, the i^{th} component of the vector $\mathbf{a} \times \mathbf{b}$ is $e_{ijk}a_jb_k$ where e_{ijk} is the Levi-Civita symbol introduced in (1.38). Thus we can write (3.9) in component form as

$$\int_{\partial\mathcal{D}} (\mathbf{y} \times \mathbf{t})_i \, \mathrm{d}A_y + \int_{\mathcal{D}} (\mathbf{y} \times \mathbf{b})_i \, \mathrm{d}V_y = \int_{\partial\mathcal{D}} e_{ijk} \, y_j \, t_k \, \mathrm{d}A_y + \int_{\mathcal{D}} e_{ijk} \, y_j b_k \, \mathrm{d}V_y = 0 \; . \tag{3.28}$$

The term involving the traction can be simplified by first using the traction-stress relation (3.19), then using the divergence theorem and finally expanding the result. This leads to

$$\begin{aligned} \int_{\partial\mathcal{D}} e_{ijk} \, y_j \, t_k \, \mathrm{d}A_y &= \int_{\partial\mathcal{D}} e_{ijk} \, y_j \, T_{km} n_m \, \mathrm{d}A_y = \int_{\mathcal{D}} e_{ijk} \, \frac{\partial}{\partial y_m}(y_j \, T_{km}) \, \mathrm{d}V_y \\ &= \int_{\mathcal{D}} e_{ijk} \left(\delta_{jm} \, T_{km} + y_j \, \frac{\partial T_{km}}{\partial y_m} \right) \, \mathrm{d}V_y \; . \end{aligned} \tag{3.29}$$

Substituting (3.29) into (3.28) and making use of the equilibrium equation (3.26) now yields

$$\int_{\mathcal{D}} e_{ijk} \, T_{kj} \, \mathrm{d}V_y = 0 \; . \tag{3.30}$$

Since (3.30) must hold for all choices of D, it follows by localization that

$$e_{ijk} \, T_{kj} = 0 \qquad \text{at each } \mathbf{y} \in \mathcal{R}. \tag{3.31}$$

One way in which to see what the three scalar equations (3.31) imply is to write them out explicitly. For example for $i = 1$ we have $e_{1jk} \, T_{kj} = e_{123} \, T_{32} + e_{132} \, T_{23}$ because all of the other e_{ijk} terms have at least two repeated subscripts and thus vanish. Since $e_{123} = 1$ and $e_{132} = -1$ it now follows that $e_{1jk} \, T_{kj} = T_{32} - T_{23}$ and therefore (3.31) implies that $T_{23} = T_{32}$. The cases $i = 2$ and $i = 3$ can be dealt with similarly. Thus we conclude that

$$T_{12} = T_{21}, \qquad T_{23} = T_{32}, \qquad T_{31} = T_{13},$$

which we can write as

$$T_{ij} = T_{ji} \qquad \text{at each } \mathbf{y} \in \mathcal{R}. \tag{3.32}$$

Thus the stress tensor $\mathbf{T}$ is *symmetric*:

$$\mathbf{T} = \mathbf{T}^T \qquad \text{at each } \mathbf{y} \in \mathcal{R}. \tag{3.33}$$

This is equivalent to (3.31) and is a local consequence of moment balance.

Exercise: To show the symmetry of the stress tensor without explicitly writing out the terms in (3.31) (as we did above) multiply (3.31) by e_{ipq} and use of the identity $e_{ijk}\, e_{ipq} = \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}$.

Conversely, when the symmetry condition (3.33), the equilibrium equation (3.27), and the traction-stress relation (3.18) all hold, then the global moment balance (3.9) holds. (Show this.)

3.4.1 Summary

In summary, the global equilibrium principles of force and moment balance hold if and only if the Cauchy stress tensor field $\mathbf{T}(\mathbf{y})$ obeys

$$\boxed{\begin{array}{rcl} \operatorname{div}\mathbf{T} + \mathbf{b} & = & \mathbf{o}, \\ & & \qquad\qquad \text{at each } \mathbf{y} \in \mathcal{R}, \\ \mathbf{T} & = & \mathbf{T}^T, \end{array}} \tag{3.34}$$

with the traction on a surface related to the stress through

$$\boxed{\mathbf{t}(\mathbf{y},\mathbf{n}) = \mathbf{T}(\mathbf{y})\mathbf{n}.} \tag{3.35}$$

In cartesian components,

$$\frac{\partial T_{ij}}{\partial y_j} + b_i = 0, \qquad T_{ij} = T_{ji}, \qquad t_i = T_{ij} n_j. \tag{3.36}$$

3.5 Principal stresses.

Since the Cauchy stress tensor $\mathbf{T}$ is symmetric, it has three real eigenvalues, τ_1, τ_2, τ_3, and a set of three corresponding orthonormal eigenvectors, $\mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3$:

$$\mathbf{T}\mathfrak{t}_i = \tau_i\ \mathfrak{t}_i \qquad\qquad \text{(no sum on i)}. \tag{3.37}$$

The eigenvalues τ_i are called the *principal stresses* and the eigenvectors $\mathfrak{t}_i$ define the *principal directions of Cauchy stress*. (Caution: Please note the distinction between the two lowercase boldface t's: $\mathbf{t}$ and $\mathfrak{t}_i$ denoting the traction and the principal stress directions respectively.)

The triplet of vectors $\{\mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3\}$ defines an orthonormal basis referred to as a principal basis of stress. The matrix of stress components in this basis is diagonal and is given by

$$[T] = \begin{pmatrix} \tau_1 & 0 & 0 \\ 0 & \tau_2 & 0 \\ 0 & 0 & \tau_3 \end{pmatrix}. \tag{3.38}$$

We can express $\mathbf{T}$ as

$$\mathbf{T} = \tau_1\, \mathfrak{t}_1 \otimes \mathfrak{t}_1 + \tau_2\, \mathfrak{t}_2 \otimes \mathfrak{t}_2 + \tau_3\, \mathfrak{t}_3 \otimes \mathfrak{t}_3 = \sum_{i=1}^{3} \tau_i\, \mathfrak{t}_i \otimes \mathfrak{t}_i. \tag{3.39}$$

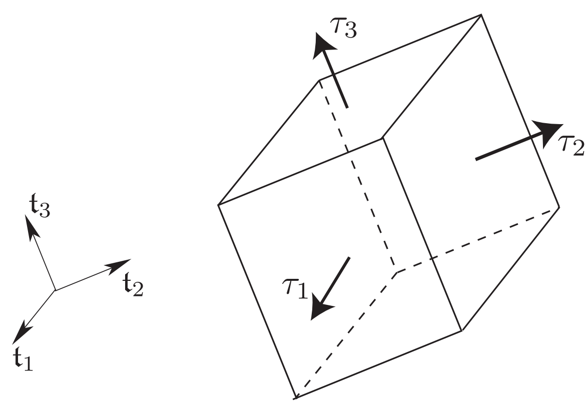


Figure 3.13: Principal stresses τ_1, τ_2, τ_3 and corresponding principal directions $\mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3$.

When the components of $\mathbf{T}$ and $\mathbf{n}$ are taken with respect to a principal basis for $\mathbf{T}$, one can show that the normal stress (3.6) and the magnitude of the resultant shear stress (3.7) can be written as (Problems 3.2 and 3.3)

$$T_{\text{normal}}(\mathbf{n}) = \mathbf{t}(\mathbf{n}) \cdot \mathbf{n} = \mathbf{Tn} \cdot \mathbf{n} = T_{ij} n_i n_j = \tau_1\, n_1^2 + \tau_2\, n_2^2 + \tau_3\, n_3^2, \tag{3.40}$$

$$T_{\text{shear}}^2 = |\mathbf{t}(\mathbf{n})|^2 - T_{\text{normal}}^2 = \tau_1^2 n_1^2 + \tau_2^2 n_2^2 + \tau_3^2 n_3^2 - (\tau_1 n_1^2 + \tau_2 n_2^2 + \tau_3 n_3^2)^2. \tag{3.41}$$

An important characteristic of the principal stresses and principal directions can be seen by asking the question "from among all planes passing through a given point, on which is $T_{\text{normal}}(\mathbf{n})$ largest? And on which is it smallest?" This requires one to consider $T_{\text{normal}}(\mathbf{n})$ as a function of the unit vector $\mathbf{n}$ and to find the specific vector(s) $\mathbf{n}$ at which it has its extrema. One can show (Problem 3.2) that the maximum value of $T_{\text{normal}}(\mathbf{n})$ over all unit vectors $\mathbf{n}$ is the largest of the principal stresses:

$$T_{\text{normal}}(\mathbf{n})\Big|_{\text{max over } \mathbf{n}} = \text{maximum of } \{\tau_1, \tau_2, \tau_3\}, \tag{3.42}$$

and that the smallest value of $T_{\text{normal}}(\mathbf{n})$ is the smallest principal stress.

The maximum value of $T_{\text{shear}}(\mathbf{n})$ over all unit vectors $\mathbf{n}$ is (Problem 3.3)

$$T_{\text{shear}}(\mathbf{n})\Big|_{\text{max over }\mathbf{n}} = \text{maximum of } \left\{\frac{1}{2}|\tau_1-\tau_2|, \quad \frac{1}{2}|\tau_2-\tau_3|, \quad \frac{1}{2}|\tau_3-\tau_1|\right\}. \tag{3.43}$$

One can also show that there is always a plane (through each point of a body) on which $T_{\text{shear}}(\mathbf{n})$ vanishes, but in general, there is no plane on which $T_{\text{normal}}(\mathbf{n})$ vanishes (though there might be in special cases) (Problem 3.4).

Finally it is worth emphasizing that the principal directions of the stress tensor $\mathbf{T}$ have no relationship, in general, to the principal directions of the stretch tensors $\mathbf{U}$ or $\mathbf{V}$. There may be a relationship between them for *particular materials*, but this depends on the constitutive law. In particular, we will find that for an isotropic elastic material, the principal directions $\{\mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3\}$ of $\mathbf{T}$ coincide with the principal directions $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ of the Eulerian stretch tensor $\mathbf{V}$. See also Problem 3.27.

3.6 Mean pressure and deviatoric stress.

It is sometimes convenient to decompose the stress additively into the sum of two parts, a hydrostatic part and a deviatoric part. By definition, the *mean pressure* is the (negative of the) average normal stress

$$= -\frac{1}{3}T_{kk} = -\frac{1}{3}\operatorname{tr}\mathbf{T},$$

and so the hydrostatic part of stress is $\frac{1}{3}(\operatorname{tr}\mathbf{T})\,\mathbf{I}$. The remaining part of the stress is called the *deviatoric part* which we denote by

$$\mathbf{T}^{(\text{dev})} := \mathbf{T} - \frac{1}{3}(\text{tr }\mathbf{T})\mathbf{I}, \qquad T^{(\text{dev})}_{ij} = T_{ij} - \frac{1}{3}T_{kk}\delta_{ij}. \tag{3.44}$$

Note that the trace of the deviatoric stress vanishes. Thus we have the decomposition

$$\mathbf{T} = \mathbf{T}^{(\text{dev})} + \frac{1}{3}(\text{tr }\mathbf{T})\mathbf{I}. \tag{3.45}$$

The principal scalar invariants[7] of the *deviatoric stress* are sometimes of interest, the second of which plays a role in the theory of plasticity.

[7]Recall our discussion (in the text surrounding (2.47)) of the three principal scalar invariants $I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C})$ of the right Cauchy-Green tensor $\mathbf{C}$. One can similarly examine the principal scalar invariants of the Cauchy stress $\mathbf{T}$.

3.7 Formulation of mechanical principles with respect to a reference configuration.

A few videos on some of the material in this section can be found here.

Thus far, our discussion of traction, stress, balance laws and field equations, did not allude to a reference configuration. Though not conceptually necessary, it is often convenient to introduce a reference configuration and to refer various kinematic quantities (e.g. the area and surface normal vector) to the geometry of that configuration. Often, the deformed configuration is not known a priori and is to be determined, while the reference configuration can be chosen in a convenient manner.

Consider an arbitrary deformation

$$\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x})$$

from some reference configuration to the deformed configuration, and let $\mathcal{R}_{\rm R}$ and $\mathcal{R}$ be the corresponding regions occupied by the body. Consider some part[8] of the body and let $\mathcal{D}_{\rm R}$ and $\mathcal{D}$ be the regions occupied by this part in the two configurations. The reference configuration *need not* coincide with a configuration that the body actually occupies, only one that it *could* occupy.

First consider the body force. If an infinitesimal part of the body has volume dV_y in the deformed configuration, the body force on this part is $\mathbf{b}\, dV_y$ since $\mathbf{b}$ is the body force per unit deformed volume. Let dV_x denote the volume of this part in the reference configuration. If we introduce the *body force per unit reference volume* and denote it by $\mathbf{b}_{\rm R}$, the body force on this part can also be expressed as $\mathbf{b}_{\rm R}\, dV_x$. Therefore we have

$$\text{Body force on infinitesimal part} = \mathbf{b}\, dV_y = \mathbf{b}_{\rm R}\, dV_x. \tag{3.46}$$

We know from Section 2.4.3 that these volumes are related by $dV_y = J\, dV_x$ where $J = \det \mathbf{F}$. Thus

$$\boxed{\mathbf{b}_{\rm R} = J\, \mathbf{b}.} \tag{3.47}$$

Keep in mind that the force $\mathbf{b}_{\rm R}\, dV_x$ acts on the *deformed* body.

Next we turn to traction and stress. First, recall that the stress tensor $\mathbf{T}$ represents the force per unit *deformed* area and the traction vector $\mathbf{Tn}$ acts on the surface whose normal

[8] Recall that a part involves a fixed set of material points, and so $\mathcal{D}_{\rm R}$ and $\mathcal{D}$ are associated with the same particles.

in the *deformed* configuration is $\mathbf{n}$. Even though the forces act on the deformed body, it is sometimes convenient to refer the *geometry* to the geometry of the reference configuration. Therefore we now consider a different stress tensor, $\mathbf{S}$, that represents force per unit *reference* area, with the associated traction vector $\mathbf{S}\mathbf{n}_{\mathrm{R}}$ acting (in the deformed body) on the surface whose normal in the *reference* configuration is $\mathbf{n}_{\mathrm{R}}$. This is illustrated in Figure 3.14.

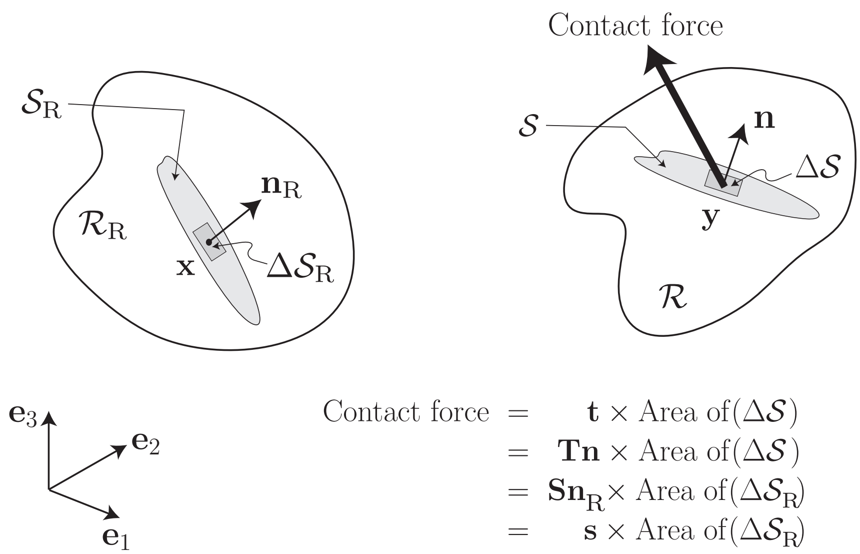


Figure 3.14: The surface $\mathcal{S}$ and a surface element $\Delta\mathcal{S}$ in the deformed configuration, and their images $\mathcal{S}_{\mathrm{R}}$ and $\Delta\mathcal{S}_{\mathrm{R}}$ in the reference configuration. The vectors $\mathbf{n}$ and $\mathbf{n}_{\mathrm{R}}$ are normal to these respective surfaces. Different (equivalent) ways for characterizing the contact force on $\Delta\mathcal{S}$ are noted in the figure. Keep in mind that the contact force acts on the deformed body.

Let $\mathcal{S}$ be a surface in $\mathcal{R}$ and let $\mathcal{S}_{\mathrm{R}}$ be its image in the reference configuration as illustrated in Figure 3.14. Let $\mathbf{y}$ be a point on $\mathcal{S}$ and let $\mathbf{x}$ be its image on $\mathcal{S}_{\mathrm{R}}$. We denote a unit vector normal to $\mathcal{S}$ at $\mathbf{y}$ by $\mathbf{n}$, and the unit vector normal to $\mathcal{S}_{\mathrm{R}}$ at $\mathbf{x}$ by $\mathbf{n}_{\mathrm{R}}$. Finally, let $\Delta\mathcal{S}$ be an infinitesimal surface element on $\mathcal{S}$ at $\mathbf{y}$ whose area is $\mathrm{d}A_y$, and let $\Delta\mathcal{S}_{\mathrm{R}}$ be its image in the reference configuration whose area is $\mathrm{d}A_x$.

The contact *force* on the surface element $\Delta\mathcal{S}$ is the product of the traction $\mathbf{t}(\mathbf{n})$ and the area $\mathrm{d}A_y$:

$$\text{Contact force on } \Delta\mathcal{S} \;=\; \mathbf{t}(\mathbf{n})\,\mathrm{d}A_y \;=\; \mathbf{T}\mathbf{n}\,\mathrm{d}A_y. \tag{3.48}$$

Keep in mind that dA_y and $\mathbf{n}$ here are the area and normal vector of the surface $\Delta\mathcal{S}$ in the deformed configuration. Our goal is to refer this contact force to the geometry (i.e. area and normal) of the reference configuration.

Recall Nanson's formula (2.38) which is the geometric relation between the vector areas $\mathrm{d}A_y\mathbf{n}$ and $\mathrm{d}A_x\mathbf{n}_\mathrm{R}$:

$$\mathrm{d}A_y\,\mathbf{n} = \mathrm{d}A_x\,J\,\mathbf{F}^{-T}\,\mathbf{n}_\mathrm{R}. \tag{3.49}$$

Combining (3.49) with (3.48) gives the

$$\text{Contact force on } \Delta\mathcal{S} \;=\; \mathbf{T}\,(J\,\mathbf{F}^{-T}\,\mathbf{n}_\mathrm{R}\,\mathrm{d}A_x) \;=\; (J\,\mathbf{T}\,\mathbf{F}^{-T})\,\mathbf{n}_\mathrm{R}\,\mathrm{d}A_x \tag{3.50}$$

It is natural therefore to define a tensor $\mathbf{S}$ and a vector $\mathbf{s}$ by the respective equations

$$\boxed{\mathbf{S} = J\,\mathbf{T}\,\mathbf{F}^{-T}\,,} \tag{3.51}$$

$$\boxed{\mathbf{s} = \mathbf{S}\,\mathbf{n}_\mathrm{R}\,,} \tag{3.52}$$

so that the contact force on $\Delta\mathcal{S}$ can then be written in the equivalent forms

$$\text{Contact force on } \Delta\mathcal{S} \;=\boxed{\mathbf{t}\,\mathrm{d}A_y = \mathbf{s}\,\mathrm{d}A_x,} \tag{3.53}$$

and

$$\text{Contact force on } \Delta\mathcal{S} \;=\boxed{\mathbf{T}\mathbf{n}\,\mathrm{d}A_y = \mathbf{S}\mathbf{n}_\mathrm{R}\,\mathrm{d}A_x\;.} \tag{3.54}$$

Thus $\mathbf{s}$ is the contact force per unit *referential area*. It acts on the surface element $\Delta\mathcal{S}$ in the deformed body. This is described by the text in Figure 3.14. The vector $\mathbf{s}$ is called the **first Piola-Kirchhoff traction vector** and the tensor $\mathbf{S}$ is called the **first Piola-Kirchhoff stress tensor**.

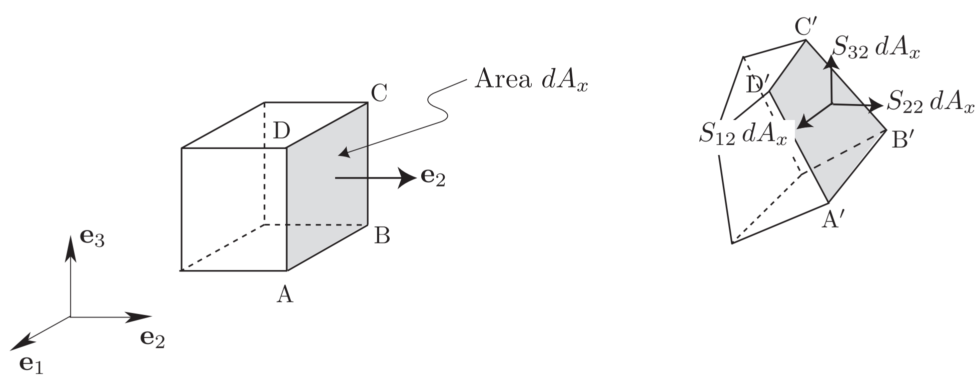


Figure 3.15: Physical significance of the components of the stress tensor $\mathbf{S}$: the shaded surface in the reference configuration is normal to $\mathbf{e}_2$ and has area dA_x. The i^{th} component of *force* acting on the image of this surface in the deformed configuration is $S_{i2} \times dA_x$.

The physical significance of the components of $\mathbf{S}$, can be deduced as follows. Let

$$\mathbf{S} = S_{ij}\,\mathbf{e}_i \otimes \mathbf{e}_j.$$

Consider a surface parallel to one of the coordinate planes, say $\mathbf{n}_\mathrm{R} = \mathbf{e}_j$. According to $\mathbf{s}(\mathbf{n}_\mathrm{R}) = \mathbf{S}\mathbf{n}_\mathrm{R}$, the traction on this surface is $\mathbf{s}(\mathbf{e}_j) = \mathbf{S}\mathbf{e}_j$. The i^{th}-component of this traction is therefore $s_i(\mathbf{e}_j) = \mathbf{s}(\mathbf{e}_j)\cdot\mathbf{e}_i = \mathbf{S}\mathbf{e}_j\cdot\mathbf{e}_i = S_{ij}$. Thus

$$S_{ij} = s_i(\mathbf{e}_j). \tag{3.55}$$

Therefore S_{ij} is the i^{th} component of force per unit referential area acting on the surface that is normal to the j^{th} direction in the reference configuration.

For example consider a surface element that is normal to $\mathbf{e}_2$ in the reference configuration as shown in Figure 3.15. Then by taking $\mathbf{n}_\mathrm{R} = \mathbf{e}_2$ in (3.52), the contact force on this element can be written as

$$\begin{aligned}\text{Contact force on } \Delta\mathcal{S} \quad &= \mathbf{t}\,\mathrm{d}A_y = \mathbf{s}\,\mathrm{d}A_x = \mathbf{S}\,\mathbf{e}_2\,\mathrm{d}A_x = (S_{12}\,\mathbf{e}_1 + S_{22}\,\mathbf{e}_2 + S_{32}\,\mathbf{e}_3)\,\mathrm{d}A_x = \\ &= (S_{12}\,\mathrm{d}A_x)\,\mathbf{e}_1 + (S_{22}\,\mathrm{d}A_x)\,\mathbf{e}_2 + (S_{32}\,\mathrm{d}A_x)\,\mathbf{e}_3\,.\end{aligned} \tag{3.56}$$

This is illustrated in Figure 3.15.

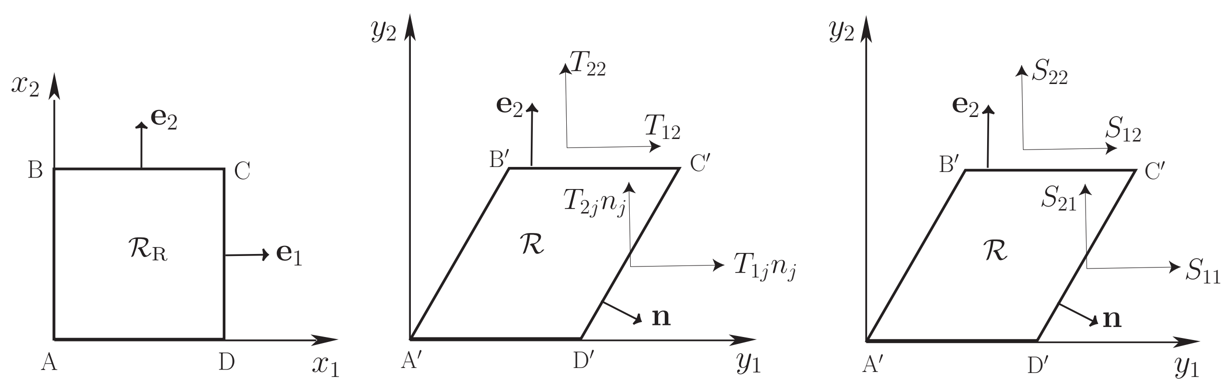


Figure 3.16: Simple shear. The middle and rightmost figures *both* show the region $\mathcal{R}$ occupied by the body in the deformed configuration. They depict the tractions on the faces $B'C'$ and $C'D'$ in two different, but equivalent, ways: the middle figure describes the traction in terms of the components of the stress $\mathbf{T}$ while the rightmost figure describes them in terms of the components of $\mathbf{S}$. The corresponding forces are found by multiplying each traction by the area of the relevant surface in either the deformed or reference configuration as appropriate.

To illustrate this further, consider a simple shear deformation of a block as shown in Figure 3.16. The leftmost figure shows the region $\mathcal{R}_\mathrm{R}$, while both the middle and right figures show the region $\mathcal{R}$. The unit outward normal vector to the face $C'D'$ is $\mathbf{n}$ and therefore the

$$\text{Contact force on } C'D' = \mathbf{T}\mathbf{n}\times|C'D'| = [(T_{1j}n_j)\mathbf{e}_1 + (T_{2j}n_j)\mathbf{e}_2 + (T_{3j}n_j)\mathbf{e}_3]\times|C'D'| \; ; \tag{3.57}$$

this is illustrated in the middle figure. Since the face CD, which is the image of $C'D'$, has a unit outward normal $\mathbf{e}_1$, we can *equivalently* write

$$\text{Contact force on } C'D' = \mathbf{S}\mathbf{e}_1 \times |CD| = [S_{11}\mathbf{e}_1 + S_{21}\mathbf{e}_2 + S_{31}\mathbf{e}_3] \times |CD|; \tag{3.58}$$

this is illustrated in the right most figure. Similarly the unit outward normal vectors to the face $B'C'$ and its image BC are both $\mathbf{e}_2$, and therefore we can write

$$\begin{aligned} \text{Contact force on } B'C' \;\; &= \;\; \mathbf{T}\mathbf{e}_2 \times |B'C'| = [T_{12}\mathbf{e}_1 + T_{22}\mathbf{e}_2 + T_{32}\mathbf{e}_3] \times |B'C'|, \\ &= \;\; \mathbf{S}\mathbf{e}_2 \times |BC| = [S_{12}\mathbf{e}_1 + S_{22}\mathbf{e}_2 + S_{32}\mathbf{e}_3] \times |BC|; \end{aligned} \tag{3.59}$$

these are also displayed in Figure 3.16.

We now turn to the equilibrium equations. Recall the discussion in Section 2.8 of the material and spatial descriptions of a field. As mentioned there, any field that is described spatially as a function of $\mathbf{y}$ can be converted to a field described referentially as a function of the reference position $\mathbf{x}$. We do this by using the deformation $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x})$ to change $\mathbf{y} \to \mathbf{x}$. It turns out that it is convenient to express the first Piola-Kirchhoff stress tensor field referentially as $\mathbf{S}(\mathbf{x})$, so that (3.51), (3.52) would read (with more detail)

$$\mathbf{S}(\mathbf{x}) = J(\mathbf{x})\, \mathbf{T}(\widehat{\mathbf{y}}(\mathbf{x}))\, \mathbf{F}^{-T}(\mathbf{x}) \;, \tag{3.60}$$

$$\mathbf{s}(\mathbf{x}, \mathbf{n}_{\mathrm{R}}) = \mathbf{S}(\mathbf{x})\, \mathbf{n}_{\mathrm{R}} \;. \tag{3.61}$$

Substituting $\mathbf{T} = \mathrm{J}^{-1}\mathbf{S}\mathbf{F}^T$ into $\operatorname{div}\mathbf{T}$ and using (2.121) from page 172 shows that

$$\operatorname{div}\mathbf{T} = \mathrm{J}^{-1}\mathrm{Div}\,\mathbf{S},$$

where $\operatorname{div}\mathbf{T}$ and $\mathrm{Div}\,\mathbf{S}$ are the vector fields whose ith cartesian components are

$$\Big(\operatorname{div}\mathbf{T}\Big)_i = \frac{\partial T_{ij}}{\partial y_j}, \qquad \Big(\mathrm{Div}\,\mathbf{S}\Big)_i = \frac{\partial S_{ij}}{\partial x_j},$$

respectively[9]. Therefore the force equilibrium equation $(3.34)_1$ can be written in the equivalent form

$$\boxed{\mathrm{Div}\,\mathbf{S} + \mathbf{b}_{\mathrm{R}} = \mathbf{0}.} \tag{3.62}$$

[9]Note the distinction between Div and div. For any tensor field $\mathbf{A}(\mathbf{x})$, the vector field $\mathrm{Div}\,\mathbf{A}$ has cartesian components $\partial A_{ij}/\partial x_j$:

$$\Big(\mathrm{Div}\,\mathbf{A}\Big)_i = \frac{\partial A_{ij}}{\partial x_j}, \qquad \mathrm{Div}\,\mathbf{A} = \frac{\partial A_{ij}}{\partial x_j}\,\mathbf{e}_i.$$

There is a parallel distinction between Grad/grad, and Curl/curl.

The moment equilibrium equation $(3.34)_2$, in view of $\mathbf{T} = \mathrm{J}^{-1}\mathbf{S}\mathbf{F}^T$, yields

$$\boxed{\mathbf{S}\mathbf{F}^T = \mathbf{F}\mathbf{S}^T.} \tag{3.63}$$

These field equations must hold at every point $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$. The traction on a surface is related to the stress through

$$\mathbf{s} = \mathbf{S}\mathbf{n}_{\mathrm{R}}. \tag{3.64}$$

In cartesian component form, equations (3.62), (3.63) and (3.64) read

$$\frac{\partial S_{ij}}{\partial x_j} + b_i^{\mathrm{R}} = 0, \qquad S_{ik}F_{jk} = F_{ik}S_{jk}, \qquad s_i = S_{ij}n_j^{\mathrm{R}}. \tag{3.65}$$

Note that the first Piola-Kirchhoff stress tensor is *not* symmetric in general. This implies in particular that $\mathbf{S}$ may not have three real eigenvalues and so we will not (usually) speak of the principal values of the first Piola-Kirchhoff stress tensor.

Before leaving this section it is instructive to express the various terms of the global balance laws for force and moment equilibrium in terms of these referential quantities. Let $\mathcal{D}_{\mathrm{R}}$ and $\mathcal{D}$ be the regions occupied by a part of the body in the reference and deformed configurations respectively. By integrating (3.53) over the body and using (3.52) we see that the resultant contact force on this part is

$$= \int_{\partial\mathcal{D}} \mathbf{t}\, dA_y \overset{(3.53)}{=} \int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{s}\, dA_x \overset{(3.52)}{=} \int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{S}\mathbf{n}_{\mathrm{R}}\, dA_x.$$

By integrating (3.46) over the body we see that the resultant body force on this part is

$$= \int_{\mathcal{D}} \mathbf{b}\, dV_y = \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{b}_{\mathrm{R}}\, dV_x.$$

Consequently, the balance law (3.8) for force equilibrium can be written equivalently as

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{s}\, \mathrm{d}A_x + \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{b}_{\mathrm{R}}\, \mathrm{d}V_x = \mathbf{0}, \tag{3.66}$$

which must hold for all $\mathcal{D}_{\mathrm{R}} \subset \mathcal{R}_{\mathrm{R}}$. An alternative (simpler) derivation of the field equation $(3.63)_1$ involves using $\mathbf{s} = \mathbf{S}\mathbf{n}_{\mathrm{R}}$ and the divergence theorem on (3.66) and then localizing the result.

Similarly, the resultant moment of the contact force is given by

$$= \int_{\partial\mathcal{D}} \mathbf{y} \times \mathbf{t}\, dA_y \overset{(3.53)}{=} \int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{y} \times \mathbf{s}\, dA_x \overset{(3.52)}{=} \int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{y} \times \mathbf{S}\mathbf{n}_{\mathrm{R}}\, dA_x$$

where it is understood that $\mathbf{y}$ in the second and third expressions is $\widehat{\mathbf{y}}(\mathbf{x})$. In this way one finds that the balance law for moment equilibrium (3.9) can be written equivalently as

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{y} \times \mathbf{S}\mathbf{n}_{\mathrm{R}}\, dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{y} \times \mathbf{b}_{\mathrm{R}}\, dV_x = \mathbf{0}\ . \tag{3.67}$$

It is worth noting that it is *not* $\mathbf{x} \times \mathbf{S}\mathbf{n}_{\mathrm{R}}$ that appears here but rather $\mathbf{y}(\mathbf{x}) \times \mathbf{S}\mathbf{n}_{\mathrm{R}}$.

3.7.1 Worked examples.

Problem 3.7.1. Bending of a block.

Consider a block undergoing a bending deformation as depicted in Figure 3.17, the kinematics of which were analyzed previously in Problem 2.5.4. We found there that the deformation that takes a particle from (x_1, x_2, x_3) to (y_1, y_2, y_3) is characterized by

$$y_1 = r(x_1)\cos\theta(x_2), \qquad y_2 = r(x_1)\sin\theta(x_2), \qquad y_3 = \Lambda x_3, \tag{i}$$

where

$$r(x_1) > 0, \qquad r'(x_1) > 0, \qquad \theta'(x_2) > 0, \qquad \theta(x_2) = -\theta(-x_2), \qquad \Lambda = \text{constant}. \tag{ii}$$

Moreover, we showed that the principal stretches were

$$\lambda_1 = r'(x_1), \qquad \lambda_2 = r(x_1)\theta'(x_2), \qquad \lambda_3 = \Lambda, \tag{iii}$$

and the deformation gradient tensor was

$$\mathbf{F} = \lambda_1(\mathbf{e}_r \otimes \mathbf{e}_1) + \lambda_2(\mathbf{e}_\theta \otimes \mathbf{e}_2) + \lambda_3(\mathbf{e}_z \otimes \mathbf{e}_3). \tag{iv}$$

Here $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ are the basis vectors associated with cylindrical polar coordinates (r, θ, z) in the deformed configuration, i.e.

$$\mathbf{e}_r = \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2, \qquad \mathbf{e}_\theta = -\sin\theta\,\mathbf{e}_1 + \cos\theta\,\mathbf{e}_2, \qquad \mathbf{e}_z = \mathbf{e}_3. \tag{v}$$

Assume that the first Piola-Kirchhoff stress tensor field is given by (cf. (iv))

$$\mathbf{S} = \sigma_1(\mathbf{e}_r \otimes \mathbf{e}_1) + \sigma_2(\mathbf{e}_\theta \otimes \mathbf{e}_2) + \sigma_3(\mathbf{e}_z \otimes \mathbf{e}_3); \tag{vi}$$

moreover, since the principal stretches are independent of x_3, assume that the same is true of the σ_i's:

$$\sigma_i = \sigma_i(x_1, x_2). \tag{vii}$$

Keep in mind that since we do not have a constitutive relation, (vi) does not follows from (iv).

Determine the restrictions on the functions $\sigma_i(x_1, x_2)$ arising from the following requirements: (a) the two curved boundaries of $\mathcal{R}$ are traction free, (b) the resultant force on the top inclined flat face $\mathcal{S}$ of $\mathcal{R}$

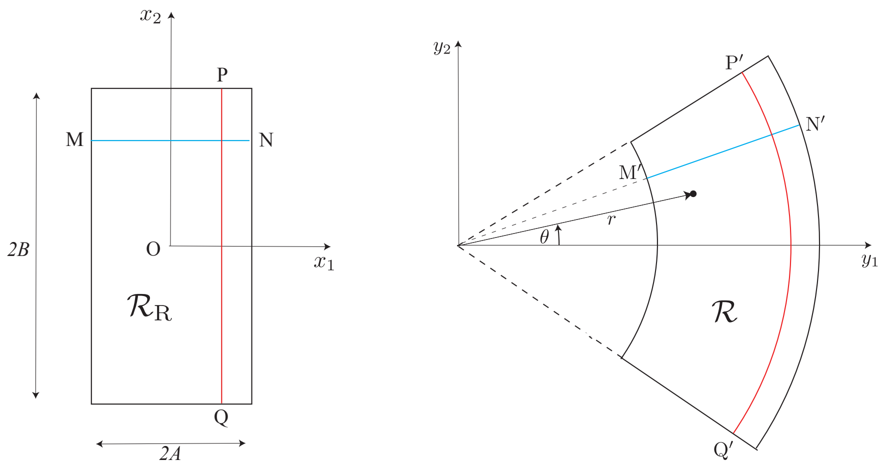


Figure 3.17: In a reference configuration the body occupies the rectangular parallelepiped region $\mathcal{R}_{\rm R} = \{(x_1, x_2, x_3) : -A \le x_1 \le A, -B \le x_2 \le B, -C \le x_3 \le C\}$ (left). The body undergoes a bending deformation in the x_1, x_2-plane . In the deformed configuration it occupies the region $\mathcal{R}$ (right).

vanishes, (c) the resultant moment about the origin of the traction distribution on $\mathcal{S}$ is $\mathbf{m} = m\mathbf{e}_3$, and (d) the body is in equilibrium with no body forces.

Remark: At the end of the solution, we will make some further simplifications to the answers obtained by using a constitutive relation (though we have not yet talked about constitutive relations!)

Solution:

(a) The outer curved boundary of $\mathcal{R}$ is the image of the flat boundary $x_1 = A$ of $\mathcal{R}_{\rm R}$. The Piola-Kirchhoff traction on this surface is given by $\mathbf{s} = \mathbf{S}\mathbf{n}_{\rm R}$ together with $(vi), \mathbf{n}_R = \mathbf{e}_1$ and $x_1 = A$:

$$\mathbf{s} = \mathbf{S}\mathbf{n}_R = \sigma_1 \mathbf{e}_r = \sigma_1(A, x_2)\mathbf{e}_r. \qquad (viii)$$

Similarly, the inner curved boundary of $\mathcal{R}$ corresponds to the flat surface $x_1 = -A$ of $\mathcal{R}_{\rm R}$. The Piola-Kirchhoff traction on this surface is given by $\mathbf{s} = \mathbf{S}\mathbf{n}_{\rm R}, (vi), \mathbf{n}_R = -\mathbf{e}_1$ and $x_1 = -A$:

$$\mathbf{s} = \mathbf{S}\mathbf{n}_R = -\sigma_1(-A, x_2)\mathbf{e}_r. \qquad (ix)$$

Observe that the traction $\mathbf{s}$ on these surfaces acts in the radial direction $\mathbf{e}_r$. Since $\mathbf{s}\, dA_x = \mathbf{t}\, dA_y$, so does the traction $\mathbf{t}$. Since these curved surfaces are traction-free, we must have

$$\sigma_1(\pm A, x_2) = 0 \qquad \text{for} \quad -B \le x_2 \le B. \qquad \Box \qquad (x)$$

(b) The top inclined flat boundary $\mathcal{S}$ of $\mathcal{R}$ is the image of the top horizontal surface $x_2 = B$ of $\mathcal{R}_{\rm R}$. The Piola-Kirchhoff traction on this surface can be calculated from $\mathbf{s} = \mathbf{S}\mathbf{n}_{\rm R}$ with $(vi), \mathbf{n}_R = \mathbf{e}_2$ and $x_2 = B$:

$$\mathbf{s} = \mathbf{S}\mathbf{n}_R = \sigma_2(x_1, B)\mathbf{e}_\theta. \qquad (xi)$$

Observe that the traction $\mathbf{s}$ on $x_2 = B$ acts in the circumferential direction $\mathbf{e}_\theta$, and therefore so does the traction $\mathbf{t}$ on the corresponding surface of $\mathcal{R}$.

Remark: The unit vector $\mathbf{e}_\theta = \mathbf{e}_\theta(\theta)$ depends in general on θ, and since $\theta = \theta(x_2)$ by (i), it depends on x_2: $\mathbf{e}_\theta = \mathbf{e}_\theta(\theta(x_2))$. However, it is constant on the surface $x_2 = B$ since the angle $\theta = \theta(B)$ is constant there.

Keeping in mind that $\mathcal{S}$ denotes the top inclined surface of $\mathcal{R}$ and $\mathcal{S}_\mathrm{R}$ is its image in the reference configuration, the resultant force on $\mathcal{S}$ is

$$
\begin{aligned}
&= \int_{\mathcal{S}} \mathbf{t}\, dA_y \overset{(3.53)}{=} \int_{\mathcal{S}_\mathrm{R}} \mathbf{s}\, dA_x = \int_{-A}^{A}\int_{-C}^{C} \mathbf{s}\, dx_3 dx_1 = 2C\int_{-A}^{A} \mathbf{s}\, dx_1 \overset{(xi)}{=} 2C\int_{-A}^{A} \sigma_2(x_1,B)\mathbf{e}_\theta\, dx_1 = \\
&= \left(2C\int_{-A}^{A} \sigma_2(x_1,B)\, dx_1\right)\ \mathbf{e}_\theta.
\end{aligned}
$$

Observe that we were able to take $\mathbf{e}_\theta$ out of the integral since $\mathbf{e}_\theta$ here is $\mathbf{e}_\theta(\theta(B))$ and so did not depend on x_1. Therefore, the resultant force on $\mathcal{S}$ vanishes when

$$
\int_{-A}^{A} \sigma_2(x_1,B)\, dx_1 = 0. \qquad \square \qquad (xii)
$$

We will be able to further simplify this boundary condition after deriving the equilibrium equations in part (d) below and using (x) and a certain constitutive relation.

(c) The position vector of a generic particle in $\mathcal{R}$ is $\mathbf{y} = r\mathbf{e}_r$. The resultant moment about O due to the traction distribution on $\mathcal{S}$ is therefore

$$
\begin{aligned}
\mathbf{m} &= \int_{\mathcal{S}} \mathbf{y}\times\mathbf{t}\, dA_y \overset{(3.53)}{=} \int_{\mathcal{S}_\mathrm{R}} \mathbf{y}\times\mathbf{s}\, dA_x \overset{(xi)}{=} \int_{\mathcal{S}_\mathrm{R}} \mathbf{y}\times\sigma_2\mathbf{e}_\theta\, dA_x = \int_{\mathcal{S}_\mathrm{R}} r\mathbf{e}_r\times\sigma_2\mathbf{e}_\theta\, dA_x = \\
&= \mathbf{e}_z\int_{\mathcal{S}_\mathrm{R}} r\sigma_2\, dA_x = \mathbf{e}_z\int_{-A}^{A} r\sigma_2\, 2C dx_1 = \left(2C\int_{-A}^{A} r(x_1)\sigma_2(x_1,B)\, dx_1\right)\mathbf{e}_z.
\end{aligned}
$$

We are told that to moment on this surface is $m\mathbf{e}_z$ and therefore we have

$$
m = 2C\int_{-A}^{A} r(x_1)\sigma_2(x_1,B)\, dx_1. \qquad \square \qquad (xiii)
$$

(d) Next we enforce the equilibrium equation $\mathrm{Div}\,\mathbf{S} = \mathbf{o}$. We will do this in rectangular cartesian coordinates as $\partial S_{ij}/\partial x_j = 0$, and for this we must first determine the components S_{ij} of $\mathbf{S}$ in the basis $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$. Substituting (v) into (vi) and simplifying leads to

$$
\mathbf{S} = \sigma_1\cos\theta(\mathbf{e}_1\otimes\mathbf{e}_1) + \sigma_1\sin\theta(\mathbf{e}_2\otimes\mathbf{e}_1) - \sigma_2\sin\theta(\mathbf{e}_1\otimes\mathbf{e}_2) + \sigma_2\cos\theta(\mathbf{e}_2\otimes\mathbf{e}_2) + \sigma_3(\mathbf{e}_3\otimes\mathbf{e}_3). \qquad (xiv)
$$

From this and $\mathbf{S} = S_{ij}\mathbf{e}_i\otimes\mathbf{e}_j$ we can read off the cartesian components of stress:

$$
\begin{gathered}
S_{11} = \sigma_1\cos\theta, \quad S_{12} = -\sigma_2\sin\theta, \quad S_{21} = \sigma_1\sin\theta, \quad S_{22} = \sigma_2\cos\theta, \quad S_{33} = \sigma_3, \\
S_{13} = S_{31} = S_{23} = S_{32} = 0.
\end{gathered} \qquad (xv)
$$

We now substitute the stress components (xv) into the equilibrium equations $\partial S_{ij}/\partial x_j = 0$ keeping in mind that $\theta = \theta(x_2)$, $\sigma_i = \sigma_i(x_1,x_2)$:

$$
\frac{\partial S_{11}}{\partial x_1} + \frac{\partial S_{12}}{\partial x_2} + \frac{\partial S_{13}}{\partial x_3} = 0 \qquad \Rightarrow \qquad \frac{\partial\sigma_1}{\partial x_1}\cos\theta - \frac{\partial\sigma_2}{\partial x_2}\sin\theta - \sigma_2\cos\theta\,\theta' = 0,
$$

$$\begin{aligned}
\frac{\partial S_{21}}{\partial x_1} + \frac{\partial S_{22}}{\partial x_2} + \frac{\partial S_{23}}{\partial x_3} = 0 \qquad &\Rightarrow \quad \frac{\partial \sigma_1}{\partial x_1}\sin\theta + \frac{\partial \sigma_2}{\partial x_2}\cos\theta - \sigma_2 \sin\theta\,\theta' = 0,\\
\frac{\partial S_{31}}{\partial x_1} + \frac{\partial S_{32}}{\partial x_2} + \frac{\partial S_{33}}{\partial x_3} = 0 \qquad &\Rightarrow \quad 0 = 0,
\end{aligned}$$

These can be combined and simplified as follows: Multiplying the first equation by $\sin\theta$, the second by $\cos\theta$ and adding, and similarly multiplying the first equation by $\cos\theta$, the second by $\sin\theta$ and subtracting leads to the following pair of partial differential equations to be obeyed by $\sigma_1(x_1, x_2), \sigma_2(x_1, x_2), \theta(x_2)$:

$$\left.\begin{aligned}
&\frac{\partial \sigma_1}{\partial x_1} - \sigma_2\theta' = 0,\\
&\frac{\partial \sigma_2}{\partial x_2} = 0,
\end{aligned}\right\} \qquad \text{for} - A \le x_1 \le A, \quad -B \le x_2 \le B. \qquad \square \tag{xvi}$$

Remark: We now simplify the preceding results further assuming a specific form of the constitutive relation. Suppose that the constitutive relation tells us that each stress component σ_i is a function of the three principal stretches:

$$\sigma_i = \widehat{\sigma}_i(\lambda_1, \lambda_2, \lambda_3), \qquad i = 1, 2, 3.$$

The specific functions $\widehat{\sigma}_i$ here will depend on the material. We shall assume only that $\widehat{\sigma}_2$ *does* depend on λ_2, or said differently, σ_2 is not independent of λ_2:

$$\frac{\partial \widehat{\sigma}_2}{\partial \lambda_2} \neq 0.$$

Equation $(xvi)_2$ can now be simplified as follows using the constitutive relation $\sigma_2 = \widehat{\sigma}_2(\lambda_1, \lambda_2, \lambda_3)$:

$$\frac{\partial \sigma_2}{\partial x_2} = \frac{\partial \widehat{\sigma}_2}{\partial \lambda_1}\cancel{\frac{\partial \lambda_1}{\partial x_2}} + \frac{\partial \widehat{\sigma}_2}{\partial \lambda_2}\frac{\partial \lambda_2}{\partial x_2} + \frac{\partial \widehat{\sigma}_2}{\partial \lambda_3}\cancel{\frac{\partial \lambda_3}{\partial x_2}} = \frac{\partial \widehat{\sigma}_2}{\partial \lambda_2}\frac{\partial \lambda_2}{\partial x_2} = 0$$

where we have used the fact that according to (iii), λ_1 and λ_3 do not depend on x_2. Since we assumed $\partial\widehat{\sigma}_2/\partial\lambda_2 \neq 0$, this gives

$$\frac{\partial \lambda_2}{\partial x_2} = 0 \qquad \stackrel{(iii)}{\Rightarrow} \qquad r(x_1)\theta''(x_2) = 0 \quad \stackrel{(ii)_1}{\Rightarrow} \quad \theta''(x_2) = 0 \quad \Rightarrow \quad \theta(x_2) \stackrel{(ii)_4}{=} \frac{\beta x_2}{B}, \tag{xvii}$$

where β is a to-be-determined constant of integration with the geometric meaning: $\theta(\pm B) = \pm\beta$.

From $(xvii)$ and (iii) we see that all three stretches are independent of x_1, and so by the constitutive relation $\sigma_i = \widehat{\sigma}_i(\lambda_1, \lambda_2, \lambda_3)$, so are the stress components σ_i:

$$\sigma_i = \sigma_i(x_1). \tag{xviii}$$

The remaining equilibrium equation $(xvi)_1$ can now be written as

$$\sigma_1'(x_1) - \frac{\beta}{B}\sigma_2(x_1) = 0. \qquad \square \tag{xvi}$$

(b) (continued) The boundary condition (xii) on the top inclined surface $\mathcal{S}$ can be now be shown to be automatic since

$$\int_{-A}^{A} \sigma_2(x_1)\,dx_1 \stackrel{(xvi)}{=} \frac{1}{\alpha}\int_{-A}^{A} \sigma_1'\,dx_1 = \frac{1}{\alpha}\Big(\sigma_1(A) - \sigma_1(-A)\Big) \stackrel{(x)}{=} 0.$$

Please revisit the last part of this solution once we have discussed constitutive relations.

3.8 Rate of working. Stress power.

We now derive a relation between the rate of external working on a part of the body and the rate of internal working within that part. This analysis, like everything else so far, is independent of the constitutive relation and is valid for *all* materials. It should also be noted that the relation to be derived here is *not* the first law of thermodynamics: it is a relation that is entirely mechanical in character and relates internal and external working.

Since we want to calculate the *rate* of working, we have to consider particle velocity, and for this we must admit time t into our analysis. Accordingly we now consider a time-dependent quasi-static motion – a family of deformations[10] $\mathbf{y}(\mathbf{x},t)$ with time t being a parameter. By saying the motion is quasi-static we mean that the *equilibrium* equations hold at each instant t, inertial effects being omitted.

The velocity of a particle $\mathbf{x}$ is the rate of change of the position of that particle with respect to time:

$$\mathbf{v}(\mathbf{x},t) = \frac{\partial \mathbf{y}}{\partial t}(\mathbf{x},t). \tag{3.68}$$

Since $F_{ij} = \partial y_i/\partial x_j$ we can write

$$\dot{F}_{ij} := \frac{\partial}{\partial t}F_{ij}(\mathbf{x},t) = \frac{\partial}{\partial t}\frac{\partial y_i}{\partial x_j}(\mathbf{x},t) = \frac{\partial}{\partial x_j}\frac{\partial y_i}{\partial t}(\mathbf{x},t) = \frac{\partial v_i}{\partial x_j}(\mathbf{x},t) \quad \Leftrightarrow \quad \dot{\mathbf{F}} = \mathrm{Grad}\,\mathbf{v}, \tag{3.69}$$

where $\dot{\mathbf{F}}$ is the time rate of change of $\mathbf{F}(\mathbf{x},t)$ at a fixed particle $\mathbf{x}$ and Grad $\mathbf{v}$ is the 2-tensor with cartesian components $\partial v_i/\partial x_j$.

Consider a part of the body that occupies a region $\mathcal{D}_t$ at time t. Let $p(\mathcal{D}_t)$ denote the *rate at which the external forces on $\mathcal{D}_t$ do work*:

$$p(\mathcal{D}_t) = \int_{\partial\mathcal{D}_t} \mathbf{t}\cdot\mathbf{v}\,\mathrm{d}A_y + \int_{\mathcal{D}_t} \mathbf{b}\cdot\mathbf{v}\,\mathrm{d}V_y\ , \tag{3.70}$$

see (3.3). By using (3.46) and (3.53), we can express $p(\mathcal{D}_t)$ in referential form as

$$p(\mathcal{D}_t) = \int_{\partial\mathcal{D}_\mathrm{R}} \mathbf{s}\cdot\mathbf{v}\,\mathrm{d}A_x + \int_{\mathcal{D}_\mathrm{R}} \mathbf{b}_\mathrm{R}\cdot\mathbf{v}\,\mathrm{d}V_x\ , \tag{3.71}$$

where $\mathcal{D}_\mathrm{R}$ is the region occupied by the part being considered in the reference configuration. It is now convenient to work in terms of components (in some fixed orthonormal basis). Then

[10] a one-parameter family of deformations,

we have

$$
\begin{aligned}
p(\mathcal{D}_t) \;&=\; \int_{\partial\mathcal{D}_{\rm R}} s_i v_i \,\mathrm{d}A_x + \int_{\mathcal{D}_{\rm R}} b_i^R v_i \,\mathrm{d}V_x \overset{(3.52)}{=} \int_{\partial\mathcal{D}_{\rm R}} S_{ij} n_j^R v_i \,\mathrm{d}A_x + \int_{\mathcal{D}_{\rm R}} b_i^R v_i \,\mathrm{d}V_x = \\
&\overset{(1.175)}{=} \int_{\mathcal{D}_{\rm R}} \frac{\partial}{\partial x_j}\Big(S_{ij} v_i\Big) \,\mathrm{d}V_x + \int_{\mathcal{D}_{\rm R}} b_i^R v_i \,\mathrm{d}V_x = \\
&= \int_{\mathcal{D}_{\rm R}} \left[\frac{\partial S_{ij}}{\partial x_j} v_i + S_{ij}\frac{\partial v_i}{\partial x_j} \;+\; b_i^R v_i\right] \mathrm{d}V_x \;=\; \int_{\mathcal{D}_{\rm R}} \left[\left(\frac{\partial S_{ij}}{\partial x_j} + b_i^R\right) v_i + S_{ij}\frac{\partial v_i}{\partial x_j}\right] \mathrm{d}V_x \\
&\overset{(3.65)_1}{=} \int_{\mathcal{D}_{\rm R}} S_{ij}\frac{\partial v_i}{\partial x_j} \mathrm{d}V_x \overset{(3.69)}{=} \int_{\mathcal{D}_{\rm R}} S_{ij}\dot{F}_{ij} \mathrm{d}V_x = \int_{\mathcal{D}_{\rm R}} \mathbf{S}\cdot\dot{\mathbf{F}} \,\mathrm{d}V_x .
\end{aligned}
\tag{3.72}
$$

Thus from (3.71) and (3.72) we have the following rate of working identity:

$$
\int_{\partial\mathcal{D}_{\rm R}} \mathbf{s}\cdot\mathbf{v}\,\mathrm{d}A_x + \int_{\mathcal{D}_{\rm R}} \mathbf{b}_{\rm R}\cdot\mathbf{v}\,\mathrm{d}V_x = \int_{\mathcal{D}_{\rm R}} \mathbf{S}\cdot\dot{\mathbf{F}}\,\mathrm{d}V_x . \tag{3.73}
$$

Equation (3.73) states that the rate of external work on a part of the body (the left-hand side) equals the rate of internal work within that part (the right-hand side). The rate of working by the internal stresses per unit reference volume, i.e. $\mathbf{S}\cdot\dot{\mathbf{F}}$, is called the **stress power**.

$$
\boxed{\text{Stress power } = \mathbf{S}\cdot\dot{\mathbf{F}}.} \tag{3.74}
$$

In general, the stress power accounts for both stored and dissipated energy. The integral involving the stress power on the right-hand side of (3.73) *cannot* in general be written as the time derivative of the volume integral of some scalar field.

Problem 3.8.1. Establish the following spatial form of the rate of working identity:

$$
\int_{\partial\mathcal{D}} \mathbf{t}\cdot\mathbf{v}\,\mathrm{d}A_y + \int_{\mathcal{D}} \mathbf{b}\cdot\mathbf{v}\,\mathrm{d}V_y = \int_{\mathcal{D}} \mathbf{T}\cdot\mathbf{D}\,\mathrm{d}V_y \;, \tag{3.75}
$$

where $\mathbf{D}$ is defined by

$$
\mathbf{D} \coloneqq \frac{1}{2}(\mathbf{L}+\mathbf{L}^T) \qquad \text{where} \quad \mathbf{L} \coloneqq \operatorname{grad}\mathbf{v}. \tag{3.76}
$$

In cartesian components

$$
D_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial y_j} + \frac{\partial v_j}{\partial y_i}\right), \qquad L_{ij} = \frac{\partial v_i}{\partial y_j}. \tag{3.77}
$$

The (kinematic) tensors $\mathbf{D}$ and $\mathbf{L}$ are known as the *stretching tensor* and the *velocity gradient tensor* respectively. Note that the gradient here is with respect to the spatial position $\mathbf{y}$ and it is understood that the velocity field has been expressed in spatial form as $\mathbf{v}(\mathbf{y}, t)$; see Section 2.8. It follows from the right-hand side of (3.75) that the stress power, the rate of internal working per unit *reference* volume, can be written as

$$\boxed{\text{Stress power } = J\,\mathbf{T} \cdot \mathbf{D}.} \tag{3.78}$$

Solution: Since the relation between $\mathbf{x}$ and $\mathbf{y}$ is one-to-one, the relation $\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{x}, t)$ can be inverted at each instant t to give $\mathbf{x} = \overline{\mathbf{x}}(\mathbf{y}, t)$. Thus the referential and spatial descriptions of the velocity field, $\widehat{\mathbf{v}}(\mathbf{x}, t)$ and $\overline{\mathbf{v}}(\mathbf{y}, t)$, are related by

$$\overline{\mathbf{v}}(\mathbf{y}, t) = \widehat{\mathbf{v}}(\overline{\mathbf{x}}(\mathbf{y}, t), t), \qquad \widehat{\mathbf{v}}(\mathbf{x}, t) = \overline{\mathbf{v}}(\widehat{\mathbf{y}}(\mathbf{x}, t), t).$$

Recall from (3.69) the relation $\dot{F}_{ij} = \frac{\partial \widehat{v}_i}{\partial x_j}$. and so by using the chain rule

$$\dot{F}_{ij} = \frac{\partial \widehat{v}_i}{\partial x_j} = \frac{\partial \overline{v}_i}{\partial y_k}\frac{\partial y_k}{\partial x_j} \overset{(3.77)_1}{=} L_{ik}F_{kj} \qquad \Leftrightarrow \qquad \dot{\mathbf{F}} = \mathbf{L}\mathbf{F}. \tag{3.79}$$

This is a relation between the time rate of change of the deformation gradient tensor (at a fixed particle $\mathbf{x}$) and the velocity gradient tensor.

It now follows that

$$\mathbf{S} \cdot \dot{\mathbf{F}} \overset{(3.79)}{=} \mathbf{S} \cdot \mathbf{L}\mathbf{F} \overset{(3.51)}{=} J\,\mathbf{T}\,\mathbf{F}^{-T} \cdot \mathbf{L}\mathbf{F} \overset{(1.120)}{=} J\mathbf{T} \cdot \mathbf{L}.$$

However,

$$\mathbf{T} \cdot \mathbf{L} = \mathbf{T} \cdot \left[\frac{1}{2}(\mathbf{L} + \mathbf{L}^T) + \frac{1}{2}(\mathbf{L} - \mathbf{L}^T)\right] = \mathbf{T} \cdot \mathbf{D} + \frac{1}{2}\mathbf{T} \cdot (\mathbf{L} - \mathbf{L}^T) = \mathbf{T} \cdot \mathbf{D},$$

where in getting to the last equality we used the result from (1.137) since $\mathbf{T}$ is symmetric and $\frac{1}{2}(\mathbf{L} - \mathbf{L}^T)$ is skew symmetric. Thus from the two preceding equations we have

$$\mathbf{S} \cdot \dot{\mathbf{F}} = J\mathbf{T} \cdot \mathbf{D}.$$

Using this in (3.73), together with $J\,dV_x = dV_y$, (3.70) and (3.71) yields (3.75)

3.8.1 Work Conjugate Stress-Strain Pairs.

Consider a body undergoing an arbitrary quasi-static motion. Suppose that the stress power $\mathbf{S} \cdot \dot{\mathbf{F}}$ can be expressed in the form $\mathbf{A} \cdot \dot{\mathbf{B}}$ where $\mathbf{B}$ is a strain measure[11] (in the sense of Section 2.6); the components of $\mathbf{A}$ will then have the dimension of stress. We say that the stress $\mathbf{A}$ and the strain $\mathbf{B}$ are conjugate[12]. This conjugacy reflects a special relationship between the stress $\mathbf{A}$ and strain $\mathbf{B}$. As we shall see when studying the constitutive behavior of an elastic material, the constitutive relation for the stress $\mathbf{A}$ is most naturally written in terms of the strain $\mathbf{B}$.

[11] Here $\mathbf{B}$ is not the left Cauchy-Green tensor.

[12] Note that $\mathbf{F}$ is not a strain. Thus one usually does not refer to the pair $\mathbf{S}, \dot{\mathbf{F}}$ as being work conjugate.

For example, consider the family of Lagrangian strain tensors

$$\mathbf{E}^{(0)} = \ln \mathbf{U}, \qquad\qquad \mathbf{E}^{(n)} = \frac{1}{n}\Big(\mathbf{U}^n - \mathbf{I}\Big), \quad n \neq 0.$$

Can one find a corresponding family of stress tensors $\mathbf{S}^{(n)}$ such that the

$$\text{stress power } = \mathbf{S}^{(n)} \cdot \dot{\mathbf{E}}^{(n)} \ ?$$

Consider the case $n = 2$, i.e. the Green Saint-Venant strain tensor $\mathbf{E}^{(2)}$. We want to find a tensor $\mathbf{S}^{(2)}$ such that the

$$\text{Stress power} = \mathbf{S} \cdot \dot{\mathbf{F}} = \mathbf{S}^{(2)} \cdot \dot{\mathbf{E}}^{(2)}. \tag{i}$$

Since $\mathbf{E}^{(2)}$ is symmetric, there is no loss of generality in assuming $\mathbf{S}^{(2)}$ to be symmetric. (Why?) Differentiating $\mathbf{E}^{(2)} = \frac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I})$ with respect to t gives

$$\dot{\mathbf{E}}^{(2)} = \frac{1}{2}\dot{\mathbf{F}}^T\mathbf{F} + \frac{1}{2}\mathbf{F}^T\dot{\mathbf{F}}. \tag{ii}$$

Now substitute (ii) into (i) and simplify:

$$\begin{aligned}\mathbf{S} \cdot \dot{\mathbf{F}} &= \mathbf{S}^{(2)} \cdot \dot{\mathbf{E}}^{(2)} = \tfrac{1}{2}\mathbf{S}^{(2)} \cdot \dot{\mathbf{F}}^T\mathbf{F} + \tfrac{1}{2}\mathbf{S}^{(2)} \cdot \mathbf{F}^T\dot{\mathbf{F}} \overset{(1.120)}{=} \tfrac{1}{2}\mathbf{S}^{(2)}\mathbf{F}^T \cdot \dot{\mathbf{F}}^T + \tfrac{1}{2}\mathbf{F}\mathbf{S}^{(2)} \cdot \dot{\mathbf{F}} = \\ &= \tfrac{1}{2}\mathbf{F}(\mathbf{S}^{(2)})^T \cdot \dot{\mathbf{F}} + \tfrac{1}{2}\mathbf{F}\mathbf{S}^{(2)} \cdot \dot{\mathbf{F}} = \mathbf{F}\mathbf{S}^{(2)} \cdot \dot{\mathbf{F}},\end{aligned}$$

where in getting to the second line we used $\mathbf{A} \cdot \mathbf{B} = \mathbf{A}^T \cdot \mathbf{B}^T$ and in the last step we used the symmetry of $\mathbf{S}^{(2)}$. Thus $\mathbf{S} \cdot \dot{\mathbf{F}} = \mathbf{F}\mathbf{S}^{(2)} \cdot \dot{\mathbf{F}}$ and so

$$\mathbf{S}^{(2)} = \mathbf{F}^{-1}\mathbf{S}. \tag{3.80}$$

The symmetric tensor $\mathbf{S}^{(2)}$ is known as the second Piola-Kirchhoff stress tensor. It is conjugate to the Green Saint-Venant strain tensor.

The case of general n is discussed in Chapter 3.5 of Ogden [5].

Exercises: Problems 3.29, 3.30, 3.31 and 3.32.

3.8.2 Some other stress tensors.

In addition to the Cauchy and 1st Piola-Kirchhoff stress tensors, various other stress measures are used in the literature. Some examples are

$$
\begin{array}{ll}
\mathbf{T} & \text{Cauchy stress tensor,} \\
\mathrm{J}\mathbf{T} & \text{Kirchhoff stress tensor,} \\
\mathbf{S} = \mathrm{J}\mathbf{T}\mathbf{F}^{-T} & 1^{\text{st}}\ \text{Piola} - \text{Kirchhoff stress tensor,} \\
\mathbf{S}^{(2)} = \mathrm{J}\mathbf{F}^{-1}\mathbf{T}\mathbf{F}^{-T} = \mathbf{F}^{-1}\mathbf{S} & 2^{\text{nd}}\ \text{Piola} - \text{Kirchhoff stress tensor,} \\
\mathbf{S}^{T} = \mathrm{J}\mathbf{F}^{-1}\mathbf{T} & \text{Nominal stress tensor,} \\
\mathbf{S}^{(1)} = \frac{1}{2}\left(\mathbf{S}^{T}\mathbf{R} + \mathbf{R}^{T}\mathbf{S}\right) & \text{Biot stress tensor.}
\end{array}
\tag{3.81}
$$

Even though many of these stress tensors have no simple physical significance, they are sometimes useful in, say, carrying out computations.

Exercise: Show that $\mathbf{S}^{(1)} \cdot \mathbf{U} = \mathbf{S} \cdot \mathbf{F}$. Note that this equation involves $\mathbf{U}$ and $\mathbf{F}$ *not* $\dot{\mathbf{U}}$ and $\dot{\mathbf{F}}$.

3.9 Linearization.

We now specialize the preceding analyses to the case where the deformed configuration is close to the reference configuration in the sense that the gradient of the displacement from the former to the latter is small, $|\mathbf{H}| = |\nabla\, \mathbf{u}| \ll 1$. It is natural therefore to work with the formulation with respect to the (fixed) reference configuration:

$$
\mathrm{Div}\, \mathbf{S} + \mathbf{b}_{\mathrm{R}} = \mathbf{0}, \qquad \mathbf{S}\mathbf{F}^{T} = \mathbf{F}\mathbf{S}^{T} \qquad \text{for all } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}. \tag{3.82}
$$

Since $\mathbf{F} = \mathbf{I} + \mathbf{H}$ we see immediately that to leading order the moment equilibrium equation $(3.82)_2$ reduces to $\mathbf{S} = \mathbf{S}^T$ and so the first Piola-Kirchhoff stress tensor is symmetric to leading order at an infinitesimal deformation. In fact, by using $\mathbf{F} = \mathbf{I} + \mathbf{H}$ and $J = \det(\mathbf{I} + \mathbf{H}) = 1 + \mathrm{tr}\,\mathbf{H} + O(|\mathbf{H}|^2)$, it follows from $\mathbf{T} = J^{-1}\mathbf{S}\mathbf{F}^T$ that, to leading order,

$$
\mathbf{S} = \mathbf{T} \; + \mathrm{O}(|\mathbf{H}|). \tag{3.83}
$$

Thus to leading order, the 1st Piola-Kirchhoff stress tensor and the Cauchy stress tensor do not differ in infinitesimal deformations. Similarly the body force density $\mathbf{b} = \mathbf{b}_{\mathrm{R}} + \mathrm{O}(|\mathbf{H}|)$.

For clarity we shall use the symbol $\boldsymbol{\sigma}$ for the stress tensor in the linearized theory. The stress component σ_{ij} is the i^{th} component of force per unit area on the surface normal to $\mathbf{e}_j$ where we do not need to distinguish between the deformed and reference configurations.

The (force) equilibrium equation now reads

$$\operatorname{Div} \boldsymbol{\sigma} + \mathbf{b}_{\mathrm{R}} = \mathbf{0} \qquad \text{for } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}, \tag{3.84}$$

and the moment equilibrium equation tells us that

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^T \qquad \text{for } \mathbf{x} \in \mathcal{R}_{\mathrm{R}}. \tag{3.85}$$

Note that these field equations hold on the region $\mathcal{R}_{\mathrm{R}}$ occupied in the reference configuration. Similarly the traction-stress relation is

$$\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}_{\mathrm{R}}. \tag{3.86}$$

Thus in conclusion, for infinitesimal deformations we will work with the stress tensor $\boldsymbol{\sigma}$ and do not need to consider the deformed configuration in formulating any of the fundamental principles for stress. Reviewing the preceding material in this chapter, we see that, for example, we can interpret the stress components σ_{ij} as in Figure 3.8 with T_{ij} replaced by σ_{ij} and we do not need to address whether the planes shown are in the reference or deformed configurations. Similarly in Problem 3.3.2 we can take the prismatic region there to be the region the body occupies in the reference configuration.

3.10 Some other coordinate systems.

3.10.1 Cylindrical polar coordinates.

In this section we derive expressions for the equilibrium equation $\operatorname{div} \mathbf{T} + \mathbf{b} = \mathbf{o}$ in terms of cylindrical polar coordinates:

– Let (y_1, y_2, y_3) denote the rectangular cartesian coordinates of a particle in the deformed configuration and let (r, θ, z) be its corresponding cylindrical polar coordinates. Then

$$y_1 = r\cos\theta, \qquad y_2 = r\sin\theta, \qquad y_3 = z, \tag{3.87}$$

and the associated basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ are related by

$$\mathbf{e}_r = \mathbf{e}_1\cos\theta + \mathbf{e}_2\sin\theta\ , \qquad \mathbf{e}_\theta = -\mathbf{e}_1\sin\theta + \mathbf{e}_2\cos\theta\ , \qquad \mathbf{e}_z = \mathbf{e}_3. \tag{3.88}$$

– The stress tensor $\mathbf{T}(\mathbf{y})$ can be written in terms of its cylindrical polar components as

$$\begin{aligned}\mathbf{T} &= T_{rr}\mathbf{e}_r \otimes \mathbf{e}_r + T_{r\theta}\mathbf{e}_r \otimes \mathbf{e}_\theta + T_{rz}\mathbf{e}_r \otimes \mathbf{e}_z + \\ &+ T_{\theta r}\mathbf{e}_\theta \otimes \mathbf{e}_r + T_{\theta\theta}\mathbf{e}_\theta \otimes \mathbf{e}_\theta + T_{\theta z}\mathbf{e}_\theta \otimes \mathbf{e}_z + \\ &+ T_{zr}\mathbf{e}_z \otimes \mathbf{e}_r + T_{z\theta}\mathbf{e}_z \otimes \mathbf{e}_\theta + T_{zz}\mathbf{e}_z \otimes \mathbf{e}_z.\end{aligned} \tag{3.89}$$

– Keeping in mind that $\operatorname{div}\mathbf{T}$ is a vector, we want to calculate its components in the $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ basis. These components are defined, just as would be the components of any vector, as

$$\mathbf{e}_r \cdot \operatorname{div}\mathbf{T}, \qquad \mathbf{e}_\theta \cdot \operatorname{div}\mathbf{T}, \qquad \mathbf{e}_z \cdot \operatorname{div}\mathbf{T}, \tag{3.90}$$

and so the vector itself can be expressed in terms of these components and basis vectors as

$$\operatorname{div}\mathbf{T} = \big(\mathbf{e}_r \cdot \operatorname{div}\mathbf{T}\big)\mathbf{e}_r + \big(\mathbf{e}_\theta \cdot \operatorname{div}\mathbf{T}\big)\mathbf{e}_\theta + \big(\mathbf{e}_z \cdot \operatorname{div}\mathbf{T}\big)\mathbf{e}_z. \tag{3.91}$$

Our goal therefore is to calculate the three terms in (3.90). We shall do this by using the vector identity established in Problem 1.8.1, viz.

$$\mathbf{v} \cdot \operatorname{div}\mathbf{T} = \operatorname{div}(\mathbf{T}^T\mathbf{v}) - \mathbf{T} \cdot \operatorname{grad}\mathbf{v}, \tag{3.92}$$

that holds for any vector field $\mathbf{v}(\mathbf{y})$ and tensor field $\mathbf{T}(\mathbf{y})$. Observe that the right-hand side of (3.92) involves the divergence and gradient of two *vector* fields, and we previously calculated expressions for these (in cylindrical polar coordinates) in Section 1.8.6; see equations (1.188) and (1.187).

First take $\mathbf{v} = \mathbf{e}_r$ in the identity (3.92) and calculate the two terms on its right-hand side. By taking $\mathbf{u} = \mathbf{e}_r$ in (1.187) (together with the obvious change of notation from (R, Θ, Z) to (r, θ, z)) we get

$$\operatorname{grad}\mathbf{e}_r = \frac{1}{r}\,\mathbf{e}_\theta \otimes \mathbf{e}_\theta \qquad \stackrel{(3.89)}{\Rightarrow} \qquad \mathbf{T} \cdot \operatorname{grad}\mathbf{e}_r = \frac{T_{\theta\theta}}{r}, \tag{i}$$

which is the second term on the right-hand side of (3.92). The first term can be evaluated as follows:

$$\operatorname{div}\mathbf{T}^T\mathbf{e}_r \stackrel{(3.89)}{=} \operatorname{div}\left(T_{rr}\mathbf{e}_r + T_{r\theta}\mathbf{e}_\theta + T_{rz}\mathbf{e}_z\right) \stackrel{(1.188)}{=} \frac{\partial T_{rr}}{\partial r} + \frac{1}{r}T_{rr} + \frac{1}{r}\frac{\partial T_{r\theta}}{\partial\theta} + \frac{\partial T_{rz}}{\partial z}. \tag{ii}$$

Substituting (i) and (ii) into (3.92) yields

$$\mathbf{e}_r \cdot \operatorname{div}\mathbf{T} = \operatorname{div}\mathbf{T}^T\mathbf{e}_r - \mathbf{T} \cdot \operatorname{grad}\mathbf{e}_r = \frac{\partial T_{rr}}{\partial r} + \frac{1}{r}\frac{\partial T_{r\theta}}{\partial\theta} + \frac{\partial T_{rz}}{\partial z} + \frac{T_{rr} - T_{\theta\theta}}{r}. \tag{iii}$$

A parallel calculation with $\mathbf{v} = \mathbf{e}_\theta$ yields

$$\operatorname{grad}\mathbf{e}_\theta = -\frac{1}{r}\,\mathbf{e}_r \otimes \mathbf{e}_\theta \qquad \overset{(3.89)}{\Rightarrow} \qquad \mathbf{T}\cdot\operatorname{grad}\mathbf{e}_\theta = -\frac{T_{r\theta}}{r}, \tag{iv}$$

$$\operatorname{div}\mathbf{T}^T\mathbf{e}_\theta \overset{(3.89)}{=} \operatorname{div}\left(T_{\theta r}\mathbf{e}_r + T_{\theta\theta}\mathbf{e}_\theta + T_{\theta z}\mathbf{e}_z\right) \overset{(1.188)}{=} \frac{\partial T_{\theta r}}{\partial r} + \frac{1}{r}T_{\theta r} + \frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial\theta} + \frac{\partial T_{\theta z}}{\partial z}, \tag{v}$$

$$\mathbf{e}_\theta\cdot\operatorname{div}\mathbf{T} = \operatorname{div}\mathbf{T}^T\mathbf{e}_\theta - \mathbf{T}\cdot\operatorname{grad}\mathbf{e}_\theta = \frac{\partial T_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial\theta} + \frac{\partial T_{\theta z}}{\partial z} + \frac{T_{r\theta}+T_{\theta r}}{r}. \tag{vi}$$

And $\mathbf{v} = \mathbf{e}_z$ leads to

$$\operatorname{grad}\mathbf{e}_z = \mathbf{o} \qquad \overset{(3.89)}{\Rightarrow} \qquad \mathbf{T}\cdot\operatorname{grad}\mathbf{e}_\theta = 0, \tag{vii}$$

$$\operatorname{div}\mathbf{T}^T\mathbf{e}_z \overset{(3.89)}{=} \operatorname{div}\left(T_{zr}\mathbf{e}_r + T_{z\theta}\mathbf{e}_\theta + T_{zz}\mathbf{e}_z\right) \overset{(1.188)}{=} \frac{\partial T_{zr}}{\partial r} + \frac{1}{r}T_{zr} + \frac{1}{r}\frac{\partial T_{z\theta}}{\partial\theta} + \frac{\partial T_{zz}}{\partial z}, \tag{$viii$}$$

$$\mathbf{e}_z\cdot\operatorname{div}\mathbf{T} = \operatorname{div}\mathbf{T}^T\mathbf{e}_z - \mathbf{T}\cdot\operatorname{grad}\mathbf{e}_z = \frac{\partial T_{zr}}{\partial r} + \frac{1}{r}\frac{\partial T_{z\theta}}{\partial\theta} + \frac{\partial T_{zz}}{\partial z} + \frac{T_{zr}}{r}. \tag{ix}$$

Finally we substitute $(iii), (vi)$ and (ix) into (3.91) to get

$$\begin{aligned}
\operatorname{div}\mathbf{T} \;&= \left(\frac{\partial T_{rr}}{\partial r} + \frac{1}{r}\frac{\partial T_{r\theta}}{\partial\theta} + \frac{\partial T_{rz}}{\partial z} + \frac{T_{rr}-T_{\theta\theta}}{r}\right)\mathbf{e}_r + \\
&+ \left(\frac{\partial T_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial\theta} + \frac{\partial T_{\theta z}}{\partial z} + \frac{T_{r\theta}+T_{\theta r}}{r}\right)\mathbf{e}_\theta + \\
&+ \left(\frac{\partial T_{zr}}{\partial r} + \frac{1}{r}\frac{\partial T_{z\theta}}{\partial\theta} + \frac{\partial T_{zz}}{\partial z} + \frac{T_{zr}}{r}\right)\mathbf{e}_z, \qquad \square
\end{aligned} \tag{3.93}$$

which gives us the divergence of a tensor field $\mathbf{T}(\mathbf{y})$ in cylindrical polar coordinates.

– **Remark:** We have (deliberately) not used the symmetry of $\mathbf{T}$ in the preceding calculations and formulae. As a result we can appropriate (3.93) (with the appropriate change of notation) to evaluate $\operatorname{Div}\mathbf{S}$.

– The equilibrium equation $\operatorname{div}\mathbf{T} + \mathbf{b} = \mathbf{o}$ obeyed by the Cauchy stress tensor field $\mathbf{T}(\mathbf{y})$ can now be written in cylindrical polar coordinates as

$$\begin{aligned}
\frac{\partial T_{rr}}{\partial r} + \frac{1}{r}\frac{\partial T_{r\theta}}{\partial\theta} + \frac{\partial T_{rz}}{\partial z} + \frac{T_{rr}-T_{\theta\theta}}{r} + b_r &= 0, \\
\frac{\partial T_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial T_{\theta\theta}}{\partial\theta} + \frac{\partial T_{\theta z}}{\partial z} + \frac{T_{r\theta}+T_{\theta r}}{r} + b_\theta &= 0, \\
\frac{\partial T_{zr}}{\partial r} + \frac{1}{r}\frac{\partial T_{z\theta}}{\partial\theta} + \frac{\partial T_{zz}}{\partial z} + \frac{T_{zr}}{r} + b_z &= 0\,,
\end{aligned} \tag{3.94}$$

where $\mathbf{b} = b_r\mathbf{e}_r + b_\theta\mathbf{e}_\theta + b_z\mathbf{e}_z$.

Exercise: Problem 3.22, Problem 3.23.

3.10.2 Spherical polar coordinates.

– Let (y_1, y_2, y_3) denote the rectangular cartesian coordinates of a particle in the deformed configuration and let (r, θ, ϕ) be its spherical polar coordinates. Then

$$y_1 = r\sin\theta\cos\phi, \qquad y_2 = r\sin\theta\sin\phi, \qquad y_3 = r\cos\theta. \tag{3.95}$$

$$0 \le r < \infty, \quad 0 \le \theta \le 2\pi, \quad 0 \le \phi < \pi.$$

The associated basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ are related by

$$\left.\begin{array}{rcl} \mathbf{e}_r &=& (\sin\theta\cos\phi)\ \mathbf{e}_1 + (\sin\theta\sin\phi)\ \mathbf{e}_2 + \cos\theta\ \mathbf{e}_3, \\ \mathbf{e}_\theta &=& (\cos\theta\cos\phi)\ \mathbf{e}_1 + (\cos\theta\sin\phi)\ \mathbf{e}_2 - \sin\theta\ \mathbf{e}_3, \\ \mathbf{e}_\phi &=& -\sin\phi\ \mathbf{e}_1 + \cos\phi\ \mathbf{e}_2, \end{array}\right\} \tag{3.96}$$

– Let $T_{rr}, T_{r\theta}, T_{r\phi}, \ldots$ be the components of the Cauchy stress tensor $\mathbf{T}$ in the basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$:

$$\mathbf{T} = T_{rr}\mathbf{e}_r \otimes \mathbf{e}_r + T_{r\theta}\mathbf{e}_r \otimes \mathbf{e}_\theta + T_{r\phi}\mathbf{e}_r \otimes \mathbf{e}_\phi + T_{\theta r}\mathbf{e}_\theta \otimes \mathbf{e}_r + T_{\theta\theta}\mathbf{e}_\theta \otimes \mathbf{e}_\theta \ldots\ .$$

– The equilibrium equation $\operatorname{div}\mathbf{T} + \mathbf{b} = \mathbf{o}$ obeyed by the Cauchy stress tensor field $\mathbf{T}(\mathbf{y})$ can be shown to be

$$\begin{array}{rcl} \dfrac{\partial T_{rr}}{\partial r} + \dfrac{1}{r}\dfrac{\partial T_{r\phi}}{\partial \phi} + \dfrac{1}{r\sin\phi}\dfrac{\partial T_{r\theta}}{\partial\theta} + \dfrac{2T_{rr} - T_{\phi\phi} - T_{\theta\theta} + T_{r\phi}\cot\phi}{r} + b_r &=& 0, \\ \dfrac{\partial T_{r\theta}}{\partial r} + \dfrac{1}{r}\dfrac{\partial T_{\phi\theta}}{\partial \phi} + \dfrac{1}{r\sin\phi}\dfrac{\partial T_{\theta\theta}}{\partial\theta} + \dfrac{3T_{r\theta} + 2T_{\theta\phi}\cot\phi}{r} + b_\theta &=& 0, \\ \dfrac{\partial T_{r\phi}}{\partial r} + \dfrac{1}{r}\dfrac{\partial T_{\phi\phi}}{\partial \phi} + \dfrac{1}{r\sin\phi}\dfrac{\partial T_{\phi\theta}}{\partial\theta} + \dfrac{3T_{r\phi} + (T_{\phi\phi} - T_{\theta\theta})\cot\phi}{r} + b_\phi &=& 0. \end{array} \tag{3.97}$$

3.10.3 Worked examples

Problem 3.10.1. *(Combined axial and azimuthal shear of a tube)* An elastic body in a reference configuration occupies a hollow circular cylindrical region of unit length and inner and outer radii A and B respectively. Its outer surface $R = B$ is held fixed. A rigid solid cylinder of radius A is inserted into the cavity, and firmly bonded to the hollow elastic cylinder on their common interface $R = A$. A force $F\mathbf{e}_z$ in the axial direction and a torque $T\mathbf{e}_z$ about the axis are applied on the rigid cylinder. Assume that the resulting traction between the cylinders is uniformly distributed on their common interface. The resulting deformation involves axial and azimuthal shear, the kinematics of which were analyzed in Problem 2.15. The deformation was $r = R, \theta = \Theta + \phi(R), z = Z + w(R)$. Here we analyze the stress field.

In view of symmetry, assume that the Cauchy stress components in cylindrical coordinates (as given by a suitable isotropic constitutive relation) depend only on the r coordinate. (a) Simplify and solve the equilibrium equations to the extent possible. (b) Determine the boundary conditions on stress at $r = A$. (c) Using your answers from (a) and (b), determine the stress fields $T_{rz}(r)$ and $T_{r\theta}(r)$ in the elastic body.

Solution:

(a) We are told that the Cauchy stress components in cylindrical polar coordinates depend only of r (and not θ and z). The equilibrium equations (3.94) (in the absence of body force) then specialize to

$$\frac{dT_{rr}}{dr} + \frac{T_{rr} - T_{\theta\theta}}{r} = 0. \qquad \frac{dT_{r\theta}}{dr} + 2\frac{T_{r\theta}}{r} = 0, \qquad \frac{dT_{rz}}{dr} + \frac{T_{rz}}{r} = 0. \tag{i}$$

The second and third of these equations can be written as

$$\frac{d}{dr}(r^2 T_{r\theta}) = 0, \qquad \frac{d}{dr}(rT_{rz}) = 0, \tag{ii}$$

which can be integrated to obtain

$$T_{r\theta}(r) = \frac{c_1}{r^2}, \qquad T_{rz}(r) = \frac{c_2}{r}, \qquad A \le r \le B, \qquad \square \tag{iii}$$

where c_1 and c_2 are constants of integration (to be found using the boundary conditions).

(b) We now consider the equilibrium of the rigid cylinder. *We shall proceed vectorially but strongly encourage the reader to derive the results (x) and (xi) below using physical arguments.* Force balance requires

$$F\mathbf{e}_z + \int_{\mathcal{S}} \mathbf{t}\, dA_y = \mathbf{o}, \tag{iv}$$

where $\mathcal{S}$ is the interface $r = A$ between the cylinders. The traction (on the rigid cylinder) at this surface can be calculated using $\mathbf{t} = \mathbf{Tn}$, $\mathbf{n} = \mathbf{e}_r$ and $r = A$:

$$\mathbf{t} = T_{rr}(A)\mathbf{e}_r + T_{r\theta}(A)\mathbf{e}_\theta + T_{rz}(A)\mathbf{e}_z. \tag{v}$$

Substituting (v) into (iv) and using $dA_y = Ad\theta$ gives

$$F\mathbf{e}_z + \int_0^{2\pi} \Big[T_{rr}(A)\mathbf{e}_r + T_{r\theta}(A)\mathbf{e}_\theta + T_{rz}(A)\mathbf{e}_z\Big] Ad\theta = \mathbf{o}.$$

Keeping in mind that the unit vectors $\mathbf{e}_r$ and $\mathbf{e}_\theta$ depend on θ but $\mathbf{e}_z$ does not, we rewrite this as

$$F\mathbf{e}_z + AT_{rr}(A)\int_0^{2\pi} \mathbf{e}_r(\theta)d\theta + AT_{r\theta}(A)\int_0^{2\pi} \mathbf{e}_\theta(\theta)d\theta + AT_{rz}(A)\mathbf{e}_z\int_0^{2\pi} d\theta = \mathbf{o}.$$

Since

$$\mathbf{e}_r(\theta) = \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2, \qquad \mathbf{e}_\theta(\theta) = -\sin\theta\mathbf{e}_1 + \cos\theta\mathbf{e}_2,$$

the first and second integrals vanish and we are left with

$$F\mathbf{e}_z + AT_{rz}(A)\mathbf{e}_z \int_0^{2\pi} d\theta = \mathbf{o} \qquad \Rightarrow \qquad F\mathbf{e}_z + 2\pi A T_{rz}(A)\mathbf{e}_z = \mathbf{o}.$$

This leads to the boundary condition

$$T_{rz}(A) = -\frac{F}{2\pi A}. \qquad \square \tag{vi}$$

We next consider moment balance of the rigid cylinder which requires

$$T\mathbf{e}_z + \int_{\mathcal{S}} \mathbf{y} \times \mathbf{t}\, dA_y = \mathbf{o}. \tag{vii}$$

Since $\mathbf{y} = r\mathbf{e}_r = A\mathbf{e}_r$ at a point on $\mathcal{S}$, we have

$$\mathbf{y} \times \mathbf{t} = A\mathbf{e}_r \times \Big[T_{rr}(A)\mathbf{e}_r + T_{r\theta}(A)\mathbf{e}_\theta + T_{rz}(A)\mathbf{e}_z\Big] = AT_{r\theta}(A)\mathbf{e}_z - AT_{rz}(A)\mathbf{e}_\theta \quad \text{on } \mathcal{S}. \tag{viii}$$

Substituting $(viii)$ into (vii) and simplifying the integrals as above leads to

$$T\mathbf{e}_z + A^2 T_{r\theta}(A)\mathbf{e}_z \int_0^{2\pi} d\theta = \mathbf{o}$$

from which we obtain the boundary condition

$$T_{r\theta}(A) = -\frac{T}{2\pi A^2}. \qquad \square \tag{ix}$$

(c) On using the boundary condition (vi) in the stress field $(iii)_2$ we get $c_2 = -F/(2\pi)$ and so the shear stress field $T_{rz}(r)$ in the elastic body is

$$T_{rz}(r) = -\frac{F}{2\pi r}, \qquad A \le r \le B. \qquad \square \tag{x}$$

Similarly from (ix) and $(iii)_1$ we find

$$T_{r\theta}(r) = -\frac{T}{2\pi r^2}, \qquad A \le r \le B. \qquad \square \tag{xi}$$

3.11 Exercises.

Unless explicitly told otherwise neglect body forces and assume all components of vectors and tensors to be with respect to a fixed orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ that has been implicitly chosen.

Problem 3.1. Suppose that the region $\mathcal{R}$ occupied by a certain body in its deformed configuration is a prismatic cylinder of length L and equilateral triangular cross section of height $3a$ as shown in Figure 3.18. The coordinate axes $\{y_1, y_2, y_3\}$ are as shown in the figure.

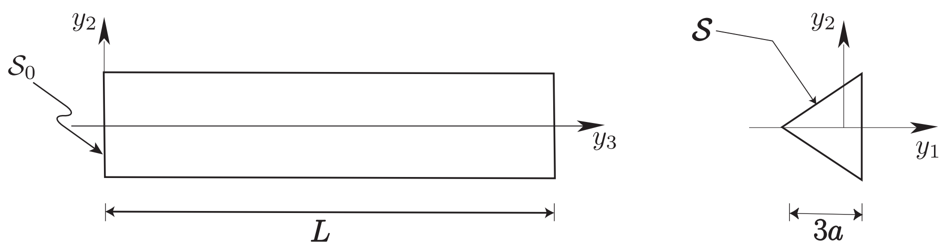


Figure 3.18: The region $\mathcal{R}$ occupied by the deformed body is a prismatic cylinder of length L and equilateral triangular cross section of height $3a$. $\mathcal{S}$ denotes one of its lateral surfaces.

The Cauchy stress field in the cylinder is known to be

$$[T(y_1, y_2, y_3)] = \begin{pmatrix} 0 & 0 & T_{13} \\ 0 & 0 & T_{23} \\ T_{31} & T_{32} & 0 \end{pmatrix} \qquad \text{where} \qquad \left. \begin{aligned} T_{13} &= T_{31} &&= K y_2 (y_1 - a), \\ T_{23} &= T_{32} &&= \tfrac{K}{2}\left(y_1^2 + 2a y_1 - y_2^2\right), \end{aligned} \right\} \qquad (i)$$

where the constant K is a given loading parameter.

Calculate

(a) the applied load distribution (traction) on the lateral surfaces, and

(b) the *resultant* force and moment on the end $y_3 = 0$?

Problem 3.2. The normal stress at a point on a surface perpendicular to the unit vector $\mathbf{n}$ is

$$T_{\text{normal}}(\mathbf{n}) = \mathbf{t}(\mathbf{n}) \cdot \mathbf{n}.$$

From among all planes through that point, on which is $T_{\text{normal}}(\mathbf{n})$ a maximum and what is its value on that plane?

Problem 3.3. The magnitude of the resultant shear stress at a point on a surface perpendicular to the unit vector $\mathbf{n}$ is

$$T_{\text{shear}}(\mathbf{n}) = \sqrt{\mathbf{t}(\mathbf{n})\cdot\mathbf{t}(\mathbf{n}) - (\mathbf{t}(\mathbf{n})\cdot\mathbf{n})^2}.$$

From among all planes through that point, on which is $T^2_{\text{shear}}(\mathbf{n})$ a maximum and what is its value on that plane?

Problem 3.4.

(a) For every stress tensor $\mathbf{T}$, is there a plane on which the magnitude of the resultant shear stress $T_{\text{shear}}(\mathbf{n})$ vanishes?

(b) For every stress tensor $\mathbf{T}$, is there a plane on which the normal stress $T_{\text{normal}}(\mathbf{n})$ vanishes?

(c) Suppose that the principal stresses τ_1, τ_2, τ_3 at some point in a body are all non-zero and $\tau_2 = \tau_3$. Find necessary and sufficient conditions on τ_1 and τ_2 under which there is a plane on which the normal stress vanishes.

Problem 3.5. Suppose the traction $\mathbf{t}(\mathbf{n})$ on every plane through a given point has the same direction $\mathbf{a}$, i.e. suppose that $\mathbf{t}(\mathbf{n}) = \alpha(\mathbf{n})\,\mathbf{a}$ for all unit vectors $\mathbf{n}$ where $\mathbf{a}$ is a constant unit vector. What is the form of the most general stress tensor $\mathbf{T}$ that is consistent with this?

Problem 3.6. Let $\mathbf{n}$ be a unit vector that is equally inclined to the principal axes of $\mathbf{T}$. The plane normal to $\mathbf{n}$ is known as the *octahedral plane*. Calculate the normal stress $T_{\text{normal}}(\mathbf{n})$ and the magnitude of the resultant shear stress $T_{\text{shear}}(\mathbf{n})$ on the octahedral plane in terms of the principal stress components τ_1, τ_2, τ_3.

Problem 3.7. In this problem you are to show that the Cauchy stress tensor at a point in a body is fully determined by the traction on any three linearly independent planes. Specifically: consider the three planes (through any one particular point in the body) normal to the respective unit vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$. Suppose the traction on each of these planes is $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3$ respectively.

(a) Write down (in terms of $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3$) the Cauchy stress tensor in the case where the unit vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ are mutually orthogonal.

(b) Write down the Cauchy stress tensor in the case where the unit vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ are linearly independent (but not necessarily mutually orthogonal).

(c) Can you write down the Cauchy stress tensor in the case where the unit vectors $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ are linearly dependent (with $\mathbf{n}_1, \mathbf{n}_2$ being linearly independent)? Explain.

Problem 3.8. (Chadwick) The Cauchy stress tensor (at a certain point in a body) is

$$\mathbf{T} = \alpha(\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3) + \beta(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_1 + \mathbf{e}_1 \otimes \mathbf{e}_3) \qquad (i)$$

where $\alpha \neq 0$ and $\beta \neq 0$ are constants.

(a) Calculate the principal stresses and corresponding principal directions.

(b) Calculate the maximum (over all $\mathbf{n}$) of the resultant shear stress magnitude $T_{\text{shear}}(\mathbf{n})$.

(c) Find necessary and sufficient conditions under which there is a plane on which the normal stress vanishes.

Problem 3.9. (Ogden) The magnitude of the resultant shear stress, $T_{\text{shear}}(\mathbf{n})$, on a plane perpendicular to the unit vector $\mathbf{n}$ was defined in (3.7) and an expression for it in a principal basis for $\mathbf{T}$ was given in equation (3.41). Show that that expression can be written in the alternative form

$$T^2_{\text{shear}} = (\tau_1 - \tau_2)^2 n_1^2 n_2^2 + (\tau_2 - \tau_3)^2 n_2^2 n_3^2 + (\tau_3 - \tau_1)^2 n_3^2 n_1^2. \qquad (3.98)$$

Show that the average value of (3.98) over all possible directions $\mathbf{n}$ is

$$\frac{1}{15}\Big[(\tau_1 - \tau_2)^2 + (\tau_2 - \tau_3)^2 + (\tau_3 - \tau_1)^2\Big].$$

Problem 3.10. (Atkin and Fox, Ogden) The stress field in a body is known to be uniaxial in the direction $\mathbf{m}$ (but not necessarily uniform) i.e. it is known to have the form

$$\mathbf{T}(\mathbf{y}) = \sigma(\mathbf{y})\, \mathbf{m} \otimes \mathbf{m}, \qquad (i)$$

where $\mathbf{m}$ is a constant unit vector and $\sigma(\mathbf{y})$ is a scalar-valued function. The body is in equilibrium and there are no body forces.

(a) Show that the vector $\operatorname{grad} \sigma$ must be perpendicular to $\mathbf{m}$.

(b) Show that $\sigma(\mathbf{y})$ must be constant on any plane parallel to $\mathbf{m}$, i.e. if you pick a basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ with $\mathbf{e}_3 = \mathbf{m}$, then $\sigma(\mathbf{y}) = \sigma(y_1, y_2)$.

(c) Show that the traction $\mathbf{t}(\mathbf{n})$ on any plane is parallel to $\mathbf{m}$.

(d) Specialize the traction from part (c) to the cases where $\mathbf{n}$ is parallel to $\mathbf{m}$ and $\mathbf{n}$ is perpendicular to $\mathbf{m}$.

(e) Show that

$$T^2_{\text{shear}}(\mathbf{n}) = \sigma^2\big[1 - (\mathbf{m}\cdot\mathbf{n})^2\big]\,(\mathbf{m}\cdot\mathbf{n})^2.$$

Show that the maximum value of this over all directions $\mathbf{n}$ (with $\mathbf{m}$ fixed) is $\frac{1}{4}\sigma^2$.

Problem 3.11. Consider two bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ related in the usual way by $\mathbf{e}'_i = Q_{ij}\mathbf{e}_j$ where $[Q]$ is an orthogonal matrix. Recall from (3.14) that the components of $\mathbf{T}$ in the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are defined by

$$T_{ij} = \mathbf{t}(\mathbf{e}_j)\cdot\mathbf{e}_i = \mathbf{T}\mathbf{e}_j\cdot\mathbf{e}_i,$$

and therefore its components in the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ are, by definition,

$$T'_{ij} = \mathbf{t}(\mathbf{e}'_j)\cdot\mathbf{e}'_i = \mathbf{T}\mathbf{e}'_j\cdot\mathbf{e}'_i.$$

Verify that the matrices $[T]$ and $[T']$ are related by the basis transformation rule for a 2-tensor.

Problem 3.12. (Ogden) Let the *mean stress* in a body be defined as

$$\overline{\mathbf{T}} := \frac{1}{\mathsf{vol}}\int_{\mathcal{R}} \mathbf{T}(\mathbf{y})\; dV_y$$

where vol is the volume of the region $\mathcal{R}$ occupied by the body.

(a) Given that the body is in equilibrium, show that one can express $\overline{\mathbf{T}}$ in the alternative form

$$\overline{\mathbf{T}} = \frac{1}{2\,\mathsf{vol}}\int_{\mathcal{R}}(\mathbf{b}\otimes\mathbf{y} + \mathbf{y}\otimes\mathbf{b})\; dV_y + \frac{1}{2\,\mathsf{vol}}\int_{\partial\mathcal{R}}(\mathbf{t}\otimes\mathbf{y} + \mathbf{y}\otimes\mathbf{t})\; dA_y. \tag{3.99}$$

This shows the following important property of the mean stress: it is *fully* determined by the traction on the boundary of the body (and the prescribed body force field).

(b) Suppose the body force vanishes and the traction on the boundary $\partial\mathcal{R}$ is $\mathbf{t}(\mathbf{y},\mathbf{n}) = \alpha\mathbf{n}$ where α is a constant. Show that $\overline{\mathbf{T}} = \alpha\mathbf{I}$.

(c) Suppose the body force vanishes and the traction on the boundary $\partial\mathcal{R}$ is $\mathbf{t}(\mathbf{y},\mathbf{n}) = \beta(\mathbf{a}\cdot\mathbf{n})\mathbf{a}$ where the unit vector $\mathbf{a}$ and scalar β are constants. Show that $\overline{\mathbf{T}} = \beta\mathbf{a}\otimes\mathbf{a}$.

Problem 3.13. Define the average of the first Piola-Kirchhoff stress tensor field in a body by

$$\overline{\mathbf{S}} = \frac{1}{\mathsf{vol}}\int_{\mathcal{R}_{\mathrm{R}}} \mathbf{S}(\mathbf{x})\, dV_x. \tag{i}$$

Show that

$$\overline{\mathbf{S}}^{\,T} = \frac{1}{\mathsf{vol}}\left[\int_{\partial\mathcal{R}_{\mathrm{R}}} \mathbf{x}\otimes\mathbf{S}\mathbf{n}_{\mathrm{R}}\, dA_x + \int_{\mathcal{R}_{\mathrm{R}}} \mathbf{x}\otimes\mathbf{b}_{\mathrm{R}}\, dV_x\right] \tag{ii}$$

and therefore that the average first Piola-Kirchhoff stress tensor field in a body depends only on the traction $\mathbf{S}(\mathbf{x})\mathbf{n}_{\mathrm{R}}(\mathbf{x})$ on the boundary $\partial\mathcal{R}_{\mathrm{R}}$ and the body force field $\mathbf{b}_{\mathrm{R}}(\mathbf{x})$ in $\mathcal{R}_{\mathrm{R}}$. Show also that

$$\int_{\mathcal{R}_{\mathrm{R}}} \mathbf{F}\mathbf{S}^T\, dV_x = \int_{\partial\mathcal{R}_{\mathrm{R}}} \mathbf{y}(\mathbf{x})\otimes\mathbf{S}\mathbf{n}_{\mathrm{R}}\, dA_x + \int_{\mathcal{R}_{\mathrm{R}}} \mathbf{y}(\mathbf{x})\otimes\mathbf{b}_{\mathrm{R}}\, dV_x. \tag{iii}$$

Problem 3.14. (Spencer) The region $\mathcal{R}$ occupied by a certain body in its deformed configuration is a right circular cylinder of length l and radius a:

$$\mathcal{R} = \{(y_1, y_2, y_3) \mid y_1^2 + y_2^2 \le a^2, \;\; -l \le y_3 \le 0\}.$$

Suppose the matrix of components of the Cauchy stress tensor field in the cylinder is

$$[T(y_1, y_2, y_3)] = \begin{pmatrix} 0 & 0 & -\alpha y_2 \\ 0 & 0 & \alpha y_1 \\ -\alpha y_2 & \alpha y_1 & \beta + \gamma y_1 + \delta y_2 \end{pmatrix}, \qquad (i)$$

where α, β, γ and δ are constants. The components here have been taken with respect to an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ where $\mathbf{e}_3$ is aligned with the axis of the cylinder.

(a) Verify that this stress field satisfies all requirements for equilibrium in the interior of the body.

(b) Verify that the curved surface of the cylinder is traction-free.

(c) Calculate the traction on the end $y_3 = 0$. Hence calculate the resultant force and couple acting on the cylinder at the end $y_3 = 0$. Show that the parameters α, β, γ and δ describe, respectively, a couple twisting the cylinder about the y_3-axis, a force pulling on the cylinder in the y_3-direction, a couple bending the cylinder about the y_2-axis, and a couple bending the cylinder about the y_1-axis.

(d) Consider the special case $\gamma = \delta = 0$ where there is no bending. Calculate the principal components of stress at an arbitrary point in the body. Calculate the value of the largest normal stress in the cylinder.

(e) Given a circular cylinder that is subjected to some prescribed traction on its boundary leading to axial loading, twisting and bending, does it necessarily follow that the stress field in the body has to be the stress field in (i)?

Problem 3.15. In a reference configuration a body occupies the unit cube

$$\mathcal{R}_{\mathrm{R}} = \{(x_1, x_2, x_3) \,:\, 0 \le x_1 \le 1, \; 0 \le x_2 \le 1, \; 0 \le x_3 \le 1\},$$

and undergoes the deformation

$$y_1 = \lambda x_1 + k x_2, \qquad y_2 = \lambda^{-1} x_2, \qquad y_3 = x_3. \qquad (i)$$

where λ and k are known constants.

(a) Sketch the region occupied by the body in the deformed configuration noting the lengths of the edges.

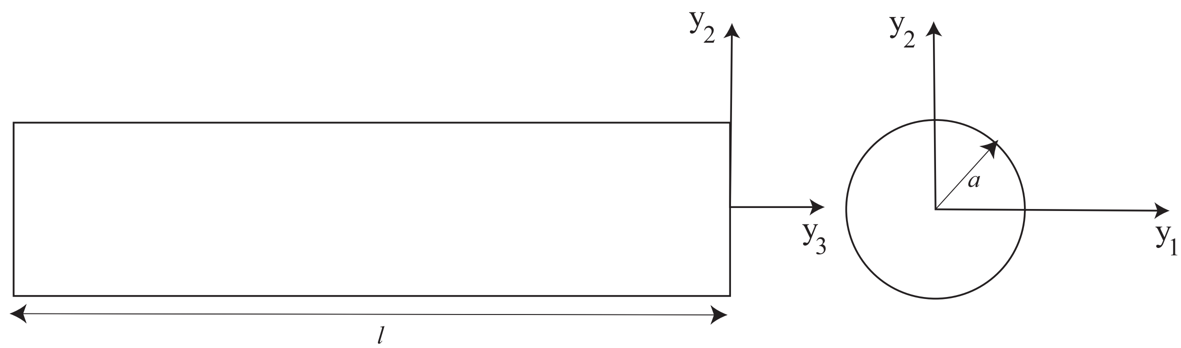


Figure 3.19: A right circular cylinder of length ℓ and radius a.

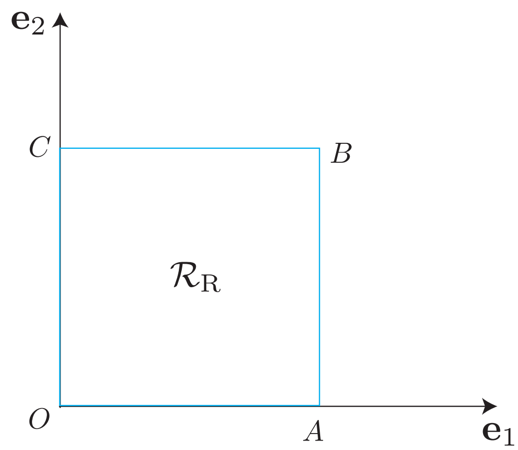


Figure 3.20: Side view of the region $\mathcal{R}_{\mathrm{R}}$.

(b) Suppose that the matrix of components of the Cauchy stress is

$$[T] = \begin{pmatrix} -p+\mu(\lambda^2+k^2) & \mu k\lambda^{-1} & 0 \\ \mu k\lambda^{-1} & -p+\mu\lambda^{-2} & 0 \\ 0 & 0 & -p+\mu \end{pmatrix}, \qquad (ii)$$

where μ is a known constant and p is an unknown constant.

The deformed images of the faces $x_3 = 0$ and $x_3 = 1$ are known to be traction-free. Simplify the expression (ii) for $[T]$ accordingly.

(c) Calculate the matrix of components of the first Piola-Kirchhoff stress tensor.

(d) Calculate the force (vector) that must be applied on the deformed image of the face $x_2 = 1$. Do this using both the first Piola-Kirchhoff and Cauchy tractions.

(e) Determine the (true) Cauchy traction that must be applied on the deformed image of the face $x_1 = 1$.

Problem 3.16. Consider a material such as a polarized dielectric solid under the action of an electric field, where (in addition to a body force $\mathbf{b}(\mathbf{y})$) there is also a *body couple* $\mathbf{c}(\mathbf{y})$ per unit deformed volume. Also, at any point $\mathbf{y}$ on a surface $\mathcal{S}$ suppose that there is (in addition to the contact force $\mathbf{t}(\mathbf{y},\mathbf{n})$) a *contact couple* $\mathbf{m}(\mathbf{y},\mathbf{n})$ per unit deformed area. Here $\mathbf{n}$ is the unit normal vector at a point on a surface in the deformed body and $\mathbf{m}$ is the couple applied by the material on the positive side of $\mathcal{S}$ on the material on the negative side. (The "positive side" of $\mathcal{S}$ is the side into which $\mathbf{n}$ points.)

Write down the global force and moment balance laws for this case. Show that in addition to the stress tensor $\mathbf{T}$ there is also a *couple stress tensor* $\mathbf{Z}(\mathbf{y})$ such that

$$\mathbf{m} = \mathbf{Z}\mathbf{n}.$$

Derive the local consequences of the force and moment equilibrium principles. Is the stress tensor $\mathbf{T}$ symmetric?

Problem 3.17. In Problem 3.16 we encountered couple stresses, and specifically, showed that there is a *couple stress tensor* $\mathbf{Z}$.

(a) Let $\mathbf{Z}_{\mathrm{R}}$ be the referential version of $\mathbf{Z}$, i.e. the tensor analogous to what the first Piola-Kirchhoff stress tensor $\mathbf{S}$ is to the Cauchy stress tensor $\mathbf{T}$. Derive a formula for $\mathbf{Z}_{\mathrm{R}}$.

(b) Similarly derive a formula for the referential body couple $\mathbf{c}_{\mathrm{R}}$.

(c) Derive the field equation obeyed by $\mathbf{Z}_{\mathrm{R}}$ and $\mathbf{c}_{\mathrm{R}}$ corresponding to moment balance in its referential form.

Problem 3.18. Consider a body that occupies a unit cube in a reference configuration:

$$\mathcal{R}_{\mathrm{R}} = \{(x_1, x_2, x_3) \mid 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 0 \le x_3 \le 1\}.$$

It is subjected to the following homogeneous deformation:

$$y_1 = x_1 + kx_2, \quad y_2 = \lambda x_2, \quad y_3 = x_3, \tag{i}$$

where k and λ are constants. The 1^{st}-Piola-Kirchhoff stress field in the body is uniform and its matrix of components is

$$[S] = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix}. \tag{ii}$$

Consider a surface $\mathcal{S}_{\mathrm{R}}$ in the reference configuration characterized by $x_1 + x_2 = 1$. The deformation carries $\mathcal{S}_{\mathrm{R}} \to \mathcal{S}$.

Without calculating the Cauchy stress tensor, determine

(a) the force (vector) that acts on $\mathcal{S}$,

(b) the true (Cauchy) traction on $\mathcal{S}$, and

(c) the normal component of true (Cauchy) traction on $\mathcal{S}$.

Next, calculate the matrix of components $[T]$ of the Cauchy stress tensor and recalculate your answers to parts $(a) - (c)$.

Problem 3.19. (Atkin and Fox) At each point $\mathbf{y}$ of a certain body the state of (Cauchy) stress is uniaxial in the radial direction with magnitude $\sigma(\mathbf{y})$, i.e.

$$\mathbf{T}(\mathbf{y}) = \sigma(\mathbf{y})\ \mathbf{m}(\mathbf{y}) \otimes \mathbf{m}(\mathbf{y}) \qquad \text{where} \quad \mathbf{m}(\mathbf{y}) = \frac{\mathbf{y}}{r}, \quad r = |\mathbf{y}|. \tag{i}$$

Assume the origin is outside the body so that $r \neq 0$. Show that

$$\mathbf{y} \cdot \operatorname{grad} \sigma + 2\sigma = 0 \qquad \text{for all} \quad \mathbf{y} \in \mathcal{R}. \tag{ii}$$

Determine $\sigma(\mathbf{y})$. Hint: In spherical polar coordinates $\mathbf{y} = r\mathbf{e}_r$.

Problem 3.20. Consider a very long circular cylindrical tube with inner and outer radii A and B respectively in the reference configuration. The tube is inflated to a pressure p, the outer wall being traction-free. The tube is made of an isotropic material and so the deformation and stress fields are axisymmetric and uniform in the axial direction. The inner and outer radii in the deformed configuration are a and b respectively. Consider a, b and p to be known. Work in cylindrical polar coordinates (r, θ, z) in the deformed configuration.

(a) What are the boundary conditions at $r = a$ and $r = b$?

(b) Specialize the general equilibrium equations (3.94) to the present setting.

(c) Now suppose that the tube is thin-walled, i.e. assume that $t \ll a, t \ll b$ where $t = b - a$ is the wall thickness. Use the equilibrium equation form part (b) to find an approximate expression for the circumferential Cauchy stress $T_{\theta\theta}$.

Problem 3.21. Reconsider Problem 3.3.4.

(a) Specialize the general equilibrium equations (in cylindrical polar coordinates) (3.94) to this setting. Assume no body forces.

(b) Show that for *any* smooth function $\phi(r, \theta)$, the stresses given by

$$T_{rr} = \frac{1}{r}\frac{\partial \phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \phi}{\partial \theta^2}, \qquad T_{r\theta} = -\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \phi}{\partial \theta}\right), \qquad T_{\theta\theta} = \frac{\partial^2 \phi}{\partial r^2}, \tag{$viii$}$$

satisfy the equilibrium equations from part (a).

(c) Suppose that the shear stress $T_{r\theta}(r,\theta)$ vanishes everywhere in the body. Determine the form of $\phi(r,\theta)$ implied by this and calculate expressions for the two nonzero normal stress components.

(d) Now impose the boundary conditions $(v), (vii), (ii)$ and (iii) (from Problem 3.3.4) and further simplify the form of the stresses.

Problem 3.22. (See also Problem 3.23.) Let (x_1, x_2, x_3) and $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be rectangular cartesian coordinates and the associated basis in the reference configuration; and let (r, θ, z) and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ be cylindrical polar coordinates and the associated basis in the deformed configuration. Let the first Piola-Kirchhoff stress tensor have components in these mixed bases

$$\begin{aligned}\mathbf{S} = \quad & S_{r1}\mathbf{e}_r \otimes \mathbf{e}_1 + S_{r2}\mathbf{e}_r \otimes \mathbf{e}_2 + S_{r3}\mathbf{e}_r \otimes \mathbf{e}_3 + \\ & + S_{\theta 1}\mathbf{e}_\theta \otimes \mathbf{e}_1 + S_{\theta 2}\mathbf{e}_\theta \otimes \mathbf{e}_2 + S_{\theta 3}\mathbf{e}_\theta \otimes \mathbf{e}_3 + \\ & + S_{z1}\mathbf{e}_z \otimes \mathbf{e}_1 + S_{z2}\mathbf{e}_z \otimes \mathbf{e}_2 + S_{z3}\mathbf{e}_z \otimes \mathbf{e}_3.\end{aligned}$$

e.g. see Problem 3.7.1 . Calculate

$$\Big(\mathrm{Div}\,\mathbf{S}\Big)\cdot\mathbf{e}_r, \qquad \Big(\mathrm{Div}\,\mathbf{S}\Big)\cdot\mathbf{e}_\theta, \qquad \Big(\mathrm{Div}\,\mathbf{S}\Big)\cdot\mathbf{e}_z.$$

and hence derive the equilibrium equations in these coordinates.

Problem 3.23. (See also Problem 3.22.) Let (R, Θ, Z) and $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ be cylindrical polar coordinates and the associated basis in the reference configuration; and let (r, θ, z) and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ be cylindrical polar coordinates and the associated basis in the deformed configuration. Let the first Piola-Kirchhoff stress tensor have components in these mixed bases

$$\begin{aligned}\mathbf{S} = \quad & S_{rR}\mathbf{e}_r \otimes \mathbf{e}_R + S_{r\Theta}\mathbf{e}_r \otimes \mathbf{e}_\Theta + S_{rZ}\mathbf{e}_r \otimes \mathbf{e}_Z + \\ & + S_{\theta R}\mathbf{e}_\theta \otimes \mathbf{e}_R + S_{\theta\Theta}\mathbf{e}_\theta \otimes \mathbf{e}_\Theta + S_{\theta Z}\mathbf{e}_\theta \otimes \mathbf{e}_Z + \\ & + S_{zR}\mathbf{e}_z \otimes \mathbf{e}_R + S_{z\Theta}\mathbf{e}_z \otimes \mathbf{e}_\Theta + S_{zZ}\mathbf{e}_z \otimes \mathbf{e}_Z.\end{aligned}$$

Calculate

$$\Big(\mathrm{Div}\,\mathbf{S}\Big)\cdot\mathbf{e}_r, \qquad \Big(\mathrm{Div}\,\mathbf{S}\Big)\cdot\mathbf{e}_\theta, \qquad \Big(\mathrm{Div}\,\mathbf{S}\Big)\cdot\mathbf{e}_z.$$

and hence derive the equilibrium equations in these coordinates.

Problem 3.24. (See Problem 2.29 for an analysis of the kinematics of a piecewise homogeneous deformation.) Consider a planar surface $\mathcal{S}$ that passes through the region $\mathcal{R}$ occupied by a body in the deformed configuration. Let $\mathbf{n}$ be a unit vector normal to $\mathcal{S}$ and let $\mathcal{R}^+$ denote the side into which $\mathbf{n}$ points, $\mathcal{R}^-$ the other side. Thus $\mathcal{S}$ is a planar interface between two parts of the body. Consider the piecewise homogeneous stress field

$$\mathbf{T}(\mathbf{y}) = \begin{cases} \mathbf{T}^{+} & \text{for } \mathbf{y} \in \mathcal{R}^{+}, \\ \mathbf{T}^{-} & \text{for } \mathbf{y} \in \mathcal{R}^{-}, \end{cases}$$

where $\mathbf{T}^{\pm}$ are constant symmetric tensors. Show that this stress field obeys force and moment balance if and only if

$$\mathbf{T}^{+}\mathbf{n} - \mathbf{T}^{-}\mathbf{n} = \mathbf{o}. \tag{3.100}$$

Let $\mathcal{S}_{\mathrm{R}}$ be the image of $\mathcal{S}$ in the reference configuration with $\mathbf{n}_{\mathrm{R}}$ being a unit vector normal to $\mathcal{S}_{\mathrm{R}}$. Assume that the associated deformation is piecewise homogeneous (see Problem 2.29). Show that force and moment balance requires the first Piola-Kirchhoff stress tensor field associated with the aforementioned stress field to obey

$$\mathbf{S}^{+}\mathbf{n}_{\mathrm{R}} - \mathbf{S}^{-}\mathbf{n}_{\mathrm{R}} = \mathbf{o}. \tag{3.101}$$

These "jump conditions" plays an important role in studying interfaces between two material phases.

Problem 3.25. Suppose one did not postulate the force balance law (3.8). The moment balance law *about an arbitrary pivot point* $\mathbf{z}$ requires

$$\int_{\partial D} (\mathbf{y} - \mathbf{z}) \times \mathbf{t}(\mathbf{n})\, dA_y + \int_{D} (\mathbf{y} - \mathbf{z}) \times \mathbf{b}\, dV_y = \mathbf{o}, \tag{i}$$

for all parts $\mathcal{D}$ of the body. Show from (i) that there exists a tensor $\mathbf{T}(\mathbf{y})$ such that $\mathbf{t}(\mathbf{n}) = \mathbf{T}(\mathbf{y})\mathbf{n}$. Moreover, by requiring (i) to hold for all pivot points $\mathbf{z}$, show that $\operatorname{div} \mathbf{T} + \mathbf{b} = \mathbf{o}$, i.e. derive the field equation corresponding to *force* balance! The symmetry of $\mathbf{T}$ can be shown in the usual way.

Problem 3.26. Establish the **Principle of Virtual Work**, i.e. show that the equilibrium equation

$$\mathrm{Div}\ \mathbf{S} + \mathbf{b}_{\mathrm{R}} = \mathbf{0} \tag{i}$$

holds at each $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$ *if and only if*

$$\int_{\partial \mathcal{R}_{\mathrm{R}}} \mathbf{S}\mathbf{n}_{\mathrm{R}} \cdot \mathbf{w}\, dA_x + \int_{\mathcal{R}_{\mathrm{R}}} \mathbf{b}_{\mathrm{R}} \cdot \mathbf{w}\, dV_x = \int_{\mathcal{R}_{\mathrm{R}}} \mathbf{S} \cdot \nabla \mathbf{w}\, dV_x \tag{ii}$$

for *all arbitrary* smooth enough vector fields $\mathbf{w}(\mathbf{x})$.

Remark: Note that $\mathbf{w}(\mathbf{x})$ is **not** required to be the actual displacement field in the body. It is called a "virtual displacement".

Problem 3.27. Two symmetric tensors are said to be *coaxial* if their principal axes coincide. Prove that the Cauchy stress tensor $\mathbf{T}$ and the left Cauchy-Green tensor $\mathbf{B}$ are coaxial if and only if the second Piola-Kirchhoff tensor $\mathbf{S}^{(2)}$ is coaxial with the right Cauchy-Green strain tensor $\mathbf{C}$.

Problem 3.28. If the Cauchy stress tensor $\mathbf{T}$ and the left Cauchy-Green tensor $\mathbf{B}$ are coaxial, show that the Biot stress tensor $\mathbf{S}^{(1)}$ is coaxial with the Biot strain tensor $\mathbf{E}^{(1)}$. Is the converse true? (Two symmetric tensors are said to be coaxial if their principal axes coincide.)

Problem 3.29. Determine the symmetric stress tensor that is work conjugate to the Biot strain tensor

$$\mathbf{E}^{(1)} = \mathbf{U} - \mathbf{I}.$$

Problem 3.30. Determine the stress tensor that is conjugate to the (Lagrangian) logarithmic strain tensor $\ln \mathbf{U}$.

Solution: See the paper by A. Hoger, The stress conjugate to logarithmic strain, *International Journal of Solids and Structures*, **23**(1987), pp. 1645-1656.

Problem 3.31. Pick any Eulerian strain tensor of your choice. Find the stress tensor that is conjugate to it.

Solution: See the paper by Andrew Norris, Eulerian conjugate stress and strain, *J. Mech. Materials Struct.*, **3**(2008), pp. 243-260. In general finding stress tensors conjugate to Eulerian strains is much more difficult than the corresponding problem for Lagrangian strains.

Problem 3.32. Write down expressions for the rate of working of the forces and couples in the settings of Problems 3.16 and 3.17. Rewrite these in the form of a volume integral of the local power.

Problem 3.33. (*Conservation of mass. Rate of change of linear momentum.)* You may find it helpful to review Section 2.11 on the material time derivative and the transport formula.

(a) A body undergoes a motion $\mathbf{y} = \mathbf{y}(\mathbf{x}, t)$. Let $\rho(\mathbf{y}, t)$ be the mass density of the body at the particle that is located at $\mathbf{y}$ at time t. Consider a part of a body, and let $\mathcal{D}_t$ be the region of space it occupies at time t. Note that the region $\mathcal{D}_t$ varies with time. The mass of this part is the integral of $\rho(\mathbf{y},t)$ over $\mathcal{D}_t$. The conservation of mass requires that the mass of every part of the body be time-independent:

$$\frac{d}{dt}\int_{\mathcal{D}_t} \rho(\mathbf{y},t)\,dV_y = 0 \qquad \text{for all parts } \mathcal{D}_t. \tag{i}$$

Show that the balance law (i) holds if and only if the following field equation holds,

$$\dot{\rho} + \rho\,\mathrm{div}\,\mathbf{v} = 0 \qquad \text{at all } \mathbf{y} \in \mathcal{R}_t, \tag{3.102}$$

where $\dot{\rho}$ is the material time derivative of ρ and the cartesian components of $\mathrm{div}\,\mathbf{v}$ are $\partial v_i/\partial y_i$.

(b) Let $\rho_R(\mathbf{x})$ be the mass density of the body in a reference configuration. Show that

$$\rho_R = \rho J. \tag{3.103}$$

(c) Let $\mathbf{v}(\mathbf{y},t)$ be (the spatial description of) the velocity field. It is defined on $\mathcal{R}_t$ at each t. Show that the rate of increase of the linear momentum of the part under consideration is

$$\frac{d}{dt}\int_{\mathcal{D}_t} \rho\mathbf{v}\,dV_y = \int_{\mathcal{D}_t} \rho\dot{\mathbf{v}}\,dV_y \tag{3.104}$$

where $\dot{\mathbf{v}}$ is the material time derivative of $\mathbf{v}$, i.e. the acceleration.

References:

1. R. Abeyaratne, Chapters 4, 5 and 6 of *Continuum Mechanics*, Volume II in the series *Lecture Notes on The Mechanics of Solids*, http://web.mit.edu/abeyaratne/lecture_notes.html.

2. R. J. Atkin and N. Fox, *An Introduction to the Theory of Elasticity*, Longman, 1980.

3. P. Chadwick, Chapter 3 of *Continuum Mechanics: Concise Theory and Problems*, Dover, 1999.

4. M.E. Gurtin, E. Fried and L. Anand, Part IV of *The Mechanics and Thermodynamics of Continua*, Cambridge University Press, 2010.

5. R.W. Ogden, Chapter 3 of *Non-Linear Elastic Deformations*, Dover, 1997.

6. A.J.M. Spencer, Chapters 5 and 7 of *Continuum Mechanics*, Longman, 1980.

7. D. J. Steigmann, Chapters 2 and 3 of *Finite Elasticity Theory*, Oxford University Press, 2017.

Chapter 4

Constitutive Relation

In principle, the analyses of deformation in Chapter 2 and stress in Chapter 3 are valid for any continuum, irrespective of the specific material of which the body is composed. However, given the loading applied on a body, the basic equations derived in those chapters are not sufficient for determining the resulting stress and deformation fields. Additional information describing how the stress depends on the deformation is needed, and this comes from considering the behavior of the specific material at hand. This is not surprising since even in the simplest case of a spring, given the force applied on the spring, one cannot determine its elongation without knowing something about the material of which it is made.

Nonlinear elasticity has been, and continues to be, successfully used to study a variety of materials such as biological tissues, "soft materials" more generally including elastomeric materials, and crystalline solids undergoing martensitic phase transformations.

There are two main approaches to constructing continuum-scale constitutive relations. One begins at the atomistic-scale and attempts to deduce the continuum-scale response by some sort of averaging across length and time scales ("coarse graining"). The other so-called "phenomenological approach" begins directly at the continuum-scale guided by experimental observations and some basic principles. A combination of these two approaches, where micro-mechanical considerations are used to infer the form of the constitutive relation, the details of which are then explored experimentally, is often particularly effective.

In Chapter 8 we shall describe one micro-mechanical model – Cauchy's beautiful derivation of the constitutive relation of a crystalline solid using a simple lattice model in which the atoms interact through a pair potential. A second micro-mechanical model that would have been natural for us to describe is a polymer chain model and its use in constructing the constitutive relation for a rubber-like material. Unfortunately, this relies crucially on calculating the entropy, a thermodynamic notion that we do not address in these notes. The interested reader can refer to Chapter 9 of Volume II.

In these notes we are concerned with *elastic materials*. We shall assume that the defining characteristic of an elastic material is that it does not dissipate energy (at least when there are no moving singularities in the body such as a propagating crack). Such elastic materials are frequently said to be *hyperelastic* (or Green elastic).

A word on notation: in order to avoid confusion, it will sometimes be helpful to distinguish between functions of different arguments, even when their values represent the same quantity. In particular, we will denote the so-called strain energy function in the various forms $W(\mathbf{x})$, $\widehat{W}(\mathbf{F})$, $\overline{W}(\mathbf{C})$, $\widetilde{W}(I_1, I_2, I_3)$ and $W^*(\lambda_1, \lambda_2, \lambda_3)$. Even though they all represent the elastic energy density, they are different *functions*. When it is not essential that we make the distinction, and there is no chance for confusion, we will simply write $W(\mathbf{x})$, $W(\mathbf{F})$, $W(\mathbf{C})$, $W(I_1, I_2, I_3)$ and $W(\lambda_1, \lambda_2, \lambda_3)$.

Occasionally, we will refer to the time t. When we do so, we will *not* be taking inertial effects into account[1]. Instead, we will simply be considering a one-parameter family of equilibrium deformations, a so-called *quasi-static motion*, with t merely being the parameter. In a quasi-static motion, the stress field obeys the equilibrium equation at each instant t.

A roadmap of this chapter is as follows. In Section 4.2 we characterize an elastic material in terms of its strain energy function $W(\mathbf{F})$. The implications of material frame indifference are explored in Section 4.3. We turn in Section 4.4 to material symmetry with Section 4.4.2 devoted to isotropic materials. (Some anisotropic materials are considered in Chapter 6.) In Section 4.5 materials with internal constraints such as incompressibility and inextensibility are considered. The response in uniaxial tension, simple shear and biaxial plane stress are explored for a general isotropic material in Section 4.7. Restrictions imposed on the strain energy function for reasons of physical reasonableness and mathematical necessity are touched upon in Section 4.6.3. In Section 4.7 we describe a few specific strain energy functions from the literature. Finally in Section 4.8 we specialize the constitutive relation to infinitesimal deformations and thus derive the stress-strain relation in linear(ized) elasticity.

The discussion of constitutive relations in this chapter, even when limited to elastic materials, is concise and incomplete. An expanded treatment can be found in the references

---

[1] We make an exception briefly when referring to strong ellipticity in Section 4.6.3.

cited at the end of this chapter and in Chapters 7, 8 (and 9) of Volume II.

4.1 Motivation.

We start by motivating why it is necessary to undertake a careful discussion of the constitutive relation since, based on our experience with linear elasticity theory as undergraduates, it may feel natural to simply write down a relationship between a stress tensor and a strain tensor, say,

$$\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{E}) \qquad \text{where} \quad \mathbf{E} = \frac{1}{2}\big(\mathbf{F}^T\mathbf{F} - \mathbf{I}\big). \tag{i}$$

Is this a reasonable constitutive relation?

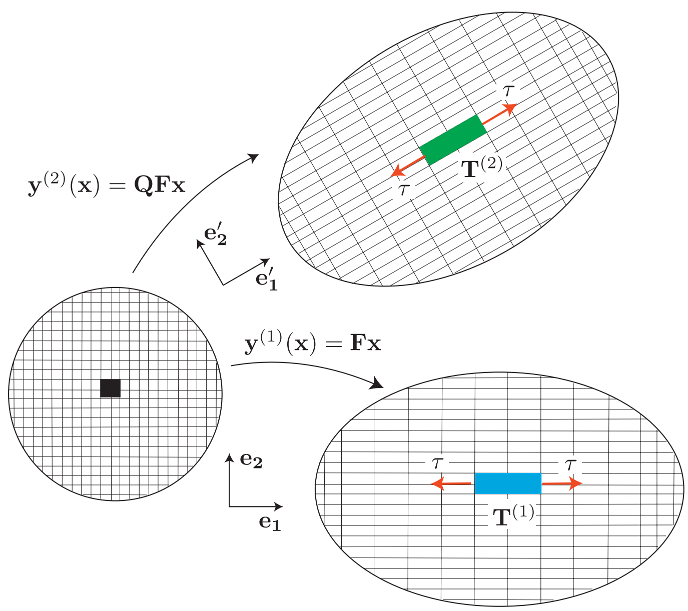


Figure 4.1: A body subjected to two deformations $\mathbf{y}^{(1)}(\mathbf{x}) = \mathbf{Fx}$ and $\mathbf{y}^{(2)}(\mathbf{x}) = \mathbf{QFx}$ that differ by a rigid rotation $\mathbf{Q}$, and the associated stress tensors.

To explore this question, first consider a homogeneous deformation

$$\mathbf{y}^{(1)}(\mathbf{x}) = \mathbf{Fx}. \tag{ii}$$

For illustrative purposes (only), suppose that $\mathbf{F}$ describes a uniaxial stretch in the $\mathbf{e}_1$-direction with equal contraction in the directions normal to it:

$$\mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + (1/\lambda_2)(\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3).$$

Suppose the associate stress is a uniaxial stress in the $\mathbf{e}_1$-direction:

$$\mathbf{T}^{(1)} = \tau\, \mathbf{e}_1 \otimes \mathbf{e}_1. \qquad (iii)$$

Next consider a second deformation, identical to the first followed by an arbitrary rigid rotation $\mathbf{Q}$:

$$\mathbf{y}^{(2)}(\mathbf{x}) = \mathbf{QFx}. \qquad (iv)$$

Figure 4.1 shows a cartoon of these two deformations where the small square in the reference configuration has been stretched in the $\mathbf{e}_1$-direction in the first deformation, and has been rotated after stretching in the $\mathbf{e}_1$-direction in the second. On physical grounds, we would expect the stress associated with the second deformation to be a uniaxial stress in the $\mathbf{e}'_1$-direction:

$$\mathbf{T}^{(2)} = \tau\, \mathbf{e}'_1 \otimes \mathbf{e}'_1. \qquad (v)$$

Here the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is obtained by rotating the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ by $\mathbf{Q}$:

$$\mathbf{e}'_i = \mathbf{Q}\mathbf{e}_i; \qquad (vi)$$

see Figure 4.1. While the two stress tensors are distinct, they are related by

$$\mathbf{T}^{(2)} \overset{(v)}{=} \tau \mathbf{e}'_1 \otimes \mathbf{e}'_1 \overset{(vi)}{=} \tau(\mathbf{Q}\mathbf{e}_1) \otimes (\mathbf{Q}\mathbf{e}_1) = \tau \mathbf{Q}(\mathbf{e}_1 \otimes \mathbf{e}_1)\mathbf{Q}^T \overset{(iii)}{=} \mathbf{Q}\mathbf{T}^{(1)}\mathbf{Q}^T. \qquad (vii)$$

Using the constitutive relation $(i)_1$ we have

$$\mathbf{T}^{(1)} = \widehat{\mathbf{T}}(\mathbf{E}^{(1)}), \qquad \mathbf{T}^{(2)} = \widehat{\mathbf{T}}(\mathbf{E}^{(2)}), \qquad (viii)$$

so that from (vii) and $(viii)$,

$$\widehat{\mathbf{T}}(\mathbf{E}^{(2)}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{E}^{(1)})\mathbf{Q}^T. \qquad (ix)$$

However the Green Saint-Venant strain tensors associated with these two deformations are

$$\mathbf{E}^{(1)} = \frac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I}), \qquad \mathbf{E}^{(2)} = \frac{1}{2}\big((\mathbf{QF})^T(\mathbf{QF}) - \mathbf{I}\big) = \frac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I}) = \mathbf{E}^{(1)}. \qquad (x)$$

Thus from (ix) and (x),

$$\widehat{\mathbf{T}}(\mathbf{E}^{(1)}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{E}^{(1)})\mathbf{Q}^T,$$

or, since $\mathbf{E}^{(1)}$ is in fact arbitrary,

$$\widehat{\mathbf{T}}(\mathbf{E}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{E})\mathbf{Q}^T \qquad \text{for all rotations } \mathbf{Q}, \qquad (xi)$$

and all strains $\mathbf{E}$. This is a restriction on the form of the function $\widehat{\mathbf{T}}(\mathbf{E})$. In fact, from the result in Problem 1.35, it follows that $\widehat{\mathbf{T}}(\mathbf{E})$ must be a scalar multiple of the identity:

$$\widehat{\mathbf{T}}(\mathbf{E}) = \tau(\mathbf{E})\,\mathbf{I}. \tag{xii}$$

Thus a constitutive relation of the general form $\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{E})$ must necessarily be of the particular form $\widehat{\mathbf{T}}(\mathbf{E}) = \tau(\mathbf{E})\,\mathbf{I}$ where the Cauchy stress tensor is hydrostatic (in all deformations) and so the material is a fluid!

Problem 4.1.1. You might say that a shortcoming of the particular constitutive relation (i) above was that the Cauchy stress tensor is associated with the deformed configuration while the Lagrangian strain tensor is associated with the reference configuration. Based on this, replace (i) with the ansatz $\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{B})$ where $\mathbf{B}$ is the (Eulerian) left Cauchy-Green tensor and carry out an analysis like the one above. What does this tell you about the form of $\widehat{\mathbf{T}}(\mathbf{B})$?

4.2 An Elastic Material.

- First, we assume that the stress at some particle $\mathbf{x}$ depends only on the deformation of the particles in the immediate neighborhood of that particle. Such a theory is said to be a local theory. We know from Chapter 2 that the deformation in the vicinity of a particle is completely characterized by the deformation gradient tensor $\mathbf{F} = \nabla\mathbf{y}$ at that particle. This implies that the stress $\mathbf{S}$ at particle $\mathbf{x}$ depends on the deformation solely through the deformation gradient tensor[2] $\mathbf{F}$ at particle $\mathbf{x}$.

- Second, we assume further that an elastic material has no memory[3], and therefore that the stress $\mathbf{S}$ at time t depends only on the value of the deformation gradient tensor $\mathbf{F}$ at that same time t. We are thus led to consider constitutive relations of the form $\mathbf{S}(\mathbf{x}, t) = \widehat{\mathbf{S}}(\mathbf{F}(\mathbf{x}, t))$, or simply
$$\mathbf{S} = \widehat{\mathbf{S}}(\mathbf{F}). \tag{4.1}$$

- Terminology: Note the distinction between $\mathbf{S}(\mathbf{x})$ and $\widehat{\mathbf{S}}(\mathbf{F})$. The former is the stress *field* in the body while the latter is the *constitutive response function* for stress. The material is characterized by $\widehat{\mathbf{S}}$.

[2]In a nonlocal theory the stress might, for example, depend on the deformation gradient *and* the gradient of the deformation gradient.

[3]unlike, say, a viscoelastic material which depends on the past history of the deformation.

- Since the Cauchy and 1st Piola-Kirchhoff stress tensors are related by $\mathbf{T} = \mathrm{J}^{-1}\mathbf{S}\mathbf{F}^T$, knowing $\widehat{\mathbf{S}}$ gives

$$\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{F}) \qquad \text{where} \quad \widehat{\mathbf{T}}(\mathbf{F}) = \mathrm{J}^{-1}\widehat{\mathbf{S}}(\mathbf{F})\mathbf{F}^T, \quad \mathrm{J} = \det \mathbf{F}. \tag{4.2}$$

  It will be convenient for us to work with the 1st Piola-Kirchhoff stress tensor.

- If the material is inhomogeneous in the reference configuration, the constitutive response function $\widehat{\mathbf{S}}$ will depend explicitly on $\mathbf{x}$: $\mathbf{S} = \widehat{\mathbf{S}}(\mathbf{F}, \mathbf{x})$.

- Third, we assume there is no energy dissipation in an elastic material in the following sense: let $W(\mathbf{x}, t)$ denote the *stored energy density*, i.e. the energy stored per unit volume in the reference configuration[4]. The total elastic energy stored in a part of the body is then

$$\int_{\mathcal{D}_{\mathrm{R}}} W \, dV_x,$$

  where $\mathcal{D}_{\mathrm{R}}$ is the region in the reference configuration occupied by the part under consideration. When we say that an elastic material is dissipation-free we mean that the rate at which external work is done on any part of the body during a quasi-static motion equals the corresponding rate of increase of stored energy[5] provided the fields are smooth:

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{s} \cdot \mathbf{v} \, dA_x + \int_{D_{\mathrm{R}}} \mathbf{b} \cdot \mathbf{v} \, dV_x = \frac{d}{dt} \int_{\mathcal{D}_{\mathrm{R}}} W \, dV_x. \tag{4.3}$$

  This must hold *in all quasi-static motions* for all parts of the body at all times.

- An important consequence of (4.3) can be deduced by combining it with (3.73). Equation (3.73) states that the rate at which the external forces do work equals the rate at which the stresses do work (the stress power). Equation (3.73) together with (4.3) yields

$$\int_{\mathcal{D}_{\mathrm{R}}} \mathbf{S} \cdot \dot{\mathbf{F}} \, dV_x \; = \; \frac{d}{dt} \int_{\mathcal{D}_{\mathrm{R}}} W \, dV_x, \tag{4.4}$$

  which says that the rate at which the "internal forces" (the stresses) do work equals the rate of increase of the stored energy. This is the sense in which the material is dissipationless.

[4]The stored energy *per unit mass*, say ψ, is related to the energy per unit reference volume W by $W = \rho_{\mathrm{R}}\psi$ where ρ_{R} is the mass density in the reference configuration. When thermal effects are accounted for W depends on both $\mathbf{F}$ and temperature.

[5]When inertial effects are taken into account, one must include the rate of increase of kinetic energy on the right-hand side of (4.3).

The preceding equation can be written as

$$\int_{\mathcal{D}_{\rm R}} \left(\mathbf{S} \cdot \dot{\mathbf{F}} \; - \dot{W} \right) \, dV_x = 0.$$

(Recall that the superior dot denotes the derivative with respect to t at a fixed particle $\mathbf{x}$ – the "material time derivative", see Section 2.11.2.) Since this must hold for every part of the body it follows by localization that

$$\mathbf{S} \cdot \dot{\mathbf{F}} \; - \; \dot{W} = 0 \qquad \text{for all } \mathbf{x} \in \mathcal{R}_{\rm R}. \tag{4.5}$$

– Finally, just as for the stress, it is natural to assume that the stored energy density W at a particle $\mathbf{x}$ at time t depends on the deformation only through the deformation gradient tensor $\mathbf{F}$ at the same particle $\mathbf{x}$ at the same time t, i.e. that there is a constitutive response function $\widehat{W}$ such that $W(\mathbf{x}, t) = \widehat{W}(\mathbf{F}(\mathbf{x}, t))$ or more simply

$$W = \widehat{W}(\mathbf{F}).$$

– On using $W = \widehat{W}(\mathbf{F})$ and $\mathbf{S} = \widehat{\mathbf{S}}(\mathbf{F})$ in (4.5) we get

$$\widehat{\mathbf{S}} \cdot \dot{\mathbf{F}} - \frac{\partial \widehat{W}}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} = 0 \qquad \Rightarrow \qquad \left(\widehat{\mathbf{S}}(\mathbf{F}) - \frac{\partial \widehat{W}}{\partial \mathbf{F}}(\mathbf{F}) \right) \cdot \dot{\mathbf{F}} = 0. \tag{4.6}$$

This must hold in *every quasi-static motion of the body.* Observe that the terms in the parenthesis only involve the deformation gradient tensor $\mathbf{F}$ and not its rate $\dot{\mathbf{F}}$. For a given $\mathbf{F}$, one can always construct a motion with an *arbitrary* $\dot{\mathbf{F}}$ at a particular particle at a particular instant[6]. It follows that (4.6) must hold for all tensors $\dot{\mathbf{F}}$ and therefore

[6]We encountered this issue previously in Problem 1.8.4. To illustrate the argument used, consider a one-dimensional continuum. Let $y = y(x,t)$ be a motion with the stretch λ and stretch-rate $\dot{\lambda}$ defined by $\lambda(x,t) = \partial y/\partial x$ and $\dot{\lambda}(x,t) = \partial \lambda / \partial t$. Then the claim is that one can always find a motion $y(x,t)$ in which the values of λ and $\dot{\lambda}$ at some *particular instant* can be arbitrarily and independently prescribed.

To see this, pick and fix an arbitrary instant t_o and let $\lambda_o > 0$ and r_o be any two constants, each chosen arbitrarily and independently of the other. Consider the motion

$$y(x,t) = \lambda_o x \exp\left(\frac{r_o (t - t_o)}{\lambda_o} \right), \tag{i}$$

and observe that

$$\lambda(x, t_0) = \lambda_o, \qquad \dot{\lambda}(x, t_o) = r_o. \tag{ii}$$

Suppose that $\widehat{\sigma}(\lambda)$ and $\widehat{W}(\lambda)$ are functions such that

$$\Big[\widehat{\sigma}(\lambda(x,t)) - \widehat{W}'(\lambda(x,t)) \Big] \, \dot{\lambda}(x,t) = 0, \tag{iii}$$

that

$$\widehat{\mathbf{S}}(\mathbf{F}) = \frac{\partial \widehat{W}}{\partial \mathbf{F}}(\mathbf{F}). \tag{4.7}$$

This tells us that given the constitutive response function $\widehat{\mathbf{S}}$ for the stress can be calculated from the the constitutive response function $\widehat{W}$ for the stored energy. We write this less formally as

$$\boxed{\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}},} \qquad \boxed{\mathbf{T} = \mathrm{J}^{-1}\mathbf{S}\mathbf{F}^T = \frac{1}{J}\frac{\partial W}{\partial \mathbf{F}}\,\mathbf{F}^T.} \tag{4.8}$$

We thus conclude that an elastic material is characterized by the constitutive response function $\widehat{W}(\mathbf{F})$ for the stored energy per unit reference volume. It is referred to as the **strain-energy function**.

If the material is inhomogeneous in the reference configuration we would have $W = \widehat{W}(\mathbf{F}, \mathbf{x})$.

4.2.1 An elastic material. Alternative approach.

The preceding analysis hinged on the balance (4.3) between the rate of work and energy, and therefore cannot be used, at least not directly, when there is dissipation. In this subsection we briefly present a modification of the preceding analysis based on the dissipation inequality. While not essential in elasticity, this approach can be used in the study of inelastic materials (and we shall do so in Chapter 9 when we touch on coupled problems).

---

for all motions $y(x,t)$, all particles x and all instants of time t. Since this holds for all t it necessarily holds at $t = t_o$:

$$\left[\widehat{\sigma}(\lambda(x,t_o)) - \widehat{W}'(\lambda(x,t_o))\right]\dot{\lambda}(x,t) = 0. \tag{iv}$$

Substituting (ii) into (iv) yields

$$\left[\widehat{\sigma}(\lambda_o) - \widehat{W}'(\lambda_o)\right] r_o = 0. \tag{v}$$

This must hold for all r_o. Since λ_o is independent of r_o, it follows that necessarily

$$\widehat{\sigma}(\lambda_o) = \widehat{W}'(\lambda_o),$$

which holds for all $\lambda_o > 0$. For a discussion of this issue in three-dimensions, see Section 3.4 of Gurtin et al. [10].

– **Dissipation inequality:** Let $W(\mathbf{x},t)$ be the *free energy* per unit reference volume. The total free energy in a part of the body is then

$$\int_{\mathcal{D}_{\mathrm{R}}} W\, dV_x,$$

where $\mathcal{D}_{\mathrm{R}}$ is the region in the reference configuration occupied by the part under consideration.

The *dissipation inequality* states that the rate of increase of free energy cannot exceed the rate at which external work is done[7]:

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{s}\cdot\mathbf{v}\, dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{b}\cdot\mathbf{v}\, dV_x \;\geq\; \frac{d}{dt}\int_{\mathcal{D}_{\mathrm{R}}} W\, dV_x. \tag{4.9}$$

This must hold *in all quasi-static motions* for all parts of the body at all times. Proceeding as above leads one to the local inequality (Exercise)

$$\mathbf{S}\cdot\dot{\mathbf{F}} \,-\, \dot{W} \geq 0 \qquad \text{for all } \mathbf{x}\in\mathcal{R}_{\mathrm{R}}. \tag{4.10}$$

– **Constitutive equations: primitive form.** We assume that the stress and free energy at particle $\mathbf{x}$ at time t depend only on the deformation of the particles in the immediate neighborhood of that particle at that same instant. This implies that $\mathbf{S}(\mathbf{x},t)$ and $W(\mathbf{x},t)$ depend on the deformation solely through $\mathbf{F}(\mathbf{x},t)$. We are thus led to consider constitutive relations of the form $\mathbf{S}(\mathbf{x},t) = \widehat{\mathbf{S}}(\mathbf{F}(\mathbf{x},t))$ and $W(\mathbf{x},t) = \widehat{W}(\mathbf{F}(\mathbf{x},t))$.

Accordingly, we now assume that the material is characterized by the constitutive equations

$$\mathbf{S} = \widehat{\mathbf{S}}(\mathbf{F}), \qquad W = \widehat{W}(\mathbf{F}). \tag{4.11}$$

This is the primitive form of the constitutive relations.

– **Constitutive equations: simplified (reduced) form.** On using (4.11) in (4.10) we get

$$\widehat{\mathbf{S}}\cdot\dot{\mathbf{F}} - \frac{\partial\widehat{W}}{\partial\mathbf{F}}\cdot\dot{\mathbf{F}} \geq 0 \qquad \Rightarrow \qquad \left[\widehat{\mathbf{S}}(\mathbf{F}) - \frac{\partial\widehat{W}}{\partial\mathbf{F}}(\mathbf{F})\right]\cdot\dot{\mathbf{F}} \geq 0. \tag{4.12}$$

This must hold in *every quasi-static motion of the body.* Observe that the terms in the square brackets only involve the deformation gradient tensor $\mathbf{F}$ and not its rate $\dot{\mathbf{F}}$. The argument used on page 249 can be generalized (and is referred to as the Coleman-Noll argument) to conclude that the terms within the square brackets must vanish:

$$\widehat{\mathbf{S}}(\mathbf{F}) = \frac{\partial\widehat{W}}{\partial\mathbf{F}}(\mathbf{F}). \tag{4.13}$$

We thus conclude that an elastic material is completely characterized by the constitutive response function $\widehat{W}(\mathbf{F})$ for the free energy.

– One can now use (4.13) to show that (4.10), and therefore (4.9), hold with *equality.* This recovers the power - energy *balance* of the preceding section and motivates us to refer to W as the stored energy.

[7] When inertial effects are taken into account, one must include the rate of increase of kinetic energy on the right-hand side of (4.9).

4.3 Material frame indifference.

Material frame indifference refers to the idea that physical laws should be independent of the observer. The reader may refer to, for example, Chapter 2 of Steigmann [20] for a careful treatment of this topic. Our treatment here is based on the idea that the stored elastic energy be unaffected by a superposed rigid body deformation.

– Without loss of generality suppose that the stored energy vanishes in the reference configuration, i.e. $\widehat{W}(\mathbf{I}) = 0$.

  Now suppose that the body is subjected to a rigid rotation characterized by the proper orthogonal tensor $\mathbf{Q}$. Such a deformation does not distort the body and so it should not store any elastic energy. Accordingly one expects $\widehat{W}(\mathbf{Q}) = 0$ for all proper orthogonal tensors $\mathbf{Q}$. Thus it would be natural to require the function $\widehat{W}$ to have this property.

– **Principle of material frame indifference:** More generally, suppose that the body is subjected to a homogeneous deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$. The associated stored energy density is $\widehat{W}(\mathbf{F})$. Suppose that this body is now subjected to a *further* rigid rotation $\mathbf{Q}$. This is equivalent to considering the deformation $\mathbf{y} = \mathbf{QF}\mathbf{x}$. The associated stored energy density is $\widehat{W}(\mathbf{QF})$. We do not expect any additional energy to be stored in the body due to the subsequent rigid rotation, and therefore require $\widehat{W}$ to have the property

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{QF}) \tag{4.14}$$

  for all nonsingular tensors $\mathbf{F}$ and proper orthogonal tensors $\mathbf{Q}$. A strain energy function $\widehat{W}(\mathbf{F})$ that conforms to (4.14) is said to be *frame indifferent* (or objective).

Problem 4.3.1. Show that the Cauchy and 1st Piola-Kirchhoff stress response functions $\widehat{\mathbf{T}}(\mathbf{F})$ and $\widehat{\mathbf{S}}(\mathbf{F})$ are frame indifferent if

$$\widehat{\mathbf{T}}(\mathbf{QF}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^T, \qquad \widehat{\mathbf{S}}(\mathbf{QF}) = \mathbf{Q}\widehat{\mathbf{S}}(\mathbf{F}), \tag{4.15}$$

for all nonsingular tensors $\mathbf{F}$ and proper orthogonal tensors $\mathbf{Q}$.

Problem 4.3.2. Consider two deformations $\mathbf{y} = \mathbf{F}\mathbf{x}$ and $\mathbf{y} = \mathbf{QF}\mathbf{x}$, and two bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ where the latter is obtained by rotating the former by $\mathbf{Q}$: $\mathbf{e}'_i = \mathbf{Q}\mathbf{e}_i$. Let T_{ij} and T'_{ij} be the components of the Cauchy stress tensor in these two bases. Show that material frame indifference is equivalent to the requirement

$$\widehat{T}_{ij}(\mathbf{F}) = \widehat{T}'_{ij}(\mathbf{QF}). \tag{4.16}$$

This is precisely the idea we used in the motivational example in Section 4.1 (and leads to the next problem).

Problem 4.3.3. It has been conjectured that the constitutive relation of a certain material has the form

$$\mathbf{T} = \mathbf{T}(\mathbf{E}),$$

where $\mathbf{E}$ is the Green Saint-Venant strain. Show that this is *not* generally consistent with the requirement of material frame indifference.

– Equation (4.14) imposes a restriction on the allowable functions $\widehat{W}(\mathbf{F})$. We now determine the most general form of $\widehat{W}$ that conforms to (4.14).

  Claim: The material frame indifference requirement (4.14) holds if and only if

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{U}) \qquad \text{where } \mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}. \tag{4.17}$$

  First suppose that (4.14) holds. Since it holds for all proper orthogonal tensors $\mathbf{Q}$ it must necessarily hold for the particular choice $\mathbf{Q} = \mathbf{R}^T$ where $\mathbf{R}$ is the rotation in the polar decomposition $\mathbf{F} = \mathbf{R}\mathbf{U}$. Thus a necessary condition for (4.14) to hold is obtained by setting $\mathbf{Q} = \mathbf{R}^T$ in it:

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{R}^T\mathbf{F}) = \widehat{W}(\mathbf{R}^T\mathbf{R}\mathbf{U}) = \widehat{W}(\mathbf{U}),$$

  where in the last step we used $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. This yields (4.17).

  Conversely, suppose that (4.17) holds for all nonsingular $\mathbf{F}$ which we may write as

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\sqrt{\mathbf{F}^T\mathbf{F}}). \tag{4.18}$$

  Since this holds for all nonsingular $\mathbf{F}$ it must hold for the tensor $\mathbf{QF}$ where $\mathbf{Q}$ is any proper orthogonal tensor. Replacing $\mathbf{F}$ by $\mathbf{QF}$ in the preceding equation yields

$$\widehat{W}(\mathbf{QF}) = \widehat{W}(\sqrt{(\mathbf{QF})^T(\mathbf{QF})}) = \widehat{W}(\sqrt{\mathbf{F}^T\mathbf{Q}^T\mathbf{QF}}) = \widehat{W}(\sqrt{\mathbf{F}^T\mathbf{F}}) = \widehat{W}(\mathbf{U}).$$

  This together with (4.17) leads to (4.14).

  We therefore conclude that (4.14) holds if and only if (4.17) holds, i.e. if the stored energy depends on the deformation through only the Lagrangian stretch tensor $\mathbf{U}$.

Problem 4.3.4. Show that the material frame indifference requirements (4.15) for the Cauchy and 1st Piola-Kirchhoff stress response functions $\widehat{\mathbf{T}}(\mathbf{F})$ and $\widehat{\mathbf{S}}(\mathbf{F})$ hold if and only if

$$\widehat{\mathbf{T}}(\mathbf{F}) = \mathbf{R}\widehat{\mathbf{T}}(\mathbf{U})\mathbf{R}^T, \qquad \widehat{\mathbf{S}}(\mathbf{F}) = \mathbf{R}\widehat{\mathbf{S}}(\mathbf{U}), \tag{4.19}$$

where $\mathbf{R}$ and $\mathbf{U}$ are the factors of $\mathbf{F}$ in the polar decomposition $\mathbf{F} = \mathbf{RU}$.

Observe from this that the constitutive relation considered in the motivation example in Section 4.1 would have been general acceptable had we started with $\mathbf{T} = \mathbf{R}\widehat{\mathbf{T}}(\mathbf{E})\mathbf{R}^T$.

– As noted previously in the discussion surrounding (2.58), there is a one-to-one relation between $\mathbf{U}$ and the right Cauchy-Green tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F} = \mathbf{U}^2$. Moreover, given $\mathbf{F}$, it is a lot easier to calculate $\mathbf{C}$ than $\mathbf{U}$. Therefore we introduce a function $\overline{W}(\mathbf{C}) = \widehat{W}(\sqrt{\mathbf{C}}) = \widehat{W}(\mathbf{U})$ and thus express the stored energy in the form

$$\widehat{W}(\mathbf{F}) = \overline{W}(\mathbf{C}) \qquad \text{where } \mathbf{C} = \mathbf{F}^T\mathbf{F}. \tag{4.20}$$

Since every Lagrangian strain tensor $\mathbf{E}$ has a one-to-one correspondence with the Lagrangian stretch tensor $\mathbf{U}$, it follows that W *can equivalently be written as a function of any Lagrangian strain tensor* $\mathbf{E}$.

By using (4.20), $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ and the chain rule, one can write the constitutive relations (4.8) for stress as (Exercise)

$$\boxed{\mathbf{S} = 2\mathbf{F}\,\frac{\partial\overline{W}}{\partial\mathbf{C}}, \qquad \mathbf{T} = \frac{2}{J}\,\mathbf{F}\,\frac{\partial\overline{W}}{\partial\mathbf{C}}\,\mathbf{F}^T.} \tag{4.21}$$

Remark: Observe that the right-hand side of $(4.21)_2$ is a symmetric tensor. Thus the value of the Cauchy stress yielded by the constitutive relation $(4.21)_2$ will be automatically symmetric.

Remark: Observe from (4.21) that $\widehat{\mathbf{S}}(\mathbf{F})$ and $\widehat{\mathbf{T}}(\mathbf{F})$ *cannot* be written in terms of $\mathbf{C}$ (or $\mathbf{U}$ or any Lagrangian strain) alone. They involve the rotational part $\mathbf{R}$ of the deformation gradient tensor as well.

Problem 4.3.5. Using the fact that the second Piola-Kirchhoff stress $\mathbf{S}^{(2)}$ is work conjugate to the Green Saint-Venant strain $\mathbf{E}$, show that the constitutive equation for it is $\mathbf{S}^{(2)} = \partial\overline{W}/\partial\mathbf{E}$.

– In the discussion above we considered subjecting a body to a rotation *after* having first deformed it, i.e. we were concerned with two deformations $\mathbf{y} = \mathbf{F}\mathbf{x}$ and $\mathbf{y} = \mathbf{Q}\mathbf{F}\mathbf{x}$. What if instead, we had rotated the body first *before* deforming it? That is, had we considered two deformations $\mathbf{y} = \mathbf{F}\mathbf{x}$ and $\mathbf{y} = \mathbf{F}\mathbf{Q}\mathbf{x}$, would we have required $\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q})$? The answer, in general, is "no". It depends on the symmetry of the material as we shall now see.

4.4 Material symmetry.

The strain energy function depends on the reference configuration:

– If we change the reference configuration, the strain energy function W changes. To see this in a simple setting consider a one-dimensional elastic bar. Suppose it has length L in some homogeneously *deformed* configuration and that the stored energy has a certain *value* in this configuration. There is nothing unique about a reference configuration. All that is required is that it be a configuration that the body *can* achieve. Accordingly consider two reference configurations and let L_1 and L_2 be the lengths of the bar in those configurations. The stretch of the bar from these respective reference configurations (to the same deformed configuration) is $\lambda_1 = L/L_1$ and $\lambda_2 = L/L_2 \neq \lambda_1$. If the material is described by a strain energy function $W(\lambda)$ that does not depend on the choice of reference configuration, then we would conclude that the energy density in the deformed body has two values, $W(\lambda_1)$ and $W(\lambda_2)$. This cannot be since we did not change the deformed configuration and the energy in the deformed body has one definite value. Thus we conclude that the function W must depend on the choice of reference configuration and so we have different strain energy functions W_1 and W_2 associated with the two reference configurations. Since W_i represents the energy per unit reference length, the total stored elastic energy in the deformed configuration can be written in the equivalent forms $L_1W_1(\lambda_1)$ and $L_2W_2(\lambda_2)$. Since these two values must be equal, it follows that the functions W_1 and W_2 must be such that

$$\frac{W_1(\lambda_1)}{\lambda_1} = \frac{W_2(\lambda_2)}{\lambda_2}. \tag{4.22}$$

We now generalize this.

– The deformation gradient tensor $\mathbf{F}$ depends on the reference configuration but the energy stored in the deformed configuration should not depend on this choice. Therefore, since

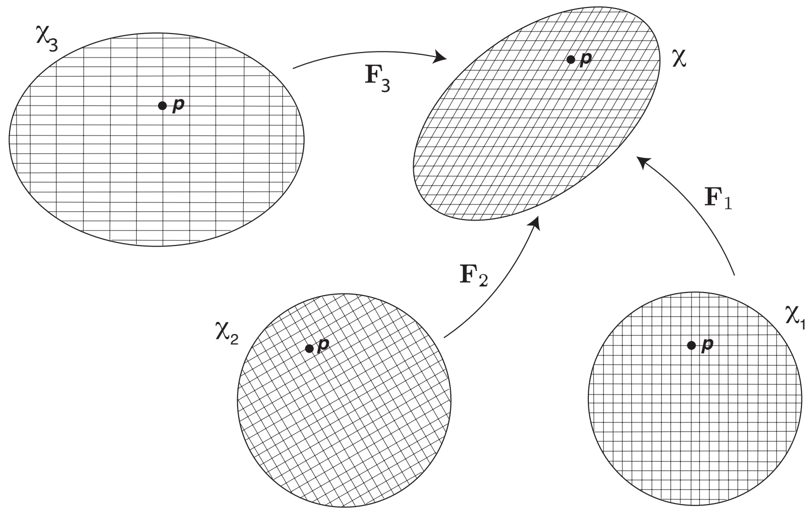


Figure 4.2: A sketch of the regions occupied by a body in a deformed configuration χ and three reference configurations χ_1, χ_2 and χ_3. The lattices are shown merely for motivational purposes where the lattice in χ_1 has been rotated to get χ_2 and stretched to get χ_3. The strain energy functions $\widehat{W}_1, \widehat{W}_2$ and $\widehat{W}_3$ with respect to these three reference configurations would be distinct in general.

$\mathbf{F}$ depends on the choice of reference configuration it is necessary that the strain energy function $\widehat{W}$ also depend on the reference configuration (in a suitable way).

This can be readily seen from Figure 4.2 which shows the regions occupied by a body in a deformed configuration χ and three reference configurations χ_1, χ_2 and χ_3. The lattices are shown for motivational purposes only where the lattice in χ_1 has been rotated to get χ_2 and stretched to get χ_3. There are three strain energy functions $\widehat{W}_1(\mathbf{F}_1), \widehat{W}_2(\mathbf{F}_2)$ and $\widehat{W}_3(\mathbf{F}_3)$ associated with these three reference configurations.

– While there is in general a different strain energy function associated with each reference configuration, if we know the strain energy function associated with one such configuration, and the gradient of the mapping from it to a second reference configuration, we can calculate the strain energy function associated with the second reference configuration. This is a consequence of the fact that the way in which $\widehat{W}$ depends on $\mathbf{F}$ must be such that the *value* of the stored energy remains unaffected by a change of reference configuration. We now show this.

Consider a (single) deformed configuration χ and let $\mathbf{F}_1$ and $\mathbf{F}_2$ denote the deformation gradient tensors of this deformed configuration with respect to two reference configurations χ_1 and χ_2. Let $\widehat{W}_1$ and $\widehat{W}_2$ be the two strain energy functions associated with the two reference configurations. Keep in mind that these function represent the energy stored per unit reference volume. Consider an infinitesimal part of volume dV_y in the deformed configuration and let the volumes of the images of this region in the respective reference configurations be dV_1 and dV_2. The energy stored in this small part can be written in either of the forms $\widehat{W}_1(\mathbf{F}_1)\,dV_1$ or $\widehat{W}_2(\mathbf{F}_2)\,dV_2$. Since the value of this stored energy cannot depend on the choice of reference configuration we must have $\widehat{W}_1(\mathbf{F}_1)\,dV_1 = \widehat{W}_2(\mathbf{F}_2)\,dV_2$. Since $dV_y = \det \mathbf{F}_1\,dV_1 = \det \mathbf{F}_2\,dV_2$, we can write this as

$$\frac{\widehat{W}_1(\mathbf{F}_1)}{\det \mathbf{F}_1} = \frac{\widehat{W}_2(\mathbf{F}_2)}{\det \mathbf{F}_2}, \tag{4.23}$$

cf. (4.22)

– Suppose the two reference configurations χ_1 and χ_2 are related by some nonsingular tensor $\mathbf{A}$ in the sense that the deformation gradient tensors $\mathbf{F}_1$ and $\mathbf{F}_2$ are related by $\mathbf{F}_1 = \mathbf{F}_2\mathbf{A}$. Then we can write (4.23) as $\widehat{W}_2(\mathbf{F}_2) = \widehat{W}_1(\mathbf{F}_2\mathbf{A})/\det \mathbf{A}$ since $\det \mathbf{F}_1 = \det(\mathbf{F}_2\mathbf{A}) = \det \mathbf{F}_2 \det \mathbf{A}$. Thus we conclude that

$$\widehat{W}_2(\mathbf{F}) = \frac{1}{\det\ \mathbf{A}}\,\widehat{W}_1(\mathbf{F}\mathbf{A}) \qquad \text{for all nonsingular } \mathbf{F}, \tag{4.24}$$

for the particular nonsingular tensor $\mathbf{A}$ that relates χ_2 to χ_1. This tells us that if we know the strain energy function $\widehat{W}_1$ associated with one reference configuration, and the gradient $\mathbf{A}$ of the mapping from it to another reference configuration, then the strain energy function $\widehat{W}_2$ associated with the second reference configuration can be determined from (4.24).

If the two reference configurations are related by a rotation, i.e. $\mathbf{A} = \mathbf{Q}$ is a proper orthogonal tensor, then

$$\widehat{W}_2(\mathbf{F}) = \widehat{W}_1(\mathbf{F}\mathbf{Q}) \qquad \text{for all nonsingular } \mathbf{F}. \tag{4.25}$$

Material symmetry:

– We now turn to a discussion of material symmetry where we restrict attention to reference configurations related by a rotation[8]. Consider the two lattices[9] associated with the two reference configurations χ_1 and χ_2 as shown in Figure 4.2. When the rotation $\mathbf{Q}$ that takes $\chi_1 \to \chi_2$ is arbitrary, these latices would be distinct in general. However for certain special

[8]In a general analysis of material symmetry, one allows for transformations between reference configurations that are more general than proper orthogonal transformations. See the discussion following (4.29). See also Section 4.2.3 of Ogden and Section 31 of Truesdell and Noll.

[9]We refer to lattices for purely motivational purposes.

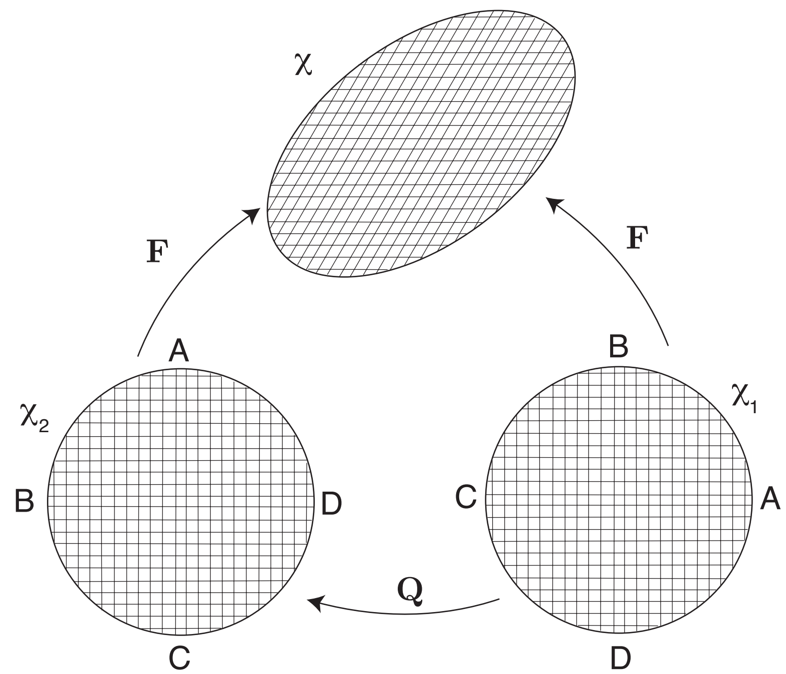


Figure 4.3: A sketch of the regions occupied by a body in a (deformed) configuration χ and two reference configurations χ_1 and χ_2. Note that the configurations χ_1 here and in Figure 4.2 are identical. The rotation $\mathbf{Q}$ in Figure 4.2 is arbitrary while here it is special. Here, it rotates the underlying square lattice through an angle $\pi/2$. The locations of 4 material points A, B, C, D in the two reference configuration are shown. Note the symmetry between the reference configurations χ_1 and χ_2 even though they are distinct. The particular transformation $\mathbf{Q}$ from $\chi_1 \to \chi_2$ here preserves the symmetry of the material.

rotations $\mathbf{Q}$, such as the 90^o rotation associated with Figure 4.3, the lattices coincide. Such a rotation $\mathbf{Q}$ *preserves material symmetry.* (Note that the reference configuration χ_1 is the same in Figures 4.2 and 4.3.)

– Thus it may so happen that two *particular* reference configurations have the same strain energy functions. In that event

$$\widehat{W}_1(\mathbf{F}) = \widehat{W}_2(\mathbf{F}) \qquad \text{for all nonsingular } \mathbf{F}, \tag{4.26}$$

and so from (4.25) and (4.26)

$$\widehat{W}_1(\mathbf{F}) = \widehat{W}_1(\mathbf{FQ}) \text{ for the \textit{particular} tensor } \mathbf{Q} \text{ relating those two configurations,} \tag{4.27}$$

and all nonsingular tensors $\mathbf{F}$. Since such a tensor $\mathbf{Q}$ preserves symmetry, the set of all rotation tensors for which (4.27) holds describes the symmetry of the material in the configuration χ_1.

– Accordingly consider a given reference configuration κ with associated strain energy function $\widehat{W}$. Let $\mathcal{G}$ denote the set of all symmetry preserving transformation of κ, i.e. the set of all rotation tensors $\mathbf{Q}$ that take this reference configuration into a configuration with the identical strain energy function:

$$\mathcal{G} = \{\mathbf{Q} : \ \mathbf{Q}\mathbf{Q}^T = \mathbf{I}, \ \det \mathbf{Q} = 1, \ \widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q}) \text{ for all nonsingular } \mathbf{F}\}.$$

This set $\mathcal{G}$ of transformations is called the *material symmetry group*[10] of the given reference configuration. Thus

$$\boxed{\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q}) \quad \text{for all } \mathbf{Q} \in \mathcal{G} \text{ and all nonsingular } \mathbf{F}.} \tag{4.28}$$

As an example, if a material has one preferred direction $\mathbf{m}_R$ (as would be the case in the presence of one family of fibers in the referential direction $\mathbf{m}_R$), the material symmetry group will contain all rotations about $\mathbf{m}_R$. We shall explore such materials – said to be *transversely isotropic* – in Chapter 6.

– The "larger" the set $\mathcal{G}$, the greater the symmetry of the reference configuration.

– Note that *symmetry is a property of the material specific to a configuration.* In general, the same body, composed of the same material, will have different symmetries in different configurations. Symmetry transformations are the particular transformations that leave the "material microstructure" invariant.

– *Terminology:* Though symmetry is a property of a material in some configuration, when there is no chance for confusion, it will be convenient (despite being imprecise) to call $\mathcal{G}$ the symmetry group of the material.

– Observe that, while the material frame indifference requirement $\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{Q}\mathbf{F})$ holds for *all* rotations $\mathbf{Q}$, the material symmetry requirement $\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q})$ holds only for the *particular* $\mathbf{Q}$'s that are in the material symmetry group. The rotation $\mathbf{Q}$ in the former is imposed on the deformed configuration, while the rotation $\mathbf{Q}$ in the latter is imposed on the reference configuration.

[10] One can readily confirm that if $\mathbf{Q}_1 \in \mathcal{G}$ and $\mathbf{Q}_2 \in \mathcal{G}$ then $\mathbf{Q}_1\mathbf{Q}_2 \in \mathcal{G}$. Moreover if $\mathbf{Q} \in \mathcal{G}$ then $\mathbf{Q}^{-1} \in \mathcal{G}$. In linear algebra, a group is a set of tensors with these two properties and so the set $\mathcal{G}$ is indeed a group in this mathematical sense.

4.4.1 Material symmetry and frame indifference: combined.

We now show that, with material frame indifference in hand, a rotation $\mathbf{Q} \in \mathcal{G}$ if and only if

$$\overline{W}(\mathbf{C}) = \overline{W}\big(\mathbf{Q}\mathbf{C}\mathbf{Q}^T\big) \quad \text{for all symmetric positive definite } \mathbf{C}. \tag{4.29}$$

– To show this, pick and fix a symmetry transformation $\mathbf{Q} \in \mathcal{G}$. Then

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q}) \quad \text{for all nonsingular } \mathbf{F}. \tag{4.30}$$

First, since this holds for all nonsingular $\mathbf{F}$ it necessarily holds with $\mathbf{F}$ replaced by $\mathbf{F}\mathbf{Q}^T$. This tells us that $\widehat{W}(\mathbf{F}\mathbf{Q}^T) = \widehat{W}(\mathbf{F}\mathbf{Q}^T\mathbf{Q}) = \widehat{W}(\mathbf{F})$:

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{Q}^T) \quad \text{for all nonsingular } \mathbf{F}. \tag{4.31}$$

(By comparing (4.31) with (4.28) we see that if $\mathbf{Q} \in \mathcal{G}$ then $\mathbf{Q}^T \in \mathcal{G}$.) Second, we turn to material frame indifference. Since (4.20) holds for all nonsingular $\mathbf{F}$ it too holds with $\mathbf{F}$ replaced by $\mathbf{F}\mathbf{Q}^T$. This yields

$$\widehat{W}(\mathbf{F}\mathbf{Q}^T) = \overline{W}\big(\mathbf{Q}\mathbf{C}\mathbf{Q}^T\big), \tag{4.32}$$

where on the right-hand side we have used $(\mathbf{F}\mathbf{Q}^T)^T\mathbf{F}\mathbf{Q}^T = \mathbf{Q}\mathbf{F}^T\mathbf{F}\mathbf{Q}^T = \mathbf{Q}\mathbf{C}\mathbf{Q}^T$. Combining (4.31) and (4.32) gives $\widehat{W}(\mathbf{F}) = \overline{W}\big(\mathbf{Q}\mathbf{C}\mathbf{Q}^T\big)$. We can now replace $\widehat{W}(\mathbf{F})$ in this by $\overline{W}(\mathbf{C})$ because of (4.20) which leads to (4.29).

Conversely if (4.29) holds, then $\mathbf{Q} \in \mathcal{G}$. (Exercise)

– Certain materials possess symmetries under "geometric transformations" that cannot be achieved by deformation. Consider for example a crystal lattice that remains invariant under a *reflection* in a plane perpendicular to a direction $\mathbf{n}_\mathrm{R}$ so that the distinction between the lattices before and after reflection cannot be detected. Recall from Problem 1.9 that this reflection is characterized by the improper orthogonal tensor $\mathbf{Q}_1 = \mathbf{I} - 2\mathbf{n}_\mathrm{R} \otimes \mathbf{n}_\mathrm{R}$. Since $\det \mathbf{Q}_1 = -1 < 0$ it cannot be achieved by deformation. At the same time, observe by specializing Problem 1.10 that the proper orthogonal tensor $\mathbf{Q}_2 = -\mathbf{I} + 2\mathbf{n}_\mathrm{R} \otimes \mathbf{n}_\mathrm{R} = -\mathbf{Q}_1$ represents a 180^o-*rotation* about the direction $\mathbf{n}_\mathrm{R}$. For it, $\det \mathbf{Q}_2 = +1 > 0$. Now observe that if (4.29) holds for $\mathbf{Q} = \mathbf{Q}_2$ then it necessarily holds for $\mathbf{Q} = -\mathbf{Q}_2 = \mathbf{Q}_1$. Therefore if mechanical testing cannot detect a difference in elastic properties before and after a 180^o-rotation about $\mathbf{n}_\mathrm{R}$, then it necessarily cannot detect a difference in properties before and after a reflection in the plane perpendicular to $\mathbf{n}_\mathrm{R}$. A symmetry group can be extended in this way to accommodate reflection symmetries.

In general, if (4.29) holds for a rotation tensor $\mathbf{Q}$, it automatically holds for the reflection tensor $-\mathbf{Q}$ and vice versa.

4.4.2 Isotropic material.

We now turn to "isotropic materials". (As mentioned above, symmetry is a property of a material in some configuration, but when there is no chance for confusion, we use the (inexact) terminology that attributes symmetry to the material. Accordingly here, what we really mean is that we are considering a reference configuration in which the material is isotropic.)

Some anisotropic materials will be considered in Chapter 6.

– **Isotropic material:** If the material symmetry group $\mathcal{G}$ contains *all* proper orthogonal tensors, we say the material is isotropic. Thus by (4.29), a material is *isotropic* if

$$\overline{W}(\mathbf{C}) = \overline{W}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T), \tag{4.33}$$

for all proper orthogonal tensors $\mathbf{Q}$ and all symmetric positive definite tensors $\mathbf{C}$. In this event it follows from the result in Problem 1.33 that W has the representation

$$\boxed{W = \widetilde{W}\big(I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C})\big)} \tag{4.34}$$

where

$$I_1(\mathbf{C}) = \operatorname{tr}\mathbf{C}, \qquad I_2(\mathbf{C}) = \frac{1}{2}\big[(\operatorname{tr}\mathbf{C})^2 - \operatorname{tr}(\mathbf{C}^2)\big], \qquad I_3(\mathbf{C}) = \det\,\mathbf{C}, \tag{4.35}$$

are the principal scalar invariants of the right Cauchy-Green tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F} = \mathbf{U}^2$.

Observe by taking $\mathbf{Q} = \mathbf{R}$ in (4.33) that $\overline{W}(\mathbf{C}) = \overline{W}(\mathbf{B})$ for an isotropic material.

– The constitutive relations (4.21) for stress can now be further reduced using (4.34), the chain rule and

$$\frac{\partial I_1}{\partial \mathbf{C}} = \mathbf{I}, \qquad \frac{\partial I_2}{\partial \mathbf{C}} = I_1\,\mathbf{I} - \mathbf{C} \qquad \frac{\partial I_3}{\partial \mathbf{C}} = J^2\mathbf{C}^{-1}, \tag{4.36}$$

(see equation (1.185) in Problem 1.8.4). This leads to the following constitutive relations for an isotropic elastic material:

$$\left.\begin{aligned} \mathbf{T} &= 2J\frac{\partial \widetilde{W}}{\partial I_3}\mathbf{I} + \frac{2}{J}\left[\frac{\partial \widetilde{W}}{\partial I_1} + I_1\frac{\partial \widetilde{W}}{\partial I_2}\right]\mathbf{B} - \frac{2}{J}\frac{\partial \widetilde{W}}{\partial I_2}\mathbf{B}^2, \\ \mathbf{S} &= 2\mathbf{F}\left[I_3\frac{\partial \widetilde{W}}{\partial I_3}\mathbf{C}^{-1} + \left[\frac{\partial \widetilde{W}}{\partial I_1} + I_1\frac{\partial \widetilde{W}}{\partial I_2}\right] - \frac{\partial \widetilde{W}}{\partial I_2}\mathbf{C}\right]. \end{aligned}\right\} \tag{4.37}$$

– In the undeformed configuration where $\mathbf{F} = \mathbf{C} = \mathbf{B} = \mathbf{I}$, the principal scalar invariants have the values $I_1 = 3, I_2 = 3, I_3 = 1$. Setting $\mathbf{F} = \mathbf{B} = \mathbf{I}$ and $I_1 = I_2 = 3, I_3 = J = 1$ in (4.37) we see that the stress in the reference configuration is

$$\mathbf{T} = \mathbf{S} = 2\left[\frac{\partial\widetilde{W}}{\partial I_1} + 2\frac{\partial\widetilde{W}}{\partial I_2} + \frac{\partial\widetilde{W}}{\partial I_3}\right]\mathbf{I} \qquad \text{evaluated at } (I_1, I_2, I_3) = (3, 3, 1). \tag{4.38}$$

Note that this stress is hydrostatic as a consequence of the material being isotropic in the reference configuration. If the reference configuration is stress-free, then it is necessary that

$$\frac{\partial\widetilde{W}}{\partial I_1} + 2\frac{\partial\widetilde{W}}{\partial I_2} + \frac{\partial\widetilde{W}}{\partial I_3} = 0 \qquad \text{at} \quad (I_1, I_2, I_3) = (3, 3, 1). \tag{4.39}$$

– Note from $(4.37)_1$ that $\mathbf{T}$ and $\mathbf{B}$ *are coaxial for an isotropic material*, i.e. they have the same principal directions. Therefore we can write

$$\mathbf{T} = \tau_1\, \boldsymbol{\ell}_1 \otimes \boldsymbol{\ell}_1 + \tau_2\, \boldsymbol{\ell}_2 \otimes \boldsymbol{\ell}_2 + \tau_3\, \boldsymbol{\ell}_3 \otimes \boldsymbol{\ell}_3, \tag{4.40}$$

where the τ_i's are the principal Cauchy stresses and the $\boldsymbol{\ell}_i$'s are the principal directions of $\mathbf{B}$ (and $\mathbf{T}$).

– In terms of the principal stretches $\lambda_1, \lambda_2, \lambda_3$ one can write the principal scalar invariants of $\mathbf{C}$ (or $\mathbf{B}$) as

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \qquad I_2 = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2, \qquad I_3 = \lambda_1^2\lambda_2^2\lambda_3^2. \tag{4.41}$$

It follows from (4.34) and (4.41) that the strain energy function for an isotropic material can be written in the form[11]

$$W = W^*(\lambda_1, \lambda_2, \lambda_3). \tag{4.42}$$

Since the I_i's remain invariant if any two of the λ's are switched, e.g. $\lambda_1 \leftrightarrow \lambda_2$, the constitutive response function W^* must also remain invariant if any two of its arguments are switched:

$$W^*(\lambda_1, \lambda_2, \lambda_3) = W^*(\lambda_2, \lambda_1, \lambda_3) = W^*(\lambda_1, \lambda_3, \lambda_2). \tag{4.43}$$

– We can now write the constitutive relation for $\mathbf{T}$ in terms of W^* by using the chain rule. For this we need an expression for $\partial\lambda_i/\partial\mathbf{C}$. Let λ_i^2 and $\mathbf{r}_i$ be an eigenvalue and corresponding

[11] Before imposing material symmetry, we had $W = \overline{W}(\mathbf{C})$. By the spectral representation we know that $\mathbf{C}$ is fully determined by its eigenvalues and eigenvectors, and so we knew at that stage that $W = W(\lambda_1, \lambda_2, \lambda_3, \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. What isotropy says is that W does not depend on the eigenvectors of $\mathbf{C}$.

eigenvector of $\mathbf{C}$. One can show by differentiating $\mathbf{C}\mathbf{r}_i = \lambda_i^2 \mathbf{r}_i$ (no sum on i) and using the fact that $\mathbf{r}_i$ is a unit vector that

$$\frac{\partial \lambda_i}{\partial \mathbf{C}} = \frac{1}{2\lambda_i} \mathbf{r}_i \otimes \mathbf{r}_i \qquad \text{(no sum on } i\text{);} \tag{4.44}$$

see Problem 2.22. The constitutive relation for the Cauchy stress can now be rewritten by changing $\overline{W}$ to W^* in $(4.21)_2$, using the chain rule, (4.44) and $\mathbf{F}\mathbf{r}_i = \lambda_i \boldsymbol{\ell}_i$. This leads to

$$\mathbf{T} = \sum_{i=1}^{3} \frac{\lambda_i}{J} \frac{\partial W^*}{\partial \lambda_i} \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i. \tag{4.45}$$

It follows from (4.40) and (4.45) that the constitutive relation for the principal Cauchy stresses is

$$\tau_1 = \frac{\lambda_1}{\lambda_1 \lambda_2 \lambda_3} \frac{\partial W^*}{\partial \lambda_1}, \qquad \tau_2 = \frac{\lambda_2}{\lambda_1 \lambda_2 \lambda_3} \frac{\partial W^*}{\partial \lambda_2}, \qquad \tau_3 = \frac{\lambda_3}{\lambda_1 \lambda_2 \lambda_3} \frac{\partial W^*}{\partial \lambda_3}. \tag{4.46}$$

Thus given $\mathbf{F}$, one can find the principal values of $\mathbf{T}$ from (4.46) and the principal directions of $\mathbf{T}$ by finding the principal directions of $\mathbf{B} = \mathbf{F}\mathbf{F}^T$:

– Since the 1st Piola-Kirchhoff stress tensor $\mathbf{S}$ is not symmetric in general, it may not have principal values. However a calculation just like the one above but now applied to $(4.21)_1$ allows us to write the constitutive relation for $\mathbf{S}$ as

$$\mathbf{S} = \sum_{k=1}^{3} \frac{\partial W^*}{\partial \lambda_k} \boldsymbol{\ell}_k \otimes \mathbf{r}_k = \sum_{k=1}^{3} \frac{\tau_k J}{\lambda_k} \boldsymbol{\ell}_k \otimes \mathbf{r}_k, \tag{4.47}$$

where $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ and $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ are the eigenvectors of the Lagrangian and Eulerian stretch tensors $\mathbf{U}$ and $\mathbf{V}$ respectively.

We will frequently consider particular deformations where $\mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3$. In this case the bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ and $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ all coincide and then

$$\mathbf{S} = \sigma_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \sigma_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \sigma_3 \mathbf{e}_3 \otimes \mathbf{e}_3 \qquad \text{where } \sigma_i = \frac{\partial W^*}{\partial \lambda_i}.$$

When we consider spherically symmetric problems for isotropic materials in Chapter 5, we will find that the bases $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ coincide, that they are the principal bases for $\mathbf{B}$ and $\mathbf{T}$, and that $[F]$ and $[S]$ are diagonal in these bases.

– In Problem 4.13 you are asked to show that, for the Biot stress tensor introduced in (3.81),

$$\mathbf{S}^{(1)} = \sum_{i=1}^{3} \frac{\partial W^*}{\partial \lambda_i} \mathbf{r}_i \otimes \mathbf{r}_i = \sum_{i=1}^{3} \frac{\tau_i J}{\lambda_i} \mathbf{r}_i \otimes \mathbf{r}_i. \tag{4.48}$$

Equation $(4.48)_1$ is a consequence of $\mathbf{S}^{(1)}$ being work conjugate to $\mathbf{U}$ (as established in Problem 3.29.)

Perhaps it is worth mentioning that if one is to conduct laboratory experiments to find W^*, it is necessary to carry out experiments that probe various paths of $\lambda_1, \lambda_2, \lambda_3$-space. Carrying out, for example a uniaxial tension test alone, would only probe a single path in this space.

Problem 4.4.1. Blatz and Ko proposed the following strain energy function for the foam rubber material they studied in their experiments:

$$\widetilde{W}(I_1, I_2, I_3) = \frac{\mu}{2}\left(\frac{I_2}{I_3} + 2\sqrt{I_3} - 5\right), \qquad \mu > 0. \tag{i}$$

Here μ is a material constant. See page 294 for a reference to their paper. Determine the response of this material in uniaxial stress and simple shear. What are the values of the Young's modulus, Poisson's ratio and shear modulus of this material at infinitesimal deformations?

See also Problem 4.7.1 concerning the bending of a block made of a Blatz-Ko material.

Solution: In terms of the principal stretches, we find from (4.41) and (i) that

$$W^*(\lambda_1, \lambda_2, \lambda_3) = \frac{\mu}{2}\big(\lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2} + 2\lambda_1\lambda_2\lambda_3 - 5\big). \tag{ii}$$

The constitutive equation for $\mathbf{T}$ is given by (i) and $(4.37)_1$ to be

$$\mathbf{T} = \frac{\mu}{J^3}\Big[(J^3 - I_2)\mathbf{I} + I_1\mathbf{B} - \mathbf{B}^2\Big], \tag{iii}$$

and from (4.46) and (ii) the principal stresses can be written in terms of the principal stretches as

$$\tau_k = \frac{\lambda_k}{J}\frac{\partial W^*}{\partial \lambda_k} = \mu\Big[1 - \lambda_k^{-2}/J\Big], \qquad J = \lambda_1\lambda_2\lambda_3. \tag{iv}$$

Consider a state of uniaxial stress in the $\mathbf{e}_1$-direction:

$$\tau_1 = \tau, \qquad \tau_2 = \tau_3 = 0, \tag{v}$$

and assume the deformation to be a homogeneous pure stretch $\mathbf{y} = \mathbf{F}\mathbf{x}$ with

$$\mathbf{F} = \lambda\mathbf{e}_1 \otimes \mathbf{e}_1 + \Lambda(\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3). \tag{vi}$$

Our aim is to calculate the longitudinal stretch λ (in the direction of the applied stress) and the transverse stretch Λ.

The principal stretches are $\lambda_1 = \lambda, \lambda_2 = \lambda_3 = \Lambda$ and the Jacobian determinant is

$$J = \lambda_1\lambda_2\lambda_3 = \lambda\Lambda^2. \qquad (vii)$$

From (iv) and $\tau_2 = 0$ we get

$$1 - \lambda_2^{-2}/J = 0 \quad \overset{(vii)}{\Rightarrow} \quad 1 - \Lambda^{-2}/(\lambda\Lambda^2) = 0 \quad \Rightarrow \quad \Lambda = \lambda^{-1/4}. \quad \Box \qquad (viii)$$

Likewise from (iv) with $k = 1$ we find

$$\tau = \mu\Big[1 - \lambda_1^{-2}/J\Big] \overset{(vii)}{=} \mu\Big[1 - \lambda^{-2}/(\lambda\Lambda^2)\Big] \overset{(viii)}{=} \mu(1 - \lambda^{-5/2}). \quad \Box \qquad (ix)$$

This relation between τ and λ is monotonic with $\tau \to -\infty$ as $\lambda \to 0^+$, $\tau = 0$ when $\lambda = 1$, and $\tau \to \mu$ as $\lambda \to \infty$.

Observe that, since $J = \lambda\Lambda^2 \overset{(viii)}{=} \lambda^{1/2}$, one has $\lambda_1 = J^2$ and $\lambda_2 = J^{-1/2}$. Therefore from (ii) it follows that $W = \frac{\mu}{2}(J^{-4} + 2J^{-1} + J - 5)$. We see from this that $W \to \infty$ when both $J \to 0^+$ and $J \to \infty$ in uniaxial stress (i.e. at extreme deformations).

For infinitesimal deformations we write the principal stretches as

$$\lambda_1 = \lambda = 1 + \varepsilon_1, \qquad \lambda_2 = \Lambda = 1 + \varepsilon_2, \qquad (x)$$

where the principal strains are small: $\varepsilon_1 \ll 1, \varepsilon_2 \ll 1$. Substituting $(x)_1$ into (ix) and linearizing gives

$$\tau = \mu\Big(1 - (1+\varepsilon_1)^{-5/2}\Big) \doteq \mu\left(1 - \left(1 - \frac{5}{2}\varepsilon_1\right)\right) = \frac{5}{2}\mu\varepsilon_1 \quad \Rightarrow \quad \tau/\varepsilon_1 = \frac{5}{2}\mu, \quad \Box \quad (xi)$$

where we have used the Taylor expansion $(1+\varepsilon)^m = 1 + m\varepsilon + \mathrm{O}(\varepsilon^2)$ as $\varepsilon \to 0$. Therefore the Young's modulus of this material is $5\mu/2$. Similarly substituting (x) into $(viii)$ and linearizing gives

$$1 + \varepsilon_2 = (1+\varepsilon_1)^{-1/4} \doteq 1 - \frac{1}{4}\varepsilon_1 \quad \Rightarrow \quad -\varepsilon_2/\varepsilon_1 = 1/4. \quad \Box \qquad (xii)$$

The Poisson's ratio of this material is therefore 0.25.

Alternatively we could have differentiated (ix) with respect to λ and used the fact that the Young's modulus equals $d\tau/d\lambda$ evaluated at $\lambda = 1$. Likewise, the Poisson's ratio is $-d\Lambda/d\lambda$ evaluated at $\lambda = 1$ which can be calculated from $(viii)$.

We can find the first Piola Kirchhoff stress by using $\mathbf{T} = J^{-1}\mathbf{S}\mathbf{F}^T$. However it is easier (and more insightful) to use the following calculation: suppose that the cross section of the undeformed specimen (normal to the stressing direction) is 1×1. The cross section of the deformed specimen is then $\Lambda \times \Lambda$. Therefore the force acting on the cross section can be written as $S_{11} \times 1$ and equivalently as $T_{11} \times \Lambda^2$. Therefore

$$S_{11} = T_{11}\Lambda^2 = \tau\Lambda^2 \overset{(viii),(ix)}{=} \mu(\lambda^{-1/2} - \lambda^{-3}). \qquad (xiii)$$

Observe that the relation between S_{11} and λ is not monotonic. As λ increase, the stress S_{11} first increases, then reaches a maximum value at $\lambda = 6^{2/5}$ and then decreases with $S_{11} \to 0$ as $\lambda \to \infty$.

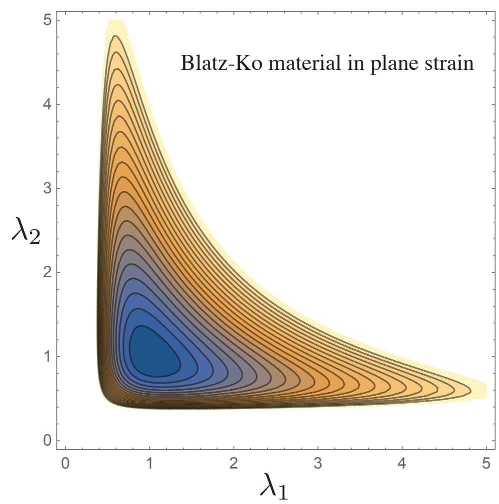


Figure 4.4: Constant energy contours for the Blatz-Ko material (ii) in plane strain $(\lambda_3 = 1)$. The reference configuration corresponds to the local minimum at $(\lambda_1, \lambda_2) = (1, 1)$.

Now consider a simple shear deformation with shearing direction $\mathbf{e}_1$ and glide plane normal $\mathbf{e}_2$:

$$y_1 = x_1 + kx_2, \qquad y_2 = x_2, \qquad y_3 = x_3. \qquad (xiv)$$

The deformation gradient tensor, left Cauchy-Green deformation tensor and its square are

$$\begin{aligned}
&\mathbf{F} = \mathbf{I} + k\,\mathbf{e}_1 \otimes \mathbf{e}_2,\\
&\mathbf{B} = \mathbf{F}\mathbf{F}^T = \mathbf{I} + k^2\,\mathbf{e}_1 \otimes \mathbf{e}_1 + k(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1), \qquad (xv)\\
&\mathbf{B}^2 = \mathbf{I} + (3k^2 + k^4)\,\mathbf{e}_1 \otimes \mathbf{e}_1 + k^2\,\mathbf{e}_2 \otimes \mathbf{e}_2 + (2k + k^3)(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1),
\end{aligned}$$

and so the principal scalar invariants of $\mathbf{B}$ are

$$I_1 = \mathrm{tr}\,\mathbf{B} = 3 + k^2, \qquad I_2 = \frac{1}{2}\big[(\mathrm{tr}\,\mathbf{B})^2 - \mathrm{tr}\,(\mathbf{B}^2)\big] = 3 + k^2 \qquad I_3 = \det\,\mathbf{B} = 1. \qquad (xvi)$$

Substituting (xv) and (xvi) into the constitutive relation (iii) and simplifying gives

$$T_{12} = \mu k, \qquad T_{22} = -\mu k^2, \qquad T_{11} = T_{33} = T_{13} = T_{23} = 0.$$

Observe that the relation between the shear stress T_{12} and the amount of shear k is linear for all deformations. The shear modulus is μ. Observe also that the normal stress $T_{22} \neq 0$ (in contrast to the linear theory for infinitesimal deformations). For small k this term is $\mathrm{O}(k^2)$ and so is an order of magnitude smaller than the shear stress. For large k however this is no longer true.

4.5 Materials with Internal Constraints.

Thus far in this chapter we have assumed that the body under consideration can undergo *any* deformation at all (provided it is subjected to suitable body forces and surface tractions). Sometimes, it is reasonable and convenient to idealize the body and permit it to *only* undergo motions of a certain restricted class. For example, a *rigid body* can only undergo rigid motions, i.e. motions in which

$$\mathbf{F}^T(\mathbf{x}, t)\mathbf{F}(\mathbf{x}, t) = \mathbf{I} \qquad \text{for all } \mathbf{x} \text{ and } t;$$

an *incompressible body* can only undergo volume-preserving motions, i.e. motions in which

$$\det \mathbf{F}(\mathbf{x}, t) = 1 \qquad \text{for all } \mathbf{x} \text{ and } t;$$

a body that is inextensible in a certain (referential) direction $\mathbf{m}_\mathrm{R}$ can only undergo motions in which

$$|\mathbf{F}(\mathbf{x}, t)\mathbf{m}_\mathrm{R}| = 1 \qquad \text{for all } \mathbf{x} \text{ and } t.$$

All of these idealizations constrain the set of possible deformation gradient tensors. Note that this is part of modeling the material's constitutive behavior.

These constraints (and several others) can be described by equations of the form

$$\widehat{\phi}(\mathbf{F}(\mathbf{x}, t)) = 0 \qquad \text{for all } \mathbf{x} \text{ and } t. \qquad (4.49)$$

For example the aforementioned constraints of rigidity, incompressibility and inextensibility correspond to the special cases

$$\widehat{\phi}(\mathbf{F}) = \mathbf{F}^T\mathbf{F} - \mathbf{I}, \qquad \widehat{\phi}(\mathbf{F}) = \det \mathbf{F} - 1, \qquad \widehat{\phi}(\mathbf{F}) = \mathbf{F}\mathbf{m}_{\mathrm{R}} \cdot \mathbf{F}\mathbf{m}_{\mathrm{R}} - 1, \tag{4.50}$$

respectively.

We now turn to the *stress* in a constrained body. In order to explain the basic idea, consider as an example, a spherical body composed of an incompressible isotropic material. It is subjected to a uniform radial pressure p on its boundary. Since the geometry, the material and the loading are all symmetric, let us restrict attention to spherically symmetric deformations. Thus the body must remain spherical when p is applied. However, due to incompressibility, the radius of the spherical body, and in fact the radius of every spherical surface within the body, cannot change. (a) Therefore irrespective of the value of p, incompressibility implies that the deformation of the sphere must be the trivial one: $\widehat{\mathbf{y}}(\mathbf{x}, t) = \mathbf{x}$ and therefore that $\mathbf{F}(\mathbf{x}, t) = \mathbf{I}$ no matter what the value of p. (b) The stress on the other hand would certainly depend on the value of the applied pressure p and will change as the value of p changes. (c) Thus the stress $\mathbf{T}$ does not vanish though $\mathbf{F} = \mathbf{I}$, and so we conclude that the stress is *not* completely determined by the deformation gradient $\mathbf{F}$, or equivalently, *different stress fields can correspond to the same deformation.* This contradicts our earlier assumption that the stress is completely determined by the deformation gradient. We must therefore modify this assumption when considering a constrained body. We choose to do this by allowing a part of the stress to be determined by the deformation and the other to be indeterminate (as far as the constitutive relation is concerned). (d) Finally, observe that the pressure does no work since the boundary does not displace and so the energy stored in the body does not increase as p increases.

Mathematically, we now assume that

$$\mathbf{S} = \widehat{\mathbf{S}}(\mathbf{F}) + \mathbf{N}, \tag{4.51}$$

where $\widehat{\mathbf{S}}(\mathbf{F})$ is the constitutively determined part of $\mathbf{S}$, and $\mathbf{N}$ is the part that arises as a reaction to the constraint. We further assume that the reactive stress $\mathbf{N}$ does no work in the sense that

$$\mathbf{N} \cdot \dot{\mathbf{F}} = 0. \tag{4.52}$$

The strain energy function on the other hand continues to have the form

$$W = \widehat{W}(\mathbf{F}),$$

and the rate of working of the stress is related to the rate of increase of stored energy by

$$\mathbf{S} \cdot \dot{\mathbf{F}} = \dot{W}.$$

Our immediate goal is to determine the form of $\mathbf{N}$. If (4.52) held for all $\dot{\mathbf{F}}$ then we would have $\mathbf{N} = \mathbf{0}$. However it only holds in all allowable motions (i.e. all motions consistent with the constraint). To determine the restriction placed on $\dot{\mathbf{F}}$ by constraint (4.49), we note that (4.49) holds at all points $\mathbf{x}$ in the body and at all times t. Differentiating it with respect to t gives $\dot{\phi} = 0$ whence

$$\frac{\partial \widehat{\phi}}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} = 0, \tag{4.53}$$

where $\partial\widehat{\phi}/\partial\mathbf{F}$ is the tensor with cartesian components $\partial\widehat{\phi}/\partial F_{ij}$. Thus (4.52) must hold for all $\dot{\mathbf{F}}$ that conform to (4.53).

Problem 1.49 describes the following algebraic result: if a vector $\mathbf{a}_1$ is orthogonal to all vectors that are orthogonal to a vector $\mathbf{a}_2$, then $\mathbf{a}_2$ must be parallel to $\mathbf{a}_1$, i.e. if $\mathbf{a}_1 \cdot \mathbf{x} = 0$ for all $\mathbf{x}$ for which $\mathbf{a}_2 \cdot \mathbf{x} = 0$, then there is a scalar q such that $\mathbf{a}_2 = -q\mathbf{a}_1$. The result in Problem 1.49 applied to an N-dimensional vector space and so, in particular, holds for the 9-dimensional vector space of all linear transformations. Thus if $\mathbf{A}_1 \cdot \mathbf{X} = 0$ for all tensors $\mathbf{X}$ for which $\mathbf{A}_2 \cdot \mathbf{X} = 0$, then there is a scalar q such that $\mathbf{A}_2 = -q\mathbf{A}_1$. Using this result with $\mathbf{X} = \dot{\mathbf{F}}, \mathbf{A}_2 = \mathbf{N}$ and $\mathbf{A}_1 = \partial\widehat{\phi}/\partial\mathbf{F}$, leads to[12]

$$\mathbf{N} = -q\, \frac{\partial \widehat{\phi}}{\partial \mathbf{F}}. \tag{4.54}$$

The stress field $\mathbf{N}(\mathbf{x})$ that arises in reaction to the constraint is referred to as the **reactive (or reaction) stress**. It should be noted that $q(\mathbf{x})$ is a scalar *field* and is not a constant in general.

The constitutive equation for the 1st Piola-Kirchhoff stress $\mathbf{S}$ for a constrained material is therefore taken to be

$$\mathbf{S} = \frac{\partial \widehat{W}}{\partial \mathbf{F}} - q\, \frac{\partial \widehat{\phi}}{\partial \mathbf{F}}. \tag{4.55}$$

The corresponding relation for the Cauchy stress tensor $\mathbf{T}$ is

$$\mathbf{T} = \frac{1}{J} \frac{\partial \widehat{W}}{\partial \mathbf{F}} \mathbf{F}^T - \frac{q}{J} \frac{\partial \widehat{\phi}}{\partial \mathbf{F}} \mathbf{F}^T. \tag{4.56}$$

[12]Many authors use the symbol p instead of q. We shall reserve p to denote the applied pressure on a body such as an inflated tube.

Note that the theory now involves an additional scalar field $q(\mathbf{x})$, but we also have an additional scalar equation $\phi(\mathbf{F}) = 0$ at our disposal. Observe also that (4.55) can be written as

$$\mathbf{S} = \frac{\partial}{\partial \mathbf{F}}\left(\widehat{W} - q\,\widehat{\phi}\right), \tag{4.57}$$

from which we see that we have, essentially, added the constraint ϕ to W using a Lagrange multiplier q.

As an example, consider an *incompressible body* in which case

$$\widehat{\phi}(\mathbf{F}) = \det \mathbf{F} - 1.$$

Differentiating this with respect to $\mathbf{F}$ using the formula (1.202) from Problem 1.45 gives

$$\frac{\partial \widehat{\phi}}{\partial \mathbf{F}} = (\det \mathbf{F})\,\mathbf{F}^{-T} \;\stackrel{J=1}{=}\; \mathbf{F}^{-T}. \tag{4.58}$$

Thus (4.55), (4.56) and (4.58) yield

$$\mathbf{S} = \frac{\partial \widehat{W}}{\partial \mathbf{F}} - q\,\mathbf{F}^{-T}, \qquad \mathbf{T} = \frac{\partial \widehat{W}}{\partial \mathbf{F}}\,\mathbf{F}^{T} - q\,\mathbf{I}. \tag{4.59}$$

Observe that the part of the Cauchy stress arising in reaction to the constraint is hydrostatic.

Exercise: In Problem 4.19 you are asked to derive the constitutive relations for $\mathbf{S}$ and $\mathbf{T}$ in the presence of the inextensibility constraint $|\mathbf{F}\mathbf{m}_{\mathrm{R}}| = 1$ and physically interpret the reactive part of the Cauchy stress.

Frame indifference: It is natural to require (4.49) to be frame indifferent. That is, if a deformation gradient tensor $\mathbf{F}$ obeys the constraint (4.49), a subsequent rigid rotation should not lead to a violation of the constraint. This requires that $\phi(\mathbf{F}) = \phi(\mathbf{QF})$ for all nonsingular tensors $\mathbf{F}$ and all rotations $\mathbf{Q}$. The earlier discussion in Section 4.3 can be readily adapted to the present context to show that (4.49) is frame indifferent if and only if the constraint can be expressed in the form

$$\phi(\mathbf{U}) = 0 \tag{4.60}$$

where $\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}$. Exercise: Verify that the constraints (4.50) can be written in this way.

Problem 4.5.1. Show that the reactive stress to be added to the Biot stress is $q\,\partial\phi/\partial\mathbf{U}$.

Material symmetry: Some care must be taken when analyzing material symmetry. Recall that we previously said that a proper orthogonal tensor $\mathbf{Q}$ is in the material symmetry group if $W(\mathbf{F}) = W(\mathbf{FQ})$ for all nonsingular $\mathbf{F}$. Now however we must limit attention to those deformation gradient tensors that obey the constraint $\phi(\mathbf{F}) = 0$. Thus a symmetry transformation $\mathbf{Q}$ must be such that *both* $W(\mathbf{F}) = W(\mathbf{FQ})$ *and* $\phi(\mathbf{FQ}) = 0$ hold. Material symmetry must therefore be compatible with the constraint. The reader is referred to Section 6.3 of Steigmann for a discussion.

In particular, Steigmann shows that an isotropic material cannot be inextensible. Therefore we will not discuss such materials in any detail in this chapter. We shall do so in Chapter 6 where we discuss anisotropic materials. On the other hand an isotropic material can be incompressible and so we now proceed to consider such materials.

In the case of an isotropic incompressible material the analysis proceeds as for an unconstrained material. In particular, one finds that the strain energy function $\widehat{W}(\mathbf{F})$ depends on the deformation only through the principal scalar invariants of $\mathbf{C}$. However, since $I_3(\mathbf{C}) = \det\,\mathbf{C} = (\det\,\mathbf{F})^2 = 1$ due to incompressibility, there are only 2 nontrivial invariants and the energy takes the form $\widetilde{W}(I_1, I_2)$. The stress tensors $\mathbf{T}$ and $\mathbf{S}$ are now found to be related to the deformation through

$$\mathbf{T} = -q\,\mathbf{I} \;+\; 2\left[\frac{\partial\widetilde{W}}{\partial I_1} + I_1\frac{\partial\widetilde{W}}{\partial I_2}\right]\mathbf{B} \;-\; 2\frac{\partial\widetilde{W}}{\partial I_2}\mathbf{B}^2, \tag{4.61}$$

$$\mathbf{S} = -q\,\mathbf{F}^{-T} \;+\; 2\left[\frac{\partial\widetilde{W}}{\partial I_1} + I_1\frac{\partial\widetilde{W}}{\partial I_2}\right]\mathbf{F} \;-\; 2\frac{\partial\widetilde{W}}{\partial I_2}\mathbf{BF}. \tag{4.62}$$

If the strain energy function is expressed in terms of the principal stretches,

$$W = W^*(\lambda_1, \lambda_2, \lambda_3), \qquad \lambda_1\lambda_2\lambda_3 = 1, \tag{4.63}$$

with W^* being invariant if any two of its arguments are switched,

$$W^*(\lambda_1, \lambda_2, \lambda_3) = W^*(\lambda_2, \lambda_1, \lambda_3) = W^*(\lambda_1, \lambda_3, \lambda_2) = \dots\;, \tag{4.64}$$

then the principal Cauchy stress components can be written as

$$T_{11} = \lambda_1\frac{\partial W^*}{\partial\lambda_1} - q, \qquad T_{22} = \lambda_2\frac{\partial W^*}{\partial\lambda_2} - q, \qquad T_{33} = \lambda_3\frac{\partial W^*}{\partial\lambda_3} - q. \tag{4.65}$$

A strain energy function $W(\mathbf{F})$ for an incompressible material is only defined on the set of all nonsingular tensors with $\det \mathbf{F} = 1$. In view of the constraint $\det \mathbf{F} = 1$, one has to explain what one means by the partial derivative $\partial W/\partial \mathbf{F}$ that enters into the constitutive equation for stress. The usual approach is to consider the following function $W^o(\mathbf{F})$ defined on the set of *all* nonsingular tensors with positive determinant:

$$W^o(\mathbf{F}) = W\left(\frac{\mathbf{F}}{(\det \mathbf{F})^{1/3}}\right). \tag{4.66}$$

Observe that the tensor argument of W on the right-hand side of (4.66) has determinant one even when the determinant of $\mathbf{F}$ is not unity. Moreover, note that $W(\mathbf{F}) = W^o(\mathbf{F})$ on the subset of tensors with $\det \mathbf{F} = 1$. Thus $W^o(\mathbf{F})$ is an extension of $W(\mathbf{F})$ to the larger set of all nonsingular tensors with positive determinant. Then by $\partial W/\partial \mathbf{F}$ we mean $\partial W^o/\partial \mathbf{F}$.

4.6 Response of Isotropic Elastic Materials.

In this section we examine the response of an isotropic elastic material in *uniaxial tension, simple shear and plane stress biaxial stretch.* We keep the strain energy function, $W(I_1, I_2)$ or $W(I_1, I_2, I_3)$ as the case may be, general, and so the results hold for any isotropic elastic material. In Section (4.7) we will turn to some specific constitutive relations, but before doing so, in Section 4.6.3, we will make some brief remarks on restrictions one might impose on the strain energy function for physical and mathematical reasons.

All of the deformations we consider in this section are homogeneous in that the deformation gradient tensor is uniform throughout the body. Therefore the stress field $\mathbf{S}(\mathbf{x})$ (resp. $\mathbf{T}(\mathbf{y})$) is also uniform and does not depend on $\mathbf{x}$ (resp. $\mathbf{y}$). The equilibrium equation without body forces, $\mathrm{Div}\,\mathbf{S} = \mathbf{o}$ (resp. $\mathrm{div}\,\mathbf{T} = \mathbf{o}$) therefore holds automatically.

It is convenient to record again the constitutive relations. For an (unconstrained) isotropic elastic material we have

$$\left.\begin{aligned} \mathbf{T} &= 2JW_3\,\mathbf{I} + \frac{2}{J}\left[W_1 + I_1W_2\right]\mathbf{B} - \frac{2}{J}W_2\mathbf{B}^2, \\ \mathbf{S} &= 2I_3W_3\,\mathbf{F}^{-T} + 2\left[W_1 + I_1W_2\right]\mathbf{F} - 2W_2\mathbf{B}\mathbf{F}, \end{aligned}\right\} \tag{4.67}$$

and for an incompressible isotropic elastic material we have

$$\begin{aligned} \mathbf{T} &= -q\,\mathbf{I} + 2\left[W_1 + I_1W_2\right]\mathbf{B} - 2W_2\mathbf{B}^2, \\ \mathbf{S} &= -q\,\mathbf{F}^{-T} + 2\left[W_1 + I_1W_2\right]\mathbf{F} - 2W_2\mathbf{B}\mathbf{F}, \end{aligned} \tag{4.68}$$

where we have set

$$W_i = \frac{\partial W}{\partial I_i}. \tag{4.69}$$

Terminology: An isotropic material whose constitutive behavior is described by (4.67) is frequently referred to as a compressible material. However a compressible material (i.e. one that is not incompressible) may involve some other constraint, for example an inextensibility constraint, in which case its constitutive relation would not be (4.67) even though it is compressible. Thus we shall speak of (4.67) as describing an **unconstrained** isotropic elastic material (rather than a compressible material).

4.6.1 Incompressible isotropic materials.

Uniaxial stress.

Consider a state of uniaxial stress in the $\mathbf{e}_1$-direction. The Cauchy stress tensor is

$$\mathbf{T} = T\,\mathbf{e}_1 \otimes \mathbf{e}_1, \tag{4.70}$$

and we assume the deformation to be a homogeneous pure stretch $\mathbf{y} = \mathbf{F}\mathbf{x}$ with

$$\mathbf{F} = \lambda \mathbf{e}_1 \otimes \mathbf{e}_1 + \Lambda(\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3). \tag{4.71}$$

The constant parameter λ is the longitudinal stretch (in the direction of T_{11}) and Λ is the transverse stretch (in the direction perpendicular to T_{11}). We have assumed[13] that $\lambda_2 = \lambda_3$. The principal stretches are

$$\lambda_1 = \lambda, \qquad \lambda_2 = \lambda_3 = \Lambda, \tag{4.72}$$

and the tensors $\mathbf{B}$ and $\mathbf{B}^2$ are

$$\mathbf{B} = \lambda_1^2\,\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2^2\,\mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3^2\,\mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \mathbf{B}^2 = \lambda_1^4\,\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2^4\,\mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3^4\,\mathbf{e}_3 \otimes \mathbf{e}_3. \tag{4.73}$$

Our aim is to calculate the normal stress T and transverse stretch Λ in terms of the longitudinal stretch λ.

Incompressibility requires $\lambda_1\lambda_2\lambda_3 = 1$ which on using (4.72) reads $\lambda\Lambda^2 = 1$. Therefore the transverse stretch Λ and longitudinal stretch λ are related by

$$\Lambda = \lambda^{-1/2}. \tag{4.74}$$

[13] Re-examine this analysis without making the a priori assumption $\lambda_3 = \lambda_2$. See Problem 4.28.

The principal scalar invariants associated with this deformation are

$$\left.\begin{aligned} I_1 &= \lambda_1^2+\lambda_2^2+\lambda_3^2 \overset{(4.72)}{=} \lambda^2+2\Lambda^2 \overset{(4.74)}{=} \lambda^2+2\lambda^{-1}, \\ I_2 &= \lambda_1^2\lambda_2^2+\lambda_2^2\lambda_3^2+\lambda_3^2\lambda_1^2 \overset{(4.72)}{=} 2\lambda^2\Lambda^2+\Lambda^4 \overset{(4.74)}{=} 2\lambda+\lambda^{-2}. \end{aligned}\right\} \tag{4.75}$$

We now turn to the constitutive relation $(4.68)_1$ for stress, keeping in mind that it involves the reaction pressure q. We first determine q by making use of the fact that $T_{22} = T_{33} = 0$, and thereafter calculate T_{11}. Substituting (4.73) into $(4.68)_1$ gives

$$T_{22} = -q + 2(W_1 + I_1 W_2)\lambda_2^2 - 2W_2\lambda_2^4 = 0,$$

which can be solved for q:

$$q = 2\lambda^{-1}W_1 + 2(\lambda+\lambda^{-2})W_2, \tag{4.76}$$

where we have made use of (4.72), (4.74) and (4.75) to eliminate I_1 and λ_2 in favor of λ. The normal stress $T_{11} = T$ is now found by substituting (4.73), (4.75) and (4.76) into $(4.68)_1$, which after simplification leads to

$$T = T_{11} = -q + 2(W_1 + I_1W_2)\lambda^2 - 2W_2\lambda^4 = 2(W_1+\lambda^{-1}W_2)(\lambda^2-\lambda^{-1}). \tag{4.77}$$

This describes the *stress-stretch response in uniaxial stress*. In both (4.76) and (4.77) the derivatives W_1 and W_2 of the strain energy function are evaluated at the values of the invariants given by (4.75), i.e. at $I_1 = \lambda^2 + 2\lambda^{-1}, I_2 = 2\lambda + \lambda^{-2}$.

We can now calculate the components of the 1st Piola-Kirchhoff stress tensor by using the formula $\mathbf{S} = J\mathbf{T}\mathbf{F}^{-T}$. However it is illuminating to do so by physical reasoning instead. Suppose the cross section of the body normal to the axis of stressing has dimensions 1×1 in the reference configuration. In the deformed configuration its dimensions are $\lambda_2 \times \lambda_3$. Thus the areas of this cross section in the undeformed and deformed configurations are 1 and $\lambda_2\lambda_3$ respectively. Therefore the axial force on this cross section can be written in the equivalent forms $S_{11}\times 1$ and $T_{11}\times\lambda_2\lambda_3$. Thus $S_{11} = T_{11}\lambda_2\lambda_3 = T\Lambda^2$:

$$S = S_{11} = 2(W_1+\lambda^{-1}W_2)(\lambda-\lambda^{-2}), \tag{4.78}$$

with all other stress components S_{ij} being zero.

Let $w(\lambda)$ be the restriction of the strain energy function W to uniaxial stress:

$$w(\lambda) := W(I_1, I_2) \qquad \text{where} \quad I_1 = \lambda^2+2\lambda^{-1}, I_2 = 2\lambda+\lambda^{-2}. \tag{4.79}$$

It is readily seen by differentiating (4.79) with respect to λ and using (4.78) that

$$S = w'(\lambda). \tag{4.80}$$

This is in fact a consequence of $\mathbf{S} \cdot \dot{\mathbf{F}} = \dot{W}$ which when specialized to the present setting reads $S_{11}\dot{\lambda} = \dot{w} = w'(\lambda)\,\dot{\lambda}$ thus resulting in (4.80).

Finally we linearize these results for an infinitesimal deformation. The normal strain components ε_1 and ε_2 of the infinitesimal strain tensor $\boldsymbol{\epsilon}$ are related to the stretches by

$$\lambda_1 = 1 + \varepsilon_1, \qquad \lambda_2 = 1 + \varepsilon_2. \tag{4.81}$$

First consider the relation between the transverse stretch Λ and the axial stretch λ. Substituting (4.81) into (4.74) and approximating the result for small ε_1 gives

$$1 + \varepsilon_2 = (1 + \varepsilon_1)^{-1/2} \approx 1 - \frac{1}{2}\varepsilon_1 + \ldots \qquad \Rightarrow \qquad \varepsilon_2 \approx -\frac{1}{2}\varepsilon_1.$$

This shows that the Poisson's ratio ν at infinitesimal deformations, $\nu = -\varepsilon_2/\varepsilon_1$, has the value $1/2$:

$$\nu = \frac{1}{2}$$

for all incompressible isotropic elastic materials. Next we linearize the stress-stretch relation (4.78) (or equivalently (4.77)). First consider the term $\lambda - \lambda^{-2}$. It can be linearized as follows:

$$\lambda - \lambda^{-2} = (1 + \varepsilon_1) - (1 + \varepsilon_1)^{-2} \approx (1 + \varepsilon_1) - (1 - 2\varepsilon_1 + \ldots) = 3\varepsilon_1.$$

Since this term is order $\mathrm{O}(\varepsilon)$, and it multiplies the remaining terms on the right hand side of (4.78), we need only approximate those other terms to $\mathrm{O}(1)$. So we set $\lambda = 1$ and write the remaining term as $2(W_1 + W_2)$ keeping in mind that the W_i's are now evaluated at $I_1 = I_2 = 3$ (which is what we get by setting $\lambda = 1$ in (4.75)). We are thus led to the linear stress-strain relation

$$S \approx 6(W_1 + W_2)\varepsilon_1.$$

The Young's modulus of any incompressible isotropic elastic material at infinitesimal deformations is therefore

$$E := 6(W_1 + W_2)\Big|_{I_1 = I_2 = 3}. \tag{4.82}$$

Simple shear.

Now consider a simple shear deformation of a unit cube with shearing direction $\mathbf{e}_1$ and glide plane normal $\mathbf{e}_2$:

$$y_1 = x_1 + kx_2, \qquad y_2 = x_2, \qquad y_3 = x_3. \tag{4.83}$$

One finds that $\det \mathbf{F} = 1$ and so the deformation is automatically volume preserving. The left Cauchy-Green deformation tensor and its square are

$$\begin{aligned} \mathbf{B} &= \mathbf{I} + k^2\, \mathbf{e}_1 \otimes \mathbf{e}_1 + k(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1), \\ \mathbf{B}^2 &= \mathbf{I} + (3k^2 + k^4)\, \mathbf{e}_1 \otimes \mathbf{e}_1 + k^2\, \mathbf{e}_2 \otimes \mathbf{e}_2 + (2k + k^3)(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1), . \end{aligned} \tag{4.84}$$

and so the principal scalar invariants of $\mathbf{B}$ are

$$I_1 = \operatorname{tr} \mathbf{B} = 3 + k^2, \qquad I_2 = \frac{1}{2}\big[(\operatorname{tr} \mathbf{B})^2 - \operatorname{tr}(\mathbf{B}^2)\big] = 3 + k^2. \tag{4.85}$$

We now turn to the constitutive relation $(4.68)_1$. Substituting (4.84) and (4.85) into $(4.68)_1$ gives the shear stress T_{12} to be

$$T_{12} = 2k(W_1 + W_2)\Big|_{I_1=3+k^2, I_2=3+k^2}. \tag{4.86}$$

This is the relation between the shear stress T_{12} and the amount of shear k. The remaining shear stress components are readily seen to vanish $T_{23} = T_{31} = 0$.

Turning to the normal components of stress, we first note that they involve the reaction pressure q which we cannot determine unless we know the stress on one pair of faces of the cube. Suppose that the faces perpendicular to the $\mathbf{e}_3$-direction are traction-free: $T_{33} = 0$. Substituting (4.84) into $(4.68)_1$ gives T_{33} and setting it equal to zero yields

$$T_{33} = -q + 2(W_1 + I_1 W_2) - 2W_2 = 0.$$

This can be solved for q, which after using (4.85), can be written as

$$q = 2W_1 + 2W_2(2 + k^2). \tag{4.87}$$

The normal stress components T_{11} and T_{22} can now be found from $(4.68)_1$, (4.84) , (4.85) and (4.87) to be

$$T_{11} = 2k^2 W_1, \qquad T_{22} = -2k^2 W_2. \tag{4.88}$$

Observe from this that in general, *normal stresses are needed in order to maintain a simple shear deformation.* This is a feature of finite deformations and is in contrast to the linearized theory where the shear stress T_{12} is the only nonzero stress. This is sometimes called the *Poynting effect*.

It is interesting to observe from (4.86) and (4.88) that

$$T_{11} - T_{22} = kT_{12},$$

which is a relation between the stress components that does not involve W. It therefore holds for all incompressible isotropic materials and is called a *universal relation*.

Problem 4.6.1. Show that this universal relation holds even if we did not take $T_{33} = 0$.

Next we define the restriction of the strain energy function W to simple shear by

$$w(k) := W(I_1, I_2)\Big|_{I_1=3+k^2,\ I_2=3+k^2}. \tag{4.89}$$

Note that w here is different to the function w introduced in (4.79). Differentiating (4.89) with respect to k and using (4.86) yields

$$T_{12} = w'(k).$$

Again, this is a consequence of $\mathbf{S} \cdot \dot{\mathbf{F}} = \dot{W}$ specialized to the present setting.

Finally consider an infinitesimal deformation (4.83) with $|k| \ll 1$. Linearizing (4.86) for small $|k|$ gives

$$T_{12} = 2k(W_1 + W_2)\Big|_{I_1=I_2=3} + \mathrm{O}(k^2).$$

Thus we have the linear stress-strain relation $T_{12} = \mu k$ where

$$\mu := 2(W_1 + W_2)\Big|_{I_1=I_2=3} \tag{4.90}$$

is the *shear modulus* of linear elasticity. Observe from (4.88) that for infinitesimal amounts of shear

$$T_{11} = \mathrm{O}(k^2), \qquad T_{22} = \mathrm{O}(k^2),$$

which is why these stress components are neglected in the linearized theory. Note from (4.82) and (4.90) that for any incompressible isotropic elastic material the Young's modulus and shear modulus are related by $E = 3\mu$.

Biaxial stretch in plane stress.

Rivlin and Saunders [19] carried out biaxial stress-stretch experiments on thin sheets of rubber and so we turn to such states of deformation and stress next. Consider a thin square sheet of dimension $1 \times 1 \times h$ in the reference configuration. The coordinate axes are aligned with the edges of the sheet with the x_3-axis being perpendicular to the square faces. The sheet is subjected to a pure stretch

$$y_1 = \lambda_1 x_1, \qquad y_2 = \lambda_2 x_2, \qquad y_3 = \lambda_3 x_3, \tag{4.91}$$

with the top and bottom faces of the sheet being traction-free:

$$T_{31} = T_{32} = T_{33} = 0 \qquad \text{for } x_3 = \pm h/2.$$

In view of incompressibility,

$$\lambda_3 = \lambda_1^{-1}\lambda_2^{-1}, \tag{4.92}$$

and so the principal scalar invariants take the form

$$\left.\begin{aligned} I_1 &= \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \overset{(4.92)}{=} \lambda_1^2 + \lambda_2^2 + \lambda_1^{-2}\lambda_2^{-2}, \\ I_2 &= \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2 \overset{(4.92)}{=} \lambda_1^2\lambda_2^2 + \lambda_1^{-2} + \lambda_2^{-2}. \end{aligned}\right\} \tag{4.93}$$

The left Cauchy-Green tensor and its square associated with the deformation (4.91) and (4.92) are

$$\begin{aligned} \mathbf{B} &= \lambda_1^2 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2^2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_1^{-2}\lambda_2^{-2} \mathbf{e}_3 \otimes \mathbf{e}_3, \\ \mathbf{B}^2 &= \lambda_1^4 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2^4 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_1^{-4}\lambda_2^{-4} \mathbf{e}_3 \otimes \mathbf{e}_3. \end{aligned} \tag{4.94}$$

It follows from $(4.68)_1$ and (4.94) that the shear stress components vanish throughout the sheet and so the traction-free boundary condition on the shear stress components, $T_{31} = T_{32} = 0$ for $x_3 = \pm h/2$, holds automatically. Calculating T_{33} from $(4.68)_1$, (4.93) and (4.94) and setting the result equal to zero allows one to solve for the reaction pressure q. This leads to

$$q = 2\lambda_1^{-2}\lambda_2^{-2}\, W_1 \; - 2\lambda_1^2\lambda_2^2\, W_2\,. \tag{4.95}$$

The in-plane normal stress components are now found from $(4.68)_1$, (4.93) and (4.94), together with (4.95), to be

$$T_{11} = 2(\lambda_1^2 - \lambda_1^{-2}\lambda_2^{-2})(W_1 + \lambda_2^2 W_2), \qquad T_{22} = 2(\lambda_2^2 - \lambda_1^{-2}\lambda_2^{-2})(W_1 + \lambda_1^2 W_2). \tag{4.96}$$

The corresponding 1st Piola-Kirchhoff stresses are readily found by noting that the dimensions of the sheet in the deformed configuration are $\lambda_1, \lambda_2, h\lambda_3$. Thus the force on a face perpendicular to the x_1-axis can be written as $T_{11} \times \textsf{deformed area} = T_{11}\lambda_2 h\lambda_3$ and equivalently as $S_{11} \times \textsf{undeformed area} = S_{11}h$. Likewise the force on a face perpendicular to the x_2-axis is $T_{22}\lambda_1 h\lambda_3$ and equivalently $S_{22}h$. On equating these one has $T_{11}\lambda_2 h\lambda_3 = S_{11}h$ and $T_{22}\lambda_1 h\lambda_3 = S_{22}h$ whence

$$S_{11} = T_{11}\lambda_2\lambda_3 = T_{11}\lambda_1^{-1} = \frac{\partial w}{\partial \lambda_1}, \qquad S_{22} = T_{22}\lambda_1\lambda_3 = T_{22}\lambda_2^{-1} = \frac{\partial w}{\partial \lambda_2}, \tag{4.97}$$

where

$$w(\lambda_1, \lambda_2) = \widetilde{W}(I_1, I_2)\Big|_{I_1=\lambda_1^2+\lambda_2^2+\lambda_1^{-2}\lambda_2^{-2},\ I_2=\lambda_1^2\lambda_2^2+\lambda_1^{-2}+\lambda_2^{-2}} .$$

Solving (4.96) for W_1 and W_2 and using (4.97) leads to

$$\begin{aligned} W_1 &= \frac{1}{2(\lambda_1^2-\lambda_2^2)}\left[\frac{\lambda_1^3 S_{11}}{\lambda_1^2-\lambda_1^{-2}\lambda_2^{-2}} - \frac{\lambda_2^3 S_{22}}{\lambda_2^2-\lambda_1^{-2}\lambda_2^{-2}}\right], \\ W_2 &= \frac{1}{2(\lambda_2^2-\lambda_1^2)}\left[\frac{\lambda_1 S_{11}}{\lambda_1^2-\lambda_1^{-2}\lambda_2^{-2}} - \frac{\lambda_2 S_{22}}{\lambda_2^2-\lambda_1^{-2}\lambda_2^{-2}}\right]. \end{aligned} \tag{4.98}$$

Rivlin and Saunders [19] carried out biaxial stress-stretch experiments on thin rubber sheets. They varied the stretches λ_1 and λ_2, keeping I_1 fixed and allowing I_2 to vary, where I_1 and I_2 are given by (4.93). The experiments were repeated for different fixed values of I_1. They measured the values of S_{11} and S_{22} and then used (4.98) to determine W_1 and W_2. They also carried out experiments in which I_2 was fixed and I_1 was varied. In this way they determined $W_1(I_1, I_2)$ and $W_2(I_1, I_2)$ along various straight lines $I_1 = \textsf{constant}$ and $I_2 = \textsf{constant}$ in the I_1, I_2-plane. This information was then used to determine the strain energy function $W(I_1, I_2)$ describing the particular material they were testing.

4.6.2 Unconstrained isotropic materials.

Uniaxial stress.

Consider a state of uniaxial stress in the $\mathbf{e}_1$-direction. The Cauchy stress tensor is again given by (4.70), the deformation by (4.71), the principal stretches by (4.72), and the tensors $\mathbf{B}$ and $\mathbf{B}^2$ by (4.73). Our aim is to calculate the normal stress T and transverse stretch Λ in terms of the longitudinal stretch λ.

In the case of an incompressible material we used the incompressibility constraint to determine Λ. Here we will use the condition $T_{22} = 0$ to determine Λ. In the incompressible case the equation $T_{22} = 0$ was used to determine the reaction pressure q which is absent for an unconstrained material.

The principal scalar invariants associated with the deformation at hand are

$$\left.\begin{aligned} I_1 &= \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \overset{(4.72)}{=} \lambda^2 + 2\Lambda^2, \\ I_2 &= \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2 \overset{(4.72)}{=} 2\lambda^2\Lambda^2 + \Lambda^4, \\ I_3 &= J^2 = \lambda_1^2\lambda_2^2\lambda_3^2 \overset{(4.72)}{=} \lambda^2\Lambda^4. \end{aligned}\right\} \tag{4.99}$$

We can now calculate the stress from the constitutive relation $(4.67)_1$. For T_{11} we get

$$T_{11} = 2\lambda\Lambda^2 W_3 + \frac{2}{\lambda\Lambda^2}\big(W_1 + (\lambda^2 + 2\Lambda^2)W_2\big)\lambda^2 - \frac{2}{\lambda\Lambda^2}W_2\lambda^4, \tag{4.100}$$

where the derivatives of W appearing in (4.100) are evaluated at the values of the invariants given in (4.99). This is an equation of the form $T_{11} = T_{11}(\lambda, \Lambda)$. Similarly we calculate T_{22} which we set equal to zero:

$$T_{22} = 2\lambda\Lambda^2 W_3 + \frac{2}{\lambda\Lambda^2}\big(W_1 + (\lambda^2 + 2\Lambda^2)W_2\big)\Lambda^2 - \frac{2}{\lambda\Lambda^2}W_2\Lambda^4 = 0. \tag{4.101}$$

Equation (4.101) is a nonlinear algebraic equation of the form $f(\lambda, \Lambda) = 0$. If (in principle) (4.101) can be solved for Λ, one has a relation of the form $\Lambda = \Lambda(\lambda)$ for finding the transverse stretch in terms of the longitudinal stretch. This is now substituted into (4.100) to obtain the stress T_{11} as a function of the stretch λ.

Remark: If we had carried out this analysis using the representation $W^*(\lambda_1, \lambda_2, \lambda_3)$ of the strain energy function together with the constitutive relation (4.46) for stress, the transverse stretch and normal stress are given by the respective equations

$$\left.\frac{\partial W^*}{\partial \lambda_2}\right|_{\lambda_1=\lambda, \lambda_2=\lambda_3=\Lambda} = 0, \qquad T_{11} = \frac{1}{\Lambda^2}\left.\frac{\partial W^*}{\partial \lambda_1}\right|_{\lambda_1=\lambda, \lambda_2=\lambda_3=\Lambda}.$$

Simple shear.

The tensors $\mathbf{B}$ and $\mathbf{B}^2$ again have the components given in (4.84) and the principal scalar invariants are

$$I_1 = 3 + k^2, \qquad I_2 = 3 + k^2, \qquad I_3 = 1. \tag{4.102}$$

Substituting these into (4.67)$_1$ and simplifying leads to

$$T_{12} = 2k(W_1 + W_2) \tag{4.103}$$

keeping in mind that the strain energy function is of the form $W(I_1, I_2, I_3)$ and its derivatives appearing in (4.103) are to be evaluated at the values of the invariants given in (4.102). Again, the normal components of stress are nonzero in general.

4.6.3 Restrictions on the strain energy function.

A nonlinearly elastic material is characterized by its strain energy function $W(\mathbf{F})$. Typically, W is determined by using (micro-mechanical reasoning to motivate) some functional form that is then fitted to experimental measurements. Since some level of judgement is used in coming up with the functional form, and since experimental data available is necessarily limited to a certain finite (even if large) number of tests, it is important to make sure that a proposed constitutive model is not fundamentally flawed in some way.

There are two types of restrictions that one might consider imposing on $W(\mathbf{F})$ in order to address this issue. The first ensures that the response predicted by the constitutive model is "physically reasonable". For example in simple shear, the shear stress τ and amount of shear k are related by $\tau = 2k(\widetilde{W}_1 + \widetilde{W}_2)$. This by itself does not ensure, for example, that a positive shear stress $\tau > 0$ is needed to deform the body by a positive amount of shear $k > 0$. That would require $\widetilde{W}$ to obey the inequality

$$\left.\left(\frac{\partial \widetilde{W}}{\partial I_1} + \frac{\partial \widetilde{W}}{\partial I_2}\right)\right|_{I_1 = I_2 = 3+k^2, I_3 = 1} > 0 \qquad \text{for all } k.$$

Similarly in uniaxial stress, equations (4.100) and (4.101) do not ensure that a tensile stress ($T_{11} > 0$) is required to elongate the body ($\lambda_1 > 1$). This would require $\widetilde{W}$ to obey a second inequality. One must of course be careful in deciding what is "reasonable". For example in some situations one might require the stress-stretch relation in uniaxial stress to be monotonically increasing for all stretches. There are however certain situations where this is not the case and the stress-stretch relation is monotonically increasing for only certain ranges of stretch. We shall encounter such a problem in Section 5.6.

The second type of restriction stems from mathematical considerations. For example, without certain requirements on W it is possible that boundary-value problems may have no solution. Or the solutions that exist maybe unstable. Again, one must be careful in what one imposes. Consider for example the strong ellipticity condition described below that is related to a certain notion of stability. When one models phase transformations in solids, it is known that the constitutive model has to violate this condition at certain deformation gradients (but not others). This does not mean that the strong ellipticity condition should be abandoned entirely; only that it not be required at *all* deformation gradients. It is also worth mentioning that one does not always expect uniqueness of solutions to boundary value problems in the nonlinear theory. If we did, we would not be able to study instabilities such as buckling. Thus uniqueness is not something one would insist on in general.

In this sub-section we list a few restrictions on the strain energy function W that have been suggested in the literature. Checking whether the specific strain energy functions presented in the next section obey these conditions is left as an exercise. See also Problems 4.5, 4.26, 4.27 and 4.31 .

Baker-Ericksen inequalities. Consider an *isotropic* material subjected to a pure homogeneous deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ where $\mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3$. Let the accompanying Cauchy stress be $\mathbf{T} = \tau_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \tau_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \tau_3 \mathbf{e}_3 \otimes \mathbf{e}_3$. Baker and Ericksen [3] suggested that if the principal stretch λ_i is larger than the principal stretch λ_j, then one would expect the principal Cauchy stress τ_i to be larger than the principal Cauchy stress τ_j, i.e. that $\lambda_i > \lambda_j$ implies $\tau_i > \tau_j$. This requires

$$(\tau_i - \tau_j)(\lambda_i - \lambda_j) > 0 \qquad \text{whenever} \quad \lambda_i \neq \lambda_j, \tag{4.104}$$

which, by using the constitutive relation $\tau_i = \lambda_i J^{-1} \partial W / \partial \lambda_i$ can be written as

$$\frac{\lambda_i \, \partial W / \partial \lambda_i - \lambda_j \, \partial W / \partial \lambda_j}{\lambda_i - \lambda_j} > 0 \qquad \text{provided} \quad \lambda_i \neq \lambda_j. \tag{4.105}$$

Problem 4.28 illustrates one consequence of this, namely that in a uniaxial tensile stress state $\mathbf{T} = \tau \mathbf{e}_1 \otimes \mathbf{e}_1$ with $\tau > 0$, the Baker-Ericksen inequalities hold if and only if the principal stretches obey $\lambda_1 > \lambda_2 = \lambda_3 > 0$.

Monotonicity: Consider again a pure homogeneous deformation of an isotropic material. Let $\mathbf{S} = \sigma_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \sigma_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \sigma_3 \mathbf{e}_3 \otimes \mathbf{e}_3$ be the associated 1st Piola-Kirchhoff stress. If one requires each stress component σ_i (which is effectively the ith component of force) to be an increasing function of the corresponding stretch λ_i, it then follows from the constitutive relation $\sigma_i = \partial W / \partial \lambda_i$ that one must have

$$\frac{\partial^2 W}{\partial \lambda_i^2} > 0. \tag{4.106}$$

This describes a certain type of convexity of $W(\lambda_1, \lambda_2, \lambda_3)$.

Convexity. Convexity plays a central role in proving the existence of solutions (to minimization problems in the Calculus of Variations). There are various notions of convexity such as quasiconvexity, polyconvexity, rank-one convexity and this is a topic that is beyond the scope of these notes. The interested reader is referred to, for example, Antman [2], Ball [5] Marsden and Hughes [15] and Steigmann [20]. It is shown in some of these references that the usual notion of convexity of $W(\mathbf{F})$ leads to various difficulties including incompatibility with material frame indifference, preclusion of buckling and the violation of the growth condition $W \to \infty$ as $\det \mathbf{F} \to 0$.

Consider the special case of a homogeneous deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ of a homogeneous elastic body whose entire boundary is subjected to a prescribed first Piola-Kirchhoff traction ("dead loading"). If $\mathbf{F}$ is a local minimizer of $W(\mathbf{F})$ one must have

$$\mathbb{A}_{ijk\ell}\, H_{ij} H_{k\ell} > 0 \tag{4.107}$$

for all tensors $\mathbf{H} \neq \mathbf{0}, |\mathbf{H}| \ll 1$, where

$$\mathbb{A}_{ijk\ell}(\mathbf{F}) \coloneqq \frac{\partial^2 W}{\partial F_{ij} \partial F_{k\ell}}(\mathbf{F}), \tag{4.108}$$

are the components of the **elasticity tensor** $\mathbb{A}$. If one limits attention to (a) isotropic materials, (b) symmetric deformation gradient tensors $\mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3$, and (c) perturbations $\mathbf{H}$ that are symmetric and coaxial with $\mathbf{F}$ (i.e. $\mathbf{H}$ that have the same principal directions as $\mathbf{F}$), the inequality (4.107) leads to the requirement that the Hessian matrix with elements

$$\frac{\partial^2 W^*}{\partial \lambda_i \partial \lambda_j} \qquad \text{be positive definite.} \tag{4.109}$$

We shall work out the details of this in the context of a particular problem in Section 5.3. Additional inequality restrictions on the $\partial W/\partial \lambda_i$'s are needed when the perturbation $\mathbf{H}$ is not coaxial with $\mathbf{F}$. For these, as well as the modification of (4.109) for incompressible materials, see Ogden [18].

Strong ellipticity (material stability): The basic idea underlying material stability is to ask whether, when a homogeneous equilibrium configuration is perturbed, does the perturbation decay to zero as $t \to \infty$ (at each point in the body)? To illustrate the idea, consider a one-dimensional setting. The deformation $y(x,t)$, stretch $\lambda(x,t)$ and stress $\sigma(x,t)$ obey the equations

$$\lambda = y_x, \qquad \sigma = W_\lambda, \qquad \sigma_x = \rho_{\mathrm{R}} y_{tt},$$

where the subscripts x, t and λ denote differentiation and the mass density ρ_{R} in the reference configuration is positive. Consider a homogeneous equilibrium configuration

$$y(x) = \lambda_0 x, \qquad \lambda(x) = \lambda_0, \qquad \sigma(x) = \sigma_0 = W_\lambda(\lambda_0),$$

where $\lambda_0 > 0$ is a constant. Now consider an infinitesimal perturbation $u(x,t)$ of this configuration: $y(x,t) = \lambda_0 x + u(x,t)$. The associated acceleration, stretch and stress are

$$y_{tt} = u_{tt}, \quad \lambda = y_x = \lambda_0 + u_x, \quad \sigma = W_\lambda(\lambda_0 + u_x) \doteq W_\lambda(\lambda_0) + W_{\lambda\lambda}(\lambda_0)u_x = \sigma_0 + W_{\lambda\lambda}(\lambda_0)u_x,$$

and so the equation of motion $\sigma_x = \rho_{\mathrm{R}} y_{tt}$ gives

$$W_{\lambda\lambda}(\lambda_0)u_{xx} = \rho_{\mathrm{R}} u_{tt} \qquad \Rightarrow \qquad \mathfrak{a} u_{xx} = \rho_{\mathrm{R}} u_{tt} \quad \text{where} \quad \mathfrak{a} := W_{\lambda\lambda}(\lambda_0).$$

We now seek a solution of this linear partial differential equation in the form $u(x,t) = a \exp ik(x - ct)$. Observe that if c is imaginary, say $c = \pm ir$, the temporal term reads $\exp ik(\mp irt) = \exp(\pm krt)$ and this becomes unbounded as $t \to \infty$. Substituting $u(x,t) = a \exp ik(x - ct)$ into $\mathfrak{a} u_{xx} = \rho_{\mathrm{R}} u_{tt}$ gives

$$c^2 = \mathfrak{a}/\rho_{\mathrm{R}}.$$

Therefore the wave speed c is real and nonzero provided

$$\mathfrak{a} = W_{\lambda\lambda}(\lambda_0) > 0.$$

Observe that when $\mathfrak{a} > 0$ the partial differential equation $\mathfrak{a} u_{xx} = \rho_{\mathrm{R}} u_{tt}$ is hyperbolic. The three-dimensional generalization of the inequality $W_{\lambda\lambda}(\lambda_0) > 0$ is the strong ellipticity condition. Note that when this holds, the energy $W(\lambda)$ is convex at $\lambda = \lambda_0$ and the slope $W_\lambda(\lambda)$ of the stress-stretch curve is positive at $\lambda = \lambda_0$. The slightly weaker inequality

$$\mathfrak{a} = W_{\lambda\lambda}(\lambda_0) \geq 0,$$

ensures a real, but possibly zero, wave speed. The three-dimensional generalization of this inequality is referred to as the Legendre-Hadamard condition.

We now turn to the three-dimensional setting. Consider a pure homogeneous equilibrium deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ of a homogenous but not necessarily isotropic elastic body, and superpose on it an infinitesimal perturbation $\mathbf{u}(\mathbf{x}, t)$. Thus we consider a motion $\mathbf{y} = \mathbf{F}\mathbf{x} + \mathbf{u}(\mathbf{x}, t)$ where $|\nabla \mathbf{u}| \ll 1$. On linearizing the equations of motion we arrive at a linear partial differential equation for the vector field $\mathbf{u}(\mathbf{x}, t)$; see Problem 4.24. Suppose the perturbed motion

$\mathbf{u}(\mathbf{x}, t)$ describes a plane harmonic wave propagating in the direction $\mathbf{n}$, with wave speed c, wave number k and particle motion in the direction $\mathbf{a}$. Then $\mathbf{u}(\mathbf{x}, t) = \mathbf{a}\, \exp[ik(\mathbf{x}\cdot\mathbf{n} - ct)]$. Observe that if the wave speed is imaginary (or complex), the perturbation $\mathbf{u}$ involves a term that grows exponentially with time t and so we would say that the homogeneous deformation is unstable. Thus **material stability**[14] requires the wave speed c to be real (for all $\mathbf{a}$ and $\mathbf{n}$). This requires the elasticity tensor (4.108) to obey the inequality

$$\mathbb{A}_{ijk\ell}(\mathbf{F}) a_i n_j a_k n_\ell > 0 \qquad \text{for all unit vectors } \mathbf{a} \text{ and } \mathbf{n}; \tag{4.110}$$

see Problem 4.24. When $W(\mathbf{F})$ satisfies (4.110) at some $\mathbf{F}$, we say that W is **strongly elliptic**[15] at that $\mathbf{F}$.

Observe that the inequality (4.110) can be written equivalently as

$$\frac{d^2}{dt^2} W(\mathbf{F} + t\mathbf{a}\otimes\mathbf{n})\bigg|_{t=0} > 0 \qquad \text{for all unit vectors } \mathbf{a} \text{ and } \mathbf{n}. \tag{4.111}$$

Viewed in this way, strong ellipticity can be thought of as requiring W to be strictly convex at $\mathbf{F}$ along certain paths in deformation gradient tensor space. If the strict inequality in (4.110) (or (4.111)) is replaced by $\geq$, the resulting inequality is the **Legendre-Hadamard condition** of the Calculus of Variations which in turn is equivalent to the so-called rank-one convexity condition (given that we have implicitly assumed W to be twice continuously differentiable on the set of all tensors with positive determinant); see Ball [5].

In Problem 4.29 you are asked to show for an isotropic material that strong ellipticity implies both the Baker-Ericksen inequalities (4.104) and the aforementioned monotonicity (4.106).

For an incompressible material, according to Problem 4.25, the strong ellipticity condition is

$$\mathbb{A}_{ijk\ell} a_i b_j a_k b_\ell > 0 \qquad \text{for all unit vectors } \mathbf{a} \text{ and } \mathbf{b} \text{ with } \mathbf{a}\cdot\mathbf{F}^{-T}\mathbf{b} = 0, \tag{4.112}$$

where the restriction $\mathbf{a}\cdot\mathbf{F}^{-T}\mathbf{b} = 0$ arises from the requirement that the motions be isochoric. This can be written equivalently as

$$\mathbb{B}_{ijk\ell} a_i n_j a_k n_\ell > 0 \qquad \text{for all unit vectors } \mathbf{a} \text{ and } \mathbf{n} \text{ with } \mathbf{a}\cdot\mathbf{n} = 0 \tag{4.113}$$

[14] One refers to this as *material stability* because no boundary or initial conditions are involved in this notion (in contrast to other notions of stability).

[15] Knowles and Sternberg [14] have shown that the strain energy function W that describes a material capable of undergoing solid-solid phase transitions cannot be strongly elliptic at *all* $\mathbf{F}$, though it would be strongly elliptic at certain $\mathbf{F}$ in general.

where

$$\mathbb{B}_{abcd}(\mathbf{F}) = \mathbb{A}_{apcq}(\mathbf{F}) F_{bp} F_{dq}. \tag{4.114}$$

For an incompressible isotropic material, the components of $\mathbb{B}$ in the principal basis $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ are given in Problem 6.1.8 of Ogden [17] where his tensor $\mathcal{A}_0^1$ is related to our tensor $\mathbb{B}$ by $\mathcal{A}^1_{0ijk\ell} = \mathbb{B}_{ji\ell k}$ (see his equations (6.1.14), (6.1.29) and keep in mind that his tensor $\mathbf{S}$ is the transpose of the first Piola-Kirchhoff stress tensor). With the summation convention suspended,

$$\mathbb{B}_{iijj} = \lambda_i \lambda_j \frac{\partial^2 W}{\partial \lambda_i \partial \lambda_j}, \tag{4.115}$$

$$\mathbb{B}_{jiji} = \frac{\lambda_i^2 \left(\lambda_i \frac{\partial W}{\partial \lambda_i} - \lambda_j \frac{\partial W}{\partial \lambda_j} \right)}{\lambda_i^2 - \lambda_j^2}, \qquad i \neq j, \ \lambda_i \neq \lambda_j, \tag{4.116}$$

$$\mathbb{B}_{jiij} = \mathbb{B}_{ijji} = \mathbb{B}_{jiji} - \lambda_i \frac{\partial W}{\partial \lambda_i}, \qquad i \neq j. \tag{4.117}$$

If $\lambda_i = \lambda_j$, (4.116) is replaced by

$$\mathbb{B}_{jiji} = \frac{1}{2} \left(\mathbb{B}_{iiii} - \mathbb{B}_{iijj} + \lambda_i \frac{\partial W}{\partial \lambda_i} \right), \qquad i \neq j, \ \lambda_i = \lambda_j. \tag{4.118}$$

Growth conditions. One might say that a deformation is extreme if (at least) one of the principal stretches tends to 0 or ∞. It is reasonable to require (for a compressible material) that the strain energy density associated with an extreme deformation be infinite:

$$W(\mathbf{F}) \to \infty \qquad \text{when } \det \mathbf{F} \to 0 \text{ or } \infty. \tag{4.119}$$

The reader is referred to Section 2, Chapter XIII of Antman [2] for a detailed discussion of growth conditions and their implications. The way in which W grows at extreme deformations plays an important role when considering the existence of certain types of discontinuous deformation fields such as those associated with fracture or cavitation. In particular, Ball [4] shows that the growth condition

$$\frac{W(\mathbf{F})}{|\mathbf{F}|^n} \to \infty \qquad \text{as} \quad |\mathbf{F}| \to \infty \tag{4.120}$$

precludes the possibility of cavitation (where $n = 2$ or 3 is the dimension of the space). Therefore in order to study cavitation one must consider strain energy functions for which this condition fails; see Section 5.4.

4.7 Some Models of Isotropic Elastic Materials.

In order to describe the detailed response of a particular isotropic elastic material one needs to know the specific strain energy function $W(I_1, I_2, I_3)$ that characterizes that material. Determining explicit forms for W must be done using laboratory experiments (together with micro-mechanical modeling when possible).

The data from the early experiments of Treloar [22, 23] continue to be important in modeling rubberlike materials. The reader is encouraged to read the classic paper by Rivlin and Saunders [19] which describes some of their experiments carried out to determine the specific W for a certain rubber-like material.

One can find a great many[16] strain energy functions in the literature. In this section we give a few specific examples of particularly simple strain energy functions W. We do not discuss how these forms were developed. Our intention is to simply give a flavor for some explicit examples. In order to determine the response according to each of these W's one simply substitutes them into the formulae we derived in Sections 4.6.1 and 4.6.2.

The Gent, 1-term Ogden and Fung models that we consider below, see (4.132), (4.138) and (4.142), each involves a single constitutive parameter J_m, n and β respectively (in addition to the shear modulus μ). According to Table 11.1 of Goriely [8] , the Gent material provides a reasonable model for elastomers when $20 < J_m < 200$ and for soft biological tissues when $1/3 < J_m < 5/2$. For the 1-term Ogden material the appropriate parameter ranges are $n \approx 3$ (elastomers) and $n \geq 9$ (soft biological tissues). For the Fung material, $3 < \beta < 20$ provides a reasonable model for soft biological tissues.

In Figure 4.5 we show the stress-stretch responses in uniaxial stress according to several of these models. In the left-hand figure we plot the Cauchy and 1st Piola-Kirchhoff stresses versus stretch for a neo-Hookean material. In the right-hand figure we plot the 1st Piola-Kirchhoff stress versus stretch for neo-Hookean, 1-term Ogden, Gent and Fung materials.

1. **A compressible inviscid fluid:** Consider an elastic material characterized by the strain energy function

$$W = W(I_3).$$

Substituting this into (4.67) yields the following constitutive relation for the Cauchy

[16] Holzapfel [11] makes this point (tongue-in-cheek) by saying "... new specific forms [of W] are published on a daily basis."

stress:

$$\mathbf{T} = -p\,\mathbf{I} \qquad \text{where} \quad p = -2JW'(I_3),\ I_3 = J^2.$$

Observe that for this class of materials, the stress tensor is hydrostatic in *every* deformation. Therefore the strain energy function $W = W(I_3)$ describes an inviscid fluid. The constitutive relation for $\mathbf{T}$ can be written in the form familiar in fluid mechanics by replacing J with the mass density ρ using

$$\rho = \rho_{\mathrm{R}}/J,$$

where ρ_{R} is the mass density in the reference configuration, and replacing $W(I_3)$ by the function $\psi(\rho)$ defined by

$$\psi(\rho) := \frac{1}{\rho_{\mathrm{R}}} W(I_3), \qquad I_3 = J^2 = (\rho_{\mathrm{R}}/\rho)^2.$$

One can then show that the constitutive relation above can be written as

$$\mathbf{T} = -p\,\mathbf{I} \qquad \text{where} \quad p = \rho^2\,\psi'(\rho).$$

2. **Generalized neo-Hookean model. (Incompressible):** A generalized neo-Hookean material is described by a strain energy function of the form

$$W(I_1, I_2) = W(I_1) \qquad \text{for } I_1 \geq 3. \tag{4.121}$$

Many constitutive models for incompressible isotropic rubber-like materials, including the neo-Hookean, Arruda-Boyce and Gent models, are special cases of the generalized neo-Hookean model. Substituting (4.121) into $(4.68)_1$ leads to the constitutive relation

$$\mathbf{T} = -q\,\mathbf{I} \; + \; 2W'(I_1)\,\mathbf{B}, \qquad I_1 = \operatorname{tr}\mathbf{B}. \tag{4.122}$$

In *uniaxial stress*, the stress-stretch relations for the Cauchy stress and the 1st Piola-Kirchhoff stress are found from (4.77) and (4.78) respectively:

$$T = 2W'(I_1)(\lambda^2 - \lambda^{-1}) \qquad \text{where } I_1 = \lambda^2 + 2\lambda^{-1}, \tag{4.123}$$

$$S = 2W'(I_1)(\lambda - \lambda^{-2}) \qquad \text{where } I_1 = \lambda^2 + 2\lambda^{-1}. \tag{4.124}$$

The relation between the transverse stretch Λ and the longitudinal stretch λ is $\Lambda = \lambda^{-1/2}$.

In **simple shear** one finds from (4.86) and (4.121) that the shear stress T_{12} is related to the amount of shear k by

$$T_{12} = 2W'(I_1)\, k, \qquad I_1 = 3 + k^2, \tag{4.125}$$

and from(4.88) that the normal stresses are

$$T_{11} = 2k^2 W'(I_1), \qquad T_{22} = T_{33} = 0, \qquad I_1 = 3 + k^2, \tag{4.126}$$

having assumed that $T_{33} = 0$. The fact that setting $T_{33} = 0$ automatically implies $T_{22} = 0$ is a peculiarity of *all* generalized neo-Hookean materials. For a more general W, one finds that $T_{22} \neq 0$ in general; see for example the response of a Mooney-Rivlin material described later in this section. If the shear stress T_{12} is to be > 0 when the amount of shear k is > 0, (4.125) shows that we must have $W' > 0$. Thus it is reasonable to require the constitutive function $W(I_1)$ for a generalized neo-Hookean material to obey

$$W'(I_1) > 0 \qquad \text{for } I_1 \geq 3. \tag{4.127}$$

From the linearized expression (4.90) we see that the infinitesimal shear modulus of the material is $\mu = 2W'(3)$.

3. **Neo-Hookean model. (Incompressible):** A neo-Hookean material is characterized by the strain energy function

$$W(I_1, I_2) = \frac{\mu}{2}(I_1 - 3), \qquad \mu > 0, \tag{4.128}$$

where μ is a material constant. This is a special case of a generalized neo-Hookean material. Substituting (4.128) into (4.122) leads to the constitutive relation

$$\mathbf{T} = -q\,\mathbf{I} \; + \; \mu\mathbf{B}. \tag{4.129}$$

The responses in *uniaxial stress* and *simple shear* are immediately found by specializing (4.123), (4.124) and (4.125):

$$T = \mu(\lambda^2 - \lambda^{-1}), \qquad S = \mu(\lambda - \lambda^{-2}), \qquad T_{12} = \mu k.$$

See Figure 4.5.

4. **Mooney-Rivlin model. (Incompressible):** A Mooney-Rivlin material is characterized by the strain energy function[17]

$$W(I_1, I_2) = \frac{\mu}{2}\Big[\alpha(I_1 - 3) + (1-\alpha)(I_2 - 3)\Big], \qquad \mu > 0,\; 0 < \alpha < 1, \tag{4.130}$$

[17]The strict inequalities on α are required by the Baker-Eriksen inequalities.

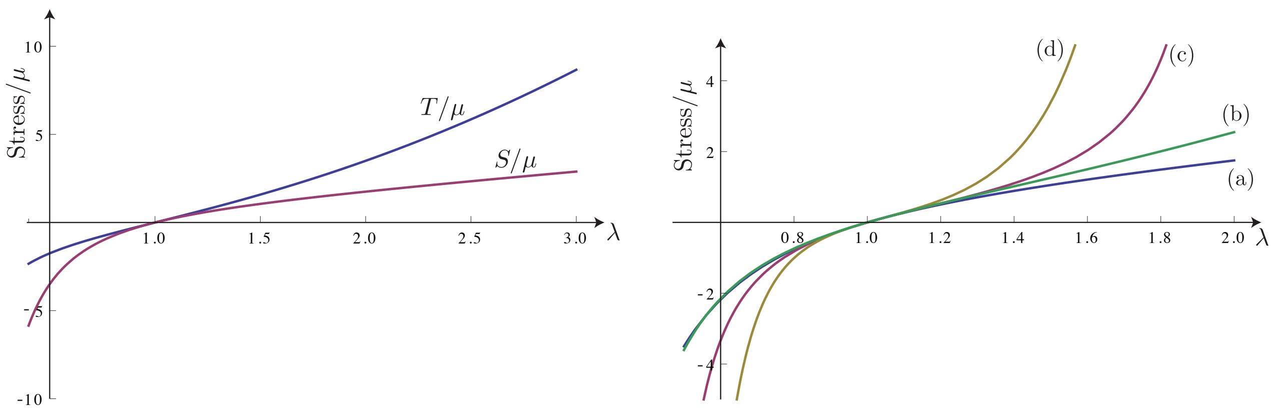


Figure 4.5: Stretch-stretch response in uniaxial stress. Left: Cauchy and 1st Piola-Kirchhoff stress for neo-Hookean material. Right: 1st Piola-Kirchhoff stress for (a) neo-Hookean, (b) 1-term Ogden material with $n_1 = 3, \mu_1 = 2\mu/n_1$, (c) Gent material with $J_m = 2$, and (d) Fung material with $Q(\mathbf{B}) = \beta(I_1 - 3), \alpha = \mu/(2\beta), \beta = 2$.

where μ and α are material constants. Due to the presence of the term I_2 this is *not* a special case of a generalized neo-Hookean material when $\alpha \neq 1$, though it specializes to a neo-Hookean material for $\alpha = 1$. Substituting (4.130) into $(4.68)_1$ gives the following constitutive relation for $\mathbf{T}$,

$$\mathbf{T} = -q\,\mathbf{I} \;+\; \mu\alpha\mathbf{B} + \mu(1-\alpha)\big[I_1\,\mathbf{B} \;-\; \mathbf{B}^2\big]. \tag{4.131}$$

The response of this material in various settings can be readily studied as above.

We shall simply make one observation here. Note from (4.122) that for a generalized neo-Hookean material, the term involving $\mathbf{B}^2$ in the general constitutive equation $(4.68)_1$ drops out while we see from (4.131) that this is not so for the Mooney-Rivlin material. To see one consequence of this consider a simple shear deformation with shearing direction $\mathbf{e}_1$ and glide plane normal $\mathbf{e}_2$, and again suppose that the boundary conditions give $T_{33} = 0$. The response in simple shear can be calculated from (4.86) and one is led to

$$T_{12} = \mu k, \qquad T_{11} = \mu\alpha k^2 \qquad T_{22} = -\mu(1-\alpha)k^2.$$

Observe that $T_{22} \neq 0$ when $0 < \alpha < 1$ which implies that one needs to apply a normal stress on the glide plane in order to maintain a simple shear deformation. This is in contrast to the behavior of a generalized neo-Hookean materials where $T_{22} = 0$; see (4.126).

Reference:

M. Mooney, A theory of large elastic deformation, *Journal of Applied Physics*, volume 11 (1940), pp. 582-592.

R.S. Rivlin, Some applications of elasticity theory to rubber engineering. Original paper 1948. See reprint in *Collected Papers of R.S. Rivlin*, edited by G.I. Barenblatt and D.D. Joseph, Springer, 1997.

5. **Gent Model. Limited Extensibility. (Incompressible):**

Rubber-like materials are composed of a network of freely-jointed randomly oriented polymer chains. As the stress increases in uniaxial tension, initially, most of the deformation arises due to the unfolding of the polymer chains and the slope of the corresponding stress-stretch curve is relatively small. As the polymer chains orient themselves in the pulling direction, the slope begins to increase rapidly until eventually when the polymer chains are "all" oriented in the axial direction, any further increase in stress requires the chains themselves to stretch. Gent modeled this by limiting the extensibility of polymer chains so that the stress tends to infinity as the stretch approaches a certain finite critical value. The strain energy function proposed by Gent is the particular generalized neo-Hookean material

$$W = W(I_1) = -\frac{\mu}{2}\, J_m \,\ln\Big(1 - \frac{I_1 - 3}{J_m}\Big), \qquad \mu > 0,\ J_m > 0, \tag{4.132}$$

where μ and J_m are positive material constants. Since the argument of the logarithm must be positive, the principal invariant I_1 cannot exceed $3 + J_m$:

$$I_1 < 3 + J_m. \tag{4.133}$$

Exercise: Show that in the limit $J_m \to \infty$, the Gent model reduces to the neo-Hookean model (4.128).

Substituting the particular form (4.132) of W into the general constitutive equation $(4.68)_1$ leads to

$$\mathbf{T} = -q\mathbf{I} + \frac{\mu J_m}{3 + J_m - I_1}\,\mathbf{B}. \tag{4.134}$$

In *uniaxial stress* the principal invariant I_1 is

$$I_1 = \lambda^2 + 2\lambda^{-1}, \tag{4.135}$$

having used $\lambda_2 = \lambda_3 = \lambda^{-1/2}$ and $I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2$. Since I_1 cannot exceed $3 + J_m$ we must have

$$\lambda^2 + 2\lambda^{-1} < 3 + J_m \qquad \Rightarrow \qquad \lambda^3 - (3 + J_m)\lambda + 2 < 0.$$

The cubic equation $\lambda^3 - (3 + J_m)\lambda + 2 = 0$ has three real roots, two of which, λ_{min} and λ_{max}, are positive with $0 < \lambda_{min} < 1 < \lambda_{max}$. This implies that the stretch λ must lie in the interval $\lambda_{min} < \lambda < \lambda_{max}$ and in particular cannot exceed λ_{max} . The stress-stretch relation in uniaxial stress is found by substituting (4.132) into (4.123):

$$T_{11} = \left(\lambda^2 - \lambda^{-1}\right)\left(\frac{\mu J_m}{3 + J_m - \lambda^2 - 2\lambda^{-1}}\right), \qquad \lambda_{min} < \lambda < \lambda_{max}. \tag{4.136}$$

Note that $T_{11} \to +\infty$ as $\lambda \to \lambda_{max}$ (and $T_{11} \to -\infty$ as $\lambda \to \lambda_{min}$). The corresponding 1st Piola-Kirchhoff stress $S = T_{11}\lambda^{-1}$ is plotted as a function of stretch in Figure 4.5.

Reference: A.N. Gent, A new constitutive relation for rubber, *Rubber Chemistry and Technology*, 69(1996), pp. 59-61

6. **Ogden model. (Incompressible):** An N-term incompressible Ogden material is characterized by the strain energy function

$$W(\lambda_1, \lambda_2, \lambda_3) = \sum_{i=1}^{N} \frac{\mu_i}{n_i}\left(\lambda_1^{n_i} + \lambda_2^{n_i} + \lambda_3^{n_i} - 3\right), \tag{4.137}$$

where $N, \mu_1, \ldots \mu_N, n_i \ldots n_N$ are material constants such that

$$\mu_i n_i > 0 \qquad \text{for each } i.$$

The material constants n_i need not be integers. Figure 4.5 shows a graph of the 1st Piola-Kirchhoff stress versus stretch in uniaxial stress for the one-term Ogden model

$$W(\lambda_1, \lambda_2, \lambda_3) = \frac{2\mu}{n^2}\left(\lambda_1^{n} + \lambda_2^{n} + \lambda_3^{n} - 3\right), \qquad \mu > 0,\ n > 0. \tag{4.138}$$

The strain energy function (4.138) reduces to the neo-Hookean form when $N = 2$ and yields the Varga model when $N = 1$:

$$W = \mu_1(\lambda_1 + \lambda_2 + \lambda_3 - 3). \tag{4.139}$$

Valanis and Landel suggested that the strain energy function for an isotropic incompressible elastic material should have the form

$$W(\lambda_1, \lambda_2, \lambda_3) = w(\lambda_1) + w(\lambda_2) + w(\lambda_3) \tag{4.140}$$

for some suitable function $w(\cdot)$. Certainly the Ogden, neo-Hookean, Mooney-Rivlin and Varga models are of this form.

Reference: R. W. Ogden, Large deformation isotropic elasticity: On the correlation of theory and experiment for incompressible rubberlike solids, *Proceedings of the Royal Society of London. Series A*, Vol. 326, (1972), issue 1567, pp. 565-584.

7. **Arruda-Boyce model. (Incompressible):**

By using statistical mechanical arguments applied to a cubic representative volume element with eight polymer chains, Arruda and Boyce developed the following particular generalized neo-Hookean material model:

$$W(I_1) = c_1 \left[\beta \mathcal{L}(\beta) - \ln \left(\frac{\sinh \beta}{\beta} \right) \right], \tag{i}$$

where

$$\beta = \mathcal{L}^{-1} \left(\sqrt{\frac{I_1}{3c_2}} \right), \tag{ii}$$

c_1 and c_2 being constant material parameters. The function $\mathcal{L}^{-1}$ in (ii) is the inverse of the Langevin function

$$\mathcal{L}(x) = \coth x - \frac{1}{x}, \qquad -\infty < x < \infty. \tag{iii}$$

The Langevin function (iii) is monotonically increasing with

$$\mathcal{L}(x) \to \pm 1 \qquad \text{as } x \to \pm\infty.$$

Therefore $\mathcal{L}^{-1}(x)$ is only defined for $-1 < x < 1$ and so according to (ii)

$$I_1 < 3c_2.$$

Thus, like the Gent model, the extensibility of this material is limited.

Reference: E. M. Arruda and M.C. Boyce, A three-dimensional model for the large stretch behavior of rubber elastic materials, *J. Mech. Phys. Solids*, 41, 1993, pp. 389-412.

8. **Fung model for soft biological tissue.**

The mechanical response of soft biological tissue is dominated by its fibrous constituents: collagen and elastin. At small strains, the collagen fibers are unstretched and the mechanical response is almost entirely due to the soft, isotropic elastin. As the load increases, the collagen fibers straighten-out and align with the direction of loading. This leads to a rapid increase in the stiffness, as well as to anisotropic material behavior due to the preferred direction induced by the alignment of collagen fibers. Both of these effects can be modeled by the strain energy function

$$W = W(\mathbf{C}) = \alpha \Big(\mathrm{e}^{Q(\mathbf{C})} - 1 \Big), \tag{4.141}$$

where α is a material constant and $Q(\mathbf{C})$ is a scalar-valued function of the right Cauchy Green tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$. The exponential term leads to rapid stiffening of the material. Different functional forms of Q have been considered in the literature, a general quadratic form being the most common. By suitably choosing the form of $Q(\mathbf{C})$, material anisotropy can be built in.

Figure 4.5 shows a graph of the 1st Piola-Kirchhoff stress versus stretch in uniaxial stress for the particular (isotropic) Fung model

$$W = \frac{\mu}{2\beta}\Big(\mathrm{e}^{\beta(I_1-3)} - 1\Big). \tag{4.142}$$

Observe that in the limit $\beta \to 0$, the strain energy function (4.142) reduces to the neo-Hookean model.

Reference: Review article by J. D. Humphrey, Continuum biomechanics of soft biological tissues, *Proceeding of the Royal Society: Series A*, Vol. 459, 2003, pp. 3 - 46.

9. **Blatz-Ko Model. (Unconstrained):** In Problem 4.4.1 (page 264) we gave the expression for the strain energy function for a Blatz-Ko material and examined its response in uniaxial stress and simple shear. This constitutive model was proposed by Blatz and Ko based on their experiments on a foam rubber.

   Reference: R. J. Blatz and W.L. Ko, Application of finite elastic theory to the deformation of rubbery materials, *Transactions of the Society of Rheology*, volume 6 (1962), pp. 223-251 .

10. **Ogden model. (Unconstrained):** The Ogden strain energy function[18] for unconstrained materials is

$$W(\lambda_1, \lambda_2, \lambda_3) = \sum_{i=1}^{M} a_i \phi(\alpha_i) + \sum_{i=1}^{N} b_i \psi(\beta_i) + h(J), \qquad J = \lambda_1\lambda_2\lambda_3,$$

where

$$\phi(\alpha) = \lambda_1^\alpha + \lambda_2^\alpha + \lambda_3^\alpha - 3, \qquad \psi(\beta) = (\lambda_1\lambda_2)^\beta + (\lambda_2\lambda_3)^\beta + (\lambda_3\lambda_1)^\beta - 3$$

and

$$a_i > 0,\ b_i > 0, \quad \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_M \geq 1, \quad \beta_1 \geq \beta_2 \geq \ldots \geq \beta_M \geq 1.$$

[18] The original form proposed by Ogden does not include the functions ψ. For the modified form shown here, see for example the paper by Ball [5].

Reference: R. W. Ogden, Large deformation isotropic elasticity - On the correlation of theory and experiment for compressible rubberlike solids, *Proceedings of the Royal Society of London. Series A*, Vol. 328, Issue 1575, (1972), pp. 567-583.

11. **Some incompressible anisotropic material models:**

As we shall see in Chapter 6, the strain energy function $W(\mathbf{C})$ characterizing an anisotropic elastic material with one preferred direction (one family of fibers) involves, in addition to the invariants I_1, I_2 and I_3, two other (pseudo) invariants I_4 and I_5. If the material has two preferred directions, its strain strain energy function involves three additional invariants I_6, I_7 and I_8.

An example of such a constitutive model that has been proposed by Holzapfel et. al. [12] as a model for soft biological tissues is

$$W(I_1, I_4, I_6) = \frac{\mu_1}{2}(I_1-3)+\frac{1}{2}\frac{\mu_4}{k_4}\Big[\exp[k_4(I_4-1)^2]-1\Big]+\frac{1}{2}\frac{\mu_6}{k_6}\Big[\exp[k_6(I_6-1)^2]-1\Big], \quad (i)$$

where $\mu_1, \mu_4, \mu_6, k_1, k_4, k_6$ are material constants and

$$I_4 = \mathbf{C}\mathbf{m}_R \cdot \mathbf{m}_R, \qquad I_6 = \mathbf{C}\mathbf{m}'_R \cdot \mathbf{m}'_R. \qquad (ii)$$

Here the unit vectors $\mathbf{m}_R$ and $\mathbf{m}'_R$ denote the preferred directions in the reference configuration. Observe that the term involving I_1 in (i) has the neo-Hookean form while the terms involving I_4, I_6 are of the Fung form. If $I_4 - 1$ and $I_6 - 1$ are small, (i) can be replaced by

$$W = \frac{\mu_1}{2}(I_1 - 3) + \frac{\mu_4}{2}(I_4 - 1)^2 + \frac{\mu_6}{2}(I_6 - 1)^2. \qquad (iii)$$

The special case of (iii) corresponding to $\mu_1 = \mu, \mu_4 = \mu_6 = \beta\mu$ is referred to as the **standard fiber reinforcing model**,

$$W = \frac{\mu}{2}(I_1 - 3) + \frac{\mu\beta}{2}\Big[(I_4 - 1)^2 + (I_6 - 1)^2\Big], \qquad \mu > 0, \beta > 0; \qquad (iv)$$

see Goriely [8].

Problem 4.7.1. Recall the bending of a block, the kinematics of which were analyzed in Problem 2.5.4 with the equilibrium equations and boundary conditions examined in Problem 3.7.1. Suppose that the material is composed a Blatz-Ko material. Use this added information to complete the solution to that problem.

Solution: Recall from Problems 2.5.4 and 3.7.1 that

$$\mathbf{F} = \lambda_1 \mathbf{e}_r \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_\theta \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_z \otimes \mathbf{e}_3, \qquad (i)$$

$$\mathbf{S} = \sigma_1 \mathbf{e}_r \otimes \mathbf{e}_1 + \sigma_2 \mathbf{e}_\theta \otimes \mathbf{e}_2 + \sigma_3 \mathbf{e}_z \otimes \mathbf{e}_3, \qquad (ii)$$

where

$$\lambda_1 = r'(x_1), \qquad \lambda_2 = \alpha r(x_1), \qquad \lambda_3 = \Lambda, \qquad (iii)$$

$$\sigma_1(x_1) = \widehat{\sigma}_1(\lambda_1, \lambda_2, \lambda_3), \qquad \sigma_2(x_1) = \widehat{\sigma}_2(\lambda_1, \lambda_2, \lambda_3), \qquad \sigma_3(x_1) = \widehat{\sigma}_3(\lambda_1, \lambda_2, \lambda_3), \qquad (iv)$$

and equilibrium required

$$\frac{d\sigma_1}{dx_1} - \alpha\sigma_2 = 0 \qquad \text{for} \quad -A \leq x_1 \leq A. \qquad (v)$$

In the absence of a constitutive relation we had been unable to proceed further.

Now that we are told that the block is composed of a Blatz-Ko material, we have

$$\sigma_1 = \mu\big(\lambda_2\lambda_3 - \lambda_1^{-3}\big), \qquad \sigma_2 = \mu\big(\lambda_1\lambda_3 - \lambda_2^{-3}\big), \qquad (vi)$$

which upon using (iii) gives

$$\sigma_1 = \mu\left(\alpha\Lambda r(x_1) - \frac{1}{[r'(x_1)]^3}\right), \qquad \sigma_2 = \mu\left(\Lambda r'(x_1) - \frac{1}{[\alpha r(x_1)]^3}\right). \qquad (vii)$$

Substituting (vii) into the equilibrium equation (v) and simplifying leads to (the nonlinear ordinary differential equation)

$$\frac{r''}{(r')^4} + \frac{1}{3r^3\alpha^2} = 0 \qquad \text{for} \quad -A \leq x_1 \leq A. \qquad (viii)$$

Integrating once gives

$$r' = \frac{\alpha r\sqrt{3}}{\sqrt{c_1 r^2 - 1}} \qquad \text{for} \quad -A \leq x_1 \leq A, \qquad (ix)$$

and integrating again yields

$$x_1 = \frac{1}{\alpha\sqrt{3}}\Big[\sqrt{c_1 r^2(x_1) - 1} - \tan^{-1}\sqrt{c_1 r^2(x_1) - 1} + c_2\Big]. \qquad \text{for} \quad -A \leq x_1 \leq A, \qquad (x)$$

The boundary conditions

$$\sigma_1(\pm A) = 0, \qquad 2C\int_{-A}^{A} r(x_1)\sigma_2(x_1)dx_1 = m,$$

are now to be used to determine the unknown constants c_1, c_2, α.

4.8 Linearized elasticity.

In each of the preceding chapters on kinematics and stress, we specialized the general theory to the case where the displacement gradient tensor $\mathbf{H} = \nabla\mathbf{u} = \mathbf{F} - \mathbf{I}$ was infinitesimal: $|\mathbf{H}| \ll 1$. In particular, we found that *all* of the general strain measures $\mathbf{E}(\mathbf{U})$ and $\mathcal{E}(\mathbf{V})$ defined in (2.67) and (2.70) reduce to the *infinitesimal strain tensor* $\boldsymbol{\varepsilon}$:

$$\boldsymbol{\varepsilon} = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T), \qquad \varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{4.143}$$

We also found that the right Cauchy-Green deformation tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ yielded

$$\mathbf{C} = \mathbf{I} + 2\boldsymbol{\varepsilon} + O(|\mathbf{H}|^2). \tag{4.144}$$

In the case of stress, the Cauchy stress and the 1st Piola-Kirchhoff stress agreed to leading order and we denoted this common *stress* by the symmetric tensor $\boldsymbol{\sigma}$.

We now approximate the general constitutive relationship for an elastic solid to this case where the displacement gradient tensor is infinitesimal.

First consider the *constitutive relation for stress*: it is convenient to start with the expression $(4.21)_1$ for the 1st Piola-Kirchhoff stress:

$$\mathbf{S} = 2\mathbf{F}\frac{\partial W}{\partial \mathbf{C}}(\mathbf{C}). \tag{4.145}$$

In the reference configuration we have $\mathbf{F} = \mathbf{I}, \mathbf{C} = \mathbf{I}$. Substituting this into the right hand side of (4.145) gives the stress in the reference configuration which we denote by $\overset{o}{\boldsymbol{\sigma}}$:

$$\overset{o}{\boldsymbol{\sigma}} = 2\left.\frac{\partial W}{\partial \mathbf{C}}\right|_{\mathbf{C}=\mathbf{I}}. \tag{4.146}$$

This is called the *residual stress*. From hereon we shall take the residual stress to vanish:

$$\left.\frac{\partial W}{\partial \mathbf{C}}\right|_{\mathbf{C}=\mathbf{I}} = \mathbf{0}. \tag{4.147}$$

We now approximate (4.145) by using $\mathbf{F} = \mathbf{I} + \mathbf{H}$, $\mathbf{C} = \mathbf{I} + 2\boldsymbol{\varepsilon}$ and carrying out a Taylor expansion of its right hand side :

$$\begin{aligned}
\sigma_{ij} = 2F_{ik}\frac{\partial W}{\partial C_{kj}}(\mathbf{C}) \quad &= \quad 2\left(\delta_{ik} + \frac{\partial u_i}{\partial x_k}\right)\frac{\partial W}{\partial C_{kj}}(\mathbf{I} + 2\boldsymbol{\varepsilon}), \\
&= \quad 2\left(\delta_{ik} + \frac{\partial u_i}{\partial x_k}\right)\left[\left.\frac{\partial W}{\partial C_{kj}}\right|_{\mathbf{C}=\mathbf{I}} + \left.\frac{\partial^2 W}{\partial C_{kj}\partial C_{pq}}\right|_{\mathbf{C}=\mathbf{I}} 2\varepsilon_{pq} + O(|\mathbf{H}|^2)\right], \\
&\overset{(4.147)}{=} \quad \mathbb{C}_{ijpq}\varepsilon_{pq} \;\; + O(|\mathbf{H}|^2),
\end{aligned}$$

where we have set

$$\mathbb{C}_{ijk\ell} = 4\frac{\partial^2 W}{\partial C_{ij} C_{k\ell}}\bigg|_{\mathbf{C}=\mathbf{I}}, \tag{4.148}$$

and used (4.146). The constitutive relation for stress in the linearized theory is thus

$$\boxed{\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon}, \qquad \sigma_{ij} = \mathbb{C}_{ijk\ell}\varepsilon_{k\ell}.} \tag{4.149}$$

The 4-tensor $\mathbb{C}$ is known as the **elasticity tensor**. Note that it does not depend on the deformation since the right-hand side of (4.148) is evaluated at $\mathbf{C} = \mathbf{I}$. Its components $\mathbb{C}_{ijk\ell}$ represent the various elastic moduli of the material. The elasticity tensor has $3^4 = 81$ components but not all of them are independent. Observe from (4.148) that

$$\mathbb{C}_{ijk\ell} = \mathbb{C}_{k\ell ij}, \qquad \mathbb{C}_{ijk\ell} = \mathbb{C}_{jik\ell}, \qquad \mathbb{C}_{ijk\ell} = \mathbb{C}_{ij\ell k}, \tag{4.150}$$

which implies that $\mathbb{C}$ has 21 independent components. Therefore the most general (anisotropic) elastic material has 21 elastic constants.

For completeness, we note the appropriate approximation for the strain energy function W. A Taylor expansion of $W(\mathbf{C})$ about $\mathbf{C} = \mathbf{I}$ is readily shown to lead to

$$\boxed{W = \frac{1}{2}\mathbb{C}_{ijk\ell}\varepsilon_{i,j}\varepsilon_{k,l}.} \tag{4.151}$$

Observe from (4.149) and (4.151) that

$$\boxed{\sigma_{ij} = \frac{\partial W}{\partial \varepsilon_{ij}}(\boldsymbol{\varepsilon}), \qquad \mathbb{C}_{ijk\ell} = \frac{\partial^2 W}{\partial \varepsilon_{ij}\varepsilon_{k\ell}}(\boldsymbol{\varepsilon}).} \tag{4.152}$$

Material symmetry in the linearized theory. Following the discussion in Section 4.4, the *material symmetry group* $\mathcal{G}$ is the collection of proper orthogonal transformations that preserves the symmetry of the material. In the present setting material symmetry tells us that

$$W(\boldsymbol{\varepsilon}) = W(\mathbf{Q}\boldsymbol{\varepsilon}\mathbf{Q}^T) \qquad \text{for each } \mathbf{Q} \in \mathcal{G} \tag{4.153}$$

and all symmetric tensors $\boldsymbol{\varepsilon}$; see (4.29). Note that the tensor $\mathbf{Q}$ (that operates on the reference configuration prior to deforming the body) need not be infinitesimal. Since

$$W(\boldsymbol{\varepsilon}) = \frac{1}{2}\mathbb{C}_{ijk\ell}\varepsilon_{ij}\varepsilon_{k\ell} \tag{4.154}$$

it follows that if $\mathbf{Q} \in \mathcal{G}$ then

$$\begin{aligned}
\tfrac{1}{2}\mathbb{C}_{ijk\ell}\varepsilon_{ij}\varepsilon_{k\ell} &= \tfrac{1}{2}\mathbb{C}_{pqrs}(\mathbf{Q}\boldsymbol{\varepsilon}\mathbf{Q}^T)_{pq}(\mathbf{Q}\boldsymbol{\varepsilon}\mathbf{Q}^T)_{rs} = \\
&= \tfrac{1}{2}\mathbb{C}_{pqrs}(Q_{pi}\varepsilon_{ij}Q_{qj})(Q_{rk}\varepsilon_{k\ell}Q_{s\ell}) = \\
&= \tfrac{1}{2}Q_{pi}Q_{qj}Q_{rk}Q_{s\ell}\mathbb{C}_{pqrs}\varepsilon_{ij}\varepsilon_{k\ell}.
\end{aligned}$$

Since this must hold for all strains $\boldsymbol{\varepsilon}$ it follows that

$$\mathbb{C}_{ijk\ell} = Q_{pi}Q_{qj}Q_{rk}Q_{s\ell}\mathbb{C}_{pqrs} \qquad \text{for each } \mathbf{Q} \in \mathcal{G}. \tag{4.155}$$

Recall that in general, the components $\mathbb{C}'_{ijk\ell}$ and $\mathbb{C}_{ijk\ell}$ of a 4-tensor $\mathbb{C}$ in two bases $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ and $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are related by $\mathbb{C}'_{ijk\ell} = Q_{pi}Q_{qj}Q_{rk}Q_{s\ell}\mathbb{C}_{pqrs}$ where the proper orthogonal matrix $[Q]$ relates the two bases by $\mathbf{e}'_i = Q_{ij}\mathbf{e}_j$. Therefore (4.155) tells us that the components of $\mathbb{C}$ in two bases are identical if the bases are related by a symmetry transformation.

For an *isotropic material*, $\mathcal{G}$ contains all proper orthogonal transformations and therefore, as in the finite deformation theory, we may conclude that W depends on the strain $\boldsymbol{\varepsilon}$ only through its scalar invariants. It is convenient to choose the invariants

$$i_i = \operatorname{tr}\boldsymbol{\varepsilon}, \qquad i_2 = \operatorname{tr}\boldsymbol{\varepsilon}^2, \qquad i_3 = \operatorname{tr}\boldsymbol{\varepsilon}^3.$$

Thus for an isotropic material we have $W(i_1, i_2, i_3)$. However W is a quadratic function of strain. Therefore it cannot depend on i_3 and it must depend linearly on i_2 and i_1^2:

$$W = \mu\, i_2 + \frac{\lambda}{2}\, i_1^2 = \mu \operatorname{tr}(\boldsymbol{\varepsilon}^2) + \frac{\lambda}{2}\,(\operatorname{tr}\boldsymbol{\varepsilon})^2 = \mu\varepsilon_{ij}\varepsilon_{ij} + \frac{\lambda}{2}\,\varepsilon_{ii}\varepsilon_{jj}, \tag{4.156}$$

where μ and λ are material parameters. (Caution: λ here is not the stretch.) It follows from (4.156) and $(4.152)_1$ that the constitutive relation for stress is

$$\boldsymbol{\sigma} = 2\mu\,\boldsymbol{\varepsilon} + \lambda \operatorname{tr}\boldsymbol{\varepsilon}\,\mathbf{I}, \qquad \sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\varepsilon_{kk}\,\delta_{ij}, \tag{4.157}$$

and from $(4.152)_2$ that

$$\mathbb{C}_{ijkl} = \mu(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) + \lambda\delta_{ij}\delta_{k\ell}. \tag{4.158}$$

Observe that an isotropic elastic material is described by 2 elastic moduli. The two moduli λ and μ are known as the Lamé constants. The more familiar material constants shear modulus G, bulk modulus κ, Young's modulus E and Poisson's ratio ν can be expressed in terms of λ and μ:

$$G = \mu, \qquad \kappa = \lambda + \frac{2}{3}\mu, \qquad E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}, \qquad \nu = \frac{\lambda}{2(\lambda + \mu)}. \tag{4.159}$$

Remark: Linearization of the strain energy function $W = W(I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C}))$ of the nonlinear theory shows, after a lengthy calculation, that the material parameters λ and μ can be expressed as

$$\mu = 2\left(\frac{\partial W}{\partial I_1} + \frac{\partial W}{\partial I_2}\right)\Big|_{I_1=I_2=3, I_3=1},$$

$$\begin{aligned}\lambda = 4\Bigg(\frac{\partial W}{\partial I_2} + \frac{\partial W}{\partial I_3} &+ \frac{\partial^2 W}{\partial I_1^2} + 4\frac{\partial^2 W}{\partial I_2^2} + \frac{\partial^2 W}{\partial I_3^2} + 4\frac{\partial^2 W}{\partial I_1 \partial I_2} + \\ &+ 2\frac{\partial^2 W}{\partial I_1 \partial I_3} + 4\frac{\partial^2 W}{\partial I_2 \partial I_3}\Bigg)\Big|_{I_1=I_2=3, I_3=1}.\end{aligned}$$

For an incompressible material one has the constraint

$$\operatorname{tr} \boldsymbol{\varepsilon} = 0, \qquad \varepsilon_{kk} = 0, \tag{4.160}$$

and a reactive stress $-q\mathbf{I}$ must be added to the expression for stress:

$$\sigma_{ij} = \mathbb{C}_{ijk\ell}\varepsilon_{k\ell} - q\delta_{ij}.$$

If the material is isotropic, this reduces to

$$\boldsymbol{\sigma} = 2\mu\, \boldsymbol{\varepsilon} - q\, \mathbf{I}, \qquad \sigma_{ij} = 2\mu\varepsilon_{ij} - q\, \delta_{ij}. \tag{4.161}$$

Note that (4.161) can be formally obtained from (4.157) by letting $\operatorname{tr} \boldsymbol{\varepsilon} \to 0$ and $\lambda \to \infty$ with $\lambda \operatorname{tr} \boldsymbol{\varepsilon}$ held constant. Observe that taking this limit in the expression (4.156) for the strain energy yields

$$W = \mu\varepsilon_{ij}\varepsilon_{ij} - \frac{1}{2}\, q\varepsilon_{ii} \overset{(4.160)}{=} \mu\varepsilon_{ij}\varepsilon_{ij},$$

where in the first expression q plays the role of a Lagrange multiplier associated with the constraint $\varepsilon_{ii} = 0$.

Exercise: Problem 4.17.

‘

4.9 Exercises.

Response of an elastic material.

Problem 4.1. Carry out the calculations described in Problem 4.1.1 and show that $\widehat{\mathbf{T}}(\mathbf{B})$ must necessarily have the form

$$\widehat{\mathbf{T}}(\mathbf{B}) = \beta_2 \mathbf{B}^2 + \beta_1 \mathbf{B} + \beta_0 \mathbf{I}, \qquad (i)$$

where the β_j's are functions of the principal scalar invariants of $\mathbf{B}$. This restricted form of $\widehat{\mathbf{T}}(\mathbf{B})$ implies the material is isotropic. (The result of Problem 1.36 will be useful.)

Problem 4.2. (Based on Holzapfel) The *Saint-Venant Kirchhoff model* of an isotropic unconstrained material is described by the strain energy function

$$W = \frac{\alpha}{2}(\operatorname{tr}\mathbf{E})^2 + \mu \operatorname{tr}(\mathbf{E}^2), \qquad (i)$$

where $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$ is the Green Saint-Venant strain tensor and $\mu > 0$ and $\alpha > 0$ are material constants. (The symbol λ is usually used for the parameter α, but since we use λ for stretch, we are using α instead.)

(a) Does this strain energy function W obey the growth conditions (4.119)? i.e. does $W \to \infty$ as $J \to \infty$ and $J \to 0^+$?

(b) Derive the constitutive law associated with (i) relating the Green Saint-Venant strain $\mathbf{E}$ to the stress that is work-conjugate to it. (Hint: Recall that the second Piola Kirchhoff stress tensor $\mathbf{S}^{(2)}$ is work-conjugate to the Green Saint-Venant strain.)

(c) Consider a uniaxial *deformation* $\mathbf{y} = \mathbf{F}\mathbf{x}$ where $\mathbf{F} = \mathbf{I} + (\lambda - 1)\mathbf{e}_1 \otimes \mathbf{e}_1$. $(c1)$: Calculate the stress-stretch relation between S_{11} and λ where S_{11} is the 1,1 component of the *first* Piola Kirchhoff stress tensor $\mathbf{S}$; $(c2)$: show that this relation loses monotonicity in compression at the stretch $\lambda = \sqrt{1/3}$; and $(c3)$ that $S_{11} \to 0$ at extreme contraction $\lambda \to 0^+$.

(d) Show for the modified Saint-Venant Kirchhoff model

$$W = \frac{\kappa}{2}(\ln J)^2 + \mu \operatorname{tr}(\mathbf{E}^2), \qquad (ii)$$

where $\mu > 0$ and κ are material constants, that $S_{11} \to -\infty$ as $\lambda \to 0^+$ in a uniaxial deformation.

Problem 4.3. (Based on Chadwick) The constitutive relation for a certain class of foam rubbers has the form

$$\mathbf{T} = \frac{1}{J^3}\Big[\big[f(J) - \beta I_2\big]\mathbf{I} + \beta I_1 \mathbf{B} - \beta \mathbf{B}^2\Big] \tag{i}$$

where β is a constitutive parameter and $f(J)$ is a constitutive function.

(a) A uniaxial stress experiment is carried out in order to determine the function $f(J)$. Let λ be the stretch in the direction of the applied stress. During the experiment λ and the transverse stretch Λ are measured. A plot of Λ versus λ on a logarithmic scale is found to be a straight line with slope $-\nu$ where ν is a positive constant. Deduce the form of the function f.

(b) Calculate the Cauchy stress - stretch relation in uniaxial stress and from it determine the Young's modulus of the material.

Problem 4.4. (Spencer) Consider a body composed of a neo-Hookean material. In a reference configuration it occupies the unit cube $0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq 1$ and undergoes the deformation

$$y_1 = \lambda x_1 + k x_2, \qquad y_2 = \lambda^{-1} x_2, \qquad y_3 = x_3. \tag{o}$$

(a) Sketch the region occupied by the body in the deformed configuration noting the lengths of the edges.

(b) Calculate the components of the Cauchy and 1st Piola-Kirchhoff stress tensors.

(c) Suppose the faces $x_3 = 0$ and $x_3 = 1$ are known to be traction-free. Simplify your answer to part (b).

(d) Calculate the force that must be applied to the face (that in the reference configuration corresponded to) $x_2 = 1$.

(e) Determine the (true) Cauchy traction that must be applied on the face (that in the reference configuration corresponded to) $x_1 = 1$.

Problem 4.5. (Ball) Consider an incompressible isotropic elastic material characterized by the strain energy function

$$W = \alpha(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) + \beta(\lambda_1^{-1} + \lambda_2^{-1} + \lambda_3^{-1} - 3), \tag{i}$$

where α and β are material constants.

(a) What restrictions (if any) do the Baker-Ericksen inequalities (4.104) impose on α and β?

(b) Assume that the restrictions determined in (a) hold. Consider a state of uniaxial stress in the x_3-direction. Show that the graph of the 1st Piola-Kirchhoff stress S_{33} versus λ_3 is not everywhere increasing if

$$\beta^2 > 64\alpha(2\alpha + \beta). \tag{ii}$$

(This is related to the phenomenon of necking, in which a material can have an instability in tension leading to a greater extension and thinner deformed cross-section.)

Problem 4.6. For a particular isotropic, incompressible, elastic material, the stress-stretch relation (for the Cauchy stress) in uniaxial stress is

$$\tau = \mu(\lambda^2 - \lambda^{-1}),$$

where $\mu > 0$ is a material parameter. Note that $\tau \to -\infty$ when $\lambda \to 0^+$ and $\tau \to \infty$ when $\lambda \to \infty$.

Determine *two* strain energy functions $W(I_1, I_2)$ that yield this same stress-stretch relation in uniaxial stress. For each W, calculate the corresponding relation between the shear stress and amount of shear in simple shear.

Problem 4.7. Can you construct an explicit example of a strain energy function for an isotropic *unconstrained* material that has a Poisson's ratio (at infinitesimal deformations) equal to 1/2 even though the material is compressible at finite deformations? Ensure that the energy and stress in the reference configuration vanish; that the Baker-Ericksen inequalities holds; that $W \to \infty$ when $J \to 0^+$ and $J \to \infty$; and the Legendre Hadamard condition (page 285) at $\mathbf{F} = \mathbf{I}$ holds. If you are able to construct such a W, determine and sketch a graph of the associated pressure-volume relation in pure dilatation.

Problem 4.8. (Anand [1]) Let

$$E_k = \ln \lambda_k, \qquad (i)$$

denote the principal logarithmic (Hencky) strains. Consider an isotropic unconstrained elastic material characterized by the strain energy function

$$W(E_1, E_2, E_3) = \alpha(E_1^2 + E_2^2 + E_3^2) + \frac{\beta}{2}\left(E_1 + E_2 + E_3\right)^2, \qquad (ii)$$

where α and β are material constants.

(a) Determine the linearized form of W at infinitesimal deformations and thus interpret α and β.

(b) Calculate the relation between the principal Cauchy stress components τ_k and the principal logarithmic strains E_k.

(c) Calculate the relation between τ_1 and E_1 in a uniaxial deformation $E_2 = E_3 = 1$.

(d) Calculate the relation between τ_1 and E_1, and also between E_2 and E_1, in uniaxial stress $\tau_1 \neq 0, \tau_2 = \tau_3 = 0$.

(e) Consider a so-called pure shear deformation $\lambda_1 = \lambda, \lambda_2 = \lambda^{-1}, \lambda_3 = 1$. Calculate the relation between τ_1 and λ.

(f) Consider the incompressible counterpart of this material and calculate the relation between τ_1 and λ in pure shear.

Some general considerations.

Problem 4.9. Show that an elastic material is isotropic if and only if

$$\mathbf{TB} = \mathbf{BT} \qquad (i)$$

where $\mathbf{T}$ is the Cauchy stress tensor and $\mathbf{B}$ is the left Cauchy-Green deformation tensor.

Problem 4.10. If an elastic material is isotropic, show that

$$\mathbf{S}^T\mathbf{R} = \mathbf{R}^T\mathbf{S}, \qquad (i)$$

and hence that the Biot stress tensor can be written (in this case) as

$$\mathbf{S}^{(1)} = \mathbf{S}^T\mathbf{R}. \qquad (ii)$$

Problem 4.11. An elastic material that is not hyperelastic is called a Cauchy elastic material. One cannot associate a strain energy function $W(\mathbf{F})$ with a Cauchy elastic material. The Cauchy stress in such a material is related to the deformation through

$$\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{F}), \qquad (i)$$

where the stress response function $\widehat{\mathbf{T}}$ is defined for all tensors with positive determinant. Explain why material frame indifference requires

$$\widehat{\mathbf{T}}(\mathbf{QF}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^T, \qquad (ii)$$

for all proper orthogonal tensors $\mathbf{Q}$ and all tensors $\mathbf{F}$ with positive determinant. Explain also why the material symmetry group $\mathcal{G}$ for such a material is defined to be the set of all proper orthogonal tensors $\mathbf{Q}$ for which

$$\widehat{\mathbf{T}}(\mathbf{F}) = \widehat{\mathbf{T}}(\mathbf{FQ}), \qquad \mathbf{Q} \in \mathcal{G}, \qquad (iii)$$

for all tensors $\mathbf{F}$ with positive determinant .

(a) Show that (ii) holds if and only if

$$\widehat{\mathbf{T}}(\mathbf{F}) = \mathbf{R}\widehat{\mathbf{T}}(\mathbf{U})\mathbf{R}^T. \qquad (v)$$

(b) Assume that (ii) holds. Show that (iii) holds for a particular $\mathbf{Q} \in \mathcal{G}$ if and only if

$$\widehat{\mathbf{T}}(\mathbf{QFQ}^T) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^T, \qquad (iv)$$

for all tensors $\mathbf{F}$ with positive determinant.

(c) Show that a material is isotropic if and only if

$$\widehat{\mathbf{T}}(\mathbf{QVQ}^T) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{V})\mathbf{Q}^T. \qquad (vi)$$

(d) If the material is isotropic show that

$$\widehat{\mathbf{T}}(\mathbf{V})\mathbf{V} = \mathbf{V}\widehat{\mathbf{T}}(\mathbf{V}). \tag{vii}$$

Problem 4.12. For an isotropic material one can express the Cauchy stress tensor $\mathbf{T}$ as

$$\mathbf{T} = \sum_{i=1}^{3} \tau_i \boldsymbol{\ell}_i \otimes \boldsymbol{\ell}_i, \tag{i}$$

where τ_1, τ_2, τ_3 are the principal Cauchy stresses and $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ are the principal directions of both $\mathbf{T}$ and the Eulerian stretch tensor $\mathbf{V}$. Show that the first Piola Kirchhoff stress tensor $\mathbf{S}$ and the Biot stress tensor $\mathbf{S}^{(1)}$ (defined in (3.81)) can be written in the respective forms

$$\mathbf{S} = \sum_{i=1}^{3} \frac{J\tau_i}{\lambda_i} \boldsymbol{\ell}_i \otimes \mathbf{r}_i, \qquad \mathbf{S}^{(1)} = \sum_{i=1}^{3} \frac{J\tau_i}{\lambda_i} \mathbf{r}_i \otimes \mathbf{r}_i. \tag{4.162}$$

Observe that this, together with (4.46), yields (4.47) and (4.48).

Problem 4.13. (*Biot stress*) It was shown in Problem 3.29 that the Biot stress

$$\mathbf{S}^{(1)} = \frac{1}{2}\left(\mathbf{S}^T\mathbf{R} + \mathbf{R}^T\mathbf{S}\right), \tag{4.163}$$

is work conjugate to the Lagrangian stretch tensor $\mathbf{U}$:

$$\mathbf{S}^{(1)} \cdot \dot{\mathbf{U}} = \mathbf{S} \cdot \dot{\mathbf{F}}. \tag{4.164}$$

(a) Show that $\mathbf{S}^{(1)}$ obeys the constitutive relation

$$\mathbf{S}^{(1)} = \frac{\partial \widehat{W}}{\partial \mathbf{U}}(\mathbf{U}). \tag{4.165}$$

(b) For an isotropic elastic material show that the principal components of the Biot stress obey

$$S_k^{(1)} = \frac{\partial W^*}{\partial \lambda_k}(\lambda_1, \lambda_2, \lambda_3), \qquad k = 1, 2, 3, \tag{4.166}$$

with associated principal directions $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ (the eigenvectors of $\mathbf{U}$), and therefore that

$$\mathbf{S}^{(1)} = \sum_{k=1}^{3} \frac{\partial W^*}{\partial \lambda_k} \mathbf{r}_k \otimes \mathbf{r}_k. \tag{4.167}$$

(c) When the material is subjected to the constitutive constraint $\phi(\mathbf{U}) = 0$, show that

$$\mathbf{S}^{(1)} = \widehat{\mathbf{S}}^{(1)}(\mathbf{U}) - q\frac{\partial \phi}{\partial \mathbf{U}} \tag{4.168}$$

(d) For an incompressible isotropic material show that the principal components of the Biot stresses are

$$S_k^{(1)} = \frac{\partial W^*}{\partial \lambda_k} - q\lambda_k^{-1}, \qquad k = 1, 2, 3. \tag{4.169}$$

Problem 4.14. In Problem 2.40 we decomposed the deformation gradient tensor multiplicatively into the product of the hydrostatic tensor $J^{1/3}\mathbf{I}$ that described the volume change, and a second tensor $\overline{\mathbf{F}}$ that characterized the shape change:

$$\mathbf{F} = J^{1/3}\overline{\mathbf{F}}, \qquad J = \det \mathbf{F}. \tag{i}$$

In this problem you are to decompose the constitutive relation for the Cauchy stress into a part due to the volume change and a part due to the shape change.

The modified left Cauchy-Green tensor $\overline{\mathbf{B}}$ associated with $\overline{\mathbf{F}}$, and its scalar invariants $\overline{I}_1, \overline{I}_2$ and $\overline{I}_3$, were shown to obey

$$\overline{\mathbf{B}} = J^{-2/3}\mathbf{B}, \qquad \overline{I}_1 = J^{-2/3}I_1, \qquad \overline{I}_2 = J^{-4/3}I_2, \qquad \overline{I}_3 = 1. \tag{ii}$$

Here I_1, I_2 and $I_3 = J^2$ are the principal scalar invariants of $\mathbf{B}$. Show that there is a one-to-one relation between $\{I_1, I_2, J\}$ and $\{\overline{I}_1, \overline{I}_2, J\}$ and therefore that the strain energy function for an unconstrained isotropic elastic material can be written as

$$W = \mathcal{W}(\overline{I}_1, \overline{I}_2, J). \tag{iii}$$

Note that $\overline{\mathbf{F}}, \overline{\mathbf{B}}, \overline{I}_1$ and $\overline{I}_2$ are associated with the shape change. Show that the constitutive relation for the Cauchy stress can be expressed as

$$\mathbf{T} = \frac{2}{J}\left[-\frac{1}{3}\left(\overline{I}_1\mathcal{W}_1 + 2\overline{I}_2\mathcal{W}_2\right)\mathbf{I} + \left(\mathcal{W}_1 + \overline{I}_1\mathcal{W}_2\right)\overline{\mathbf{B}} - \mathcal{W}_2\overline{\mathbf{B}}^2\right] + \frac{\partial \mathcal{W}}{\partial J}\mathbf{I} \tag{iv}$$

where we have written

$$\mathcal{W}_\alpha = \frac{\partial \mathcal{W}}{\partial \overline{I}_\alpha}, \qquad \alpha = 1, 2.$$

Observe the additive decomposition of the constitutive relation (iv) into a part determined by $\overline{\mathbf{F}}$ and a part determined by J.

Problem 4.15. (Related to Ericksen's problem on universal deformations.)

Consider a body composed of an isotropic incompressible elastic material that is in equilibrium with no body forces. The traction boundary conditions lead to a deformation $\mathbf{y} = \mathbf{y}(\mathbf{x})$ for which the associated principal scalar invariants $I_1(\mathbf{B})$ and $I_2(\mathbf{B})$ are constants independent of $\mathbf{x}$. Does this imply that the deformation is homogeneous?

Problem 4.16. *Energy-Momentum Tensor*

The tensor

$$\mathbf{P} = W(\mathbf{F})\,\mathbf{I} - \mathbf{F}^T\mathbf{S} \tag{4.170}$$

is known as the **Energy-Momentum Tensor**. Consider a homogeneous elastic body that is in equilibrium (with no body forces).

(a) Show that

$$\operatorname{Div}\mathbf{P} = \mathbf{o} \qquad \text{at all } \mathbf{x} \in \mathcal{R}_R. \tag{i}$$

(b) Suppose that the body contains a cavity in its interior. Let $\mathcal{S}$ be an arbitrary closed surface in the body that encloses the cavity. Show that the value of the integral

$$\int_{\mathcal{S}} \mathbf{P}\mathbf{n}_R\, dA_x \tag{ii}$$

is the same for all such surfaces. (*Remark:* This result underlies the path-independent nature of the famous J-integral of fracture mechanics.)

(c) If the material is isotropic, show that

$$\mathbf{P} = \mathbf{P}^T. \tag{iii}$$

Problem 4.17. In terms of the strain energy function $W^*(\lambda_1, \lambda_2, \lambda_3)$ for an isotropic unconstrained material, show that

$$\left.\frac{\partial^2 W^*}{\partial \lambda_i^2}\right|_{\lambda_1=\lambda_2=\lambda_3=1} = \lambda + 2\mu = \kappa + \frac{4}{3}\mu, \qquad \left.\frac{\partial^2 W^*}{\partial \lambda_i \partial \lambda_j}\right|_{\lambda_1=\lambda_2=\lambda_3=1} = \lambda = \kappa - \frac{2}{3}\mu \qquad i \neq j,$$

where λ is a Lamè constant and κ and μ are the bulk and shear moduli respectively at infinitesimal deformations.

Problem 4.18. Show that the constitutive relation (4.37)$_1$ for an isotropic unconstrained material can be written equivalently as

$$\mathbf{T} = \frac{2}{J}\left(I_2 \frac{\partial \widetilde{W}}{\partial I_2} + I_3 \frac{\partial \widetilde{W}}{\partial I_3} \right)\mathbf{I} \;+\; \frac{2}{J}\frac{\partial \widetilde{W}}{\partial I_1}\mathbf{B} \;-\; 2J\frac{\partial \widetilde{W}}{\partial I_2}\mathbf{B}^{-1}. \tag{4.171}$$

Materials with internal constraints.

Problem 4.19.

(a) Consider a body that is inextensible in a direction $\mathbf{m}_R$ (in the reference configuration). It may, for example, involve a family of very stiff fibers in that direction. Determine the corresponding reactive stress that needs to be added to the constitutively determined part of the stress for both the Cauchy stress and the 1st Piola-Kirchhoff stress. Physically interpret the reactive part of the Cauchy stress.

(b) Now reconsider Problem 2.3 where the body involved two families of inextensible fibers in directions $\mathbf{m}_R^+$ and $\mathbf{m}_R^-$ and the material was incompressible. Write down the reaction stress that needs to be added to the constitutively determined part of the Cauchy stress.

Problem 4.20. Reconsider the problem considered previously in Chapter 2 and illustrated in Figure 4.6. The material is incompressible and there are two families of inextensible fibers. In the reference configuration

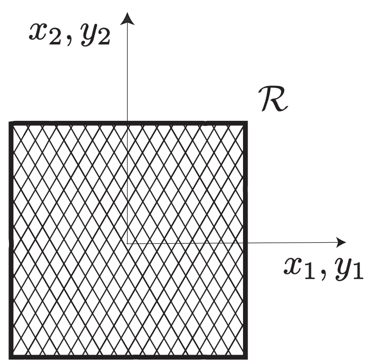


Figure 4.6: Region occupied (in a reference configuration) by an incompressible rectangular block with two families of inextensible fibers.

the fiber directions $\mathbf{m}_1$ and $\mathbf{m}_2$ are

$$\mathbf{m}_1 = \cos\Theta\,\mathbf{e}_1 + \sin\Theta\,\mathbf{e}_2, \qquad \mathbf{m}_2 = \cos\Theta\,\mathbf{e}_1 - \sin\Theta\,\mathbf{e}_2, \qquad 0 < \Theta < \pi/2. \tag{i}$$

The two faces perpendicular to the y_3-axis are traction-free. The four faces perpendicular to the y_1- and y_2-axes are free of shear traction. Normal tractions are applied on these faces leading to the homogeneous stress state

$$\mathbf{T} = T_{11}\mathbf{e}_1 \otimes \mathbf{e}_1 + T_{22}\mathbf{e}_2 \otimes \mathbf{e}_2, \tag{ii}$$

in the body. Assume the resulting deformation gradient tensor to have the form

$$\mathbf{F} = \lambda_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2\mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3\mathbf{e}_3 \otimes \mathbf{e}_3. \tag{iii}$$

The constitutive relation for this material is such that the stress is *only* due to the reaction stresses, i.e. assume that

$$\mathbf{T} = q_0\,\mathbf{I} + q_1\mathbf{F}\mathbf{m}_1 \otimes \mathbf{F}\mathbf{m}_1 + q_2\mathbf{F}\mathbf{m}_2 \otimes \mathbf{F}\mathbf{m}_2, \tag{iv}$$

for some q_0, q_1, q_2. Given T_{11}, T_{22} and Θ, calculate λ_1, λ_2 and λ_3. Discuss your results.

Problem 4.21. Here we consider a body subjected to the following kinematic constraint: let the unit vector $\mathbf{m}_\mathrm{R}$ denote a direction in the reference configuration, and suppose that the area of any plane normal to $\mathbf{m}_\mathrm{R}$ cannot change. (Though the body is treated as a homogeneous continuum, it might, for example, be a solid that has a family of stiff parallel planes aligned normal to the direction $\mathbf{m}_\mathrm{R}$.) Determine the corresponding reaction stress.

Problem 4.22. Ericksen has suggested that certain elastic crystals obey the kinematic constraint

$$\operatorname{tr}\mathbf{C} = 3.$$

(a) Determine the associated reaction stress that should be added to the Cauchy stress. (b) Write down the relation between $\mathbf{T}$ and $\mathbf{B}$ for an isotropic Ericksen material.

(c) Show that for infinitesimal deformations Ericksen's constraint is equivalent to the incompressibility constraint. (d) Show for finite plane strain deformations that the only deformation that simultaneously satisfies the incompressibility constraint and Ericksen's constraint is a rigid deformation. (e) Is this true for all finite deformations?

Reference: J. L. Ericksen, Constitutive theory for some constrained elastic crystals, *International Journal of Solids and Structures*, Vol. 22, 1986, pp. 951-964.

Problem 4.23. (Chadwick) The only deformations that a particular body can undergo are those that preserve the angle between pairs of material fibers that, in the reference configuration, lie in the directions $\mathbf{m}_\mathrm{R}$ and $\mathbf{n}_\mathrm{R}$. Determine the associated reactive stress to be added to the Cauchy stress tensor.

Restrictions on W.

Problem 4.24. (*Strong ellipticity.*) Consider a homogeneous body that is in equilibrium under the deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ where $\mathbf{F}$ is a constant tensor with positive determinant. Consider the (time-dependent) motion

$$\mathbf{y}(\mathbf{x}, t) = \mathbf{F}\mathbf{x} + \mathbf{u}(\mathbf{x}, t) \qquad \text{where} \quad |\nabla\mathbf{u}| \ll 1. \tag{i}$$

This is a small perturbation superposed on the given homogeneous deformation. Show that the equation of motion $\operatorname{Div}\mathbf{S} = \rho_\mathrm{R}\ddot{\mathbf{y}}$ when linearized about the deformation $\mathbf{y} = \mathbf{F}\mathbf{x}$ reads

$$\mathbb{A}_{pqrs}\frac{\partial^2 u_r}{\partial x_q \partial x_s} = \rho_\mathrm{R}\ddot{u}_p \qquad \text{where} \quad \mathbb{A}_{pqrs}(\mathbf{F}) := \frac{\partial^2 W(\mathbf{F})}{\partial F_{pq}\partial F_{ra}}. \tag{ii}$$

Suppose that the motion $\mathbf{u}(\mathbf{x}, t)$ is a plane harmonic wave propagating in a direction $\mathbf{n}$ with wave speed c and wave number k, the particle motion being in a direction $\mathbf{a}$:

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{a}\ \mathrm{e}^{ik(\mathbf{n}\cdot\mathbf{x}-ct)}. \qquad (iii)$$

Here i is the unit imaginary number ($i^2 = -1$, not an integer), $\mathbf{a}$ and $\mathbf{n}$ are constant unit vectors and the scalars c and k are constants.

Show using (ii) that the wave speed c is given by

$$\rho_{\mathrm{R}} c^2 = \mathbb{A}_{pqrs} a_p n_q a_r n_s,$$

and therefore that material stability requires the strong ellipticity condition (4.110) to hold for all directions $\mathbf{a}$ and $\mathbf{n}$.

See Problem 4.25 for the case of an incompressible material.

Problem 4.25. (*Strong ellipticity for incompressible material.*)

By carrying out calculations analogous to those in Problem 4.24, derive the conditions for strong ellipticity for an incompressible elastic material. The calculations in Chapter 5.7.2 will be helpful.

Problem 4.26. Consider an (isotropic, incompressible) generalized neo-Hookean material characterized by the strain energy function $W(I_1)$.

(a) Determine the restrictions imposed on W by the Baker-Ericksen inequalities.

(b) Find necessary and sufficient conditions for strong ellipticity.

Problem 4.27. *Baker-Ericksen:* Determine the restrictions imposed by the Baker-Ericksen inequalities (4.104) on a Mooney-Rivlin material

$$W = \frac{\mu}{2}\alpha(I_1 - 3) + \frac{\mu}{2}(1-\alpha)(I_1 - 3).$$

Do **not** assume the inequalities given in (4.130) (i.e. $\mu > 0, 0 < \alpha < 1$) to hold.

Problem 4.28. *Baker-Ericksen:* Consider a uniaxial tensile stress state

$$\mathbf{T} = T\mathbf{e}_3 \otimes \mathbf{e}_3, \qquad T > 0,$$

in an isotropic elastic material. Show that the Baker-Ericksen inequalities (4.104) hold if and only if the principal stretches obey

$$\lambda_3 > \lambda_1 = \lambda_2 > 0. \qquad (o)$$

Problem 4.29. *Strong ellipticity:* The notion of strong ellipticity was introduced previously in (4.108), (4.110).

Suppose the material is isotropic: $W(\mathbf{F}) = W(\lambda_1, \lambda_2, \lambda_3)$. By taking

$$\mathbf{F} = \lambda_1\, \mathbf{e}_1 \otimes \mathbf{e}_1 \ + \lambda_2\, \mathbf{e}_2 \otimes \mathbf{e}_2 \ + \lambda_3\, \mathbf{e}_3 \otimes \mathbf{e}_3, \quad \mathbf{a} = \mathbf{e}_1, \quad \mathbf{b} = \mathbf{e}_2,$$

in the strong ellipticity condition show that the Baker-Ericksen inequalities are implied by strong ellipticity. By taking

$$\mathbf{F} = \lambda_1\, \mathbf{e}_1 \otimes \mathbf{e}_1 \ + \lambda_2\, \mathbf{e}_2 \otimes \mathbf{e}_2 \ + \lambda_3\, \mathbf{e}_3 \otimes \mathbf{e}_3, \quad \mathbf{a} = \mathbf{e}_1, \quad \mathbf{b} = \mathbf{e}_1,$$

in the strong ellipticity condition show that

$$\frac{\partial^2 W}{\partial \lambda_1^2} > 0 \qquad \Rightarrow \qquad S_{11} \text{ is an increasing function of } \lambda_1.$$

It will be useful to note that by replacing $\mathbf{b}$ by $\mathbf{F}^T\mathbf{b}$ in the strong ellipticity inequality (4.110), it can be written equivalently as

$$\mathbb{B}_{ijk\ell}a_i b_j a_k b_\ell > 0 \qquad \text{for all vectors } \mathbf{a} \text{ and } \mathbf{b} \tag{4.172}$$

where

$$\mathbb{B}_{ijk\ell}(\mathbf{F}) = \mathbb{A}_{ipkq}(\mathbf{F}) F_{jp} F_{\ell q}. \tag{4.173}$$

For an isotropic unconstrained material the components of $\mathbb{B}$ in the principal basis $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ are given in Problem 6.1.7 of Ogden [17] keeping in mind that his $\mathcal{A}^1_{0ijk} = \mathbb{B}_{ji\ell k}$:

$$\left.\begin{aligned} &\mathbb{B}_{iijj} = \lambda_j \frac{\partial \tau_i}{\partial \lambda_j} + (1 - \delta_{ij})\tau_i,\\ &\mathbb{B}_{jiji} = \frac{\lambda_i^2\,(\tau_i - \tau_j)}{\lambda_i^2 - \lambda_j^2}, \qquad i \neq j,\ \lambda_i \neq \lambda_j,\\ &\mathbb{B}_{jiij} = \mathbb{B}_{ijji} = \mathbb{B}_{jiji} - \tau_i, \qquad i \neq j, \end{aligned}\right\} \qquad \text{where} \quad \tau_i = \frac{\lambda_i}{J}\frac{\partial W}{\partial \lambda_i}. \tag{4.174}$$

Problem 4.30. The components of the elasticity tensor $\mathbb{A}$ are defined by

$$\mathbb{A}_{ijk\ell}(\mathbf{F}) = \frac{\partial^2 W(\mathbf{F})}{\partial F_{ij} \partial F_{k\ell}}.$$

Calculate these components for an unconstrained isotropic material characterized by the strain energy function

$$W(\mathbf{F}) = f(I_1) + g(J) \qquad \text{where} \quad I_1 = \mathbf{F}\cdot\mathbf{F},\ J = \det \mathbf{F}.$$

Linearize your answer for infinitesimal deformations and identify the Lamé constants.

Problem 4.31. Show that the "compressible neo-Hookean" material[19],

$$W(I_1, I_2, I_3) = \frac{\mu}{2}(I_1 - 3) + h(J),$$

[19]This is a special case of the Hadamard material given in Problem 4.32.

(a) conforms to the Baker-Ericksen inequalities provided $\mu > 0$;

(b) is strongly elliptic at a given deformation if and only if

$$\mu > 0, \qquad J^2 h''(J) + \mu\lambda_i^2 > 0, \quad i = 1, 2, 3; \qquad \text{and}$$

(c) is strongly elliptic at all deformations if and only if

$$\mu > 0, \qquad h''(J) > 0 \quad \text{for } J > 0.$$

Problem 4.32. Determine necessary and sufficient conditions for strong ellipticity of the "compressible Mooney-Rivlin" material (also known as a Hadamard material)

$$W(I_1, I_2, I_3) = \frac{\mu}{2}\Big[\alpha(I_1 - 3) + (1-\alpha)(I_2 - 3)\Big] + h(I_3).$$

Problem 4.33. This problem asks you to construct from first principles the one-dimensional counterparts of the theory developed in Chapters 2, 3 and 4 in the setting of a *deformable elastic string*. When a body is modeled as a string, its cross-section is invisible and the body is treated as a one-dimensional object. Assume that the string lies in a plane.

Part A: Kinematics. In a reference configuration the string occupies a curve $\mathcal{R}_{\rm R}$ as illustrated in the left-hand figure in Figure 4.7. The position vector of a particle in this configuration is denoted by $\mathbf{x}(x), 0 \leq x \leq \ell_{\rm R}$, where x is arc length along this curve and $\ell_{\rm R}$ is the corresponding length of the string. A *particle* can be identified (i.e. labeled) by its coordinate x in this configuration. Thus, rather than saying "the particle located at $\mathbf{x}(x)$ in the reference configuration" we can simply say "the particle x". Since the choice of reference configuration is arbitrary, provided only that it be a configuration the body *can* occupy, the string could, for example, be straight and horizontal in this configuration in which case $\mathbf{x}(x) = x\mathbf{e}_1$.

During a motion, the position vector of particle x at time t is $\mathbf{y}(x,t)$ and the string occupies the curve $\mathcal{R}_t$ as depicted in the right-hand figure in Figure 4.7. Time is not important in this part of the problem where all calculations are to be carried out at the same instant t. Thus we need not keep referring to time.

Let $\boldsymbol{\ell}_{\rm R}$ and $\boldsymbol{\ell}$ be unit vectors tangent to the string in the reference and current configurations, and let x and s be arc lengths along the string in these respective configurations. A material fiber in the reference and current configurations can then be expressed as

$$d\mathbf{x} = dx\,\boldsymbol{\ell}_{\rm R}, \qquad d\mathbf{y} = ds\,\boldsymbol{\ell}; \tag{ia}$$

see Figure 4.7. The stretch λ of the fiber, being the ratio between the deformed and undeformed lengths ds and dx, obeys

$$ds = \lambda dx. \tag{iia}$$

Derive expressions for $\boldsymbol{\ell}_{\rm R}, \boldsymbol{\ell}, \lambda$ and s in terms of $\mathbf{x}(x), \mathbf{y}(x,t)$ and their derivatives.

Determine a tensor[20] $\mathbf{F}$ that takes an undeformed material fiber $d\mathbf{x}$ into its deformed image $d\mathbf{y}$; a tensor $\mathbf{U}$ that stretches $d\mathbf{x}$ by λ without rotating it; a tensor $\mathbf{R}$ that rotates a fiber in the direction $\boldsymbol{\ell}_{\rm R}$ without stretching it; and a tensor $\mathbf{V}$ that stretches a fiber in the direction $\boldsymbol{\ell}$ by λ without rotating it. Express you answers in terms of $\lambda, \boldsymbol{\ell}$ and $\boldsymbol{\ell}_R$.

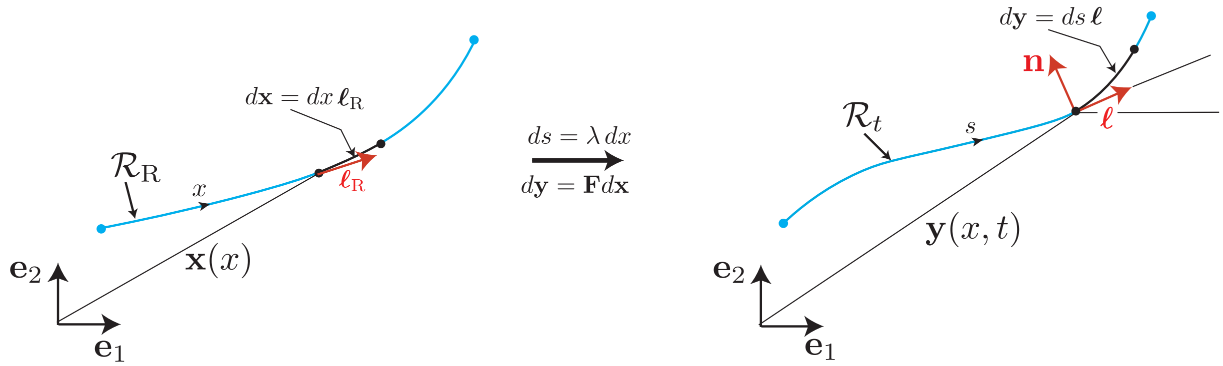


Figure 4.7: An elastic string in a reference configuration (left) and at time t (right). In the reference configuration an infinitesimal material fiber $d\mathbf{x}$ has length dx and direction $\boldsymbol{\ell}_{\rm R}$. Its image $d\mathbf{y}$ in the current configuration has length $ds = \lambda\,dx$ and direction $\boldsymbol{\ell}$.

Part B: Forces. Equilibrium. Now focus attention on the string in the current configuration at some fixed instant t, and identify points along the string by the arc length s. The string is subjected to a distributed (body) force per unit *deformed* length $\mathbf{b}(s,t)$. This induces an internal force field within the string, i.e. if we make a hypothetical cut through the string at some s, the part of the string on one side of the cut applies a force on the part on the other side (and vice versa) due to contact at s. Do *not* assume the contact force to be tangent to the string. Assume that the string in the undeformed and current configurations as well as all forces lie in the same plane. Neglect inertial effects and enforce force and moment equilibrium for an arbitrary part of the string, and thus derive the associated equilibrium equations that the contact force must obey.

Part C: Rate of working. Energy. Constitutive relation. Thus far did not say anything about the constitutive relation of the string.

(c1) Derive the relation between stretch rate $\dot{\lambda}$ and velocity gradient $\mathbf{v}'$ where the dot denotes the derivative with respect to t at fixed x (and the prime denotes the derivative with respect to x at fixed t). Here $\mathbf{v} = \dot{\mathbf{y}}$ is particle velocity.

(c2) Calculate the rate at which the external forces acting on a part of the string do work.

(c3) If the material is dissipation-free, the work done is stored in the string. Let $\mathcal{W}$ be the energy stored per unit deformed length of the string. Write down the elastic power identity, i.e. the equation that

[20] Since we are studying a one-dimensional body in a two-dimensional space, the tensor $\mathbf{F}$ will not be unique since there is no information about "what happens in the second dimension". This is why the question asks you to find a tensor $\mathbf{F}$ rather than the tensor $\mathbf{F}$.

describes the balance between the rate of work and the rate of change of the stored energy for a part of the string. If W denotes the energy stored per unit *reference length*, what is the relation between W and $\mathcal{W}$? Write the elastic power identity in terms of W.

(c4) Finally, if the string is elastic in the sense that there is a constitutive function $\widetilde{W}(\lambda)$ such that $W(x,t) = \widetilde{W}(\lambda(x,t))$, derive the constitutive relation for t. (Here $\mathbf{t} = t\boldsymbol{\ell}$.)

Problem 4.34. *Axisymmetric deformation of an elastic membrane.* (See Problem 4.33 for an analogous problem for an elastic string.) In an unstressed reference configuration an elastic membrane[21] is a circular cylinder of radius R and length L. The Z-axis coincides with the axis of the cylinder and the origin is at one end. The membrane is subjected to a circumferentially uniform but axially varying internal pressure $p(Z)$ per unit deformed area; it acts in a direction perpendicular to the deformed membrane. Formulate the complete theory for the kinematic and force fields in the membrane. Do so by working from first principles, addressing the kinematics, balance laws of equilibrium and constitutive principles. (A different way in which to approach this problem is by taking the limit as the thickness tends to zero of the corresponding problem for a thick-walled tube. Do NOT take this approach here but you may want to try that approach as an exercise.)

---

[21] In the membrane model of an elastic body, its thickness is invisible and the membrane is treated as a two-dimensional object. Moreover, a membrane has no bending stiffness and so the internal forces are assumed to be tangential to the deformed membrane.

References:

1. L. Anand, On H. Hencky's approximate strain-energy function for moderate deformations, *ASME Journal of Applied Mechanics*, 46(1), 1979, pp. 78-82.

2. S. S. Antman, *Nonlinear Problems of Elasticity*, Chapter 2 of Section 13, Springer-Verlag, 1995.

3. M. Baker and J.L. Ericksen, Inequalities restricting the form of the stress-deformation relations for isotropic elastic solids and Reiner-Rivlin fluids, *Journal of the Washington Academy of Sciences*, Volume 44, 1954, pp. 33- 35.

4. J. M. Ball, Discontinuous Equilibrium Solutions and Cavitation in Nonlinear Elasticity, *Philosophical Transactions of the Royal Society of London. Series A.*, Volume 306, 1982, issue 1496, pp. 557-611.

5. J. M. Ball, Some recent developments in nonlinear elasticity and its applications to materials science, in *Nonlinear Mathematics and Its Applications*, edited by P.J. Aston, pp. 93–119. Cambridge University Press, 1996.

6. P. Chadwick, *Continuum Mechanics: Concise Theory and Problems*, Chapter 4, Dover, 1999.

7. J. C. Criscone, Rivlin's Representation Formula is Ill-Conceived for the Determination of Response Functions via Biaxial Testing, in The Rational Spirit in Modern Continuum Mechanics, edited by C-S. Man and R. L. Fosdick, Springer, 2004.

8. A. Goriely, The Mathematics and Mechanics of Biological Growth, Springer, 2017, Chapter 11.

9. A. Goriely, A. Erlich and C. Goodbrake, C5.1 Solid Mechanics: Online problem sheets, https://courses.maths.ox.ac.uk/node/36846/materials, Oxford University, 2018.

10. M.E. Gurtin, E. Fried and L. Anand, *The Mechanics and Thermodynamics of Continua*, Chapters 47-51, Cambridge University Press, 2010.

11. G. A. Holzapfel, *Nonlinear Solid Mechanics. A Continuum Approach for Engineering*, Wiley, 2000.

12. G.A. Holzapfel, T.C. Gasser and Ogden, R. W. A new constitutive framework for arterial wall mechanics and a comparative study of material models, *Journal of Elasticity*, 61(2000), pp. 1-48.

13. J. K. Knowles, The finite anti-plane shear field near the tip of a crack for a class of incompressible elastic solids, *International Journal of Fracture*, Volume 11, 1977, pp. 611-639.

14. J.K. Knowles and E. Sternberg, On the failure of ellipticity and the emergence of discontinuous deformation gradients in plane finite elastostatics, *Journal of Elasticity*, 8 (1978), pp. 329-379.

15. J.E. Marsden and T.J.R. Hughes, *Mathematical Foundations of Elasticity*, Dover 1994 (Prentice-Hall 1983). Especially section 5 of the chapter "A point of departure" preceding chapter 1.

16. R.W. Ogden, Large deformation isotropic elasticity - on the correlation of theory and experiment for incompressible rubberlike materials, *Proceedings of the Royal Society of London*, A326(1972), pp. 565-584.

17. R.W. Ogden, *Non-Linear Elastic Deformations*, Chapter 4, Dover, 1997.

18. R.W. Ogden, Elements of the theory of elasticity, in *Nonlinear Elasticity: Theory and Applications*, Edited by Y.B. Fu and R.W Ogden, London Mathematical Society Lecture Note Series 283, Cambridge University Press, 2001.

19. R. S. Rivlin and D. W. Saunders, Large elastic deformations of isotropic materials. VII Experiments on the deformation of rubber, *Philosophical Transactions of the Royal Society of London. Series A.*, Volume 243, 1951, pp. 34 - 288.

20. D. J. Steigmann, *Finite Elasticity Theory,* Chapters 2-4, Oxford, 2017.

21. R. A. Toupin, Elastic materials with couple stresses, *Archive for Rational Mechanics and Analysis*, Volume 11, 1962, pp. 385-414.

22. L. R. G. Treloar, Stress-strain data for vulcanized rubber under various types of deformation, *Transactions of the Faraday Society*, 40(1944), pp. 59-70.

23. L. R. G. Treloar, The Physics of Rubber Elasticity, Oxford University Press, 1975.

Chapter 5

Some Nonlinear Effects: Illustrative Examples

Nonlinearity can lead to phenomena that are not seen in the linearized theory, phenomena that may even be totally unexpected and counterintuitive. In this chapter we describe some examples of this, but first we summarize the basic field equations and make a few remarks on boundary conditions.

5.1 Summary and boundary conditions.

5.1.1 Field equations.

The strain energy function $W(\mathbf{F})$ characterizing the material is given. The region $\mathcal{R}_{\mathrm{R}}$ occupied by the body in the reference configuration is known and the body force field $\mathbf{b}_{\mathrm{R}}(\mathbf{x})$ is prescribed on $\mathcal{R}_{\mathrm{R}}$. (It might vanish.) The stress field $\mathbf{S}(\mathbf{x})$, deformation gradient tensor field $\mathbf{F}(\mathbf{x})$ and deformation field $\mathbf{y}(\mathbf{x})$ must satisfy the following equations at each point $\mathbf{x} \in \mathcal{R}_{\mathrm{R}}$:

$$\mathrm{Div}\ \mathbf{S} + \mathbf{b}_{\mathrm{R}} = \mathbf{o}, \qquad \mathbf{S} = \frac{\partial W}{\partial \mathbf{F}}, \qquad \mathbf{F} = \mathrm{Grad}\,\mathbf{y}. \tag{5.1}$$

(Question: What about moment equilibrium?)

– Observe that the equilibrium equations

$$\frac{\partial S_{ij}}{\partial x_j} + b_i^R = 0, \tag{5.2}$$

comprise a set of three scalar partial differential equations involving the 9 unknown stress components. Thus the equilibrium equations are *not* sufficient for determining $S_{ij}(\mathbf{x})$ in the body. Stated differently, in general there will be many stress fields that satisfy the equilibrium equations (and traction boundary conditions).

– When the constitutive equations
$$S_{ij} = \frac{\partial W}{\partial F_{ij}} \tag{5.3}$$
are also taken into account, one has an additional set of 9 scalar equations but they involve the 9 components of the deformation gradient tensor.

– Finally, the kinematic equations
$$F_{ij} = \frac{\partial y_i}{\partial x_j} \tag{5.4}$$
provide a set of 9 more scalar equations. They involve the 3 components of the deformation y_i.

– Thus taken together, the system (5.2), (5.3), (5.4) comprises a set of 21 $(= 3 + 9 + 9)$ scalar equations for the 21 $(= 9 + 9 + 3)$ unknown scalar fields $S_{ij}(\mathbf{x}), F_{ij}(\mathbf{x}), y_i(\mathbf{x})$.

– Given the deformation field $\mathbf{y}(\mathbf{x})$, one can calculate the 9 components of the deformation gradient tensor using (5.4). However given a tensor field $\mathbf{F}(\mathbf{x})$ with positive determinant, the equations
$$\frac{\partial y_i}{\partial x_j} = F_{ij} \tag{5.5}$$
for finding $y_i(\mathbf{x})$ comprise a set of 9 equations for the three fields $y_i(\mathbf{x})$. For an arbitrary $\mathbf{F}(\mathbf{x})$, this would be an over-determined set of equations. If a solution $y_i(\mathbf{x})$ is to exist, the nine fields $F_{ij}(\mathbf{x})$ must be suitably restricted. These are known as the **compatibility conditions** and were looked at previously in Problem 2.24.

– For a homogeneous material[1], substituting (5.3) into (5.2) and using (5.4) leads one to
$$\mathbb{A}_{ijk\ell}(\mathbf{F}) \frac{\partial^2 y_k}{\partial x_j \partial x_\ell} + b_i^R = 0, \tag{5.6}$$
where
$$\mathbb{A}_{ijk\ell}(\mathbf{F}) = \frac{\partial^2 W(\mathbf{F})}{\partial F_{ij} \partial F_{k\ell}} \tag{5.7}$$

[1]For an inhomogeneous material the strain energy function would depend on the particle and would have the form $W(\mathbf{F}, \mathbf{x})$. Thus the strain energy would depend on $\mathbf{x}$ both through $\mathbf{F}(\mathbf{x})$ and explicitly.

are the components of the 4-tensor $\mathbb{A}$ that we encountered previously when looking at strong ellipticity.

Equation (5.6) is a set of 3 scalar second-order partial differential equations involving the 3 deformation component fields $y_i(\mathbf{x})$.

5.1.2 Boundary conditions

Let $\partial\mathcal{R}_R$ be the boundary of the region $\mathcal{R}_R$ occupied by the body in a reference configuration. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be complementary parts of $\partial\mathcal{R}_R$ so that $\partial\mathcal{R}_R = \mathcal{S}_1 \cup \mathcal{S}_2$. Then the boundary conditions associated with the **mixed boundary-value problem** of elastostatics are

$$\mathbf{y}(\mathbf{x}) = \widehat{\mathbf{y}}(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathcal{S}_1, \qquad\qquad \mathbf{S}\mathbf{n}_R = \widehat{\mathbf{s}}(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in \mathcal{S}_2, \tag{5.8}$$

where $\widehat{\mathbf{y}}(\mathbf{x})$ and $\widehat{\mathbf{s}}(\mathbf{x})$ are the given deformation and traction on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. Observe that three scalar quantities are prescribed, either y_1, y_2, y_3 or s_1, s_2, s_3 at (almost) every point on the entire boundary.

- In the special case where $\mathcal{S}_2$ is empty, so that the deformation is prescribed on the entire boundary, i.e. $\mathbf{y}(\mathbf{x}) = \widehat{\mathbf{y}}(\mathbf{x})$ for $\mathbf{x} \in \partial\mathcal{R}_R$, one has the deformation boundary-value problem. When the deformation is prescribed on $\partial\mathcal{R}_R$, the displacement is also known on $\partial\mathcal{R}_R$ since $\mathbf{u}(\mathbf{x}) = \widehat{\mathbf{u}}(\mathbf{x}) \coloneqq \widehat{\mathbf{y}}(\mathbf{x}) - \mathbf{x}$ on $\partial\mathcal{R}_R$. Thus this is equivalently a **displacement boundary-value problem.**

- In the complementary case where $\mathcal{S}_1$ is empty, so that the traction is prescribed on the entire boundary, $\mathbf{S}\mathbf{n}_R = \widehat{\mathbf{s}}(\mathbf{x})$ for $\mathbf{x} \in \partial\mathcal{R}_R$, one has a **traction boundary-value problem**. This loading is referred to as **dead loading**. Observe that in dead loading, the prescribed (Piola-Kirchhoff) traction $\widehat{\mathbf{s}}(\mathbf{x})$ does not depend on the deformation itself. Since the body is in equilibrium, it is necessary that $\widehat{\mathbf{s}}(\mathbf{x})$ be such that

  $$\int_{\partial\mathcal{R}_R} \widehat{\mathbf{s}}(\mathbf{x})\, dA_x + \int_{\partial\mathcal{R}} \mathbf{b}_R(\mathbf{x})\, dV_x = 0. \tag{5.9}$$

  Question: what about moment balance?

- **Configuration dependent traction boundary condition.** Let $\mathcal{R}$ be the region occupied by the body in the deformed configuration and let $\partial\mathcal{R}$ be its boundary. Suppose, as an example, that a pressure $-p\mathbf{n}$ is applied on $\partial\mathcal{R}$. Then one has the boundary condition $\mathbf{Tn} = -p\mathbf{n}$ for $\mathbf{y} \in \partial\mathcal{R}$. By using Nanson's formula $\mathbf{n}dA_y = J\mathbf{F}^{-T}\mathbf{n}_R dA_x$ and $\mathbf{Tn}dA_y = \mathbf{Sn}_R dA_x$, this boundary condition can be written as

$$\mathbf{Sn}_R = -pJ\mathbf{F}^{-T}\mathbf{n}_R \qquad \text{for all } \mathbf{x} \in \partial\mathcal{R}_R. \tag{5.10}$$

Observe that for this loading, the (Piola Kirchhoff) traction on $\partial\mathcal{R}_R$ depends on the deformation through $J\mathbf{F}^{-T}$. It is *not* dead loading.

5.2 Example (1): Torsion of a circular cylinder.

References:

1. R. S. Rivlin, Torsion of a Rubber Cylinder, Journal of Applied Physics, 18(1947), pp. 444-449.

2. R.S. Rivlin, Large elastic deformations of isotropic materials. III. Some simple problems in cylindrical polar co-ordinates, *Philosophical Transactions of the Royal Society A*, 240(1948), pp. 509-525.

3. R.S. Rivlin, Large elastic deformations of isotropic materials VI. Further results in the theory of torsion, shear and flexure, *Philosophical Transactions of the Royal Society A*, 242(1949), pp. 173-195.

The region $\mathcal{R}_{\rm R}$ occupied by the body in a reference configuration is a solid circular cylinder of radius A and length L. It is composed of a generalized neo-Hookean material. This cylindrical body is welded onto two rigid plates at its ends. The plate at $x_3 = 0$ is held fixed while that at $x_3 = L$ is rotated through an angle αL about the axis of the cylinder. Here we have in mind rectangular cartesian coordinates with the origin at the center of the fixed end and the x_3-axis being along the centerline of the cylinder. The curved lateral boundary is traction-free.

Let (R, Θ, Z) and (r, θ, z) be the cylindrical polar coordinates of a particle in the reference and deformed configurations respectively:

$$\left.\begin{aligned} x_1 = R\cos\Theta, \qquad x_2 = R\sin\Theta, \qquad x_3 = Z,\\ y_1 = r\cos\theta, \qquad y_2 = r\sin\theta, \qquad y_3 = z. \end{aligned}\right\} \tag{i}$$

The associated basis vectors are $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ respectively. A deformation that takes $(R, \Theta, Z) \to (r, \theta, z)$ can, in general, be characterized by

$$r = \widehat{r}(R, \Theta, Z), \qquad \theta = \widehat{\theta}(R, \Theta, Z), \qquad z = \widehat{z}(R, \Theta, Z). \tag{ii}$$

Here, since the end $Z = 0$ is held fixed, the boundary condition there requires

$$\widehat{r}(R, \Theta, 0) = R, \qquad \widehat{\theta}(R, \Theta, 0) = \Theta, \qquad \widehat{z}(R, \Theta, 0) = 0. \tag{iii}$$

Since the other end $Z = L$ is rigidly rotated through an angle αL about the Z-axis, the boundary condition there necessitates

$$\widehat{r}(R, \Theta, L) = R, \qquad \widehat{\theta}(R, \Theta, L) = \Theta + \alpha L, \qquad \widehat{z}(R, \Theta, L) = L. \tag{iv}$$

The curved lateral boundary of the body is traction-free.

Motivated by the boundary conditions (iii) and (iv), we make the following ansatz[2] that the deformation of the body is given by

$$r = \widehat{r}(R, \Theta, Z) = R, \qquad \theta = \widehat{\theta}(R, \Theta, Z) = \Theta + \alpha Z, \qquad z = \widehat{z}(R, \Theta, Z) = Z, \tag{v}$$

where α is the given angle of twist per unit length. Note that the region $\mathcal{R}$ that the body occupies in the deformed configuration is also a circular cylinder[3] of radius A and length L.

The deformation (v) satisfies the boundary conditions (iii) and (iv) automatically. In this deformation, the cross section at $x_3 = Z$ rotates rigidly through an angle αZ about the x_3-axis. Though each cross section rotates rigidly, different cross sections rotate through different angles and so the body does not undergo a purely rigid rotation. Consider the cross sections at $x_3 = Z$ and $x_3 = Z + \Delta Z$. If we draw an infinitesimal square on the surface of the cylinder touching these two sections, one observes that the square will rotate rigidly through an angle αZ, and that the additional rotation $\alpha \Delta Z$ will shear the square in the circumferential direction $\mathbf{e}_\theta$ (the normal to the shearing plane being $\mathbf{e}_z$).

We want to inquire whether (a) the assumed deformation (v) is possible, i.e. whether the stress field associated with (v) obeys the equilibrium equations and the traction-free

[2]If, based on this assumed form of the deformation, we can satisfy all of the field equations and boundary conditions, then we certainly have a solution of the problem, though we have no assurance that it is unique. This approach to solving problems in solid mechanics is referred to as the "semi-inverse method".

[3]This illustrates why we must not confuse a configuration of a body with the region it occupies. In the present setting, the regions occupied by the body in the undeformed and deformed configurations are identical, though they correspond to different configurations of the body. See Volume II for a careful discussion of this issue.

boundary condition on $r = A$; and if it is possible, (b) to calculate the loading that must be applied on the ends of the cylinder in order to maintain this deformation. Body forces will be ignored.

Substituting the deformation (v) into the formula (2.77) for the deformation gradient tensor in polar coordinates yields

$$\mathbf{F} = \mathbf{e}_r \otimes \mathbf{e}_R + \mathbf{e}_\theta \otimes \mathbf{e}_\Theta + \mathbf{e}_z \otimes \mathbf{e}_Z + \gamma \mathbf{e}_\theta \otimes \mathbf{e}_Z, \qquad (vi)$$

where we have set

$$\gamma := r\alpha. \qquad (vii)$$

The expression (vi) for the deformation gradient tensor can be written in the illuminating form

$$\mathbf{F} = (\mathbf{I} + \gamma \mathbf{e}_\theta \otimes \mathbf{e}_z)(\mathbf{e}_r \otimes \mathbf{e}_R + \mathbf{e}_\theta \otimes \mathbf{e}_\Theta + \mathbf{e}_z \otimes \mathbf{e}_Z). \qquad (viii)$$

Observe that the second factor in $(viii)$ is the rotation tensor that carries the basis $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_Z\}$ into the basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$. The first factor is a simple shear with shearing direction $\mathbf{e}_\theta$, glide plane normal $\mathbf{e}_z$ and amount of shear $\gamma = r\alpha$.

The left Cauchy-Green tensor corresponding to (vi) is

$$\mathbf{B} = \mathbf{F}\mathbf{F}^T = \mathbf{I} + \gamma^2 \mathbf{e}_\theta \otimes \mathbf{e}_\theta + \gamma(\mathbf{e}_\theta \otimes \mathbf{e}_z + \mathbf{e}_z \otimes \mathbf{e}_\theta), \qquad (ix)$$

and it can be readily verified that $\det \mathbf{B} = 1$. Therefore

$$\det\, \mathbf{F} = 1, \qquad (x)$$

and so the torsional deformation (v) is automatically locally volume preserving (isochoric). Incompressibility does not impose any restrictions on it. From (ix) we get

$$I_1(\mathbf{B}) = \operatorname{tr} \mathbf{B} = 3 + \gamma^2, \qquad I_2(\mathbf{B}) = \frac{1}{2}\big[(\operatorname{tr} \mathbf{B})^2 - \operatorname{tr} \mathbf{B}^2\big] = 3 + \gamma^2. \qquad (xi)$$

Exercise: Calculate the components of $\mathbf{B}$ in rectangular cartesian coordinates by first writing (i) and (v) as

$$\left.\begin{aligned} y_1 &= r\cos\theta = R\cos(\Theta + \alpha Z) = R\cos\Theta\cos\alpha Z - R\sin\Theta\sin\alpha Z = x_1\cos\alpha x_3 - x_2\sin\alpha x_3,\\ y_2 &= r\sin\theta = R\sin(\Theta + \alpha Z) = R\sin\Theta\cos\alpha Z + R\cos\Theta\sin\alpha Z = x_2\cos\alpha x_3 + x_1\sin\alpha x_3,\\ y_3 &= x_3. \end{aligned}\right\}$$

Then calculate the components of $\mathbf{B}$ in the cylindrical polar basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ by using the basis change formula for 2-tensors.

The body is composed of a generalized neo-Hookean material characterized by its strain energy function

$$W = W(I_1), \qquad W(3) = 0, \quad W'(3) > 0, \tag{xii}$$

the corresponding constitutive relation for the Cauchy stress being

$$\mathbf{T} = -q\mathbf{I} + 2W'(I_1)\mathbf{B}. \tag{$xiii$}$$

In view of (vi), the deformation is characterized by the single kinematic parameter $\gamma = \alpha r$ and so it is natural to express the strain energy function $W(I_1)$ in terms of γ. Thus let $w(\gamma)$ be the restriction of $W(I_1)$ to a torsional deformation:

$$w(\gamma) := W(3 + \gamma^2). \tag{xiv}$$

Substituting (ix) and (xiv) into $(xiii)$ gives the Cauchy stress tensor to be

$$\mathbf{T} = (2W'(I_1) - q)\mathbf{I} + \gamma w'(\gamma)\mathbf{e}_\theta \otimes \mathbf{e}_\theta + w'(\gamma)(\mathbf{e}_\theta \otimes \mathbf{e}_z + \mathbf{e}_z \otimes \mathbf{e}_\theta), \tag{xv}$$

and so the cylindrical polar components of the Cauchy stress are[4]

$$\begin{aligned} &T_{rr} = -q + 2W'(3+\gamma^2), \qquad T_{\theta\theta} = -q + 2W'(3+\gamma^2) + \gamma w'(\gamma), \\ &T_{zz} = -q + 2W'(3+\gamma^2), \qquad T_{r\theta} = T_{rz} = 0, \qquad T_{\theta z} = w'(\gamma). \end{aligned} \tag{xvi}$$

Observe from (xvi) that the shear stress component $T_{z\theta}$ is known, and that once q has been determined, the normal stress components will also be known.

Since the torque about the z-axis is due to the shear stress $T_{z\theta}$, and this stress component is known, we can calculate the component m_z of the moment on the end $x_3 = L$ of the cylinder to be

$$m_z = \int_0^A r\, T_{\theta z}\, 2\pi r\, dr = 2\pi \int_0^A r^2 w'(\gamma)\, dr \overset{(vii)}{=} \frac{2\pi}{\alpha^3} \int_0^{A\alpha} \gamma^2 w'(\gamma)\, d\gamma. \qquad \square$$

Given the twist angle α and the material $w(\gamma)$, this equation gives the value of m_z. This is conditional of course on being able to satisfy all the field equations and boundary conditions, which requires that the stress field (xvi) satisfy the equilibrium equations and the following traction-free boundary conditions on the curved lateral boundary:

$$T_{rr} = T_{r\theta} = T_{rz} = 0 \qquad \text{at } r = A. \tag{$xvii$}$$

[4]Observe that these expressions can be simplified by absorbing the term $2W'(3+\gamma^2)$ into q.

Note from (xvi) that the second and third boundary conditions hold trivially and so only the first has to be enforced.

It would be natural to assume that q is independent of θ and z and depends only on r. However it is easy to show as follows that this is necessarily so (without having to make this assumption): the equilibrium equations in cylindrical polar coordinates were given in (3.94). Substituting (xvi) into $(3.94)_2$ and $(3.94)_3$ yields

$$\frac{\partial q}{\partial \theta} = 0, \qquad \frac{\partial q}{\partial z} = 0. \tag{$xviii$}$$

Thus q does not depend on θ and z and so is a function of r alone: $q = q(r)$. All stress components are now functions of r only. Finally, the radial equilibrium equation $(3.94)_1$ specializes to the ordinary differential equation

$$\frac{dT_{rr}}{dr} + \frac{T_{rr} - T_{\theta\theta}}{r} = 0, \qquad 0 \le r < A. \tag{xix}$$

A straightforward approach in which to proceed would be to substitute (xvi) into (xix) to obtain a differential equation for $q(r)$. After this has been solved for $q(r)$, one can then calculate T_{rr} from (xvi) and finally enforce the boundary condition. However, since we are not particularly interested in $q(r)$ we shall proceed in a way that avoids having to find it. Since the boundary conditions are given on T_{rr} it is natural to use (xix) to solve for T_{rr} directly if possible. To this end, observe from (xvi) that $T_{\theta\theta} - T_{rr}$ does not involve q:

$$T_{\theta\theta} - T_{rr} = \gamma w'(\gamma). \tag{xx}$$

Substituting (xx) into (xix) gives

$$\frac{dT_{rr}}{dr} = \alpha w'(\gamma), \qquad 0 \le r < A. \tag{xxi}$$

We now integrate this from an arbitrary radius r to the outer radius $r = A$ and use the boundary condition $T_{rr}(A) = 0$:

$$\cancel{T_{rr}(A)} - T_{rr}(r) = \int_r^A \alpha w'(\gamma) dr \overset{(vii)}{=} \int_{r\alpha}^{A\alpha} w'(\gamma) d\gamma = w(A\alpha) - w(r\alpha). \tag{$xxii$}$$

Thus the equilibrium equations and boundary conditions have been satisfied and the radial stress field is given by

$$T_{rr}(r) = w(r\alpha) - w(A\alpha), \qquad 0 \le r \le A. \tag{$xxiii$}$$

Therefore we conclude that the assumed deformation (v) does indeed satisfy all of the requirements of the problem.

Observe from $(xvi)_1$ and $(xxiii)$ that

$$q(r) = 3W'(3+\gamma^2) - w(r\alpha) + w(A\alpha), \qquad 0 \le r \le A,$$

and note, as mentioned previously in Chapter 4.5, that $q(r)$ is a field, not a constant (in general).

Exercise: Suppose the shaft was hollow with the inner and outer boundaries being traction-free. Can you satisfy *both* boundary conditions $T_{rr}(A) = 0$ and $T_{rr}(B) = 0$ (where B is the inner radius)? If not, how would you proceed?

Next we calculate the loading on the ends (having already calculated the torque above).

Observe from (xvi) that due to the non-zero stress T_{zz} there is a normal traction on the the ends of the cylinder. Thus we now calculate the resultant force on a cross section. This is simply the integral of T_{zz} over the cross section. From (xvi) we see that $T_{zz} = T_{rr}$ and so the axial stress is

$$T_{zz} = w(r\alpha) - w(A\alpha), \qquad 0 \le r \le A. \tag{xxiv}$$

The axial force to be applied on a cross-section can now be calculated:

$$f_z = \int_0^A T_{zz}\, 2\pi r dr \overset{(xxi),(iii)}{=} \frac{2\pi}{\alpha^2}\int_0^{A\alpha} \gamma\big[w(\gamma) - w(A\alpha)\big]\, d\gamma. \qquad \square \tag{xxv}$$

Exercise: Use symmetry arguments to infer that the components f_x and f_y of the resultant force and the components m_x and m_y of the resultant torque vanish.

5.2.1 Discussion.

– When the preceding results are specialized to the neo-Hookean material $W(I_1) = \frac{\mu}{2}(I_1 - 3)$, one has $w(\gamma) = \frac{\mu}{2}\gamma^2$ and so

$$m_z \overset{(\square)}{=} \frac{2\pi\mu}{\alpha^3}\int_0^{A\alpha} \gamma^3\, d\gamma = \frac{\pi}{2}\mu\alpha A^4, \tag{xxvii}$$

$$f_z \overset{(xxv)}{=} \frac{\pi\mu}{\alpha^2}\int_0^{A\alpha} [\gamma^3 - A^2\alpha^2\gamma]\, d\gamma = -\frac{\pi}{4}\mu A^4\alpha^2. \tag{xxviii}$$

– Our analysis in this section was limited to a generalized neo-Hookean material. For a general isotropic incompressible material characterized by the strain energy function $W(I_1, I_2)$, one can show that

$$f_z = -2\pi\alpha^2 \int_0^A r^3 \left(\frac{\partial W}{\partial I_1} + 2\frac{\partial W}{\partial I_2} \right) \Big|_{I_1=I_2=3+\alpha^2 r^2} dr, \qquad (xxix)$$

$$m_z = 4\pi\alpha \int_0^A r^3 \left(\frac{\partial W}{\partial I_1} + \frac{\partial W}{\partial I_2} \right) \Big|_{I_1=I_2=3+\alpha^2 r^2} dr. \qquad (xxx)$$

Again, since the torsional deformation involves a single parameter γ, it is convenient to express the strain energy function $W(I_1, I_2)$ in terms of it. Thus, let $w(\gamma)$ be the restriction of $W(I_1, I_2)$ to a torsional deformation:

$$w(\gamma) := W(3+\gamma^2, 3+\gamma^2), \qquad (xxxi)$$

where we have used (xi). One can show that equation (xxx) can be written in terms of w as

$$m_z = 2\pi \int_0^A r^2 w'(\alpha r)\ dr = \frac{2\pi}{\alpha^3} \int_0^{A\alpha} \gamma^2 w'(\gamma)\ d\gamma, \qquad (xxxii)$$

which coincides with what we got for the generalized neo-Hookean material except that here, w is defined by $(xxxi)$.

– If a normal force was not applied on the two end plates, then the plates would displace in the x_3-direction and so the length of the body would change. In this case one would consider a deformation

$$r = \widehat{r}(R,\Theta,Z) = R, \qquad \theta = \widehat{\theta}(R,\Theta,Z) = \Theta + \alpha\lambda Z, \qquad z = \widehat{z}(R,\Theta,Z) = \lambda Z, \qquad (xxxiii)$$

where the stretch λ is to be determined (from the zero resultant axial force condition).

Exercise: In Problem 5.3 you are asked to carry out the calculations underlying the preceding remark and determine λ.

– Recall that according to the classical linearized theory of elasticity, in order to subject a circular cylindrical body to a torsional deformation one need only apply a torque about the shaft axis; an axial force is not required. It is not surprising that the finite deformation theory says that one must also apply an axial force. Recall from Section 4.6.1 that in order to maintain a finite simple shear deformation one must apply both shear and normal stresses. Locally, at each point of the shaft, a torsional deformation is just a simple shear together with a rigid rotation. The need for an axial force here is simply a manifestation of the need for normal stresses in simple shear.

– To see another consequence of the presence of normal stresses in simple shear, consider a large thin sheet that contains a small planar crack in its interior. Far from the crack the sheet is subjected to a simple shear deformation in the plane of the sheet with the direction of shearing being parallel to the crack – a so-called Mode II loading. If the crack was not present, the sheet would undergo a simple shear deformation and in particular, there would be a normal stress acting on the plane where we intend there to be a crack. If we now introduce a crack with traction-free faces, it follows that in addition to sliding, the crack faces will either move apart and so the crack will open up, or the crack faces will press together and be in contact. Which of these occurs depends on whether the normal stress in the direction perpendicular to the crack faces is tensile or compressive in the absence of the crack. In contrast, in the linearized theory, the crack faces simply slide parallel to each other.

– Finally we remark that normal stresses are also present in the shear flow of non-Newtonian fluids. If such a fluid is placed between two vertical coaxial circular cylindrical tubes with a closed horizontal base, and one of the tubes is rotated about its axis, the fluid will climb up along the tubes. This is because, in order to maintain a shear flow, a suitable normal stress must be applied, and such a stress was not applied at the free surface of the fluid in the aforementioned experiment. Thus the fluid moves in the vertical direction.

Exercises: Problems 5.2, 5.3, 5.4, 5.5, 5.13, 5.14 and 5.15.

5.3 Example (2): Deformation of an Incompressible Cube Under Prescribed Tensile Forces.

References:

1. J. L. Ericksen, Introduction to the Thermodynamics of Solids, Chapman Hall, 1991, Chapter 6.

2. E. A. Kearsley, Asymmetric Stretching of a Symmetrically Loaded Elastic Sheet, *International Journal of Solids and Structures*, 22(1986), issue 2, pp. 111-119.

3. R.S. Rivlin, Large elastic deformations of isotropic materials. II Some uniqueness theorems for pure homogeneous deformation, *Philosophical Transactions of Royal Society*

(London), Series A, **240**(1948), pp. 491-508.

4. R.S. Rivlin, Stability of pure homogeneous deformations of an elastic cube under dead loading, *Quarterly of Applied Mathematics*, **32**(1974), pp. 265-271

5. L.R.G. Treloar, Stresses and birefringence in rubber subjected to general homogeneous strain, *Proceedings of the Physical Society*, 60(1948), pp. 135-144.

6. L.R.G. Treloar, *The Physics of Rubber Elasticity*, Clarendon Press, Oxford, 1949, pp. 114–120.

Equilibrium configurations of a cube: Consider an incompressible isotropic elastic body that occupies a unit cube in a reference configuration. The cube is composed of a neo-Hookean material[5] characterized by the strain-energy function

$$W = \frac{\mu}{2}(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3), \qquad \mu > 0. \tag{i}$$

The principal Cauchy stresses are related to the principal stretches by the constitutive relation

$$\tau_i = \lambda_i \frac{\partial W}{\partial \lambda_i} - q = \mu\lambda_i^2 - q, \quad i = 1, 2, 3, \qquad \text{(no sum on } i\text{)}, \tag{ii}$$

where q arises due to the incompressibility constraint.

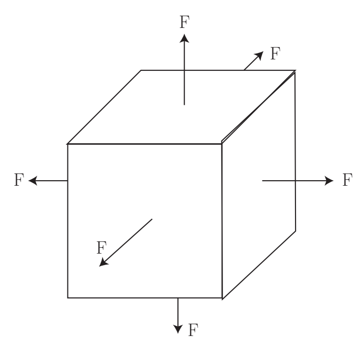


Figure 5.1: A unit cube in the reference configuration. All six of its faces are subjected to uniformly distributed normal tractions whose resultant force, on each face, is F. The figure only shows the resultant force and not the distributed traction.

Each of the six faces of the cube is subjected to a uniformly distributed normal traction whose resultant is a tensile force $F(> 0)$. This is illustrated in Figure 5.1 where the uniform

[5] In Problem 5.9 you will generalize this to an arbitrary isotropic material.

distribution of normal traction is not shown, only the resultant forces are. We wish to determine the resulting pure homogeneous deformation of the body. This problem is frequently referred to as the "Rivlin cube problem".

It should be noted that in the problem we are considering it is the force F, or equivalently the first Piola-Kirchhoff traction (why?), that is prescribed and so we have dead loading on the entire boundary of the body. The associated Cauchy (true) tractions on the faces of the cube will depend on the areas of the faces in the deformed configuration. One could consider the problem in which the Cauchy tractions are prescribed on each face. This is a different problem to the one we study here.

Because of the symmetry of the body, the loading and the material, one may be inclined to assume that the deformation will also be symmetric. However, we wish to look at the possibility of not-necessarily symmetric pure homogeneous deformations, and so we shall not assume a priori that the cube deforms symmetrically. If it does, then we will find this to be the case. Thus suppose that the cube undergoes a pure homogeneous deformation

$$y_1 = \lambda_1 x_1, \qquad y_2 = \lambda_2 x_2, \qquad y_3 = \lambda_3 x_3. \tag{iii}$$

Incompressibility of the material requires

$$\lambda_1\lambda_2\lambda_3 = 1. \tag{iv}$$

The deformed faces of the body have areas $\lambda_2\lambda_3$, $\lambda_3\lambda_1$ and $\lambda_1\lambda_2$ and so the prescribed boundary conditions tell us that the Cauchy stress components are

$$T_{11} = \frac{F}{\lambda_2\lambda_3} \overset{(iv)}{=} F\lambda_1, \qquad T_{22} = \frac{F}{\lambda_3\lambda_1} \overset{(iv)}{=} F\lambda_2, \qquad T_{33} = \frac{F}{\lambda_1\lambda_2} \overset{(iv)}{=} F\lambda_3. \tag{v}$$

The problem at hand is to find the principal stretches λ_i, given F (and μ). Note that the first Piola-Kirchhoff stress tensor (field throughout the body) is

$$\mathbf{S}(\mathbf{x}) = F\mathbf{e}_1 \otimes \mathbf{e}_1 + F\mathbf{e}_2 \otimes \mathbf{e}_2 + F\mathbf{e}_3 \otimes \mathbf{e}_3, \tag{vi}$$

and observe that it does not depend on the deformation.

Since the deformation is homogeneous, and assuming the reaction pressure field q to be constant, the stress field will also be homogeneous throughout the body. Therefore (ignoring body forces) the equilibrium equations are satisfied automatically. The boundary conditions have already been accounted for in (v) above. All that remains is to enforce the constitutive law (ii):

$$T_{11} = \mu\lambda_1^2 - q, \qquad T_{22} = \mu\lambda_2^2 - q, \qquad T_{33} = \mu\lambda_3^2 - q. \tag{vii}$$

Combining (vii) with (v) and using (iv) leads to

$$F\lambda_1 = \mu\lambda_1^2 - q, \qquad F\lambda_2 = \mu\lambda_2^2 - q, \qquad F\lambda_3 = \mu\lambda_3^2 - q. \tag{viii}$$

Equations $(viii)$ and (iv) provide four (nonlinear) algebraic equations involving $\lambda_1, \lambda_2, \lambda_3$ and q.

In order to solve these equations systematically it is convenient to first eliminate q. Thus, subtracting the second of $(viii)$ from the first, and similarly the third from the second leads to

$$\left.\begin{aligned} [F - \mu(\lambda_1 + \lambda_2)](\lambda_1 - \lambda_2) &= 0, \\ [F - \mu(\lambda_2 + \lambda_3)](\lambda_2 - \lambda_3) &= 0. \end{aligned}\right\} \tag{ix}$$

Equations (iv) and (ix) are to be solved for the principal stretches $\lambda_1, \lambda_2, \lambda_3$. There are three cases to consider:

<u>Case (1)</u>: Suppose first that all of the λ's are distinct: $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1$. Then (ix) yields

$$F = \mu(\lambda_1 + \lambda_2), \qquad F = \mu(\lambda_2 + \lambda_3),$$

which implies that $\lambda_1 + \lambda_2 = \lambda_2 + \lambda_3$ whence

$$\lambda_1 = \lambda_3.$$

This contradicts the assumption that the $\lambda's$ are all distinct. Thus there is no solution in which the three λ's are distinct. (If such a solution had existed, it would have described an *orthorhombic configuration* of the body.)

<u>Case (2)</u>: Suppose next that all of the λ's are equal: $\lambda_1 = \lambda_2 = \lambda_3$. This describes a *cubic configuration.* In this case equations (ix) are automatically satisfied and (iv) tells us that

$$\lambda_1 = \lambda_2 = \lambda_3 = 1. \tag{x}$$

Thus one solution of the problem, for every value of the applied force F, is given by (iii), (x). This corresponds to a configuration of the body in which, geometrically, it remains a unit cube, but one that is under stress.

<u>Case (3)</u>: Finally, consider the remaining possibility that two λ's are equal and different to the third. This describes a *tetragonal configuration.* Suppose that

$$\lambda_1 = \lambda_2 = \lambda \text{ (say)}, \qquad \lambda_3 \neq \lambda. \tag{xi}$$

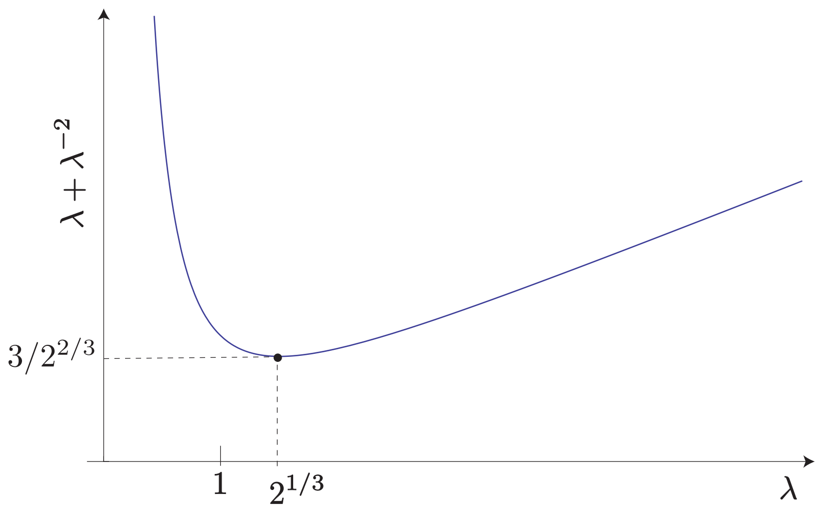


Figure 5.2: Graph of $\lambda + \lambda^{-2}$ versus λ.

Incompressibility (iv) together with (xi) requires

$$\lambda_3 = \lambda^{-2}, \quad \lambda \neq 1, \tag{xii}$$

while the pair of equations (ix) reduce to $F = \mu(\lambda_2 + \lambda_3) = \mu(\lambda + \lambda^{-2})$, i.e.

$$\lambda + \lambda^{-2} = f \qquad \text{where we have set} \quad f := F/\mu. \tag{xiii}$$

Given f, if $(xiii)$ can be solved for one or more real roots $\lambda > 0$, then $(iii), (xi)$, $(xiii)$ provides the corresponding solution to the problem[6]. Whether $(xiii)$ can be solved or not depends on the value of f.

In order to examine the solvability of $(xiii)$ let $h(\lambda) := \lambda + \lambda^{-2}$ for $\lambda > 0$, and observe that $h \to \infty$ as $\lambda \to 0^+$ and $h \to \infty$ as $\lambda \to \infty$. Moreover, $h'(\lambda) = 1 - 2\lambda^{-3}$ and so $h'(\lambda) < 0$ for $0 < \lambda < 2^{1/3}$, $h'(\lambda) = 0$ for $\lambda = 2^{1/3}$, and $h'(\lambda) > 0$ for $\lambda > 2^{1/3}$. Thus the graph of $h(\lambda)$ versus λ is as shown in Figure 5.2. The minimum value of h is $h(2^{1/3}) = 3/2^{2/3}$.

From Figure 5.2 and $(xiii)$ we see that

if $f < 3/2^{2/3}$, equation $(xiii)$ has no roots,

if $f = 3/2^{2/3}$, equation $(xiii)$ has one root $\lambda = 2^{1/3}$, and

if $f > 3/2^{2/3}$, equation $(xiii)$ has two roots.

[6] There are of course additional configurations corresponding to permutations of the λ's, e.g. $\lambda_3 = \lambda_1 = \lambda$, $\lambda_2 = \lambda^{-2}$.

For a solution with $\lambda > 1$ one has $\lambda_2 = \lambda_3 > 1, \lambda_1 < 1$ and so the deformed body has two relatively long equal edges and one relatively short unequal edge, i.e. the block has a flattened shape as depicted by the upper inset in Figure 5.3. On the other hand $\lambda < 1$ describes configurations in which $\lambda_2 = \lambda_3 < 1, \lambda_1 > 1$ where the deformed body has two relatively short equal edges and one relatively long unequal edge, i.e. the block has a pillar-like shape as in the lower inset.

Thus in *summary*, there are two types of configurations which the body can adopt. In one, the body remains a unit cube in the deformed configuration and this is possible for all values of the applied force F. The other is possible only if $F/\mu \geq 3/2^{2/3}$ and here the deformed body is no longer a cube. Rather, it has a tetragonal shape where two sides are equal and the third is different, and there are two possibilities of this form corresponding to the two roots of $(xiii)$.

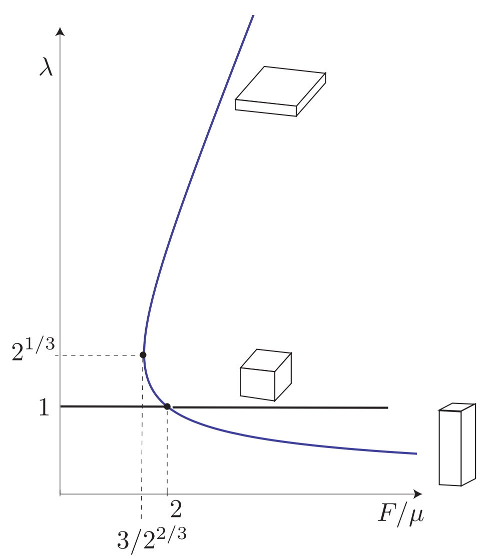


Figure 5.3: Equilibrium configurations of the cube: the symmetric (cubic) configuration corresponds to the line $\lambda = 1$. The curve corresponds to the asymmetric (tetragonal) configurations given by $(xiii)$.

Both types of solutions are depicted in Figure 5.3. The cubic solution corresponds to the horizontal line $\lambda = 1$ which extends indefinitely to the right. The tetragonal solution corresponds to the curve (which is the same curve as in Figure 5.2 but with the axes switched). The figure shows that if $f < 3/2^{2/3}$ the body must be in the cubic configuration,

if $f > 3/2^{2/3}$ the body can be in either a cubic or tetragonal configuration there being two configurations of the latter type.

Thus the solution to the equilibrium problem is *non-unique*.

Stability of the cube: The lack of uniqueness prompts us to examine the stability of the various equilibrium configurations.

Remarks on the stability of an equilibrium configuration: An alternative approach for studying equilibrium configurations of an elastic solid/structure is via the minimization of the potential energy. One considers *all geometrically possible deformation fields* $\mathbf{z}(\mathbf{x})$, and minimizes the potential energy Φ over this class of functions. If the potential energy has an extremum at say $\mathbf{z}(\mathbf{x}) = \mathbf{y}(\mathbf{x})$ then $\mathbf{y}(\mathbf{x})$ describes an equilibrium configuration of the body. If this extremum corresponds to a minimum of the potential energy, then this configuration is stable.

Suppose that an elastic body occupies a region $\mathcal{R}_{\rm R}$ in a reference configuration and that the deformation is prescribed on a portion $\mathcal{S}_1$ of its boundary and the first Piola-Kirchhoff traction ("dead load") $\widehat{\mathbf{s}}(\mathbf{x})$ is prescribed on the remaining portion $\mathcal{S}_2$ of the boundary where $\partial\mathcal{R}_{\rm R} = \mathcal{S}_1 \cup \mathcal{S}_2$. A kinematically possible deformation field ("a virtual deformation field") is any smooth enough vector field $\mathbf{z}(\mathbf{x})$ defined on $\mathcal{R}_{\rm R}$ that obeys all geometric constraints. One geometric requirement is that $\mathbf{z}(\mathbf{x})$ coincide with the prescribed deformation on $\mathcal{S}_1$. If there are internal kinematic constraints such as incompressibility, then these too must be enforced. The potential energy associated with a geometrically possible deformation field $\mathbf{z}(\mathbf{x})$ is

$$\Phi = \int_{\mathcal{R}_{\rm R}} W(\nabla\mathbf{z})\; dV_x \;-\; \int_{\mathcal{S}_2} \widehat{\mathbf{s}}\cdot\mathbf{z}\; dA_x; \qquad (xiv)$$

Appendix 5.3.1 on page 341 explains how this expression arises. The first term on the right hand side describes the elastic energy stored in the body while the second term corresponds to the potential energy of the applied dead loading[7]. One seeks to minimize this functional over the set of all geometrically possible deformation fields $\mathbf{z}(\mathbf{x})$.

In the case of the tri-axially loaded cube, dead-loading is prescribed on the entire bound-

[7]If a body force field is present, one would take its potential energy into account by including the volume integral of $-\mathbf{b}_{\rm R}\cdot\mathbf{z}$ over the body on the right-hand side of (xiv).

ary and so one can replace $\mathcal{S}_2$ by $\partial\mathcal{R}_{\rm R}$ in (xiv) and write

$$\Phi = \int_{\mathcal{R}_{\rm R}} W(\nabla\mathbf{z})\ dV_x \ - \int_{\partial\mathcal{R}_R} \widehat{\mathbf{s}}\cdot\mathbf{z}\ dA_x. \qquad (xv)$$

This is to be minimized over the set of all smooth enough vector field $\mathbf{z}(\mathbf{x})$ subject to the incompressibility requirement $\det\nabla\mathbf{z} = 1$.

Rather than minimizing this over the set of all geometrically possible kinematic fields suppose that we minimize over the smaller class of all geometrically possible *homogeneous* deformation fields: $\mathbf{z}(\mathbf{x}) = \mathbf{F}\mathbf{x}$ where $\mathbf{F}$ is an arbitrary constant tensor with unit determinant. (In Problem 5.8 you will consider all virtual deformations.) Then the potential energy specializes to

$$\Phi = \int_{\mathcal{R}_{\rm R}} W(\mathbf{F})\ dV_x \ - \int_{\partial\mathcal{R}_{\rm R}} \mathbf{S}\mathbf{n}_{\rm R}\cdot\mathbf{F}\mathbf{x}\ dA_x = \int_{\mathcal{R}_{\rm R}} \Big[W(\mathbf{F})\ -\ \mathbf{S}\cdot\mathbf{F}\Big]\ dV_x,$$

where we have used the divergence theorem in getting to the second equality. Since $\mathbf{F}$ and $\mathbf{S}$ are constants, this leads to

$$\Phi = W(\mathbf{F}) - \mathbf{S}\cdot\mathbf{F}, \qquad (xvi)$$

which is to be minimized over all tensors $\mathbf{F}$ with $\det\mathbf{F} = 1$.

Finally suppose that we further limit attention to geometrically possible deformation fields of the even more restricted form

$$z_1 = \lambda_1 x_1, \quad z_2 = \lambda_2 x_2, \quad z_3 = \lambda_3 x_3, \qquad \lambda_1\lambda_2\lambda_3 = 1; \qquad (xvii)$$

here $\lambda_1, \lambda_2, \lambda_3$ *are arbitrary* subject only to $(xvii)_4$ and $\lambda_k > 0$. The potential energy (xvi) (for the neo-Hookean material with $\mathbf{S}$ given by (vi)) now takes the explicit form

$$\frac{\mu}{2}(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) - F(\lambda_1 + \lambda_2 + \lambda_3). \qquad (xviii)$$

We are to minimize this over all $(\lambda_1, \lambda_2, \lambda_3)$ subject to the constraint $\lambda_1\lambda_2\lambda_3 = 1$. We can simplify this by eliminating λ_3 using the incompressibility equation. After dropping an inessential constant, dividing by μ and letting $f = F/\mu$ as before, we can write the function to be minimized as

$$\Phi(\lambda_1, \lambda_2) = \frac{1}{2}(\lambda_1^2 + \lambda_2^2 + \lambda_1^{-2}\lambda_2^{-2}) - f\,(\lambda_1 + \lambda_2 + \lambda_1^{-1}\lambda_2^{-1}). \qquad (xix)$$

This function is to be minimized over all $\lambda_1 > 0, \lambda_2 > 0$.

For convenience let

$$\Phi_\alpha = \frac{\partial \Phi}{\partial \lambda_\alpha}, \qquad \Phi_{\alpha\beta} = \frac{\partial^2 \Phi}{\partial \lambda_\alpha \partial \lambda_\beta}, \qquad [P] = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12} & \Phi_{22} \end{pmatrix}, \qquad D = \det [P]. \qquad (xx)$$

The equilibrium configurations correspond to the extrema of the potential energy function $\Phi(\lambda_1, \lambda_2)$. They are found by setting

$$\Phi_1 = \Phi_2 = 0.$$

An equilibrium configuration is locally stable if the corresponding extremum is a local minimum of $\Phi(\lambda_1, \lambda_2)$. To study the character of an extremum one evaluates the Hessian matrix $[P]$ at that extremum. The extremum is a local minimum if both eigenvalues of $[P]$ are positive; it is a local maximum if both eigenvalues of $[P]$ are negative; and it is a saddle if one eigenvalue of $[P]$ is positive and the other is negative. Thus

$$\begin{array}{lll} \text{An extremum is a local minimum if} & D > 0, & \Phi_{11} > 0, \\ \text{An extremum is a local maximum if} & D > 0, & \Phi_{11} < 0, \\ \text{An extremum is a saddle if} & D < 0. & \end{array}$$

Differentiating (xix) yields

$$\left.\begin{aligned} \Phi_1 &= \lambda_1 - \lambda_1^{-3}\lambda_2^{-2} - f(1 - \lambda_2^{-1}) = [\lambda_1 + \lambda_1^{-1}\lambda_2^{-1} - f](\lambda_2 - \lambda_1^{-2})\lambda_2^{-1}, \\ \Phi_2 &= \lambda_2 - \lambda_2^{-3}\lambda_1^{-2} - f(1 - \lambda_2^{-2}\lambda_1^{-1}) = [\lambda_2 + \lambda_2^{-1}\lambda_1^{-1} - f](\lambda_1 - \lambda_2^{-2})\lambda_1^{-1}. \end{aligned}\right\} \qquad (xxi)$$

Setting $\Phi_1 = \Phi_2 = 0$ leads to a pair of algebraic equations. There are four cases to consider since each equation involves two factors, each of which could vanish. This tells us that the roots of $\Phi_1 = \Phi_2 = 0$ (i.e. the equilibrium configurations) are:

$$\begin{array}{ll} (a) & \lambda_1 = \lambda_2 = 1, \ (\lambda_3 = 1), \\ (b) & \lambda_1 + \lambda_1^{-2} = f, \quad \lambda_1 = \lambda_2, \ (\lambda_3 = \lambda_1^{-2}), \\ (c) & \lambda_1 + \lambda_1^{-2} = f, \quad \lambda_2 = \lambda_1^{-2}, \ (\lambda_3 = \lambda_1), \\ (d) & \lambda_2 + \lambda_2^{-2} = f, \quad \lambda_1 = \lambda_2^{-2}, \ (\lambda_3 = \lambda_2). \end{array}$$

By using the incompressibility condition $\lambda_1\lambda_2\lambda_3 = 1$, we see that solution (a) corresponds to the cubic configuration $\lambda_1 = \lambda_2 = \lambda_3 = 1$; solution (b) corresponds to the tetragonal configuration $\lambda_1 = \lambda_2 \neq \lambda_3$; solution (c) corresponds to the tetragonal configuration $\lambda_1 = \lambda_3 \neq \lambda_2$; and solution (d) corresponds to the tetragonal configuration $\lambda_2 = \lambda_3 \neq \lambda_1$. Thus the equilibrium configurations corresponding to solutions (a) and (b) are the same ones we found earlier, while solutions (c) and (d) are permutations of solution (b).

Since (c) and (d) are simply permutations of (b), we will shortly ignore them (as we did before). However, in order to better understand these extrema we shall plot energy contours on the λ_1, λ_2-plane, and when we do this we will be forced to confront all of the extrema. Thus for a short while longer we will continue to consider all extrema $(a) - (d)$. The λ_1, λ_2-plane with these seven extrema marked is shown in Figure 5.4.

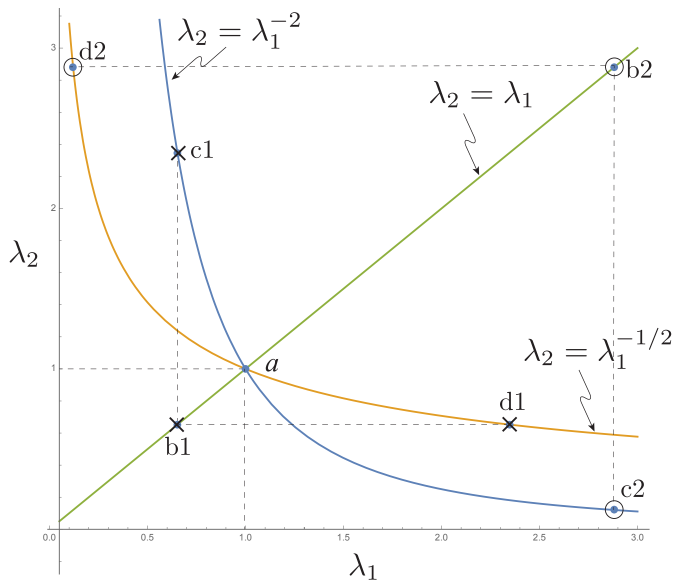


Figure 5.4: Extrema of the potential energy function $\Phi(\lambda_1, \lambda_2)$ on the λ_1, λ_2-plane. The point a corresponds to solution (a); points $b1, b2$ to solution (b); points $c1, c2$ to solution (c); and points $d1, d2$ to solution (d). The figure has been drawn for $f = 3$. Figures 5.5 and 5.6 tell us which of these are local minima, (local maxima and saddle points).

To examine the stability of these equilibrium configurations we next calculate the second derivatives of Φ:

$$\left.\begin{aligned}
\Phi_{11} &= 1 + 3\lambda_1^{-4}\lambda_2^{-2} - 2f\ \lambda_1^{-3}\lambda_2^{-1}, \\
\Phi_{22} &= 1 + 3\lambda_2^{-4}\lambda_1^{-2} - 2f\ \lambda_2^{-3}\lambda_1^{-1}, \\
\Phi_{12} &= 2\lambda_1^{-3}\lambda_2^{-3} - f\lambda_1^{-2}\lambda_2^{-2}.
\end{aligned}\right\} \qquad (xxii)$$

Considering the cubic configuration, we set $\lambda_1 = \lambda_2 = \lambda_3 = 1$ in $(xxii)$ and $(xx)_4$ which

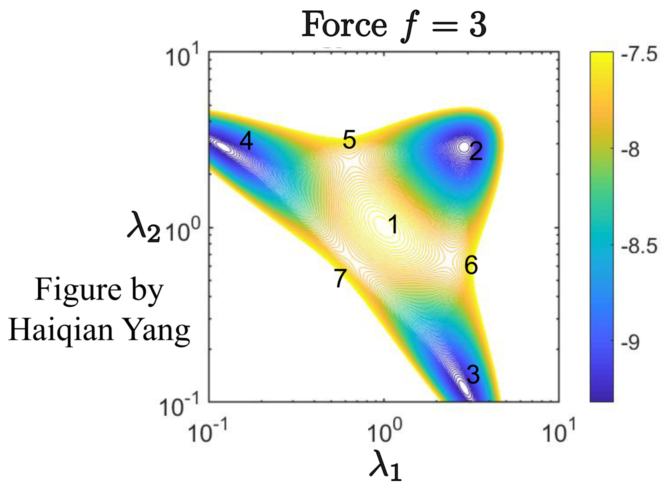


Figure 5.5: Energy contours of potential energy $\Phi(\lambda_1, \lambda_2)$ on the λ_1, λ_2-plane for $f = 3$. The energy has a local maximum at point 1 (corresponding to point a in Figure 5.4). It has saddle points at 7, 5 and 6 (corresponding to $b1, c1, d1$ in Figure 5.4). The energy has local minima at 2, 3 and 4 (corresponding to $b2, c2, d2$ in Figure 5.4). Figure provided by Haiqian Yang (student in 2.074 in 2020).

gives

$$\Phi_{11} = 2(2 - f), \qquad \Phi_{22} = 2(2 - f), \qquad \Phi_{12} = 2 - f, \qquad D = 3(2 - f)^2.$$

Therefore by the statement above (xxi), this extremum is a local minimum and *the cubic configuration is stable* if $\Phi_{11} > 0$:

$$f < 2.$$

(It is a local maximum for $f > 2$.) At a tetragonal solution (say (b) above) we have

$$\lambda_3 = \lambda^{-2}, \qquad \lambda_1 = \lambda_2 = \lambda \neq 1, \qquad f = \lambda + \lambda^{-2}, \qquad (xxiii)$$

and to examine its stability we evaluate $(xxii)$ and $(xx)_4$ at $(xxiii)$:

$$\Phi_{11} = (1 - \lambda^{-3})^2 > 0, \qquad \Phi_{22} = (1 - \lambda^{-3})^2, \qquad \Phi_{12} = -(1 - \lambda^{-3})\lambda^{-3},$$

$$D = \lambda^{-3}(1 - \lambda^{-3})^2 \,(\lambda^3 - 2)\,.$$

By statement above (xxi), this extremum is a minimum and *the tetragonal configuration is stable* if $D > 0$:

$$\lambda^3 > 2 \qquad \Rightarrow \qquad \lambda > 2^{1/3}. \qquad (xxiv)$$

(It is a saddle point when it is not a minimum.)

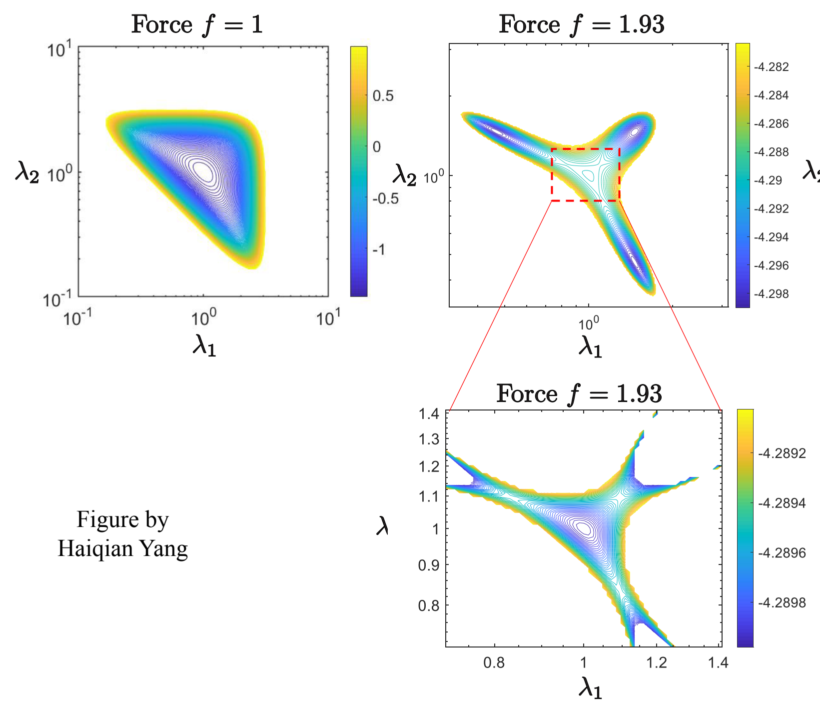


Figure 5.6: Energy contours of the potential energy at three values of force. For $f = 1$ it has a single energy well at the cubic configuration. The force $f = 1.93$ corresponds to the intermediate range in Figure 5.3. The energy has four energy wells corresponding to the cubic and three flattened configurations. When f exceeds 2, the cubic configuration is no longer a local minimum. Figure provided by Haiqian Yang (student in 2.074 in 2020).

Figure 5.5 shows the energy contours of the potential energy $\Phi(\lambda_1, \lambda_2)$. The figure has been drawn for $f = 3$. At this value of force, the energy has a local maximum at point 1 (corresponding to point a in Figure 5.4); saddle points at 7, 5 and 6 (corresponding to $b1, c1, d1$ in Figure 5.4); and energy wells (local minima) at 2, 3 and 4 (corresponding to $b2, c2, d2$ in Figure 5.4). The saddle points $b1, c1$ and $d1$ correspond to unstable pillar like configurations, while the local minima $b2, c2$ and $d2$ correspond to flattened stable configurations.

Figure 5.6 shows the energy contours at several values of f. For $f = 1$ it has a single energy well and it occurs at the cubic configuration. Since $3/2^{2/3} < 1.93 < 2$, the force $f = 1.93$ corresponds to the intermediate range in Figure 5.3, and the energy in Figure 5.6 has four energy wells corresponding to the cubic and three flattened configurations. When f exceeds 2, the cubic configuration is no longer a local minimum.

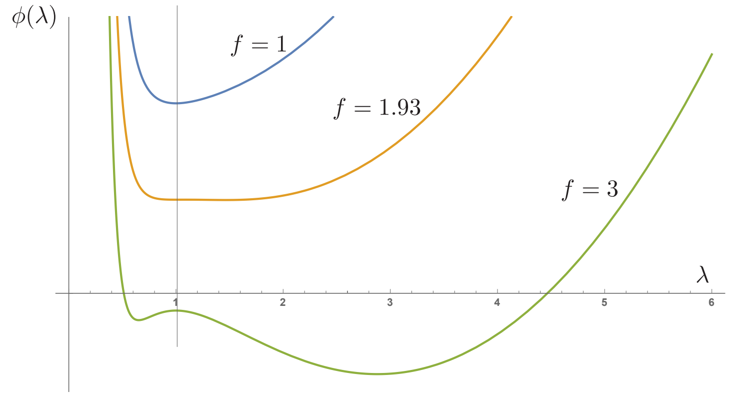


Figure 5.7: Graphs of the potential energy $\phi(\lambda) = \Phi(\lambda, \lambda)$ versus λ for three fixed values of the force f, one > 2, the second in the interval $(3/2^{2/3}, 2)$ and the third $< 3/2^{2/3}$. The leftmost local minimum of the curve corresponding to $f = 3$ is in fact a saddle point of the energy $\Phi(\lambda_1, \lambda_2)$ on the two-dimensional λ_1, λ_2-plane.

Figure 5.7 shows a plot of the potential energy

$$\phi(\lambda) := \Phi(\lambda_1, \lambda_2)\Big|_{\lambda_1=\lambda, \lambda_2=\lambda} = \frac{1}{2}(2\lambda^2 + \lambda^{-4}) - f(2\lambda + \lambda^{-2})$$

versus the stretch λ for three different values of force. This is the slice of the graph of $\Phi(\lambda_1, \lambda_2)$ on the λ_1, λ_2-plane along the straight line $\lambda_2 = \lambda_1$. In keeping with Figure 5.6, observe that the extremum at $\lambda = 1$ is an energy well (local minimum) for $f < 2$ while it is a local maximum for $f > 2$. For $3/2^{1/3} < f < 2$ the energy has two energy wells, one at $\lambda = 1$ and the other at a value of $\lambda > 2^{1/3}$, and a local maximum between them. For $f > 2$ the graph in Figure 5.7 shows two energy wells with a local maximum at $\lambda = 1$. We know from the preceding analysis that what appears to be an energy-well at the smaller stretch here is in fact a saddle point when viewed on the λ_1, λ_2-plane. It merely appears to be a minimum along the particular slice of the energy shown in Figure 5.7.

Finally we return to considering solutions (a) and (b) only and ignore the various permutations. Figure 5.8 depicts the solutions again, now with the solid line/curve corresponding to the stable solutions and the dashed line/curve the unstable ones. For $f < 3/2^{1/3}$ we have a unique stable cubic configuration. For $f > 2$ we have a unique (to within permutations) stable tetragonal configuration; see also Figure 5.4. On the intermediate range $3/2^{1/3} < f < 2$ there is one stable cubic configuration and one stable tetragonal configuration. (Strictly, there are 3 asymmetric solutions since permutations of the λ_i's should be considered.) Observe that λ is > 1 for the stable tetragonal configurations and these, as noted previously, correspond to configurations where the deformed body is flattened.

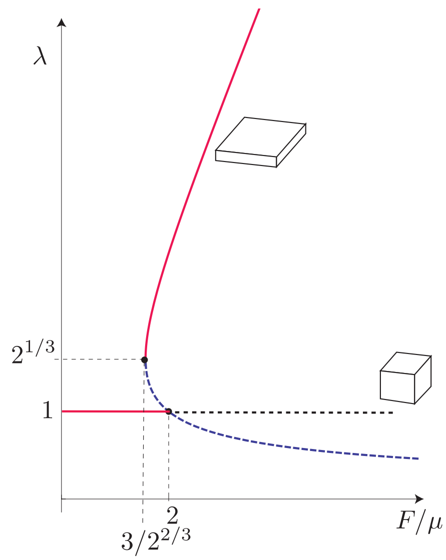


Figure 5.8: The stable and unstable solutions are depicted by the solid and dashed curves respectively.

Remark: We were supposed to minimize the potential energy (xvi) over all $\mathbf{F}$ with $\det \mathbf{F} = 1$. However we only considered admissible deformations of the form $(xvii)$ where $\mathbf{F}$ is coaxial with $\mathbf{S}$, i.e. $\mathbf{F}$ and $\mathbf{S}$ are diagonal in the same basis. When the more restricted analysis we carried out claims an equilibrium state to be a local minimizer, it may or may not be a minimizer under the wider class of all deformations. However if the restricted analysis shows an equilibrium state to be not a local minimizer, then it is not a minimizer even under the wider class of all deformations. Therefore equilibrium states that we found to be unstable are indeed unstable. Those we found to be stable may not be stable in the context of a wider class of admissible deformations.

Remark: The analogous plane stress problem was analyzed by Kearsley in the context of the experiments of Treloar. In this case the stress-free body is a thin sheet and it is subjected to equal in-plane biaxial forces F, the other two faces of the sheet being traction-free. See references given at the beginning of this section (including the monograph by Ericksen). The neo-Hookean material model does not exhibit asymmetric configurations in this case but a Mooney-Rivlin material does; see Problem 5.7.

Exercises: Problems 5.6 and 5.7.

5.3.1 Appendix: Potential energy of an elastic body subjected to conservative loading:

A brief video on potential energy can be found here.

Consider a motion $\mathbf{y}(\mathbf{x}, t)$. By taking the scalar product of the equation of motion $\mathrm{Div}\,\mathbf{S} + \mathbf{b}_R = \rho_R \dot{\mathbf{v}}$ with the particle velocity $\mathbf{v}$, and integrating over $\mathcal{R}_R$, one can readily show for an elastic body that

$$\int_{\partial\mathcal{R}_R} \mathbf{S}\mathbf{n}_R \cdot \mathbf{v}\, dA_x + \int_{\mathcal{R}_R} \mathbf{b}_R \cdot \mathbf{v}\, dV_x = \frac{d}{dt}\int_{\mathcal{R}_R} \left(\frac{1}{2}\rho_R \mathbf{v}\cdot\mathbf{v} + W\right) dV_x, \qquad (i)$$

where ρ_R is the mass density in the reference configuration. This is a statement of the fact that the rate at which work is done on the body by the traction on $\partial\mathcal{R}_R$ and the body force on $\mathcal{R}_R$ equals the rate at which the kinetic energy and the potential energy due to deformation (the strain energy) increase. When the loading is conservative, the rate at which the loading does work can also be expressed as the rate of increase of an associated potential energy.

Suppose that the deformation is prescribed on a part $\mathcal{S}_1$ of the boundary to be $\widehat{\mathbf{y}}(\mathbf{x})$, and the Piola-Kirchhoff traction is prescribed on the complementary part $\mathcal{S}_2$ to be $\widehat{\mathbf{s}}(\mathbf{x})$. Thus for all t,

$$\mathbf{y}(\mathbf{x}, t) = \widehat{\mathbf{y}}(\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathcal{S}_1, \qquad \mathbf{S}(\mathbf{x}, t)\mathbf{n}_R = \widehat{\mathbf{s}}(\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathcal{S}_2. \qquad (ii)$$

A body force $\mathbf{b}_R(\mathbf{x})$ is applied on $\mathcal{R}_R$. Observe that the prescribed deformation $\widehat{\mathbf{y}}$, the (dead load) traction $\widehat{\mathbf{s}}$ and the body force $\mathbf{b}_R$ have all been assumed to be time-independent[8]. By $(ii)_1$ the particle velocity vanishes on $\mathcal{S}_1$ and so we can write (i) as

$$\int_{\mathcal{S}_2} \widehat{\mathbf{s}} \cdot \mathbf{v}\, dA_x + \int_{\mathcal{R}_R} \mathbf{b}_R \cdot \mathbf{v}\, dV_x = \frac{d}{dt}\int_{\mathcal{R}_R} \left(W + \frac{1}{2}\rho_R \mathbf{v}\cdot\mathbf{v}\right) dV_x,$$

having also used $(ii)_2$. Since $\widehat{\mathbf{s}}$ and $\mathbf{b}_R$ (as well as $\mathcal{S}_2$ and $\mathcal{R}_R$) are time independent, and $\mathbf{v} = \partial\mathbf{y}/\partial t$, this can be written as

$$\frac{d}{dt}\int_{\mathcal{S}_2} \widehat{\mathbf{s}} \cdot \mathbf{y}\, dA_x + \frac{d}{dt}\int_{\mathcal{R}_R} \mathbf{b}_R \cdot \mathbf{y}\, dV_x = \frac{d}{dt}\int_{\mathcal{R}_R} \left(W + \frac{1}{2}\rho_R \mathbf{v}\cdot\mathbf{v}\right) dV_x,$$

which yields

$$\frac{d}{dt}\left[\int_{\mathcal{R}_R} \frac{1}{2}\rho_R \mathbf{v}\cdot\mathbf{v}\, dV_x + \int_{\mathcal{R}_R} W\, dV_x - \int_{\mathcal{S}_2} \widehat{\mathbf{s}} \cdot \mathbf{y}\, dA_x - \int_{\mathcal{R}_R} \mathbf{b}_R \cdot \mathbf{y}\, dV_x\right] = 0.$$

[8]This does not imply that the body is in equilibrium. It could, for example, be vibrating while the loading remains constant.

This states that the sum of the kinetic and potential energies is conserved, the first term being the kinetic energy and the next three terms the potential energy Φ:

$$\Phi := \int_{\mathcal{R}_R} W\, dV_x - \int_{\mathcal{S}_2} \widehat{\mathbf{s}} \cdot \mathbf{y}\, dA_x - \int_{\mathcal{R}_R} \mathbf{b}_R \cdot \mathbf{y}\, dV_x. \qquad (iii)$$

The first term in (iii) is the elastic potential energy of the body due to deformation and the next two terms represent the potential energy of the loading.

5.4 Example (3): Growth of a Cavity.

References:

1. A.N. Gent and P.B. Lindley, Internal rupture of bonded rubber cylinders in tension, *Proceedings of the Royal Society (London)*, **A249**, (1958), 195-205.
2. J.M. Ball, Discontinuous equilibrium solutions and cavitation in nonlinear elasticity, *Philosophical Transactions of the Royal Society (London)*, **A306**, (1982), 557-611.
3. H. Wang and S. Cai, Drying-induced cavitation in a constrained hydrogel, *Soft Matter*, 11(2015), pp. 1058-1061.

A body occupies a hollow spherical region of inner radius A and outer radius B in a reference configuration. It is composed of a generalized neo-Hookean material. A uniformly distributed radial tensile dead load (1st Piola Kirchhoff traction) of magnitude σ is applied on the outer surface of the body while the inner surface remains traction-free. We wish to determine the deformation and stress fields in the body. *Our particular interest is in the radius a of the deformed cavity as a function of the applied stress σ.*

Let (R, Θ, Φ) and (r, θ, ϕ) be the spherical polar coordinates of a particle in the reference and deformed configurations respectively with associated basis vectors $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi\}$ and $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$. The geometric, material and loading symmetries, suggest that we assume the deformation to have the spherically symmetric form

$$r = \widehat{r}(R, \Theta, \Phi) = r(R), \qquad \theta = \widehat{\theta}(R, \Theta, \Phi) = \Theta, \qquad \phi = \widehat{\phi}(R, \Theta, \Phi) = \Phi. \qquad (i)$$

Thus the displacement of a particle is in the radial direction and its magnitude depends only on the radial coordinate. Moreover, $\{\mathbf{e}_R, \mathbf{e}_\Theta, \mathbf{e}_\Phi\} = \{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$. In Chapter 2.7.2 we derived

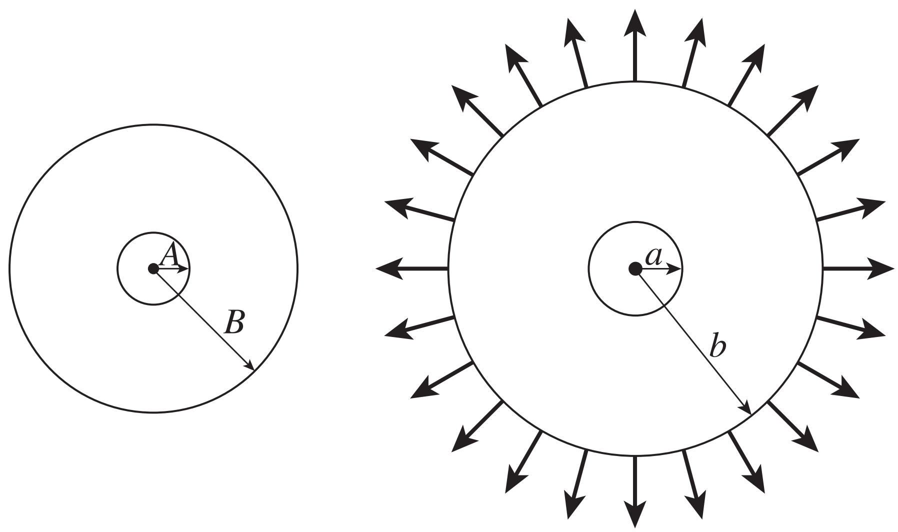


Figure 5.9: A hollow sphere in a reference configuration (left) and in the deformed configuration (right). A uniform radial dead load is applied on the outer surface of the sphere.

formulae for the components of the left Cauchy-Green deformation tensor $\mathbf{B}$ in spherical polar coordinates. Substituting (i) into (2.87) yields

$$\mathbf{B} = \big[r'(R)\big]^2 \mathbf{e}_r \otimes \mathbf{e}_r + \left[\frac{r(R)}{R}\right]^2 (\mathbf{e}_\theta \otimes \mathbf{e}_\theta + \mathbf{e}_\phi \otimes \mathbf{e}_\phi). \qquad (ii)$$

Since $\mathbf{B}$ is diagonal in the basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ we can read off the principal stretches to be

$$\lambda_r = r'(R), \qquad \lambda_\theta = \lambda_\phi = r(R)/R,$$

having assumed $r'(R) > 0$. Since the material is incompressible, $\lambda_r\lambda_\theta\lambda_\phi = \lambda_r\lambda_\theta^2 = 1$ and so it is convenient to work in terms of the circumferential stretch which we denote by λ. Then

$$\lambda_\theta = \lambda_\phi = \lambda := \frac{r}{R}, \qquad \lambda_r = \lambda^{-2} = r'. \qquad (iii)$$

Incompressibility requires

$$\lambda_r\lambda_\theta\lambda_\phi = r'\,\frac{r^2}{R^2} = 1. \qquad (iv)$$

Integrating (iv) gives $r^3 = R^3 + \mathsf{constant}$. Let a denote the as-yet-unknown radius of the cavity in the deformed configuration. Then, since $r = a$ when $R = A$ we can write this as

$$r(R) = \big[R^3 + a^3 - A^3\big]^{1/3}. \qquad (v)$$

The value of the constant a is to be determined. We could have written (v) directly by equating the volume between the spherical surfaces of radii A and R in the reference configuration to the volume between the surfaces of radii a and r in the deformed configuration: $\frac{4}{3}\pi(R^3 - A^3) = \frac{4}{3}\pi(r^3 - a^3)$. The outer radius of the deformed body, $b = r(B)$, is

$$b = \left[B^3 + a^3 - A^3\right]^{1/3}. \qquad (vi)$$

The deformation $r = r(R)$ is completely determined by (i) and (v) once we determine a.

Turning to the constitutive relation we first note from (iii) that

$$I_1 = \operatorname{tr}\mathbf{B} = \lambda_r^2 + \lambda_\theta^2 + \lambda_\phi^2 = 2\lambda^2 + \lambda^{-4}. \qquad (vii)$$

From (ii), (iii) and the constitutive relation (4.65) for a generalized neo-Hookean material, we find the Cauchy stress components to be

$$T_{rr} = -q + 2\lambda^{-4}W'(I_1), \qquad T_{\theta\theta} = T_{\phi\phi} = -q + 2\lambda^2 W'(I_1), \qquad T_{r\theta} = T_{\theta z} = T_{zr} = 0, \quad (viii)$$

where q arises from the incompressibility constraint. Note that the stress components are fully determined by $(viii)$ when $q(r, \theta, \phi)$ and a are known.

We assume that the stress components are functions of the radial coordinate r alone, which, by $(viii)$, implies that q depends only on r: $q = q(r)$. The equilibrium equations (3.97) in spherical polar coordinates now reduce to the single equation

$$\frac{dT_{rr}}{dr} + \frac{2}{r}(T_{rr} - T_{\theta\theta}) = 0. \qquad (ix)$$

The direct way in which to proceed is to substitute $(viii)$ into (ix) to obtain a differential equation for $q(r)$, and after it has been solved, to then calculate the stresses from $(viii)$. However, since we are not particularly interested in q, and the boundary conditions are given on the radial stress, it is more natural to work with T_{rr}.

Observe from $(viii)$ that $T_{rr} - T_{\theta\theta}$ does not depend on q. In fact,

$$T_{rr} - T_{\theta\theta} = 2(\lambda^{-4} - \lambda^2)W'(I_1), \qquad (x)$$

and so we can write (ix) as

$$\frac{dT_{rr}}{dr} = \frac{4}{r}(\lambda^2 - \lambda^{-4})W'(I_1). \qquad (xi)$$

While we can integrate both sides with respect to r, it turns out to be preferable to integrate the right-hand side with respect to λ instead. This requires us to convert the left-hand side to $dT_{rr}/d\lambda$ which we achieve as follows:

$$\frac{dT_{rr}}{dr} = \frac{dT_{rr}}{d\lambda}\frac{d\lambda}{dr} = \frac{dT_{rr}}{d\lambda}\frac{d\lambda}{dR}\frac{dR}{dr} \overset{(iii)}{=} \frac{dT_{rr}}{d\lambda}\frac{d(r/R)}{dR}\frac{1}{r'} = \frac{dT_{rr}}{d\lambda}\frac{r' - r/R}{Rr'} \overset{(iii)}{=} \frac{\lambda^{-2}-\lambda}{R\lambda^{-2}}\frac{dT_{rr}}{d\lambda}. \tag{xii}$$

Therefore from the two preceding equation we obtain

$$\frac{dT_{rr}}{d\lambda} = \frac{4(\lambda-\lambda^{-5})}{1-\lambda^3}W'(I_1) \qquad \text{where } I_1 = 2\lambda^2 + \lambda^{-4}. \tag{xiii}$$

It is convenient to express the strain energy as a function of the stretch λ, i.e. to define a function $w(\lambda)$ by

$$w(\lambda) \coloneqq W(I_1)\Big|_{I_1 = 2\lambda^2+\lambda^{-4}}. \tag{xiv}$$

Differentiating (xiv) with respect to λ gives

$$w'(\lambda) = 4(\lambda - \lambda^{-5})W'(I_1), \tag{xv}$$

and so we can finally write $(xiii)$ as

$$\frac{dT_{rr}}{d\lambda} = \frac{w'(\lambda)}{1-\lambda^3}. \tag{xvi}$$

On integrating (xvi) from the inner boundary to the outer boundary we get

$$T_{rr}(b) - T_{rr}(a) = \int_{\lambda_a}^{\lambda_b} \frac{w'(\lambda)}{1-\lambda^3}\,d\lambda = \int_{\lambda_b}^{\lambda_a} \frac{w'(\lambda)}{\lambda^3 - 1}\,d\lambda, \tag{xvii}$$

where λ_a and λ_b are the circumferential stretches at the corresponding boundaries:

$$\lambda_a \coloneqq \frac{a}{A}, \qquad \lambda_b \coloneqq \frac{b}{B} \overset{(vi)}{=} \left[1 - \frac{A^3}{B^3} + \frac{a^3}{B^3}\right]^{1/3}. \tag{xviii}$$

As for the boundary conditions, we are told the inner boundary is traction-free:

$$T_{rr} = 0 \qquad \text{at } r = a. \tag{xix}$$

We are also told that there is a radial 1st Piola Kirchhoff stress $S_{rR}(B) = \sigma$ at the outer boundary. The corresponding radial Cauchy stress $T_{rr}(b)$ can be calculated using the general tensor relation between $\mathbf{S}$ and $\mathbf{T}$. Alternatively, since the area of the outer surface in the reference configuration is $4\pi B^2$ while the corresponding area in the deformed configuration is $4\pi b^2$ we must have $4\pi B^2 S_{rR}(B) = 4\pi b^2 T_{rr}(b)$ and so the boundary condition at the outer surface is

$$T_{rr} = \sigma\frac{B^2}{b^2} \overset{(xviii)}{=} \frac{\sigma}{\lambda_b^2} \qquad \text{at } r = b. \tag{xx}$$

On substituting (xix) and (xx) into $(xvii)$,

$$\sigma = \lambda_b^2 \int_{\lambda_b}^{\lambda_a} \frac{w'(\lambda)}{\lambda^3 - 1}\, d\lambda. \tag{xxi}$$

By $(xviii)$, λ_a and λ_b are functions of the unknown deformed cavity radius a which is therefore the only unknown in equation (xxi). Thus given σ, equation (xxi) constitutes an algebraic equation for a.

To illustrate the behavior predicted by (xxi) consider a neo-Hookean material. In this case

$$W = \frac{\mu}{2}\big(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3\big) \qquad \Rightarrow \qquad w(\lambda) = \frac{\mu}{2}(\lambda^{-4} + 2\lambda^2 - 3).$$

Substituting this into (xxi) and evaluating the integral leads to

$$\frac{\sigma}{2\mu} = \left[\frac{1}{\lambda_b} + \frac{1}{4}\frac{1}{\lambda_b^4} - \frac{1}{\lambda_a} - \frac{1}{4}\frac{1}{\lambda_a^4}\right]\lambda_b^2. \tag{$xxii$}$$

This equation together with $(xviii)$ tells us how the cavity radius a in the deformed configuration depends on the stress σ.

Equation $(xxii)$ with $(xviii)$ is of the form $\sigma/2\mu = h(a)$. One can show that $h(a)$ increases monotonically with a, and moreover that $h(A) = 0$ and $h(a) \to \infty$ as $a \to \infty$. Thus, for each given value of the stress $\sigma > 0$, the equation $\sigma/2\mu = h(a)$ can be solved for a unique root a. The graph in Figure 5.10 shows the variation of the cavity radius a/B with the stress $\sigma/2\mu$ according to $(xxii)$; the figure has been drawn for a cavity of initial radius $A/B = 0.3$.

Thus far, this problem in the nonlinear theory, has not been qualitatively different to the corresponding problem in linear elasticity. However, if we plot graphs of a versus σ for progressively decreasing initial cavity radii A, the family of curves obtained shows an interesting trend as seen in Figure 5.11. As $A/B \to 0$, the graph of a/B versus $\sigma/2\mu$ approaches the curve $\mathcal{C}$. Note that $\mathcal{C}$ is composed of two segments: the straight line segment $a = 0$ for $0 < \sigma/2\mu \leq 1.25$ and the curved portion for $\sigma/2\mu \geq 1.25$. The curve $\mathcal{C}$ describes the growth of a cavity whose radius is infinitesimal in the undeformed configuration. According to it, for $\sigma/2\mu \leq 1.25$ the cavity radius in the deformed configuration is also zero. However when $\sigma/2\mu > 1.25$, the cavity has a positive radius and has opened. Thus "cavitation" occurs at the critical stress $\sigma/2\mu = 1.25$.

To examine this analytically, we first determine the curved portion of $\mathcal{C}$ by taking the limit of the right-hand side of (xxi) as $A/B \to 0$ at fixed $a/B > 0$. Observe from $(xviii)$

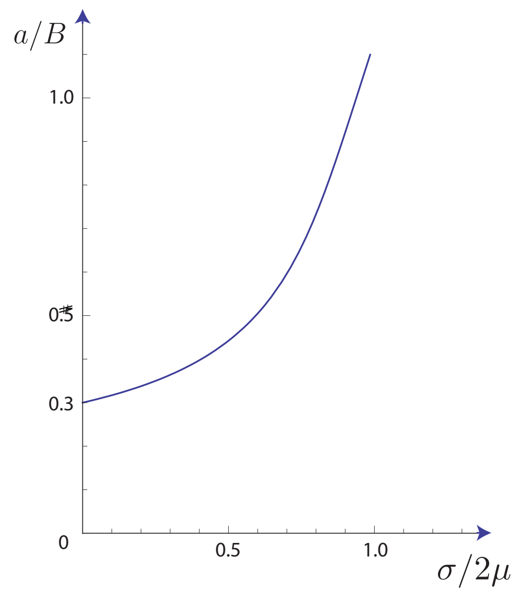


Figure 5.10: Variation of the deformed cavity radius a with applied stress σ for a neo-Hookean material. The figure has been drawn for the case $A/B = 0.3$.

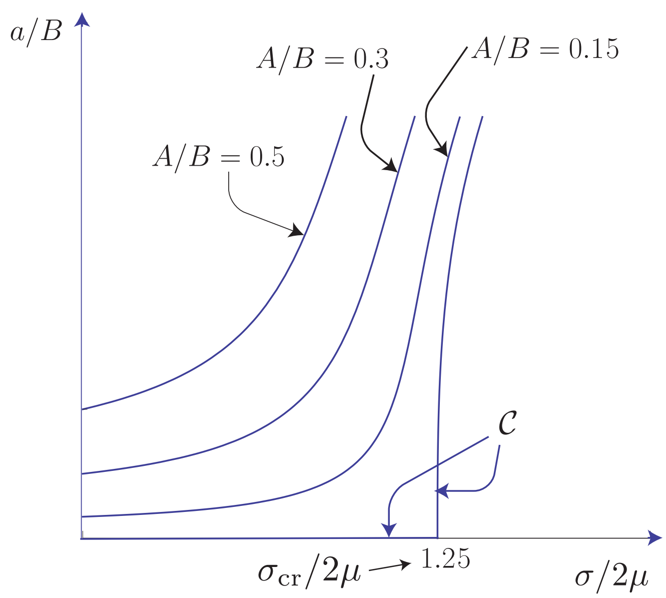


Figure 5.11: Variation of the deformed cavity radius a with applied stress σ for a neo-Hookean material. The different curves correspond to different values of the undeformed cavity radius A. Observe that as $A/B \to 0$ these curves approach the curve $\mathcal{C}$.

that in this limit

$$\lambda_a = \frac{a}{A} = \frac{a}{B}\,\frac{B}{A} \to \infty, \qquad \lambda_b \to \left[1 + \frac{a^3}{B^3}\right]^{1/3}. \tag{xxiii}$$

Substituting these limiting values into (xxi) gives

$$\frac{\sigma}{2\mu} = \lambda_b + \frac{1}{4}\frac{1}{\lambda_b^2} = \frac{5/4 + a^3/B^3}{(1 + a^3/B^3)^{2/3}},$$

which is the equation of the curved portion of $\mathcal{C}$. To find $\sigma_{\rm cr}$, i.e. the point at which $\mathcal{C}$ departs from the horizontal axis, we let $a/B \to 0$ in this equation which yields

$$\frac{\sigma_{\rm cr}}{2\mu} = 5/4. \tag{xxiv}$$

This is indicated in Figure 5.11.

Remark: Observe from $(xxiii)$ that $\lambda_b \to 1$ in the limit $a/B \to 0$ and therefore from (xx) that

$$T_{cr} = \sigma_{cr}.$$

In summary, we have shown for a neo-Hookean material that a cavity that is infinitesimally small in the undeformed configuration remains infinitesimally small as the stress σ increases until it reaches the critical value σ_{cr}. When σ exceeds σ_{cr}, the cavity opens and grows (i.e. $a > 0$) in the manner described by the curved portion of $\mathcal{C}$. This describes the phenomenon of *cavitation*.

We now return to the generalized neo-Hookean material (from the neo-Hookean material) and concern ourselves with (xxi). On taking the limit $A/B \to 0$ at fixed $a/B > 0$ and keeping $(xxiii)$ in mind, (xxi) yields

$$\sigma = \lambda_b^2 \int_{\lambda_b}^{\infty} \frac{w'(\lambda)}{\lambda^3 - 1}\, d\lambda \qquad \text{where} \quad \lambda_b = \left[1 + \frac{a^3}{B^3}\right]^{1/3}. \tag{xxv}$$

This relates the deformed radius a of a cavity that was infinitesimal in the reference configuration to the stress σ. To find the critical stress σ_{cr} for cavitation we let $a/B \to 0$ in (xxv) to obtain the formal expression

$$\sigma_{cr} = \int_1^{\infty} \frac{w'(\lambda)}{\lambda^3 - 1} d\lambda. \tag{xxvi}$$

As noted in the exercise below, this expression for the cavitation stress continues to hold for an arbitrary isotropic incompressible material (with $w(\lambda)$ defined by $(xxix)$ below). Observe that the integrand in $(xxvi)$ has a potential singularity at $\lambda = 1$ unless $w'(1)$ behaves suitably. Moreover, since the range of this integral is infinite, its convergence depends on the behavior of $w(\lambda)$ as $\lambda \to \infty$. If $w(\lambda) \sim \lambda^m$ for large λ, the integral will converge if $m < 3$ and not otherwise. This is essentially Ball's condition (4.120). Thus for example for the one-term Ogden material (4.138), cavitation will not occur, i.e. the critical stress at cavitation will be infinite, if the constitutive parameter $n \geq 3$. In summary, for certain elastic materials, i.e. certain functions W, the integral in $(xxvi)$ will not converge and so an infinitesimally small void will remain infinitesimally small for all values of applied stress. For other materials (this integral will converge and) an infinitesimally small cavity will begin to grow when σ exceeds the critical value given by $(xxvi)$.

Remark: Ball [2] studied this problem using energy minimization which therefore addresses the issue of stability as well. Moreover, he studied this as a bifurcation problem for an (initially) solid sphere.

Exercise: *Arbitrary incompressible isotropic material:* The preceding analysis (for a generalized neo-Hookean material) can be carried over quite easily to an arbitrary incompressible isotropic material. Rather than working with $W(I_1, I_2)$ work with the form $W(\lambda_1, \lambda_2, \lambda_3)$ and use the constitutive relation (4.65) to show that

$$T_{rr} = \lambda_1 \frac{\partial W}{\partial \lambda_1} - q = \lambda^{-2} \frac{\partial W}{\partial \lambda_1}\Big|_{\lambda_1=\lambda^{-2}, \lambda_2=\lambda_3=\lambda} - q, \qquad (xxvii)$$

$$T_{\theta\theta} = T_{\phi\phi} = \lambda_2 \frac{\partial W}{\partial \lambda_2} - q = \lambda \frac{\partial W}{\partial \lambda_2}\Big|_{\lambda_1=\lambda^{-2}, \lambda_2=\lambda_3=\lambda} - q. \qquad (xxviii)$$

Let $w(\lambda)$ be the restriction of the strain energy function to deformations of the sort at hand:

$$w(\lambda) := W(\lambda^{-2}, \lambda, \lambda). \qquad (xxix)$$

Show that

$$T_{rr} - T_{\theta\theta} = -\frac{1}{2}\lambda w'(\lambda), \qquad (xxx)$$

and that equation $(xxvi)$ continues to hold except that w is now defined by $(xxix)$.

Exercises: Problems 5.10, 5.11, 5.13, 5.16.

5.5 Example (4): Limit point instability of a thin-walled hollow sphere.

The limit point instability of a thin-walled hollow sphere (or cylinder) refers to the loss of monotonicity of the function $p(\lambda)$ that gives the pressure p as a function of the mean circumferential stretch λ. Since we studied the deformation of a hollow spherical body in Section 5.4, it is convenient to make use of those results here, specializing them to the thin-walled case[9]. The two differences, (at least to start with), between the problem in Section 5.4 and that considered here is that one, the loading here is an internal pressure p per unit deformed area applied on the inner boundary with the outer boundary being traction-free, and two, the body is composed of an arbitrary isotropic incompressible elastic material. (We will want to specialize the results to a Mooney-Rivlin material, and so it is not sufficient to limit attention to a generalized neo-Hookean material.)

We start be rewriting the relevant equations from Section 5.4 (making use of the results in the Exercise at the end of that section): the deformation is given by either of the equivalent forms

$$r^3(R) = R^3 + a^3 - A^3 = R^3 + b^3 - B^3, \tag{i}$$

where a and b are the inner and outer radii in the deformed configuration; and the circumferential stretch $\lambda(R)$ is

$$\lambda(R) = \frac{r(R)}{R}. \tag{ii}$$

From the constitutive relation we obtain

$$T_{\theta\theta} - T_{rr} = \frac{1}{2}\lambda w'(\lambda), \tag{iii}$$

and the equilibrium equations reduce to

$$\frac{dT_{rr}}{d\lambda} = -\frac{w'(\lambda)}{\lambda^3 - 1}, \tag{iv}$$

where

$$w(\lambda) := W(\lambda^{-2}, \lambda, \lambda). \tag{v}$$

Here $W(\lambda_1, \lambda_2, \lambda_3)$ is the strain energy function characterizing the incompressible isotropic material. The boundary conditions in the problem considered here are

$$T_{rr}(a) = -p, \qquad T_{rr}(b) = 0. \tag{vi}$$

Integrating (iv) from $r = a$ to $r = b$ and using the boundary conditions (vi) leads to

$$p = \int_{\lambda_b}^{\lambda_a} \frac{w'(\lambda)}{\lambda^3 - 1}\, d\lambda, \tag{vii}$$

[9]Alternatively one can derive the results pertinent to this case directly as in the next problem in Section 5.6.

where λ_a and λ_b are the circumferential stretches at the inner and outer surfaces respectively:

$$\lambda_a = a/A, \qquad \lambda_b = b/B. \tag{viii}$$

Equation (vii) together with $(viii)$ and $b^3 - B^3 = a^3 - A^3$ provides a relation between the pressure p and the deformed inner radius a.

We now specialize the preceding results to the case where the body is *thin-walled*, i.e. when the wall thickness $T = B - A$ is small compared to the mean radius $\overline{R} = (A+B)/2$:

$$\epsilon := T/\overline{R} \ll 1. \tag{ix}$$

The inner and outer undeformed radii can be expressed in terms of $\overline{R}$ and ε as

$$A = \overline{R} - T/2 = \overline{R}(1-\epsilon/2), \qquad B = \overline{R} + T/2 = \overline{R}(1+\epsilon/2). \tag{x}$$

If $\overline{\lambda} = \lambda(\overline{R})$ is the mean circumferential stretch we can approximate (vii) as

$$p \approx \frac{w'(\overline{\lambda})}{\overline{\lambda}^3 - 1}\,(\lambda_a - \lambda_b).$$

Thus our immediate task is to find an approximate expression for $\lambda_a - \lambda_b$ for small ε. We shall do this by deriving expressions of the form $\lambda_a = \overline{\lambda} + \varepsilon\alpha + \mathrm{O}(\varepsilon^2)$ and $\lambda_b = \overline{\lambda} + \varepsilon\beta + \mathrm{O}(\varepsilon^2)$.

From (i), together with (ii) and $(viii)$, we have

$$\lambda_a^3 = \frac{a^3}{A^3} = \frac{R^3}{A^3}(\lambda^3 - 1) + 1, \qquad \lambda_b^3 = \frac{b^3}{B^3} = \frac{R^3}{B^3}(\lambda^3 - 1) + 1, \tag{xi}$$

where $\lambda = \lambda(R)$ is the stretch at the radius R. Now evaluate $(xi)_1$ at the mean radius $\overline{R}$ where the stretch is $\overline{\lambda}$, use $(x)_1$ and drop terms smaller than $\mathrm{O}(\epsilon)$:

$$\begin{aligned}
\lambda_a \ &\overset{(xi)_1}{=} \left[\frac{\overline{R}^3}{A^3}(\overline{\lambda}^3 - 1) + 1\right]^{1/3} \overset{(x)_1}{=} \left[\left(1 - \frac{\epsilon}{2}\right)^{-3}(\overline{\lambda}^3 - 1) + 1\right]^{1/3} = \\
&= \left[\left(1 + \frac{3\epsilon}{2}\right)(\overline{\lambda}^3 - 1) + 1\right]^{1/3} + \mathrm{O}(\epsilon^2) = \overline{\lambda}\left[1 + \frac{3\epsilon}{2}\frac{\overline{\lambda}^3 - 1}{\overline{\lambda}^3}\right]^{1/3} + \mathrm{O}(\epsilon^2) = \\
&= \overline{\lambda} + \frac{\overline{\lambda}^3 - 1}{2\overline{\lambda}^2}\epsilon + \mathrm{O}(\epsilon^2).
\end{aligned} \tag{xii}$$

Likewise from $(xi)_2$ and $(x)_2$ one obtains

$$\lambda_b = \overline{\lambda} - \frac{\overline{\lambda}^3 - 1}{2\overline{\lambda}^2}\epsilon + \mathrm{O}(\epsilon^2). \tag{xiii}$$

We can now approximate (vii) as

$$p = \frac{w'(\overline{\lambda})}{\overline{\lambda}^3 - 1}\,(\lambda_a - \lambda_b) + \mathrm{O}(\epsilon^2) \overset{(xii),(xiii)}{=} \epsilon\frac{w'(\overline{\lambda})}{\overline{\lambda}^2} + \mathrm{O}(\epsilon^2). \qquad (xiv)$$

Therefore to leading order, the pressure p is related to the mean circumferential stretch $\overline{\lambda}$ by

$$p = p(\overline{\lambda}) \;:=\; \epsilon\,\frac{w'(\overline{\lambda})}{\overline{\lambda}^2} = \frac{T}{R}\,\frac{w'(\overline{\lambda})}{\overline{\lambda}^2}. \qquad (xv)$$

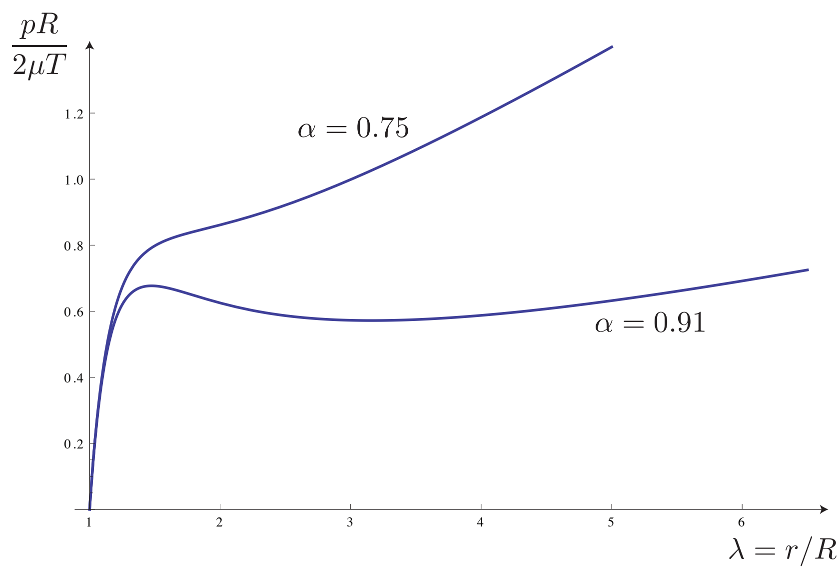


Figure 5.12: Pressure p versus the mean circumferential stretch $\overline{\lambda}$ for a thin-walled spherical shell of a Mooney-Rivlin material. The figure has been drawn for $\alpha = 0.75$ and 0.91

To illustrate the response of the thin-walled shell, consider a Mooney-Rivlin material:

$$W = \frac{\mu}{2}\Big[\alpha(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) + (1-\alpha)(\lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2} - 3)\Big], \qquad \mu > 0, \quad 0 < \alpha < 1.$$

The associated energy function $w(\lambda)$ defined by (v) is

$$w(\lambda) = \frac{\mu}{2}\Big[\alpha(\lambda^{-4} + 2\lambda^2 - 3) + (1-\alpha)(\lambda^4 + 2\lambda^{-2} - 3)\Big],$$

and so from (xv) we find that

$$\frac{pR}{T} = \frac{w'(\overline{\lambda})}{\overline{\lambda}^2} = 2\mu(\overline{\lambda} - \overline{\lambda}^{-5})(1 - \alpha + \alpha\overline{\lambda}^{-2}). \qquad (xvi)$$

Figure 5.12 shows a plot of the pressure p versus the circumferential stretch $\overline{\lambda}$ based on (xvi) for two values of α. For $\alpha = 0.75$, the pressure increases monotonically with increasing stretch but the curvature is seen to change. The curve for $\alpha = 0.91$ is quite different (once $\overline{\lambda}$ exceeds about 1.25). In particular, the relation between p and r/R loses monotonicity – the so-called limit point instability. The curve now has two rising branches connected by a declining branch. The multiple branches of the curve indicate the existence of multiple equilibrium configurations if the pressure is in a suitable range. The configuration associated with the left-most branch has a relatively small deformed radius while the configuration associated with the right-most branch has a significantly larger radius. Transition between these configurations occurs at the local maximum and minimum.

Remark 1: Before leaving this problem we shall calculate the radial and circumferential stress components T_{rr} and $T_{\theta\theta}$. For convenience we shall drop the overline on the mean stretch. First, considering the thick-walled shell, the radial stress is found by integrating (iv) from an arbitrary radius to the outer boundary leading to

$$T_{rr}(r) = -\int_{\lambda_b}^{r/R} \frac{w'(\lambda)}{\lambda^3 - 1}\, d\lambda, \qquad (xvii)$$

and then the circumferential stress is given by (iii) and $(xvii)$ to be

$$T_{\theta\theta}(r) = \frac{1}{2}\lambda w'(\lambda) - \int_{\lambda_b}^{r/R} \frac{w'(\lambda)}{\lambda^3 - 1}\, d\lambda. \qquad (xviii)$$

Approximating the expression $(xvii)$ for the radial stress gives

$$T_{rr} = -\frac{w'(\lambda)}{\lambda^3 - 1}\,(\lambda - \lambda_b) + \mathrm{O}(\epsilon^2) \overset{(xiii)}{=} -\frac{1}{2}\epsilon\frac{w'(\lambda)}{\lambda^2} + \mathrm{O}(\epsilon^2) \overset{(xiv)}{=} -\frac{p}{2} + \mathrm{O}(\epsilon^2). \qquad (xix)$$

Observe from (xvi) that

$$w'(\lambda) = \frac{p\lambda^2}{\epsilon} + \mathrm{O}(\epsilon). \qquad (xx)$$

Finally we turn to the circumferential stress $T_{\theta\theta}$. Substituting (xix) and (xx) into (iii) yields

$$T_{\theta\theta} = T_{rr} + \frac{1}{2}\lambda w'(\lambda) = -\frac{p}{2} + \frac{p\lambda^3}{2\epsilon} + \mathrm{O}(\epsilon) = \frac{p\lambda^3}{2\epsilon} + \mathrm{O}(1). \qquad (xxi)$$

The familiar expression for $T_{\theta\theta}$ in a thin-walled spherical shell involves the mean radius r and wall thickness t in the *deformed* configuration. Setting $\lambda = r/R$ and $\epsilon = T/R$ in (xxi) gives

$$T_{\theta\theta} \approx \frac{pr^3}{2TR^2}. \qquad (xxii)$$

However by incompressibility, the volumes $4\pi R^2 T$ and $4\pi r^2 t$ of the undeformed and deformed shell must be equal whence

$$R^2 T = r^2 t. \qquad (xxiii)$$

Observe from this and $\lambda = r/R$ that

$$t = T/\lambda^2. \qquad (xxiv)$$

Using $(xxiii)$ in $(xxii)$ gives the circumferential stress in the familiar form

$$T_{\theta\theta} \approx \frac{pr}{2t}, \qquad (xxv)$$

where r and t are the mean radius and wall thickness in the deformed configuration.

In summary, given the geometric parameters R and T, the constitutive description of the material $w(\lambda)$ and the applied loading p, one solves (xv) for the stretch λ, determines the deformed radius and wall-thickness from (ii) and $(xxiv)$, and calculates the circumferential stress $T_{\theta\theta}$ from (xxv).

Remark 2: Volume is the work-conjugate kinematic variable corresponding to pressure. Accordingly if v denotes the volume contained within the deformed spherical shell, it is related to the stretch by

$$v = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi R^3\lambda^3.$$

Let $\widehat{w}(v) = w(\lambda)$. Since w denotes the energy per unit reference volume, the total elastic energy in the shell is

$$\widehat{E}(v) = 4\pi R^2 T\, \widehat{w}(v).$$

It is readily seen by differentiating the two preceding equations that

$$\widehat{E}\,'(v) = \frac{dE}{dv} = 4\pi R^2 T\,\frac{d\widehat{w}(v)}{dv} = 4\pi R^2 T\,\frac{dw(\lambda)}{d\lambda}\frac{d\lambda}{dv} = 4\pi R^2 T\, w'(\lambda)\,\frac{1}{4\pi R^3\lambda^2} = \frac{T}{R}\,\frac{w'(\lambda)}{\lambda^2} \overset{(xv)}{=} p.$$

Thus we have

$$p = \widehat{p}(v) = \widehat{E}\,'(v). \qquad (xxvi)$$

Remark 3: In the thin-walled limit, where the pressurized sphere is like an inflated membrane, it is natural to absorb the wall-thickness into the other variables. Accordingly let $\mathcal{W}(\lambda)$ denote the elastic energy per unit deformed *area* so that then, since w is the energy per unit reference volume,

$$(4\pi r^2)\mathcal{W} = (4\pi R^2 T)w \qquad \Rightarrow \qquad \mathcal{W} = \frac{Tw(\lambda)}{\lambda^2}. \qquad (xxvii)$$

Also, let τ be the circumferential (hoop) force per unit deformed *length* – the "surface tension". Since $T_{\theta\theta}$ is the circumferential force per unit deformed area, they are related by

$$(2\pi r)\tau = (2\pi rt)T_{\theta\theta} \qquad \Rightarrow \qquad \tau = tT_{\theta\theta}. \qquad (xxviii)$$

Therefore we can rewrite $T_{\theta\theta} = pr/(2t)$ as

$$p = \frac{2\tau}{r}. \qquad (xxix)$$

This is the well-known Young-Laplace equation for a spherical bubble relating the pressure p, radius r and surface tension τ. Moreover the equation $p = (T/R)w'(\lambda)/\lambda^2$ can be rewritten as $\tau = (\lambda^2\mathcal{W})'/(2\lambda)$, i.e.

$$\tau = \mathcal{W} + \frac{\lambda\mathcal{W}'}{2}. \qquad (xxx)$$

In the special case where the energy per unit deformed area $\mathcal{W}$ is a constant (corresponding to $w(\lambda) = \mathsf{constant}\ \lambda^2$), then $\tau = \mathcal{W}$. This is a second commonly used relation in the study of bubbles: the surface tension equals the surface energy per unit area. The results $(xxix)$ and (xxx) are derived directly in Problem 5.1.

5.6 Example (5): Two-Phase Configurations of a Thin-Walled Tube.

References:

1. J.L. Ericksen, Equilibrium of Bars, *Journal of Elasticity*, **5**(1975), pp. 191-201.
2. J.L. Ericksen, *Introduction to the Thermodynamics of Solids*, Chapman & Hall, 1991, Chapters 3 and 5.
3. S. Kyriakides and Y-C. Chang, On the initiation and propagation of a localized instability in an inflated elastic tube, *International Journal of Solids and Structures*, **27**(1991), 1085-1111.
4. S. Kyriakides and L-H. Lee, *Mechanics of Offshore Pipelines, Volume 2: Buckle Propagation and Arrest*, Elsevier, 2020.

The phenomenon to be studied in this example is a toy model for a *phase transition* where a body can exist in one of several phases and transform between them as the (mechanical or thermal) loading is varied. This is because different phases are "preferred" (stable) at different load-levels. Often, for some intermediate range of loading, multiple phases will[10] *co-exist*. An introduction to this topic can be found in Chapter 7.

Consider a long *thin-walled* circular cylindrical tube of length L, mean radius R and wall thickness $T\,(\ll R)$ in a stress-free reference configuration. The tube is composed of an incompressible isotropic elastic material and is subjected to an internal pressure p (per unit deformed area). In the deformed configuration the tube has (an unknown) mean radius r and wall thickness t. We assume a state of plane strain so that particles do not undergo any displacement in the axial direction, and so in particular, the length of the tube in the deformed configuration is also L.

Let $\lambda = r/R$ denote the mean circumferential stretch. Then the principal stretches are

$$\lambda_R = \lambda^{-1}, \qquad \lambda_\Theta = \lambda, \qquad \lambda_Z = 1 \qquad \text{where} \quad \lambda = \frac{r}{R}, \tag{i}$$

[10]One could say that the Rivlin cube problem in Section 5.3 involves cubic and tetragonal phases, though (a) one does not encounter configurations in which both these phases co-exist and (b) at zero stress, one only encounters the cubic phase and not the tetragonal phase.

having used the plane strain assumption to write λ_Z and incompressibility to write λ_R. Let $w(\lambda)$ be the restriction of the strain energy function to deformations of the present form, i.e. let

$$w(\lambda) := W^*(\lambda^{-1}, \lambda, 1). \tag{ii}$$

The pressure-stretch relation $p = p(\lambda)$ can be shown to be

$$p = p(\lambda) = \frac{T}{R}\,\frac{w'(\lambda)}{\lambda}. \tag{iii}$$

This can be derived by specializing the solution to the thick-walled tube to the case $T/R \ll 1$ (as we did for the thin-walled sphere in Section 5.5, the corresponding equation being (xv) on page 352), or directly, as in the appendix at the end of this section.

The total elastic energy per unit length of the tube is $E = 2\pi RTw$ since w is the strain energy per unit reference volume and the reference volume of the tube per unit length is $2\pi RT$. Since volume is work-conjugate to pressure, it is convenient to let

$$v := \pi r^2 \overset{(i)}{=} \pi R^2\lambda^2 \tag{iv}$$

be the *volume* enclosed by *a unit length* of the tube in the deformed configuration, and to express E as a function of v:

$$\widehat{E}(v) = 2\pi RTw(\lambda)\Big|_{\lambda=(v/\pi R^2)^{1/2}}, \tag{v}$$

Then, (iii) and (v) gives the following pressure-volume relation:

$$p = \widehat{p}(v) := \widehat{E}\,'(v). \tag{vi}$$

Observe that the pressure is the gradient of the energy with respect to volume reflecting the work-conjugacy of p and v.

Given a specific strain energy function, one can work out the details above and obtain an explicit expression for the pressure-volume relation $p = \widehat{p}(v)$. For example for the neo-Hookean strain energy function,

$$W^*(\lambda_1, \lambda_2, \lambda_3) = \frac{\mu}{2}\,(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3),$$

we find that

$$w(\lambda) \overset{(ii)}{=} \frac{\mu}{2}(\lambda^{-2} + \lambda^2 - 2), \qquad \widehat{E}(v) \overset{(iv),(v)}{=} \mu\pi RT\left(\frac{\pi R^2}{v} + \frac{v}{\pi R^2} - 2\right).$$

Then (vi) takes the explicit form

$$p = \widehat{p}(v) = \mu \, \frac{T}{R} \left\{ 1 - \left(\frac{\pi R^2}{v} \right)^2 \right\}.$$

For certain strain energy functions such as the neo-Hookean and Gent models, the pressure is found to be a monotonically increasing function of volume. Consequently given the pressure p, there is a unique corresponding value of volume v.

However for certain other strain energy functions this relation is non-monotonic, an example of which is the following 3-term Ogden strain energy function that models a certain latex rubber material, see Kyriakides and Chang [3]:

$$W^* = \sum_{n=1}^{3} \frac{\mu_n}{\alpha_n} (\lambda_1^{\alpha_n} + \lambda_2^{\alpha_n} + \lambda_3^{\alpha_n} - 3),$$

where

$$\mu_1 = 617\,\text{kPa}, \quad \mu_2 = 1.86\,\text{kPa}, \quad \mu_3 = -9.79\,\text{kPa}, \quad \alpha_1 = 1.30, \quad \alpha_2 = 5.08, \quad \alpha_3 = -2.00.$$

For this material, as the volume v increases, the pressure p first increases until it reaches a

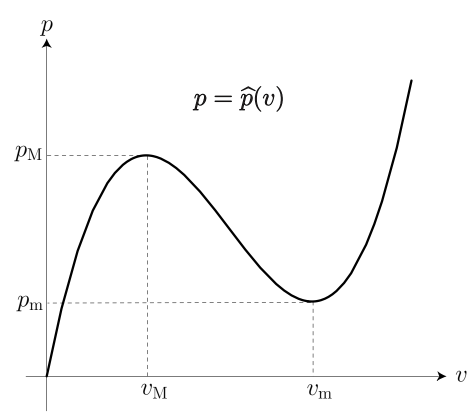


Figure 5.13: Schematic graph of $p = \widehat{p}(v) = \widehat{E}'(v)$ versus v for a certain class of strain energy functions. The pressure reaches a (local) maximum value p_M at $v = v_M$ and a (local) minimum value p_m at $v = v_m$.

value p_M, it then decreases until it reaches a value p_m, and finally increases again. Figure 5.13 depicts this *schematically* where the (local) maximum value of pressure $p = p_M$ is attained at $v = v_M$, and the (local) minimum value of pressure $p = p_m$ is attained at $v = v_m$.

We shall now discuss the consequences of having a rising-falling-rising pressure-volume curve.

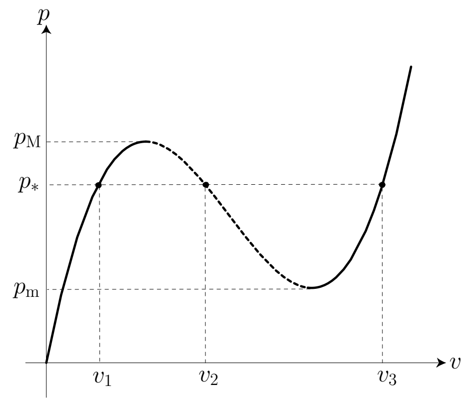


Figure 5.14: Three values of volume, v_1, v_2 and v_3, correspond to the given pressure p_*. The tube has a relatively small radius in the configuration associated with v_1 and a large radius in the configuration associated with v_3. The configuration associated with v_2 is unstable.

Loading: pressure controlled (soft device). If the given value of pressure is $< p_m$ or $> p_M$ we see from Figure 5.13 that there is a unique corresponding value of v. However if the pressure lies in the intermediate range $p_m < p < p_M$, there are three values of v, say v_1, v_2 and v_3, corresponding to the three branches of the pressure-volume curve as depicted in Figure 5.14. Thus the solution to the equilibrium problem is non-unique. Additional considerations must be taken into account in order to resolve this non-uniqueness.

Equilibrium solutions that are observed in the laboratory must be stable. Thus it is natural to look at the stability of these multiple equilibrium states. In order to examine this, one must describe more carefully the manner in which the loading is controlled. Suppose that the pressure is controlled – often called loading by a "soft device". This can be achieved, for example, by inflating the tube with an incompressible fluid using a piston carrying a weight. Changing the magnitude of the weight changes the pressure. The potential energy of the elastic tube and loading device is

$$\Phi(v; p) = \widehat{E}(v) - pv, \qquad v > 0. \tag{vii}$$

(In Section 5.3 we discussed the potential energy of an elastic body in a general setting.)

At each value of p, equation (vii) defines Φ for all positive v. The particular values of v corresponding to equilibrium configurations are given by the extrema of $\Phi(\cdot; p)$; moreover, we say that such an equilibrium configuration is stable against small disturbances (*locally stable*) if the extremum is a local minimum. Thus the volume v associated with a stable equilibrium configuration at a given value of p is determined from

$$\frac{d\Phi}{dv} = \widehat{E}\,'(v) - p = 0, \qquad \frac{d^2\Phi}{dv^2} = \widehat{E}\,''(v) > 0. \tag{viii}$$

Equation $(viii)$ tells us that at a locally stable equilibrium configuration

$$p = \widehat{E}\,'(v) = \widehat{p}(v), \qquad \widehat{p}\,'(v) > 0. \tag{ix}$$

Thus according to $(ix)_2$, the p, v-curve must be rising at a stable equilibrium configuration. This implies that the configuration associated with v_2 in Figure 5.14 is unstable, but those associated with v_1 and v_3 are both stable.

Figure 5.15 illustrates this in terms of the potential energy function: $\Phi(\cdot; p)$ has a single energy-well (local minimum) for $p < p_m$ corresponding to the first branch of the p, v-curve; a single energy-well for $p > p_M$ corresponding to the third branch of the p, v-curve; and for $p_m < p < p_M$, it has two energy-wells, one associated with the first branch and the other the third branch.

Observe that there are two (locally) stable configurations corresponding to each value of pressure in the range $p_m < p < p_M$ and so the solution to the equilibrium problem (continues to be) nonunique[11].

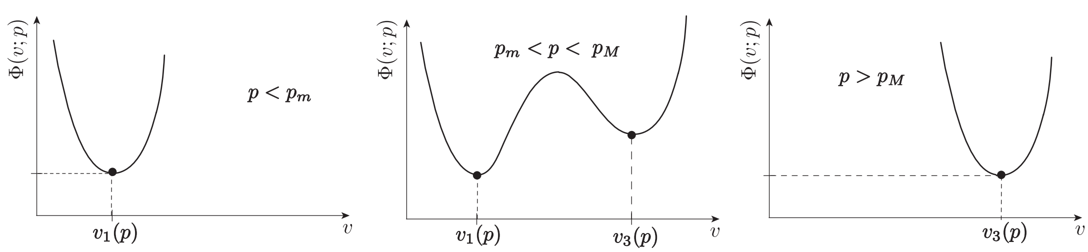


Figure 5.15: Potential energy $\Phi(v; p)$ versus volume v at three values of the pressure. Though the middle figure shows the energy-well on the left being lower than that on the right, this need not be the case; it depends on the value of $p \in (p_m, p_M)$; see Figure 5.19.

One approach to understanding this non-uniqueness is to consider the *process* by which the tube is pressurized instead of considering a strictly equilibrium problem. Say the pressure in the tube is p_* as shown in Figure 5.16. The observed value of the corresponding volume v may depend on the process by which the pressure p_* is reached: one might reasonably

[11]As can be seen from Figure 5.8, this also occurs in the problem studied in Section 5.3 for applied force values in the range $3/2^{2/3} < F/\mu < 2$.

expect based on Figure 5.16 that if the pressure had increased monotonically from 0 to p_* the associated volume would be v_1; but that if instead the pressure had decreased monotonically from some large value to p_*, the associated volume would be v_3.

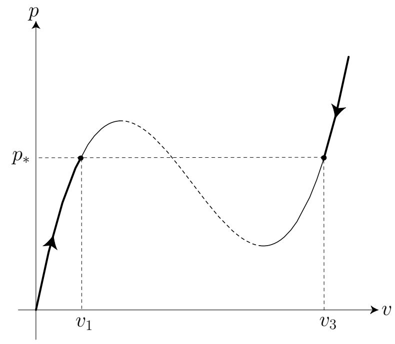


Figure 5.16: A process during which the pressure increases from 0 to p_*; and a second process during which the pressure decreases from some large value to p_*. The first process necessarily starts on the first branch of the p, v-curve, while the second process necessarily starts on the third branch of the p, v-curve.

An alternative approach would be to require an equilibrium configuration to be a *global* minimizer of the potential energy. Suppose there are two local minimizers v_1 and v_3 as shown in Figure 5.17:

$$p = \widehat{p}(v_1) = \widehat{p}(v_3), \qquad \widehat{p}\,'(v_1) > 0, \quad \widehat{p}\,'(v_3) > 0. \tag{x}$$

In the current problem, the global minimizer would be the solution with the smaller value of potential energy. Thus we must compare $\Phi(v_3; p) = \widehat{E}(v_3) - pv_3$ with $\Phi(v_1; p) = \widehat{E}(v_1) - pv_1$:

$$\Phi(v_3; p) - \Phi(v_1; p) = \Big[\widehat{E}(v_3) - pv_3\Big] - \Big[\widehat{E}(v_1) - pv_1\Big] = \int_{v_1}^{v_3} \widehat{p}(v)\, dv - p(v_3 - v_1), \tag{xi}$$

where in getting to the second equality we used the fact that $\widehat{p}(v) = \widehat{E}'(v)$. The first term represents the area below the p, v-curve between v_1 and v_3 (see Figure 5.17) while the second is the area of the rectangle with base $v_3 - v_1$ and height p. Therefore in terms of the areas shown in Figure 5.17,

$$\Phi(v_3; p) - \Phi(v_1; p) = \textsf{Area } A \ - \textsf{Area } B. \tag{xii}$$

Thus $\Phi(v_1; p) < \Phi(v_3; p)$ when $\textsf{Area } A > \textsf{Area } B$ in which case the configuration associated with v_1 is the global minimizer. On the other hand $\Phi(v_3; p) < \Phi(v_1; p)$ when $\textsf{Area } A < \textsf{Area } B$ and so the configuration associated with v_3 is the global minimizer in this case. It is readily

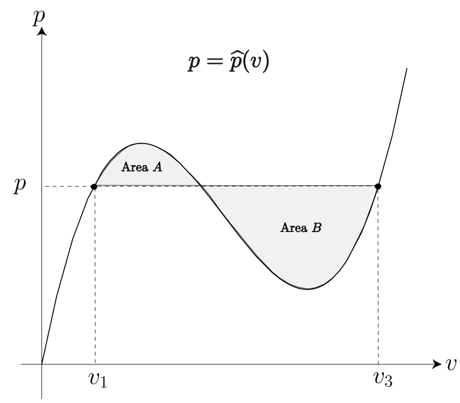


Figure 5.17: Areas A and B of two lobes cut off by the $p =$ constant line.

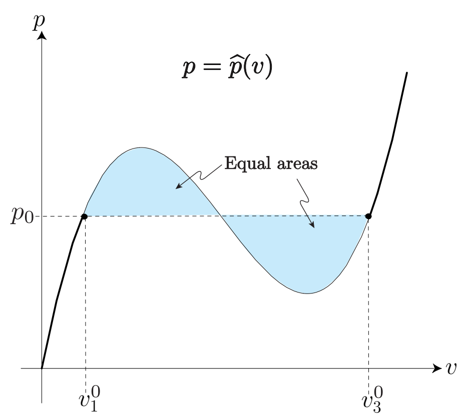


Figure 5.18: The Maxwell pressure p_0 cuts off lobes of equal area.

shown that there is a unique value of pressure, say p_0, at which these areas are equal; see Figure 5.18. It is called the *Maxwell pressure* and is given by

$$\int_{v_1(p_0)}^{v_3(p_0)} \widehat{p}(v)\, dv = p_0 \Big[v_3(p_0) - v_1(p_0) \Big]. \qquad (xiii)$$

Here the functions $v_1(p)$ and $v_3(p)$ are the inverses of $\widehat{p}(v)$ when it is restricted to the first and third branch respectively. Thus we conclude that the solutions $v = v_1$ and $v = v_3$ are the respective global minimizers for $p < p_0$ and $p > p_0$. This is illustrated by the bold curves in Figure 5.18.

Figure 5.19 shows plots of the potential energy $\Phi(v;p) = \widehat{E}(v) - pv$ versus v at three different values of $p \in (p_m, p_M)$; see Figure 5.15 for plots corresponding to $0 < p < p_m$ and $p > p_M$. The energy-well associated with the first branch of the p, v-curve is lower than that

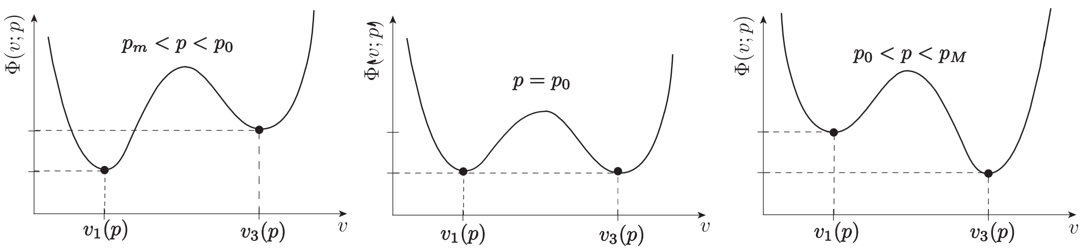


Figure 5.19: Potential energy $\Phi(v;p) = \widehat{E}(v) - pv$ versus volume v at three values of the pressure $p \in (p_m, p_M)$; see Figure 5.15 for plots corresponding to $0 < p < p_m$ and $p > p_M$

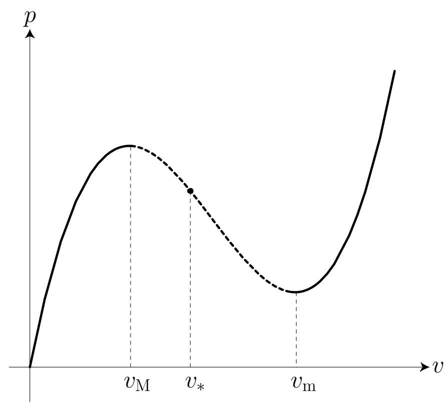


Figure 5.20: There is only one equilibrium configuration of the type discussed here corresponding to the prescribed volume v_* and it is unstable.

associated with the third branch for $p_m < p < p_0$, and the reverse is true for $p_0 < p < p_M$.

Loading: volume controlled (hard device). In this case the total volume within the tube is controlled – often called loading by a "hard device". This can be achieved, for example, by inflating the tube with an incompressible fluid using a screw: moving the screw in or out would increase or decrease the prescribed volume.

As seen from Figure 5.20, there is a unique value of pressure corresponding to *any* value of the prescribed volume v. The solutions corresponding to $v < v_M$ and $v > v_m$, i.e. the two rising branches of the pressure-volume curve, are stable against small disturbances. The solution associated with the intermediate range $v_M < v < v_m$, i.e. the falling branch of the pressure-volume curve, is unstable; see Ericksen [1] for a proof of this. Thus if the given value of v lies in the range $v_M < v < v_m$ there is no stable solution to the problem within the class of solutions we have considered (*non-existence*).

However, since we control the volume in this experiment, we are free to prescribe a value such as v_* shown in Figure 5.20. What configuration does the tube take, given that the homogeneously deformed configuration is unstable?

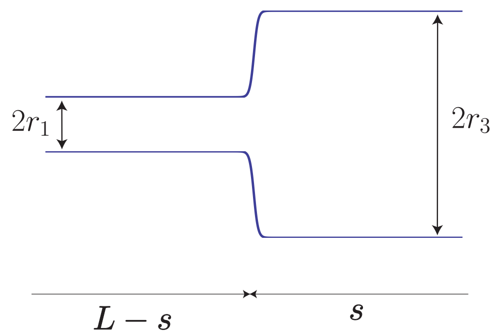


Figure 5.21: A configuration of the tube in which a length $L-s$ of the tube has a radius r_1 (where $v_1 = \pi r_1^2$ is associated with the first branch of the p, v-curve); and the remaining length s of the tube has a radius r_3 (where $v_3 = \pi r_3^2$ is associated with the third branch of the p, v-curve);

Based on the experiments of Kyriakides and Chang [3], it turns out that when $v = v_*$ the deformed configuration of the tube involves two segments, one segment associated with the first branch of the pressure-volume curve with some volume per unit length v_1, and the remaining segment being associated with the third branch of the pressure-volume curve with some volume per unit length v_3. See Figure 5.22. These two values of volume average out to give the value v_*. Thus the equilibrium configuration of the tube involves lengths of two different radii (with a transition zone joining them) as shown in Figure 5.21 (*co-existence*).

When v_* is close to v_M (see Figure 5.20), most of the tube will be associated with the first branch (and the remaining short segment with the third branch). When v_* is close to v_m, most of the tube will be associated with the third branch. Thus as the value of v_* increases from v_M to v_m, the segment associated with the first branch gets monotonically shorter as the tube transforms from the first to the third branch. See Kyriakides and Chang [3] and Kyriakides and Lee [4] for experiments that exhibit this behavior.

To make this quantitative, now consider *piecewise homogeneous* configurations of the tube in which a length s of the tube is associated with the third branch of the pressure-volume curve and has volume (per unit length) v_3. The rest of the tube of length $L - s$ is associated with the first branch of the pressure-volume curve and has volume (per unit length) v_1. Let the (prescribed) total volume in the tube be $V = v_* L$. Then

$$(L - s)v_1 + sv_3 = V. \qquad (xiv)$$

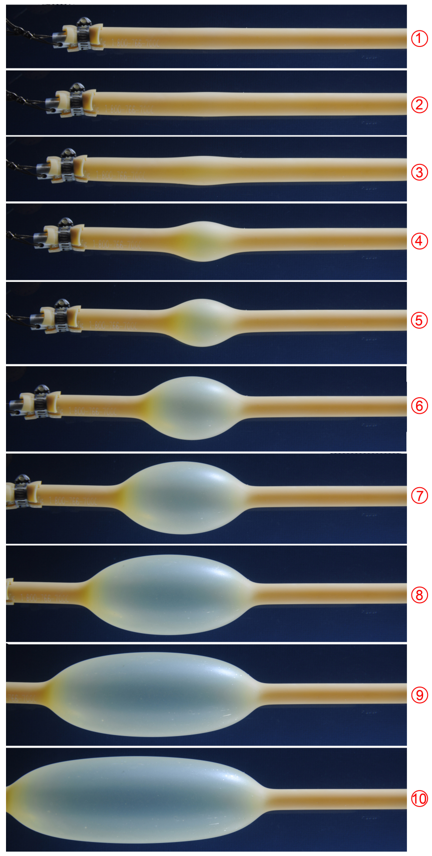


Figure 5.22: Sequence of photographs provided by Stelios Kyriakides (see also Kyriakides and Chang [3]) showing a "two-phase" equilibrium configuration of an inflated latex rubber tube with the larger radius phase growing at the expense of the other.

(We can add a constant to account for any volume inside the loading device if we wish.) Since $V = (v_3 - v_1)s + Lv_1$ and $v_3 > v_1$, it follows that s will increase as V increases (at constant v_1 and v_3). In fact, when V goes from Lv_1 to Lv_3, the length s will go from 0 to L. Thus the length of the segment of the tube associated with branch-3 increases as stated in the preceding paragraph. The potential energy of the system is[12]

$$\Phi(v_1, v_3, s) = \widehat{E}(v_1)(L - s) + \widehat{E}(v_3)s, \tag{xv}$$

there being no potential energy associated with the hard loading device. We are to minimize (xv) subject to the constraint (xiv). The constraint can be accounted for in the usual way

[12]Since the configuration of interest is not homogeneous we have to calculate the potential energy of the entire length of the tube. In contrast previously, we only needed to consider the potential energy per unit length.

through a Lagrange multiplier q and so we consider the modified function:

$$\Psi(v_1, v_3, s) = \widehat{E}(v_1)(L - s) + \widehat{E}(v_3)s - q[(L - s)v_1 + sv_3 - V].$$

Setting $\partial\Psi/\partial v_1 = \partial\Psi/\partial v_3 = 0$ gives

$$q = \widehat{E}'(v_1) = \widehat{E}'(v_3),$$

while the requirement $\partial\Psi/\partial s = 0$ leads to

$$\frac{\partial\Psi}{\partial s} = -\widehat{E}(v_1) + \widehat{E}(v_3) - q[-v_1 s + v_3] = \int_{v_1}^{v_3} \widehat{p}(v)\, dv - q(v_3 - v_1) = 0.$$

From this we conclude that $q = p_0$ is the Maxwell pressure and v_1 and v_3 have the values

$$v_1 = v_1^0 = v_1(p_0), \qquad v_3 = v_3^0 = v_3(p_0),$$

shown in Figure 5.18. The length s of the tube associated with the third branch is given by (xix) with $v_1 = v_1^0$ and $v_3 = v_3^0$:

$$s = \frac{V - v_1^0 L}{v_3^0 - v_1^0}.$$

These solutions are relevant for $v_1^0 L < V < v_3^0 L$.

Thus, as the prescribed volume V increases from $v_1^0 L$ to $v_3^0 L$ the length s increases from 0 to L, and so the length of the segment associated with the third branch gradually increases as the tube progressively transforms from the first to the third branch. For $0 < V < v_1^0 L$ there is only the homogeneous solution and it involves the first branch. Likewise for $V > v_3^0 L$ there is only the homogeneous solution and it involves the third branch. However for values of volume in the intermediate range $v_1^0 L < V < v_M L$ we have a homogeneous solution associated with the first branch *and* a piecewise homogeneous solution. Likewise for values of volume in $v_m L < V < v_3^0 L$ we have a homogeneous solution associated with the third branch and a piecewise homogeneous solution.

The experiments of Kyriakides and Chang [3] and Kyriakides and Lee [4] show that, as the prescribed volume is increased, the pressure rises along the first branch of the pressure-volume curve all the way until the pressure reaches the value p_M. Two-phase configurations then emerge and the pressure drops to the Maxwell pressure p_0. As the volume continues to increase, the pressure remains constant at the value p_0. This can be seen in the videos here provided to us by Kyriakides taken from [4].

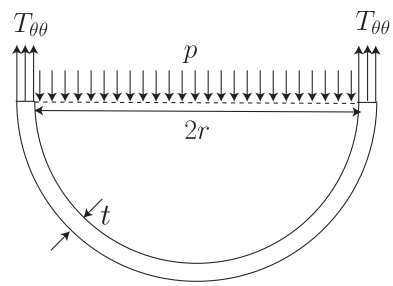


Figure 5.23: Free body diagram of a longitudinal section of the tube including the fluid it contains (in the deformed configuration). The tube has thickness t and the hoop stress is $T_{\theta\theta}$. The relevant portion of fluid has length $2r$ and pressure p. The force $2 \times (T_{\theta\theta}\, t)$ must balance the force $p(2r)$.

Appendix: Derivation of equation (vi)**:** Our approach will be to directly construct an *approximate solution* exploiting the fact that the tube is thin-walled. The results can be justified by taking the limit $T/R \to 0$ of the results of an exact analysis of a thick-walled tube just as we did for the thin-walled spherical shell in Section 5.5.

Incompressibility tells us that the volumes of material (per unit axial length of the tube), $2\pi RT$ in the undeformed configuration and $2\pi rt$ in the deformed configuration, must be equal: $2\pi RT = 2\pi rt$. Thus $t/T = R/r$ and so we can write the deformed radius r and deformed wall-thickness t in terms of the stretch λ as

$$r = \lambda R, \qquad t = T/\lambda. \tag{xvi}$$

By symmetry, the principal Cauchy stresses are $T_{rr}, T_{\theta\theta}$ and T_{zz}. The radial stress T_{rr} varies from the value $-p$ at the inner wall to the value zero at the outer wall over a small distance t. Thus we approximate T_{rr} to be

$$T_{rr} \approx -\frac{p}{2}. \tag{xvii}$$

Next consider the equilibrium of the longitudinal section of the tube shown in Figure 5.23. Note that the figure shows the tube in the *deformed configuration*. Equilibrium requires the resultant force on this free body diagram to vanish, i.e. that $p \times 2r = 2(T_{\theta\theta} \times t)$ where $T_{\theta\theta}$ is the mean Cauchy hoop stress. Thus

$$T_{\theta\theta} \approx \frac{pr}{t} \overset{(xvi)}{=} \frac{pR}{T}\,\lambda^2. \tag{xviii}$$

Observe that $T_{\theta\theta} = \mathrm{O}\big((T/R)^{-1}\big)$ while $T_{rr} = O(1)$ as $T/R \to 0$ and so $T_{\theta\theta} \gg T_{rr}$.

The principal Cauchy stress τ_i is related to the principal stretches by the constitutive relation $\tau_i = \lambda_i \partial W/\partial\lambda_i - q$ (no sum on i). By taking the 1- and 2-directions to refer to the radial and circumferential directions respectively we have

$$T_{rr} = \lambda_R \frac{\partial W}{\partial \lambda_1} - q \overset{(i)}{=} \lambda^{-1}\frac{\partial W}{\partial \lambda_1} - q, \qquad T_{\theta\theta} = \lambda_\Theta \frac{\partial W}{\partial \lambda_2} - q \overset{(i)}{=} \lambda \frac{\partial W}{\partial \lambda_2} - q.$$

The reaction pressure q can be eliminated by subtracting the first equation from the second leading to

$$T_{\theta\theta} - T_{rr} = \lambda \frac{\partial W}{\partial \lambda_2} - \lambda^{-1}\frac{\partial W}{\partial \lambda_1}.$$

Since $T_{\theta\theta} \gg T_{rr}$ we drop T_{rr} and write

$$T_{\theta\theta} \approx \lambda \frac{\partial W}{\partial \lambda_2} - \lambda^{-1} \frac{\partial W}{\partial \lambda_1}. \tag{xix}$$

Next, differentiating (ii) with respect to λ gives

$$w'(\lambda) = -\lambda^{-2} \frac{\partial W}{\partial \lambda_1} + \frac{\partial W}{\partial \lambda_2},$$

and so we can write (xix) as

$$T_{\theta\theta} = \lambda w'(\lambda). \tag{xx}$$

On combining (iv) and (vii) we get

$$p = p(\lambda) := \frac{T}{R} \lambda^{-1} w'(\lambda). \tag{xxi}$$

Finally, differentiating (v) with respect to v and using (iv) and (xxi) yields (vi).

5.7 Example(6): Surface instability of a neo-Hookean half-space.

References:

1. M.A. Biot, Surface instability of rubber in compression. *Applied Scientific Research Section A*, 12(1963), pp. 168-182. (doi:10.1007/BF03184638).
2. Yi-chao Chen, Shengyou Yang and Lewis Wheeler, Surface instability of elastic half-spaces by using the energy method, *Proceedings of the Royal Society Series A*, Volume 474 (2018).
3. M. A. Dowaikh and R. W. Ogden, On Surface Waves and Deformations in a Pre-stressed Incompressible Elastic Solid, *IMA Journal of Applied Mathematics*, 44(1990), pp. 261-284.
4. Yanping Cao and J.W. Hutchinson, From wrinkles to creases in elastomers: the instability and imperfection sensitivity of wrinkling, *Proceedings of the Royal Society Series A*, Volume 468 (2012), pp. 94-115.
5. M.K. Kang and R. Huang, Effect of surface tension on swell-induced surface instability of substrate-confined hydrogel layers, *Soft Matter*, Volume 22, 2010, pp. 5736-5742. (doi: 10.1039/c0sm00335b)

In this section we study a homogeneously deformed body that is in equilibrium under a certain loading, and inquire as to the conditions under which there may exist a second equilibrium configuration "close" to the homogeneous one, satisfying *the same loading*. If such a configuration exists, and if it is energetically preferred, this would indicate the instability of the first.

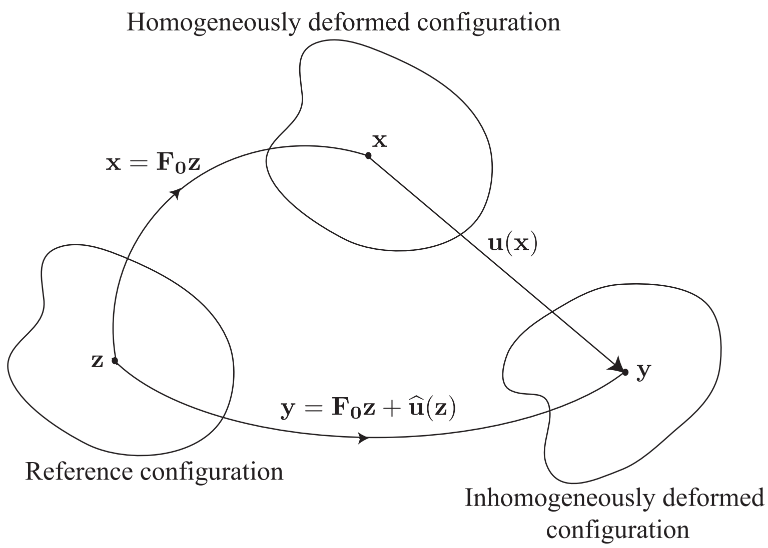


Figure 5.24: A homogeneous deformation takes $\mathbf{z} \to \mathbf{x}$. An inhomogeneous deformation takes $\mathbf{z} \to \mathbf{y}$. The displacement of a particle from the homogeneously deformed configuration to the inhomogeneously deformed configuration is $\mathbf{u}(\mathbf{x})$. In the problem of interest to us here, the two deformed configurations are "close" to each other.

In analyzing the question at hand there are three configurations to consider as shown schematically in Figure 5.24: a stress-free reference configuration, a homogeneously deformed configuration, and an inhomogeneously deformed configuration. We shall let $\mathbf{z}, \mathbf{x}$ and $\mathbf{y}$ denote the respective positions of a particle in these three configurations. Since we will have to calculate the gradients of various fields with respect to different configurations, we shall append a subscript to refer to the configuration. Thus for example $\text{grad}_z \mathbf{f}, \text{grad}_x \mathbf{f}$ and $\text{grad}_y \mathbf{f}$ will denote the gradients of $\mathbf{f}(\mathbf{z}), \mathbf{f}(\mathbf{x})$ and $\mathbf{f}(\mathbf{y})$ with respect to the reference configuration, the homogeneously deformed configuration and the inhomogeneously deformed configuration respectively; they have cartesian components $\partial f_i/\partial z_j, \partial f_i/\partial x_j$ and $\partial f_i/\partial y_j$.

We start in Section 5.7.1 by carrying out all calculations explicitly in the context of a specific boundary-value problem. This analysis will be generalized in Section 5.7.2 where we consider an arbitrary small deformation superimposed on an arbitrary homogeneous finite deformation. Problem 5.20 is concerned with an arbitrary small deformation superposed on an arbitrary (not necessarily homogeneous) finite deformation.

5.7.1 Example: Surface instability of a neo-Hookean half-space.

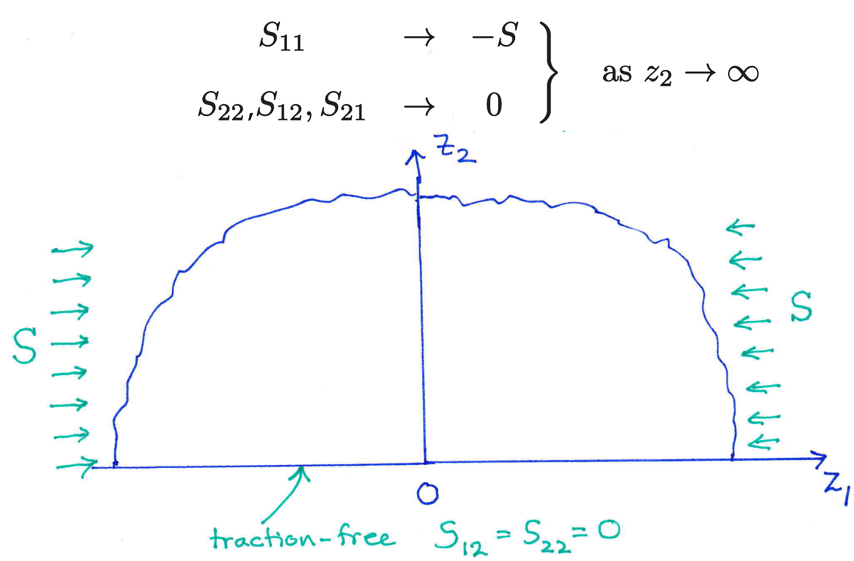


Figure 5.25: Semi-infinite neo-Hookean body with traction-free surface subjected to a uniaxial compression.

The wrinkling and creasing of surfaces under compression are of interest in various applications as described in the papers listed above and the reference is them. The particular problem we consider is the following: in a stress-free reference configuration the body occupies the half-space $z_2 > 0$ as depicted in Figure 5.25. The rectangular cartesian coordinates of a generic particle in this configuration are denoted by (z_1, z_2, z_3). All components are taken with respect to a fixed basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The surface $z_2 = 0$ is traction-free which implies that $\mathbf{S}\mathbf{e}_2 = \mathbf{0}$:

$$S_{12} = S_{22} = S_{32} = 0 \qquad \text{for } z_2 = 0. \tag{i}$$

A uniform compressive first Piola-Kirchhoff normal stress of magnitude S parallel to the z_1-axis is applied remotely as depicted in Figure 5.25. We model this by requiring

$$S_{11} \to -\mathrm{S}, \quad \text{all other } \ S_{ij} \text{ (except } S_{33}) \ \to 0 \qquad \text{as } |\mathbf{z}| \to \infty; \tag{ii}$$

the stress component S_{33} tends to some finite value as will be discussed below.

We **first** consider a homogeneous deformation that is consistent with (the field equations and) the preceding boundary conditions. Let (x_1, x_2, x_3) be the rectangular cartesian coordinates of a particle in the homogeneously deformed configuration, the deformation that takes $(z_1, z_2, z_3) \to (x_1, x_2, x_3)$ being

$$x_1 = \lambda_1 z_1, \quad x_2 = \lambda_2 z_2, \quad x_3 = \lambda_3 z_3. \tag{iii}$$

The constant stretches λ_1, λ_2 and λ_3 are to be determined. The deformation gradient tensor associated with (iii) is

$$\mathbf{F}_0 = \mathrm{Grad}_z \mathbf{x}(\mathbf{z}) = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{iv}$$

For a tensor $\mathbf{A}$ associated with the homogeneous configuration we will interchangeably use the notation $\mathbf{A}^0$ and $\mathbf{A}_0$. (The former is more convenient when, for example, we want to show its components A^0_{ij}, whereas the latter is preferred when, say, we want to write $\mathbf{A}_0^{-1}$). Incompressibility requires

$$\det \mathbf{F}_0 = \lambda_1 \lambda_2 \lambda_3 = 1. \tag{v}$$

Assume the material to be neo-Hookean. The first Piola-Kirchhoff stress tensor $\mathbf{S}_0$ is then related to the deformation gradient tensor $\mathbf{F}_0$ by the constitutive relation

$$\mathbf{S}_0 = \mu \mathbf{F}_0 - q_0 \mathbf{F}_0^{-T}, \tag{vi}$$

where the constant q_0 is the reactive pressure associated with the incompressibility constraint. From (iv) and (vi), the first Piola-Kirchhoff stress tensor $\mathbf{S}_0$ is

$$\mathbf{S}_0 = (\mu\lambda_1 - q_0\lambda_1^{-1})\, \mathbf{e}_1 \otimes \mathbf{e}_1 + (\mu\lambda_2 - q_0\lambda_2^{-1})\, \mathbf{e}_2 \otimes \mathbf{e}_2 + (\mu\lambda_3 - q_0\lambda_3^{-1})\, \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{vii}$$

Since this stress field is uniform, the equilibrium equation $\mathrm{Div}_z\, \mathbf{S}_0(\mathbf{z}) = \mathbf{o}$ holds automatically. In order to conform with the boundary conditions (i) we must have $S^0_{22} = 0$ and so from (vii),

$$q_0 = \mu\lambda_2^2. \tag{$viii$}$$

The remote prescribed loading condition (ii) requires $S^0_{11} = -S$, which by (vii) and $(viii)$ leads to the stress- stretch relation

$$S = \mu \left(\lambda_2^2 \lambda_1^{-1} - \lambda_1 \right). \tag{ix}$$

Thus we have

$$\mathbf{S}_0 = -S\, \mathbf{e}_1 \otimes \mathbf{e}_1 + \mu(\lambda_3 - \lambda_2^2\lambda_3^{-1})\, \mathbf{e}_3 \otimes \mathbf{e}_3.$$

The stretches are to be determined from $(v), (ix)$ and one more condition pertaining to either λ_3 or S^0_{33}. By leaving this condition unspecified we are able to describe several sub-cases. For example, if the homogeneous deformation is one of *plane strain* in the z_1, z_2-plane, one has $\lambda_3 = 1$ and therefore it follows from (v) and (vii) that

$$\lambda_2 = \lambda_1^{-1}, \qquad \lambda_3 = 1, \qquad S^0_{33} = \mu(1 - \lambda_1^{-2}). \tag{x}$$

On the other hand if the body is in a state of uniaxial stress in the z_1-direction, one similarly finds using $S^0_{33} = 0$ that

$$\lambda_2 = \lambda_1^{-1/2}, \qquad \lambda_3 = \lambda_1^{-1/2}, \qquad S^0_{33} = 0. \tag{xi}$$

If instead the body is in a state of equi-biaxial stretch $\lambda_1 = \lambda_3$ one finds

$$\lambda_2 = \lambda_1^{-2}, \qquad \lambda_3 = \lambda_1, \qquad S^0_{33} = \mu(\lambda_1 - \lambda_1^{-5}). \tag{xii}$$

The Cauchy stress tensor corresponding to (vii) is given by $\mathbf{T}_0 = \mathbf{S}_o \mathbf{F}_0^T$:

$$\mathbf{T}_0 = T\,\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_3 S^0_{33}\,\mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \text{where} \quad T := -S\lambda_1 = \mu(\lambda_1^2 - \lambda_2^2). \tag{$xiii$}$$

Given $S > 0$, we seek $\lambda_1, \lambda_2, \lambda_3$ from $(v), (ix)$ and either $(x)_1, (xi)_1$ or $(xii)_1$.

Since the deformation has the form (iii), we see that the region occupied by the body in the homogeneously deformed configuration is

$$x_2 > 0, \qquad -\infty < x_1 < \infty; \tag{xiv}$$

see the middle figure in Figure 5.26.

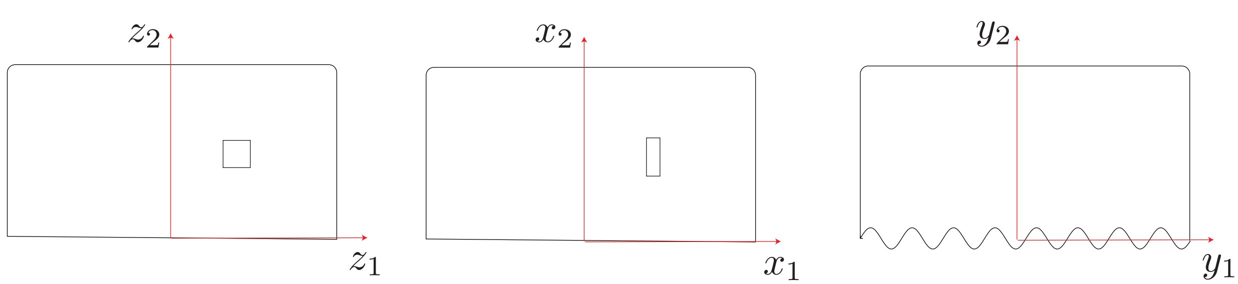


Figure 5.26: Coordinate systems: Left: reference configuration. Middle: Homogeneously deformed configuration. Right: Inhomogeneously deformed configuration.

We now seek a **second** deformation corresponding to the same loading (i.e. boundary conditions). This deformation, which is necessarily inhomogeneous, takes the particle located at $\mathbf{z}$ in the reference configuration to the location $\mathbf{y}$ in the deformed configuration:

$$\mathbf{y} = \mathbf{F}_0\mathbf{z} + \widehat{\mathbf{u}}(\mathbf{z}). \tag{xv}$$

We continue to use the preceding stress-free configuration as the reference configuration. Then the deformation gradient tensor, equilibrium equation and constitutive relation are

$$\mathbf{F} = \mathrm{Grad}_z\mathbf{y}, \qquad \mathrm{Div}_z\mathbf{S} = \mathbf{o}, \qquad \mathbf{S} = \mu\mathbf{F} - q\mathbf{F}^{-T}. \tag{xvi}$$

One could of course use the homogeneously deformed configuration as the reference configuration in which case the constitutive relation has to be modified in order to take into account the stress in this reference configuration.

Observe from $(iii), (iv)$ and (xv) that $\mathbf{y} = \mathbf{x} + \widehat{\mathbf{u}}(\mathbf{z})$ and so $\widehat{\mathbf{u}}$ is the displacement *from the homogeneously deformed* configuration to the inhomogeneously deformed configuration, see Figure 5.24. As one might expect, it is convenient to change variables and express this displacement field as a function of $\mathbf{x}$ rather than $\mathbf{z}$ by introducing the function

$$\mathbf{u}(\mathbf{x}) = \widehat{\mathbf{u}}(\mathbf{z}) \quad \text{where} \quad \mathbf{z} = \mathbf{F}_0^{-1}\mathbf{x}. \tag{$xvii$}$$

The deformation (xv) can then be expressed as

$$\mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x}). \tag{$xviii$}$$

We shall limit attention to superposed displacement fields $\mathbf{u}$ of the plane strain form

$$\mathbf{u}(\mathbf{x}) = u_1(x_1, x_2)\mathbf{e}_1 + u_2(x_1, x_2)\mathbf{e}_2. \tag{xix}$$

The inhomogeneous deformation can therefore be written out as

$$\left.\begin{aligned} y_1 &= \lambda_1 z_1 + u_1(x_1, x_2), \\ y_2 &= \lambda_2 z_2 + u_2(x_1, x_2), \\ y_3 &= \lambda_3 z_3, \end{aligned}\right\} \tag{xx}$$

where the x_i's are related to the z_i's by (iii). In the analysis going forward, we assume that the inhomogeneous deformation is close to the homogeneous deformation in the sense that

$$\epsilon := |\mathrm{grad}_x\mathbf{u}| \ll 1. \tag{xxi}$$

Accordingly we shall consistently drop terms that are of $\mathrm{O}(\epsilon^2)$. Observe that in previous chapters we linearized the equations of finite elasticity about the reference configuration whereas here we will be linearizing about the homogeneously deformed configuration; the former therefore corresponds to a special case of the latter. In order to minimize the cumbersomeness of various expressions to follow, it will convenient to use a comma followed by a subscript to indicate partial differentiation with respect to the corresponding x-coordinate, for example to write

$$u_{i,j} = \frac{\partial u_i}{\partial x_j}. \qquad (xxii)$$

We shall use this convention from hereon in the rest of this section.

The components of the deformation gradient tensor with respect to the reference configuration are given by $F_{ij} = \partial y_i/\partial z_j$. From this, (xx) and (iii) one finds[13]

$$[F] = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} + \begin{pmatrix} u_{1,1}\lambda_1 & u_{1,2}\lambda_2 & 0 \\ u_{2,1}\lambda_1 & u_{2,2}\lambda_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad (xxiii)$$

Incompressibility requires[14]

$$1 = \det \mathbf{F} \stackrel{(xxiii)}{=} 1 + u_{1,1} + u_{2,2} + \mathrm{O}(\epsilon^2), \qquad (xxiv)$$

where we have used $\det \mathbf{F}_0 = 1$. Thus to leading order, incompressibility requires

$$u_{1,1} + u_{2,2} = 0, \qquad \text{for} -\infty < x_1 < \infty,\ x_2 \geq 0. \qquad (xxv)$$

The components of $\mathbf{F}^{-1}$ can be readily calculated from $(xxiii)$ to be[15]

$$[F]^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & 0 \\ 0 & \lambda_2^{-1} & 0 \\ 0 & 0 & \lambda_3^{-1} \end{pmatrix} + \begin{pmatrix} -\lambda_1^{-1}u_{1,1} & -\lambda_1^{-1}u_{1,2} & 0 \\ -\lambda_2^{-1}u_{2,1} & -\lambda_2^{-1}u_{2,2} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathrm{O}(\epsilon^2)\ . \qquad (xxvi)$$

[13]Note that $F_{ij} = \partial y_i/\partial z_j = F^0_{ij} + \partial u_i/\partial z_j = F^0_{ij} + (\partial u_i/\partial x_k)(\partial x_k/\partial z_j) = F^0_{ij} + (\partial u_i/\partial x_k)F^0_{kj}$ and so $\mathbf{F} = \mathbf{F}_0 + \mathbf{H}\mathbf{F}_0$ where $\mathbf{H} = \mathrm{grad}_x\mathbf{u}$.

[14]Note that $\det \mathbf{F} = \det(\mathbf{I}+\mathbf{H})\mathbf{F}_0 = \det \mathbf{F}_0 \det(\mathbf{I}+\mathbf{H}) = \det(\mathbf{I}+\mathbf{H}) = 1 + \mathrm{tr}\mathbf{H} + \mathrm{O}(\epsilon^2)$.

[15]Note that $\mathbf{F}^{-1} = [(\mathbf{I}+\mathbf{H})\mathbf{F}_0]^{-1} = \mathbf{F}_0^{-1}(\mathbf{I}+\mathbf{H})^{-1} = \mathbf{F}_0^{-1}(\mathbf{I}-\mathbf{H}) + \mathrm{O}(\epsilon^2)$.

Using $(xxiii)$ and $(xxvi)$ in the constitutive law $\mathbf{S} = \mu\mathbf{F} - q\mathbf{F}^{-T}$ and linearizing leads to

$$\left.\begin{aligned}
&S_{11} = -S + \mu(\lambda_1^2 + \lambda_2^2)\lambda_1^{-1}u_{1,1} - \lambda_1^{-1}\tilde{q},\\
&S_{12} = \mu\lambda_2\left(u_{1,2} + u_{2,1}\right),\\
&S_{21} = \mu(\lambda_1^2 u_{2,1} + \lambda_2^2 u_{1,2})\lambda_1^{-1},\\
&S_{22} = 2\mu\lambda_2 u_{2,2} - \tilde{q}\lambda_2^{-1},\\
&S_{33} = S_{33}^0 - \widetilde{q}\lambda_3^{-1}, \qquad S_{13} = S_{23} = S_{31} = S_{32} = 0.
\end{aligned}\right\} \tag{xxvii}$$

In arriving at $(xxvii)$ we have dropped terms of $\mathrm{O}(\epsilon^2)$ and approximated the reactive pressure as

$$q(\mathbf{x}) = q_0 + \widetilde{q}(\mathbf{x}), \qquad \widetilde{q} = \mathrm{O}(\epsilon). \tag{xxviii}$$

Upon using the chain rule we can write the equilibrium equation $\mathrm{Div}_z\,\mathbf{S} = \mathbf{o}$ as $\mathrm{Div}_x\,(\mathbf{S}\mathbf{F}_0^T) = \mathbf{o}$. Substituting $(xxvii)$ into this leads to

$$\left.\begin{aligned}
&\widetilde{q}_{,1} = \mu(\lambda_1^2 + \lambda_2^2)u_{1,11} + \mu\lambda_2^2\left(u_{1,22} + u_{2,12}\right),\\
&\widetilde{q}_{,2} = \mu\lambda_1^2 u_{2,11} + \mu\lambda_2^2\left(u_{1,12} + 2u_{2,22}\right).
\end{aligned}\right\} \tag{xxix}$$

The third equilibrium equation $\partial S_{31}/\partial z_1 + \partial S_{32}/\partial z_2 + \partial S_{33}/\partial z_3 = 0$ yields $\partial\widetilde{q}/\partial z_3 = 0$ which tells us that

$$\widetilde{q}(\mathbf{x}) = \widetilde{q}(x_1, x_2). \tag{xxx}$$

In the far-field the first Piola-Kirchhoff stress tensor $\mathbf{S} \to \mathbf{S}_0$ and so we must have

$$u_{\alpha,\beta} \to 0 \ \text{ for } \ \alpha,\beta = 1,2, \qquad \widetilde{q} \to 0 \qquad \text{as} \quad |\mathbf{x}| \to \infty. \tag{xxxi}$$

The traction-free boundary condition requires $\mathbf{S}\mathbf{e}_2 = \mathbf{0}$ on $z_2 = 0$ which by $(xxvii)$ leads to

$$\left.\begin{aligned}
&u_{1,2} + u_{2,1} = 0,\\
&2\mu\lambda_2^2 u_{2,2} - \tilde{q} = 0,
\end{aligned}\right\} \qquad \text{for} \quad x_2 = 0. \tag{xxxii}$$

In summary, the unknown fields $u_1(x_1,x_2), u_2(x_1,x_2), \widetilde{q}(x_1,x_2)$ must obey the incompressibility equation (xxv), the equilibrium equations $(xxix)_1$ and $(xxix)_2$, the boundary conditions $(xxxii)$ at $x_2 = 0$ and the decay condition $(xxxi)$ in the far-field. Observe that $u_1(x_1,x_2) = 0, u_2(x_1,x_2) = 0, \widetilde{q}(x_1,x_2) = 0$ is one solution of this problem (corresponding to the homogeneous deformation). If any non-trivial solutions exist we expect them to do so at particular values of the applied stretch λ_1 in which case this would be an eigenvalue problem.

Simplification: Before solving the boundary value problem just formulated, it is possible to simplify it in two ways.

First we eliminate the pressure field $\widetilde{q}$ from the problem as follows. Differentiating $(xxix)_1$ with respect to x_2 and $(xxix)_2$ with respect to x_1 and equating the resulting expressions eliminates $\widetilde{q}$ from the field equations and leads to the differential equation

$$\lambda_1^2 u_{1,112} + \lambda_2^2 u_{1,222} - \lambda_1^2 u_{2,111} - \lambda_2^2 u_{2,122} = 0. \qquad (xxxiii)$$

Similarly $\widetilde{q}$ can be eliminated from the boundary conditions as follows: since $(xxxii)$ holds along the boundary, i.e. for all x_1, it may be differentiated with respect to x_1. Thereafter $(xxix)_1$ can be used to eliminate $\partial\widetilde{q}/\partial x_1$ from the result. This leads to

$$(\lambda_1^2 + \lambda_2^2) u_{1,11} + \lambda_2^2 u_{1,22} - \lambda_2^2 u_{2,12} = 0 \qquad \text{for} \quad x_2 = 0. \qquad (xxxiv)$$

Thus in summary the displacement fields $u_1(x_1, x_2), u_2(x_1, x_2)$ must obey the field equations (xxv) and $(xxxiii)$, i.e.

$$\left.\begin{aligned} u_{1,1} + u_{2,2} &= 0, \\ \lambda_1^2 u_{1,112} + \lambda_2^2 u_{1,222} - \lambda_1^2 u_{2,111} - \lambda_2^2 u_{2,122} &= 0 \end{aligned}\right\} \quad \text{for} \quad -\infty < x_1 < \infty, \ x_2 \geq 0, \quad (xxxv)$$

and the boundary conditions $(xxxii)_1$ and $(xxxiv)$, i.e.

$$\left.\begin{aligned} u_{1,2} + u_{2,1} &= 0, \\ (\lambda_1^2 + \lambda_2^2) u_{1,11} + \lambda_2^2 u_{1,22} - \lambda_2^2 u_{2,12} &= 0, \end{aligned}\right\} \quad \text{for} \quad x_2 = 0. \qquad (xxxvi)$$

In addition, in the far field

$$u_{i,j} \to 0 \qquad \text{as} \quad x_2 \to \infty \text{ at each fixed } x_1. \qquad (xxxvii)$$

The problem can be simplified even further since the general solution of the incompressibility equation $(xxxv)_1$ can be written down in terms of an arbitrary scalar potential $\phi(x_1, x_2)$ as

$$u_1 = \phi_{,2}, \quad u_2 = -\phi_{,1}. \qquad (xxxviii)$$

Substituting this into the differential equation $(xxxiii)$ leads to

$$\lambda_1^2 \ \phi_{,1111} + (\lambda_1^2 + \lambda_2^2) \ \phi_{,1122} + \lambda_2^2 \ \phi_{,2222} = 0 \qquad \text{for } x_2 > 0, \ -\infty < x_1 < \infty, \qquad (xxxix)$$

while the boundary conditions $(xxxvi)$ yield

$$(\lambda_1^2 + 2\lambda_2^2)\phi_{,112} + \lambda_2^2\phi_{,222} = 0 \qquad \text{for } x_2 = 0, \tag{XL}$$

$$\phi_{,22} - \phi_{,11} = 0 \qquad \text{for } x_2 = 0. \tag{XLI}$$

The boundary value problem $(xxxix), (XL), (XLI)$ is in fact an eigenvalue problem. We are interested in the values of λ_1/λ_2 for which it has a nontrivial solution $\phi(x_1, x_2)$.

Solution: We look for solutions of the differential equation $(xxxix)$ that are (a) periodic in the x_1-direction (and therefore of the form e^{ikx_1}) and (b) exponentially decaying away from the free-surface (and therefore of the form e^{skx_2} where $ks < 0$.) Thus we seek solutions of the form

$$\phi = \mathrm{e}^{skx_2 + ikx_1}, \tag{$XLII$}$$

where s and k are unknown constants. Without loss of generality we can assume $k > 0$. Substituting $(XLII)$ into $(xxxix)$ leads to the following quartic equation for determining s:

$$\lambda_2^2 s^4 - (\lambda_1^2 + \lambda_2^2)s^2 + \lambda_1^2 = 0, \tag{$XLIII$}$$

the roots of which are

$$s = \pm 1, \qquad \pm\lambda_1/\lambda_2. \tag{$XLIV$}$$

Since we want the displacement field to decay as $x_2 \to +\infty$, and since the displacement arising from $(xxxviii), (XLII)$ involves the term $\mathrm{e}^{skx_2},\ k > 0$, we discard the two positive roots $s = +1, +\lambda_1/\lambda_2$.

Thus we have two linearly independent solutions, $\mathrm{e}^{-kx_2+ikx_1}$ and $\mathrm{e}^{-(\lambda_1/\lambda_2)kx_2+ikx_1}$, of the differential equation $(xxxix)$. This leads us to seek a solution of the complete boundary-value problem of the form

$$= \Big[A\mathrm{e}^{-kx_2} + B\mathrm{e}^{-(\lambda_1/\lambda_2)kx_2}\Big]\mathrm{e}^{ikx_1}. \tag{XLV}$$

Equation (XLV) satisfies the equilibrium equation for any choice of the constants A and B. Substituting (XLV) into the boundary conditions (XL) and (XLI) yields a pair of algebraic equations which we write in matrix form as

$$\begin{pmatrix} \lambda_1^2 + \lambda_2^2 & 2\lambda_1\lambda_2 \\ & \\ 2 & 1 + \lambda_1^2/\lambda_2^2 \end{pmatrix} \begin{pmatrix} A \\ \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ \\ 0 \end{pmatrix}. \tag{$XLVI$}$$

If $(XLVI)$ is to have a nontrivial solution for A and B, the determinant of the 2×2 matrix must vanish and this requires

$$\left(\frac{\lambda_1}{\lambda_2}\right)^4 + 2\left(\frac{\lambda_1}{\lambda_2}\right)^2 - 4\left(\frac{\lambda_1}{\lambda_2}\right) + 1 = 0. \tag{$XLVII$}$$

One root of this equation is $\lambda_1/\lambda_2 = 1$. However by (ix), this corresponds to $S = 0$ and so we discard this root. We therefore cancel out the factor $\lambda_1/\lambda_2 - 1$ and obtain

$$\left(\frac{\lambda_1}{\lambda_2}\right)^3 + \left(\frac{\lambda_1}{\lambda_2}\right)^2 + 3\left(\frac{\lambda_1}{\lambda_2}\right) - 1 = 0. \tag{XLVIII}$$

We are interested in the real positive roots λ_1/λ_2 of (L). Consider the function $f(\xi) = \xi^3 + \xi^2 + 3\xi - 1$ for $\xi \geq 0$ and note that $f'(\xi) > 0$ for $\xi > 0$. Therefore f increases monotonically with ξ. Since $f(0) = -1 < 0$ and $f(1) = 4 > 0$ it follows that $f(\xi)$ has a unique positive zero in the interval $(0, 1)$. This proves that equation $(XLVIII)$ has exactly one real positive root. Numerical solution gives this root to be

$$\frac{\lambda_1}{\lambda_2} \approx 0.295598 \qquad \text{and so by (v)}, \quad \lambda_1 \approx \frac{0.543689}{\sqrt{\lambda_3}}. \tag{XLIX}$$

As noted previously in the context of $(x), (xi)$ and (xii), our analysis covers several sub-cases. Consider for example *the case of plane strain* where by (x),

$$\lambda_2 = \lambda_1^{-1}, \qquad \lambda_3 = 1. \tag{L}$$

Then $(XLIX)$ gives

$$\lambda_1 = \lambda_{\text{cr}} \approx 0.543689. \tag{LI}$$

Thus we conclude that an inhomogeneous deformation of the form (xv) is possible if the stretch λ_1 has the value λ_{cr}, the corresponding value of stress S_{cr} being given by (ix).

To determine the complete deformation we obtain the ratio A/B from $(XLVI)$ and then the displacement field from $(xxxviii)$ and (XLV). Observe that the constant k (which represents the reciprocal of the wave length of the oscillations in the x_1-direction) remains arbitrary. This is because our problem statement for the semi-infinite body involves no length scale.

The original solution of this problem is due to Biot. For a discussion of stability, see Chen, Yang and Wheeler, and for an analysis not limited to neo-Hookean materials, see Dowaikh and Ogden. For other modes of surface instability (such as creasing), see Cao and Hutchinson. See M.K. Kang and R. Huang, Soft Matter, (2010), DOI: 10.1039/c0sm00335b, for a treatment of wrinkling in a hydrogel.

5.7.2 An arbitrary small deformation superimposed on an arbitrary homogeneous finite deformation.

The preceding analysis can be carried out rather generally (and then specialized to the specific problem of interest, whether it be the problem studied in the previous section or some other problem). We now illustrate this by considering an arbitrary small deformation superimposed on an arbitrary homogeneous deformation. Many of the results below hold even if the deformation about which we are linearizing is not homogenous, see Problem 5.20.

We know that the natural stress measures to use when working on the reference configuration and the current configuration are the first Piola-Kirchhoff stress tensor field $\mathbf{S}(\mathbf{z})$ and the Cauchy stress tensor field $\mathbf{T}(\mathbf{y})$ respectively. As we shall see below, it will be convenient to work with a stress measure $\boldsymbol{\Sigma}(\mathbf{x})$ when working on the intermediate configuration.

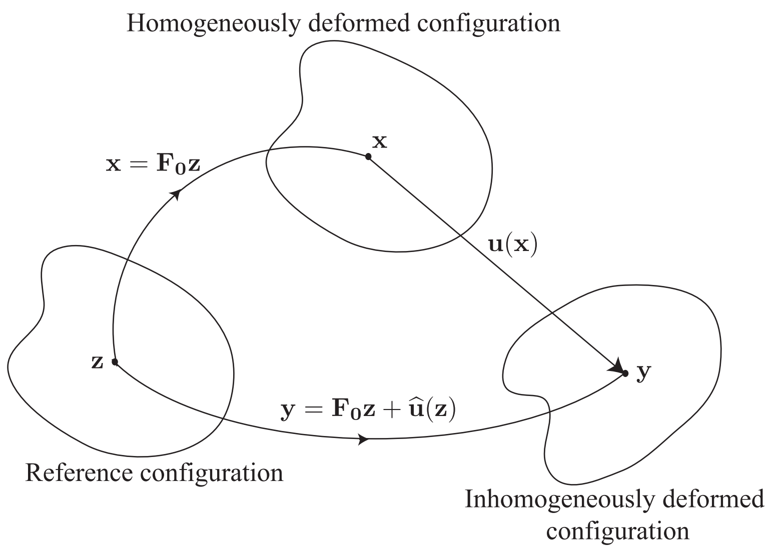


Figure 5.27: A homogeneous deformation takes $\mathbf{z} \to \mathbf{x}$. An inhomogeneous deformation takes $\mathbf{z} \to \mathbf{y}$. The displacement of a particle from the homogeneously deformed configuration to the inhomogeneously deformed configuration is $\mathbf{u}(\mathbf{x})$. The two deformed configurations are "close" to each other.

Consider a homogeneous deformation of the body

$$\mathbf{x} = \mathbf{F}_0\mathbf{z} \tag{5.11}$$

that takes a particle located at $\mathbf{z}$ in the reference configuration to the location $\mathbf{x}$ in the deformed configuration, the deformation gradient tensor $\mathbf{F}_0$ being constant. The material is incompressible and so

$$\det \mathbf{F}_0 = 1. \tag{5.12}$$

Assuming the material to be an arbitrary homogeneous, incompressible elastic material, the first Piola-Kirchhoff stress tensor $\mathbf{S}_0$ is related to the deformation gradient tensor $\mathbf{F}_0$ by

$$\mathbf{S}_0 = \left.\frac{\partial W}{\partial \mathbf{F}}\right|_{\mathbf{F}=\mathbf{F_0}} - q_0 \mathbf{F}_0^{-T}, \tag{5.13}$$

where the constant q_0 is the reactive pressure associated with the incompressibility constraint. Since the stress field is uniform, the equilibrium equations hold automatically.

Now consider an inhomogeneous deformation

$$\mathbf{y} = \mathbf{F}_0\mathbf{z} + \widehat{\mathbf{u}}(\mathbf{z}), \tag{5.14}$$

in which the particle located at $\mathbf{z}$ in the reference configuration is carried to the location $\mathbf{y}$ in the deformed configuration. Observe from (5.11) and (5.14) that $\mathbf{y} = \mathbf{x} + \widehat{\mathbf{u}}(\mathbf{z})$ and so $\widehat{\mathbf{u}}$ is the displacement *from the homogeneously deformed configuration* to the inhomogeneously deformed configuration, see Figure 5.27. As one might expect, it is convenient to change variables and express this displacement field as a function of $\mathbf{x}$ rather than $\mathbf{z}$ by introducing the function

$$\mathbf{u}(\mathbf{x}) = \widehat{\mathbf{u}}(\mathbf{z}) \quad \text{with} \quad \mathbf{x} = \mathbf{F}_0\mathbf{z}. \tag{5.15}$$

The deformation (5.14) can now be written

$$\mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x}). \tag{5.16}$$

When the context makes clear as to whether we are working with $\mathbf{u}(\mathbf{x})$ or $\widehat{\mathbf{u}}(\mathbf{z})$, we will omit the hat.

Let $\nabla_z\mathbf{u}$ and $\nabla_x\mathbf{u}$ denote the displacement gradient tensors whose cartesian components are

$$(\nabla_z\mathbf{u})_{ij} = \frac{\partial u_i}{\partial z_j}, \qquad (\nabla_x\mathbf{u})_{ij} = \frac{\partial u_i}{\partial x_j}. \tag{5.17}$$

It will be convenient to let

$$\mathbf{H} = \nabla_x\mathbf{u}. \tag{5.18}$$

By the chain rule,

$$\nabla_z\mathbf{u} = \nabla_x\mathbf{u}\,\mathbf{F}_0 = \mathbf{H}\mathbf{F}_0. \tag{5.19}$$

The deformation gradient tensor associated with (5.14) (with respect to the reference configuration) is

$$\mathbf{F} = \nabla_z \mathbf{y} = \mathbf{F}_0 + \nabla_z \mathbf{u} = \mathbf{F}_0 + \mathbf{H}\,\mathbf{F}_0 \;=\; (\mathbf{I}+\mathbf{H})\mathbf{F}_0, \tag{5.20}$$

and the Jacobian determinant is

$$\det \mathbf{F} = \det\Big[(\mathbf{I}+\mathbf{H})\mathbf{F}_0\Big] = \;\det(\mathbf{I}+\mathbf{H}), \tag{5.21}$$

having used (5.12).

Going forward, we shall assume that

$$\epsilon := |\nabla_x \mathbf{u}| = |\mathbf{H}| \ll 1 \tag{5.22}$$

and approximate all equations based on this, neglecting terms that are quadratic or smaller in ϵ. Thus in particular the incompressibility requirement $\det \mathbf{F} = \det(\mathbf{I}+\mathbf{H}) = 1$ gives, to leading order,

$$\operatorname{tr}\mathbf{H} + \mathrm{O}(\epsilon^2) = 0 \qquad \Rightarrow \quad \operatorname{div}_x \mathbf{u} + \mathrm{O}(\epsilon^2) = 0. \tag{5.23}$$

Likewise from (5.20)

$$\mathbf{F}^{-1} = \mathbf{F}_0^{-1}(\mathbf{I}+\mathbf{H})^{-1} = \mathbf{F}_0^{-1}(\mathbf{I}-\mathbf{H}) + \mathrm{O}(\epsilon^2), \qquad \mathbf{F}^{-T} = (\mathbf{I}-\mathbf{H}^T)\mathbf{F}_0^{-T} + \mathrm{O}(\epsilon^2). \tag{5.24}$$

Turning next to the constitutive relation, we first let $\mathbb{A}(\mathbf{F})$ and $\mathbb{B}(\mathbf{F})$ be the 4-tensor functions of $\mathbf{F}$ defined as the tensors with components

$$\mathbb{A}_{ijk\ell}(\mathbf{F}) := \frac{\partial^2 W(\mathbf{F})}{\partial F_{ij}\partial F_{k\ell}}, \qquad \mathbb{B}_{ipkq}(\mathbf{F}) := \mathbb{A}_{ijk\ell}(\mathbf{F})F_{pj}F_{q\ell}, \tag{5.25}$$

Observe that when evaluated at $\mathbf{F} = (\mathbf{I}+\mathbf{H})\mathbf{F}_0$ and linearized,

$$\frac{\partial W}{\partial F_{ij}}(\mathbf{F})\bigg|_{\mathbf{F}=(\mathbf{I}+\mathbf{H})\mathbf{F_0}} = \frac{\partial W}{\partial F_{ij}}(\mathbf{F}_0) + \mathbb{A}_{ijk\ell}(\mathbf{F}_0)\ \overset{o}{F}_{q\ell}\ H_{kq} + \mathrm{O}(\epsilon^2).$$

In what follows we shall omit the argument $\mathbf{F}_0$ from the partial derivatives of W (including $\mathbb{A}$). Setting $\mathbf{F} = (\mathbf{I}+\mathbf{H})\mathbf{F}_0$ and

$$q = q_0 + \widetilde{q}, \tag{5.26}$$

where $\widetilde{q}$ is assumed to be $\mathrm{O}(\epsilon)$, in the constitutive equation

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}) - q\mathbf{F}^{-T},$$

and linearizing gives

$$
\begin{aligned}
S_{ij} &= \frac{\partial W}{\partial F_{ij}} + \mathbb{A}_{ijk\ell} \overset{o}{F}_{q\ell} H_{kq} - (q_0 + \widetilde{q})(\overset{o}{F}{}^{-1}_{ji} - H_{pi} \overset{o}{F}{}^{-1}_{jp}) = \\
&= S^o_{ij} + \mathbb{A}_{ijk\ell} \overset{o}{F}_{q\ell} H_{kq} - \widetilde{q} \overset{o}{F}{}^{-1}_{ji} + q_0 H_{pi} \overset{o}{F}{}^{-1}_{jp} = \\
&= S^o_{ij} + \left[\mathbb{A}_{isk\ell} \overset{o}{F}_{q\ell} \overset{o}{F}_{ps} H_{kq} - \widetilde{q}\delta_{pi} + q_0 H_{pi}\right] \overset{o}{F}{}^{-1}_{jp} = \\
&= S^o_{ij} + [\mathbb{B}_{ipkq} H_{kq} + q_0 H_{pi} - \widetilde{q}\delta_{pi}] \overset{o}{F}{}^{-1}_{jp} = \\
&= S^o_{ij} + \Sigma_{ip} \overset{o}{F}{}^{-1}_{jp}.
\end{aligned}
\tag{5.27}
$$

Thus

$$\mathbf{S} = \mathbf{S}^o + \boldsymbol{\Sigma}\mathbf{F}_0^{-T},$$

where

$$\boldsymbol{\Sigma} = \mathbb{B}\mathbf{H} + q_0 \mathbf{H}^T - \widetilde{q}\mathbf{I}; \tag{5.28}$$

$\boldsymbol{\Sigma}\mathbf{F}_0^{-T}$ is the perturbation of the 1st Piola-Kirchhoff stress. Equation (5.28) is effectively the constitutive relation for $\boldsymbol{\Sigma}$.

It is easy to show to $O(\epsilon^2)$ that $\text{div}_z\mathbf{S} = \text{div}_z\mathbf{S}_0 + \text{div}_z(\boldsymbol{\Sigma}\mathbf{F}_0^{-T}) = \text{div}_z(\boldsymbol{\Sigma}\mathbf{F}_0^{-T}) = \text{div}_x(\boldsymbol{\Sigma})$, noting that in the last step the divergence is taken with respect to the homogeneously deformed configuration. Thus the equilibrium equation $\text{div}_z\mathbf{S} = \mathbf{0}$ to leading order can be written as [16]

$$\text{div}_x\boldsymbol{\Sigma} = 0, \tag{5.29}$$

for the stress tensor field $\boldsymbol{\Sigma}(\mathbf{x})$.

Thus in summary, the perturbed problem involves the fields $\mathbf{u}(\mathbf{x}), \mathbf{H}(\mathbf{x})$ and $\boldsymbol{\Sigma}(\mathbf{x})$ and they obey the incompressibility equation (5.23), the equilibrium equation (5.29) and the constitutive equation (5.28).

In order to recover what we had before, we now specialize (5.28) for the neo-Hookean material. It is readily found from $(5.25)_1$ and

$$W = \frac{\mu}{2}(\mathbf{F}\cdot\mathbf{F} - 1)$$

that

$$\mathbb{A}_{ijk\ell} = \mu\delta_{j\ell}\delta_{ik},$$

[16] As an exercise show that $\text{div}_z\widetilde{\mathbf{S}} = \text{div}_x\boldsymbol{\Sigma}$ where $\widetilde{\mathbf{S}} = \boldsymbol{\Sigma}\mathbf{F}_0^{-T}$ even if $\mathbf{F}_0$ is not a constant, i.e. if it is a field $\mathbf{F}_0(\mathbf{z})$.

and therefore from $(5.25)_2$ that

$$\mathbb{B}_{ijpq} = \mu B_{jq}\delta_{ip},$$

whence

$$\mathbb{B}_{ijpq}H_{pq} = \mu B^0_{jq}H_{iq} = \mu(\mathbf{H}\mathbf{B}_0)_{ij}.$$

Equation (5.28) for the stress $\boldsymbol{\Sigma}$ now specializes to

$$\boldsymbol{\Sigma} = \mu\mathbf{H}\mathbf{B}_0 + q_0\mathbf{H}^T - \widetilde{q}\mathbf{I} = \mathbf{H}(\mu\mathbf{B}_0 - q_0\mathbf{I}) + q_0(\mathbf{H}+\mathbf{H}^T) - \widetilde{q}\mathbf{I} = \mathbf{H}\mathbf{T}_0 + q_0(\mathbf{H}+\mathbf{H}^T) - \widetilde{q}\mathbf{I}$$

and so we recover (xix).

When these equations are specialized to the plane strain perturbation (ixb), the traction-free boundary condition on $z_2 = 0$ and $\mathbf{S} \to -S\mathbf{e}_1 \otimes \mathbf{e}_1$ as $|\mathbf{z}| \to \infty$ one recovers the equations we had above.

In terms of components in an arbitrary fixed cartesian basis, (5.28) reads

$$\Sigma_{ij} = \mathbb{B}_{ijk\ell}\frac{\partial u_k}{\partial x_\ell} + q_0\frac{\partial u_j}{\partial x_i} - \widetilde{q}\,\delta_{ij}. \tag{5.30}$$

Note that since the deformation (5.11) is homogeneous, the elastic moduli $\mathbb{B}_{ijk\ell}$ and the scalar q_0 are constants. Substituting (5.30) into the equilibrium equations $\partial\Sigma_{ij}/\partial x_j = 0$ and using the incompressibility equation $\partial u_i/\partial x_i = 0$ leads to

$$\mathbb{B}_{ijk\ell}\,\frac{\partial^2 u_k}{\partial x_j \partial x_\ell} - \frac{\partial \widetilde{q}}{\partial x_i} = 0. \tag{5.31}$$

This, together with the incompressibility equation $\partial u_i/\partial x_i = 0$, are to be solved for $u_i(\mathbf{x})$ and $\widetilde{q}(\mathbf{x})$.

Exercises: Problems 5.17, 5.18, 5.19 and 5.20.

5.8 Exercises.

Problem 5.1. Consider a spherical elastic shell that undergoes a large (spherically symmetric) deformation when subjected to an internal pressure p.

(I) First, model the shell as a two-dimensional entity (a membrane). Let r and R denote the radii of the membrane in the deformed and reference configurations respectively; let σ be the circumferential *force* in the membrane per unit deformed (circumferential) *length*; and let $\mathcal{W}$ be the elastic energy in the membrane per unit deformed *area*. The constitutive relation for the energy is $\mathcal{W} = \mathcal{W}(\lambda)$ where $\lambda = r/R$ is the circumferential stretch of the membrane.

(a) Use force balance to show that $p = 2\sigma/r$.

(b) Balance the rate of external working with the rate of increase of stored energy in a quasi-static motion and show that

$$p = \frac{2}{r}\mathcal{W} + \frac{1}{R}\mathcal{W}', \qquad \sigma = \frac{1}{2}\lambda\mathcal{W}' + \mathcal{W}. \tag{i}$$

Remark: You may have seen these results in the context of a soap bubble with surface tension σ and the surface energy $\mathcal{W} = \mathsf{constant}$.

(II) Now model the shell as a hollow spherical solid composed of an incompressible, isotropic material, whose wall-thickness is small but positive. The quantities r and R introduced above now represent the *mean* radii in the deformed and undeformed configurations, and let $T\,(\ll R)$ denote the wall-thickness in the undeformed configuration. Note by symmetry and incompressibility that $\lambda_\theta = \lambda_\phi = \lambda, \lambda_r = \lambda^{-2}$ where $\lambda = r/R$ is the mean circumferential stretch. Let $w(\lambda)$ denote the elastic energy in the shell per unit *reference volume*. Show that $w(\lambda) = \lambda^2\mathcal{W}(\lambda)/T$. Substituting this into $(i)_1$ gives

$$p = \frac{T}{R}\,\frac{w'(\lambda)}{\lambda^2}. \tag{ii}$$

which coincides with equation (xv) on page 352 as it should.

Problem 5.2. *Pressurized hollow circular cylinder.* A thick-walled hollow circular tube has inner and outer radii A and B respectively in the undeformed configuration. It is composed of an isotropic incompressible elastic material. Determine the radii a and b of the tube in the deformed configuration when it is subjected to an internal pressure p on the inner curved surface. Assume that particles do not displace in the axial direction. Calculate the forces that must be applied on the two end faces of the tube in order to prevent axial deformation.

Problem 5.3. *Torsion.* Consider the torsional deformation of an isotropic incompressible solid circular cylindrical body as in Section 5.2 but now assume that no *resultant* axial force is applied on the two rigid plates at its ends. (The axial stress T_{zz} need not be zero everywhere, only its resultant on a cross section must vanish.) Under these conditions, in addition to twisting about the z-axis, the cylinder will change in length (and therefore so will its radius). Thus now consider the deformation

$$r = \widehat{r}(R, \Theta, \Phi) = r(R), \qquad \theta = \widehat{\theta}(R, \Theta, \Phi) = \Theta + \alpha\lambda Z, \qquad z = \widehat{z}(R, \Theta, \Phi) = \lambda Z.$$

Assume the material is neo-Hookean and that in the undeformed configuration the body has radius A. Derive an algebraic equation in which λ is the only unknown. (The parameters μ, α and A are given.)

Problem 5.4. *Combined axial and azimuthal shear of a tube.* In Problems 2.15 and 3.10.1 we considered a hollow elastic tube of inner and outer radii A and B respectively, whose outer surface $R = B$ was held fixed. A rigid solid cylinder of radius A was inserted into the cavity, and firmly bonded to the hollow elastic cylinder on their common interface $R = A$. An axial force F and a torque T were applied on the rigid cylinder.

In Problem 2.15 we examined the kinematics of the resulting shear deformation (with both axial and azimuthal parts)

$$r = R, \qquad \theta = \Theta + \phi(R), \qquad z = Z + w(R). \tag{i}$$

In Problem 3.10.1 we analyzed the equilibrium of the associated axisymmetric stress field.

Suppose the hollow elastic tube is composed of a neo-Hookean material.

(a) Determine the rotation and axial displacement of the solid rigid cylinder.

(b) Calculate the radial stress field $T_{rr}(r)$ to within an unknown constant. Explain how you might find the constant but do not carry out any calculations to find it.

Problem 5.5. *Internal pressure and axial loading of a circular tube.* In a stress-free reference configuration, a hollow circular cylindrical tube has inner and outer radii A and B respectively. It is composed of an isotropic incompressible material. The tube is subjected to an internal pressure p and an axial force N and undergoes a deformation of the axi-symmetric form

$$r = r(R), \qquad \theta = \Theta, \qquad z = \Lambda Z, \tag{i}$$

where Λ is the (unknown) axial stretch of the tube. The pressure p and force N are given. Denote the (unknown) inner and outer radii of the tube in the deformed configuration by a and b.

Derive three equations in which a, b and Λ are the only unknowns. You will find it convenient to let $\lambda = r/R$ and introduce the function

$$w(\lambda, \Lambda) = W^*(\lambda^{-1}\Lambda^{-1}, \lambda, \Lambda), \tag{ii}$$

because the principal stretches are $\lambda^{-1}\Lambda^{-1}, \lambda, \Lambda$.

Problem 5.6. *Instability of a cube.* Reconsider the "Rivlin cube problem" that we considered in Section 5.3 where a unit cube of homogeneous incompressible isotropic elastic material was subjected to normal dead-load forces of magnitude F. Show that for a Mooney-Rivlin material with stored-energy function

$$W(\lambda_1, \lambda_2, \lambda_3) = \alpha(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) + \beta(\lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2} - 3), \tag{i}$$

where $\alpha > 0$ and $\beta > 0$ are constants, there can be solutions of the form

$$y_1 = \lambda_1 x_1, \quad y_2 = \lambda_2 x_2, \quad y_3 = \lambda_3 x_3, \tag{ii}$$

with *all three* λ_i*'s different*, provided α and β are suitably restricted. What is the restriction?

Problem 5.7. *Instability of a thin sheet.* [Experiments of this nature have been carried out by Treloar.] Consider a body that in an unstressed reference configuration is a square sheet $a \times a \times t$. The long edges of the sheet are parallel to the x_1- and x_2-axes. Tensile normal forces F act on the four edges $x_1 = \pm a/2$ and $x_2 = \pm a/2$ of the sheet (through the application of uniform normal traction distributions). The two faces $x_3 = \pm t/2$ are traction-free. The sheet is made of a Mooney-Rivlin material

$$W(\lambda_1, \lambda_2, \lambda_3) = \frac{\mu}{2}\Big[\alpha\big(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3\big) + (1-\alpha)\big(\lambda_1^{-2} + \lambda_2^{-2} + \lambda_3^{-2} - 3\big)\Big], \qquad \mu > 0, \ 0 \le \alpha \le 1. \tag{i}$$

Assume the deformation to be a pure homogeneous stretch

$$\mathbf{y} = \mathbf{F}\mathbf{x}, \qquad \mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 \ + \ \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 \ + \ \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{ii}$$

(a) Show that for all values of the the force F there is a symmetric equilibrium configuration with $\lambda_1 = \lambda_2$.

(b) Show for a neo-Hookean material ($\alpha = 1$) that there are no asymmetric equilibrium configurations with $\lambda_1 \neq \lambda_2$.

(c) Show that for sufficiently large values of F there is an equilibrium configuration with $\lambda_1 \neq \lambda_2$ (when $\alpha \neq 1$). What is the smallest value of F at which this second form of solution becomes possible?

(d) For what range of values of F is the symmetric solution stable?

Problem 5.8. *Stability of the "Rivlin Cube" with respect to arbitrary perturbations.* Reconsider the stability of the "Rivlin cube" studied in Section 5.3. There, we first determined the various pure homogeneous deformations the body could undergo, and second, investigated whether these deformations minimized the potential energy. In this latter calculation, we limited attention to virtual deformations that were homogeneous and coaxial with the pure homogeneous deformations we were studying. In the present problem, you are asked to consider all virtual deformations.

In the "Rivlin cube" problem the unit cube is subjected to the dead loading $\mathbf{s} = \overline{\mathbf{S}}\mathbf{n}_{\rm R}$ on $\partial\mathcal{R}_{\rm R}$ where $\overline{\mathbf{S}}$ is a given constant tensor. The associated deformation whose stability we want to study is

$$\mathbf{y}(\mathbf{x}) = \mathbf{F}\mathbf{x} \qquad \text{for} \quad \mathbf{x} \in \mathcal{R}_{\rm R}, \tag{i}$$

where the constant tensor $\mathbf{F}$ has $\det \mathbf{F} = 1$ and

$$\overline{\mathbf{S}} = \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}) - q\,\mathbf{F}^{-T}. \tag{ii}$$

In order to study the stability of a deformation (i), consider virtual deformations of the form

$$\mathbf{z}(\mathbf{x}) = \mathbf{y}(\mathbf{x}) + \epsilon\boldsymbol{\eta}(\mathbf{x}) = \mathbf{F}\mathbf{x} + \epsilon\boldsymbol{\eta}(\mathbf{x}) \qquad \text{for} \quad \mathbf{x} \in \mathcal{R}_{\rm R}. \tag{iii}$$

Here $\mathbf{z}(\mathbf{x})$ is the virtual deformation, $\mathbf{y}(\mathbf{x})$ is the deformation whose stability we wish the study, and $\epsilon\boldsymbol{\eta}(\mathbf{x})$ is the virtual displacement. The associated virtual deformation gradient tensor is

$$\mathbf{G} = \nabla_x \mathbf{z} = \mathbf{F} + \epsilon\nabla_x\boldsymbol{\eta}. \tag{iv}$$

In (iii), ϵ is a scalar parameter and $\boldsymbol{\eta}(\mathbf{x})$ is an arbitrary smooth function subject only to the incompressibility requirement

$$\det \mathbf{G} = 1. \tag{v}$$

The potential energy associated with a virtual deformation $\mathbf{z}(\mathbf{x})$ is

$$\Phi = \int_{\mathcal{R}_{\rm R}} W(\nabla_x \mathbf{z}) dV_x - \int_{\partial\mathcal{R}_{\rm R}} \overline{\mathbf{S}}\mathbf{n}_{\rm R} \cdot \mathbf{z}\, dA_x.$$

It is convenient to incorporate the kinematic constraint (v) into the potential energy through a Lagrange multiplier q, and to therefore consider

$$\Phi = \int_{\mathcal{R}_{\rm R}} \left(W(\nabla_x \mathbf{z}) - q\det(\nabla_x \mathbf{z})\right) dV_x - \int_{\partial\mathcal{R}_{\rm R}} \overline{\mathbf{S}}\mathbf{n}_{\rm R} \cdot \mathbf{z}\, dA_x. \tag{vi}$$

On evaluating (vi) at a virtual deformation (iii), and keeping $\boldsymbol{\eta}(\mathbf{x})$ fixed for the moment, the potential energy takes the form

$$\Phi = \Phi(\epsilon). \tag{vii}$$

which is a function of the single scalar ϵ. Since $\mathbf{z}(\mathbf{x}) = \mathbf{y}(\mathbf{x})$ when $\epsilon = 0$, see (iii), it follows that if $\mathbf{y}(\mathbf{x})$ is a minimizer of the potential energy then $\epsilon = 0$ is a minimizer of $\Phi(\epsilon)$. This requires

$$\left.\frac{\mathrm{d}\Phi}{d\epsilon}\right|_{\epsilon=0} = 0, \qquad \left.\frac{\mathrm{d}^2\Phi}{d\epsilon^2}\right|_{\epsilon=0} \geq 0. \tag{$viii$}$$

It will be convenient in what follows to let

$$\mathbf{H} \coloneqq \nabla_x\boldsymbol{\eta}\,\mathbf{F}^{-1} = \nabla_y\boldsymbol{\eta}. \tag{ix}$$

(a) Show that

$$\det \mathbf{G} = 1 + \operatorname{tr}\mathbf{H} + \mathrm{O}(\epsilon^2) \qquad \text{as} \quad \epsilon \to 0, \tag{x}$$

so that the incompressibility requirement (v) tells us that $\operatorname{tr}\mathbf{H} = 0 + \mathrm{O}(\epsilon)$ as $\epsilon \to 0$.

(b) Evaluate $d\Phi/d\epsilon$ and show that, in view of (ii), the first requirement $(viii)_1$ holds automatically.

(c) Show that

$$\left.\frac{d^2\Phi}{d\epsilon^2}\right|_{\epsilon=0} = \int_{\mathcal{R}_{\rm R}} \left[\frac{\partial^2 W}{\partial F_{ij}\partial F_{k\ell}}(\mathbf{F})\eta_{i,j}\eta_{k,\ell} - q(H_{ii}H_{jj} - H_{ij}H_{ji})\right] dV_x. \qquad (xi)$$

(d) Next consider a neo-Hookean material:

$$W = \frac{\mu}{2}(\mathbf{F}\cdot\mathbf{F} - 1), \qquad (xii)$$

and show that (xi) now specializes to

$$\left.\frac{d^2\Phi}{d\epsilon^2}\right|_{\epsilon=0} = \int_{\mathcal{R}_{\rm R}} \left[\mu B_{kj}H_{ij}H_{ik} + qH_{ij}H_{ji}\right] dV_x, \qquad (xiii)$$

where $\mathbf{B} = \mathbf{F}\mathbf{F}^T$.

(e) Now consider the stability of the cubic solution $\mathbf{F} = \mathbf{I}$. Recall that the loading is in fact an equi-triaxial dead loading, i.e. $\overline{\mathbf{S}} = S\,\mathbf{I}$. In this case (io) gives $q = \mu - S$. Show that

$$\left.\frac{d^2\Phi}{d\epsilon^2}\right|_{\epsilon=0} = \int_{\mathcal{R}_{\rm R}} \Big[(2\mu - S)\varepsilon_{i,j}\varepsilon_{i,j} + S\omega_{i,j}\omega_{i,j})\Big] dV_x, \qquad (xiv)$$

where we have set $\varepsilon_{ij} := \frac{1}{2}(\eta_{i,j} + \eta_{j,i})$ and $\omega_{ij} := \frac{1}{2}(\eta_{i,j} - \eta_{j,i})$.

Thus far we kept $\boldsymbol{\eta}(\mathbf{x})$ fixed. But in fact it is arbitrary, subject only to the requirement stemming from incompressibility. Thus, for stability, it is necessary that the expression in the previous equation be non-negative for all such $\eta_{i,j}$. Show from this that the cubic deformation is stable for $0 < S < 2\mu$. And unstable for $S > 2\mu$ and $S < 0$. What is the nature of a virtual deformation that makes the cubic configuration unstable in the case $S < 0$?

References:

– R, Hill, On uniqueness and stability in the theory of finite elastic strain, *Journal of the Mechanics and Physics of Solids*, 5 (1957), pp. 229–241.
– R.S. Rivlin, Stability of pure homogeneous deformations of an elastic cube under dead loading, *Quarterly Journal of Applied Mathematics*, 1974, pp. 265-271.

Problem 5.9. *Stability of the "Rivlin cube" problem for an arbitrary isotropic material.* In Section 5.3 we examined the stability of a neo-Hookean cube subjected to an equi-triaxial dead loading (1st Piola-Kirchhoff traction). Generalize that analysis to a cube composed of an arbitrary isotropic elastic material by extremizing

$$\Phi(\mathbf{F}) = W(\mathbf{F}) - \mathbf{S}\cdot\mathbf{F} \qquad (i)$$

over all geometrically admissible homogeneous deformations. Assume the dead loading to be

$$\mathbf{S} = \sum_{i=1}^{3} S_i\mathbf{e}_i, \qquad (ii)$$

and consider only deformations of the form $\mathbf{y} = \mathbf{F}\mathbf{x}$ where

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \mathbf{e}_i \otimes \mathbf{e}_i, \tag{iii}$$

i.e. where $\mathbf{F}$ and $\mathbf{S}$ are coaxial.

How would your analysis change if the material is incompressible?

Problem 5.10. *Cavitation.* Derive a formula for the cavitation stress for a general incompressible isotropic elastic material in plane strain by considering the growth of the cylindrical cavity in a tube of undeformed inner and outer radii A and B. Under what conditions on $w(\lambda) = W(\lambda^{-1}, \lambda, 1)$ is the critical stress for cavitation finite?

Problem 5.11. *Pressurized spherical shell with radial inextensibility constraint.* A hollow spherical shell has inner and outer radii A and B respectively in the undeformed configuration. It is composed of an isotropic elastic material. Very stiff fibers oriented in the radial direction have been embedded throughout the body. The shell is subjected to an internal pressure p on the inner curved surface, the outer surface being traction-free. Derive two algebraic equations in which the only unknowns are the radii a and b in the deformed configuration (which one could in principle solve for a and b.)

Specialize your solution to the case when the strain energy function is

$$W = \frac{\mu}{2}(I_1 - 3). \tag{i}$$

Though this strain energy function has the form of the neo-Hookean model, do not assume the material to be incompressible. (Why?)

Problem 5.12. *Steadily rotating cylinder.* An incompressible solid circular cylinder has radius A and length L in a reference configuration. Its mass density is ρ. The cylinder undergoes a steady rotation about its axis of symmetry at the constant angular speed ω. Assume the motion to be described by

$$r = r(R), \qquad \theta = \Theta + \omega t, \qquad z = \lambda Z. \tag{i}$$

Calculate the acceleration of a particle by differentiating $\mathbf{y}(\mathbf{x}, t)$ with respect to time t at a fixed particle $\mathbf{x}$ (i.e. by differentiating $\mathbf{y}(R, \Theta, Z, t) = r\,\mathbf{e}_r(\theta) + z\mathbf{e}_z$ at fixed R, Θ, Z).

Use incompressibility to determine $r(R)$.

Suppose that the curved boundary of the cylinder is traction-free, and the resultant force on its two ends are zero. Moreover, suppose the cylinder is composed of a generalized neo-Hookean material. Derive an algebraic equation relating the axial stretch λ to the angular speed ω. Specialize your answer to a neo-Hookean material. *Note*: Since inertial effects are being taken into account, you must use the equations of motion (i.e. the equilibrium equations (3.94) with the term $\rho\mathbf{a}$ added to the right-hand side where $\mathbf{a} = \ddot{\mathbf{y}}$ is the acceleration). Neglect the body force due to gravity.

Problem 5.13. *Eversion of a hollow sphere.* In a stress-free reference configuration a body occupies a spherical shell of inner radius $R_1 > 0$ and outer radius $R_2 > R_1$. It is composed of an incompressible isotropic elastic material with strain energy function $W(I_1, I_2)$. The body is everted – turned inside out – (say by cutting the body in half, everting each part, and then gluing the two halves together). In the deformed configuration it occupies a spherical shell of inner radius $r_2 > 0$ and outer radius $r_1 > r_2$. *The outer surface $R = R_2$ of the undeformed shell maps into the inner surface $r = r_2$ of the deformed shell and the inner surface $R = R_1$ of the undeformed shell maps into the outer surface $r = r_1$ of the deformed shell.* Here R and r are the radial spherical polar coordinates in the undeformed and deformed configurations. The inner and outer surfaces of the deformed sphere, $r = r_2$ and $r = r_1$ are traction-free. Determine r_1 and r_2.

J. L. Ericksen, Inversion of a perfectly elastic spherical shell, *Zeitschrift für Angewandte Mathematik und Mechanik*, 35(1955), issue 9-10, pp. 382-385.

Problem 5.14. *Eversion of a cylinder.* In the reference configuration a body occupies the region $A \le R \le B$, $0 \le \Theta \le 2\pi$, $0 < Z < L$. Consider a deformation in which the body is turned inside out (everted). (Imagine turning a sock inside out.) In the deformed configuration it occupies the region $a \le r \le b$, $0 \le \theta \le 2\pi$, $-\ell \le z \le 0$. The position of a particle in the reference and deformed configurations are

$$\mathbf{x} = R\,\mathbf{e}_R \;+\; Z\,\mathbf{e}_Z, \qquad \mathbf{x} = r\,\mathbf{e}_r \;+\; z\,\mathbf{e}_z.$$

Assume the deformation describing the eversion to have the form

$$r = \widehat{r}(R, \Theta, Z) = r(R), \qquad \theta = \widehat{\theta}(R, \Theta, Z) = \Theta, \qquad z = -\Lambda Z,$$

with

$$r(A) = b, \qquad r(B) = a,$$

$$T_{rr}(a) = T_{rr}(b) = 0.$$

$$2\pi \int_a^b T_{zz}(r, z)\, dr = 0 \qquad \text{for } z = 0, -\ell$$

The material is incompressible.

(a) Determine $r(R)$.

(b) Assuming the body to be composed of a neo-Hookean material, calculate the components of Cauchy stress (in terms of $p(r)$.

(c) Can the body be maintained in equilibrium under the assumed loading?

Problem 5.15. *Harmonic material.* (Goriely et al.)

(a) Consider the cylindrically symmetric deformation of a hollow circular tube that has inner and outer radii A and B in a stress-free reference configuration. It is composed of an arbitrary unconstrained isotropic elastic material. The deformation is described by $r = r(R), \theta = \Theta, z = Z$ where (R, Θ, Z) and (r, θ, z) are the cylindrical polar coordinates of a particle in the reference and deformed configurations respectively. Show that $r(R)$ satisfies the ordinary differential equation

$$\frac{d}{dR}\left(R\frac{\partial W}{\partial \lambda_1}\right) - \frac{\partial W}{\partial \lambda_2} = 0 \qquad \text{for} \quad A \le R \le B, \qquad (i)$$

where $W(\lambda_1, \lambda_2) \coloneqq W^*(\lambda_1, \lambda_2, 1)$, $\lambda_1 = r'(R)$ and $\lambda_2 = r(R)/R$.

(b) The so-called Harmonic strain energy function is a model for a homogeneous, unconstrained, isotropic elastic material. It is given by

$$W^*(\lambda_1, \lambda_2, \lambda_3) = F(j_1) + \xi(j_2 - 3) + \eta(j_3 - 1), \qquad (ii)$$

where ξ and η are material constants and $F(j_1)$ is a constitutive function. Here

$$j_1 = \lambda_1 + \lambda_2 + \lambda_3, \qquad j_2 = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1, \qquad j_3 = \lambda_1\lambda_2\lambda_3. \qquad (iii)$$

Determine the restrictions on F, ξ and η needed to ensure that the strain energy and stress vanish in the reference configuration. Determine also the restrictions imposed by the Baker-Ericksen inequality.

(c) Show that in plane strain, this model can be reduced to

$$W(\lambda_1, \lambda_2) = f(i_1) - \alpha(i_2 - 1), \qquad (iv)$$

where α is a material constant, $f(i_1)$ is a constitutive function, and

$$i_1 = \lambda_1 + \lambda_2, \qquad i_2 = \lambda_1\lambda_2. \qquad (v)$$

Determine (by specializing your answers to part (b)) the restrictions on f and α needed to ensure that the strain energy and stress vanish in the reference configuration and that the Baker-Ericksen inequalities hold.

(d) Solve equation (i) for the Harmonic material.

(e) Suppose that the outer radius B is infinite and that $T_{rr}(R) \to 0$ as $R \to \infty$. Moreover, let $T_{rr}(A) = -p < 0$. Determine $T_{\theta\theta}(R)$. At which point R in the body, and at what values of pressure p, does $T_{\theta\theta}(R)$ become unbounded.

Problem 5.16. (Goriely et al.) The strain energy function for the simplified isotropic Fung model for soft tissue is given in (4.142) where $\mu > 0$ and $\beta > 0$ are material constants.

Consider a thin-walled hollow sphere composed of this material. The "limit-point instability" refers to the loss of monotonicity in the function of pressure p as a function of the mean circumferential stretch λ.

Determine the critical value β_{cr} above which the limit point instability disappears. Plot four curves of p versus λ corresponding to the four choices (a) $\beta = 0$, (b) $\beta = \beta_{cr}/2$, (c) $\beta = \beta_{cr}$ and (d) $\beta = 1.5\beta_{cr}$.

Determine a realistic value of β for soft tissue from the literature, and reach a conclusion about the existence of this instability in such a material (proposed as a model for aneurysm rupture).

For analyses of the limit point instability for other constitutive models, including less simplified Fung models, see Chapter 8 of *Cardiovascular Solid Mechanics* by Jay Humphrey, Springer, 2002.

Problem 5.17. *Surface instability.* Reconsider the surface instability of a half-space as in see Section 5.7 but now consider an arbitrary isotropic incompressible material. Determine conditions for the onset of a surface instability. Specialize your results to a Gent material.

Problem 5.18. *Surface instability.* Reconsider the surface instability of a half-space, see Section 5.7. Using the same basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ as there, now consider the following cases:

(a) Suppose that the homogeneous configuration involves an *equibiaxial stretch*, i.e.

$$\mathbf{F}_0 = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \text{with } \lambda_1 = \lambda_3.$$

(b) Suppose the homogeneous configuration involves a uniaxial stress, i.e.

$$\mathbf{F}_0 = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3, \qquad \mathbf{T}_0 = T \mathbf{e}_1 \otimes \mathbf{e}_1,$$

where $T_{22} = T_{33} = 0$.

Assume the material to be neo-Hookean. Determine the conditions (if any) under which a surface instability occurs. Consider only plane strain perturbations where the displacement from the homogeneously deformed configuration has the form $\mathbf{u}(\mathbf{x}) = u_1(x_1, x_2)\mathbf{e}_1 + u_2(x_1, x_2)\mathbf{e}_2$.

Problem 5.19. *Surface instability with surface tension.* Recall that the "Biot Problem" – the instability of a half-space under compression – did not involve a length-scale and so the critical value of stretch at instability was independent of the wave number of the sinusoidal surface undulation. Thus the analysis did not pick a particular wave number of the undulation at instability. In this problem you are to endow the free-surface with (constant) surface tension γ (force per unit length). The ratio $\ell_0 := \gamma/\mu$ then has the dimension of length. You are to determine the critical stretch λ_1 at instability. (Limit attention plane strain deformations and a neo-Hookean solid.)

In the simplest model of *surface tension*, the traction-free boundary condition $\mathbf{Tn} = \mathbf{o}$ at a free surface is replaced by

$$\mathbf{Tn} = -\gamma\kappa\mathbf{n}, \tag{i}$$

where κ is the (mean) curvature and $\mathbf{n}$ the unit outward normal vector, both associated with the deformed surface.

Problem 5.20. *Small deformation superposed on an arbitrary finite deformation.* A body is composed of an arbitrary unconstrained elastic material. It occupies a region $\mathcal{R}_{\mathrm{R}}$ in a homogeneous reference configuration and a region $\mathcal{R}$ in deformed configuration-1. A particle $\mathbf{z} \in \mathcal{R}_{\mathrm{R}}$ is taken to $\mathbf{x}_0(\mathbf{z}) \in \mathcal{R}$ by the equilibrium deformation-1:

$$\mathbf{x} = \mathbf{x}_0(\mathbf{z}).$$

The associated deformation gradient tensor is

$$\mathbf{F}_0 = \nabla_z \mathbf{x}_0(\mathbf{z}).$$

The associated Cauchy stress tensor field $\mathbf{T}_0(\mathbf{x})$ obeys the equilibrium equation

$$\mathrm{div}_x\, \mathbf{T}_0 = \mathbf{o}.$$

Deformation-2 takes $\mathbf{z} \to \widehat{\mathbf{y}}(\mathbf{z})$ where

$$\mathbf{y} = \widehat{\mathbf{y}}(\mathbf{z}) = \mathbf{x}_0(\mathbf{z}) + \widehat{\mathbf{u}}(\mathbf{z}),$$

with associated deformation gradient tensor

$$\mathbf{F} = \nabla_z \widehat{\mathbf{y}}(\mathbf{z}) = \mathbf{F}_0 + \nabla_z \widehat{\mathbf{u}}(\mathbf{z}).$$

Introduce the following representation of the displacement field (from configuration-1 to configuration-2)

$$\mathbf{u}(\mathbf{x}) := \widehat{\mathbf{u}}(\mathbf{z}) \qquad \text{where} \qquad \mathbf{x} = \mathbf{x}_0(\mathbf{z}),$$

with associated gradient

$$\mathbf{H} = \nabla_x \mathbf{u}(\mathbf{x}).$$

Let deformation-2 be close to deformation-1 in the sense that

$$\varepsilon := |\mathbf{H}| \ll 1.$$

Let $\mathbf{T}$ be the Cauchy stress tensor associated with deformation-2 and it obeys the equilibrium equation

$$\mathrm{div}_y \mathbf{T} = \mathbf{o}.$$

Show that

(a) $\mathbf{T} = \mathbf{T}_0 + \widetilde{\mathbf{T}} + \mathrm{O}(\varepsilon^2)$ where

$$\widetilde{\mathbf{T}} = -\mathrm{tr}\,\mathbf{H}\ \mathbf{T}_0 + \mathbf{H}\mathbf{T}_0 + \mathbf{T}_0\mathbf{H}^T + \mathbb{C}\mathbf{E},$$

$$\mathbb{C} = \frac{4}{J_0}\mathbb{F}\mathbb{A}\mathbb{F}^T, \qquad \mathbb{F} := \mathbf{F}_0 \boxtimes \mathbf{F}_0, \qquad \mathbb{A}_{ijk\ell} := \frac{\partial^2 W(\mathbf{C}_0)}{\partial C_{ij}\partial C_{k\ell}}.$$

See Problem 1.60 for the definition and properties of the tensor product of two 2-tensors.

(b) $\mathrm{div}_x \widetilde{\mathbf{T}}(\mathbf{x}) = \mathbf{o}$.

(c) the 4-tensor $\mathbb{C}$ has the first and second minor symmetries and the major symmetry. See Problem 1.60 for the definition and some results concerning the minor and major symmetries of a 4-tensor.

Problem 5.21. *Universal deformation.* A deformation that can be maintained in equilibrium by the application of surface tractions only, and no body forces, *for arbitrary W* is said to be a *universal deformation*. Ericksen showed that for an unconstrained material the most general universal deformation is a homogenous one. For incompressible materials however there are certain additional universal deformations (essentially because of the presence of the reaction pressure field $q(\mathbf{x})$). This problem is concerned with one of them.

When studying the inflation, extension and twisting of a tube of an incompressible *isotropic* material we encountered the following deformation:

$$r(R) = \sqrt{c + R^2/\Lambda}, \qquad \theta = \Theta + \alpha\Lambda Z, \qquad z = \Lambda Z,$$

where c, α and Λ are constants. Here (R, Θ, Z) and (r, θ, z) are cylindrical polar coordinates in the reference and deformed configurations respectively. Show that *any* incompressible elastic body in equilibrium can sustain this deformation purely by applying suitable tractions on its boundaries (with no body forces)[17] . This, therefore, is a universal deformation for an incompressible elastic material. In view of this, this is also a possible deformation field for an incompressible *anisotropic* material.

References:

1. A. Goriely, A. Erlich and C. Goodbrake, C5.1 Solid Mechanics: Online problem sheets, `https://courses.maths.ox.ac.uk/node/36846/materials`, Oxford University, 2018.

2. R.W. Ogden, Chapter 5 of *Non-Linear Elastic Deformations*, Chapter 3, Dover, 1997.

3. D. J. Steigmann, Chapters 7 and 8 of Finite Elasticity Theory, Oxford, 2017.

[17] Of course the tractions that have to be applied will depend on the material.

Chapter 6

Anisotropic Elastic Solids.

Our treatment of anisotropy in this chapter is limited to that arising from the presence of either one or two preferred directions, the former corresponding to transverse isotropy. For a more complete treatment of anisotropy, see, for example, Spencer [7, 8].

6.1 One family of fibers. Transversely isotropic material.

Consider a material with one preferred direction $\mathbf{m}_{\mathrm{R}}$ in the reference configuration. For example this may be due to the presence of a family of fibers in that direction. For simplicity, we will sometimes refer to a preferred direction as a "fiber direction" (even if there are no fibers). In general, the fibers will *not* be inextensible.

The constitutive response functions for the first Piola-Kirchhoff stress and strain energy function for such a material will depend on both the deformation gradient tensor $\mathbf{F}$ and the fiber direction $\mathbf{m}_{\mathrm{R}}$:

$$\mathbf{S} = \widehat{\mathbf{S}}(\mathbf{F}, \mathbf{m}_{\mathrm{R}}), \qquad W = \widehat{W}(\mathbf{F}, \mathbf{m}_{\mathrm{R}}).$$

The power-energy balance $\mathbf{S} \cdot \dot{\mathbf{F}} = \dot{W}$ implies, as before, that

$$\widehat{\mathbf{S}}(\mathbf{F}, \mathbf{m}_{\mathrm{R}}) = \frac{\partial \widehat{W}}{\partial \mathbf{F}}(\mathbf{F}, \mathbf{m}_{\mathrm{R}}).$$

Material frame indifference requires that

$$\widehat{W}(\mathbf{F}, \mathbf{m}_{\mathrm{R}}) = \widehat{W}(\mathbf{QF}, \mathbf{m}_{\mathrm{R}}) \qquad \text{for all orthogonal } \mathbf{Q},$$

keeping in mind that $\mathbf{m}_{\rm R}$ is a direction in the reference configuration and so $\mathbf{Q}$ does not transform it. By the same argument as in Chapter 4, this holds if and only if the strain energy function depends on the deformation through the right Cauchy-Green tensor. Thus, there is a function $\overline{W}$ such that

$$\widehat{W}(\mathbf{F}, \mathbf{m}_{\rm R}) = \overline{W}(\mathbf{C}, \mathbf{m}_{\rm R}), \qquad \mathbf{C} = \mathbf{F}^T\mathbf{F},$$

and the Cauchy stress is related to the deformation through

$$\mathbf{T} = \frac{2}{J}\mathbf{F}\frac{\partial \overline{W}}{\partial \mathbf{C}}\mathbf{F}^T. \tag{6.1}$$

Up to this point the anisotropy of the material has not had any (substantive) effect.

Now consider material symmetry. In the particular reference configuration at hand, the material with the single preferred direction $\mathbf{m}_{\rm R}$ is *transversely isotropic* with respect to that direction in the sense that the strain energy function is invariant under all rotations (of the reference configuration) about $\mathbf{m}_{\rm R}$ and under reflection in the plane perpendicular to $\mathbf{m}_{\rm R}$. (The latter implies it is invariant to replacing $\mathbf{m}_{\rm R}$ by $-\mathbf{m}_{\rm R}$). Thus the material symmetry group for such a material is

$$\mathcal{G} = \{\mathbf{Q} \ : \ \mathbf{Q}\mathbf{Q}^T = \mathbf{I}, \quad \mathbf{Q}\mathbf{m}_{\rm R} = \pm\mathbf{m}_{\rm R}\}, \tag{vi}$$

see Problem 1.10(b), and material symmetry tells us that

$$\overline{W}(\mathbf{C}, \mathbf{m}_{\rm R}) = \overline{W}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T, \mathbf{Q}\mathbf{m}_{\rm R}) \qquad \text{for all } \mathbf{Q} \in \mathcal{G}, \tag{vii}$$

and all symmetric positive definite tensors $\mathbf{C}$. Note that in this step, $\mathbf{Q}$ acts on the reference configuration and so it does transform $\mathbf{m}_{\rm R}$. However in view of (vi) we can write this equivalently as

$$\overline{W}(\mathbf{C}, \mathbf{m}_{\rm R}) = \overline{W}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T, \pm\mathbf{m}_{\rm R}) \qquad \text{for all } \mathbf{Q} \in \mathcal{G}. \tag{$viii$}$$

It is shown in Problem 6.1 that the following group of orthogonal tensors,

$$\mathcal{G}' = \{\mathbf{Q} \ : \ \mathbf{Q}\mathbf{Q}^T = \mathbf{I}, \quad \mathbf{Q}\mathbf{M}\mathbf{Q}^T = \mathbf{M}\} \qquad \text{where} \quad \mathbf{M} := \mathbf{m}_{\rm R} \otimes \mathbf{m}_{\rm R}, \tag{6.2}$$

is identical[1] to the group $\mathcal{G}$ defined by (vi). Thus we can replace $\mathcal{G}$ by $\mathcal{G}'$ in $(viii)$. It is then shown in Problem 6.2 that the strain energy function $\overline{W}(\mathbf{C}, \mathbf{m}_{\rm R})$ obeys the invariance $(viii)$ over the set $\mathcal{G}'$ if and only if the function $\check{W}(\mathbf{C}, \mathbf{M})$ obeys the invariance

$$\check{W}(\mathbf{C}, \mathbf{M}) = \check{W}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T, \mathbf{Q}\mathbf{M}\mathbf{Q}^T) \qquad \text{for all orthogonal tensors } \mathbf{Q}, \tag{6.3}$$

[1] When $\mathbf{m}_R \to \mathbf{Q}\mathbf{m}_R$ note that $\mathbf{M} \to (\mathbf{Q}\mathbf{m}_R) \otimes (\mathbf{Q}\mathbf{m}_R) = \mathbf{Q}(\mathbf{m}_R \otimes \mathbf{m}_R)\mathbf{Q}^T = \mathbf{Q}\mathbf{M}\mathbf{Q}^T$.

where

$$\overline{W}(\mathbf{C}, \mathbf{m}_{\rm R}) = \breve{W}(\mathbf{C}, \mathbf{M}), \qquad \mathbf{M} := \mathbf{m}_{\rm R} \otimes \mathbf{m}_{\rm R}.$$

Note that (6.3) holds for *all* orthogonal $\mathbf{Q}$ not just those in $\mathcal{G}'$. Therefore the function $\breve{W}$ is jointly isotropic in both arguments.

The tensor $\mathbf{M} := \mathbf{m}_{\rm R} \otimes \mathbf{m}_{\rm R}$ is referred to as the **structural tensor** (for transverse isotropy). It characterizes the "internal structure" of the material in the reference configuration.

Finally it is claimed in Problem 6.3 that a strain energy function that obeys (6.3) can be expressed as

$$\overline{W}(\mathbf{C}, \mathbf{M}) = \widetilde{W}(I_1, I_2, I_3, I_4, I_5), \tag{6.4}$$

where

$$\begin{gathered} I_1(\mathbf{C}) = \mathrm{tr}\,\mathbf{C}, \qquad I_2(\mathbf{C}) = \tfrac{1}{2}\Big[(\mathrm{tr}\,\mathbf{C})^2 - \mathrm{tr}\,\mathbf{C}^2\Big], \qquad I_3(\mathbf{C}) = \det\,\mathbf{C}, \\ I_4(\mathbf{C}, \mathbf{M}) = \mathbf{C} \cdot \mathbf{M}, \qquad I_5(\mathbf{C}, \mathbf{M}) = \mathbf{C}^2 \cdot \mathbf{M}; \end{gathered} \tag{6.5}$$

see Ericksen and Rivlin [1]. We would expect this list to also include $I_1(\mathbf{M}), I_2(\mathbf{M})$ and $I_3(\mathbf{M})$, but recall from Problem 1.18 that $I_1(\mathbf{M}) = 1$ and $I_2(\mathbf{M}) = I_3(\mathbf{M}) = 0$. Moreover it does not include $\mathbf{C} \cdot \mathbf{M}^2$ since it is readily seen that $\mathbf{C} \cdot \mathbf{M}^2 = \mathbf{C} \cdot \mathbf{M}$; this follows because $\mathbf{M}^n = \mathbf{M}$ for any positive integer n. It is sometimes convenient to express the invariants in terms of $\mathbf{m}_{\rm R}$ by substituting $\mathbf{M} = \mathbf{m}_{\rm R} \otimes \mathbf{m}_{\rm R}$ into (6.5). This yields

$$\begin{gathered} I_1 = \mathrm{tr}\,\mathbf{C}, \qquad I_2 = \tfrac{1}{2}\Big[(\mathrm{tr}\,\mathbf{C})^2 - \mathrm{tr}\,\mathbf{C}^2\Big], \qquad I_3 = \det\,\mathbf{C}, \\ I_4 = \mathbf{C}\mathbf{m}_{\rm R} \cdot \mathbf{m}_{\rm R}, \qquad I_5 = \mathbf{C}^2\mathbf{m}_{\rm R} \cdot \mathbf{m}_{\rm R}. \end{gathered} \tag{6.6}$$

Strictly, the scalar-valued functions I_4 and I_5 are not invariants in the usual sense of invariance (i.e. invariance over the set of *all* orthogonal tensors) though that term is often used. Other authors refer to them as *pseudo-invariants*. Observe that

$$I_4 = \mathbf{C}\mathbf{m}_{\rm R} \cdot \mathbf{m}_{\rm R} = \mathbf{F}^T\mathbf{F}\mathbf{m}_{\rm R} \cdot \mathbf{m}_{\rm R} = \mathbf{F}\mathbf{m}_{\rm R} \cdot \mathbf{F}\mathbf{m}_{\rm R} = |\mathbf{F}\mathbf{m}_{\rm R}|^2, \tag{6.7}$$

and so I_4 denotes the (square of the) stretch in the fiber direction $\mathbf{m}_{\rm R}$. The fiber direction in the deformed configuration is

$$\mathbf{m} = \frac{\mathbf{F}\mathbf{m}_{\rm R}}{|\mathbf{F}\mathbf{m}_{\rm R}|} \overset{(6.7)}{=} \frac{\mathbf{F}\mathbf{m}_{\rm R}}{\sqrt{I_4}}. \tag{6.8}$$

Observe that:

$$I_5 = \mathbf{C}\mathbf{m}_{\rm R} \cdot \mathbf{C}\mathbf{m}_{\rm R} = \mathbf{F}^T\mathbf{F}\mathbf{m}_{\rm R} \cdot \mathbf{F}^T\mathbf{F}\mathbf{m}_{\rm R} \overset{(6.8)}{=} I_4\, \mathbf{F}^T\mathbf{m} \cdot \mathbf{F}^T\mathbf{m} = I_4\, \mathbf{F}\mathbf{F}^T\mathbf{m} \cdot \mathbf{m} = I_4\, \mathbf{B}\mathbf{m} \cdot \mathbf{m}$$

where $\mathbf{m}$ is the fiber direction in the deformed configuration.

From (6.1), (6.4) and the chain rule, the constitutive relation for the Cauchy stress tensor can be written as

$$\mathbf{T} = \frac{2}{J}\mathbf{F}\left[\sum_{i=1}^{5} \frac{\partial \widetilde{W}}{\partial I_i}\frac{\partial I_i}{\partial \mathbf{C}}\right]\mathbf{F}^T. \tag{6.9}$$

The terms $\partial I_i/\partial \mathbf{C}$ for $i = 1, 2, 3$ were calculated previously when we considered isotropic materials, while it is readily shown from $(6.5)_{4,5}$ and $(6.6)_{4,5}$ that

$$\frac{\partial I_4}{\partial \mathbf{C}} = \mathbf{M} = \mathbf{m}_{\mathrm{R}} \otimes \mathbf{m}_{\mathrm{R}}, \qquad \frac{\partial I_5}{\partial \mathbf{C}} = \mathbf{CM} + \mathbf{MC} = \mathbf{Cm}_{\mathrm{R}} \otimes \mathbf{m}_{\mathrm{R}} + \mathbf{m}_{\mathrm{R}} \otimes \mathbf{Cm}_{\mathrm{R}}. \tag{6.10}$$

Observe that the two tensors $\partial I_4/\partial \mathbf{C}$ and $\partial I_5/\partial \mathbf{C}$ are symmetric (as they must be since $I_4(\cdot, \mathbf{M})$ and $I_5(\cdot, \mathbf{M})$ are defined on the set of symmetric tensors). On substituting the expressions for $\partial I_i/\partial \mathbf{C}$ into (6.9) we get the following explicit form for the constitutive relation for $\mathbf{T}$:

$$\begin{aligned}\mathbf{T} \ &= 2JW_3\,\mathbf{I} \ + \ \frac{2}{J}\left[W_1 + I_1 W_2\right]\mathbf{B} \ - \ \frac{2}{J}W_2\mathbf{B}^2 + \\ &+\frac{2}{J}W_4(\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}}) \ + \ \frac{2}{J}W_5\left[(\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{BFm}_{\mathrm{R}}) + (\mathbf{BFm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}})\right],\end{aligned} \tag{6.11}$$

where $\mathbf{B} = \mathbf{FF}^T$ and we have used the notation

$$W_i = \frac{\partial \widetilde{W}}{\partial I_i}, \qquad i = 1, 2, \ldots, 5. \tag{6.12}$$

The Cauchy stress tensor given by (6.11) is automatically symmetric (as it must be since it stemmed from (6.9) that yields a symmetric stress tensor). Observe also that the principal directions of $\mathbf{T}$ and $\mathbf{B}$ *no longer coincide* in general. One can use (6.8) to express the constitutive relation in terms of the direction $\mathbf{m}$ of the (stretched) fiber in the deformed configuration.

If the material is *incompressible*, then

$$I_3 = J^2 = \det \mathbf{F} = 1$$

and the strain energy function has the form $W = \widetilde{W}(I_1, I_2, \cancel{I_3}, I_4, I_5)$. We must drop the term involving W_3 from the constitutive relation (6.11) and replace it with the reaction stress associated with the incompressibility constraint (i.e. a pressure $-q\,\mathbf{I}$). This leads to

$$\begin{aligned}\mathbf{T} \ &= -q\mathbf{I} + 2W_1\mathbf{B} + 2W_2(I_1\mathbf{B} - \mathbf{B}^2) + 2W_4(\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}}) + \\ &+2W_5\left[(\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{BFm}_{\mathrm{R}}) + (\mathbf{BFm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}})\right],\end{aligned} \tag{6.13}$$

having set $J = 1$ throughout.

If the fibers are *inextensible*, then

$$I_4 = |\mathbf{F}\mathbf{m}_{\mathrm{R}}|^2 = 1$$

and the strain energy function has the form $W = \widetilde{W}(I_1, I_2, I_3, \cancel{I_4}, I_5)$. Now we must drop the term involving W_4 from the constitutive relation (6.11) and replace it with the reaction stress associated with the inextensibility constraint, i.e. a uniaxial stress $q\,\mathbf{m}\otimes\mathbf{m}$ in the direction $\mathbf{m}$ – the fiber direction in the deformed configuration (see Problem 4.19 (a)). This leads to

$$\begin{aligned}\mathbf{T} &= 2JW_3\,\mathbf{I} + \frac{2}{J}\left[W_1 + I_1 W_2\right]\mathbf{B} - \frac{2}{J}W_2\mathbf{B}^2 + \\ &+ q\,\mathbf{m}\otimes\mathbf{m} + \frac{2}{J}W_5\Big[(\mathbf{F}\mathbf{m}_{\mathrm{R}}\otimes\mathbf{B}\mathbf{F}\mathbf{m}_{\mathrm{R}}) + (\mathbf{B}\mathbf{F}\mathbf{m}_{\mathrm{R}}\otimes\mathbf{F}\mathbf{m}_{\mathrm{R}})\Big]\,,\end{aligned} \tag{6.14}$$

where $\mathbf{m} = \mathbf{F}\mathbf{m}_{\mathrm{R}}$.

6.1.1 Example: pure homogeneous stretch of a cube.

Consider a rectangular block with its edges parallel to the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The material is incompressible and transversely isotropic with respect to the direction

$$\mathbf{m}_{\mathrm{R}} = \cos\Theta\,\mathbf{e}_1 + \sin\Theta\,\mathbf{e}_2, \qquad 0 \le \Theta \le \pi/2. \tag{6.15}$$

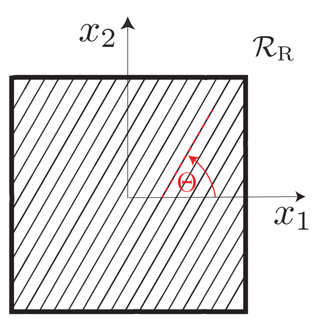


Figure 6.1: Rectangular block in reference configuration occupied by a material involving one family of fibers in the x_1, x_2-plane.

The body is subjected to the pure homogeneous deformation

$$\mathbf{y} = \mathbf{F}\mathbf{x} \qquad \text{where} \qquad \mathbf{F} = \lambda_1\mathbf{e}_1\otimes\mathbf{e}_1 + \lambda_2\mathbf{e}_2\otimes\mathbf{e}_2 + \lambda_3\mathbf{e}_3\otimes\mathbf{e}_3. \tag{6.16}$$

Since the material is incompressible

$$\lambda_3 = \lambda_1^{-1}\lambda_2^{-1}. \tag{6.17}$$

The image of the fiber $\mathbf{m}_\mathrm{R}$ in the deformed configuration is

$$\mathbf{F}\mathbf{m}_\mathrm{R} \overset{(6.15),(6.16)}{=} \lambda_1 \cos\Theta\, \mathbf{e}_1 + \lambda_2 \sin\Theta\, \mathbf{e}_2. \tag{6.18}$$

Let

$$\mathbf{m} = \cos\theta\, \mathbf{e}_1 + \sin\theta\, \mathbf{e}_2, \tag{6.19}$$

denote the direction of a fiber in the deformed configuration. Then, since $\mathbf{F}\mathbf{m}_\mathrm{R}$ is parallel to $\mathbf{m}$, it follows from the two preceding equations that

$$\frac{\lambda_1 \cos\Theta}{\cos\theta} = \frac{\lambda_2 \sin\Theta}{\sin\theta} \qquad \Rightarrow \qquad \tan\theta = \frac{\lambda_2}{\lambda_1}\tan\Theta. \tag{6.20}$$

This gives the fiber orientation θ in the deformed configuration. In particular it tells us how θ varies with the deformation (unless, $\Theta = 0$ or $\pi/2$ in which case the angle θ in the deformed configuration does not depend on the deformation and remains at $\theta = \Theta$).

We now calculate the terms in the constitutive relation (6.13). From $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ and (6.18) we find

$$\mathbf{F}\mathbf{m}_\mathrm{R} \otimes \mathbf{F}\mathbf{m}_\mathrm{R} = \lambda_1^2 \cos^2\Theta \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_1\lambda_2 \cos\Theta \sin\Theta(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + \lambda_2^2 \sin^2\Theta \mathbf{e}_2 \otimes \mathbf{e}_2,$$

$$\begin{aligned}\mathbf{F}\mathbf{m}_\mathrm{R} \otimes \mathbf{B}\mathbf{F}\mathbf{m}_\mathrm{R} + \mathbf{B}\mathbf{F}\mathbf{m}_\mathrm{R} \otimes \mathbf{F}\mathbf{m}_\mathrm{R} = {}& 2\lambda_1^4 \cos^2\Theta\ \mathbf{e}_1 \otimes \mathbf{e}_1 + 2\lambda_2^4 \sin^2\Theta\ \mathbf{e}_2 \otimes \mathbf{e}_2 + \\ & + \lambda_1\lambda_2(\lambda_1^2 + \lambda_2^2)\cos\Theta \sin\Theta\ (\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1).\end{aligned}$$

The invariants (6.6) specialize for the deformation (6.16), (6.17) and fiber direction (6.15) to

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_1^{-2}\lambda_2^{-2}, \qquad I_2 = \lambda_1^{-2} + \lambda_2^{-2} + \lambda_1^2\lambda_2^2, \tag{6.21}$$

$$I_4 = \lambda_1^2 \cos^2\Theta + \lambda_2^2 \sin^2\Theta, \qquad I_5 = \lambda_1^4 \cos^2\Theta + \lambda_2^4 \sin^2\Theta. \tag{6.22}$$

Note that I_4 and I_5 are *not* symmetric in λ_1, λ_2 in general. Consequently, in contrast to the isotropic case, if we replace the I's in the strain energy function W with the λ's using (6.21), (6.22), the resulting expression will *not* be invariant to a change $\lambda_1 \leftrightarrow \lambda_2$.

The constitutive relation (6.13) now gives

$$\begin{aligned}
T_{11} &= -q + 2W_1\lambda_1^2 + 2W_2(I_1\lambda_1^2 - \lambda_1^4) + 2W_4\lambda_1^2\cos^2\Theta + 4W_5\lambda_1^4\cos^2\Theta,\\
T_{22} &= -q + 2W_1\lambda_2^2 + 2W_2(I_1\lambda_2^2 - \lambda_2^4) + 2W_4\lambda_2^2\sin^2\Theta + 4W_5\lambda_2^4\sin^2\Theta,\\
T_{33} &= -q + 2W_1\lambda_3^2 + 2W_2(I_1\lambda_3^2 - \lambda_3^4),\\
T_{12} &= 2[W_4 + W_5(\lambda_1^2 + \lambda_2^2)]\lambda_1\lambda_2 \sin\Theta\cos\Theta, \qquad T_{23} = T_{31} = 0.
\end{aligned} \tag{6.23}$$

Observe that in contrast to the isotropic case, the shear stress[2] $T_{12} \neq 0$. This shear stress is required in order to maintain the deformation (6.16). Thus the deformation (6.16) cannot be sustained (for example) by a state of uniaxial stress $\mathbf{T} = T_{11}\mathbf{e}_1 \otimes \mathbf{e}_1$. Note also that the directions $\mathbf{e}_1, \mathbf{e}_2$ are *not* principal directions for $\mathbf{T}$ though they are principal directions for $\mathbf{B}$. This too is a consequence of anisotropy.

As we have done repeatedly in the isotropic case, it is natural to introduce the restriction of the strain energy function W to the setting at hand by introducing the function

$$w(\lambda_1, \lambda_2, \Theta) := W(I_1, I_2, I_4, I_5), \tag{6.24}$$

where the invariants have the expressions in (6.21) and (6.22). As noted previously, $w(\lambda_1, \lambda_2, \Theta) \neq w(\lambda_2, \lambda_1, \Theta)$ due to the anisotropy. Differentiating w with respect to its arguments and keeping (6.24), (6.21), (6.22) and (6.23) in mind shows that (Ogden [5])

$$T_{11} - T_{33} = \lambda_1 \frac{\partial w}{\partial \lambda_1}, \qquad T_{22} - T_{33} = \lambda_2 \frac{\partial w}{\partial \lambda_2}, \qquad T_{12} = \frac{\lambda_1 \lambda_2}{\lambda_2^2 - \lambda_1^2} \frac{\partial w}{\partial \Theta}, \tag{6.25}$$

where it should be kept in mind that T_{11} and T_{22} are not principal stresses.

To illustrate the response described by (6.23) consider the particular material

$$W = \frac{\mu}{2}(I_1 - 3) + \frac{\mu\beta}{2}(I_4 - 1)^2, \qquad \mu > 0,\ \beta > 0, \tag{i}$$

and the special case of plane strain with vanishing normal stress $T_{22} = 0$:

$$\lambda_3 = 1, \qquad T_{22} = 0. \tag{ii}$$

Note that $T_{33} \neq 0$ and therefore this is not a state of uniaxial stress. Incompressibility gives $\lambda_2 = \lambda^{-1}$ where we have set $\lambda_1 = \lambda$, and (6.21) and (6.22) specialize to

$$I_1 = \lambda^2 + \lambda^{-2} + 1, \qquad I_4 = \lambda^2 \cos^2 \Theta + \lambda^{-2} \sin^2 \Theta. \tag{iii}$$

Setting $T_{22} = 0$ in $(6.23)_2$ allows one to solve for q and use the result to eliminate q from $(6.23)_1$. This leads to

$$\begin{aligned} T_{11}/\mu \quad &= \lambda^2 - \lambda^{-2} + 2\beta\big[\lambda^4 \cos^4 \Theta - \lambda^{-4} \sin^4 \Theta - \lambda^2 \cos^2 \Theta + \lambda^{-2} \sin^2 \Theta\big], \\ T_{12}/(2\mu\beta) \quad &= \big[\lambda^2 \cos^2 \Theta + \lambda^{-2} \sin^2 \Theta - 1\big] \sin\Theta \cos\Theta = \lambda^{-2}(\lambda^2 - 1)(\lambda^2 - \tan^2 \Theta) \sin\Theta \cos^3 \Theta. \end{aligned} \tag{iv}$$

[2]It does vanish in the special cases $\Theta = 0$ or $\Theta = \pi/2$, i.e. when the fibers are oriented in one of the principal directions for $\mathbf{B}$.

The stress component T_{33} needed to maintain the plane strain deformation is given by $(6.23)_3$. Figure 6.2(a) shows plots of the normal stress T_{11}/μ versus λ for three different values of the anisotropy parameter β; while Figure 6.2(b) shows plots of the shear stress $T_{12}/(2\mu\beta)$ versus λ for three different values of the fiber angle Θ.

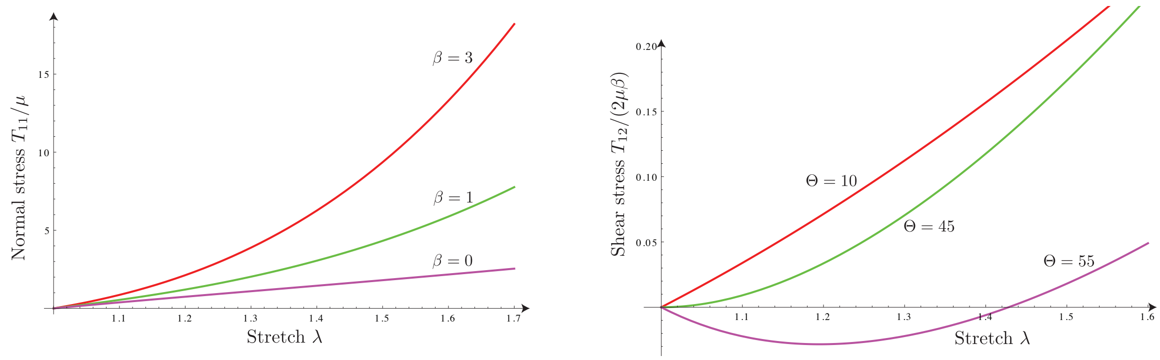


Figure 6.2: Variation of the normal and shear stress components T_{11} and T_{12} with the stretch $\lambda_1 = \lambda$ for the material described by (i) under conditions where $\lambda_3 = 1$ and $T_{22} = 0$.

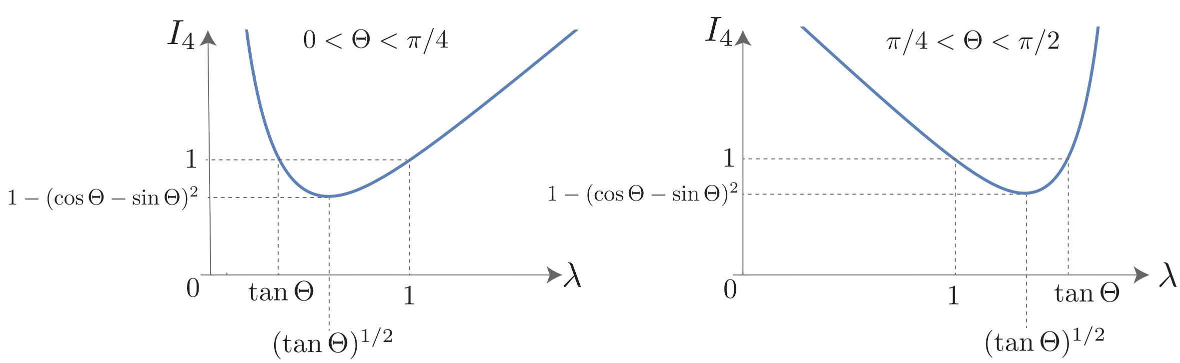


Figure 6.3: Schematic plots of the square of the fiber stretch (I_4) versus the imposed stretch (λ). The left and right figures correspond to the cases $0 < \Theta < \pi/4$ and $\pi/4 < \Theta < \pi/2$ respectively.

Equation $(iii)_2$ can be rewritten equivalently as

$$I_4 = 1 + (1 - \lambda^{-2})(\lambda^2 - \tan^2\Theta)\cos^2\Theta. \tag{v}$$

Figure 6.3 shows plots of I_4 – the square of the fiber stretch – versus the imposed stretch $\lambda\,(= \lambda_1)$. Consider the left-hand figure corresponding to the case $0 < \Theta < \pi/4$ (in which event $\tan\Theta < 1$). When λ increases monotonically from unity, I_4 starts from the value 1 and increases monotonically, eventually tending to infinity as $\lambda \to \infty$. On the other hand when λ decreases monotonically from the value 1, the fiber stretch starts from the value 1 and initially decreases until it reaches a minimum at $\lambda = \sqrt{\tan\Theta}$. It increases monotonically thereafter, passing through the value 1 (again) when $\lambda = \tan\Theta$ and tending to infinity as $\lambda \to 0^+$; the fact that the fibers initially contract in this case (before subsequently elongating) can be intuitively seen by visualizing fibers that are almost horizontal initially. The right-hand figure corresponds to the case $\pi/4 < \Theta < \pi/2$ ($\tan\Theta > 1$) where a similar sort of behavior is seen, with $\lambda > 1$ and $\lambda < 1$ reversed. When $\Theta = \pi/4$, the fiber elongates for all values of λ.

The minimum value of I_4 in Figure 6.3 is $I_4 = 1 - (\cos\Theta - \sin\Theta)^2$. Note from (6.20) that the fiber angle in the deformed configuration is $\theta = \pi/2 - \Theta$ when $\lambda = \tan\Theta$.

Observe from $(iv)_2$ that the shear stress T_{12} vanishes at the two values of stretch $\lambda = 1$ and $\lambda = \tan\Theta$ at which the fiber stretch is unity.

6.2 Two families of fibers.

There are many examples of materials involving two families of fibers, biological tissue with collagen fibers being one. Let the fiber directions (in the reference configuration) be $\mathbf{m}_\mathrm{R}$ and $\mathbf{m}'_\mathrm{R}$. Then the strain energy function will depend on, see Spencer [7, 8], the invariants I_1, I_2, I_3, the invariants I_4 and I_5 associated with the first family of fibers, the analogous invariants I_6 and I_7 for the second family of fibers,

$$I_6 = \mathbf{C}\mathbf{m}'_\mathrm{R} \cdot \mathbf{m}'_\mathrm{R} \qquad I_7 = \mathbf{C}^2\mathbf{m}'_\mathrm{R} \cdot \mathbf{m}'_\mathrm{R}, \tag{6.26}$$

and the invariants I_8 and I_9 *that couple* $\mathbf{m}_\mathrm{R}$ and $\mathbf{m}'_\mathrm{R}$:

$$I_8 = \mathbf{C}\mathbf{m}'_\mathrm{R} \cdot \mathbf{m}_\mathrm{R}, \qquad I_9 = (\mathbf{m}_\mathrm{R} \cdot \mathbf{m}'_\mathrm{R})^2. \tag{6.27}$$

Thus the strain energy function has the form

$$W = \widetilde{W}(I_1, I_2, \ldots, I_8), \tag{6.28}$$

where we have omitted I_9 since it does not involve the deformation gradient[3]. We remark that the form in which I_8 has been written is *not* invariant[4] to the replacement of $\mathbf{m}_\mathrm{R}$ by $-\mathbf{m}_\mathrm{R}$

[3] In view of the forms of I_4 to I_8, one might expect the quantity $\mathbf{C}^2\mathbf{m}_\mathrm{R} \cdot \mathbf{m}'_\mathrm{R}$ to also appear in this list of invariants. In Problem 6.4 you are asked to show that this quantity can be expressed in terms of the other invariants.

[4] Or said differently, it has not been written as a function of $\mathbf{C}, \mathbf{M}$ and $\mathbf{M}'$.

(while keeping $\mathbf{m}'_{\rm R}$ fixed). This can be addressed by working with, for example, $(\mathbf{C}\mathbf{m}'_{\rm R}\cdot\mathbf{m}_{\rm R})^2$ or $(\mathbf{C}\mathbf{m}'_{\rm R}\cdot\mathbf{m}_{\rm R})(\mathbf{m}_{\rm R}\cdot\mathbf{m}'_{\rm R})$. Note also that, in contrast to I_4, I_5, I_6 and I_7, the invariant $I_8 \neq 1$ in the reference configuration.

From (6.1), (6.28) and the chain rule, together with (6.6), (6.26) and (6.27), the constitutive equation for $\mathbf{T}$ (assuming the material to be incompressible) reads

$$
\begin{aligned}
\mathbf{T} \;&= -q\mathbf{I} + 2W_1\mathbf{B} + 2W_2(I_1\mathbf{B} - \mathbf{B}^2) + \\
&+2W_4\mathbf{F}\mathbf{m}_{\rm R}\otimes\mathbf{F}\mathbf{m}_{\rm R} + 2W_6\mathbf{F}\mathbf{m}'_{\rm R}\otimes\mathbf{F}\mathbf{m}'_{\rm R} + \\
&+2W_5(\mathbf{F}\mathbf{m}_{\rm R}\otimes\mathbf{B}\mathbf{F}\mathbf{m}_{\rm R} + \mathbf{B}\mathbf{F}\mathbf{m}_{\rm R}\otimes\mathbf{F}\mathbf{m}_{\rm R}) + 2W_7(\mathbf{F}\mathbf{m}'_{\rm R}\otimes\mathbf{B}\mathbf{F}\mathbf{m}'_{\rm R} + \mathbf{B}\mathbf{F}\mathbf{m}'_{\rm R}\otimes\mathbf{F}\mathbf{m}'_{\rm R}) + \\
&+W_8(\mathbf{F}\mathbf{m}_{\rm R}\otimes\mathbf{F}\mathbf{m}'_{\rm R} + \mathbf{F}\mathbf{m}'_{\rm R}\otimes\mathbf{F}\mathbf{m}_{\rm R})
\end{aligned}
\tag{6.29}
$$

where we have set $W_i = \partial\widetilde{W}/\partial I_i,\ i = 1, \ldots 8$.

Keep in mind that $I_4 = \mathbf{C}\mathbf{m}_{\rm R}\cdot\mathbf{m}_{\rm R}$ and $I_6 = \mathbf{C}\mathbf{m}'_{\rm R}\cdot\mathbf{m}'_{\rm R}$ are corresponding quantities for the two families of fibers. Thus if the two families are *mechanically equivalent* then the energy should be unaffected by an exchange of I_4 and I_6. The same goes for I_5 and I_7. Thus for two mechanically equivalent families of fibers the strain energy function must have the property

$$\widetilde{W}(I_1, I_2, I_4, I_5, I_6, I_7, I_8) = \widetilde{W}(I_1, I_2, I_6, I_5, I_4, I_7, I_8) = \widetilde{W}(I_1, I_2, I_4, I_7, I_6, I_5, I_8). \tag{6.30}$$

There are many strain energy functions that have been proposed in the literature for modeling soft biological tissues. One example is

$$\widetilde{W}(I_1, I_4, I_6) = \frac{\mu_1}{2}(I_1 - 3) + \frac{1}{2}\frac{\mu_4}{k_4}\Big[\exp[k_4(I_4-1)^2] - 1\Big] + \frac{1}{2}\frac{\mu_6}{k_6}\Big[\exp[k_6(I_6-1)^2] - 1\Big], \tag{6.31}$$

where $\mu_1, \mu_4, \mu_6, k_1, k_4, k_6$ are material constants, see Holzapfel et. al. [4]. If the two families of fibers are mechanically equivalent, one would take $\mu_4 = \mu_6$ and $k_4 = k_6$. Observe that (6.31) has the neo-Hookean form for the I_1 term while the terms involving I_4, I_6 are of the Fung form (Section 4.7). If $I_4 - 1$ and $I_6 - 1$ are small, this can be replaced by

$$\widetilde{W} = \frac{\mu_1}{2}(I_1 - 3) + \frac{\mu_4}{2}(I_4 - 1)^2 + \frac{\mu_6}{2}(I_6 - 1)^2; \tag{6.32}$$

the special case of this described by (6.51) below is known as the “standard fiber reinforcing model”.

6.2.1 Example: pure homogeneous stretch of a cube.

Consider a rectangular block with its edges parallel to the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The material is incompressible and involves two families of fibers that, in the reference configuration, are

$$\mathbf{m}_{\mathrm{R}} = \cos\Theta\,\mathbf{e}_1 + \sin\Theta\,\mathbf{e}_2, \qquad \mathbf{m}'_{\mathrm{R}} = \cos\Theta\,\mathbf{e}_1 - \sin\Theta\,\mathbf{e}_2; \tag{6.33}$$

see Figure 6.4. This material is *orthotropic* with the symmetry planes coinciding with the three coordinate planes.

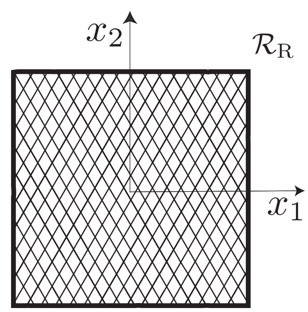


Figure 6.4: Region occupied (in a reference configuration) by an incompressible rectangular block with two families of fibers in the x_1, x_2-plane.

The body is subjected to a homogeneous deformation

$$\mathbf{y} = \mathbf{F}\mathbf{x} \qquad \text{where} \qquad \mathbf{F} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 \otimes \mathbf{e}_3. \tag{6.34}$$

Since the material is incompressible,

$$\lambda_3 = \lambda_1^{-1}\lambda_2^{-1}. \tag{6.35}$$

Let the fiber directions in the deformed configuration be denoted by

$$\mathbf{m} = \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2, \qquad \mathbf{m}' = \cos\theta\,\mathbf{e}_1 - \sin\theta\,\mathbf{e}_2. \tag{6.36}$$

Since $\mathbf{F}\mathbf{m}_{\mathrm{R}} = \lambda_1 \cos\Theta\,\mathbf{e}_1 + \lambda_2 \sin\Theta\,\mathbf{e}_2$ and $\mathbf{F}\mathbf{m}'_{\mathrm{R}} = \lambda_1 \cos\Theta\,\mathbf{e}_1 - \lambda_2 \sin\Theta\,\mathbf{e}_2$ it is readily shown that the fiber angle in the deformed configuration is given by

$$\tan\theta = \frac{\lambda_2}{\lambda_1}\,\tan\Theta; \tag{6.37}$$

see (6.20).

The invariants specialize, for the deformation (6.34), (6.35) and fiber directions (6.33), to

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_1^{-2}\lambda_2^{-2}, \qquad I_2 = \lambda_1^{-2} + \lambda_2^{-2} + \lambda_1^2\lambda_2^2, \tag{6.38}$$

$$I_4 = I_6 = \lambda_1^2 \cos^2\Theta + \lambda_2^2 \sin^2\Theta, \tag{6.39}$$

$$I_5 = I_7 = \lambda_1^4 \cos^2\Theta + \lambda_2^4 \sin^2\Theta, \tag{6.40}$$

$$I_8 = \lambda_1^2 \cos^2\Theta - \lambda_2^2 \sin^2\Theta. \tag{6.41}$$

The constitutive relation (6.29) now gives

$$T_{11} = -q+2W_1\lambda_1^2+2W_2(I_1\lambda_1^2-\lambda_1^4)+2(W_4+W_6+W_8)\lambda_1^2\cos^2\Theta+4(W_5+W_7)\lambda_1^4\cos^2\Theta, \tag{6.42}$$

$$T_{22} = -q+2W_1\lambda_2^2+2W_2(I_1\lambda_2^2-\lambda_2^4)+2(W_4+W_6-W_8)\lambda_2^2\sin^2\Theta+4(W_5+W_7)\lambda_2^4\sin^2\Theta, \tag{6.43}$$

$$T_{33} = -q + 2W_1\lambda_3^2 + 2W_2(I_1\lambda_3^2 - \lambda_3^4). \tag{6.44}$$

$$T_{12} = 2[W_4 - W_6 + (W_5 - W_7)(\lambda_1^2 + \lambda_2^2)]\lambda_1\lambda_2 \sin\Theta\cos\Theta, \qquad T_{23} = T_{31} = 0, \tag{6.45}$$

Observe from (6.45) that the shear stress $T_{12} \neq 0$ in general and so the principal directions of $\mathbf{T}$ do not coincide with those of $\mathbf{B}$. However note from (6.39) and (6.40) that $I_4 = I_6$ and $I_5 = I_7$. Therefore if the two fiber families are mechanically equivalent, i.e. if (6.30) holds, then $W_4 = W_6$ and $W_5 = W_7$. In this case (6.45) gives $T_{12} = 0$ (as one would expect).

Again, it is natural to introduce the restriction of the strain energy function W to the setting at hand by introducing the function

$$w(\lambda_1, \lambda_2, \Theta) = W(I_1, I_2, I_4, I_5, I_6, I_7, I_8), \tag{6.46}$$

where the invariants are expressed in terms of $\lambda_1, \lambda_2, \Theta$ by (6.38), (6.39), (6.40) and (6.41). Note that $w(\lambda_1, \lambda_2, \Theta) \neq w(\lambda_2, \lambda_1, \Theta)$ due to anisotropy. Differentiating w with respect to λ_1 and λ_2 shows that (Ogden [5])

$$T_{11} - T_{33} = \lambda_1 \frac{\partial w}{\partial \lambda_1}, \qquad T_{22} - T_{33} = \lambda_2 \frac{\partial w}{\partial \lambda_2}, \tag{6.47}$$

keeping in mind that T_{11} and T_{22} are not principal stresses. The direct analog of $(6.25)_3$ does not hold.

Remark: In contrast to Problem 2.3, the fibers here are not constrained to being inextensible. In Section 6.2.2 we shall consider the inextensible case.

Remark: In certain materials one may wish to allow fibers to elongate but not to shorten. Since the fiber stretch is $|\mathbf{F}\mathbf{m}_{\mathrm{R}}| = |\mathbf{F}\mathbf{m}'_{\mathrm{R}}| = [\lambda_1^2\cos^2\Theta + \lambda_2^2\sin^2\Theta]^{1/2}$, in such a situation one must constrain the stretches to obey

$$\lambda_1^2\cos^2\Theta + \lambda_2^2\sin^2\Theta \geq 1. \tag{6.48}$$

Suppose now that the fiber families are mechanically equivalent and therefore that the shear stress T_{12} vanishes automatically. Suppose further that the boundary conditions on the block lead to a state of *uniaxial stress*[5]: $\mathbf{T} = T_{11}\,\mathbf{e}_1\otimes\mathbf{e}_1$. Setting $T_{22} = T_{33} = 0$ in (6.47) gives

$$T_{11} = \lambda_1\frac{\partial w}{\partial\lambda_1}, \qquad \frac{\partial w}{\partial\lambda_2} = 0. \tag{6.49}$$

The second of these is an equation involving the two stretches λ_1 and λ_2 (and Θ). If it can be solved (in principle) for λ_2, we would have a relation $\lambda_2 = \lambda_2(\lambda_1,\Theta)$ between the axial stretch λ_1 and the transverse stretch λ_2. (The third principal stretch is $\lambda_3 = \lambda_1^{-1}\lambda_2^{-1}$.) Substituting this back into the first equation in (6.49) gives the following stress-stretch relation between T_{11} and λ_1

$$T_{11} = T_{11}(\lambda_1) = \lambda_1\frac{\partial w}{\partial\lambda_1}(\lambda_1,\lambda_2,\Theta)\bigg|_{\lambda_2=\lambda_2(\lambda_1,\Theta)}. \tag{6.50}$$

To illustrate the response in uniaxial stress[6] consider the so-called "standard fiber reinforcing model",

$$W = \frac{\mu}{2}(I_1-3) + \frac{\mu\beta}{2}\Big[(I_4-1)^2 + (I_6-1)^2\Big], \qquad \mu>0, \beta>0. \tag{6.51}$$

This is the special case of (6.32) with

$$\mu_1 = \mu, \qquad \mu_4 = \mu_6 = \beta\mu. \tag{6.52}$$

Since this strain energy function obeys (6.30) the two fiber families are mechanically equivalent. Observe that the parameter $\beta > 0$ effectively characterizes the stiffness of the fibers (due to both the actual fiber stiffness and the concentration of fibers). Large β corresponds to stiff fibers.

The relation between the transverse stretch λ_2 and the longitudinal stretch λ_1 as given by $(6.49)_2$ specializes for the material (6.51) to the following cubic equation for λ_2^2:

$$4\beta\sin^4\Theta\,\lambda_2^6 + \Big[1 + 4\beta(\lambda_1^2\cos^2\Theta - 1)\sin^2\Theta\Big]\,\lambda_2^4 - \lambda_1^{-2} = 0. \tag{6.53}$$

---

[5]Note that the deformation (6.34) would not be compatible with a state of *uniaxial stress* in the $\mathbf{e}_1$-direction if $T_{12} \neq 0$.

[6]Based on Chapter 11 of Goriely [2] .

In principle, one solves this for λ_2 to get

$$\lambda_2 = \lambda_2(\lambda_1, \Theta). \tag{6.54}$$

Suppose we want to calculate the Poisson's ratio at infinitesimal deformations that measures the contraction in the x_2-direction with respect to the stretch in the x_1-direction. This is given by $-d\lambda_2/d\lambda_1$ evaluated at the undeformed configuration $\lambda_2 = 1$. (Why?) Differentiating (6.53) with respect to λ_1 gives

$$24\beta \sin^4\Theta\, \lambda_2^5 \frac{d\lambda_2}{d\lambda_1} + \left[8\beta\lambda_1 \cos^2\Theta \sin^2\Theta\right] \lambda_2^4 + \left[1 + 4\beta(\lambda_1^2\cos^2\Theta - 1)\sin^2\Theta\right] 4\lambda_2^3 \frac{d\lambda_2}{d\lambda_1} + 2\lambda_1^{-3} = 0.$$

Solving this for $d\lambda_2/d\lambda_1$ and evaluating the result at $\lambda_1 = \lambda_2 = 1$ leads to

$$\left.\frac{d\lambda_2}{d\lambda_1}\right|_{\lambda_1=1} = -\frac{1 + 4\beta\cos^2\Theta\sin^2\Theta}{2 + 4\beta\sin^4\Theta}. \tag{6.55}$$

The (negative) of this gives the particular Poisson's ratio we sought as a function of the fiber angle Θ. Observe that if $\beta = 0$, corresponding to a neo-Hookean material, this reduces to the classical value 1/2. At the other extreme, in the limit $\beta \to \infty$, this reduces to the value found previously in Problem 2.3 for rigid fibers. Observe that for *all values* of the anisotropy parameter β, this Poisson ratio has the value 1/2 at the particular fiber angle Θ_* given by[7]

$$\tan\Theta_* = \sqrt{2} \qquad \Theta_* \approx 54.74^o. \tag{6.56}$$

Since the material is anisotropic, (6.55) is *not* the Poisson's ratio that measures the contraction in the x_3-direction with respect to the x_1-direction. To determine that, we differentiate $\lambda_3 = \lambda_1^{-1}\,\lambda_2^{-1}(\lambda_1)$ with respect to λ_1 and evaluate the result at $\lambda_1 = 1$. This leads to

$$\left.\frac{d\lambda_3}{d\lambda_1}\right|_{\lambda_1=1} = -\frac{1 - 4\beta\cos^2\Theta\sin^2\Theta + 4\beta\sin^4\Theta}{2 + 4\beta\sin^4\Theta}. \tag{6.57}$$

Again, observe that if $\beta = 0$, corresponding to a neo-Hookean material, this reduces to the classical value 1/2. At the other extreme, in the limit $\beta \to \infty$, it reduces to the value we found in Problem 2.3 for rigid fibers. Again, for *all values* of the anisotropy parameter β, this Poisson ratio also has the value 1/2 at the particular fiber angle Θ_* given by (6.56).

The relation between the stress T_{11} and stretch λ_1 is now found from (6.50), (6.32), (6.46) and (6.52) :

$$T_{11} = T_{11}(\lambda_1) = \mu\Big[\lambda_1^2 - \lambda_1^{-2}\lambda_2^{-2} + 4\beta(\lambda_1^2\cos^2\Theta + \lambda_2^2\sin^2\Theta - 1)\lambda_1^2\cos^2\Theta\Big] \tag{6.58}$$

[7] This angle appears in various other contexts and is referred to as the "magic angle".

with $\lambda_2 = \lambda_2(\lambda_1, \Theta)$ given by (6.54). The graph of $\widehat{T}_{11}(\lambda_1)$ versus λ_1 describes the stress-stretch behavior in uniaxial stress.

To find the *effective Young's modulus* at infinitesimal deformations (for stress in the particular direction under consideration) we differentiate (6.58) with respect to λ_1 and evaluate the result at $\lambda_1 = \lambda_2 = 1$. This yields

$$E_{\text{eff}}(\Theta) = \left.\frac{dT_{11}}{d\lambda_1}\right|_{\lambda_1=1} = 4\mu + 8\beta\mu\cos^4\Theta + 2\mu(1 + 4\beta\sin^2\Theta\cos^2\Theta)\left.\frac{d\lambda_2}{d\lambda_1}\right|_{\lambda_1=1},$$

which upon using (6.55) simplifies to

$$E_{\text{eff}}(\Theta) = \left.\frac{dT_{11}}{\partial\lambda_1}\right|_{\lambda_1=1} = \mu\,\frac{4[3 + 5\beta + 3\beta\cos 4\Theta]}{4 + 3\beta - 4\beta\cos 2\Theta + \beta\cos 4\Theta}. \tag{6.59}$$

This gives the effective Young's modulus of the material (in the particular direction under consideration) as a function of the fiber angle Θ. If $\beta = 0$, corresponding to a neo-Hookean material, (6.59) yields

$$E_{\text{eff}}(\Theta) = 3\mu, \tag{6.60}$$

which coincides with the value of Young's modulus we found previously; see discussion below in (4.90). At the other extreme when $\beta \to \infty$, corresponding to rigid fibers, the limiting value of $E_{\text{eff}}(\Theta)$ agrees with what we will find in Section 6.2.2. One can readily verify that the effective Young's modulus (6.59) is *independent of the anisotropy parameter* β at the particular fiber angle given by (6.56) and has the value 3μ.

It is interesting to examine the variation of the effective Young's modulus with the fiber angle. It is readily found that the maximum value of $E_{\text{eff}}(\Theta)$ (as a function of the fiber angle Θ) occurs at $\Theta = 0$ corresponding to the case when the fibers are parallel to the stressing direction. Its value is

$$\left.E_{\text{eff}}\right|_{\max} = E_{\text{eff}}(0) = \mu(3 + 8\beta). \tag{6.61}$$

When the fibers are perpendicular to the stressing direction, $\Theta = \pi/2$, one finds

$$E_{\text{eff}}(\pi/2) = \mu\,\frac{3 + 8\beta}{1 + 2\beta}. \tag{6.62}$$

This, however, turns out not to be the minimum value of $E_{\text{eff}}(\Theta)$. It is not difficult to show that the minimum value occurs when the fiber angle has the value Θ_* given by (6.56) and that this value is

$$\left.E_{\text{eff}}\right|_{\min} = E_{\text{eff}}(\Theta_*) = 3\mu. \tag{6.63}$$

Figure (6.5) shows the graph of $E_{\text{eff}}(\Theta)$ versus Θ where the figure has been drawn for $\beta = 1$.

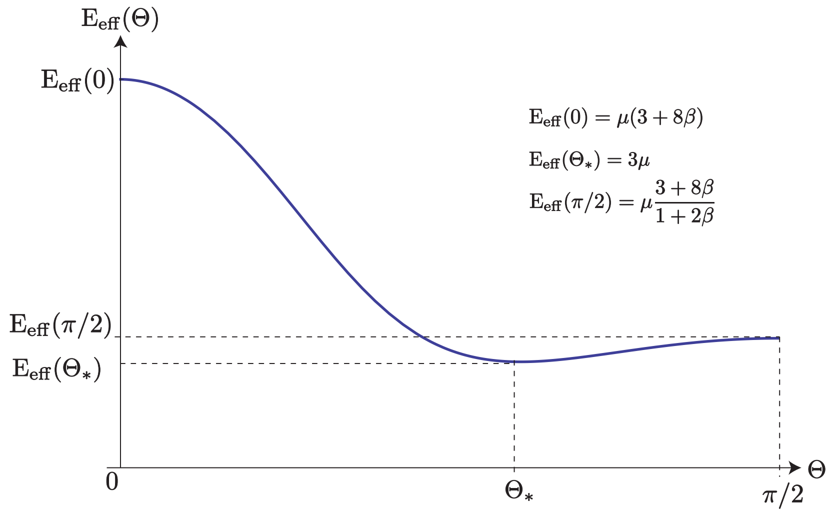


Figure 6.5: Effective Young's modulus as a function of fiber angle. Figure drawn for $\beta = 1$.

6.2.2 Inextensible fibers.

Consider again the problem studied in Section 6.2.1 but now assume the fibers to be inextensible. The kinematics of this problem was analyzed previously in Problem 2.3.

The constitutive relation for a material with two families of fibers is given by (6.29). If the fibers are inextensible, then $I_4 = |\mathbf{F}\mathbf{m}_\mathrm{R}|^2 = 1$ and $I_6 = |\mathbf{F}\mathbf{m}'_\mathrm{R}|^2 = 1$ and so the strain energy function has the form $W = \widetilde{W}(I_1, I_2, I_5, I_7, I_8)$ having also assumed the material to be incompressible. The terms involving W_3, W_4 and W_6 in the constitutive relation must be omitted, and the reaction stresses arising due to the constraints must be included. We know from Problem 4.19(a) that the reaction stress associated with inextensibility is a uniaxial stress in the direction of the deformed fibers (and a hydrostatic stress due to incompressibility). Therefore (6.29) is replace by

$$\begin{aligned}\mathbf{T} &= -q\mathbf{I} + q_4\mathbf{F}\mathbf{m}_\mathrm{R}\otimes\mathbf{F}\mathbf{m}_\mathrm{R} + q_6\mathbf{F}\mathbf{m}'_\mathrm{R}\otimes\mathbf{F}\mathbf{m}'_\mathrm{R} + 2W_1\mathbf{B} + 2W_2(I_1\mathbf{B} - \mathbf{B}^2)+\\ &\quad +2W_5(\mathbf{F}\mathbf{m}_\mathrm{R}\otimes\mathbf{B}\mathbf{F}\mathbf{m}_\mathrm{R} + \mathbf{B}\mathbf{F}\mathbf{m}_\mathrm{R}\otimes\mathbf{F}\mathbf{m}_\mathrm{R}) + 2W_7(\mathbf{F}\mathbf{m}'_\mathrm{R}\otimes\mathbf{B}\mathbf{F}\mathbf{m}'_\mathrm{R} + \mathbf{B}\mathbf{F}\mathbf{m}'_\mathrm{R}\otimes\mathbf{F}\mathbf{m}'_\mathrm{R})+\\ &\quad +W_8(\mathbf{F}\mathbf{m}_\mathrm{R}\otimes\mathbf{F}\mathbf{m}'_\mathrm{R} + \mathbf{F}\mathbf{m}'_\mathrm{R}\otimes\mathbf{F}\mathbf{m}_\mathrm{R})\end{aligned} \tag{6.64}$$

where q is due to incompressibility, q_4 is due to inextensibility of the $\mathbf{m}_\mathrm{R}$-fibers and q_6 is due to inextensibility of the $\mathbf{m}'_\mathrm{R}$-fibers.

For illustrative purposes consider the strain energy function (6.31), but now omit the terms involving I_4 and I_6:

$$W = \frac{\mu}{2}(I_1 - 3), \qquad \mu > 0. \tag{6.65}$$

The constitutive relation (6.64) for the Cauchy stress now simplifies to

$$\mathbf{T} = -q\mathbf{I} + \mu\mathbf{B} + q_4\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}} + q_6\mathbf{Fm}'_{\mathrm{R}} \otimes \mathbf{Fm}'_{\mathrm{R}}. \tag{6.66}$$

As in Section 6.2.1, consider a rectangular block of the material with its edges parallel to the basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The material is incompressible and involves two families of inextensible fibers that, in the reference configuration are oriented as in (6.33); see Figure 6.4.

The body is subjected to a homogeneous deformation $\mathbf{y} = \mathbf{Fx}$ where $\mathbf{F}$ is given by (6.34). It follows from (6.33) and (6.34) that

$$\mathbf{Fm}_{\mathrm{R}} = \lambda_1 \cos\Theta\, \mathbf{e}_1 + \lambda_2 \sin\Theta\, \mathbf{e}_2, \qquad \mathbf{Fm}'_{\mathrm{R}} = \lambda_1 \cos\Theta\, \mathbf{e}_1 - \lambda_2 \sin\Theta\, \mathbf{e}_2, \tag{6.67}$$

and so inextensibility requires

$$I_4 = I_6 = |\mathbf{Fm}_{\mathrm{R}}| = |\mathbf{Fm}'_{\mathrm{R}}| = \lambda_1^2 \cos^2\Theta + \lambda_2^2 \sin^2\Theta = 1. \tag{6.68}$$

Since the material is incompressible,

$$\lambda_3 = \lambda_1^{-1}\lambda_2^{-1}. \tag{6.69}$$

Thus we have *two* constraints on the principal stretches. The kinematics of this problem was analyzed in detail in Problem 2.3. In particular, we found that the stretch λ_1 is restricted to the range $0 < \lambda_1 < 1/\cos\Theta$ and solving (6.68) and (6.69) for λ_2 and λ_3 in terms of λ_1 led to

$$\lambda_2 = \frac{(1 - \lambda_1^2\cos^2\Theta)^{1/2}}{\sin\Theta}, \qquad \lambda_3 = \frac{\sin\Theta}{\lambda_1(1 - \lambda_1^2\cos^2\Theta)^{1/2}}, \qquad 0 < \lambda_1 < 1/\cos\Theta. \tag{6.70}$$

Graphs of λ_2 and λ_3 versus λ_1 for $1 < \lambda_1 < 1/\cos\Theta$ are shown in Figure 6.6. Observe that λ_3 is not a monotonic function of λ_1. The slopes of these curves at $\lambda_1 = 1$ are the negatives of the Poisson's ratios (which, for the extensible case, we found previously).

We next turn to the relation between T_{11} and λ_1. It follows from (6.34) and (6.67) that

$$\mathbf{B} = \mathbf{FF}^T = \lambda_1^2\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2^2\mathbf{e}_2 \otimes \mathbf{e}_2 + \lambda_3^2\mathbf{e}_3 \otimes \mathbf{e}_3, \tag{6.71}$$

$$\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}} = \lambda_1^2\cos^2\Theta\, \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_1\lambda_2 \sin\Theta\cos\Theta(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1) + \lambda_2^2 \sin^2\Theta\, \mathbf{e}_2 \otimes \mathbf{e}_2, \tag{6.72}$$

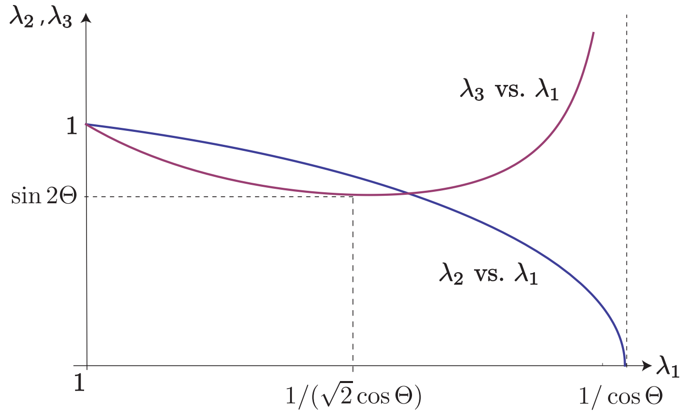


Figure 6.6: Transverse stretches λ_2 and λ_3 versus longitudinal stretch λ_1 in uniaxial stress according to (6.70) for the block shown in Figure 6.4. Figure has been drawn for $\Theta = 3\pi/8$.

$$\mathbf{F}\mathbf{m}'_{\rm R}\otimes\mathbf{F}\mathbf{m}'_{\rm R} = \lambda_1^2\cos^2\Theta\,\mathbf{e}_1\otimes\mathbf{e}_1 - \lambda_1\lambda_2\sin\Theta\cos\Theta(\mathbf{e}_1\otimes\mathbf{e}_2+\mathbf{e}_2\otimes\mathbf{e}_1)+\lambda_2^2\sin^2\Theta\,\mathbf{e}_2\otimes\mathbf{e}_2, \quad (6.73)$$

which when substituted into the constitutive relation (6.66) gives

$$T_{11} = -q + \mu\lambda_1^2 + (q_4+q_6)\lambda_1^2\cos^2\Theta, \quad (6.74)$$

$$T_{22} = -q + \mu\lambda_2^2 + (q_4+q_6)\lambda_2^2\sin^2\Theta, \quad (6.75)$$

$$T_{33} = -q + \mu\lambda_3^2, \quad (6.76)$$

$$T_{12} = [q_4 - q_6]\lambda_1\lambda_2\sin\Theta\cos\Theta, \qquad T_{23} = T_{31} = 0. \quad (6.77)$$

Suppose now that the boundary conditions on the body lead to a state of uniaxial stress:

$$\mathbf{T} = T\,\mathbf{e}_1\otimes\mathbf{e}_1, \quad (6.78)$$

where T is the component of Cauchy stress in the loading direction. On setting $T_{22} = 0, T_{33} = 0$ and $T_{12} = 0$ in (6.75), (6.76), (6.77) we find

$$q_4 = q_6 = \frac{\mu(\lambda_3^2 - \lambda_2^2)}{2\lambda_2^2\sin^2\Theta}, \qquad q = \mu\lambda_3^2. \quad (6.79)$$

Substituting (6.79) into (6.74) gives the stress T:

$$T/\mu = (\lambda_1^2 - \lambda_3^2) + \left[\frac{\lambda_3^2}{\lambda_2^2} - 1\right]\frac{\lambda_1^2\cos^2\Theta}{\sin^2\Theta} \quad (6.80)$$

Finally we substitute for λ_2 and λ_3 from (6.70) to get

$$T/\mu = \frac{\sin^2\Theta - \cos^2\Theta}{\sin^2\Theta}\lambda_1^2 \;+\; \frac{\sin^2\Theta(2\lambda_1^2\cos^2\Theta - 1)}{\lambda_1^2(1-\lambda_1^2\cos^2\Theta)^2}. \quad (6.81)$$

Figure 6.7 shows a plot of T versus λ_1. The figure has been drawn for $\Theta = \pi/3$.

Differentiating (6.81) with respect to λ_1 and evaluating the result at $\lambda_1 = 1$ gives the effective Young's modulus (in the direction under consideration):

$$E_{\text{eff}}(\Theta) = 4\mu\ \frac{5 + 3\cos 4\Theta}{3 - 4\cos 2\Theta + \cos 4\Theta}. \tag{6.82}$$

This agrees with the limit $\beta \to \infty$ of the result (6.59) we got in the case of the extensible fibers. Observe that $E_{\text{eff}}(\Theta) \to \infty$ as $\Theta \to 0$ corresponding to the rigid fibers being in the x_1–direction. When the fibers are perpendicular to the x_1-direction we get $E_{\text{eff}}(\pi/2) = 4\mu$.

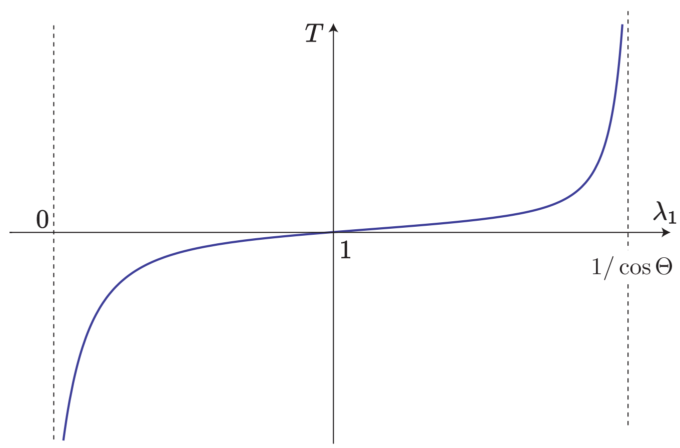


Figure 6.7: Stress T_{11}-stretch λ_1 curve in uniaxial stress according to (6.81) for the block shown in Figure 6.4.

6.2.3 Inflation, extension and twisting of a *thin*-walled tube.

Consider a *thin-walled* tube of mean radius R, wall thickness T and length L in the reference configuration[8]. Its two ends are closed. We are told that $T \ll R$. The tube wall involves two in-plane families of fibers. They are inclined, in the reference configuration, at angles Φ and $-\Psi$ from the *circumferential direction* as shown in Figure 6.10:

$$\mathbf{m}_{\text{R}} = \cos\Phi\,\mathbf{e}_\Theta + \sin\Phi\,\mathbf{e}_Z, \qquad \mathbf{m}'_{\text{R}} = -\cos\Psi\,\mathbf{e}_\Theta + \sin\Psi\,\mathbf{e}_Z. \tag{6.83}$$

The fibers are not necessarily mechanically equivalent and they are not necessarily oriented symmetrically with respect to the tube, i.e. $\Phi \neq \Psi$. The tube is subjected to an internal

[8]The corresponding problem for a thick-walled cylinder is considered in Problem 6.7.

pressure p, twisting moment (torque) M and axial force F. This loading expands the tube to a mean radius r, elongates it to a length ℓ and rotates one end of the tube with respect to the other by an angle $\alpha\ell$. Let t be the wall thickness in the deformed configuration.

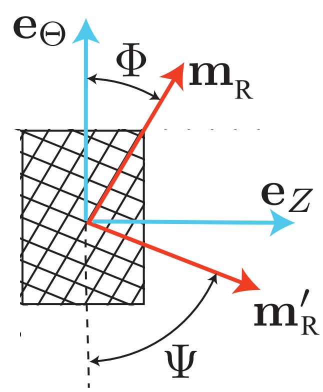


Figure 6.8: In the reference configuration the fiber directions $\mathbf{m}_\mathrm{R}$ and $\mathbf{m}'_\mathrm{R}$, locally, at each point in the tube, lie in the Θ, Z-plane as shown.

Let $\lambda = r/R$ and $\Lambda = \ell/L$ be the mean circumferential and axial stretches of the tube. The volumes of the circular part of the tube before and after deformation are $2\pi RTL$ and $2\pi rt\ell$ respectively and so, since the material is incompressible, $2\pi RTL = 2\pi rt\ell$. Thus

$$r = \lambda R, \qquad \ell = \Lambda L, \qquad t = \lambda^{-1}\Lambda^{-1}T. \tag{6.84}$$

Assume that the deformation that takes $(R, \Theta, Z) \to (r, \theta, z)$ has the form

$$r = r(R), \qquad \theta = \Theta + \alpha\Lambda Z, \qquad z = \Lambda Z. \tag{6.85}$$

The associated deformation gradient tensor is

$$\mathbf{F} = \lambda^{-1}\Lambda^{-1}\mathbf{e}_r \otimes \mathbf{e}_R + \lambda\,\mathbf{e}_\theta \otimes \mathbf{e}_\Theta + \alpha\Lambda\, r\mathbf{e}_\theta \otimes \mathbf{e}_Z + \Lambda\mathbf{e}_z \otimes \mathbf{e}_Z, \tag{6.86}$$

where incompressibility has been used in writing the first term. The left Cauchy Green tensor is

$$\mathbf{B} = \mathbf{F}\mathbf{F}^T = \lambda^{-2}\Lambda^{-2}\mathbf{e}_r\otimes\mathbf{e}_r + (\lambda^2+\alpha^2\Lambda^2 r^2)\mathbf{e}_\theta\otimes\mathbf{e}_\theta + \Lambda^2\mathbf{e}_z\otimes\mathbf{e}_z + \alpha\Lambda^2\, r(\mathbf{e}_\theta\otimes\mathbf{e}_z + \mathbf{e}_z\otimes\mathbf{e}_\theta), \tag{6.87}$$

The vectors $\mathbf{F}\mathbf{m}_\mathrm{R}, \mathbf{F}\mathbf{m}'_\mathrm{R}$, tensors $\mathbf{F}\mathbf{m}_\mathrm{R} \otimes \mathbf{F}\mathbf{m}_\mathrm{R}, \mathbf{F}\mathbf{m}'_\mathrm{R} \otimes \mathbf{F}\mathbf{m}'_\mathrm{R}$ and invariants I_1, I_4 and I_6 can now be readily calculated:

$$\mathbf{F}\mathbf{m}_\mathrm{R} = (\lambda\cos\Phi + r\Lambda\alpha\sin\Phi)\mathbf{e}_\theta + \Lambda\sin\Phi\mathbf{e}_z, \tag{6.88}$$

$$\mathbf{Fm}'_{\mathrm{R}} = (-\lambda \cos\Psi + r\Lambda\alpha \sin\Psi)\mathbf{e}_\theta + \Lambda \sin\Psi \mathbf{e}_z, \tag{6.89}$$

$$\begin{aligned} \mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}} &= (\lambda\cos\Phi + r\Lambda\alpha\sin\Phi)^2 \mathbf{e}_\theta \otimes \mathbf{e}_\theta + \Lambda^2 \sin^2\Phi \mathbf{e}_z \otimes \mathbf{e}_z + \\ &+ (\lambda\cos\Phi + r\Lambda\alpha\sin\Phi)\Lambda\sin\Phi(\mathbf{e}_z \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_z). \end{aligned} \tag{6.90}$$

$$\begin{aligned} \mathbf{Fm}'_{\mathrm{R}} \otimes \mathbf{Fm}'_{\mathrm{R}} &= (-\lambda\cos\Psi + r\Lambda\alpha\sin\Psi)^2 \mathbf{e}_\theta \otimes \mathbf{e}_\theta + \Lambda^2 \sin^2\Psi \mathbf{e}_z \otimes \mathbf{e}_z + \\ &+ (-\lambda\cos\Psi + r\Lambda\alpha\sin\Psi)\Lambda\sin\Psi(\mathbf{e}_z \otimes \mathbf{e}_\theta + \mathbf{e}_\theta \otimes \mathbf{e}_z). \end{aligned} \tag{6.91}$$

$$\left. \begin{aligned} I_1 &= \operatorname{tr}\mathbf{B} = \lambda^{-2}\Lambda^{-2} + \lambda^2 + \alpha^2\lambda^2\Lambda^2R^2 + \Lambda^2, \\ I_4 &= |\mathbf{Fm}_{\mathrm{R}}|^2 = \lambda^2(\cos\Phi + \alpha\Lambda R\sin\Phi)^2 + \Lambda^2\sin^2\Phi, \\ I_6 &= |\mathbf{Fm}'_{\mathrm{R}}|^2 = \lambda^2(-\cos\Psi + \alpha\Lambda R\sin\Psi)^2 + \Lambda^2\sin^2\Psi. \end{aligned} \right\} \tag{6.92}$$

The constitutive relation

$$\mathbf{T} = -q\mathbf{I} + 2W_1\mathbf{B} + 2W_4\mathbf{Fm}_{\mathrm{R}} \otimes \mathbf{Fm}_{\mathrm{R}} + 2W_6\mathbf{Fm}'_{\mathrm{R}} \otimes \mathbf{Fm}'_{\mathrm{R}} \tag{6.93}$$

together with (6.87), (6.90), (6.91) and (6.92) can be used to calculate the stress components $T_{rr}, T_{\theta\theta}, T_{zz}$ and $T_{z\theta}$ (the remaining two stress components $T_{r\theta}$ and T_{rz} vanish). This leads to the following expression where we have eliminated the reactive pressure q by subtracting the normal stress T_{rr} from the other two normal stresses $T_{\theta\theta}$ and T_{zz}:

$$\begin{aligned} T_{\theta\theta} - T_{rr} &= 2(\lambda^2 + \alpha^2\lambda^2\Lambda^2R^2 - \lambda^{-2}\Lambda^{-2})W_1 + 2\lambda^2(\cos\Phi + R\Lambda\alpha\sin\Phi)^2W_4 + \\ &+ 2\lambda^2(-\cos\Psi + R\Lambda\alpha\sin\Psi)^2W_6, \\ T_{zz} - T_{rr} &= 2(\Lambda^2 - \lambda^{-2}\Lambda^{-2})W_1 + 2\Lambda^2\sin^2\Phi W_4 + 2\Lambda^2\sin^2\Psi W_6 \\ T_{\theta z} &= 2\alpha R\lambda\Lambda^2\left[W_1 + \sin^2\Phi W_4 + \sin^2\Psi W_6\right] + \lambda\Lambda\left[\sin 2\Phi W_4 - \sin 2\Psi W_6\right]. \end{aligned} \tag{6.94}$$

By exploiting the symmetry of the problem and the fact that the tube is thin-walled, we can use the equilibrium equations to derive approximate expressions for the stress components, exactly as we did in the isotropic case. This leads to

$$T_{\theta\theta} \approx \frac{pr}{t}, \qquad T_{zz} \approx \frac{F + \pi r^2 p}{2\pi r t} \qquad T_{z\theta} \approx \frac{M}{2\pi r^2 t}, \qquad T_{rr} \approx -\frac{p}{2}, \tag{6.95}$$

where the term $\pi r^2 p$ in T_{zz} arises because the two ends of the tube are closed. Note that, due to the factor $1/t$ in $T_{\theta\theta}$ and T_{zz} these two normal stress components are significantly larger than T_{rr}.

Finally, combining (6.94) with (6.95) and dropping T_{rr} leads to

$$
\begin{aligned}
\frac{pr}{t} &= 2(\lambda^2 + \alpha^2\lambda^2\Lambda^2R^2 - \lambda^{-2}\Lambda^{-2})W_1 + 2\lambda^2(\cos\Phi + \alpha\Lambda R\sin\Phi)^2W_4 + \\
&\qquad +2\lambda^2(-\cos\Psi + \alpha\Lambda R\sin\Psi)^2W_6, \\
\frac{F + \pi r^2 p}{2\pi r t} &= 2(\Lambda^2 - \lambda^{-2}\Lambda^{-2})W_1 + 2\Lambda^2\sin^2\Phi W_4 + 2\Lambda^2\sin^2\Psi W_6, \\
\frac{M}{2\pi r^2 t} &= 2\alpha\lambda\Lambda^2R\left[W_1 + \sin^2\Phi W_4 + \sin^2\Psi W_6\right] + \lambda\Lambda\left[\sin 2\Phi W_4 - \sin 2\Psi W_6\right].
\end{aligned}
\tag{6.96}
$$

Given p, F and M, the three equations (6.96) are to be solved for λ, Λ and α. Various special cases can be examined. For example, suppose we do not apply a twisting moment M. The third equation in (6.96) must still hold with zero on its left hand side; the rotation α of the tube will not be zero in general.

6.3 Worked Examples and Exercises.

Problem 6.1. Let $\mathbf{M}$ denote the tensor

$$\mathbf{M} = \mathbf{m}_{\rm R} \otimes \mathbf{m}_{\rm R}, \tag{6.97}$$

where $\mathbf{m}_{\rm R}$ is a unit vector. Show that

$$\mathbf{Q}\mathbf{m}_{\rm R} = \pm\mathbf{m}_{\rm R}, \tag{6.98}$$

for an orthogonal tensor $\mathbf{Q}$ if and only if

$$\mathbf{Q}\mathbf{M}\mathbf{Q}^T = \mathbf{M}. \tag{6.99}$$

An important consequence of this is that the groups

$$\mathcal{G} = \{\mathbf{Q} \ : \ \mathbf{Q}\mathbf{Q}^T = \mathbf{I},\ \mathbf{Q}\mathbf{m}_{\rm R} = \pm\mathbf{m}_{\rm R}\ \} \quad \text{and} \quad \mathcal{G}' = \{\mathbf{Q} \ : \ \mathbf{Q}\mathbf{Q}^T = \mathbf{I},\ \mathbf{Q}\mathbf{M}\mathbf{Q}^T = \mathbf{M}\ \} \tag{6.100}$$

are identical.

Problem 6.2. Keeping in mind that the strain energy function is invariant to replacing $\mathbf{m}_R$ by $-\mathbf{m}_R$, let

$$\overline{W}(\mathbf{C}, \mathbf{m}_R) = \breve{W}(\mathbf{C}, \mathbf{m}_R \otimes \mathbf{m}_R). \tag{i}$$

Show that

$$\overline{W}(\mathbf{C}, \mathbf{m}_R) = \overline{W}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T, \mathbf{Q}\mathbf{m}_R) \qquad \text{for all } \mathbf{Q} \in \mathcal{G}, \tag{ii}$$

if and only if

$$\breve{W}(\mathbf{C}, \mathbf{m}_R \otimes \mathbf{m}_R) = \breve{W}\big(\mathbf{Q}\mathbf{C}\mathbf{Q}^T, \mathbf{Q}(\mathbf{m}_R \otimes \mathbf{m}_R)\mathbf{Q}^T\big) \qquad \text{for all orthogonal } \mathbf{Q}. \tag{iii}$$

Here $\mathcal{G}$ is the material symmetry group for transverse isotropy given in (6.100). Note that the second statement holds for *all* orthogonal $\mathbf{Q}$ not just those in $\mathcal{G}$. Therefore the function $\breve{W}$ is jointly isotropic in both arguments.

Problem 6.3. Show that $\breve{W}(\mathbf{C}, \mathbf{M}) = \breve{W}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T, \mathbf{Q}\mathbf{M}\mathbf{Q}^T)$ for all orthogonal $\mathbf{Q}$ if and only if there is a function $\widetilde{W}$ such that

$$\breve{W}(\mathbf{C}, \mathbf{M}) = \widetilde{W}(I_1, I_2, I_3, I_4, I_5), \tag{6.101}$$

where

$$\begin{gathered} I_1(\mathbf{C}) = \operatorname{tr}\mathbf{C}, \qquad I_2(\mathbf{C}) = \tfrac{1}{2}\Big[(\operatorname{tr}\mathbf{C})^2 - \operatorname{tr}\mathbf{C}^2\Big], \qquad I_3(\mathbf{C}) = \det\,\mathbf{C}, \\ I_4(\mathbf{C}, \mathbf{M}) = \mathbf{C}\cdot\mathbf{M}, \qquad I_5(\mathbf{C}, \mathbf{M}) = \mathbf{C}^2\cdot\mathbf{M}. \end{gathered} \tag{6.102}$$

Remark 1: Observe that one can equivalently write

$$I_4 = \mathbf{C}\cdot\mathbf{M} = \mathbf{C}\mathbf{m}_{\rm R}\cdot\mathbf{m}_{\rm R} \qquad I_5 = \mathbf{C}^2\cdot\mathbf{M} = \mathbf{C}^2\mathbf{m}_{\rm R}\cdot\mathbf{m}_{\rm R}. \tag{6.103}$$

For a proof, see Chapter 5 of Steigmann [9].

Problem 6.4. In the case of a material involving two families of fibers, it was claimed in Section 6.2 that the strain energy function depended on (in addition to I_1, I_2, I_3) the invariants

$$I_4 = \mathbf{C}\mathbf{m}_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}, \quad I_5 = \mathbf{C}^2\mathbf{m}_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}, \quad I_6 = \mathbf{C}\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}'_{\mathrm{R}}, \quad I_7 = \mathbf{C}^2\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}'_{\mathrm{R}},$$

and the coupling term

$$I_8 = \mathbf{C}\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}.$$

In view of the forms of I_4, I_5, I_6, I_7 and I_8, one might expect the quantity $\mathbf{C}^2\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}$ to also appear in this list of invariants. Show that $\mathbf{C}^2\mathbf{m}'_{\mathrm{R}} \cdot \mathbf{m}_{\mathrm{R}}$ can be written in terms of the other invariants.

Problem 6.5. An incompressible body occupies a unit cube in the reference configuration with its edges parallel to the coordinate axes. The body contains one family of fibers in planes parallel to the x_1, x_2-plane, oriented at an angle Θ from the x_1-axis. The body is subjected to a uniform stress $\mathbf{T} = T\mathbf{e}_1 \otimes \mathbf{e}_1 + T_{33}\mathbf{e}_3 \otimes \mathbf{e}_3$. The value of T is given while that of T_{33} is such that the deformation is a plane strain in the x_1, x_2-plane, i.e. $\lambda_3 = 1$. Note that, in contrast to the problem considered in Section 6.1.1, here we have $T_{12} = 0$. Because of anisotropy, the deformation will involve a (to-be-determined) amount of shear k and so assume that the deformation has the form

$$\mathbf{y} = \mathbf{F}\mathbf{x} \qquad \text{where} \quad \mathbf{F} = \lambda_1\mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 + k\mathbf{e}_1 \otimes \mathbf{e}_2. \tag{i}$$

Assuming the material to be characterized by the strain energy function

$$W = \frac{\mu}{2}(I_1 - 3) + \frac{\mu\beta}{2}(I_4 - 1)^2, \qquad \mu > 0, \ \beta > 0. \tag{ii}$$

Derive two algebraic equations involving λ_1 and k as the only unknowns.

Now suppose the deformation is infinitesimal. Linearize the pair of equations you found and thus calculate the amount of shear k as a function of T (and μ, β and Θ).

Explore various limiting cases, e.g. $\beta \to 0, \beta \to \infty, \Theta \to 0$, etc. Assuming $T > 0$, when is $k > 0$ and when is it < 0?

Problem 6.6. In the reference configuration a *thin-walled* tube has mean radius R, wall thickness $T \ll R$ and length L. Its two ends are closed. The tube wall involves a single in-plane family of fibers. At each point of the tube wall, they lie in the Θ, Z-plane, oriented at an angle Φ with respect to the circumferential Θ-direction as shown in Figure 6.9:

$$\mathbf{m}_{\mathrm{R}} = \cos\Phi\,\mathbf{e}_\Theta + \sin\Phi\,\mathbf{e}_Z.$$

The tube is subjected to an internal pressure p and an axial force F. This loading will, in general, expand the tube to a mean radius r, elongate it to a length ℓ and rotate one end of the tube with respect to the other by an angle $\alpha\ell$. Assume the material to be incompressible and characterized by the strain energy function

$$W(I_1, I_4) = \frac{\mu}{2}(I_1 - 3) + \frac{\mu\beta}{2}(I_4 - 1)^2, \qquad \mu > 0, \ \beta > 0.$$

With $F = 0$, plot a graph of the twist angle α versus p. Does α vary monotonically with p (i.e. does the tube reverse its twist direction at some p)? Is there a value of F for which $\alpha = 0$?

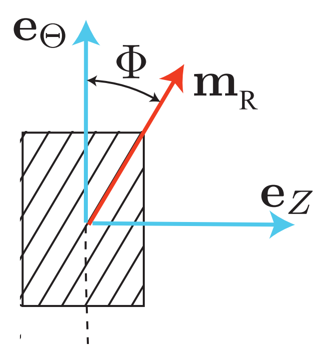


Figure 6.9: In the reference configuration the fiber direction $\mathbf{m}_\mathrm{R}$, locally, at each point in the tube wall, lies in the Θ, Z-plane as shown.

Problem 6.7. Consider a hollow (*thick-walled*) circular cylindrical tube with closed ends. It has inner radius A, outer radius B and length L in a reference configuration and is subjected to an internal pressure p, axial force F and twisting moment (torque) M. The tube is made of an incompressible material involving two families of fibers that lie, locally at each point, in the Θ, Z-plane, oriented at *different* angles Φ and Ψ with respect to the circumferential Θ-direction as shown in Figure 6.10. Formulate the problem and derive the equations to be solved to determine the resulting radial expansion, axial elongation and twist angle of the tube.

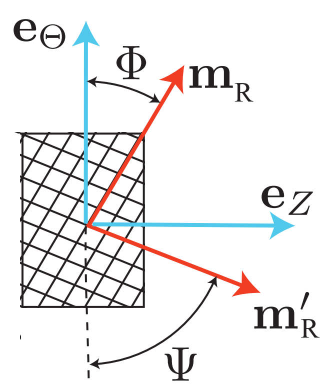


Figure 6.10: In the reference configuration the fiber directions $\mathbf{m}_\mathrm{R}$ and $\mathbf{m}'_\mathrm{R}$, locally, at each point in the tube, lie in the Θ, Z-plane as shown.

References:

1. J. L. Ericksen, and R.S. Rivlin, Large Elastic Deformations of Homogeneous Anisotropic Materials, *Journal of Rational Mechanics and Analysis*, 3(1954), issue 3, pp. 281-301.

2. A. Goriely, The Mathematics and Mechanics of Biological Growth, Springer, 2017, Chapter 11.

3. G. A. Holzapfel, *Nonlinear Solid Mechanics. A Continuum Approach for Engineering*, Wiley, 2000.

4. G.A. Holzapfel, T.C. Gasser and Ogden, R. W. A new constitutive framework for arterial wall mechanics and a comparative study of material models, *Journal of Elasticity*, 61(2000), pp. 1-48.

5. R.W. Ogden, Nonlinear Elasticity, Material Stability and Residual Stresses in Soft Tissue, in *Biomechanics of Soft Tissue in Cardiovascular Systems*, Edited by G. A. Holzapfel and R.W Ogden, CISM Courses and Lecture Series No. 441, Springer, 2003.

6. R. W. Ogden, Nonlinear Continuum Mechanics and Modeling the Elasticity of Soft Biological Tissues with a Focus on Artery Walls, in Biomechanics: Trends and Simulation, edited by G. A. Holzapfel and R. W. Ogden, Springer, 2017.

7. A.J.M. Spencer *Deformation of Fiber-Reinforced Materials*, Clarendon Press, Oxford, 1972.

8. A.J.M. Spencer, Constitutive theory for strongly anisotropic solids, in *Continuum Theory of the Mechanics of Fibre-Reinforced Composites*, edited by A.J.M. Spencer et al., CISM Courses and Lectures, No. 282, Springer-Verlag pp.1-32.

9. D. J. Steigmann, Finite Elasticity Theory, Oxford, 2017, Chapter 5.

Chapter 7

A Two-Phase Elastic Material: An Example.

7.1 A material with cubic and tetragonal phases.

We shall use the terms "energy-well" and "local minimum" interchangeably in this chapter and so we will say that the strain energy function $W(\mathbf{C})$ has an **energy-well** at $\mathbf{C} = \mathbf{C}_*$ if $W(\mathbf{C})$ has a local minimum at $\mathbf{C} = \mathbf{C}_*$:

$$\left.\frac{\partial W}{\partial C_{ij}}\right|_{\mathbf{C}=\mathbf{C}_*} = 0, \qquad \left.\frac{\partial^2 W}{\partial C_{ij}\partial C_{k\ell}}\right|_{\mathbf{C}=\mathbf{C}_*} H_{ij}H_{k\ell} > 0, \tag{7.1}$$

for all symmetric tensors $\mathbf{H} \neq \mathbf{0}$. Note from the constitutive relation for stress and $(7.1)_1$ that the body is necessarily stress-free at an energy-well. The particular strain energy functions considered previously had energy-wells at $\mathbf{C} = \mathbf{I}$ corresponding to the reference configuration.

In this section we construct a strain energy function with *multiple* energy-wells, i.e. multiple local minima. Each energy-well describes a "phase" of the material and multi-well strain energy functions describe materials that can exist in more than one phase. It should be noted that the multiple phases here are all solid phases. Figure 7.1 depicts a particular crystalline solid that has a cubic lattice in one phase and a tetragonal lattice in another phase as is the case for certain alloys of In-Tl, Mn-Ni and Mn-Cu.

Temperature plays an important role in the study of such materials. In our present discussion we hold the temperature fixed at the so-called *transformation temperature*. At this temperature, the heights of the different energy-wells are the same, i.e. the values of W

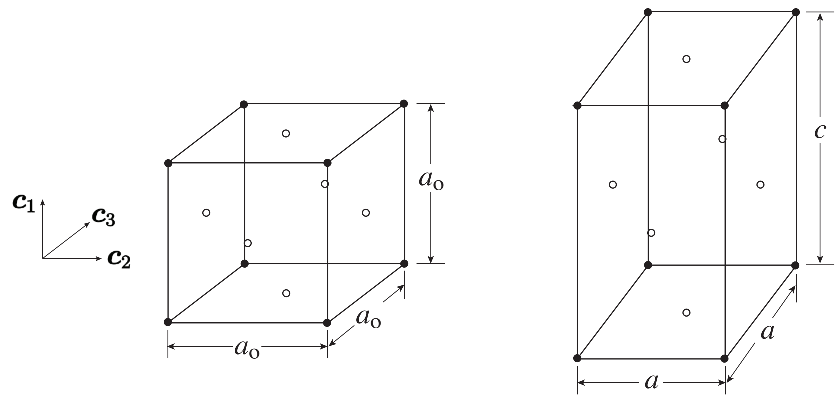


Figure 7.1: One phase has a face-centered cubic lattice with lattice parameters $a_o \times a_o \times a_o$, the other a face-centered tetragonal lattice with lattice parameters $a \times a \times c$. The unit vectors $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ are associated with the cubic directions. Solely for purposes of clarity, the atoms at the vertices are depicted by filled dots while those at the centers of the faces are shown by open dots.

are the same at all local minima. Thus we might say that there is no preferred phase at this temperature. Above the transformation temperature, one phase will have lower energy than the other, while this will switch below the transformation temperature. This is illustrated in Figure 7.5 where, at high temperatures the energy-well at $\varepsilon = 0$ has lower energy than the other two energy-wells, while the reverse is true at low temperatures. The temperature-dependent version of the material in this chapter can be found in Abeyaratne et al. [1].

Consider a material that can exist in two phases, a cubic phase and a tetragonal phase. Let $a_0 \times a_0 \times a_0$ and $a \times a \times c$ be the respective lattice parameters of the cubic and tetragonal lattices, and let $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ denote fixed unit vectors in the cubic directions. The deformation from the cubic phase to the tetragonal phase is achieved by stretching the cubic lattice equally in two of the cubic directions and unequally in the third direction, see Figure 7.1, the associated stretches being α, α and β:

$$\alpha = \frac{a}{a_0} > 1, \qquad \beta = \frac{c}{a_0}.$$

Since we can impose the unequal stretch β on any one of the $\mathbf{c}_k$-directions, $k = 1, 2, 3$, (and the stretch α on the remaining two cubic directions), there are three ways in which to stretch the cubic lattice into the tetragonal lattice as depicted in Figure 7.2. We say there are three *variants* of the tetragonal phase. With the cubic phase taken to be the reference configuration, the three variants are characterized by the stretch tensors $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3$ whose components in the basis $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ are

$$[U_1] = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}, \qquad [U_2] = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \alpha \end{pmatrix}, \qquad [U_3] = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}. \tag{7.2}$$

It will be useful for subsequent calculations to note that

$$\mathbf{U}_1 \;=\; \beta\, \mathbf{c}_1 \otimes \mathbf{c}_1 + \alpha\, \mathbf{c}_2 \otimes \mathbf{c}_2 + \alpha\, \mathbf{c}_3 \otimes \mathbf{c}_3. \tag{7.3}$$

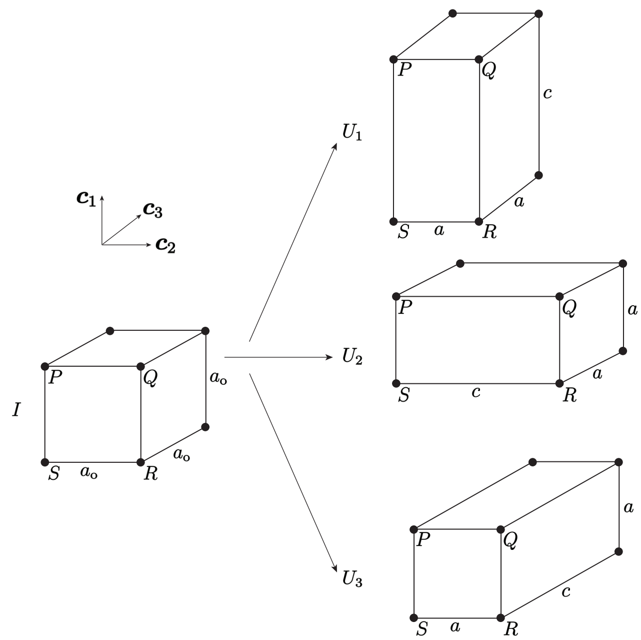


Figure 7.2: Cubic phase and three variants of the tetragonal phase. Observe that one *cannot* rigidly rotate one tetragonal variant in such a way as to make the "atoms" P,Q,R,S coincide with the locations of these same atoms in a different tetragonal variant.

We now construct an explicit strain-energy function $W(\mathbf{C})$ that can be used to describe such a material. It therefore has local minima at $\mathbf{C} = \mathbf{I}, \mathbf{U}_1^2, \mathbf{U}_2^2$ and $\mathbf{U}_3^2$. According to the claim in the exercise below, a function $W(\mathbf{C})$ that has an energy-well at $\mathbf{C} = \mathbf{U}_1^2$, will automatically have energy-wells at $\mathbf{C} = \mathbf{U}_2^2$ and $\mathbf{U}_3^2$ in view of material symmetry.

Exercise: Show using material symmetry, that if $W(\mathbf{C})$ has a local minimum at any one of the tensors $\mathbf{C}_k$, then W will automatically have energy-wells at the remaining two $\mathbf{C}_k$'s where $\mathbf{C}_k = \mathbf{U}_k^2$, $k = 1, 2, 3$.

Following Ericksen [2, 3, 4] we write the strain-energy function as a function of the Green Saint-Venant strain tensor $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$. It must have energy-wells at $\mathbf{E} = \mathbf{0}, \mathbf{E}_1, \mathbf{E}_2$ and

$\mathbf{E}_3$ where $\mathbf{E}_k = \frac{1}{2}(\mathbf{C}_k - \mathbf{I}) = \frac{1}{2}(\mathbf{U}_k^2 - \mathbf{I})$. Since we have taken the reference configuration to coincide with the cubic phase, the strain-energy function $W(\mathbf{E})$ (with respect to that configuration), must possess cubic symmetry. It is known (see, for example, Smith and Rivlin [5] and Green and Adkins [6]) that, to have cubic symmetry, W must be a function of the "cubic invariants" $i_k(\mathbf{E})$:

$$\begin{aligned} &i_1 = E_{11} + E_{22} + E_{33}, && i_2 = E_{11}E_{22} + E_{22}E_{33} + E_{33}E_{11}, \\ &i_3 = E_{11}E_{22}E_{33}, && i_4 = E_{12}E_{23}E_{31}, \\ &i_5 = E_{12}^2 + E_{23}^2 + E_{31}^2, && \ldots \text{etc.}, \end{aligned} \tag{7.4}$$

where E_{ij} refers to the i,j-component of $\mathbf{E}$ in the cubic basis $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$. The number of invariants in this list depends on the particular cubic sub-class under consideration (see, e.g., Section 1.11 of Green and Adkins [6]) but the analysis below is valid for all of them.

For temperatures close to the transformation temperature, Ericksen [2, 3, 4] has argued based on experimental observations that, as a first approximation, (a) all of the shear strain components (in the cubic basis) vanish and (b) the sum of the normal strains also vanishes. Accordingly he suggested a kinematically constrained theory based on the two constraints

$$i_1 = E_{11} + E_{22} + E_{33} = 0, \qquad i_5 = E_{12}^2 + E_{23}^2 + E_{31}^2 = 0. \tag{7.5}$$

In this case, the only two nontrivial strain invariants among those in the preceding list are i_2 and i_3 and so we take $W = W(i_2, i_3)$. The constraint $i_1 = 0$ can be written as $E_{33} = -E_{11} - E_{22}$ and so we can eliminate E_{33} from i_2 and i_3 to obtain

$$i_2 = -E_{11}^2 - E_{11}E_{22} - E_{22}^2, \qquad i_3 = -E_{11}^2E_{22} - E_{11}E_{22}^2. \tag{7.6}$$

Note that the constraint $i_1 = 0$ implies $\operatorname{tr} \mathbf{E} = 0$ and so in particular $\operatorname{tr} \mathbf{E}_1 = 0$. Therefore the lattice stretches α and β must be related by $2\alpha^2 + \beta^2 = 3$.

Let us start by considering a one-phase material involving only the cubic phase. Then the only local minimum of $W(\mathbf{E})$ is at $\mathbf{E} = \mathbf{0}$. The simplest form of W to consider is a polynomial in the components of $\mathbf{E}$, and in order to have one local minimum its degree must be (at least) quadratic. Thus consider

$$W(\mathbf{E}) = c_0 + c_2 i_2, \tag{7.7}$$

where c_0 and c_2 are constants and the invariant i_2 is given by $(7.6)_1$. The first derivatives of $W(E_{11}, E_{22})$ are

$$\frac{\partial W}{\partial E_{11}} = -c_2(2E_{11} + E_{22}), \qquad \frac{\partial W}{\partial E_{22}} = -c_2(2E_{22} + E_{11}), \tag{7.8}$$

and they vanish at $\mathbf{E} = \mathbf{0}$. Thus W has an extremum at $\mathbf{E} = \mathbf{0}$. To ensure that this is a local minimum, we calculate the second derivatives of $W(E_{11}, E_{22})$:

$$\frac{\partial^2 W}{\partial E_{11}^2} = -2c_2, \qquad \frac{\partial^2 W}{\partial E_{22}^2} = -2c_2, \qquad \frac{\partial^2 W}{\partial E_{11}\partial E_{22}} = -c_2. \tag{7.9}$$

The extremum is a local minimum if the Hessian matrix

$$\begin{pmatrix} W_{11} & W_{12} \\ W_{12} & W_{22} \end{pmatrix} \qquad \text{where } W_{ij} = \frac{\partial^2 W}{\partial E_{ii}\partial E_{jj}} \text{ (no sum on } i \text{ and } j), \tag{7.10}$$

evaluated at $\mathbf{E} = \mathbf{0}$ is positive definite. It is easily seen that this requires $c_2 < 0$. Thus the strain energy function (7.7), $(7.6)_1$ with $c_2 < 0$ has a local minimum at $\mathbf{E} = \mathbf{0}$ (and this local minimum is unique). A contour plot of this energy is shown in Figure 7.3. Note the energy-well at $\mathbf{E} = \mathbf{0}$.

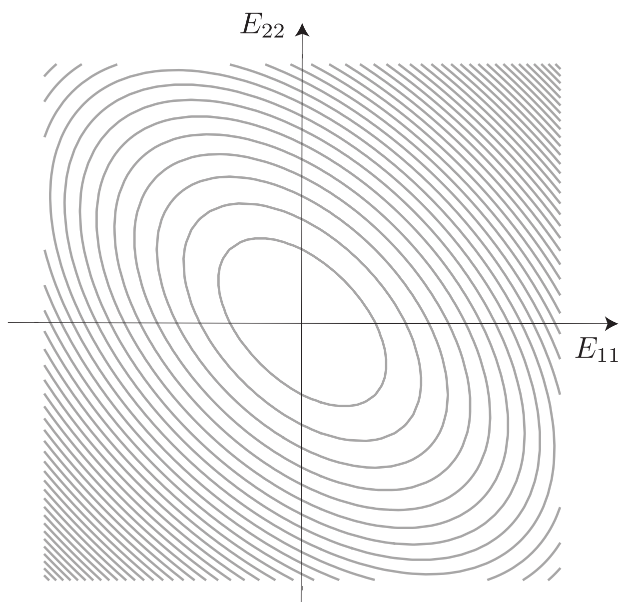


Figure 7.3: Constant energy contours on the E_{11}, E_{22}-plane for the one-phase strain energy function (7.7), $(7.6)_1$ with $c_2 < 0$. Observe the single energy-well at $\mathbf{E} = \mathbf{0}$.

We now return to the two-phase material of interest. This energy function must have one energy well at $\mathbf{E} = \mathbf{0}$ (the cubic phase) and another at $\mathbf{E} = \mathbf{E}_1$ (the tetragonal phase). Here

$$\mathbf{E}_1 = \frac{1}{2}(\mathbf{U}_1^2 - \mathbf{I}) \overset{(7.3)}{=} \frac{1}{2}(\beta^2 - 1)\,\mathbf{c}_1 \otimes \mathbf{c}_1 + \frac{1}{2}(\alpha^2 - 1)\,\mathbf{c}_2 \otimes \mathbf{c}_2 + \frac{1}{2}(\alpha^2 - 1)\,\mathbf{c}_3 \otimes \mathbf{c}_3. \tag{7.11}$$

The simplest form of W to consider is again a polynomial in the components of $\mathbf{E}$, but in order to endow it with two distinct local minima, its degree must be at least quartic. It follows from (7.4) and (7.5) that the most general quartic polynomial of the form $W = W(i_2, i_3)$ is

$$W = c_0 + c_2 i_2 + c_3 i_3 + c_{22} i_2^2, \tag{7.12}$$

where the c_i's are constants[1]. From (7.6) and (7.12), the first and second partial derivatives of $W(E_{11}, E_{22})$ with respect to the strain components are

$$\begin{aligned}
\frac{\partial W}{\partial E_{11}} &= -c_2(2E_{11} + E_{22}) - c_3(2E_{11}E_{22} + E_{22}^2) + 2c_{22}(E_{11}^2 + E_{11}E_{22} + E_{22}^2)(2E_{11} + E_{22}),\\
\frac{\partial W}{\partial E_{22}} &= -c_2(2E_{22} + E_{11}) - c_3(2E_{11}E_{22} + E_{11}^2) + 2c_{22}(E_{11}^2 + E_{11}E_{22} + E_{22}^2)(2E_{22} + E_{11}),
\end{aligned} \tag{7.13}$$

$$\begin{aligned}
\frac{\partial^2 W}{\partial E_{11}^2} &= -2c_2 - 2c_3E_{22} + 6c_{22}(2E_{11}^2 + 2E_{11}E_{22} + E_{22}^2),\\
\frac{\partial^2 W}{\partial E_{22}^2} &= -2c_2 - 2c_3E_{11} + 6c_{22}(2E_{22}^2 + 2E_{11}E_{22} + E_{11}^2),\\
\frac{\partial^2 W}{\partial E_{11}\partial E_{22}} &= -c_2 - 2c_3(E_{11} + E_{22}) + 6c_{22}(E_{11}^2 + 2E_{11}E_{22} + E_{22}^2).
\end{aligned} \tag{7.14}$$

Of the multiple local minima of W, one must be at $\mathbf{E} = \mathbf{0}$ corresponding to the cubic phase. It is seen immediately from (7.13) that the first derivatives of $W(E_{11}, E_{22})$ vanish automatically at $\mathbf{E} = \mathbf{0}$. Evaluating the Hessian matrix at $\mathbf{E} = \mathbf{0}$ shows that it is positive definite provided

$$c_2 < 0. \tag{7.15}$$

A second local minimum of $W(\mathbf{E})$ is at the strain $\mathbf{E}_1$ given by (7.11). A straightforward calculation based on substituting (7.11) into (7.13) shows that the first derivatives of $W(E_{11}, E_{22})$ vanish at $\mathbf{E}_1$ provided

$$c_2 = -pc_3 + 6p^2 c_{22}, \tag{7.16}$$

where we have set

$$p := \frac{1}{2}(\alpha^2 - 1) > 0. \tag{7.17}$$

In order to ensure that this extremum is a local minimum, we evaluate the second derivatives of W at $\mathbf{E} = \mathbf{E}_1$ using (7.14). The positive definiteness of the associated Hessian matrix is found to require

$$12pc_{22} > c_3 > 0. \tag{7.18}$$

[1] When thermal effects are taken into account the c_i's will be functions of temperature.

Finally, since we are working at the transformation temperature where the heights of the energy-wells are the same, we need $W(\mathbf{0}) = W(\mathbf{E}_1)$. Calculating these two values of the energy from (7.12) and equating them leads to

$$c_3 = 9pc_{22}. \tag{7.19}$$

Combining all of the preceding requirements (7.15), (7.16), (7.18) and (7.19) leads to the strain energy function

$$W = c_0 + c_{22}\Big[-3p^2 i_2 + 9p i_3 + i_2^2 \Big], \tag{7.20}$$

where

$$c_{22} > 0, \qquad p = \frac{1}{2}(\alpha^2 - 1). \tag{7.21}$$

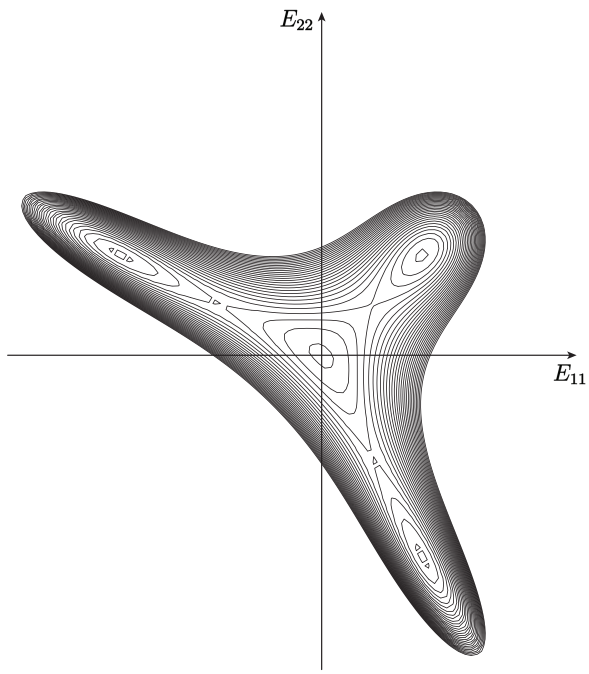


Figure 7.4: Constant energy contours on the E_{11}, E_{22}-plane for the strain energy function (7.20), (7.21). Observe the presence of the cubic phase energy-well at $\mathbf{E} = \mathbf{0}$ surrounded by the three tetragonal phase energy-wells at $\mathbf{E} = \mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$.

Figure 7.4 shows the equal energy contours of the strain-energy function (7.20), (7.21) on the E_{11}, E_{22}-plane. It shows the cubic energy-well at the origin $\mathbf{E} = \mathbf{0}$ surrounded by the three tetragonal energy-wells at $\mathbf{E} = \mathbf{E}_1, \mathbf{E}_2$ and $\mathbf{E}_3$. The values of the material parameters

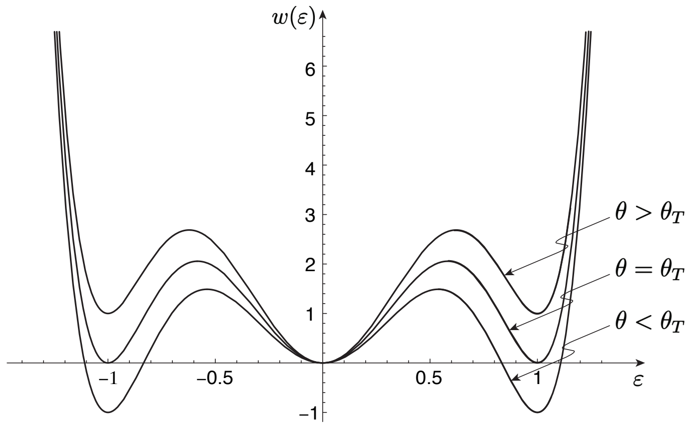


Figure 7.5: A one-dimensional cross-section of the energy (7.20), (7.21) along the path (7.22) in strain-space. The figure plots $w(\varepsilon) = W(\widehat{\mathbf{E}}(\varepsilon))$ versus strain ε at three constant temperatures. The analysis in these notes were limited to the transformation temperature θ_T where the wells have the same height. We have included the corresponding plots, both above and below the temperature θ_T, taken from [1]. The material parameters underlying this plot are the same as those associated with the previous figure.

associated with this figure were chosen solely on the basis of obtaining a fairly clear contour plot.

In order to draw a one-dimensional cross-section of the energy (7.20), (7.21) consider the following path in strain space

$$\mathbf{E} = \widehat{\mathbf{E}}(\varepsilon) = p\varepsilon^2 \mathbf{c}_1 \otimes \mathbf{c}_1 + \frac{1}{2} p\varepsilon(3-\varepsilon)\mathbf{c}_2 \otimes \mathbf{c}_2 - \frac{1}{2} p\varepsilon(3+\varepsilon)\mathbf{c}_3 \otimes \mathbf{c}_3, \qquad -1.5 < \varepsilon < 1.5, \quad (7.22)$$

where ε is a parameter. Observe that $\widehat{\mathbf{E}}(0) = \mathbf{0}$, $\widehat{\mathbf{E}}(-1) = \mathbf{E}_2, \widehat{\mathbf{E}}(1) = \mathbf{E}_3$ so that this path passes through the cubic well and two of the tetragonal wells. The energy along this path is $w(\varepsilon) := W(\widehat{\mathbf{E}}(\varepsilon))$ and Figure 7.5 shows a plot of $w(\varepsilon)$ versus ε (the curve labelled $\theta = \theta_T$). The calculations in this chapter were carried out at the transformation temperature ($\theta = \theta_T$) where the three energy-wells have the same height. Figure 7.5 taken from Abeyaratne et al. [1] shows the results for two other temperatures as well, one $> \theta_T$ and the other $< \theta_T$.

The strain-energy function (7.20) captures the key qualitative characteristics of a material that exists in cubic and tetragonal phases. However, due to the restrictive nature of the kinematic constraints (7.5), it fails to provide a *quantitatively* accurate model. The natural generalization of (7.20) is therefore to relax these constraints by using Lagrange multipliers

c_5 and c_{11} and to replace (7.12) by

$$W = c_0 + c_2 I_2 + c_3 I_3 + c_{22} I_2^2 + c_5 I_5 + c_{11} I_1^2. \tag{7.23}$$

R.D. James (unpublished[2]) has shown that the response predicted by this generalized form of W is in reasonable quantitative agreement with the observed behavior of In-Tl.

Remark: A strain energy function for a Cu-Al-Ni alloy can be found in Vedantam and Abeyaratne [8]. This alloy can exist in a cubic phase as well as *six* variants of a orthorhombic phase.

Remark: In Problem 2.31 you looked at the kinematics of a piecewise homogeneous deformation that involved the cubic phase on one side of a planar interface and a tetragonal variant on the other.

References:

1. R. Abeyaratne, K. Bhattacharya and J.K. Knowles, , in *Nonlinear Elasticity: Theory and Applications*, Edited by Y.B. Fu and R.W Ogden, London Mathematical Society Lecture Note Series 283, Cambridge University Press, 2001.

2. J.L. Ericksen, On the symmetry and stability of thermoelastic solids. *ASME Journal of Applied Mechanics* **45**(1978), 740–744.

3. J.L. Ericksen, Some phase transitions in crystals. *Archive for Rational Mechanics and Analysis* **73**(1980), 99–124.

4. J.L. Ericksen, Constitutive theory for some constrained elastic crystals. *International Journal of Solids Structures* **22**(1986), 951–964.

5. G.F. Smith and R.S. Rivlin, The strain energy function for anisotropic elastic materials. *Transactions of the American Mathematical Society* **88**(1958), 175–193.

6. A.E. Green and J.E. Adkins, *Large Elastic Deformations*. Oxford University Press, 1970.

7. P. Kloućek and M. Luskin, The computation of the dynamics of the martensitic transformation. *Continuum Mech. Thermodyn.* **6**(1994), 209–240.

[2]See Kloućek and Luskin [7].

8. S. Vedantam and R. Abeyaratne, A Helmholtz free-energy function for a Cu-Al-Ni shape memory alloy, *International Journal of Nonlinear Mechanics*, 40(2005), pp. 177-193.

Chapter 8

A Micromechanical Constitutive Model

Continuum theory says that an elastic material is characterized by a strain energy function $W(\mathbf{C})$. If additional information on material symmetry is available, this can be reduced further, for example to the form $W(I_1(\mathbf{C}), I_2(\mathbf{C}), I_3(\mathbf{C}))$ for an isotropic material. However, that is as far as the theory goes. The examples of explicit functions W given in Section 4.7 (corresponding, for example, to the Blatz-Ko or Fung models) are "phenomenological models" of particular elastic materials, i.e. the functional form of W is laid out at the outset at the continuum level, and subsequent laboratory experiments are used to refine it.

On the other hand the macroscopic (or continuum) behavior of a material reflects its underlying microscopic behavior. If one could describe the processes at the microscopic scale, and knew how to homogenize them across scales, one could then infer the response at the macroscopic scale. When this is possible, the continuum model so developed captures the microscopic physics.

In this chapter we start at the atomistic scale and develop an explicit form for the strain energy function $W(\mathbf{C})$ for a crystalline solid. The atomistic model we use is the simplest conceivable one, and our purpose is *merely to illustrate* how one might develop continuum models from microscale models.

A second example that rightfully belongs here is the derivation of the strain energy function for rubber-like materials based on a polymer chain model. Unfortunately, the strain energy of such materials turns out to be dominated by its entropy, and since we have not considered thermodynamics in these notes (at least not beyond the brief discussion in Section 9.2), we are not able to describe those calculations here, not without a lot of preliminary work. The interested reader may look at Volume II in this series on notes.

8.1 Example: Lattice Theory of Elasticity.

The notes in this section closely follow the unpublished lecture notes of Professor Kaushik Bhattacharya of Caltech. I am most grateful to him for sharing them with me. The original calculations are due to Cauchy, see Love [3].

The aim of this section is to illustrate how a simple atomistic model of a crystalline solid can be used to *derive* explicit continuum scale constitutive response functions $\widehat{\mathbf{T}}$ and $\widehat{W}$ for the Cauchy stress and the strain energy function in terms of the deformation gradient tensor. We will see that the expressions to be derived *automatically satisfy the requirements of material frame indifference, material symmetry and the dissipation inequality.* Moreover we find that the traction - stress relation $\mathbf{t} = \mathbf{Tn}$ and the symmetry condition $\mathbf{T} = \mathbf{T}^T$ hold automatically. The expressions for $\widehat{\mathbf{T}}$ and $\widehat{W}$ that we derive are explicit in terms of the lattice geometry and the interatomic force potential; see (8.11) and (8.16).

8.1.1 A Bravais Lattice. Pair Potential.

A Bravais lattice $\mathcal{L}$ is an infinite set of points in $\mathbb{R}^3$ generated by translating a point $\mathbf{y}_o$ through three linearly independent vectors $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$: i.e.,

$$\mathcal{L}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3) = \left\{ \mathbf{y} : \ \mathbf{y} \in \mathbb{R}^3, \ \mathbf{y} = \mathbf{y}_o + \nu_i \boldsymbol{\ell}_i \quad \text{for all integers } \nu_1, \nu_2, \nu_3 \ \right\} \ . \tag{8.1}$$

The *lattice vectors* $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ define a *unit cell.* Note the distinction between the lattice $\mathcal{L}$, which is an infinite set of periodically arranged points in space, and the lattice vectors. In particular, it is generally possible to generate the same lattice $\mathcal{L}$ from more than one set of lattice vectors, i.e., a given set of lattice vectors generates a unique lattice, but the converse is not necessarily true. More on this later. We shall take the orientation of the lattice vectors to be right-handed so that in particular, the volume of the unit cell is

$$\text{vol}\,(\text{unit cell}) = (\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) \cdot \boldsymbol{\ell}_3 \ > \ 0. \tag{8.2}$$

The neighborhood of any lattice point, say $\mathbf{y}_A$, is identical to that of any other lattice point, say $\mathbf{y}_C$. To see this we simply note that if $\mathbf{y}_B$ is any third lattice point, then there necessarily is a fourth lattice point $\mathbf{y}_D$ such that $\mathbf{y}_D - \mathbf{y}_C = \mathbf{y}_B - \mathbf{y}_A$. Thus the position of $\mathbf{y}_B$ relative to $\mathbf{y}_A$ is the same as the position of $\mathbf{y}_D$ relative to $\mathbf{y}_C$. Any two lattice points $\mathbf{y}_A$ and $\mathbf{y}_C$ of a Bravais lattice are therefore *geometrically equivalent*. Bravais lattices can represent only monoatomic lattices; in particular, no alloy is a Bravais lattice[1].

We will ignore lattice vibrations and assume that the atoms are located at the lattice points. Therefore the calculations we carry out are valid at zero degrees Kelvin.

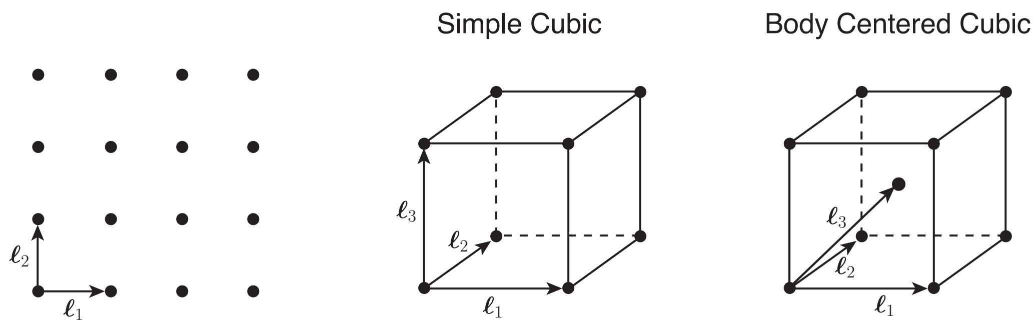


Figure 8.1: Examples of lattices in $\mathbb{R}^2$ and $\mathbb{R}^3$.

In the simplest model of interatomic interactions one assumes the existence of a *pair potential* $\phi(\rho)$ such that the force exerted by atom A on atom B, say $\mathbf{f}_{A,B}$, is the gradient of this potential:

$$\mathbf{f}_{A,B} = -\boldsymbol{\nabla}_y \phi(|\mathbf{y}|)\Big|_{\mathbf{y}=\mathbf{y}_B-\mathbf{y}_A} = -\phi'(|\mathbf{y}_B - \mathbf{y}_A|)\ \frac{\mathbf{y}_B - \mathbf{y}_A}{|\mathbf{y}_B - \mathbf{y}_A|}\ . \tag{8.3}$$

In this model the force exerted by one atom on the other depends solely on the relative positions of *those* two atoms and is independent of the positions of the surrounding atoms. Note that the force (8.3) is a *central force* in that it acts along the line joining those two atoms. Also observe that if the distance ρ between the atoms is such that $\phi'(\rho) < 0$, then the force between them is repulsive; if $\phi'(\rho) > 0$ it is attractive. Finally, observe from (8.3) that $\mathbf{f}_{A,B} = -\mathbf{f}_{B,A}$ so that the force exerted by atom A on atom B is equal in magnitude and opposite in direction to the force applied by atom B on atom A.

Figure 8.2 shows a graph of a typical pair-potential $\phi(\rho)$ versus the distance ρ between the pair of atoms. Note that the associated force is repulsive at short distances ($< \rho_o$) and attractive at large distances ($> \rho_o$).

[1]Even some monoatomic lattices – e.g. a hexagonal close-packed lattice – cannot be represented as a Bravais lattice.

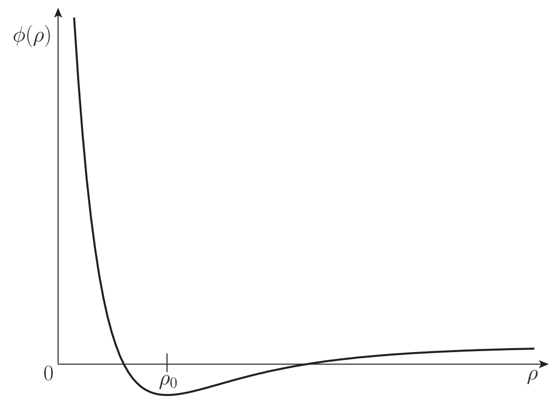


Figure 8.2: Typical graph of the pair-potential ϕ.

Several of the calculations to follow will involve infinite sums of terms involving ϕ, ϕ' and ϕ'' over all lattice points; it is necessary to ensure that these sums converge to finite values. This requires that $\phi(\rho) \to 0$ fast enough as $\rho \to \infty$. We assume that ϕ possesses the requisite[2] decay rate.

Because of the periodicity and symmetry of a Bravais lattice, if $\mathbf{y}_A$ and $\mathbf{y}_B$ are any two lattice points, there necessarily is a third lattice point $\mathbf{y}_C$ which is such that $\mathbf{y}_A - \mathbf{y}_B = -(\mathbf{y}_A - \mathbf{y}_C)$. Therefore according to the force law (8.3), the forces exerted on atom A by atoms B and C are equal in magnitude and opposite in direction. Consequently for each atom B that exerts a force on atom A, there is another atom C that exerts an equal and opposite force on A. Thus a Bravais lattice is *always in equilibrium.*

8.1.2 Homogenous Deformation of a Bravais Lattice.

Most, but not all, of the discussion to follow will be carried out entirely on the deformed lattice. There will however be a few occasions when we wish to consider a reference lattice

[2]To determine the required decay rate, one can consider a sphere of radius, say $\overline{\rho}$, and separate the infinite sum over the entire lattice into a finite sum over the finite number of lattice points in the interior of the sphere plus a sum over the infinite number of lattice points in the exterior of the sphere. An upper bound for the second term can then be written by replacing the sum by an integral (over the entire three dimensional region exterior to the sphere). Convergence of the integral guarantees convergence of the sum. For example the energy (8.16) will converge if the integral of $\rho^2\phi(\rho)$ over the interval $[\overline{\rho}, \infty)$ converges, which would be true if $\phi \to 0$ faster than ρ^{-3} as $\rho \to \infty$.

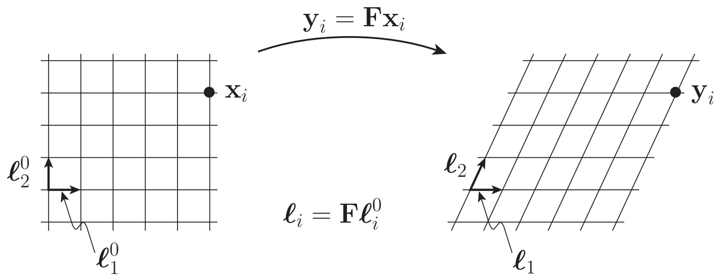


Figure 8.3: Homogeneous deformation of a lattice. The lattice vectors $\{\boldsymbol{\ell}_1^o, \boldsymbol{\ell}_2^o\}$ of the reference lattice are mapped by $\mathbf{F}$ into the lattice vectors $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2\}$ of the deformed lattice.

and for this purpose we consider a second Bravais lattice $\mathcal{L}_0$:

$$\mathcal{L}(\boldsymbol{\ell}_1^o, \boldsymbol{\ell}_2^o, \boldsymbol{\ell}_3^o) = \left\{ \mathbf{x} : \ \mathbf{x} \in \mathbb{R}^3, \ \mathbf{x} = \mathbf{x}_o + \nu_i \boldsymbol{\ell}_i^o \quad \text{for all integers } \nu_1, \nu_2, \nu_3 \ \right\}$$

where the lattice vectors $\{\boldsymbol{\ell}_1^o, \boldsymbol{\ell}_2^o, \boldsymbol{\ell}_3^o\}$ define a unit cell of the reference lattice. Since each set of lattice vectors is linearly independent, there is a nonsingular tensor $\mathbf{F}$ that maps $\{\boldsymbol{\ell}_1^o, \boldsymbol{\ell}_2^o, \boldsymbol{\ell}_3^o\}$ $\to \{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$:

$$\boldsymbol{\ell}_i = \mathbf{F}\boldsymbol{\ell}_i^o, \quad i = 1, 2, 3. \tag{8.4}$$

This is illustrated in Figure 8.3.

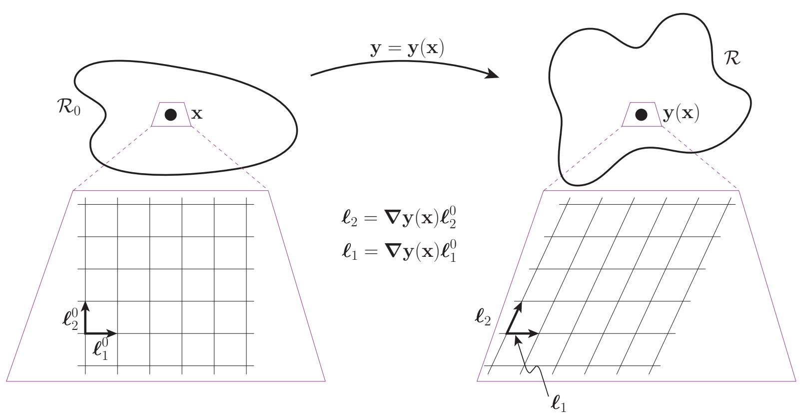


Figure 8.4: The deformation $\mathbf{y}(\mathbf{x})$ carries the three dimensional region $\mathcal{R}_0$ into $\mathcal{R}$. The figure shows blown-up views of infinitesimal neighborhoods of $\mathbf{x}$ and $\mathbf{y}(\mathbf{x})$. The mapping of the lattice vectors is assumed to be described by Grad $\mathbf{y}(\mathbf{x})$ $(= \nabla \mathbf{y}(\mathbf{x}))$ as depicted in the figure.

Suppose that we associate a (continuum) body with the lattice. The lattices $\mathcal{L}_0$ and $\mathcal{L}$ are associated with two configurations of the body. Let $\mathbf{y}(\mathbf{x})$ be the deformation of the continuum that maps $\mathcal{R}_0$ into $\mathcal{R}$. The deformation gradient tensor is $\nabla\mathbf{y}(\mathbf{x})$. As discussed in Section 2.3, $\nabla\mathbf{y}(\mathbf{x})$ maps material fibers of the continuum from the reference to the deformed configurations. The tensor $\mathbf{F}$ introduced above maps the reference lattice vectors to the deformed lattice vectors through (8.4). The Cauchy-Born hypothesis (rule) states that the "continuum deforms with the lattice" in the sense that $\nabla\mathbf{y}(\mathbf{x}) = \mathbf{F}$. This is illustrated in Figure 8.4.

8.1.3 Traction and Stress.

We now establish a notion of traction and then derive an explicit expression for it in terms of the interatomic forces. Let $\mathcal{P}$ be an arbitrary plane through the lattice and let $\mathbf{n}$ denote a unit vector normal to $\mathcal{P}$. Let $\mathcal{L}^+$ and $\mathcal{L}^-$ denote the two subsets of the lattice $\mathcal{L}$ that are on, respectively, the side into which and the side away from which $\mathbf{n}$ points; see Figure 8.5. Let $\mathcal{A}$ be a subregion of the plane $\mathcal{P}$. Consider two lattice points $\mathbf{y}_+ \in \mathcal{L}^+$ and $\mathbf{y}_- \in \mathcal{L}^-$ such that the line joining them intersects the subregion $\mathcal{A}$; see Figure 8.5. By summing the forces between all such pairs of atoms, we can associate a force with the region $\mathcal{A}$. The traction $\mathbf{t}$ can then be defined as the normalization of this force by the area of $\mathcal{A}$:

$$\mathbf{t}(\mathcal{A}) = \frac{1}{\text{area }(\mathcal{A})}\sum \mathbf{f}_{i,j} \;=\; \frac{1}{\text{area }(\mathcal{A})}\sum -\phi'(|\mathbf{y}_- - \mathbf{y}_+|)\,\frac{\mathbf{y}_- - \mathbf{y}_+}{|\mathbf{y}_- - \mathbf{y}_+|}, \tag{8.5}$$

where the summation is carried out over all $\mathbf{y}_+ \in \mathcal{L}^+$ and $\mathbf{y}_- \in \mathcal{L}^-$ which are such that the line joining $\mathbf{y}_+$ to $\mathbf{y}_-$ intersects $\mathcal{A}$.

For (8.5) to be useful, we need to characterize the range of summation in a simpler form. First, since $\mathbf{y}_-$ and $\mathbf{y}_+$ are lattice points, it follows that there are integers $\{\nu_1, \nu_2, \nu_3\}$ for which $\mathbf{y}_- - \mathbf{y}_+ = \nu_i\boldsymbol{\ell}_i$. Conversely, given any three integers $\{\nu_1, \nu_2, \nu_3\}$ which are such that $(\nu_i\boldsymbol{\ell}_i)\cdot\mathbf{n} < 0$ (which simply means that the vector $\nu_i\boldsymbol{\ell}_i$ points in the $-\mathbf{n}$ direction), there exist pairs (note plural) of lattice points $\mathbf{y}_+ \in \mathcal{L}^+$ and $\mathbf{y}_- \in \mathcal{L}^-$ such that $\mathbf{y}_- - \mathbf{y}_+ = \nu_i\boldsymbol{\ell}_i$; of these, the number of pairs whose line of connection intersects $\mathcal{A}$ can be estimated to be

$$\begin{aligned} N &= \frac{\text{volume of the (non-prismatic) cylinder with base } \mathcal{A} \text{ and generator } \nu_i\boldsymbol{\ell}_i}{\text{volume of the unit cell}} \\ &= \frac{\text{area }(\mathcal{A})\,|(\nu_i\boldsymbol{\ell}_i)\cdot\mathbf{n}|}{(\boldsymbol{\ell}_1\times\boldsymbol{\ell}_2)\cdot\boldsymbol{\ell}_3} = -\text{area }(\mathcal{A})\frac{(\nu_i\boldsymbol{\ell}_i)\cdot\mathbf{n}}{(\boldsymbol{\ell}_1\times\boldsymbol{\ell}_2)\cdot\boldsymbol{\ell}_3}, \end{aligned} \tag{8.6}$$

when the area of $\mathcal{A}$ is sufficiently large[3]; see Figure 8.6. In the last step we have used the

[3]In a homogeneously deformed continuum, the traction on the plane $\mathcal{P}$ would be uniform, i.e. it would

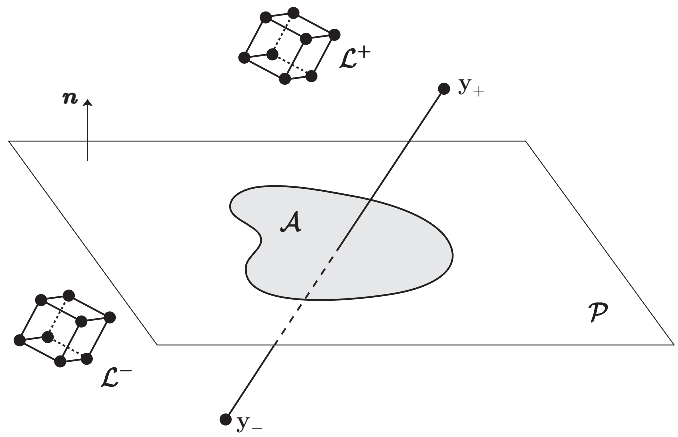


Figure 8.5: A plane $\mathcal{P}$ separating the lattice into two parts $\mathcal{L}^+$ and $\mathcal{L}^-$. The two lattice points $\mathbf{y}_+ \in \mathcal{L}^+$ and $\mathbf{y}_- \in \mathcal{L}^-$ are such that the line joining them intersects the subregion $\mathcal{A} \subset \mathcal{P}$.

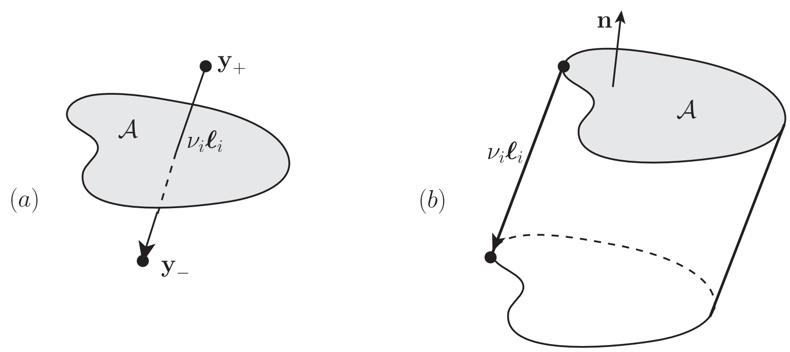


Figure 8.6: (a) Two lattice points $\mathbf{y}_+ \in \mathcal{L}^+$ and $\mathbf{y}_- \in \mathcal{L}^-$: $\mathbf{y}_- - \mathbf{y}_+ = \nu_i \boldsymbol{\ell}_i$ for some integers ν_1, ν_2, ν_3. (b) Non-prismatic cylinder whose base is $\mathcal{A}$ and generator is $\nu_i \boldsymbol{\ell}_i$.

fact that $(\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n} < 0$. Given the triplet of integers $\{\nu_1, \nu_2, \nu_3\}$, equation (8.6) gives the corresponding number of pairs of points whose line of connection intersects $\mathcal{A}$.

We can now evaluate the summation in (8.5) in two steps: first, for given $\{\nu_1, \nu_2, \nu_3\}$ with $(\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n} < 0$, we sum over all pairs of lattice points $\mathbf{y}_-$ and $\mathbf{y}_+$ which have $\mathbf{y}_- - \mathbf{y}_+ = \nu_i \boldsymbol{\ell}_i$ and where the line connecting them intersects $\mathcal{A}$. Then, we sum over all triplets of integers

be the same at all point on $\mathcal{P}$. The lattice at hand has a uniform geometry and we want (8.5) to be related to the continuum notion of traction. This requires that the right-hand side of (8.5) be independent of the size of $\mathcal{A}$. This in turn requires that the subregion $\mathcal{A}$ be sufficiently large.

$\{\nu_1, \nu_2, \nu_3\}$ obeying $(\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n} < 0$. This leads to

$$\mathbf{t}(\mathcal{A}) = \frac{1}{\text{area } (\mathcal{A})} \sum_{\substack{\{\nu_1, \nu_2, \nu_3\} \ni \\ (\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n} < 0}} -\phi'(|\nu_p \boldsymbol{\ell}_p|) \, \frac{\nu_i \boldsymbol{\ell}_i}{|\nu_k \boldsymbol{\ell}_k|} \, N \ . \tag{8.7}$$

Substituting (8.6) into this yields

$$\mathbf{t}(\mathcal{A}) = \frac{1}{(\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) \cdot \boldsymbol{\ell}_3} \sum_{\substack{\{\nu_1, \nu_2, \nu_3\} \ni \\ (\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n} < 0}} \phi'(|\nu_p \boldsymbol{\ell}_p|) \, \frac{\nu_i \boldsymbol{\ell}_i}{|\nu_k \boldsymbol{\ell}_k|} \, (\nu_j \boldsymbol{\ell}_j) \cdot \mathbf{n} \ . \tag{8.8}$$

Finally, observe that if we change $\{\nu_1, \nu_2, \nu_3\} \to \{-\nu_1, -\nu_2, -\nu_3\}$, the term within the summation sign remains unchanged though $(\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n}$ changes sign. Therefore, the sum above with the restriction $(\nu_i \boldsymbol{\ell}_i) \cdot \mathbf{n} < 0$ equals one-half the sum without this restriction. Therefore we obtain the following expression for the *traction* on the plane $\mathcal{P}$:

$$\mathbf{t}(\mathcal{A}) = \left[\frac{1}{2(\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) \cdot \boldsymbol{\ell}_3} \sum_{\{\nu_1, \nu_2, \nu_3\}} \phi'(|\nu_p \boldsymbol{\ell}_p|) \, \frac{(\nu_i \boldsymbol{\ell}_i) \otimes (\nu_j \boldsymbol{\ell}_j)}{|\nu_k \boldsymbol{\ell}_k|} \right] \mathbf{n} \tag{8.9}$$

where the summation is taken over all triplets of integers $\{\nu_1, \nu_2, \nu_3\}$.

Observe that the traction given by (8.9) depends linearly on the unit normal vector $\mathbf{n}$. This suggests that we define the *Cauchy stress tensor* $\mathbf{T}$ by

$$\mathbf{T} = \frac{1}{2(\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) \cdot \boldsymbol{\ell}_3} \sum_{\{\nu_1, \nu_2, \nu_3\}} \phi'(|\nu_p \boldsymbol{\ell}_p|) \, \frac{(\nu_i \boldsymbol{\ell}_i) \otimes (\nu_j \boldsymbol{\ell}_j)}{|\nu_k \boldsymbol{\ell}_k|} \ . \tag{8.10}$$

Note that $\mathbf{t} = \mathbf{T}\mathbf{n}$. Moreover $\mathbf{T} = \mathbf{T}^T$ as required by the balance of angular momentum. Given a Bravais lattice and a pair potential, equation (8.10) provides *an explicit formula for the stress*. It involves the geometry of the deformed lattice and the pair-potential.

Finally we provide a representation for $\mathbf{T}$ in terms of a referential lattice by replacing the deformed lattice vectors $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ in (8.10) by reference lattice vectors. To this end, consider a reference lattice defined by lattice vectors $\{\boldsymbol{\ell}^o_1, \boldsymbol{\ell}^o_2, \boldsymbol{\ell}^o_3\}$. The lattice vectors $\{\boldsymbol{\ell}^o_1, \boldsymbol{\ell}^o_2, \boldsymbol{\ell}^o_3\}$ of the reference lattice are related to the lattice vectors $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ of the deformed lattice through the nonsingular tensor $\mathbf{F}$ where $\boldsymbol{\ell}_i = \mathbf{F}\boldsymbol{\ell}^o_i$. The stress (in the deformed lattice) given by (8.10) can now be written in terms of the referential lattice vectors and $\mathbf{F}$ as

$$\boxed{\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{F}) = \frac{1}{2(\mathbf{F}\boldsymbol{\ell}^o_1 \times \mathbf{F}\boldsymbol{\ell}^o_2) \cdot \mathbf{F}\boldsymbol{\ell}^o_3} \sum_{\{\nu_1, \nu_2, \nu_3\}} \phi'(|\nu_p \, F\boldsymbol{\ell}^o_p|) \frac{(\nu_i \, \mathbf{F}\boldsymbol{\ell}^o_i) \otimes (\nu_j \, \mathbf{F}\boldsymbol{\ell}^o_j)}{|\nu_k \, \mathbf{F}\boldsymbol{\ell}^o_k|} \ .} \tag{8.11}$$

This provides an explicit formula for the *stress response function* $\widehat{\mathbf{T}}$ in terms of the referential lattice, the tensor $\mathbf{F}$, and the pair potential. If we associate a continuum with this lattice and invoke the Cauchy-Born hypothesis, $\mathbf{F}$ would be the deformation gradient tensor.

8.1.4 Energy.

We begin by calculating the energy of a single atom located at a lattice point $\mathbf{y}$. The energy associated with the pair of atoms located at $\mathbf{y}$ and $\boldsymbol{\xi}$ is $\phi(|\mathbf{y}-\boldsymbol{\xi}|)$. Assume that this energy is equally shared by the two atoms. Then, the energy of the atom located at $\mathbf{y}$ due to its interaction with all other atoms of the lattice is

$$\frac{1}{2}\sum_{\substack{\boldsymbol{\xi}\,\in\,\mathcal{L}\\ \boldsymbol{\xi}\,\neq\,\mathbf{y}}} \phi(|\mathbf{y}-\boldsymbol{\xi}|) \;=\; \frac{1}{2}\sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\,\boldsymbol{\ell}_i|). \tag{8.12}$$

Observe that this energy does not depend on $\mathbf{y}$, reflecting the fact that the lattice is uniform and the energy of each atom is the same. Now consider the energy associated with some region $\mathcal{R}$ of three dimensional space. If $\mathcal{R}$ is sufficiently large, the number of lattice points in $\mathcal{R}$ is

$$\frac{\text{vol }(\mathcal{R})}{(\boldsymbol{\ell}_1\times\boldsymbol{\ell}_2)\cdot\boldsymbol{\ell}_3} \tag{8.13}$$

where the denominator denotes the volume of the unit cell. Therefore the energy associated with the region $\mathcal{R}$ is given by the product of the two preceding expressions:

$$\frac{\text{vol }(\mathcal{R})}{(\boldsymbol{\ell}_1\times\boldsymbol{\ell}_2)\cdot\boldsymbol{\ell}_3}\quad\frac{1}{2}\sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\,\boldsymbol{\ell}_i|). \tag{8.14}$$

On dividing by vol $(\mathcal{R})$, we get the energy per unit deformed volume. Thus, given a Bravais lattice and a pair potential, equation (8.14) provides *an explicit formula for the energy per unit deformed volume*. It involves the geometry of the deformed lattice and the pair-potential.

Finally we express this in terms of a referential lattice. Consider the lattice defined by lattice vectors $\{\boldsymbol{\ell}_1^o,\boldsymbol{\ell}_2^o,\boldsymbol{\ell}_3^o\}$ that are related to the deformed lattice vectors by $\boldsymbol{\ell}_i=\mathbf{F}\boldsymbol{\ell}_i^o$. Substituting $\boldsymbol{\ell}_i=\mathbf{F}\boldsymbol{\ell}_i^o$ and using the fact that the volumes of $\mathcal{R}$ and its pre-image $\mathcal{R}_o$ in the reference configuration are related by $\text{vol}(\mathcal{R})=\det\mathbf{F}\,\text{vol}(\mathcal{R}_\text{o})$ allows us to write (8.14) as

$$\frac{\text{vol }(\mathcal{R}_o)\,\det\mathbf{F}}{(\mathbf{F}\boldsymbol{\ell}_1^o\times\mathbf{F}\boldsymbol{\ell}_2^o)\cdot\mathbf{F}\boldsymbol{\ell}_3^o}\quad\frac{1}{2}\sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\,\mathbf{F}\boldsymbol{\ell}_i^o|). \tag{8.15}$$

Finally, on using the identity $(\mathbf{Aa} \times \mathbf{Ab}) \cdot \mathbf{Ac} = \det \mathbf{A}\ (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ and dividing by vol$(\mathcal{R}_o)$, we obtain the following expression for the *energy per unit referential volume*:

$$\boxed{\widehat{W}(\mathbf{F}) \ = \ \frac{1}{(\boldsymbol{\ell}_1^o \times \boldsymbol{\ell}_2^o) \cdot \boldsymbol{\ell}_3^o} \quad \frac{1}{2} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\, \mathbf{F} \boldsymbol{\ell}_i^o|).} \tag{8.16}$$

This provides an explicit formula for the *strain energy response function* $\widehat{W}$ in terms of the referential lattice, the tensor $\mathbf{F}$, and the pair potential. If we associate a continuum with this lattice and invoke the Cauchy-Born hypothesis, $\mathbf{F}$ would be the deformation gradient tensor.

Note from (8.16) and (8.4) that the function $\widehat{W}$ and the tensor $\mathbf{F}$ both depend on the choice of reference lattice vectors. However, the energy of the deformed lattice does not depend on the choice of reference lattice vectors. Therefore the way in which $\mathbf{F}$ and $\widehat{W}$ depend on the reference lattice vectors must balance each other out such that the value of $\widehat{W}$ is independent of the choice of reference lattice vectors.

It is shown in Problem 8.2 that the stress response function (8.11) derived previously and the energy response function (8.16) are related *automatically* through the relation

$$\widehat{\mathbf{T}}(\mathbf{F}) = \frac{1}{\det \mathbf{F}} \frac{\partial \widehat{W}}{\partial \mathbf{F}}(\mathbf{F})\, \mathbf{F}^T \tag{8.17}$$

which is precisely what the continuum theory requires based on an elastic material being dissipation-free.

8.1.5 Material Frame Indifference.

It is shown in Problem 8.1 that the constitutive response function $\widehat{\mathbf{T}}(\mathbf{F})$ defined by (8.11) *automatically* obeys the relation

$$\widehat{\mathbf{T}}(\mathbf{QF}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^T$$

for all proper orthogonal tensors $\mathbf{Q}$ as would be required by material frame indifference in the continuum theory.

It can similarly be verified that the energy response function (8.16) has the property that

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{QF})$$

for all proper orthogonal tensors $\mathbf{Q}$. This shows that $\widehat{W}$ is automatically consistent with material frame indifference.

8.1.6 Linearized Elastic Moduli. Cauchy Relations.

In Problem 8.3 we shall linearize the constitutive quantities (8.16) and (8.17) to the special case of infinitesimal deformations. This leads to the constitutive relation of linear elasticity with the material being characterized by an elasticity tensor $\mathbb{C}$. In fact, equation (8.35) of Problem 8.3 provides an explicit formula for the components $\mathbb{C}_{ijk\ell}$ of the elasticity tensor in terms of the referential lattice and the pair potential.

The elastic moduli obtained in this way exhibit the symmetries

$$\mathbb{C}_{ijk\ell} = \mathbb{C}_{k\ell ij} = \mathbb{C}_{jik\ell} = \mathbb{C}_{ij\ell k}, \tag{8.18}$$

just as required by the continuum theory; see (4.150). However *in addition*, $\mathbb{C}_{ijk\ell}$ given by (8.35) also possesses the symmetry

$$\mathbb{C}_{ijk\ell} = \mathbb{C}_{i\ell kj} \tag{8.19}$$

which is not required by the continuum theory. The symmetries (8.19) obtained from the present lattice model are known as the *Cauchy relations*. The Cauchy relations are known to be not obeyed by most elastic materials[4] and this is therefore a limitation of the lattice theory formulated here. This limitation is directly related to the use of a pair-potential to model interatomic interactions. More realistic interatomic interaction models remove this limitation.

8.1.7 Lattice and Continuum Symmetry.

Since the stress and strain energy response functions $\widehat{\mathbf{T}}$ and $\widehat{W}$ given by (8.11) and (8.16) were derived from lattice considerations, they inherit the appropriate invariance characteristics associated with the symmetry of the underlying lattice. In this section we address three issues:

1. We examine the geometric invariance characteristics of a Bravais lattice and construct its "lattice symmetry group".

2. We show that the lattice symmetry group plays the role of the material symmetry group for the response functions $\widehat{\mathbf{T}}$ and $\widehat{W}$ derived above.

[4]For example, for an isotropic material, the Cauchy relations imply that the Poisson ratio must always be 0.25.

3. We remark on the suitability of using the lattice symmetry group to characterize the symmetry of a continuum.

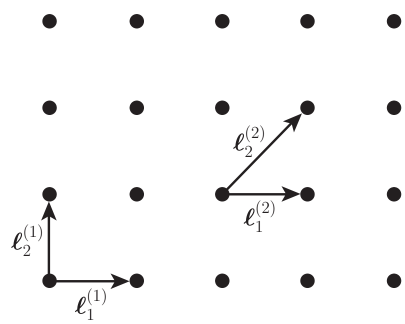


Figure 8.7: Two sets of lattice vectors that describe the same lattice.

Characterizing the symmetry of a Bravais lattice: First observe that because of its inherent symmetry, more than one set of lattice vectors may generate the same lattice. For example, the two-dimensional lattice shown in Figure 8.7 is generated by both $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}\}$ and $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}\}$. Observe that

$$\begin{aligned} \boldsymbol{\ell}_1^{(2)} &= \boldsymbol{\ell}_1^{(1)}, \\ \boldsymbol{\ell}_2^{(2)} &= \boldsymbol{\ell}_1^{(1)} + \boldsymbol{\ell}_2^{(1)}, \end{aligned} \tag{8.20}$$

so that the 2×2 matrix $[\mu]$, whose elements relate the two sets of lattice vectors through $\boldsymbol{\ell}_i^{(2)} = \mu_{ij} \boldsymbol{\ell}_j^{(1)}$, is

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{8.21}$$

Note that the elements of $[\mu]$ are integers and that $\det [\mu] = 1$.

In general, let $\mathcal{L}(\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)})$ be the lattice generated by a given set of lattice vectors $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$. Suppose that $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}\}$ is a second[5] set of lattice vectors that generates *this same lattice*, i.e.

$$\mathcal{L}(\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}) = \mathcal{L}(\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}).$$

One can show that two sets of lattice vectors generate the same lattice if and only if the matrix $[\mu]$, whose elements relate the two sets of lattice vectors through

$$\boldsymbol{\ell}_i^{(2)} = \mu_{ij} \boldsymbol{\ell}_j^{(1)}, \tag{8.22}$$

[5] We shall only consider lattice vector sets that have the same orientation.

has elements that are integers and whose determinant is 1.

An alternative more useful way in which to characterize symmetry is as follows: given a set of lattice vectors $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ and the associated Bravais lattice $\mathcal{L} = \mathcal{L}(\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)})$, let $\mathcal{G}(\mathcal{L})$ denote the set of all nonsingular tensors $\mathbf{H}$ that map $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ into a set of vectors $\{\mathbf{H}\boldsymbol{\ell}_1^{(1)}, \mathbf{H}\boldsymbol{\ell}_2^{(1)}, \mathbf{H}\boldsymbol{\ell}_3^{(1)}\}$ that generate the same lattice:

$$\mathcal{L}(\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}) = \mathcal{L}(\mathbf{H}\boldsymbol{\ell}_1^{(1)}, \mathbf{H}\boldsymbol{\ell}_2^{(1)}, \mathbf{H}\boldsymbol{\ell}_3^{(1)}) \qquad \text{for all } \mathbf{H} \in \mathcal{G}(\mathcal{L}).$$

It follows from (8.22) that $\mathcal{G}(\mathcal{L})$ admits the representation

$$\mathcal{G}(\mathcal{L}) = \Big\{\mathbf{H}\colon\ \mathbf{H}\boldsymbol{\ell}_i^{(1)} = \mu_{ij}\, \boldsymbol{\ell}_j^{(1)} \text{ for all } \mu_{ij} \text{ that are integers with } \det[\mu] = 1\Big\}\ . \tag{8.23}$$

Two sets of lattice vectors generate the same lattice if and only if

$$\boldsymbol{\ell}_i^{(2)} = \mathbf{H}\boldsymbol{\ell}_i^{(1)}, \quad i = 1, 2, 3, \tag{8.24}$$

where $\mathbf{H} \in \mathcal{G}(\mathcal{L})$. This is equivalent to (8.22). Despite the presence of the lattice vectors on the right hand side of (8.23), by its definition, $\mathcal{G}(\mathcal{L})$ depends on the lattice but not on the particular set of lattice vectors used to represent it. The set $\mathcal{G}(\mathcal{L})$ can be shown to be a group. It characterizes the symmetry of the lattice $\mathcal{L}$ and may be referred to as the "lattice symmetry group".

It is shown in Problem 8.5 that

$$\det\ \mathbf{H} = 1 \qquad \text{for all } \mathbf{H} \in \mathcal{G}(\mathcal{L}). \tag{8.25}$$

As a consequence, note that the volumes of the unit cells formed by lattice vectors $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ and $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}\}$ are equal if the lattice vectors are related through a symmetry transformation:

$$(\boldsymbol{\ell}_1^{(2)} \times \boldsymbol{\ell}_2^{(2)}) \cdot \boldsymbol{\ell}_3^{(2)} = (\boldsymbol{\ell}_1^{(1)} \times \boldsymbol{\ell}_2^{(1)}) \cdot \boldsymbol{\ell}_3^{(1)} \tag{8.26}$$

provided

$$\boldsymbol{\ell}_i^{(2)} = \mathbf{H}\boldsymbol{\ell}_i^{(1)}, \qquad \mathbf{H} \in \mathcal{G}(\mathcal{L}). \tag{8.27}$$

Symmetry of the response functions $\widehat{\mathbf{T}}$ and $\widehat{W}$: As noted at the beginning of this subsection, since the stress and strain energy response functions $\widehat{\mathbf{T}}$ and $\widehat{W}$ given by (8.11) and (8.16) were derived from lattice considerations, they inherit the appropriate invariance characteristics associated with the symmetry of the underlying lattice. We shall now verify this claim and show, for example, that

$$\widehat{W}(\mathbf{F}) = \widehat{W}(\mathbf{F}\mathbf{H}) \qquad \text{for all } \mathbf{H} \in \mathcal{G}(\mathcal{L}_o) \tag{8.28}$$

where $\mathcal{G}(\mathcal{L}_o)$ is the lattice symmetry group (8.23) of the reference lattice $\mathcal{L}_o$ and $\widehat{W}$ is the strain energy response function (8.16).

Recall from Section 4.4 that when examining symmetry in the continuum theory, we considered a deformed configuration χ and two reference configurations χ_1 and χ_2. We were interested in the special case when a symmetry transformation took $\chi_1 \to \chi_2$. In the lattice theory we analogously consider a deformed lattice $\mathcal{L}$ that is generated by lattice vectors $\{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3\}$ and two reference lattices $\mathcal{L}_1$ and $\mathcal{L}_2$ that are generated by lattice vectors $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ and $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}\}$. We are interested in the special case when a symmetry transformation takes $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ to $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}\}$ in which case the reference lattices $\mathcal{L}_1$ and $\mathcal{L}_2$ are identical: $\mathcal{L}_1 = \mathcal{L}_2$.

Let $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ be a set of lattice vectors characterizing a reference lattice $\mathcal{L}_1$, and let $\widehat{W}_1$ be the stored energy response function with respect to this reference lattice:

$$\widehat{W}_1(\mathbf{F}) \;=\; \frac{1}{(\boldsymbol{\ell}_1^{(1)} \times \boldsymbol{\ell}_2^{(1)}) \cdot \boldsymbol{\ell}_3^{(1)}} \quad \frac{1}{2} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^{(1)}|). \tag{8.29}$$

Let $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}\}$ be another set of lattice vectors characterizing a (possibly different) reference lattice $\mathcal{L}_2$, and let $\widehat{W}_2$ be the stored energy response function with respect to this reference lattice:

$$\widehat{W}_2(\mathbf{F}) \;=\; \frac{1}{(\boldsymbol{\ell}_1^{(2)} \times \boldsymbol{\ell}_2^{(2)}) \cdot \boldsymbol{\ell}_3^{(2)}} \quad \frac{1}{2} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^{(2)}|). \tag{8.30}$$

If the two sets of reference lattice vectors are related by (8.22), or equivalently by (8.24), then they generate the same reference lattice ($\mathcal{L}_1 = \mathcal{L}_2$) in which case

$$\widehat{W}_1(\mathbf{F}) = \widehat{W}_2(\mathbf{F}). \tag{8.31}$$

It then follows from (8.29) and (8.24) that

$$\begin{aligned}
\widehat{W}_1(\mathbf{F}\mathbf{H}) \;&=\; \frac{1}{(\boldsymbol{\ell}_1^{(1)} \times \boldsymbol{\ell}_2^{(1)}) \cdot \boldsymbol{\ell}_3^{(1)}} \quad \frac{1}{2} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\, \mathbf{F}\mathbf{H}\boldsymbol{\ell}_i^{(1)}|) \\
&=\; \frac{1}{(\boldsymbol{\ell}_1^{(1)} \times \boldsymbol{\ell}_2^{(1)}) \cdot \boldsymbol{\ell}_3^{(1)}} \quad \frac{1}{2} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^{(2)}|) \\
&=\; \frac{1}{(\boldsymbol{\ell}_1^{(2)} \times \boldsymbol{\ell}_2^{(2)}) \cdot \boldsymbol{\ell}_3^{(2)}} \quad \frac{1}{2} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(|\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^{(2)}|) \\
&=\; \widehat{W}_2(\mathbf{F})
\end{aligned}$$

where in the penultimate step we have used (8.26) and in the ultimate step we have used (8.30). It follows from this and (8.31) that

$$\widehat{W}_1(\mathbf{FH}) = \widehat{W}_1(\mathbf{F}) \qquad \text{for all nonsingular } \mathbf{F} \text{ and all } \mathbf{H} \in \mathcal{G}(\mathcal{L}_1).$$

Similarly one can show that

$$\widehat{\mathbf{T}}_1(\mathbf{F}) = \widehat{\mathbf{T}}_1(\mathbf{FH}) \qquad \text{for all } \mathbf{H} \in \mathcal{G}(\mathcal{L}_1)\ . \tag{8.32}$$

Thus the stress response function $\widehat{\mathbf{T}}$ and the energy response function $\widehat{W}$ derived from the present lattice theory, i.e. (8.11) and (8.16), are invariant under the group of transformations $\mathcal{G}(\mathcal{L}_o)$ that map the reference lattice back onto itself.

The lattice symmetry group and the symmetry of a continuum. Suppose that the lattice underlying the reference configuration of some elastic solid is a known Bravais lattice $\mathcal{L}_0$. However, suppose that one does not adopt the elementary pair potential model for interatomic interactions but arrives at a form for the strain energy function $\widehat{W}(\mathbf{F})$ by some other method, i.e. consider a strain energy response function $\widehat{W}(\mathbf{F})$ for the lattice that is *not* given by (8.16).

Even though the pair potential model for interatomic interactions was not used, the underlying lattice is (by assumption) a known Bravais lattice. Thus in particular the symmetry of the lattice is characterized by a known lattice symmetry group $\mathcal{G}(\mathcal{L}_o)$. Should one require that the continuum model exhibit all of the symmetries of the lattice? i.e. should we require

$$\widehat{W}(\mathbf{FH}) = \widehat{W}(\mathbf{F}) \quad \text{for all } \mathbf{H} \in \mathcal{G}(\mathcal{L}_o)? \tag{8.33}$$

The generally accepted answer is "no": the material symmetry group of the continuum should be a suitable subgroup of $\mathcal{G}(\mathcal{L}_o)$. This is based on the fact that in addition to rotations and reflections, the lattice symmetry group $\mathcal{G}(\mathcal{L}_o)$ contains finite shears as well. For example consider (8.20), (8.21). In view of (8.21) one sees that the transformation from $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}\}$ to $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}\}$ is a simple shear. Such shears cause large distortions of the lattice and are usually associated with lattice slip and plasticity. It is natural therefore to exclude these large shears when modeling elastic materials.

Based on the work of Ericksen & Pitteri (see Bhattacharya) the appropriate material symmetry group of the continuum should be the subgroup of rotations in $\mathcal{G}(\mathcal{L}_o)$:

$$\mathsf{P}(\mathcal{L}_o) = \{\mathbf{R} :\ \mathbf{R} \in SO(3),\, \mathbf{R} \in \mathcal{G}(\mathcal{L}_o)\}\ . \tag{8.34}$$

Thus we would require $\widehat{W}(\mathbf{FR}) = \widehat{W}(\mathbf{F})$ for all $\mathbf{R} \in \mathsf{P}(\mathcal{L}_o)$ instead of the more stringent requirement (8.33). $\mathsf{P}(\mathcal{L}_o)$ is called the "point group" or " Laue group" of the lattice. It is the group of rotations which map the lattice[6] back into itself. The point group associated with any Bravais lattice is a finite group.

8.1.8 Worked Examples and Exercises.

Problem 8.1. Show that the stress response function $\widehat{\mathbf{T}}(\mathbf{F})$ given explicitly in (8.11) *automatically* satisfies the condition $\widehat{\mathbf{T}}(\mathbf{QF}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^T$ for all proper orthogonal tensors $\mathbf{Q}$. (Therefore this $\widehat{\mathbf{T}}(\mathbf{F})$ is automatically material frame indifferent.)

Solution: From (8.11),

$$\mathbf{T}(\mathbf{QF}) = \frac{1}{2(\mathbf{QF}\boldsymbol{\ell}_1^o \times \mathbf{QF}\boldsymbol{\ell}_2^o)\cdot \mathbf{QF}\boldsymbol{\ell}_3^o} \sum_{\{\nu_1,\nu_2,\nu_3\}} \left[\phi'(|\nu_p\, \mathbf{QF}\boldsymbol{\ell}_p^o|) \ . \ \frac{(\nu_i\, \mathbf{QF}\boldsymbol{\ell}_i^o)\otimes(\nu_j\, \mathbf{QF}\boldsymbol{\ell}_j^o)}{|\nu_k\, \mathbf{QF}\boldsymbol{\ell}_k^o|} \right] . \qquad (a)$$

By using the vector identity $(\mathbf{Aa}\times\mathbf{Ab})\cdot\mathbf{Ac} = \det\mathbf{A}\ (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ and the fact that $\det\mathbf{Q} = 1$ we can write

$$(\mathbf{QF}\boldsymbol{\ell}_1^o \times \mathbf{QF}\boldsymbol{\ell}_2^o)\cdot \mathbf{QF}\boldsymbol{\ell}_3^o \ = \ (\mathbf{F}\boldsymbol{\ell}_1^o \times \mathbf{F}\boldsymbol{\ell}_2^o)\cdot \mathbf{F}\boldsymbol{\ell}_3^o \ . \qquad (b)$$

Next, since $\mathbf{Q}$ is orthogonal, it preserves length, i.e. $|\mathbf{Qy}| = |\mathbf{y}|$ for all vectors $\mathbf{y}$, and consequently

$$|\nu_i\, \mathbf{QF}\boldsymbol{\ell}_i^o| = \ |\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o| \ . \qquad (c)$$

Finally, in view of the vector identity $(\mathbf{Aa})\otimes(\mathbf{Bb}) = \mathbf{A}(\mathbf{a}\otimes\mathbf{b})\mathbf{B}^T$ can write

$$(\nu_i\mathbf{QF}\boldsymbol{\ell}_i^o)\otimes(\nu_i\mathbf{QF}\boldsymbol{\ell}_i^o) = \mathbf{Q}\left((\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o)\otimes(\nu_j\, \mathbf{F}\boldsymbol{\ell}_j^o)\right)\mathbf{Q}^T. \qquad (d)$$

Therefore we can simplify (a) by using $(b), (c)$ and (d) to get

$$\begin{aligned} \mathbf{T}(\mathbf{QF}) \ &= \ \frac{1}{2(\mathbf{F}\boldsymbol{\ell}_1^o \times \mathbf{F}\boldsymbol{\ell}_2^o)\cdot \mathbf{F}\boldsymbol{\ell}_3^o}\, \mathbf{Q} \left(\sum_{\{\nu_1,\nu_2,\nu_3\}} \left[\phi'(|\nu_p\, F\boldsymbol{\ell}_p^o|) \frac{(\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o)\otimes(\nu_j\, \mathbf{F}\boldsymbol{\ell}_j^o)}{|\nu_k\, \mathbf{F}\boldsymbol{\ell}_k^o|} \right] \right) \mathbf{Q}^T \\ &= \ \mathbf{Q}\mathbf{T}(\mathbf{F})\mathbf{Q}^T \ . \end{aligned}$$

Problem 8.2. Show that the Cauchy stress response function $\widehat{\mathbf{T}}(\mathbf{F})$ given explicitly by (8.11) and the strain energy response function $\widehat{W}(\mathbf{F})$ given explicitly by (8.16) are *automatically* related by

$$\widehat{\mathbf{T}}(\mathbf{F}) = \frac{1}{\det\mathbf{F}}\,\frac{\partial\widehat{W}}{\partial\mathbf{F}}(\mathbf{F})\,\mathbf{F}^T.$$

---

[6] For example, the point group of a simple cubic lattice consists of the 24 rotations that map the unit cube back into itself.

(Therefore the stress and strain energy response functions provided by the lattice theory automatically satisfy the relation imposed by the dissipation inequality; see Section 4.2.1 on page 250.)

Solution: Differentiating (8.16) with respect to $\mathbf{F}$ gives

$$2(\boldsymbol{\ell}_1^o \times \boldsymbol{\ell}_2^o) \cdot \boldsymbol{\ell}_3^o \left(\frac{\partial \widehat{W}}{\partial \mathbf{F}} \right) = \sum_{\{\nu_1,\nu_2,\nu_3\}} \left[\phi'(|\nu_p\, \mathbf{F}\boldsymbol{\ell}_p^o|) \left(\frac{\partial}{\partial \mathbf{F}} (|\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o|) \right) \right] . \qquad (a)$$

The following identity can be readily verified for an arbitrary vector $\mathbf{y}$:

$$\frac{\partial}{\partial \mathbf{F}}(|\mathbf{F}\mathbf{y}|) = \frac{1}{2|\mathbf{F}\mathbf{y}|} \frac{\partial}{\partial \mathbf{F}}(|\mathbf{F}\mathbf{y}|^2) = \frac{1}{2|\mathbf{F}\mathbf{y}|} \frac{\partial}{\partial \mathbf{F}}\Big(\mathbf{F}\mathbf{y}\cdot\mathbf{F}\mathbf{y}\Big) = \frac{1}{2|\mathbf{F}\mathbf{y}|}\Big(2\mathbf{F}\mathbf{y}\otimes\mathbf{y}\Big),$$

from which it follows that

$$\frac{\partial}{\partial \mathbf{F}}(|\nu_p\, \mathbf{F}\boldsymbol{\ell}_p^o|) = \frac{1}{|\nu_k\, \mathbf{F}\boldsymbol{\ell}_k^o|} \left[(\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o) \otimes \nu_j\, \boldsymbol{\ell}_j^o \right] . \qquad (b)$$

Substituting (b) into (a) yields

$$2(\boldsymbol{\ell}_1^o \times \boldsymbol{\ell}_2^o) \cdot \boldsymbol{\ell}_3^o \left(\frac{\partial \widehat{W}}{\partial \mathbf{F}} \right) = \sum_{\{\nu_1,\nu_2,\nu_3\}} \left[\phi'(|\nu_p\, \mathbf{F}\boldsymbol{\ell}_p^o|) \frac{(\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o) \otimes \nu_j\, \boldsymbol{\ell}_j^o}{|\nu_k\, \mathbf{F}\boldsymbol{\ell}_k^o|} \right],$$

from which it follows that

$$2(\boldsymbol{\ell}_1^o \times \boldsymbol{\ell}_2^o) \cdot \boldsymbol{\ell}_3^o \left(\frac{\partial \widehat{W}}{\partial \mathbf{F}} \right) \mathbf{F}^T = \sum_{\{\nu_1,\nu_2,\nu_3\}} \left[\phi'(|\nu_p\, \mathbf{F}\boldsymbol{\ell}_p^o|) \frac{(\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^o) \otimes \nu_j\, \mathbf{F}\boldsymbol{\ell}_j^o}{|\nu_k\, \mathbf{F}\boldsymbol{\ell}_k^o|} \right] . \qquad (c)$$

Finally, because of the identity $(\mathbf{A}\mathbf{a} \times \mathbf{A}\mathbf{b}) \cdot \mathbf{A}\mathbf{c} = \det \mathbf{A}\ (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$, we see from (c) and (8.11) that the relation (8.17) between $\widehat{\mathbf{T}}$ and $\widehat{W}$ holds.

Problem 8.3. Derive an explicit expression for the elasticity tensor $\mathbb{C}$ of linear elasticity by linearization of the results of this chapter. Show that the resulting components $\mathbb{C}_{ijk\ell}$ posses the usual symmetries

$$\mathbb{C}_{ijk\ell} = \mathbb{C}_{k\ell ij} = \mathbb{C}_{jik\ell} = \mathbb{C}_{ij\ell k}, \qquad (a)$$

as well as the additional symmetry

$$\mathbb{C}_{ijk\ell} = \mathbb{C}_{i\ell kj} \qquad (b)$$

known as the Cauchy relations.

Solution: We first show that the energy response function $\widehat{W}$ given by (8.16) depends on $\mathbf{F}$ only through the Cauchy-Green tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$; thereafter we determine the components of the elasticity tensor $\mathbb{C}$ by recalling that

$$\mathbb{C}_{ijk\ell} = \left. \frac{\partial^2 W(\mathbf{C})}{\partial C_{ij}\, \partial C_{kl}} \right|_{\mathbf{C}=\mathbf{I}} .$$

The fact that $\widehat{W}$ depends on $\mathbf{F}$ only through $\mathbf{C}$ follows from

$$|\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^0| = \Big((\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^0) \cdot (\nu_i\, \mathbf{F}\boldsymbol{\ell}_i^0) \Big)^{1/2} = \Big(\mathbf{F}^T\mathbf{F}(\nu_i\, \boldsymbol{\ell}_i^0) \cdot \nu_i\, \boldsymbol{\ell}_i^0 \Big)^{1/2} = \Big(\mathbf{C}(\nu_i\, \boldsymbol{\ell}_i^0) \cdot \nu_i\, \boldsymbol{\ell}_i^0 \Big)^{1/2}$$

whence we can write (8.16) as

$$W(\mathbf{C}) = \frac{1}{2(\boldsymbol{\ell}_1^0 \times \boldsymbol{\ell}_2^0) \cdot \boldsymbol{\ell}_3^0} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi\Big((\mathbf{C}(\nu_i\, \boldsymbol{\ell}_i^0) \cdot \nu_i\, \boldsymbol{\ell}_i^0)^{1/2} \Big) .$$

In order to calculate the elasticity tensor we must calculate the second derivative of W with respect to $\mathbf{C}$ and then evaluate it in the reference configuration where $\mathbf{C} = \mathbf{I}$. In order to simplify the writing it is convenient to introduce the notation

$$\alpha = 2(\boldsymbol{\ell}_1^0 \times \boldsymbol{\ell}_2^0) \cdot \boldsymbol{\ell}_3^0, \qquad \mathbf{y} = \nu_i\, \boldsymbol{\ell}_i^0, \qquad \rho(\mathbf{C}) = (\mathbf{C}\mathbf{y} \cdot \mathbf{y})^{1/2} ,$$

so that

$$W(\mathbf{C}) = \frac{1}{\alpha} \sum_{\{\nu_1,\nu_2,\nu_3\}} \phi(\rho(\mathbf{C})).$$

It is straightforward to show that

$$\frac{\partial \rho(\mathbf{C})}{\partial C_{k\ell}} = \frac{y_k\, y_\ell}{2\rho(\mathbf{C})}.$$

Therefore the first derivative of W is

$$\frac{\partial W}{\partial C_{k\ell}}(\mathbf{C}) = \sum_{\{\nu_1,\nu_2,\nu_3\}} \frac{1}{\alpha}\, \phi'(\rho(\mathbf{C}))\, \frac{\partial \rho(\mathbf{C})}{\partial C_{k\ell}} = \sum_{\{\nu_1,\nu_2,\nu_3\}} \frac{1}{2\alpha\rho(\mathbf{C})}\, \phi'(\rho(\mathbf{C}))\, y_k\, y_\ell .$$

The second derivative can be calculated similarly by differentiating this once more, which leads after some calculation to

$$\frac{\partial^2 W(\mathbf{C})}{\partial C_{ij}\, \partial C_{k\ell}} = \sum_{\{\nu_1,\nu_2,\nu_3\}} \frac{1}{4\alpha\rho^2(\mathbf{C})} \left(\phi''(\rho(\mathbf{C})) - \frac{1}{\rho(\mathbf{C})} \phi'(\rho(\mathbf{C})) \right) y_i\, y_j\, y_k\, y_\ell .$$

In order to calculate the components of the elasticity tensor we set $\mathbf{C} = \mathbf{I}$, $\rho(\mathbf{C}) = \rho(\mathbf{I}) = |\mathbf{y}|$ in the preceding expression to obtain

$$\mathbb{C}_{ijk\ell} = \left. \frac{\partial^2 W(\mathbf{I})}{\partial C_{ij}\, \partial C_{k\ell}} \right|_{\mathbf{C}=\mathbf{I}} = \frac{1}{2(\boldsymbol{\ell}_1^0 \times \boldsymbol{\ell}_2^0) \cdot \boldsymbol{\ell}_3^0} \sum_{\{\nu_1,\nu_2,\nu_3\}} \frac{1}{4|\mathbf{y}|^2} \left(\phi''(|\mathbf{y}|) - \frac{1}{|\mathbf{y}|} \phi'(|\mathbf{y}|) \right) y_i\, y_j\, y_k\, y_\ell \tag{8.35}$$

where the vector $\mathbf{y} = \nu_i \boldsymbol{\ell}_i$. The right-hand side of this is invariant with respect to the change of any pair of subscripts, and therefore so is the left-hand side. This establishes the symmetries (a) and (b).

Problem 8.4. Show that two sets of lattice vectors $\{\boldsymbol{\ell}_1^{(1)}, \boldsymbol{\ell}_2^{(1)}, \boldsymbol{\ell}_3^{(1)}\}$ and $\{\boldsymbol{\ell}_1^{(2)}, \boldsymbol{\ell}_2^{(2)}, \boldsymbol{\ell}_3^{(2)}\}$ generate the same lattice if and only if the matrix $[\mu]$, whose elements relate the lattice vectors through

$$\boldsymbol{\ell}_i^{(2)} = \mu_{ij} \boldsymbol{\ell}_j^{(1)},$$

has elements that are integers and has determinant 1.

Problem 8.5. Let $\mathbf{H}$ be any member of the lattice symmetry group $\mathcal{G}(\mathcal{L})$ defined in (8.23). Show that $\det \mathbf{H} = 1$.

Solution: Substitute

$$\mathbf{H}\boldsymbol{\ell}_i^{(1)} = \mu_{ij}\boldsymbol{\ell}_j^{(1)}$$

into the vector identity

$$(\mathbf{H}\boldsymbol{\ell}_1^{(1)} \times \mathbf{H}\boldsymbol{\ell}_2^{(1)}) \cdot \mathbf{H}\boldsymbol{\ell}_3^{(1)} = \det \mathbf{H}\ (\boldsymbol{\ell}_1^{(1)} \times \boldsymbol{\ell}_2^{(1)}) \cdot \boldsymbol{\ell}_3^{(1)}$$

and expand out the result. This leads to

$$\det\ [\mu] = \det \mathbf{H}$$

after making use of the fact that

$$(\boldsymbol{\ell}_1^{(1)} \times \boldsymbol{\ell}_2^{(1)}) \cdot \boldsymbol{\ell}_3^{(1)} = (\boldsymbol{\ell}_2^{(1)} \times \boldsymbol{\ell}_3^{(1)}) \cdot \boldsymbol{\ell}_1^{(1)} = (\boldsymbol{\ell}_3^{(1)} \times \boldsymbol{\ell}_1^{(1)}) \cdot \boldsymbol{\ell}_2^{(1)}$$

where each of these expressions represents the volume of the unit cell. Finally, since $\det\ [\mu] = 1$ it follows that $\det \mathbf{H} = 1$.

References:

1. K. Bhattacharya, Unpublished lecture notes, California Institute of Technology,
2. K. Bhattacharya, *Microstructure of Martensite*, Chapter 3, Oxford, 2003.
3. A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity*, Appendix B, Cambridge University Press, 1927.

Chapter 9

Brief Remarks on Coupled Problems

While the theory and problems dealt with in these notes have been focused on the purely mechanical theory of an elastic solid, there are many settings in which elasticity is coupled with other physical phenomena. For example thermoelasticity involves the coupling of mechanical and thermal effects, while mechanical and electrical effects are coupled in piezoelectricity.

To adequately deal with such coupled phenomena would require several more chapters which is beyond the scope of these notes. However, given the current interest of many students in "coupled problems", I will provide a very brief introduction to **the formalism** by which one sets-up such theories. Common to each of these settings is

(a) identifying the additional fields involved;

(b) identifying the additional governing physical principles;

(c) deriving the additional field equations from (b);

(d) stating the (primitive) form of the constitutive relations; and

(e) simplifying the form of the constitutive relations using the dissipation inequality. Recall that in Section 4.2.1 we briefly touched on using the dissipation inequality to simplify a constitutive relation.

For example, modeling thermoelasticity requires that in addition to the deformation $\mathbf{y}(\mathbf{x}, t)$, the deformation gradient $\mathbf{F}(\mathbf{x}, t)$ and stress $\mathbf{S}(\mathbf{x}, t)$, one also consider the heat flux $\mathbf{q}(\mathbf{x}, t)$, internal energy $\epsilon(\mathbf{x}, t)$, temperature $\theta(\mathbf{x}, t)$ and entropy $\eta(\mathbf{x}, t)$, at least some of which would have to be determined when solving an initial-boundary value problem. The additional physical principles in this case are the first and second laws of thermodynamics.

In general, one cannot solve a mechanical problem to find $\mathbf{y}(\mathbf{x},t), \mathbf{F}(\mathbf{x},t)$ and $\mathbf{S}(\mathbf{x},t)$ and a separate thermal problem to find the remaining fields. This is because the mechanical and thermal fields are coupled: for example, the constitutive relation for stress depends on both the deformation and the temperature, and the first law of thermodynamics involves both the rate of working of the stress and the rate of heat flux.

As a second example consider a piezoelectric solid. Here one has the following additional fields (in electrostatics): the electric potential $\varphi(\mathbf{x},t)$, electric field $\mathbf{E}(\mathbf{x},t)$ and electric displacement $\mathbf{D}(\mathbf{x},t)$ and one must take into account Maxwell's equations as well as the laws of thermodynamics when setting up the theory.

In the rest of this chapter I will briefly illustrate the underlying mathematical formalism through two examples: hydrogels and thermoelasticity. In each example we will go through steps $(a)-(e)$. Piezoelectric materials will not be touched on since most such materials are ceramics that undergo very small strains. They are treated in Chapter 4 of Volume IV which concerns the linear(ized) theory of elasticity.

We will be using a referential formulation throughout. As usual, the body occupies a region $\mathcal{R}_{\rm R}$ in a reference configuration and an arbitrary part of the body occupies a region $\mathcal{D}_{\rm R} \subset \mathcal{R}_{\rm R}$. The position vector of a particle in the reference configuration is $\mathbf{x}$ and at time t during a motion is $\mathbf{y}(\mathbf{x},t)$. Whenever we say "per unit volume" or "per unit area" we mean "per unit reference volume" or "per unit reference area" unless explicitly stated otherwise. The deformation gradient tensor is $\mathbf{F} = \nabla\mathbf{y}$ and the first Piola-Kirchhoff stress tensor is $\mathbf{S}$. A superior dot, such as in $\dot{\mathbf{F}}$, denotes the time derivative (with $\mathbf{x}$ held fixed – the material time derivative). All processes are assumed to be quasi-static in that inertial effects are neglected. Thus we do not account for linear and angular momentum and kinetic energy.

9.1 Hydrogels:

References:

1. L.R.G. Treloar, The Physics of Rubber Elasticity, Chapter 7, third edition, Oxford, 2005

2 W. Hong, X. Zhao, J. Zhou and Z. Suo, *A theory of coupled diffusion and large deformation in polymeric gels*, J. Mech. Phys. Solids, , **56**(2008), pp. 1779-1793.

3. S.A. Chester and L. Anand, *A coupled theory of fluid permeation and large deformations for elastomeric materials*, J. Mech. Phys. Solids, **58** (2010), pp. 1879-1906.

4. F.P. Duda, A.C. Souza and E. Fried, *A theory for species migration in a finitely strained solid with application to polymer network swelling*, J. Mech. Phys. Solids, **58** (2010), pp. 515-529.

A hydrogel is essentially an elastic polymer through which a solvent, usually water, diffuses. In what follows we shall speak of a "polymer" (an elastic solid) and a "solvent" (an inviscid fluid) that can move in and out of the polymer. One speaks of the swelling of the hydrogel as it absorbs the solvent. We do not use mixture theory (at least not explicitly) and so speak only of a single effective continuum.

A part $\mathcal{P}$ of the body is "material" with respect to the elastic solid in the sense that the same set of polymer particles are associated with $\mathcal{P}$ at all times. Solvent particles however may diffuse in and out of $\mathcal{D}_{\rm R}$. Often, the reference configuration is taken to correspond to a dry stress-free state of the polymer, but that particular choice doesn't affect the development below.

In addition to the basic fields of deformation $\mathbf{y}(\mathbf{x},t)$, deformation gradient $\mathbf{F}(\mathbf{x},t)$ and stress $\mathbf{S}(\mathbf{x},t)$, the theory now involves the *concentration* $c_{\rm R}(\mathbf{x},t)$ of the solvent, the *flux* $\mathbf{j}_{\rm R}(\mathbf{x},t)$ of the solvent across a surface, and the *chemical potential* of the solvent $\mu(\mathbf{x},t)$. Let $c_{\rm R}(\mathbf{x},t)$ be the referential solvent concentration so that the number of solvent molecules in $\mathcal{D}_{\rm R}$ at time t is

$$\int_{\mathcal{D}_{\rm R}} c_{\rm R}(\mathbf{x},t)\,dV_x.$$

The number of solvent molecules crossing a unit area of the boundary $\partial\mathcal{D}_{\rm R}$ in unit time and entering $\mathcal{D}_{\rm R}$ is $-\mathbf{j}_{\rm R}\cdot\mathbf{n}_{\rm R}$ where $\mathbf{j}_{\rm R}(\mathbf{x},t)$ is the referential solvent flux vector (and $\mathbf{n}_{\rm R}$ is a unit outward pointing normal vector on $\partial\mathcal{D}_{\rm R}$). The total rate at which the solvent enters $\mathcal{D}_{\rm R}$ across $\partial\mathcal{D}_{\rm R}$ is thus

$$\int_{\partial\mathcal{D}_{\rm R}} -\mathbf{j}_{\rm R}\cdot\mathbf{n}_{\rm R}\,dA_x.$$

The *bulk supply* of solvent molecules per unit volume per unit time (from sources outside the body) is denoted by $r_{\rm R}(\mathbf{x},t)$, and so the rate at which solvent molecules directly enter the interior of $\mathcal{D}_{\rm R}$ (in contrast to diffusing across its boundary) is

$$\int_{\mathcal{D}_{\rm R}} r_{\rm R}(\mathbf{x},t)\,dV_x.$$

The role of $r_{\rm R}$ is similar to that of the mechanical body force $\mathbf{b}_{\rm R}$ in that it is usually prescribable. Finally, let $\mu(\mathbf{x},t)$ denote the chemical potential of the solvent. It represents the energy per solvent molecule and so the rate of increase of the chemical energy of $\mathcal{D}_{\rm R}$ due to the influx of solvent molecules is

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} -\mu\,\mathbf{j}_{\mathrm{R}}\cdot\mathbf{n}_{\mathrm{R}}\,dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mu\, r_{\mathrm{R}}\,dV_x.$$

9.1.1 Basic mechanical equations. Balance laws and field equations.

Force and moment equilibrium require the usual balance laws

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{S}\mathbf{n}_{\mathrm{R}}\,dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{b}_{\mathrm{R}}\,dV_x = \mathbf{0}, \qquad \int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{y}\times(\mathbf{S}\mathbf{n}_{\mathrm{R}})\,dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{y}\times\mathbf{b}_{\mathrm{R}}\,dV_x = \mathbf{0},$$

that lead to the usual field equations

$$\operatorname{Div}\mathbf{S} + \mathbf{b}_{\mathrm{R}} = \mathbf{0}, \qquad \mathbf{S}\mathbf{F}^T = \mathbf{F}\mathbf{S}^T. \tag{9.1}$$

9.1.2 Basic chemical equation. Balance law and field equation.

The conservation of solvent molecules requires that

$$\frac{d}{dt}\int_{\mathcal{D}_{\mathrm{R}}} c_{\mathrm{R}}\,dV_x = \int_{\partial\mathcal{D}_{\mathrm{R}}} -\mathbf{j}_{\mathrm{R}}\cdot\mathbf{n}_{\mathrm{R}}\,dA_x \;+\; \int_{\mathcal{D}_{\mathrm{R}}} r_{\mathrm{R}}\,dV_x. \tag{9.2}$$

The left-hand side represents the rate of increase of the number of solvent molecules in $\mathcal{D}_{\mathrm{R}}$. The first term on the right-hand side denotes the number of solvent molecules entering $\mathcal{D}_{\mathrm{R}}$ across its boundary while the second term is the number of solvent molecules added to its interior, both per unit time. Equation (9.2) must hold for all $\mathcal{D}_R$ and so localization leads to the field equation

$$\dot{c}_{\mathrm{R}} + \operatorname{Div}\mathbf{j}_{\mathrm{R}} = r_{\mathrm{R}}. \tag{9.3}$$

This must hold at each particle at each time.

9.1.3 Dissipation inequality.

The *dissipation inequality* states that the rate of increase of free energy of $\mathcal{D}_{\mathrm{R}}$ cannot exceed the rate at which mechanical work is done on $\mathcal{D}_{\mathrm{R}}$ plus the rate at which chemical energy is added to $\mathcal{D}_{\mathrm{R}}$. Thus the dissipation inequality requires that

$$\int_{\partial\mathcal{D}_{\mathrm{R}}} \mathbf{S}\mathbf{n}_{\mathrm{R}}\cdot\mathbf{v}\,dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mathbf{b}_{\mathrm{R}}\cdot\mathbf{v}\,dV_x + \int_{\partial\mathcal{D}_{\mathrm{R}}} -\mu\mathbf{j}_{\mathrm{R}}\cdot\mathbf{n}_{\mathrm{R}}\,dA_x + \int_{\mathcal{D}_{\mathrm{R}}} \mu r_{\mathrm{R}}\,dV_x \ge \frac{d}{dt}\int_{\mathcal{D}_{\mathrm{R}}} \psi\,dV_x, \tag{9.4}$$

for all subregions $\mathcal{D}_\mathrm{R} \subset \mathcal{R}_\mathrm{R}$. Here $\mathbf{v}(\mathbf{x},t) = \dot{\mathbf{y}}(\mathbf{x},t)$ is the particle velocity. The first two terms represent the rate of working by the boundary traction and the body force, while the next two terms denote the influx of chemical energy. The *free energy* (per unit volume) has been denoted by ψ so that

$$\int_{\mathcal{D}_\mathrm{R}} \psi \, dV_x$$

is the total free energy of $\mathcal{D}_\mathrm{R}$ at time t. Its counterpart in the purely mechanical theory is the strain energy function W.

We can simplify (9.4) by converting the surface integrals to volume integrals using the divergence theorem and then making use of the field equations (9.1) and (9.3). This leads to

$$\int_{\mathcal{D}_\mathrm{R}} \Big(\mathbf{S} \cdot \dot{\mathbf{F}} - \mathbf{j}_\mathrm{R} \cdot \mathrm{Grad}\, \mu + \mu \dot{c}_\mathrm{R} - \dot{\psi} \Big) \, dV_x \geq 0. \tag{9.5}$$

Localization of (9.5) yields

$$\mathbf{S} \cdot \dot{\mathbf{F}} - \mathbf{j}_\mathrm{R} \cdot \nabla \mu + \mu \dot{c}_\mathrm{R} - \dot{\psi} \geq 0. \tag{9.6}$$

This is the local form of the dissipation inequality and it is required to hold at each point in the body at each time.

9.1.4 Constitutive equations:

Suppose that the material is characterized by the following set of constitutive relations[1]:

$$\psi = \psi(\mathbf{F}, c_\mathrm{R}), \qquad \mathbf{S} = \mathbf{S}(\mathbf{F}, c_\mathrm{R}), \qquad \mu = \mu(\mathbf{F}, c_\mathrm{R}), \qquad \mathbf{j}_\mathrm{R} = \mathbf{j}_\mathrm{R}(\mathbf{F}, c_\mathrm{R}, \nabla \mu). \tag{9.7}$$

This is the primitive form of the constitutive equations. It can be simplified (reduced) using the dissipation inequality as follows:

First note from $(9.7)_1$ that

$$\dot{\psi} = \frac{\partial \psi}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} + \frac{\partial \psi}{\partial c_\mathrm{R}} \dot{c}_\mathrm{R}.$$

Therefore we can write the dissipation inequality (9.6) as

$$\left[\mathbf{S}(\mathbf{F}, c_\mathrm{R}) - \frac{\partial \psi}{\partial \mathbf{F}}(\mathbf{F}, c_\mathrm{R}) \right] \cdot \dot{\mathbf{F}} + \left[\mu(\mathbf{F}, c_\mathrm{R}) - \frac{\partial \psi}{\partial c_\mathrm{R}}(\mathbf{F}, c_\mathrm{R}) \right] \dot{c}_\mathrm{R} - \mathbf{j}_\mathrm{R}(\mathbf{F}, c_\mathrm{R}, \nabla \mu) \cdot \nabla \mu \geq 0. \tag{9.8}$$

[1]See also Problem 9.1.1 on page 458.

Since (9.8) must hold in all processes, and therefore for all $\dot{\mathbf{F}}$ and $\dot{c}_{\mathrm{R}}$, and since the terms within the square brackets do not involve $\dot{\mathbf{F}}$ and $\dot{c}_{\mathrm{R}}$, it follows that those terms must vanish[2]. This leads to the following *constitutive relations* for $\mathbf{S}$ and μ:

$$\left.\begin{aligned} \mathbf{S} &= \frac{\partial \psi}{\partial \mathbf{F}}(\mathbf{F}, c_{\mathrm{R}}), \\ \mu &= \frac{\partial \psi}{\partial c_{\mathrm{R}}}(\mathbf{F}, c_{\mathrm{R}}). \end{aligned}\right\} \tag{9.9}$$

The dissipation inequality (9.8) now reduces to

$$\mathbf{j}_{\mathrm{R}}(\mathbf{F}, c_{\mathrm{R}}, \mathbf{g}) \cdot \mathbf{g} \leq 0 \tag{9.10}$$

for all vectors $\mathbf{g}$. The argument used in getting to (9.9) and (9.10) from (9.7) and (9.6) is known as the *Coleman-Noll argument*.

Thus a hydrogel is characterized by the free energy $\psi(\mathbf{F}, c_{\mathrm{R}})$ and the solvent flux law $\mathbf{j}_{\mathrm{R}}(\mathbf{F}, c_{\mathrm{R}}, \nabla\mu)$ where the latter must be consistent with (9.10). An example of the latter is

$$\mathbf{j}_{\mathrm{R}}(\mathbf{F}, c_{\mathrm{R}}, \nabla\mu) = -\mathbf{M}(\mathbf{F}, c_{\mathrm{R}})\, \nabla\mu, \tag{9.11}$$

where (9.10) requires the "mobility tensor" $\mathbf{M}$ to be positive semi-definite. This is the well-known Fick's law.

Exercise: Show that material frame indifference implies

$$\psi = \psi(\mathbf{C}, c_{\mathrm{R}}), \qquad \mu = \mu(\mathbf{C}, c_{\mathrm{R}}),$$

where $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ and that

$$\mathbf{S} = 2\mathbf{F}\frac{\partial \psi}{\partial \mathbf{C}}(\mathbf{C}, c_{\mathrm{R}}).$$

As a consequence of this, moment balance, $(9.1)_2$, holds automatically.

Thus **in summary**, a boundary-initial value problem for a body composed of a hydrogel involves specifying the body force $\mathbf{b}_{\mathrm{R}}(\mathbf{x}, t)$, the solvent supply $r_{\mathrm{R}}(\mathbf{x}, t)$, the material characterizations $\psi(\mathbf{F}, c_{\mathrm{R}})$ and $\mathbf{j}_{\mathrm{R}}(\mathbf{F}, c_{\mathrm{R}}, \nabla\mu)$ and suitable initial and boundary conditions (both mechanical and chemical)[3]. Substituting $(9.9)_2$ and (9.11) into (9.3) gives a scalar partial differential equation involving $\mathbf{y}(\mathbf{x}, t)$ and $c_{\mathrm{R}}(\mathbf{x}, t)$, often referred to casually as the "diffusion equation". Similarly substituting $(9.9)_1$ into the equilibrium equation $(9.1)_1$ gives three scalar (or one vector) partial differential equation. These equations are to be solved for $\mathbf{y}(\mathbf{x}, t)$ and $c_{\mathrm{R}}(\mathbf{x}, t)$.

[2] See Section 4.2.1 for more details on this argument.

[3] See Problem 9.1 for an example.

9.1.5 Alternative form of the constitutive relation.

Exercise: Suppose that you wanted to express the constitutive relations as functions of $\mathbf{F}$ and μ (instead of $\mathbf{F}$ and c_{R}) and therefore took the primitive form of the constitutive relations to be

$$\psi = \psi(\mathbf{F}, \mu), \qquad \mathbf{S} = \mathbf{S}(\mathbf{F}, \mu), \qquad c_{\mathrm{R}} = c_{\mathrm{R}}(\mathbf{F}, \mu), \qquad \mathbf{j}_{\mathrm{R}} = \mathbf{j}_{\mathrm{R}}(\mathbf{F}, \mu, \mathrm{Grad}\, \mu). \tag{9.12}$$

Can you use the dissipation inequality (9.6) to simplify these constitutive relations?

As you would have discovered from this exercise, one cannot directly use (9.6) to reduce constitutive response functions of the form (9.12). Since (9.6) involves $\dot{\mathbf{F}}$ and $\dot{c}_{\mathrm{R}}$ it favors constitutive characterizations in terms of $\mathbf{F}$ and c_{R}. This suggests that if we want to consider constitutive relations that are functions of $\mathbf{F}$ and μ we should seek to rewrite the dissipation inequality in terms $\dot{\mathbf{F}}$ and $\dot{\mu}$.

This is achieved by introducing the function

$$\omega := \psi - \mu c_{\mathrm{R}}. \tag{9.13}$$

The transformation from $\psi \to \omega$ is called a Legendre transformation and ω is the *Legendre transform* of ψ with respect to μ and c_{R}. In the present setting ω is referred to as the *grand canonical energy*. The dissipation inequality (9.6) can be written in terms of ω as

$$\mathbf{S} \cdot \dot{\mathbf{F}} - c_{\mathrm{R}} \dot{\mu} - \dot{\omega} - \mathbf{j}_{\mathrm{R}} \cdot \mathrm{Grad}\, \mu \geq 0. \tag{9.14}$$

Now consider constitutive relations of the primitive form

$$\omega = \omega(\mathbf{F}, \mu), \qquad \mathbf{S} = \mathbf{S}(\mathbf{F}, \mu), \qquad c_{\mathrm{R}} = c_{\mathrm{R}}(\mathbf{F}, \mu), \qquad \mathbf{j}_{\mathrm{R}} = \mathbf{j}_{\mathrm{R}}(\mathbf{F}, \mu, \nabla\mu). \tag{9.15}$$

Since,

$$\dot{\omega} = \frac{\partial \omega}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} + \frac{\partial \omega}{\partial \mu} \dot{\mu},$$

(9.14) can be written as

$$\left[\mathbf{S} - \frac{\partial \omega}{\partial \mathbf{F}}\right] \cdot \dot{\mathbf{F}} - \left[c_{\mathrm{R}} + \frac{\partial \omega}{\partial \mu}\right] \dot{\mu} - \mathbf{j}_{\mathrm{R}} \cdot \nabla\mu \geq 0.$$

The Coleman-Noll argument now tells us that

$$\mathbf{S} = \frac{\partial \omega}{\partial \mathbf{F}}(\mathbf{F}, \mu), \qquad c_{\mathrm{R}} = -\frac{\partial \omega}{\partial \mu}(\mathbf{F}, \mu),$$

together with $\mathbf{j}_{\mathrm{R}}(\mathbf{F}, \mu, \mathbf{g}) \cdot \mathbf{g} \leq 0$.

Problem 9.1.1. In the constitutive ansatz (9.15) we did not treat all of the constitutive response functions symmetrically: we allowed $\mathbf{j}_{\rm R}$ to depend on $\nabla\mu$ but not $\omega, \mathbf{S}$ and $c_{\rm R}$. Carry out an analysis that starts from the following set of primitive constitutive relations:

$$\omega = \omega(\mathbf{F}, \mu, \nabla\mu), \qquad \mathbf{S} = \mathbf{S}(\mathbf{F}, \mu, \nabla\mu), \qquad c_{\rm R} = c_{\rm R}(\mathbf{F}, \mu, \nabla\mu), \qquad \mathbf{j}_{\rm R} = \mathbf{j}_{\rm R}(\mathbf{F}, \mu, \nabla\mu),$$

and use the Coleman-Noll argument to show that $\omega, \mathbf{S}$ and $c_{\rm R}$ must in fact be independent of $\nabla\mu$.

9.2 Thermoelasticity.

References:

1 R. Abeyaratne, *Continuum Mechanics*, Volume II in the series *Lecture Notes on The Mechanics of Solids*, http://web.mit.edu/abeyaratne/lecture_notes.html.

2 P. Chadwick, *Continuum Mechanics*, Dover, 1999.

3 M.E. Gurtin, E. Fried and L. Anand, *The Mechanics and Thermodynamics of Continua*, Cambridge University Press, 2010.

4 C. Truesdell and W. Noll, *The Non-Linear Field Theories of Mechanics*, in *Handbuch der Physik* III/3, edited by S. Flügge, Springer, Berlin, 1965.

In addition to the basic fields of deformation $\mathbf{y}(\mathbf{x}, t)$, deformation gradient $\mathbf{F}(\mathbf{x}, t)$ and stress $\mathbf{S}(\mathbf{x}, t)$, one now has to account for the following fields: *temperature* $\theta(\mathbf{x}, t)$, *heat flux* $\mathbf{q}_{\rm R}(\mathbf{x}, t)$, *entropy* $\eta(\mathbf{x}, t)$ and *internal energy* $\epsilon(\mathbf{x}, t)$. The additional physical principles to be considered are the first and second laws of thermodynamics. The amount of heat that crosses a unit area of the boundary $\partial\mathcal{D}_{\rm R}$ in unit time and enters $\mathcal{D}_{\rm R}$ is $-\mathbf{j}_{\rm R}\cdot\mathbf{n}_{\rm R}$ where $\mathbf{j}_{\rm R}(\mathbf{x}, t)$ is the referential heat flux vector (and $\mathbf{n}_{\rm R}$ is a unit outward pointing normal vector on $\partial\mathcal{D}_{\rm R}$). The total rate at which heat enters $\mathcal{D}_{\rm R}$ across $\partial\mathcal{D}_{\rm R}$ is thus

$$\int_{\partial\mathcal{D}_{\rm R}} -\mathbf{q}_{\rm R}\cdot\mathbf{n}_{\rm R}\, dA_x.$$

The *bulk supply* of heat per unit volume per unit time (from sources outside the body) is denoted by $r_{\rm R}(\mathbf{x}, t)$ and so the rate at which heat directly enters the interior of $\mathcal{D}_{\rm R}$ is

$$\int_{\mathcal{D}_{\rm R}} r_{\rm R}(\mathbf{x}, t)\, dV_x.$$

Its role is similar to that of the mechanical body force $\mathbf{b}_{\rm R}$ in that we can usually take it to be prescribable. Let $\eta(\mathbf{x}, t)$ be the entropy per unit volume so that the total entropy in $\mathcal{D}_{\rm R}$

at time t is
$$\int_{\mathcal{D}_R} \eta(\mathbf{x}, t)\, dV_x.$$
Finally, let $\epsilon(\mathbf{x}, t)$ denote the internal energy per unit volume. The total internal energy in $\mathcal{D}_R$ is
$$\int_{\mathcal{D}_R} \epsilon\, dV_x.$$

9.2.1 Basic mechanical equations.

As before, force and moment equilibrium require the usual balance laws
$$\int_{\partial\mathcal{D}_R} \mathbf{S}\mathbf{n}_R\, dA_x + \int_{\mathcal{D}_R} \mathbf{b}_R\, dV_x = \mathbf{0}, \qquad \int_{\partial\mathcal{D}_R} \mathbf{y}\times(\mathbf{S}\mathbf{n}_R)\, dA_x + \int_{\mathcal{D}_R} \mathbf{y}\times\mathbf{b}_R\, dV_x = \mathbf{0},$$
that lead to the associated field equations
$$\mathrm{Div}\,\mathbf{S} + \mathbf{b}_R = \mathbf{0}, \qquad \mathbf{S}\mathbf{F}^T = \mathbf{F}\mathbf{S}^T. \tag{9.16}$$

9.2.2 First law of thermodynamics.

The *first law of thermodynamics* requires
$$\int_{\partial\mathcal{D}_R} \mathbf{S}\mathbf{n}_R\cdot\mathbf{v}\, dA + \int_{\mathcal{D}_R} \mathbf{b}_R\cdot\mathbf{v}\, dV \;+ \int_{\partial\mathcal{D}_R} -\mathbf{q}_R\cdot\mathbf{n}_R\, dA + \int_{\mathcal{D}_R} r_R\, dV \;=\; \frac{d}{dt}\int_{\mathcal{D}_R} \epsilon\, dV\,, \tag{9.17}$$
where $\mathbf{v}(\mathbf{x}, t) = \dot{\mathbf{y}}(\mathbf{x}, t)$ is the particle velocity. The first two terms on the left-hand side represent the rate of mechanical working by the traction on $\partial\mathcal{D}_R$ and the body force on $\mathcal{D}_R$ respectively. The third and fourth terms quantify the heat supplied to $\mathcal{D}_R$ across its boundary and to its interior respectively. The right-hand side is the rate of increase of internal energy. (Since we are considering quasi-static processes and accordingly neglected inertia in the equations of motion, we must neglect kinetic energy in the right-hand side of (9.17).)

We can localize (9.17) by converting the surface integrals to volume integrals using the divergence theorem and then making use of the field equations (9.16). This leads to the local form of the first law,
$$\mathbf{S}\cdot\dot{\mathbf{F}} - \mathrm{Div}\,\mathbf{q}_R + r_R \;=\; \dot{\epsilon}, \tag{9.18}$$
which is to hold at each point in the body and each time. Equation (9.18) is frequently referred to as the *energy equation*.

9.2.3 Dissipation inequality. The second law of thermodynamics.

Note that the temperature θ did not enter into the first law of thermodynamics.

Turning next to the second law, entropy accompanies the flow of heat. Specifically, we take the flow of entropy per unit area per unit time into $\mathcal{D}_{\rm R}$ across its boundary to be $(-\mathbf{q}_{\rm R}/\theta)\cdot\mathbf{n}_{\rm R}$ and the flow directly into the interior of $\mathcal{D}_{\rm R}$ per unit volume per unit time to be $r_{\rm R}/\theta$. Here $\theta > 0$ is the absolute temperature. The total entropy in $\mathcal{D}_{\rm R}$ is given by the volume integral of η. However the rate of increase of entropy and the inflow of entropy need not be balanced. The second law of thermodynamics states that the increase in entropy cannot be less than the inflow of entropy, i.e.

$$\frac{d}{dt}\int_{\mathcal{D}_{\rm R}} \eta \, dV_x \geq \int_{\partial\mathcal{D}_{\rm R}} \frac{(-\mathbf{q}_{\rm R}\cdot\mathbf{n}_{\rm R})}{\theta}\, dA_x + \int_{\mathcal{D}_{\rm R}} \frac{r_{\rm R}}{\theta}\, dV_x \,. \tag{9.19}$$

Thus (when the strict inequality holds) there is a net production of entropy. The entropy inequality plays the role here that the dissipation inequality played in the preceding section.

The entropy inequality (9.19) can be written after using the divergence theorem as

$$\int_{\mathcal{D}_{\rm R}} \left\{\dot\eta + {\rm Div}(\mathbf{q}_{\rm R}/\theta) - r_{\rm R}/\theta\right\}\; dV_x \geq 0. \tag{9.20}$$

This must hold for all $\mathcal{D}_{\rm R}$ and so may be localized to obtain

$$\theta\dot\eta - \frac{1}{\theta}(\mathbf{q}_{\rm R}\cdot\nabla\theta) + {\rm Div}\,\mathbf{q}_{\rm R} - r_{\rm R} \geq 0. \tag{9.21}$$

This local form of the entropy inequality must hold at each point in the body at each time.

9.2.4 Constitutive equations:

We now turn to the constitutive relations. In order to make use of (9.21) in our analysis we need to eliminate the heat supply term $r_{\rm R}$ (and ideally bring in the stress $\mathbf{S}$). This can be achieved by using the energy equation which leads us to

$$\theta\dot\eta + \mathbf{S}\cdot\dot{\mathbf{F}} - \dot\epsilon - \frac{1}{\theta}(\mathbf{q}_{\rm R}\cdot\nabla\theta) \geq 0. \tag{9.22}$$

Note the presence of $\dot{\mathbf{F}}$ and $\dot\eta$ in (9.22) and recall the remark in the first paragraph of Section 9.1.5. Accordingly, suppose that the material is characterized by the following set of constitutive relations:

$$\epsilon = \epsilon(\mathbf{F}, \eta), \qquad \mathbf{S} = \mathbf{S}(\mathbf{F}, \eta), \qquad \theta = \theta(\mathbf{F}, \eta), \qquad \mathbf{q}_{\mathrm{R}} = \mathbf{q}_{\mathrm{R}}(\mathbf{F}, \eta, \nabla\theta). \tag{9.23}$$

The entropy inequality (9.22) can now be used to reduce these relations to simpler forms using the Coleman-Noll argument. By $(9.23)_1$,

$$\dot{\epsilon} = \frac{\partial \epsilon}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} + \frac{\partial \epsilon}{\partial \eta}\dot{\eta},$$

and so we can write the entropy inequality as

$$\left[\theta(\mathbf{F}, \eta) - \frac{\partial \epsilon}{\partial \eta}(\mathbf{F}, \eta)\right] \dot{\eta} + \left[\mathbf{S}(\mathbf{F}, \eta) - \frac{\partial \epsilon}{\partial \mathbf{F}}(\mathbf{F}, \eta)\right] \cdot \dot{\mathbf{F}} - \frac{1}{\theta}(\mathbf{q}_{\mathrm{R}} \cdot \nabla\theta) \geq 0. \tag{9.24}$$

Since (9.24) must hold in all processes, and therefore for all $\dot{\mathbf{F}}$ and $\dot{\eta}$, and since the terms within the square brackets do not involve $\dot{\mathbf{F}}$ and $\dot{\eta}$, it follows that those terms must vanish. This leads to the following constitutive relations for θ and $\mathbf{S}$:

$$\theta = \frac{\partial \epsilon}{\partial \eta}(\mathbf{F}, \eta), \qquad \mathbf{S} = \frac{\partial \epsilon}{\partial \mathbf{F}}(\mathbf{F}, \eta), \tag{9.25}$$

and the entropy inequality (9.24) reduces to

$$\mathbf{q}_{\mathrm{R}}(\mathbf{F}, \eta, \mathbf{g}) \cdot \mathbf{g} \leq 0, \tag{9.26}$$

which is to hold for all vectors $\mathbf{g}$.

Thus a thermoelastic material is characterized by the internal energy $\epsilon(\mathbf{F}, \eta)$ and the heat flux law $\mathbf{q}_{\mathrm{R}}(\mathbf{F}, \eta, \nabla\theta)$ where the latter must be consistent with (9.26). An example of the latter is

$$\mathbf{q}_{\mathrm{R}}(\mathbf{F}, \eta, \nabla\theta) = -\mathbf{K}(\mathbf{F}, \eta)\, \nabla\theta, \tag{9.27}$$

where (9.26) requires the heat conductivity tensor $\mathbf{K}$ to be positive semi-definite.

9.2.5 Alternative form of the constitutive relation.

Suppose that we wish to express the constitutive relations as functions of $\mathbf{F}$ and θ (instead of $\mathbf{F}$ and η). In order to develop this form of the constitutive relations we must trade the term $\dot{\eta}$ in (9.22) for $\dot{\theta}$ and this is achieved by introducing the Legendre transform of ϵ with respect to η and θ defined by

$$\psi \;=\; \epsilon - \theta\eta; \tag{9.28}$$

ψ is called the *Helmholtz free energy* per unit volume. In terms of ψ, one may rewrite (9.22) as

$$-\eta\dot{\theta} + \mathbf{S}\cdot\dot{\mathbf{F}} - \dot{\psi} - \frac{1}{\theta}(\mathbf{q}_{\mathrm{R}}\cdot\nabla\theta) \geq 0.$$

Now suppose that we want to construct constitutive relations in the form

$$\psi = \psi(\mathbf{F},\theta), \qquad \mathbf{S} = \mathbf{S}(\mathbf{F},\theta), \qquad \eta = \eta(\mathbf{F},\theta), \qquad \mathbf{q}_{\mathrm{R}} = \mathbf{q}_{\mathrm{R}}(\mathbf{F},\theta,\nabla\theta).$$

Since

$$\dot{\psi} = \frac{\partial\psi}{\partial\mathbf{F}}\cdot\dot{\mathbf{F}} + \frac{\partial\psi}{\partial\theta}\dot{\theta},$$

the preceding entropy inequality yields

$$\left[-\eta - \frac{\partial\psi}{\partial\theta}\right]\dot{\theta} + \left[\mathbf{S} - \frac{\partial\psi}{\partial\mathbf{F}}\right]\cdot\dot{\mathbf{F}} - \frac{1}{\theta}(\mathbf{q}_{\mathrm{R}}\cdot\nabla\theta) \geq 0. \tag{9.29}$$

The Coleman-Noll argument thus allows us to write the constitutive relations as

$$\eta = -\frac{\partial\psi}{\partial\theta}(\mathbf{F},\theta), \qquad \mathbf{S} = \frac{\partial\psi}{\partial\mathbf{F}}(\mathbf{F},\theta), \tag{9.30}$$

and the entropy inequality reduces to

$$\mathbf{q}_{\mathrm{R}}(\mathbf{F},\theta,\mathbf{g})\cdot\mathbf{g} \leq 0. \tag{9.31}$$

Thus a thermoelastic material is characterized by the Helmholtz free energy $\psi(\mathbf{F},\theta)$ and the heat flux law $\mathbf{q}_{\mathrm{R}}(\mathbf{F},\theta,\nabla\theta)$ where the latter must be consistent with (9.31). An example of the latter is

$$\mathbf{q}_{\mathrm{R}}(\mathbf{F},\theta,\nabla\theta) = -\mathbf{K}(\mathbf{F},\theta)\,\nabla\theta, \tag{9.32}$$

where $\mathbf{K}$ is the positive semi-definite heat conductivity tensor. This is the familiar *Fourier's law.* The Helmholtz free energy function $\psi(\mathbf{F},\theta)$ here is the counterpart of the strain energy function $W(\mathbf{F})$ of the purely mechanical theory.

Thus **in summary**, a boundary-initial value problem for a thermoelastic body involves specifying the body force $\mathbf{b}_{\mathrm{R}}(\mathbf{x},t)$, the heat supply $r_{\mathrm{R}}(\mathbf{x},t)$, the material characterizations $\psi(\mathbf{F},\theta)$ and $\mathbf{q}_{\mathrm{R}}(\mathbf{F},\theta,\nabla\theta)$ and suitable initial and boundary conditions (both mechanical and thermal). One then solves the field equations (9.16), (9.18) subject to the constitutive relations (9.30), (9.32) in order to determine the deformation $\mathbf{y}(\mathbf{x},t)$ and the temperature $\theta(\mathbf{x},t)$.

9.2.6 Worked examples.

Problem 9.2.1. Consider a one-dimensional setting where a thermoelastic material is characterized by the Helmholtz free energy

$$\psi = \psi(\lambda, \theta), \tag{i}$$

and the heat conduction law

$$q = -K(\lambda, \theta)\frac{\partial \theta}{\partial x}, \tag{ii}$$

where $K > 0$ is a constant and we simplicity have dropped the R from $q_{\rm R}$. Here λ is the stretch. Write down and simplify the one-dimensional counterparts of the general equations developed above.

Solution: Partial differentiation with respect to λ and θ will be denoted by subscripts while partial differentiation with respect to x and t will be displayed explicitly.

The counterparts of the constitutive equations $(9.30)_2$ and $(9.30)_1$ for stress and entropy are

$$\sigma = \psi_\lambda(\lambda, \theta), \tag{iii}$$

$$\eta = -\psi_\theta(\lambda, \theta), \tag{iv}$$

and the equilibrium equation and energy equation corresponding to $(9.16)_1$ and (9.18) read

$$\frac{\partial \sigma}{\partial x} + b = 0, \tag{v}$$

$$\sigma\frac{\partial \lambda}{\partial t} - \frac{\partial q}{\partial x} = \frac{\partial \epsilon}{\partial t}. \tag{vi}$$

For convenience we have taken the bulk heat supply to vanish, $r_{\rm R} = 0$, and have dropped the subscript R from $b_{\rm R}$. Here ϵ is the internal energy and it is related to the Helmholtz free energy ψ by

$$\psi = \epsilon - \theta\eta. \tag{vii}$$

The one-dimensional counterpart of the entropy inequality (9.31) is

$$q\frac{\partial \theta}{\partial x} \le 0. \tag{$viii$}$$

We first simplify the energy equation (vi):

$$\sigma\frac{\partial \lambda}{\partial t} - \frac{\partial q}{\partial x} \stackrel{(vi)}{=} \frac{\partial \epsilon}{\partial t} \stackrel{(vii)}{=} \frac{\partial \psi}{\partial t} + \eta\frac{\partial \theta}{\partial t} + \theta\frac{\partial \eta}{\partial t} \stackrel{(i)}{=} \psi_\lambda\frac{\partial \lambda}{\partial t} + \psi_\theta\frac{\partial \theta}{\partial t} + \eta\frac{\partial \theta}{\partial t} + \theta\frac{\partial \eta}{\partial t} =$$

$$\stackrel{(iii),(iv)}{=} \sigma\frac{\partial \lambda}{\partial t} - \eta\frac{\partial \theta}{\partial t} + \eta\frac{\partial \theta}{\partial t} + \theta\frac{\partial \eta}{\partial t},$$

which reduces to

$$-\frac{\partial q}{\partial x} = \theta \frac{\partial \eta}{\partial t}.$$

Substituting (ii) and (iv) into this gives

$$\frac{\partial}{\partial x}\left(K\frac{\partial \theta}{\partial x}\right) = -\theta\psi_{\theta\theta}\frac{\partial \theta}{\partial t} - \theta\psi_{\lambda\theta}\frac{\partial \lambda}{\partial t}. \qquad \square \qquad (ix)$$

The equilibrium equation (v) in view of the constitutive relation (iii) can be written as

$$\psi_{\lambda\lambda}\frac{\partial \lambda}{\partial x} + \psi_{\lambda\theta}\frac{\partial \theta}{\partial x} + b = 0. \qquad \square \qquad (x)$$

Equations $(ix), (x)$ are two partial differential equations for the stretch $\lambda(x,t)$ and the temperature $\theta(x,t)$.

Problem 9.2.2. Consider the particular thermoelastic material characterized by

$$\psi(\lambda,\theta) = \frac{1}{2}\mu(\lambda-1)^2 - c\beta(\theta-\theta_0)(\lambda-1) - c\,\theta\ln(\theta/\theta_0), \qquad (xi)$$

$$K(\lambda,\theta) = K \text{ (constant)}, \qquad (xii)$$

where θ_0 is a fixed ("reference") temperature and μ, c, β and K are constant material parameters. Further reduce the equations obtained in Problem 9.2.1. Interpret the four material parameters μ, c, K and β.

Solution: Observe that $\psi = 0$ when $\lambda = 1$ and $\theta = \theta_0$. From (iii) and (ix) the constitutive relation for stress is

$$\sigma = \mu(\lambda-1) - c\beta(\theta-\theta_0). \qquad (xiii)$$

The parameter μ is therefore the *elastic modulus*. On writing $(xiii)$ as

$$\lambda - 1 = \frac{\sigma}{\mu} + \frac{c\beta}{\mu}(\theta-\theta_0)$$

we see that $c\beta/\mu$ is the *coefficient of thermal expansion*. Observe from $(xiii)$ that $\sigma = 0$ when $\lambda = 1$ and $\theta = \theta_0$.

Since $q = -K\,\partial\theta/\partial x$, K is the *thermal conductivity*.

The various second derivatives of $\psi(\lambda,\theta)$ can be calculated from (xi). They are

$$\psi_{\lambda\lambda} = \mu, \qquad \psi_{\lambda\theta} = -c\beta, \qquad \psi_{\theta\theta} = -c/\theta. \qquad (xiv)$$

Substituting (xiv) into the energy equation (ix) yields

$$K\frac{\partial^2\theta}{\partial x^2} = c\frac{\partial\theta}{\partial t} + c\beta\theta\frac{\partial\lambda}{\partial t}. \qquad \square \tag{xv}$$

Observe that in general, the energy equation involves both the stretch and the temperature (and that the stretch drops out only if $\beta = 0$). If $\beta = 0$ this simplifies to

$$K\frac{\partial^2\theta}{\partial x^2} = c\frac{\partial\theta}{\partial t},$$

the so-called heat equation.

If one calculates the internal energy using $\epsilon = \psi + \theta\eta = \psi - \theta\psi_\theta$ and (i), one finds

$$\epsilon = \frac{1}{2}\mu(\lambda - 1)^2 - c\beta\theta_0(\lambda - 1) + c\,\theta.$$

Therefore $c = \partial\epsilon(\lambda, \theta)/\partial\theta$ and so c can be identified with the *specific heat* of the material (or more accurately the specific heat at constant stretch).

Substituting (xi) into (x) allows us to write the equilibrium equation as

$$\mu\frac{\partial\lambda}{\partial x} - c\beta\frac{\partial\theta}{\partial x} + b = 0. \qquad \square \tag{xvi}$$

Finally, if we write the one-dimensional motion as

$$y = y(x, t) = x + u(x, t)$$

where $u(x, t)$ is the displacement, the stretch is

$$\lambda = 1 + \frac{\partial u}{\partial x}.$$

Therefore the energy equation (xv) and the equilibrium equation (xvi) can be written as

$$K\frac{\partial^2\theta}{\partial x^2} = c\frac{\partial\theta}{\partial t} + c\beta\theta\frac{\partial^2 u}{\partial t\partial x}, \qquad \mu\frac{\partial^2 u}{\partial x^2} - c\beta\frac{\partial\theta}{\partial x} + b = 0. \qquad \square$$

This is a pair of coupled partial differential equations for $u(x, t)$ and $\theta(x, t)$.

9.3 Exercises.

Problem 9.1. Consider a one-dimensional setting where the body (the hydrogel) occupies the region $\mathcal{R}_{\mathrm{R}} = [0, L]$ in a reference configuration. Perhaps it is a slab of thickness L that is infinite in its other dimensions. On one side of the body, the region $x > L$, is an infinite reservoir of solvent at the fixed chemical potential μ_∞ and pressure p_∞. Therefore

$$\mu(L,t) = \mu_\infty, \qquad \sigma(L,t) = -p_\infty \qquad \text{for} \quad t > 0. \tag{i}$$

The left-hand end of the body is fixed and so the displacement there is zero; moreover it is impermeable to the solvent and so the solvent flux there vanishes:

$$u(0,t) = 0, \qquad j_{\mathrm{R}}(0,t) = 0 \qquad \text{for} \quad t > 0. \tag{ii}$$

Initially, the body is undeformed which tell us that the displacement vanishes; it is also dry which implies that the solvent concentration vanishes:

$$u(x,0) = 0, \qquad c_{\mathrm{R}}(x,0) = 0 \qquad \text{for} \quad 0 < x < L. \tag{iii}$$

Equations (i) and (ii) are the boundary conditions, (iii) are the initial conditions. (Question: do you also need to know whether the body is at rest at the initial instant so that $\dot{u}(x,0) = 0$?) Ignore any solvent supply and body force: $r_{\mathrm{R}} = 0, b_{\mathrm{R}} = 0$.

Write down the one-dimensional counterparts of the general equations for a hydrogel given in Section 9.1. Take the hydrogel to be characterized by a free energy function $\psi(\lambda, c_{\mathrm{R}})$ and a solvent flux law $j_{\mathrm{R}} = -M\,\frac{\partial \mu}{\partial x}$ where the mobility $M > 0$ is a constant. You may choose an explicit (not unreasonable) function $\psi(\lambda, c_{\mathrm{R}})$. Note: if you decide to use the function ψ given in Problem 9.5 keep in mind that it is for a material with the constraint $\det \mathbf{F} = 1 + \nu c_{\mathrm{R}}$.

Calculate the displacement and solvent concentration $u(x,t)$ and $c_{\mathrm{R}}(x,t)$ for $0 \le x \le L, t > 0$. What happens when $t \to \infty$?

Problem 9.2. Formulation of the equations for a hydrogel with respect to the *current configuration.* A part $\mathcal{P}$ of the body occupies a region $\mathcal{D}_t$ at time t and $\mathbf{y}$ is the position vector of a particle in that configuration.

(a) Let $c(\mathbf{y}, t)$ denote the solvent concentration per unit current volume so that the number of solvent molecules in $\mathcal{D}_t$ at time t is

$$\int_{\mathcal{D}_t} c(\mathbf{y}, t)\, dV_y.$$

Show that

$$c_{\mathrm{R}} = c\mathrm{J} \qquad \text{where } \mathrm{J} = \det \mathbf{F}.$$

(b) Let $\mathbf{j}(\mathbf{y}, t)$ be the solvent flux vector characterizing the number of solvent molecules per unit current area crossing into $\mathcal{D}_t$ in unit time across its boundary so that the total solvent flux across $\partial\mathcal{D}_t$ is

$$\int_{\partial\mathcal{D}_t} -\mathbf{j}\cdot\mathbf{n}\, dA_y.$$

Here $\mathbf{n}$ is the unit outward-pointing normal vector on the boundary $\partial\mathcal{D}_t$. Show that

$$\mathbf{j}_{\mathrm{R}} = \mathrm{J}\,\mathbf{F}^{-1}\mathbf{j}.$$

(c) Let $r(\mathbf{y}, t)$ be the number of solvent molecules per unit current volume directly entering the interior of $\mathcal{D}_t$ per unit time (from sources outside the body). Show that

$$r_{\mathrm{R}} = r\mathrm{J} \qquad \text{where } \mathrm{J} = \det\mathbf{F}.$$

(d) Show that the conservation of solvent molecules requires the balance law

$$\frac{d}{dt}\int_{\mathcal{D}_t} c\, dV_y = \int_{\partial\mathcal{D}_t} -\mathbf{j}\cdot\mathbf{n}\, dA_y \;+\; \int_{\mathcal{D}_t} r\, dV_y, \tag{9.33}$$

whose associated field equation is

$$\frac{\partial}{\partial t}c(\mathbf{y}, t) + \mathrm{div}\,(c\mathbf{v}) + \mathrm{div}\,\mathbf{j} = r. \tag{9.34}$$

(e) If $\mu(\mathbf{y}, t)$ is the chemical potential of a solvent molecule show that the rate of increase of chemical energy of $\mathcal{D}_t$ due to the influx of solvent molecules is

$$\int_{\partial\mathcal{D}_t} -\mu\mathbf{j}\cdot\mathbf{n}\, dA_y + \int_{\mathcal{D}_t} \mu r\, dV_y,$$

and thus show that the dissipation per unit current volume, $\Delta(\mathbf{y}, t)$, obeys the following global inequality

$$\begin{aligned} \textstyle\int_{\mathcal{D}} \Delta\, dV_y \quad &\stackrel{\text{def}}{=} \quad \int_{\partial\mathcal{D}} \mathbf{T}\mathbf{n}\cdot\mathbf{v}\, dA_y + \int_{\mathcal{D}} \mathbf{b}\cdot\mathbf{v}\, dV_y + \\ &+ \quad \int_{\partial\mathcal{D}} -\mu\mathbf{j}\cdot\mathbf{n}\, dA_y + \int_{\mathcal{D}} \mu r\, dV_y \;-\; \frac{d}{dt}\int_{\mathcal{D}} \psi/\mathrm{J}\, dV_y \;\geq 0 \end{aligned} \tag{9.35}$$

where $\mathbf{b}_{\mathrm{R}} = \mathrm{J}\mathbf{b}$ and $\mathbf{T}$ is the Cauchy stress tensor. Show that the local version of this inequality is

$$\mathbf{T}\cdot\mathbf{L} - \mathbf{j}\cdot\mathrm{grad}\,\mu + \mu\dot{c} + \mu c\,\mathrm{div}\,\mathbf{v} \geq \frac{1}{\mathrm{J}}\,\dot{\psi} \tag{9.36}$$

Problem 9.3. *Current polymer volume fraction.* Let ϕ be the volume of polymer per unit volume in the current configuration. Show that

$$\phi = 1 - \frac{\nu c_{\mathrm{R}}}{\det\mathbf{F}}, \tag{9.37}$$

where ν is the volume of a solvent molecule (assumed to be the same in any configuration).

Problem 9.4. Suppose that the polymeric matrix itself is incompressible (and the volume ν of a solvent molecule is the same in all configurations). This does *not* mean that the volume of a material region $\mathcal{D}_t$ does not increase, rather, that it increases solely due to the addition of solvent molecules. Show that the following relation between the deformation and solvent concentration,

$$\det \mathbf{F} = 1 + \nu c_{\mathrm{R}}, \tag{9.38}$$

is consistent with this constraint.

Problem 9.5. Recall that in the purely mechanical theory, the constitutive relation for stress,

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}},$$

had to be modified if the body could only undergo motions in which $\mathrm{J} = \det \mathbf{F} = 1$ (i.e. if it was incompressible). Carry out a similar modification to the constitutive relations for a hydrogel if it can only undergo processes in which $\det \mathbf{F} = 1 + \nu c_{\mathrm{R}}$ and show that

$$\mathbf{S} = \frac{\partial \psi}{\partial \mathbf{F}} - p\mathrm{J}\mathbf{F}^{-T}, \qquad \mu = \frac{\partial \psi}{\partial c_{\mathrm{R}}} + p\nu,$$

where the constant ν is the volume of a solvent molecule.

Remark 1: Note that the constraint $\det \mathbf{F} = 1 + \nu c_{\mathrm{R}}$ is not purely kinematic since it involves both the deformation and the solvent concentration.

Remark 2: The Helmholtz free energy function for a hydrogel is frequently taken to have the specific form

$$\psi(\mathbf{F}, c_{\mathrm{R}}) = W(\mathbf{F}) + \mu_{\mathrm{R}} c_{\mathrm{R}} + \psi_m(c_{\mathrm{R}}) \qquad \text{where} \quad \psi_m(c_{\mathrm{R}}) = k_B \theta c_{\mathrm{R}} \left[\ln\left(\frac{\nu c_{\mathrm{R}}}{1 + \nu c_{\mathrm{R}}}\right) + \frac{\chi}{1 + \nu c_{\mathrm{R}}} \right].$$

Here the constant parameter μ_{R} is the chemical potential of the pure solvent (in the absence of the polymer), $\psi_m(c_{\mathrm{R}})$ is the chemical potential "due to mixing" and the constant χ is known as the Flory-Huggins parameter. The separable form of ψ into one part that depends on $\mathbf{F}$ and not c_{R} and a second part that depends on c_{R} and not $\mathbf{F}$ might lead one to expect the mechanical and chemical problems to be decoupled. This is incorrect – they are in fact coupled through the constraint $\det \mathbf{F} = 1 + \nu c_{\mathrm{R}}$.

Index