HANDBOOK OF GEOPHYSICAL EXPLORATION
SEISMIC EXPLORATION

Klaus Helbig *and* Sven Treitel, Editors

VOLUME 34

Seismic Waves and Rays in Elastic Media

by M.A. SLAWINSKI

PERGAMON

HANDBOOK OF GEOPHYSICAL EXPLORATION

SEISMIC EXPLORATION

VOLUME 34

SEISMIC WAVES AND RAYS IN
ELASTIC MEDIA

HANDBOOK OF GEOPHYSICAL EXPLORATION

SEISMIC EXPLORATION

Editors: Klaus Helbig and Sven Treitel

PUBLISHED VOLUMES

1984 – Mathematical Aspects of Seismology. 2nd Enlarged Edition (M. Båth and A.J. Berkhout)*
1984 – Seismic Instrumentation (M. Pieuchot) ISBN 0-08-036944-8
1984 – Seismic Inversion and Deconvolution (a) Classical Methods (E.A. Robinson)*
1985 – Vertical Seismic Profiling (a) Principles. 2nd Enlarged Edition (B.A. Hardage)*
1987 – Pattern Recognition & Image Processing (F. Aminzadeh)*
1987 – Seismic Stratigraphy (B.A. Hardage)*
1987 – Production Seismology (J.E. White and R.L. Sengbush)*
1989 – Supercomputers in Seismic Exploration (E. Eisner)*
1994 – Seismic Coal Exploration (b) In-Seam Seismics (L. Dresen and H. Rüter)*
1994 – Foundations of Anisotropy for Exploration Seismics (K. Helbig) ISBN 0-08-037224-4
1998 – Physical Properties of Rocks: Fundamentals and Principles of Petrophysics (J.H. Schön) ISBN 0-08-041008-1
1998 – Shallow High-Resolution Reflection Seismics (J. Brouwer and K. Helbig) ISBN 0-08-043197-6
1999 – Seismic Inversion and Deconvolution (b) Dual-Sensor Technology (E.A. Robinson) ISBN 0-08-043627-7
2000 – Vertical Seismic Profiling: Principles. 3d Updated and Revised Edition (B.A. Hardage) ISBN 0-08-043518-1
2001 – Seismic Signatures and Analysis of Reflection Data in Anisotropic Media (I. Tsvankin) ISBN 0-08-043649-8
2001 – Computational Neural Networks for Geophysical Data Processing (M.M. Poulton) ISBN 0-08-043986-1
2001 – Wave Fields in Real Media: Wave Propagation in Anisotropic, Anelastic and Porous Media (J.M. Carcione) ISBN 0-08-043929-2
2002 – Multi-Component VSP Analysis for Applied Seismic Anisotropy (C. MacBeth) ISBN 0-08-0424439-2
2002 – Nuclear Magnetic Resonance. Petrophysical and Logging Applications (K.J. Dunn, D.J. Bergman and G.A. LaTorraca) ISBN 0-08-043880-6
2003 – Seismic Amplitude Inversion in Reflection Tomography (Y. Wang) ISBN 0-08-044243-9

* *Book out of print.*

SEISMIC EXPLORATION

Volume 34

SEISMIC WAVES AND RAYS IN ELASTIC MEDIA

by

M.A. Slawinski
Department of Earth Sciences
Memorial University of Newfoundland
St. John's, NF
Canada

2003

PERGAMON
An Imprint of Elsevier Science
Amsterdam – Boston – Heidelberg – London – New York – Oxford
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo

ELSEVIER SCIENCE Ltd
The Boulevard, Langford Lane
Kidlington, Oxford OX5 1GB, UK

First edition 2003

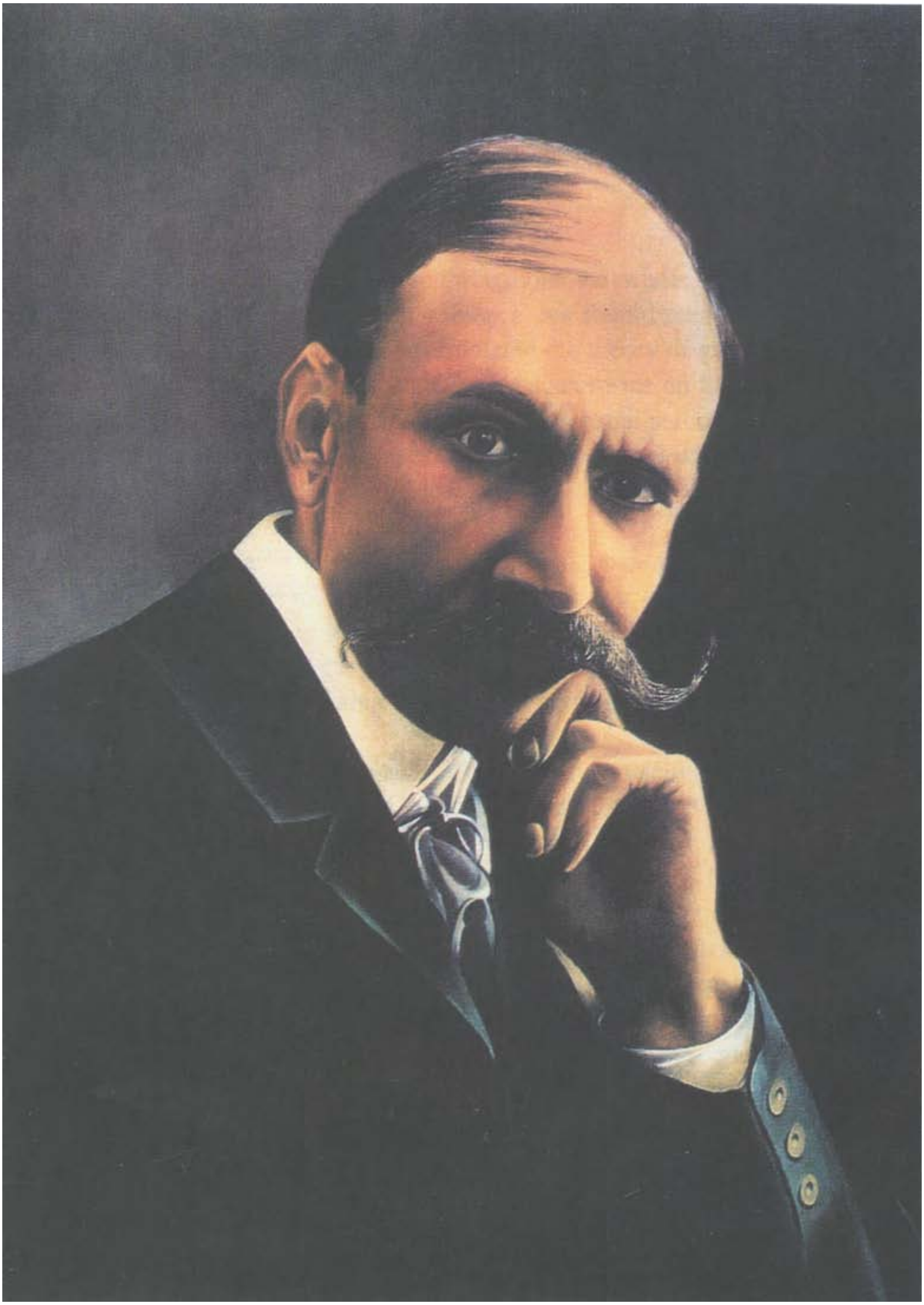
Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

British Library Cataloguing in Publication Data
A catalogue record from the British Library has been applied for.

ISBN: 0-08-043930-6
ISSN: 0950-1401 (Series)

Portrait of M.P. Rudzki on page vi reproduced by permission of Ida Czechowska
Symbol appearing at the end of each chapter by permission of Dwight Braunberger
Photograph of M.A. Slawinski on page 400 by permission of Ned Pratt

⊗ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.



This Page Intentionally Left Blank

Dedication

This book is dedicated to the scientific spirit and accomplishments of Maurycy Pius Rudzki, Chair of Geophysics at the Jagiellonian University, where, in 1895, the first Institute of Geophysics was created.

In order to interpret the wealth of detail contained in the seismographic record of a distant earthquake, we must know the path or trajectory of each ray that leaves the focus or origin in any given direction. An indirect solution of this problem was attempted by several earlier investigators, prominent among whom were Rudzki and Benndorf.

James B. Macelwane (1936) Introduction to theoretical seismology: Geodynamics

Seismological studies appear to have stimulated Rudzki to make the first quantitative calculations on elastic waves.

Michael J.P. Musgrave (1970) Crystal acoustics: Introduction to the study of elastic waves and vibrations in crystals

In the first decade of the [twentieth] century M.P. Rudzki in Cracow began to investigate the consequences of anisotropy in the earth for seismic waves. As far as I can make out, he was the first to determine the wave surface for elastic waves in an anisotropic solid.

Klaus Helbig (1994) Foundations of anisotropy for exploration seismics

This Page Intentionally Left Blank

Acknowledgements

The author wishes to acknowledge the substantial improvements of this book that have resulted from numerous collaborations.

The manuscript of this book originated as the notes for a graduate course in the Department of Mechanical Engineering at the University of Calgary. The quality of these notes was improved by collaborations with Rachid Ait-Haddou, Dan Calistrate, Nilanjan Ganguli, Hugh Geiger, Bill Goodway, Jeff Grossman, Marcelo Epstein, Xinxiang Li, John Parkin and Paul Webster.

The manuscript of this book has reached its final form as a result of collaborations with Andrej Bóna, Len Bos, Nelu Bucataru, David Dalton, Peter Gibson, Andrzej Hanyga, Klaus Helbig, Misha Kotchetov, Ed Krebes, Michael Rochester, Yves Rogister, Raphaël Slawinski, Jędrzej Śniatycki and Chad Wheaton.

Editorial revisions of the text with Cathy Beveridge, from the original sketch to the final draft, enhanced the structure and coherence of this book.

This Page Intentionally Left Blank

Preface

Il ne suffit pas d'observer, il faut se servir de ses observations, et pour cela il faut généraliser. [...] Le savant doit ordonner; on fait la science avec des faits comme une maison avec des pierres; mais une accumulation de faits n'est pas plus une science qu'un tas de pierres n'est une maison.[1]

Henri Poincaré (1902) La Science et l'Hypothèse

Theoretical formulations of applied seismology are substantiated by observable phenomena. Reciprocally, our perception and understanding of these phenomena necessitate rigorous descriptions of physical behaviours. This book emphasizes the interdependence of mathematical formulation and physical meaning in the description of seismic phenomena. The purpose of this book is to use aspects of continuum mechanics, wave theory and ray theory to explain phenomena resulting from the propagation of seismic waves.

The book is divided into three main parts: *Elastic continua*, *Waves and rays* and *Variational formulation of rays*. There is also a fourth part, which consists of *Appendices*. In *Part I*, we use continuum mechanics to describe the material through which seismic waves propagate, and to formulate a system of equations to study the behaviour of such a material. In *Part II*, we use these equations to identify the types of body waves propagating in elastic continua as well as to express their velocities and displacements in terms of the properties of these continua. To solve the equations of motion in anisotropic inhomogeneous continua, we use the high-frequency approximation and, hence, establish the concept of a ray. In *Part III*, we show that, in elastic continua, a ray is tantamount to a trajectory along which a seismic signal propagates in accordance with the variational principle of stationary

[1] It is not enough to observe. One must use these observations, and for this purpose one must generalize. [...] The scientist must organize [knowledge]; science is composed of facts as a house is composed of bricks; but an accumulation of facts is no more a science than a pile of bricks is a house.

traveltime. Consequently, many seismic problems in elastic continua can be conveniently formulated and solved using the calculus of variations. In *Part IV*, we describe two mathematical concepts that are used in the book, namely, homogeneity of a function and Legendre's transformation. This part also contains a *List of symbols.*

The book contains an *Index* that focuses on technical terms. The purpose of this index is to contribute to the coherence of the book and to facilitate its use as a study manual and a reference text. Numerous terms are grouped to indicate the relations among their meanings and nomenclatures. Some references to selected pages are marked in bold font. These pages contain a defining statement of a given term.

This book is intended for senior undergraduate and graduate students as well as scientists interested in quantitative seismology. We assume that the reader is familiar with linear algebra, differential and integral calculus, vector algebra and vector calculus, tensor analysis, as well as ordinary and partial differential equations. The chapters of this book are intended to be studied in sequence. In that manner, the entire book can be used as a manual for a single course. If the variational formulation of ray theory is not to be included in such a course, the entire *Part III* can be omitted.

Each part begins with an *Introduction*, which situates the topics discussed therein in the overall context of the book as well as in a broader scientific context. Each chapter begins with *Preliminary remarks*, which state the motivation for the specific concepts discussed therein, outline the structure of the chapter and provide links to other chapters in the book. Each chapter ends with *Closing remarks*, which specify the limitations of the concepts discussed and direct the reader to related chapters. Each chapter is followed by *Exercises* and their solutions, some of which are referred to in the main text. Reciprocally, the footnotes attached to these exercises refer the reader to the sections in the main text, where a given exercise is mentioned. Also, throughout the book, footnotes refer the reader to specific sources included in the *Bibliography.*

"Seismic waves and rays in elastic media" strives to respect the scientific spirit of Rudzki, described in the following statement[2] of Marian Smoluchowski, Rudzki's colleague and friend.

> Tematyka geofizyczna musiała nęcić Rudzkiego, tak wielkiego, fantastycznego miłośnika przyrody, z drugiej zaś strony ta właśnie

[2]Smoluchowski, M., (1916) Maurycy Rudzki jako geofizyk / Maurycy Rudzki as a geophysicist: Kosmos, **41**, 105 – 119

nauka odpowiadała najwybitniejszej właściwości umysłu Rudzkiego, jego dążeniu do matematycznej ścisłości w rozumowaniu.[3]

St. John's, Newfoundland
Spring 2003

[3]The subject of geophysics must have attracted Rudzki, a great lover of nature. Also, this very science accommodated the most outstanding quality of Rudzki's mind, his striving for mathematical rigour in reasoning.

This Page Intentionally Left Blank

Contents

This Page Intentionally Left Blank

Part I

Elastic continua

This Page Intentionally Left Blank

Introduction to Part I

> One conceives the causes of all natural effects in terms of mechanical motion. This, in my opinion, we must necessarily do, or else renounce all hopes of ever comprehending anything in Physics.[4]
>
> *Christian Huygens* (1690) *Treatise on light: In which are explained the causes of that which occurs in reflection and refraction*

Our focus in this book is the description of seismic phenomena in elastic media.

The physical basis of seismic wave propagation lies in the interaction of grains within the material through which deformations propagate. It is difficult to individually describe all these interactions among the grains. However, since our experimental data are the result of a large number of such interactions, we can consider these interactions as an ensemble and describe seismic wave propagation through a granular material in terms of wave propagation through a medium that is continuous. We refer to such a medium as a continuum.

Consequently, in this book, we follow the concepts of continuum mechanics where any material is described by a continuum. A continuum is formulated mathematically in terms of continuous functions representing the average properties of many microscopic objects forming the actual material. In this context, all the associated quantities become scalar, vector or tensor fields, and the formulated problems are governed by differential equations.

Using the methods of continuum mechanics, we adhere to the following statement of Kennett from his book "The seismic wavefields".

[4]Readers interested in the modern view of this statement, in the context of analytical mechanics, might refer to Born, M., and Wolf, E., (1999) Principles of optics (*7th* edition): Cambridge University Press, p. xxix.

> We adopt a viewpoint in which the details of the microscopic structure of the medium through which seismic waves propagate is ignored. The material is supposed to comprise a continuum of which every subdivision possesses the macroscopic properties.

At the beginning of *Part I*, we formulate the methods for describing deformations of continua and we introduce the concept of strain. This is followed by a description of forces acting within the continuum and the introduction of the concept of stress. We also derive the fundamental equations, namely, the equation of continuity and the equations of motion, which result from the conservation of mass and the balance of linear momentum, respectively.

To supplement these equations and, hence, to formulate a determined system that governs the behaviour of a continuum, we consider a particular class of continua, which is general enough to be of significance in applied seismology. Our attention focuses on elastic continua. In general, a continuum is characterized by its deformation in response to applied loads. In this book, we assume that this response can be adequately described by linear stress-strain equations. Also, we assume that all the energy expended on deformation is transformed into potential energy, which is stored in the deformed continuum. Consequently, upon the removal of the load, the stored energy allows this continuum to return to its undeformed state.

The original formulation of the theory of continuum mechanics can be dated to the second half of the eighteenth century and is associated with the work of Leonhard Euler. At the beginning of the nineteenth century, further development was achieved by Augustin-Louis Cauchy and George Green, as well as several other European scientists. The modern development of the theory of continuum mechanics is mainly associated with the work of American scientists, in particular, the work of Walter Noll, Ronald Rivlin and Clifford Truesdell, in the second half of the twentieth century.

We should also note that a too literal interpretation of the concept of continuum can lead to inaccurate conclusions. This can be illustrated by an example given by Schrödinger in his book entitled "Nature and the Greeks".

> Let a cone be cut in two by a plane parallel to its base; are the two circles, produced by the cut on the two parts equal or unequal? If unequal, then, since this would hold for any such a cut, the ascending part of the cone's surface would not be smooth but covered with indentations; if you say equal, then for the same reason, would it not mean that all these parallel sections are equal and thus the cone is a cylinder?

Also, in view of the abstract nature of continuum mechanics, we must carefully consider the definition of exactness of a solution. While exact mathematical solutions to the equations formulated in continuum mechanics exist, the equations themselves are not exact representations of nature since they rely on abstract formulations. Hardy expresses a similar thought in his book entitled "A mathematician's apology".

> It is quite common for a physicist to claim that he has found a 'mathematical proof' that the physical universe must behave in a particular way. All such claims, if interpreted literally, are strictly nonsense. It cannot be possible to prove mathematically that there will be an eclipse tomorrow, because eclipses, and other physical phenomena, do not form part of the abstract world of mathematics.

Nevertheless, the notion of continuum, as it pertains to the theory of elasticity, is particularly useful for seismological purposes because it permits convenient mathematical analysis that gives rise to scientific theory validated by experimental data.

This Page Intentionally Left Blank

Chapter 1

Deformations

> ... au lieu de considérer la masse donnée comme un assemblage d'une infinité de points contigus, il faudra, suivant l'esprit du calcul infinitésimal, la considérer plutôt comme composée d'éléments infiniment petits, qui soient du même ordre de dimension que la masse entière;[1]
>
> *Joseph-Louis Lagrange (1788) Mécanique Analytique*

Preliminary remarks

We begin our study of seismic wave propagation by considering the materials through which these waves propagate. Physical materials are composed of atoms and, hence, the fundamental treatment of this propagation would require the study of interactions among the atoms. At present, such an approach is impractical and, perhaps, impossible with the available mathematical tools. Consequently, we seek a more convenient approach. An alternative approach is offered by continuum mechanics, which allows us to obtain results consistent with observable phenomena without dealing directly with the discrete properties of the materials through which seismic waves propagate.

As all mathematical physics, continuum mechanics utilizes abstract concepts to model physical reality. In a seismological context, the Earth is regarded as a continuum that transmits mechanical disturbances. The notion

[1]... instead of considering a given mass as an assembly of an infinity of neighbouring points, one shall – following the spirit of calculus – consider rather the mass as composed of infinitely small elements, which would be of the same dimension as the entire body;

of continuum allows us to describe the deformations and forces experienced by a deformable body in terms of stresses and strains within a continuum.

We begin this chapter with an explanation of the notion of continuum followed by a description of deformations within it. In particular, we derive the strain tensor, which allows us to describe both a relative change in volume and a change in shape within the continuum.

1.1 Notion of continuum

In continuum mechanics, we choose to disregard the atomic structure of matter and the explicit interactions among particles. The notion of continuum is justified by the assumption that a material is composed of sufficiently closely spaced particles, so that its descriptive functions can be considered to be continuous. In other words, the infinitesimal elements of the material are assumed to possess the same physical properties as the properties observed in macroscopic studies. Although the microscopic structure of real materials is not consistent with the concept of continuum, this idealization provides a useful platform for mathematical analysis, which in turn permits us to model physical reality using abstract concepts.[2]

The concept of continuum allows us to consider materials in such a way that their descriptive functions are continuous and differentiable. In particular, we can define stress at a given point, thereby enabling us to apply calculus to the study of forces within a continuum. This definition and the subsequent application of calculus is associated with the work of Augustin-Louis Cauchy in the first half of the nineteenth century. Instead of studying atomic forces among individual particles, he introduced the notions of stress and strain in a continuum, which resulted in the equations associated with the theory of elasticity.

Using a continuum-mechanics approach to describe seismic wave propagation raises some concerns. In continuum mechanics, the behaviour of a multitude of grains in a portion of a material is discussed by studying the behaviour of the whole ensemble. Consequently, information relating to the grains themselves is lost in the averaging process. In other words, the application of continuum mechanics raises the question whether the loss of

[2]Readers interested in rigorous mathematical foundations of elasticity might refer to Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover. For general aspects of continuum-mechanics formulations, readers might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall.

information about the granular structures of the material allows us to properly represent the macroscopic behaviour of that material. To answer this question, we state that our ability to formulate a coherent theory to accurately describe and predict observable seismic phenomena is a key criterion to justify our usage of the notion of continuum.

1.2 Material and spatial descriptions

1.2.1 Fundamental concepts

While using the concept of continuum, which does not involve any discrete particles, we must carefully consider methods that allow us to describe the displacement of material points within the continuum, where we define a material point as an infinitesimal element of volume that possesses the same physical properties as the properties observed in macroscopic studies. In the context of continuum mechanics, this element of volume is sufficiently large that it contains enough discrete particles of matter to allow us to establish a concept of continuum, while it is sufficiently small to be perceived as a mathematical point.

In continuum mechanics, we can describe such a displacement in at least two ways, namely, by studying material and spatial descriptions.[3] We can observe the displacement either by following a given material point — in other words, following an infinitesimal element of the continuum, which is analogous to following a particle in particle mechanics — or by studying the flow of the continuum across a fixed position, which does not have an analogue in particle mechanics. The first approach is called the material description of the motion while the second one is called the spatial description of the motion. These approaches are also known as the Lagrangian description and the Eulerian description, respectively.[4]

In global geodynamics, the fundamental laws that govern deformations of the Earth necessitate the distinction between the equations derived using the material and the spatial formulations. However, in applied seismology, we can often accurately analyze observable phenomena while ignoring the distinction between the material and the spatial descriptions.

[3]Material and spatial descriptions correspond to the referential and spatial descriptions of Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, p. 138, where the relative description is also discussed.

[4]Readers interested in detailed descriptions of these approaches and their consequences might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, pp. 138 - 145.

To gain insight into the meaning of the material and spatial descriptions, consider a moving continuum and let the observer focus attention on a given material point within the continuum. Suppose the position of a material point at initial time t_0 is given by vector $\mathbf{X}$. Although position vector $\mathbf{X}$ is not a material point, we will refer to a given material point as "material point $\mathbf{X}$", which is a concise way of referring to a material point that at time t_0 occupied position $\mathbf{X}$, as shown in Remark 1.1, which follows Exercise 1.1. At a later time t, the position vector of the material point $\mathbf{X}$ is given by $\mathbf{x}$. Mapping $\mathbf{x}(\mathbf{X}, t)$ gives position, $\mathbf{x}$, of material point $\mathbf{X}$ at time t. This is the material description of the motion, where the value of the independent variable, $\mathbf{X}$, identifies the material point. We assume that, for a given time t, this mapping is one-to-one and is continuous, as well as possessing the continuous inverse. Also, we have to assume that this mapping and its inverse have continuous partial derivatives to whatever order is required. Since we assume that the transition of the material point from the initial position to the present one occurs in a smooth fashion, vector $\mathbf{x}$ is a continuous function of time and, by symmetry, its inverse is also continuous. This inverse can be written as $\mathbf{X}(\mathbf{x}, t)$, which fixes our attention on a given region in space and takes position, $\mathbf{x}$, and time, t, as independent variables.

To define the material and spatial descriptions, consider an orthonormal coordinate system, where

$$x_i = x_i(X_1, X_2, X_3, t), \qquad i \in \{1,2,3\},$$

and

$$X_i = X_i(x_1, x_2, x_3, t), \qquad i \in \{1,2,3\},$$

with the components x_i and X_i being the spatial and material coordinates, respectively. If the arguments of a given function are given in terms of the x_i, we are dealing with a spatial description, while, if they are given in terms of the X_i, we are dealing with a material description.

In general, a physical quantity that characterizes a continuum can be described by a function $f(\mathbf{x}, t)$, which is a spatial description of this quantity, or by a function $F(\mathbf{X}, t)$, which is a material description of this quantity. The material and spatial descriptions are consistent with one another. The relation between f and F is given by $f[\mathbf{x}(\mathbf{X}, t), t] = F(\mathbf{X}, t)$, or by $f(\mathbf{x}, t) = F[\mathbf{X}(\mathbf{x}, t), t]$.

1.2.2 Material time derivative

In view of the previous section, we see that either the material or the spatial description can be used to describe the temporal variation of a given physical quantity. Let us consider time derivatives in the context of either description.

The material description consists of fixing our attention on a given material point $\mathbf{X}$ and observing the variation of the quantity F with time. The time derivative associated with this viewpoint can be written as

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \left.\frac{\mathrm{d}F(\mathbf{X},t)}{\mathrm{d}t}\right|_{\mathbf{X}},$$

which implies that $\mathbf{X}$ is kept constant when we take the derivative of $F(\mathbf{X},t)$ with respect to time, t. Symbol $|_{\mathbf{X}}$ means that the function is evaluated at $\mathbf{X}$.

The spatial description consists of fixing our attention on a given spatial point $\mathbf{x}$ and observing the variation of the quantity f with time. The time derivative associated with this viewpoint can be written as

$$\frac{\partial f}{\partial t} = \left.\frac{\partial f(\mathbf{x},t)}{\partial t}\right|_{\mathbf{x}}.$$

In the context of temporal variations, the material and spatial descriptions are related by the chain rule of differentiation. Considering a three-dimensional continuum, we can write

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}t} &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x_1}\frac{\mathrm{d}x_1}{\mathrm{d}t} + \frac{\partial f}{\partial x_2}\frac{\mathrm{d}x_2}{\mathrm{d}t} + \frac{\partial f}{\partial x_3}\frac{\mathrm{d}x_3}{\mathrm{d}t}\\ &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x_1}v_1 + \frac{\partial f}{\partial x_2}v_2 + \frac{\partial f}{\partial x_3}v_3.\end{aligned}$$

Denoting $\mathbf{v} = [v_1, v_2, v_3]$, where $\mathbf{v}$ is the velocity vector, and invoking the gradient operator, we obtain

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial f}{\partial t} + (\mathbf{v}\cdot\nabla)f = \left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)f, \tag{1.1}$$

where the material and spatial coordinates are related by

$$\mathbf{x} = \mathbf{x}(\mathbf{X},t). \tag{1.2}$$

The term in parentheses on the right-hand side of expression (1.1) is the material time-derivative operator. The material time derivative is a rate of change associated with particular elements of the continuum and measured by an observer travelling with its flow. In other words, the material time derivative is the time derivative with material coordinates held constant. The first term of $(\partial/\partial t + \mathbf{v}\cdot\nabla)$ describes the rate of change at the location $\mathbf{x}$, while the second term describes the rate of change associated with the motion of material points. In general, the material-derivative operator can be applied to a scalar, to a vector, or to a tensor function of position and time coordinates.

1.2.3 Conditions of linearized theory

In general, equations governing wave phenomena in elastic media are nonlinear. However, seismic experiments indicate that important aspects of wave propagation can be adequately described by linear equations, which greatly simplify mathematical formulations. The process of going from nonlinear equations to linear ones is called the linearization process and the resulting theory is the linearized theory. This linearization is achieved by the fact that, under certain assumptions that appear to be satisfied for many seismological studies, the material and spatial descriptions are equivalent to one another.

The linearization allows us to formulate mathematical statements of seismic wave phenomena in a form that is simpler than it would be otherwise possible. In this section, we briefly discuss the conditions that allow us to use linearization. A more detailed description of the linearization process is beyond the scope of this book.[5]

In applied seismology, we often assume that the displacements of material elements resulting from the propagation of seismic waves can be considered as infinitesimal. Such an assumption is used in this entire book. As a consequence of this assumption and in view of the material time derivative, discussed in Section 1.2.2, we can conclude that, while considering displacements, it is unnecessary to distinguish between the material and spatial descriptions.

To arrive at this conclusion, let us consider the notion of displacement using both the material and spatial descriptions. Displacement is the difference between the final position and the initial position. Using the material description, we can write the displacement vector as

$$\mathbf{U}(\mathbf{X},t) = \mathbf{x}(\mathbf{X},t) - \mathbf{X}, \tag{1.3}$$

while using the spatial description, we note that the displacement vector is

$$\mathbf{u}(\mathbf{x},t) = \mathbf{x} - \mathbf{X}(\mathbf{x},t). \tag{1.4}$$

Note that at the initial time, $\mathbf{x} = \mathbf{X}$.

[5] Readers interested in a thorough analysis of physical quantities in the material and spatial descriptions, and the subsequent linearization might refer to Achenbach, J.D., (1973) Wave propagation in elastic solids: North Holland, pp. 11 – 21 and 46 – 47, to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, pp. 497 – 565, and to Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover, pp. 9 – 10 and 226 – 246.

Since the same quantity is given by expressions (1.3) and (1.4), we can write

$$\mathbf{U}(\mathbf{X},t) = \mathbf{u}(\mathbf{x},t), \tag{1.5}$$

where the material and spatial coordinates are related by equation (1.2).

We can develop each component of $\mathbf{U}(\mathbf{X},t)$ into Taylor's series about $\mathbf{x}$ to obtain

$$U_i(\mathbf{X},t) = U_i(\mathbf{X},t)|_{\mathbf{X}=\mathbf{x}} + (\mathbf{X}-\mathbf{x}) \cdot \left[\left.\frac{\partial U_i(\mathbf{X},t)}{\partial X_1}\right|_{\mathbf{X}=\mathbf{x}}, \left.\frac{\partial U_i(\mathbf{X},t)}{\partial X_2}\right|_{\mathbf{X}=\mathbf{x}}, \left.\frac{\partial U_i(\mathbf{X},t)}{\partial X_3}\right|_{\mathbf{X}=\mathbf{x}} \right] + \ldots,$$

where $i \in \{1,2,3\}$. Assuming that the displacement is infinitesimal, namely, $\mathbf{X}-\mathbf{x}$ is vanishingly small, we can consider only the first term of the series. Thus, we can write

$$\mathbf{U}(\mathbf{X},t) \approx \mathbf{U}(\mathbf{x},t). \tag{1.6}$$

Hence, expression (1.3) can be written as

$$\mathbf{U}(\mathbf{x},t) \approx \mathbf{x}(\mathbf{X},t) - \mathbf{X}. \tag{1.7}$$

Since in expression (1.7), $\mathbf{U}$ is a function of $\mathbf{x}$, we rewrite the displacement as a function of $\mathbf{x}$ to get

$$\mathbf{U}(\mathbf{x},t) \approx \mathbf{x} - \mathbf{X}(\mathbf{x},t). \tag{1.8}$$

Comparing expressions (1.4) and (1.8), we see that

$$\mathbf{U}(\mathbf{x},t) \approx \mathbf{u}(\mathbf{x},t).$$

Thus, in view of expression (1.6), we conclude that — for infinitesimal displacements — we can write

$$\mathbf{U}(\mathbf{X},t) \approx \mathbf{u}(\mathbf{x},t). \tag{1.9}$$

To gain insight into the meaning of this result, we examine equations (1.5) and (1.9). Equation (1.5) states that $\mathbf{U} = \mathbf{u}$, with $\mathbf{x}$ related to $\mathbf{X}$ by equation (1.2). Equation (1.9) states that $\mathbf{U} \approx \mathbf{u}$, where we can simply replace $\mathbf{x}$ by $\mathbf{X}$, without invoking equation (1.2). This approximation is illustrated in Exercise 1.2.

Now, let us consider the velocity using both the material and spatial descriptions. To do so, let the physical quantity considered in the material time derivative be given by displacement. In such a case, expression (1.1) becomes

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}=\frac{\partial \mathbf{u}}{\partial t}+(\mathbf{v}\cdot\nabla)\,\mathbf{u}.$$

If both the gradient of the displacement $\mathbf{u}$ and the velocity $\mathbf{v}$ are infinitesimal, we can ignore the second term on the right-hand side to obtain

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}\approx\frac{\partial \mathbf{u}}{\partial t}.$$

Also, let us consider the acceleration using both the material and spatial descriptions. To do so, let the physical quantity considered in the material time derivative be given by velocity. In such a case, expression (1.1) becomes

$$\frac{\mathrm{d}^2\mathbf{U}}{\mathrm{d}t^2}=\frac{\partial^2 \mathbf{u}}{\partial t^2}+(\mathbf{v}\cdot\nabla)\,\frac{\partial \mathbf{u}}{\partial t}.$$

If both the gradient of $\partial\mathbf{u}/\partial t$ and the velocity $\mathbf{v}$ are infinitesimal, we can ignore the second term on the right-hand side to obtain

$$\frac{\mathrm{d}^2\mathbf{U}}{\mathrm{d}t^2}\approx\frac{\partial^2 \mathbf{u}}{\partial t^2}.$$

This property of the time derivative of displacement, which results from the linearized theory, is used, for instance, in the derivation of equations of motion (2.34).

Thus, we can conclude that, under the assumption of infinitesimal displacements of a given element of the continuum, we do not need to distinguish between either the material and spatial coordinates or the material and spatial descriptions of displacements. In other words, $\mathbf{X}\approx\mathbf{x}$ and $\mathbf{U}\approx\mathbf{u}$. Furthermore, if we also assume that the velocities of these displacements are infinitesimal, that the gradients of these displacements are infinitesimal, and that the gradients of these velocities are also infinitesimal, there is no need to distinguish between the material and spatial descriptions while studying velocities and accelerations. In other words, $\mathrm{d}\mathbf{U}/\mathrm{d}t\approx\partial\mathbf{u}/\partial t$ and $\mathrm{d}^2\mathbf{U}/\mathrm{d}t^2\approx\partial^2\mathbf{u}/\partial t^2$, respectively.

Note that the assumptions about the properties of the displacements, gradients of displacements, velocities and gradients of velocities are independent of each other. They result from the physical context in which we consider a given mathematical formulation. For instance, in the context of applied seismology, we assume that the displacement amplitude of a material point is small compared to the wavelength. Also, we assume that the velocity of this displacement is small compared to the wave propagation velocity.[6]

Following our decision to make no distinction between the material and spatial descriptions, we follow the customary notation to describe the coordinates as well as the displacements of a given element of the continuum using lower-case letters. Also, to avoid any confusion, we note that the velocities denoted by v and V, in Parts II and III of the book, refer to the phase velocity and the ray velocity, respectively. They are not directly associated with the velocities of displacements of a given element of the continuum, which we discuss herein.

1.3 Strain

Introductory comments

[7]Seismic waves consist of the propagation of deformations through a material. To study these waves, we wish to describe the associated deformations of the continuum in the context of infinitesimal displacements.

Deformation of a continuum is a change of positions of material points within it relative to each other. If such a change occurs, a continuum is said to be strained. This strain is accompanied by stress. The produced stress resists deformation and attempts to restore the continuum to its unstrained state. The resistance of a continuum to the deformation and the continuum's tendency to restore itself to its undeformed state account for the propagation of seismic waves.

The relation between stress and strain is one of mutual dependence and is an intrinsic concept of elasticity theory. In this theory, applied forces are formulated in terms of a stress tensor, discussed in Chapter 2, while the associated deformations are formulated in terms of a strain tensor, discussed below.

1.3.1 Derivation of strain tensor

In this section we show that the strain tensor relates the states of strain prior to and after the deformation. We also show that the strain tensor is a

[6]Readers interested in details of this linearization might refer to Achenbach, J.D., (1973) Wave propagation in elastic solids: North Holland, pp. 17 – 21.

[7]Readers interested in a thorough description of strain and deformation might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, Chapter 4: Strain and Deformation.

second-rank tensor.[8]

To derive the strain tensor in a three-dimensional continuum, consider two infinitesimally close material points with the coordinates given by $[x, y, z]$ and $[x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z]$. The square of the distance between these points is given by

$$(\mathrm{d}s)^2 = (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 . \tag{1.10}$$

Let the continuum be subjected to deformation. After the deformation, which is described by displacement vector

$$\mathbf{u} = [u_x(x, y, z), u_y(x, y, z), u_z(x, y, z)],$$

the coordinates of the first point are given by

$$\left[x + u_x|_{x,y,z}, y + u_y|_{x,y,z}, z + u_z|_{x,y,z}\right], \tag{1.11}$$

while the coordinates of the second point are given by

$$\begin{aligned} &\left[x + \mathrm{d}x + u_x|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z},\right. \\ &y + \mathrm{d}y + u_y|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z}, \\ &\left.z + \mathrm{d}z + u_z|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z}\right], \end{aligned} \tag{1.12}$$

where the arguments in the subscripts are the values at which the components of function $\mathbf{u}$ are evaluated. Subtracting the components given in expression (1.11) from the corresponding components given in expression (1.12), we obtain the difference between the corresponding coordinates of the two points, namely,

$$\begin{aligned} &\left[\mathrm{d}x + u_x|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z} - u_x|_{x,y,z},\right. \\ &\mathrm{d}y + u_y|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z} - u_y|_{x,y,z}, \\ &\left.\mathrm{d}z + u_z|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z} - u_z|_{x,y,z}\right]. \end{aligned} \tag{1.13}$$

In view of infinitesimal displacements, the components of $\mathbf{u}$ that are evaluated at $(x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z)$ can be approximated by the first two terms of Taylor's series about (x, y, z), namely,

$$u_x|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z} \approx u_x|_{x,y,z} + \left.\frac{\partial u_x}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_x}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_x}{\partial z}\right|_{x,y,z} \mathrm{d}z,$$

[8]Both terms "rank" and "order" are commonly used to describe the number of indices of a tensor. In this book, we use the former term since it does not appear in any other context, while the latter term is used in the context of differential equations. Note that although the term "rank" also has a specific meaning in matrix algebra, we do not use it in such a context in this book.

$$u_y\big|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z} \approx u_y\big|_{x,y,z} + \left.\frac{\partial u_y}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_y}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_y}{\partial z}\right|_{x,y,z} \mathrm{d}z$$

and

$$u_z\big|_{x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z} \approx u_z\big|_{x,y,z} + \left.\frac{\partial u_z}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_z}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_z}{\partial z}\right|_{x,y,z} \mathrm{d}z.$$

Inserting these Taylor's series terms into expression (1.13) and simplifying, we obtain the approximation for the difference of the corresponding coordinates of the two points after the deformation, namely,

$$\begin{aligned}\Bigg[&\mathrm{d}x + \left.\frac{\partial u_x}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_x}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_x}{\partial z}\right|_{x,y,z} \mathrm{d}z,\\ &\mathrm{d}y + \left.\frac{\partial u_y}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_y}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_y}{\partial z}\right|_{x,y,z} \mathrm{d}z,\\ &\mathrm{d}z + \left.\frac{\partial u_z}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_z}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_z}{\partial z}\right|_{x,y,z} \mathrm{d}z\Bigg].\end{aligned}$$

Hence, the square of the distance between the two points after the deformation can be approximated by

$$\begin{aligned}(\mathrm{d}\breve{s})^2 \approx &\left(\mathrm{d}x + \left.\frac{\partial u_x}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_x}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_x}{\partial z}\right|_{x,y,z} \mathrm{d}z\right)^2\\ &+ \left(\mathrm{d}y + \left.\frac{\partial u_y}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_y}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_y}{\partial z}\right|_{x,y,z} \mathrm{d}z\right)^2\\ &+ \left(\mathrm{d}z + \left.\frac{\partial u_z}{\partial x}\right|_{x,y,z} \mathrm{d}x + \left.\frac{\partial u_z}{\partial y}\right|_{x,y,z} \mathrm{d}y + \left.\frac{\partial u_z}{\partial z}\right|_{x,y,z} \mathrm{d}z\right)^2.\end{aligned}$$

Squaring the parentheses on the right-hand side and — in view of infinitesimal gradients of the displacement — neglecting the terms that contain the products of two derivatives, we obtain

$$\begin{aligned}(\mathrm{d}\breve{s})^2 \approx\ &(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 \qquad (1.14)\\ &+ 2\left(\left.\frac{\partial u_x}{\partial x}\right|_{x,y,z} (\mathrm{d}x)^2 + \left.\frac{\partial u_y}{\partial y}\right|_{x,y,z} (\mathrm{d}y)^2 + \left.\frac{\partial u_z}{\partial z}\right|_{x,y,z} (\mathrm{d}z)^2\right.\\ &+ \left.\frac{\partial u_x}{\partial y}\right|_{x,y,z} \mathrm{d}x\mathrm{d}y + \left.\frac{\partial u_x}{\partial z}\right|_{x,y,z} \mathrm{d}x\mathrm{d}z + \left.\frac{\partial u_y}{\partial x}\right|_{x,y,z} \mathrm{d}x\mathrm{d}y\\ &\left.+ \left.\frac{\partial u_y}{\partial z}\right|_{x,y,z} \mathrm{d}y\mathrm{d}z + \left.\frac{\partial u_z}{\partial x}\right|_{x,y,z} \mathrm{d}x\mathrm{d}z + \left.\frac{\partial u_z}{\partial y}\right|_{x,y,z} \mathrm{d}y\mathrm{d}z\right),\end{aligned}$$

which is the expression for the square of the distance between the two points after the deformation.

Using expressions (1.10) and (1.14), we obtain the difference in the square of the distance between the two points that results from the deformation, namely,

$$(\mathrm{d}\check{s})^2 - (\mathrm{d}s)^2 \approx 2\left[\left.\frac{\partial u_x}{\partial x}\right|_{x,y,z}(\mathrm{d}x)^2 + \left.\frac{\partial u_y}{\partial y}\right|_{x,y,z}(\mathrm{d}y)^2 + \left.\frac{\partial u_z}{\partial z}\right|_{x,y,z}(\mathrm{d}z)^2 + \left(\left.\frac{\partial u_x}{\partial y}\right|_{x,y,z} + \left.\frac{\partial u_y}{\partial x}\right|_{x,y,z}\right)\mathrm{d}x\mathrm{d}y + \left(\left.\frac{\partial u_y}{\partial z}\right|_{x,y,z} + \left.\frac{\partial u_z}{\partial y}\right|_{x,y,z}\right)\mathrm{d}y\mathrm{d}z + \left(\left.\frac{\partial u_x}{\partial z}\right|_{x,y,z} + \left.\frac{\partial u_z}{\partial x}\right|_{x,y,z}\right)\mathrm{d}x\mathrm{d}z\right].$$

Letting $x_1 = x$, $x_2 = y$ and $x_3 = z$, we can can concisely write this expression as

$$(\mathrm{d}\check{s})^2 - (\mathrm{d}s)^2 \approx \sum_{i=1}^{3}\sum_{j=1}^{3}\left(\left.\frac{\partial u_{x_i}}{\partial x_j}\right|_{x_1,x_2,x_3} + \left.\frac{\partial u_{x_j}}{\partial x_i}\right|_{x_1,x_2,x_3}\right)\mathrm{d}x_i\mathrm{d}x_j,$$

The left-hand side is a scalar while $\mathrm{d}x_i$ and $\mathrm{d}x_j$ are components of a vector. The term in parentheses on the right-hand side is a component of a second-rank tensor, as shown in Exercise 1.4.

In elasticity theory, the term in parentheses is used in the definition of the strain tensor for infinitesimal displacements, namely,

$$\varepsilon_{ij} := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad i,j \in \{1,2,3\}, \tag{1.15}$$

[9]where $u_i = u_{x_i}$, $u_j = u_{x_j}$ and the partial derivatives are evaluated at $\mathbf{x} = [x_1, x_2, x_3]$.[10]

Thus, if we suppose that a continuum is deformed in such a way that a material point originally located at $\mathbf{x}$ is displaced by vector $\mathbf{u}(\mathbf{x})$, then, the strain tensor is defined by expression (1.15). Considering infinitesimal displacements in a three-dimensional continuum, the components of this tensor allow us to describe the deformation associated with any such a displacement.

[9]In this book, symbol := denotes definition. Equivalently, we could write it as $\stackrel{def}{=}$.

[10]Readers interested in formulation of the strain tensor leading to its form that is valid for curvilinear coordinates might refer to Synge, J.L., and Schild, A., (1949/1978) Tensor calculus: Dover, pp. 202 – 205.

In view of its definition, the strain tensor is symmetric, namely, $\varepsilon_{ij} = \varepsilon_{ji}$. Consequently, in a three-dimensional continuum, there are only six independent components. Also, in view of its definition, the strain tensor is dimensionless.

Note the following analogy between vector calculus and tensor calculus. The gradient operator applied to the scalar field $f(x_1, x_2, x_3)$ results in a vector field described by three components $(\partial f/\partial x_1, \partial f/\partial x_2, \partial f/\partial x_3)$. As shown in the derivation of expression (1.15), the gradient operator applied to the vector field $\mathbf{u} = [u_1, u_2, u_3]$ results in a second-rank tensor field described by nine components of the form $\partial u_i/\partial x_j$, where $i, j \in \{1, 2, 3\}$.

1.3.2 Physical meaning of strain tensor

Introductory comments

The strain tensor describes two types of deformation. Firstly, the sides of a volume element within a continuum can change in length. This can result in a change of volume without, necessarily, a change in shape. Components of the strain tensor, which we use to describe such deformations, are dimensionless quantities given by a change in length per unit length. Secondly, the sides of an element within a continuum can change orientation with respect to each other. This results in a change of shape without, necessarily, a change in volume. Components of the corresponding strain tensor are measured in radians and describe the change in angles before and after the deformation. Thus, the strain tensor describes relative linear displacement and relative angular displacement.

Relative change in length

To illustrate a length change expressed by a strain tensor, we revisit the derivation shown in Section 1.3.1 and consider the one-dimensional case.

Let $\mathbf{x} = [x_1, 0, 0]$ and $\mathbf{x}+\mathrm{d}\mathbf{x} = [x_1 + \mathrm{d}x_1, 0, 0]$ be two close points on the x_1-axis prior to deformation. During deformation, these points may be removed from the x_1-axis, however, their coordinates along this axis after the deformation are

$$\breve{x}_1 = x_1 + u_1|_{x_1,0,0}, \tag{1.16}$$

and

$$\breve{x}_1 + \mathrm{d}\breve{x}_1 = x_1 + \mathrm{d}x_1 + u_1|_{x_1+\mathrm{d}x_1,0,0}. \tag{1.17}$$

The distance between their components along the x_1-axis after the deformation is given by the difference between expressions (1.16) and (1.17), namely,

$$\mathrm{d}\breve{x}_1 = \mathrm{d}x_1 + u_1|_{x_1+\mathrm{d}x_1,0,0} - u_1|_{x_1,0,0}\,. \tag{1.18}$$

Taylor's series of the middle term on the right-hand side can be written as

$$u_1|_{x_1+\mathrm{d}x_1,0,0} = u_1|_{x_1,0,0} + \left.\frac{\partial u_1}{\partial x_1}\right|_{x_1,0,0} \mathrm{d}x_1 + \frac{1}{2}\left.\frac{\partial^2 u_1}{\partial x_1^2}\right|_{x_1,0,0} (\mathrm{d}x_1)^2 + \ldots .$$

Using the approximation consisting of the first two terms, we can write expression (1.18) as

$$\mathrm{d}\breve{x}_1 \approx \mathrm{d}x_1 + \left.\frac{\partial u_1}{\partial x_1}\right|_{x_1,0,0} \mathrm{d}x_1,$$

which can be restated as

$$\mathrm{d}\breve{x}_1 \approx \left(1 + \frac{\partial u_1}{\partial x_1}\right) \mathrm{d}x_1.$$

Hence, in view of definition (1.15), we can write the distance between the two points after deformation as

$$\mathrm{d}\breve{x}_1 \approx (1 + \varepsilon_{11})\, \mathrm{d}x_1, \tag{1.19}$$

where $\mathrm{d}x_1$ is the distance between these two points prior to deformation.

Thus, ε_{11} is a relative elongation or contraction along the x_1-axis. Similarly, $\partial u_2/\partial x_2 = \varepsilon_{22}$ and $\partial u_3/\partial x_3 = \varepsilon_{33}$ correspond to relative elongations or contractions along the x_2-axis and the x_3-axis, respectively.

To pictorially see the meaning of ε_{ii}, where $i \in \{1, 2, 3\}$, consider Figure 1.1 with axes defined in terms of the material coordinates that correspond to the configuration of the element of the continuum before deformation. The relative elongation along the X_1-axis can be written as

$$\frac{\Delta X_1 + \Delta u_1}{\Delta X_1} = 1 + \frac{\Delta u_1}{\Delta X_1}. \tag{1.20}$$

Considering infinitesimal gradients of the displacement, discussed in Section 1.2.3, and in view of Exercise 1.5, we can restate expression (1.20) as

$$1 + \frac{\partial u_1}{\partial x_1}. \tag{1.21}$$

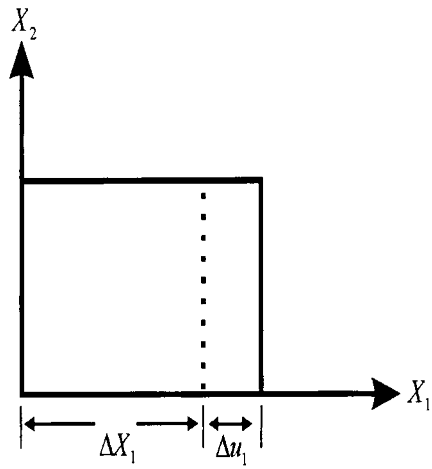


Figure 1.1: Uniaxial extension in the X_1-axis direction.

Expression (1.21) is a relative change in length due to deformation. Now, recall equation (1.19), which we can restate as

$$\frac{\mathrm{d}\breve{x}_1}{\mathrm{d}x_1} \approx 1 + \varepsilon_{11}, \tag{1.22}$$

to describe a relative change in length due to deformation. Hence, examining expressions (1.21) and (1.22), we conclude that $\varepsilon_{11} \equiv \partial u_1/\partial x_1$, as expected.

Relative change in volume

Having formulated the relative change in length, we can express a relative change in volume.

Consider a rectangular box with edge lengths Δx_1, Δx_2, and Δx_3, along the x_1-axis, the x_2-axis and the x_3-axis, respectively. Its volume is

$$V = \Delta x_1 \Delta x_2 \Delta x_3. \tag{1.23}$$

After the deformation, following expression (1.19), the edge lengths become $(1 + \varepsilon_{11})\,\Delta x_1$, $(1 + \varepsilon_{22})\,\Delta x_2$, and $(1 + \varepsilon_{33})\,\Delta x_3$, respectively, and, the volume becomes

$$\breve{V} = (1 + \varepsilon_{11})\,(1 + \varepsilon_{22})\,(1 + \varepsilon_{33})\,V.$$

Note that to express $\breve{V}$, we assume that after the deformation, the original rectangular box remains rectangular.

Assuming small deformations and, consequently, retaining only first-order strain-component terms resulting from the triple product, the volume of the deformed rectangular box can be written as

$$\breve{V} \approx (1 + \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33})\, V. \tag{1.24}$$

Thus, using expressions (1.23) and (1.24), we can state the relative change in volume as

$$\frac{\breve{V} - V}{V} \approx \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} := \varphi. \tag{1.25}$$

We refer to φ as dilatation.

Using vector calculus, we can conveniently state the relative change in volume in terms of the displacement vector, $\mathbf{u}$. In view of definition (1.15), expression (1.25) can be stated as divergence, since we can write

$$\varphi = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = \nabla \cdot \mathbf{u}. \tag{1.26}$$

The dilatation will appear in stress-strain equations (5.65), and, expressed in terms of divergence, it will appear again in wave equations for P waves, given in expression (6.12). Since dilatation is associated with a change in volume, P waves can be viewed as the propagation of local compression within the continuum.

Note that, in terms of tensor algebra, expression (1.25) is the trace of the strain tensor, $tr\,(\varepsilon_{ij})$, namely, the sum of the diagonal terms. The trace of a second-rank tensor is a scalar; hence, it is invariant under the coordinate transformations, as proven in Exercise 1.6. Thus, as expected, the description of the change in volume is independent of the choice of the coordinate system. Relative change in volume, in the context of material properties, is shown in Exercise 5.8.

Change in shape

The strain tensor also describes deformations leading to a change in shape. To gain geometrical insight, consider Figure 1.2 with axes defined in terms of the material coordinates that correspond to the configuration of the element of the continuum before deformation. A rectangular element of the continuum is deformed into a parallelogram. In other words, the original right angle is reduced to angle α. We can write this reduction as

$$\frac{\pi}{2} - \alpha = \beta_1 + \beta_2,$$

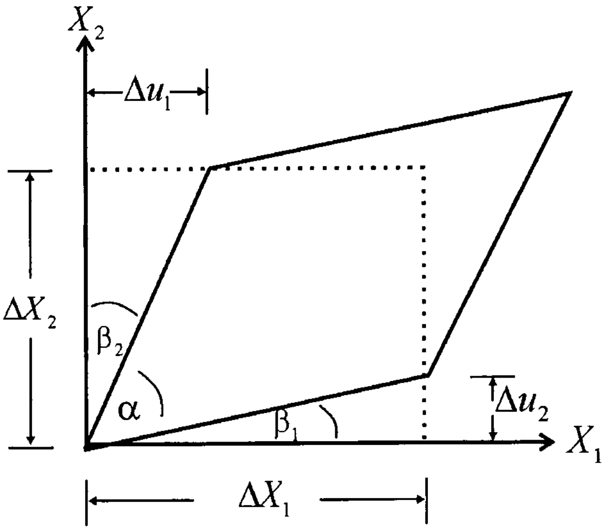


Figure 1.2: Relative change in angles.

where β_1 and β_2 are the angles measured with respect to the X_1-axis and the X_2-axis, respectively. Assuming that angles β_1 and β_2 are small and measured in radians, we can approximate them by the corresponding tangents. Hence, examining Figure 1.2, we can write

$$\beta_1 + \beta_2 \approx \frac{\Delta u_2}{\Delta X_1} + \frac{\Delta u_1}{\Delta X_2}. \tag{1.27}$$

Considering infinitesimal displacements, discussed in Section 1.2.3, and in view of Exercise 1.5, we can write equation (1.27) as

$$\beta_1 + \beta_2 \approx \frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2} = 2\varepsilon_{21} = 2\varepsilon_{12}, \tag{1.28}$$

where we assume the equivalence of X_i and x_i. In other words, a function of coordinates that is evaluated at a point corresponding to the original configuration is approximately equal to this function evaluated at a point corresponding to the final position.[11]

Examining Figure 1.2, we see that equation (1.28) implies that the original segments are deviated by small angles β_1 and β_2 that can be stated

[11]Readers interested in more details associated with the strain tensor in the context of the material and the spatial coordinates might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, pp. 120 – 135.

as $\partial u_2/\partial x_1$ and $\partial u_1/\partial x_2$, respectively. Consequently, the initial right angle between segments, coinciding with the two axes, is changed by the sum of these two angles.[12]

1.4 Rotation tensor and rotation vector

In Section 1.3.2, we defined dilatation, φ, which allows us to describe a relative change in volume using the divergence operator and the displacement vector, as shown in expression (1.26). In this section, we will associate a change in shape with the displacement vector by using the curl operator.

Let us define a tensor given by

$$\xi_{ij} := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right), \qquad i,j \in \{1,2,3\}. \tag{1.29}$$

In view of definition (1.29), $\xi_{11} = \xi_{22} = \xi_{33} = 0$, and tensor ξ_{ij} has only three independent components, namely, $\xi_{23} = -\xi_{32}$, $\xi_{13} = -\xi_{31}$ and $\xi_{12} = -\xi_{21}$. Thus, ξ_{ij} is an antisymmetric tensor. We refer to ξ_{ij} as the rotation tensor. As discussed in Section 1.3.2 and illustrated in Figure 1.2, the quantities $\partial u_i/\partial x_j$, where $i \neq j$, are tantamount to the small deviation angles. Following the properties of the curl operator, we can associate tensor (1.29) with a vector given by

$$\boldsymbol{\Psi} = \nabla \times \mathbf{u}, \tag{1.30}$$

as shown in Exercise 1.7. We refer to $\boldsymbol{\Psi}$ as the rotation vector.[13]

Rotation vector (1.30) will be used in formulating the wave equation involving S waves, as shown in expression (6.16). In other words, S waves can be viewed as the propagation of local rotation within the continuum.

Note that we can use tensor calculus to relate the components of the strain tensor, the components of the rotation tensor and the components of the gradient of the displacement vector. Using expressions (1.15) and (1.29), we can write the partial derivative of a component of displacement as

$$\frac{\partial u_i}{\partial x_j} = \varepsilon_{ij} + \xi_{ij}, \qquad i,j \in \{1,2,3\}. \tag{1.31}$$

[12]Readers interested in a geometrical interpretation of the strain-tensor components might refer to Fung, Y.C., (1977) A first course in continuum mechanics: Prentice-Hall, Inc., pp. 129 - 130.

[13]Readers interested in a relation between the rotation tensor and rotation vector might also refer to Fung, Y.C., (1977) A first course in continuum mechanics: Prentice-Hall, Inc., pp. 130 - 132.

Equation (1.31) corresponds to the fact that any second-rank tensor can be written as a sum of symmetric and antisymmetric tensors.

Closing remarks

Formulations of continuum mechanics allow us to describe deformation in a three-dimensional continuum. In subsequent chapters, these formulations will allow us to study and describe phenomena associated with wave propagation. In this study, we will use the linearized theory of elasticity. Although linearization results in a loss of subtle details, the agreement between the theory and experiments is satisfactory for our purposes.

Exercises

Exercise 1.1 [14]*Given a material description of motion,*

$$\mathbf{x}(\mathbf{X},t)=\left\{\begin{array}{l} x_1 = X_1 e^t + X_3\left(e^t-1\right) \\ x_2 = X_2 + X_3\left(e^t-e^{-t}\right) \\ x_3 = X_3 \end{array}\right. , \tag{1.32}$$

verify that the transformation between the material, $\mathbf{X}$*, and spatial,* $\mathbf{x}$*, coordinates exists, and find the spatial description of this motion.*[15]

Solution 1.1 *The transformation between the material and spatial coordinates exists if and only if the Jacobian, which is given by*

$$J := \det\begin{bmatrix} \dfrac{\partial x_1}{\partial X_1} & \dfrac{\partial x_1}{\partial X_2} & \dfrac{\partial x_1}{\partial X_3} \\ \dfrac{\partial x_2}{\partial X_1} & \dfrac{\partial x_2}{\partial X_2} & \dfrac{\partial x_2}{\partial X_3} \\ \dfrac{\partial x_3}{\partial X_1} & \dfrac{\partial x_3}{\partial X_2} & \dfrac{\partial x_3}{\partial X_3} \end{bmatrix}, \tag{1.33}$$

does not vanish. Using equations (1.32), we obtain

$$J = \det\begin{bmatrix} e^t & 0 & e^t-1 \\ 0 & 1 & e^t-e^{-t} \\ 0 & 0 & 1 \end{bmatrix} = e^t \neq 0.$$

[14] See also Section 1.2.1.

[15] In this book, $e^{(\cdot)}$ and $\exp(\cdot)$ are used as synonymous notations.

Thus, the transformation exists. Since, in this exercise, mapping $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ *is linear, we can write it using matrix notation* $\mathbf{x} = \mathbf{AX}$. *We can explicitly write,*

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} e^t & 0 & e^t - 1 \\ 0 & 1 & e^t - e^{-t} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix},$$

where $\mathbf{A}$ *is the transformation matrix. Since* $\det \mathbf{A} = J \neq 0$, *transformation matrix* $\mathbf{A}$ *has an inverse. Thus, the spatial description of motion, namely,* $\mathbf{X} = \mathbf{X}(\mathbf{x}, t)$, *is* $\mathbf{X} = \mathbf{A}^{-1}\mathbf{x}$. *In other words,*

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} e^{-t} & 0 & e^{-t}\left(1 - e^t\right) \\ 0 & 1 & e^{-t}\left(1 - e^{2t}\right) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Remark 1.1 *Note that at* $t = 0$, $\mathbf{A} = \mathbf{A}^{-1} = \mathbf{I}$; *hence,* $\mathbf{X}(0) = \mathbf{x}(0)$. *In other words, at the initial time, both material and spatial descriptions of motion coincide. At a later time, the material point that occupied position* $\mathbf{X}$ *at time* $t = 0$, *occupies position* $\mathbf{x}$.

Exercise 1.2 [16] *Consider*

$$F(X) = a \sin \frac{X}{b}, \tag{1.34}$$

where a *and* b *are constants. Let the change of variables be given by* $X = x - u(x)$. *Show that if both* a *and* $u(x)$ *are infinitesimal while* b *is finite, we obtain*

$$F(X) = F(x).$$

Solution 1.2 *Considering the given change of variables, we can write expression (1.34) as*

$$\begin{aligned} F(X(x)) &= a \sin \frac{x - u(x)}{b} \\ &= a\left(\sin \frac{x}{b} \cos \frac{u(x)}{b} - \sin \frac{u(x)}{b} \cos \frac{x}{b}\right). \end{aligned}$$

Since $u(x)$ *is an infinitesimal quantity and* b *is finite*

$$\lim_{u/b \to 0} \cos \frac{u(x)}{b} = 1.$$

[16] See also Section 1.2.3.

Also, assuming that $u(x)/b$ is expressed in radians,

$$\sin\frac{u(x)}{b} \approx \frac{u(x)}{b}.$$

Thus, we can write

$$\begin{aligned} F(X(x)) &\approx a\left(\sin\frac{x}{b} - \frac{u(x)}{b}\cos\frac{x}{b}\right) \\ &= a\sin\frac{x}{b} - a\frac{u(x)}{b}\cos\frac{x}{b}. \end{aligned}$$

Again, since both a and $u(x)$ are infinitesimal, we ignore their product to obtain

$$F(X(x)) \approx a\sin\frac{x}{b} = F(x),$$

as required.

Remark 1.2 *The result of Exercise 1.2, as well as the equivalence of the material and spatial coordinates for the infinitesimal displacements, is quite intuitive. In other words, considering the change of variables given by $X = x - u(x)$, we get $X \approx x$, for infinitesimal values of $u(x)$.*

Exercise 1.3 *A bar of length l would have an elongation u_1 due to strain ε_1, that is, $u_1 = \varepsilon_1 l$. The same bar would have another elongation u_2 due to strain ε_2, that is, $u_2 = \varepsilon_2 l$. Show that considering only linear terms, under the assumption of small strains, the total elongation due to both strains is equal to the sum of both elongations.*

Solution 1.3 *Assume that ε_1 is applied first. This results in the elongation,*

$$u_1 = \varepsilon_1 l.$$

Hence, the new length of the bar is

$$l + u_1 = l + \varepsilon_1 l = l(1 + \varepsilon_1).$$

Subsequently, applying strain, ε_2, we obtain the final elongation,

$$u_f = u_1 + \varepsilon_2 l(1 + \varepsilon_1) = u_1 + \varepsilon_2 l + \varepsilon_1\varepsilon_2 l = u_1 + u_2 + \varepsilon_1\varepsilon_2 l.$$

Assuming that the value of the product, $\varepsilon_1\varepsilon_2 l$, is small compared with the values of both $\varepsilon_1 l$ and $\varepsilon_2 l$ — in other words, both ε_1 and ε_2 are much smaller than unity — we obtain

$$u_f \approx u_1 + u_2.$$

Remark 1.3 *The same result is obtained if the order is reversed, or if ε_1 and ε_2 are applied simultaneously. This is the illustration of the fact that the principle of superposition is applicable to all linear systems — a commonly used property in mathematical physics.*

Exercise 1.4 [17] *Using definition (1.15) and considering orthonormal coordinate systems, show that strain, ε_{ij}, which is given in terms of first partial derivatives of a vector, is a second-rank tensor.*

Notation 1.1 *The repeated-index summation notation is used in this solution. Any term in which an index appears twice stands for the sum of all such terms as the index assumes all the values between* 1 *and* 3.

Solution 1.4 *Following definition (1.15), consider $\partial \hat{u}_i / \partial \hat{x}_j$, where $\hat{u}_i$ are the components of the displacement vector,* **u**, *in the transformed coordinates $\hat{x}_j$. The transformation rule of the coordinate points is given by*

$$\hat{x}_j = a_{jl} x_l, \qquad j \in \{1, 2, 3\}, \tag{1.35}$$

where the entries of matrix **a** *are the projections between the transformed and original axes. Matrix* **a** *is an orthogonal matrix; in other words, its inverse is equal to its transpose. Hence,*

$$x_j = a_{lj} \hat{x}_l, \qquad j \in \{1, 2, 3\}.$$

Consequently, we obtain

$$\frac{\partial x_j}{\partial \hat{x}_l} = a_{lj}. \tag{1.36}$$

Since **u** *is a vector, its components transform according to the rule*

$$\hat{u}_i = a_{ik} u_k, \qquad i \in \{1, 2, 3\}.$$

Thus, we can write

$$\frac{\partial \hat{u}_i}{\partial \hat{x}_l} = a_{ik} \frac{\partial u_k}{\partial \hat{x}_l}, \qquad i, l \in \{1, 2, 3\},$$

which can be restated as

$$\frac{\partial \hat{u}_i}{\partial \hat{x}_l} = a_{ik} \frac{\partial u_k}{\partial x_j} \frac{\partial x_j}{\partial \hat{x}_l}, \qquad i, l \in \{1, 2, 3\}.$$

[17] See also Sections 1.3.1, 5.1.1 and 5.2.2.

Hence, in view of equation (1.36), we can write

$$\frac{\partial \hat{u}_i}{\partial \hat{x}_l} = a_{ik} a_{lj} \frac{\partial u_k}{\partial x_j}, \qquad i, l \in \{1, 2, 3\}, \tag{1.37}$$

which is a transformation rule for the second-rank tensors. Consequently, since the sum of second-rank tensors is a second-rank tensor, an entity given by $\varepsilon_{ij} := (\partial u_i / \partial x_j + \partial u_j / \partial x_i)/2$ *is a second-rank tensor.*

Exercise 1.5 [18]*Considering the one-dimensional case and assuming infinitesimal displacement gradients, in view of expressions (1.3) and (1.4), show that*

$$\frac{\partial u}{\partial x} \approx \frac{\partial U}{\partial X}. \tag{1.38}$$

Solution 1.5 *Consider the one-dimensional case of expressions (1.3) and (1.4), namely*

$$\left\{ \begin{array}{l} U(X, t) = x(X, t) - X \\ \\ u(x, t) = x - X(x, t) \end{array} \right. .$$

Taking partial derivatives with respect to the first arguments, we obtain

$$\left\{ \begin{array}{l} \dfrac{\partial U}{\partial X} = \dfrac{\partial x}{\partial X} - 1 \\ \\ \dfrac{\partial u}{\partial x} = 1 - \dfrac{\partial X}{\partial x} \end{array} \right. . \tag{1.39}$$

Since $x(X, t)$ *and* $X(x, t)$ *are inverses of one another, we use the properties of the derivative of an inverse to obtain*

$$\frac{\partial x}{\partial X} = \frac{1}{\dfrac{\partial X}{\partial x}}.$$

Hence, we can write expression (1.39) as

$$\left\{ \begin{array}{l} \dfrac{\partial U}{\partial X} = \dfrac{1}{\dfrac{\partial X}{\partial x}} - 1 \\ \\ \dfrac{\partial u}{\partial x} = 1 - \dfrac{\partial X}{\partial x} \end{array} \right. .$$

[18]See also Section 1.3.2.

Solving both equations for $\partial X/\partial x$, we obtain

$$\begin{cases} \dfrac{\partial X}{\partial x} = \dfrac{1}{\dfrac{\partial U}{\partial X}+1} \\ \\ \dfrac{\partial X}{\partial x} = 1 - \dfrac{\partial u}{\partial x} \end{cases}.$$

Equating the right-hand sides and solving for $\partial u/\partial x$, we get

$$\frac{\partial u}{\partial x} = \frac{\dfrac{\partial U}{\partial X}}{\dfrac{\partial U}{\partial X}+1}. \tag{1.40}$$

Examining equation (1.40), we notice that for the infinitesimal displacement gradients, namely, $\partial U/\partial X << 1$, we can write $\partial u/\partial x \approx \partial U/\partial X$, which is expression (1.38), as required.

Exercise 1.6 [19]*Prove the following theorem.*

Theorem 1.1 *The sum of diagonal elements of a second-rank tensor is a scalar. Hence, it is invariant under transformations of the coordinate system.*

Solution 1.6 .

Proof. By definition, the components of the second-rank tensor ε_{lm} transform to the components $\hat{\varepsilon}_{ik}$, which are expressed in another coordinate system, according to the rule

$$\hat{\varepsilon}_{ik} = \sum_{l=1}^{3}\sum_{m=1}^{3} a_{il}a_{km}\varepsilon_{lm}, \qquad i,k \in \{1,2,3\},$$

where $\mathbf{a}$ is an orthogonal transformation matrix. Setting $k = i$, we obtain the sum of the components along the main diagonal, namely,

$$\sum_{i=1}^{3}\hat{\varepsilon}_{ii} = \sum_{i=1}^{3}\sum_{l=1}^{3}\sum_{m=1}^{3} a_{il}a_{im}\varepsilon_{lm}.$$

[19] See also Section 1.3.2.

Hence, by orthogonality of $\mathbf{a}$, we have

$$\sum_{i=1}^{3} a_{il}a_{im} = \delta_{lm}, \qquad l,m \in \{1,2,3\}.$$

Thus, we can write

$$\sum_{i=1}^{3} \hat{\varepsilon}_{ii} = \sum_{l=1}^{3}\sum_{m=1}^{3} \delta_{lm}\varepsilon_{lm} = \sum_{m=1}^{3} \varepsilon_{mm}.$$

Since both i and m are the summation indices, we are allowed to write

$$\sum_{j=1}^{3} \hat{\varepsilon}_{jj} = \sum_{j=1}^{3} \varepsilon_{jj}.$$

This means that the sum of the diagonal elements of a second-rank tensor is a scalar. This implies that the value of the sum of the diagonal elements is invariant under transformations of the coordinate system. ■

Remark 1.4 *Following Exercise 1.4, we can see that dilatation, φ, defined by expression (1.26) is the sum of diagonal elements of the second-rank tensor, namely, the trace of the strain tensor, $\varepsilon_{ij} := (\partial u_i/\partial x_j + \partial u_j/\partial x_i)/2$. Consequently, as shown in Exercise 1.6, we can prove that dilatation is a scalar quantity. This is expected because of the physical meaning of dilatation. In other words, the change of volume must be independent of the coordinate system.*

Exercise 1.7 [20]*In view of the properties of vector operators, show that the components of the second-rank tensor, given by expression (1.29), namely,*

$$\xi_{ij} := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right), \qquad i,j \in \{1,2,3\},$$

are associated with rotation vector (1.30).

Solution 1.7 *Consider the displacement vector $\mathbf{u} = [u_1, u_2, u_3]$. We can write its curl as*

$$\nabla \times \mathbf{u} = \left(\frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}\right)\mathbf{e}_1 + \left(\frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}\right)\mathbf{e}_2 + \left(\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}\right)\mathbf{e}_3,$$

[20]See also Section 1.4.

where $\mathbf{e}_i$ *denotes a unit vector along the* x_i*-axis. Following expression (1.29), we can rewrite the curl as*

$$\nabla \times \mathbf{u} = [2\xi_{32}, 2\xi_{13}, 2\xi_{21}] .$$

Thus, ξ_{ij} *can be viewed as the components of the vector that results from the rotation of* $\mathbf{u}/2$*. Denoting* $\boldsymbol{\Psi} = [2\xi_{32}, 2\xi_{13}, 2\xi_{21}]$*, we obtain definition (1.30).*

Remark 1.5 *The association between the components of the second-rank tensor* ξ_{ij} *and the components of vector* $\boldsymbol{\Psi}$ *is due to the antisymmetry of this tensor that results in only three independent components.*

Chapter 2

Forces and balance principles

> It is as necessary to science as to pure mathematics that the fundamental principles should be clearly stated and that the conclusions shall follow from them. But in science it is also necessary that the principles taken as fundamental should be as closely related to observation as possible.
>
> *Harold Jeffreys and Bertha Jeffreys (1946) Methods of mathematical physics*

Preliminary remarks

In the context of continuum mechanics, seismic waves are deformations that propagate in a continuum. These deformations are associated with forces. In order to describe the propagation of deformations, we now seek to formulate the equations that relate these deformations to forces acting within the continuum.

We begin this chapter with the study of the conservation of mass, which is a fundamental balance principle of continuum mechanics and which is associated with the motion of mass within the continuum. Using the conservation of mass, we derive the equation of continuity. Then we formulate the balance of linear momentum. Subsequently, in order to take into account the forces acting within the continuum, we formulate the stress tensor. Using the balance of linear momentum and the concept of stress, we derive Cauchy's equations of motion. To obtain all fundamental equations that relate the unknowns that appear in the equation of continuity and in Cauchy's equations of motion, we also invoke the balance of angular momentum. These three balance principles lead to a system of equations that is associated with the propagation of deformations in an elastic continuum.

2.1 Conservation of mass

Introductory comments

A fundamental principle in which our description of continuum mechanics must be rooted is the conservation of mass. We use this principle to derive an equation that relates mass density, ρ, and displacement vector, $\mathbf{u}$.

Note that, in general, conservation principles are special cases of the corresponding balance principles. Herein, discussing the balance of mass, we wish to emphasize that we do not consider production or destruction of mass and, hence, the total amount of mass is conserved. Discussing the balance of linear momentum and the balance of angular momentum in Sections 2.4 and 2.7, respectively, we wish to emphasize that — for a given portion of continuum — the total amount of these momenta changes and these changes are balanced by forces acting within the continuum.

2.1.1 Integral equation

The amount of mass, m, occupying a fixed volume, V, at an instant of time is given by

$$m(t) = \iiint\limits_V \rho(\mathbf{x}, t)\, \mathrm{d}V, \tag{2.1}$$

where ρ denotes mass density.

The rate of change of mass contained in this volume is given by the differentiation of equation (2.1) with respect to time, namely,

$$\frac{\mathrm{d}}{\mathrm{d}t} m(t) = \frac{\mathrm{d}}{\mathrm{d}t} \iiint\limits_V \rho(\mathbf{x}, t)\, \mathrm{d}V. \tag{2.2}$$

Furthermore, for an arbitrary fixed volume V, we can rewrite expression (2.2) as

$$\frac{\mathrm{d}}{\mathrm{d}t} m(t) = \iiint\limits_V \frac{\partial \rho(\mathbf{x}, t)}{\partial t}\, \mathrm{d}V. \tag{2.3}$$

We can also express $\mathrm{d}m/\mathrm{d}t$ in a different way. Since, in classical physics, there is no production or destruction of mass, the rate of change of mass contained in a fixed volume is only a function of the mass flowing through this volume. In other words, the rate of change of mass contained in volume V is equal to the amount of mass that passes through the surface, S, bounding this volume. This can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}m(t) = -\iint_S \rho(\mathbf{x},t)\,\mathbf{v}\cdot\mathbf{n}\,\mathrm{d}S, \tag{2.4}$$

where $\mathbf{v}$ represents the velocity of a portion of mass that passes through this surface, and where $\mathbf{n}$ denotes an outward normal vector to this surface. Herein, we assume the element $\mathrm{d}S$ to be sufficiently small that it might be considered as a plane and to have the same mass flow across all its points.

Expressions (2.3) and (2.4) describe the same quantity. To equate them, we express the right-hand side of equation (2.4) as a volume integral. Following the divergence theorem — where the surface integral of vector $\rho(\mathbf{x},t)\,\mathbf{v}(\mathbf{x})$ over a closed surface equals the volume integral of the divergence of that vector integrated over the volume enclosed by this surface — we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}m(t) = -\iiint_V \nabla\cdot(\rho\mathbf{v})\,\mathrm{d}V. \tag{2.5}$$

Now, equating expressions (2.3) and (2.5), we obtain

$$\iiint_V \frac{\partial\rho}{\partial t}\,\mathrm{d}V = -\iiint_V \nabla\cdot(\rho\mathbf{v})\,\mathrm{d}V,$$

where the negative sign results from the fact that the vector normal to the surface points away from the volume V. Combining the two volume integrals, we can write

$$\iiint_V \left[\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v})\right]\mathrm{d}V = 0, \tag{2.6}$$

which states the conservation of mass for a fixed volume V.

Note that in this derivation of equation (2.6), the change in the amount of mass in a volume at any instant is balanced by the mass flowing through the surface that encloses this volume. Consequently, considering a given volume, one could refer to equation (2.6) as a balance-of-mass equation rather than a conservation-of-mass equation. However, as stated above, we choose to use only the latter term. Our choice is also justified by the fact that discussing the balance of linear momentum and the balance of angular momentum in Sections 2.4 and 2.7, respectively, we consider a moving volume that consistently contains the same portion of the continuum, as discussed in Section 2.2; in such a case, there is no mass flowing through the surface that encloses this moving volume.

2.1.2 Equation of continuity

The equation of continuity is a differential equation expressing the conservation of mass within the continuum. To derive the equation of continuity, consider integral equation (2.6). For this equation to be true for an arbitrary fixed volume, the integrand must be identically zero. Thus, we require

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{2.7}$$

Note that if there were a point where the integrand were nonzero, we could consider a sufficiently small volume around that point. This would result in a nonzero value of the integral, as illustrated in Exercise 2.1.

Since $\mathbf{v} = \partial \mathbf{u} / \partial t$, where $\mathbf{u}$ denotes the displacement vector, we can write equation (2.7) as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left(\rho \frac{\partial \mathbf{u}}{\partial t} \right) = 0. \tag{2.8}$$

This is the equation of continuity. The equation of continuity equates the rate of change of the amount of material inside a closed surface to the net rate at which the material flows through this surface.

2.2 Time derivative of volume integral

To derive the remaining two balance principles, namely, the balance of linear momentum and the balance of angular momentum, we use the concept of the time derivative of a moving-volume integral, which is associated with the conservation of mass. For this purpose, let us consider a moving volume that consistently contains the same portion of the continuum. In other words, there is no mass transport through the surface encompassing this volume. In such a case, the portion of the continuum possessing a given velocity and acceleration is identifiable. Hence, such a description lends itself to a convenient extension of particle mechanics and, therefore, allows us to use Newton's laws of motion.

To consider the temporal variation of a physical quantity enclosed in a moving volume, we must consider the time derivative of

$$I(\mathbf{x},t) = \iiint\limits_{V(t)} \rho(\mathbf{x},t)\,\mathcal{A}(\mathbf{x},t)\,\mathrm{d}V,$$

where ρ is mass density, $\mathcal{A}$ is a scalar, vector or tensor, while $V(t)$ is a volume that varies with time but always contains the same portion of the continuum. This derivative can be formulated in the following way.

Consider

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)} \rho\mathcal{A}\mathrm{d}V = \iiint\limits_{V(t)} \frac{\partial}{\partial t}(\rho\mathcal{A})\,\mathrm{d}V + \iint\limits_{S(t)} \rho\mathcal{A}\sum_{j=1}^{3} v_j n_j \mathrm{d}S,$$

where $\mathcal{A}$ describes a physical quantity of interest. The left-hand side is the rate of change of the total amount of $\mathcal{A}$ within the moving volume, the first integral on the right-hand side is the change of $\mathcal{A}$ associated with this volume while the second integral on the right-hand side is the change associated with the surface enclosing this volume. Using the divergence theorem, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)} \rho\mathcal{A}\mathrm{d}V = \iiint\limits_{V(t)} \frac{\partial}{\partial t}(\rho\mathcal{A})\,\mathrm{d}V + \iiint\limits_{V(t)} \sum_{j=1}^{3} \frac{\partial(\rho\mathcal{A}v_j)}{\partial x_j}\mathrm{d}V.$$

Differentiating and rearranging, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)} \rho\mathcal{A}\mathrm{d}V = \iiint\limits_{V(t)} \left\{\rho\left(\frac{\partial\mathcal{A}}{\partial t} + \sum_{j=1}^{3} v_j\frac{\partial\mathcal{A}}{\partial x_j}\right) + \mathcal{A}\left[\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v})\right]\right\}\mathrm{d}V.$$

We note that the term in brackets vanishes due to equation of continuity (2.7). Thus, we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)} \rho\mathcal{A}\mathrm{d}V = \iiint\limits_{V(t)} \rho\left(\frac{\partial\mathcal{A}}{\partial t} + \sum_{j=1}^{3} v_j\frac{\partial\mathcal{A}}{\partial x_j}\right)\mathrm{d}V,$$

which can be restated as

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)} \rho\mathcal{A}\mathrm{d}V = \iiint\limits_{V(t)} \rho\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\mathcal{A}\mathrm{d}V. \qquad (2.9)$$

In view of expression (1.1), we note that the operator in parentheses in expression (2.9) is the material time-derivative operator. Hence, we obtain the desired result, namely,

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)} \rho(\mathbf{x},t)\,\mathcal{A}(\mathbf{x},t)\,\mathrm{d}V = \iiint\limits_{V(t)} \rho(\mathbf{x},t)\,\frac{\mathrm{d}\mathcal{A}(\mathbf{x},t)}{\mathrm{d}t}\mathrm{d}V. \tag{2.10}$$

Note that, in the case of moving volume, the time derivative and the volume integral do not commute, while, as shown in equations (2.2) and (2.3), they do commute for a fixed volume.

2.3 Stress

2.3.1 Stress as description of surface forces

We wish to analyze the internal forces acting among the adjacent material elements within the continuum. For this purpose, we introduce the concept of stress.

The presence of stress, as described below, sets continuum mechanics apart from particle mechanics. Stress, as a mathematical entity, was introduced by Cauchy in 1827 to express the interaction of a material with the surrounding material in terms of surface forces.[1]

When a material is subjected to loads, internal forces are induced within it. Deformation of this material is a function of the distribution of these forces. In a continuum, stress is associated with internal surface forces that an element of the continuum exerts on another element of the continuum across an imaginary surface that separates them. Stress is a system of surface forces producing strain within a continuum. Owing to the mutual dependence of stress and strain, strains cannot be produced without inducing stresses, and stresses cannot be induced without producing, or tending to produce, strains. This interrelation between stress and strain is an intrinsic property of the elasticity theory.

2.3.2 Traction

As a result of forces being transmitted within the continuum, the portion of the continuum within an arbitrary volume enclosed by an imaginary surface interacts with the portion of the continuum on the other side of this surface.

[1]Interested readers might refer to Cauchy, A. L., (1827) De la pression ou tension dans un corps solide: Ex. de Math, **2**, pp. 42 – 56.

Let $\Delta\mathbf{F}$ be the force exerted on the surface element ΔS by the continuum on either side of this surface. The average force per unit area can be written in terms of the ratio

$$\bar{\mathbf{T}} = \frac{\Delta\mathbf{F}}{\Delta S}. \tag{2.11}$$

Cauchy's stress principle — the fundamental principle of continuum mechanics — asserts that as $\Delta S \to 0$, ratio (2.11) tends to a finite limit.[2] The resulting vector is called the traction and is given by

$$\mathbf{T}^{(\mathbf{n})} = \lim_{\Delta S \to 0} \frac{\Delta\mathbf{F}}{\Delta S} = \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}S}, \tag{2.12}$$

where the superscript $\mathbf{n}$ refers to the surface element, ΔS, upon which the traction is acting and which is defined by its unit normal, $\mathbf{n}$. Thus, traction is a vector that describes the contact force with which the elements at each side of an internal surface within the continuum act upon each other.

Note that since the value of traction is finite while the element of the surface area becomes infinitesimal, we can describe a distribution of forces at every point within the continuum. Also note that the traction is explicitly dependent on the particular choice of the surface element, as indicated by the unit vector, $\mathbf{n}$, and, consequently, we can describe a distribution of forces associated with any given direction.

2.4 Balance of linear momentum

In general, the forces acting within a continuum are classified as either surface forces or body forces according to their mode of application. Surface forces are transmitted by direct mechanical contacts across imaginary surfaces separating given portions of the continuum. Body forces, such as gravitational force, are associated with action at a distance.

Consider a portion of a continuum contained in volume V and subjected to time-varying and space-varying forces. The surface forces are given as the traction vector shown in expression (2.12), namely,

$$\mathbf{T} = \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}S},$$

and the body forces are given by

$$\mathbf{f} = \mathbf{f}(\mathbf{x}, t).$$

[2] Interested readers might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, p. 70.

Consequently, the total force is

$$\mathbf{F}_T = \iint\limits_S \mathbf{T}\,\mathrm{d}S + \iiint\limits_V \mathbf{f}\,\mathrm{d}V, \tag{2.13}$$

where S is the surface containing volume V.

Note that $\mathbf{T}$ and $\mathbf{f}$ have units of force per area and force per volume, respectively.

To study the effect of this force, we choose to consider a moving volume that consistently contains the same portion of the continuum. Hence, invoking Newton's second law of motion, we can write

$$\frac{\mathrm{d}}{\mathrm{d}t} \iiint\limits_{V(t)} \rho \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\,\mathrm{d}V = \iint\limits_{S(t)} \mathbf{T}\,\mathrm{d}S + \iiint\limits_{V(t)} \mathbf{f}\,\mathrm{d}V, \tag{2.14}$$

which is an integral equation that states balance of linear momentum and where the displacement,

$$\mathbf{u} = \left[u_1(\mathbf{x},t), u_2(\mathbf{x},t), u_3(\mathbf{x},t)\right], \tag{2.15}$$

is a function of both space and time. Invoking expression (2.10) and letting $\mathcal{A} = \mathrm{d}\mathbf{u}/\mathrm{d}t$, we can restate the balance of linear momentum as

$$\iiint\limits_{V(t)} \rho \frac{\mathrm{d}^2\mathbf{u}}{\mathrm{d}t^2}\mathrm{d}V = \iint\limits_{S(t)} \mathbf{T}\,\mathrm{d}S + \iiint\limits_{V(t)} \mathbf{f}\,\mathrm{d}V. \tag{2.16}$$

Note that $\mathrm{d}^2/\mathrm{d}t^2$ refers to the material time-derivative operator, which is shown in expression (1.1).

Integral equation (2.16) states that the rate of change of linear momentum of an element within the continuum is equal to the sum of the forces acting upon this element. This statement is analogous to Newton's second law of motion in particle mechanics.

In Section 2.5.2, we use the balance of linear momentum to formulate the stress tensor. In Section 2.6, following the formulation of the stress tensor, we use equation (2.16) to derive Cauchy's equations of motion.

2.5 Stress tensor

2.5.1 Traction on coordinate planes

We wish to describe the state of stress at a given point in a continuum. At an arbitrary point within a continuum, Cauchy's stress principle associates a traction and a unit normal of a surface element on which this vector is acting.

Consider a fixed coordinate system with the orthonormal vectors given by $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. The traction acting on the ith coordinate plane is represented by a vector, which can be written as

$$\mathbf{T}^{(\mathbf{e}_i)} = T_1^{(\mathbf{e}_i)}\mathbf{e}_1 + T_2^{(\mathbf{e}_i)}\mathbf{e}_2 + T_3^{(\mathbf{e}_i)}\mathbf{e}_3,$$

where $T_j^{(\mathbf{e}_i)}$ are the components of this vector along the x_j-axis. The three tractions associated with the three mutually orthogonal planes can be explicitly written as three vectors given by

$$\begin{bmatrix} \mathbf{T}^{(\mathbf{e}_1)} \\ \mathbf{T}^{(\mathbf{e}_2)} \\ \mathbf{T}^{(\mathbf{e}_3)} \end{bmatrix} = \begin{bmatrix} T_1^{(\mathbf{e}_1)} & T_2^{(\mathbf{e}_1)} & T_3^{(\mathbf{e}_1)} \\ T_1^{(\mathbf{e}_2)} & T_2^{(\mathbf{e}_2)} & T_3^{(\mathbf{e}_2)} \\ T_1^{(\mathbf{e}_3)} & T_2^{(\mathbf{e}_3)} & T_3^{(\mathbf{e}_3)} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix}. \tag{2.17}$$

Considering the traction components shown in the 3×3 matrix, we see that the subscript refers to the component of a given traction, while the superscript identifies the plane on which this traction is acting. For instance, $T_2^{(\mathbf{e}_1)}$ is the x_2-component of a traction acting on the plane normal to the x_1-axis.

For convenience, we write the square matrix in equations (2.17) as

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}. \tag{2.18}$$

By examining expressions (2.17) and (2.18), we immediately see that σ_{ij} represents the jth component of the surface force acting on the surface whose normal is parallel to the x_i-axis. This index convention, which allows us to describe the direction of the force and the orientation of the surface on which it is acting, is also illustrated in Figure 2.1.[3]

We also wish to distinguish between tension and compression for the traction components normal to a given face, as well as denote the direction

[3]This index convention is consistent with Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, p. 80, and with Aki, K., and Richards, P.G., (2002) Quantitative seismology (2*nd* edition): University Science Books, pp. 17 – 18.

One can also use the opposite convention; for instance, Kolsky, H., (1953/1963) Stress waves in solids: Dover, p. 5, and Sheriff, R.E., and Geldart, L.P., (1982) Exploration Seismology: Cambridge University Press, Vol. I, p. 33.

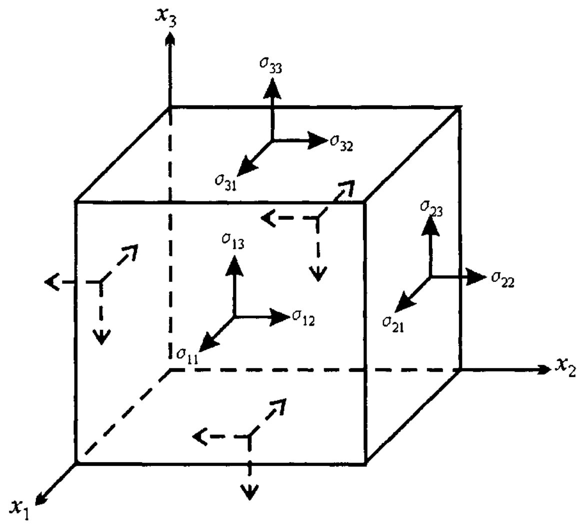


Figure 2.1: The index convention for the σ_{ij} components. σ_{ij} represents the jth component of the surface force acting on the surface whose normal is parallel to the x_i-axis. All components shown herein are positive.

of the traction components tangential to a given face. For this purpose, we adopt the following sign convention. On a surface whose outward normal points in the positive direction of the corresponding coordinate axis, all traction components that act in the positive direction of a given axis are positive. On a surface whose outward normal points in the negative direction of the corresponding coordinate axis, all traction components that act in the negative direction of a given axis are positive. This convention applies to both the normal and the tangential components. Examining Figure 2.1, we see that all the traction components on each of the six faces illustrated therein are positive. In the context of the normal components, our sign convention implies that tension is positive while compression is negative.[4]

Note that, if we wished, we could reverse our sign convention without affecting Newton's third law. In other words,

$$\mathbf{T}^{(-\mathbf{n})} = -\mathbf{T}^{(\mathbf{n})} \tag{2.19}$$

[4]This sign convention is consistent with Aki, K. and Richards, P.G., (2002) Quantitative seismology (2*nd* edition): University Science Books, p. 15.

is always true.

As formulated herein, the entries of matrix (2.18) determine the stress state within a continuum at a given point with respect to the coordinate planes. As shown in Section 2.5.2, these entries can also be used to describe the stress state with respect to an arbitrary plane within the continuum.

2.5.2 Traction on arbitrary planes

To study forces within the continuum, we wish to describe them with respect to a plane of arbitrary orientation. For this purpose, consider an element of a continuum in the form of a tetrahedron. Let the tetrahedron be spanned by four points $O\,(0,0,0)$, $A\,(a,0,0)$, $B\,(0,b,0)$ and $C\,(0,0,c)$, as shown in Figure 2.2. Thus, the four faces of the tetrahedron consist of the oblique face, namely, ABC, and of three orthogonal faces, namely, OAB, OBC and OAC.

We seek to determine the force $\Delta\mathbf{F}$ acting on the oblique face whose area is ΔS and whose unit normal is $\mathbf{n}$.

The key statement of this derivation relies on the balance of linear momentum, discussed in Section 2.4, and the fact that the tetrahedron is subjected to both surface and body forces. In view of equation (2.16), for a finite-size tetrahedron, we can write

$$\Delta\mathbf{F} + \Delta\mathbf{F}^{(\mathbf{e}_1)} + \Delta\mathbf{F}^{(\mathbf{e}_2)} + \Delta\mathbf{F}^{(\mathbf{e}_3)} + \bar{\mathbf{f}}\Delta V = \bar{\rho}\Delta V\frac{\mathrm{d}\bar{\mathbf{v}}}{\mathrm{d}t}, \tag{2.20}$$

where $\Delta\mathbf{F}$ is the surface force acting on the oblique face, $\Delta\mathbf{F}^{(\mathbf{e}_i)}$ is the surface force acting on the orthogonal face normal to the x_i-axis, and $\bar{\mathbf{f}}$ refers to the body force acting on the tetrahedron with volume ΔV and mass density $\bar{\rho}$. Thus, the left-hand side of equation (2.20) gives the sum of forces, while the right-hand side gives the rate of change of linear momentum with $\bar{\mathbf{v}}$ denoting velocity. The bars above a given symbol denote the average value of the corresponding quantity for this finite-size tetrahedron.

In view of expression (2.11), we can write

$$\Delta\mathbf{F} = \bar{\mathbf{T}}^{(\mathbf{n})}\Delta S. \tag{2.21}$$

Using expressions (2.19) and (2.21), we can rewrite equation (2.20) as

$$\bar{\mathbf{T}}^{(\mathbf{n})}\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_1)}\Delta S_1 - \bar{\mathbf{T}}^{(\mathbf{e}_2)}\Delta S_2 - \bar{\mathbf{T}}^{(\mathbf{e}_3)}\Delta S_3 + \bar{\mathbf{f}}\Delta V = \bar{\rho}\Delta V\frac{\mathrm{d}\bar{\mathbf{v}}}{\mathrm{d}t}, \tag{2.22}$$

where ΔS is the area of the oblique face and ΔS_i is the area of an orthogonal face normal to the x_i-axis. In equation (2.22), $\bar{\mathbf{T}}^{(\cdot)}$ is a resultant traction that corresponds to a given face.

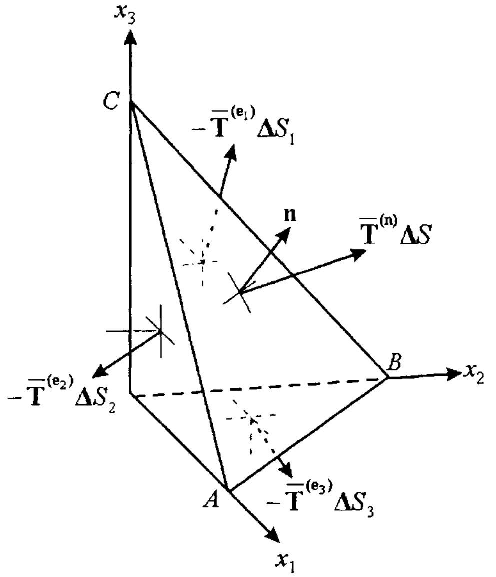


Figure 2.2: Tetrahedron used in the formulation of the stress tensor. This construction is also called Cauchy's tetrahedron.

Note that the orthogonal faces have unit outward normals parallel and opposite in sign to the unit vectors of the coordinate axes, $\mathbf{e}_i$. Hence, in view of Newton's third law, we introduced the negative signs in the summation.

The surface forces, which are used in equation (2.22), are illustrated in Figure 2.2.

To study equation (2.22), we wish to geometrically relate the surface areas of the tetrahedron, ΔS and ΔS_i, where $i \in \{1, 2, 3\}$, and its volume, V.

The areas of the orthogonal faces are

$$\Delta S_i = n_i \Delta S, \qquad i \in \{1, 2, 3\}, \tag{2.23}$$

where n_i is the component of the unit vector, $\mathbf{n}$, that is normal to the oblique face. Using expression (2.23), we can rewrite equation (2.22) as

$$\bar{\mathbf{T}}^{(\mathbf{n})}\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_1)}n_1\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_2)}n_2\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_3)}n_3\Delta S + \bar{\mathbf{f}}\Delta V = \bar{\rho}\Delta V\frac{\mathrm{d}\bar{\mathbf{v}}}{\mathrm{d}t}. \tag{2.24}$$

Now, we wish to relate the volume, ΔV, to the area of the oblique face, ΔS. Considering the oblique face as the base of the tetrahedron, we can state its volume as

$$\Delta V = \frac{h}{3}\Delta S, \tag{2.25}$$

where h is the height of the tetrahedron. Hence, using expression (2.25), we can rewrite equation (2.24) as

$$\bar{\mathbf{T}}^{(\mathbf{n})}\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_1)}n_1\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_2)}n_2\Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_3)}n_3\Delta S + \frac{\bar{\mathbf{f}}h}{3}\Delta S = \frac{\bar{\rho}h}{3}\Delta S\frac{\mathrm{d}\bar{\mathbf{v}}}{\mathrm{d}t}. \tag{2.26}$$

Dividing both sides of equation (2.26) by ΔS, we obtain

$$\bar{\mathbf{T}}^{(\mathbf{n})} - \bar{\mathbf{T}}^{(\mathbf{e}_1)}n_1 - \bar{\mathbf{T}}^{(\mathbf{e}_2)}n_2 - \bar{\mathbf{T}}^{(\mathbf{e}_3)}n_3 + \frac{\bar{\mathbf{f}}h}{3} = \frac{\bar{\rho}h}{3}\frac{\mathrm{d}\bar{\mathbf{v}}}{\mathrm{d}t}. \tag{2.27}$$

To describe the state of stress at a point within the continuum, we let $h \to 0$ in such a way that the areas of all faces simultaneously approach zero, the orientation of the height, h, does not change, and the origin of the coordinate system does not move. In other words, the finite-size tetrahedron reduces to an infinitesimal tetrahedron at point $O\,(0,0,0)$. Thus, we obtain

$$\mathbf{T}^{(\mathbf{n})} = \mathbf{T}^{(\mathbf{e}_1)}n_1 + \mathbf{T}^{(\mathbf{e}_2)}n_2 + \mathbf{T}^{(\mathbf{e}_3)}n_3. \tag{2.28}$$

Note that in equation (2.28), the tractions no longer correspond to the average values but to the local values at point $O\,(0,0,0)$. This also implies that equation (2.28) is valid for any coordinate system.

Equation (2.28) can be viewed as an equilibrium equation of an infinitesimal element within the continuum. Note, however, that the derivation of this equation stems from the balance of linear momentum without *a priori* assuming such an equilibrium.[5]

Expressing the orthogonal-face tractions in terms of their components, equation (2.28) can be explicitly written as

$$\begin{aligned}\mathbf{T}^{(\mathbf{n})} &= \begin{bmatrix} T_1^{(\mathbf{e}_1)} \\ T_2^{(\mathbf{e}_1)} \\ T_3^{(\mathbf{e}_1)} \end{bmatrix} n_1 + \begin{bmatrix} T_1^{(\mathbf{e}_2)} \\ T_2^{(\mathbf{e}_2)} \\ T_3^{(\mathbf{e}_2)} \end{bmatrix} n_2 + \begin{bmatrix} T_1^{(\mathbf{e}_3)} \\ T_2^{(\mathbf{e}_3)} \\ T_3^{(\mathbf{e}_3)} \end{bmatrix} n_3 \\ &= \begin{bmatrix} T_1^{(\mathbf{e}_1)} & T_1^{(\mathbf{e}_2)} & T_1^{(\mathbf{e}_3)} \\ T_2^{(\mathbf{e}_1)} & T_2^{(\mathbf{e}_2)} & T_2^{(\mathbf{e}_3)} \\ T_3^{(\mathbf{e}_1)} & T_3^{(\mathbf{e}_2)} & T_3^{(\mathbf{e}_3)} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}. \end{aligned} \tag{2.29}$$

[5] Readers interested in the theorem relating the stress tensor and the balance of linear momentum might refer to Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover, pp. 132 – 135.

Equation (2.29) states that, at a given point, we can determine traction $\mathbf{T}^{(\mathbf{n})}$ that acts on an arbitrary plane through that point, provided we know the tractions at this point that act on the three mutually orthogonal planes.

Examining expressions (2.17) and (2.29), we conclude that

$$\mathbf{T}^{(\mathbf{n})} = \sigma^T \mathbf{n}, \tag{2.30}$$

where σ is given in expression (2.18). In the context of an arbitrary plane, we see that the entries of matrix σ are the components of a second-rank tensor. The fact that σ is a second-rank tensor is shown in Exercise 2.3.

Tensor σ_{ij} is called the stress tensor. This tensor is also known as Cauchy's stress tensor. The stress tensor allows us to determine the stress state associated with an infinitesimal plane of arbitrary orientation. The stress tensor takes into account both the direction of the traction and the orientation of the surface upon which the traction is acting.

In view of expression (2.18), we can rewrite equation (2.30) as

$$\mathbf{T}^{(\mathbf{n})} = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} \\ \sigma_{12} & \sigma_{22} & \sigma_{32} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix},$$

which can be concisely stated as

$$T_i^{(\mathbf{n})} = \sum_{j=1}^{3} \sigma_{ji} n_j, \qquad i \in \{1, 2, 3\}. \tag{2.31}$$

Expression (2.31) is an important statement of elasticity theory in the context of continuum mechanics. It relates the components of forces acting within the continuum to the orientation of the plane upon which the forces are acting. In other words, two vectorial properties, namely, traction, $\mathbf{T}^{(\mathbf{n})}$, and surface-normal vector, $\mathbf{n}$, are uniquely related by the stress tensor, σ_{ij}. The derivation performed in this section shows that in order to describe a traction related to an arbitrary plane, it is enough to consider tractions on three planes with linearly independent normals.[6]

[6] Readers interested in formulation of the stress tensor as a generalization of the concept of hydrostatic pressure might refer to Synge, J.L., and Schild, A., (1949/1978) Tensor calculus: Dover, pp. 205 – 208.

2.6 Cauchy's equations of motion

2.6.1 General formulation

In order to formulate the equations of motion, we consider the balance of linear momentum and the concept of the stress tensor.

In view of expression (2.31), we can write the balance of linear momentum, stated in equation (2.16), in terms of components, as

$$\iiint\limits_{V(t)} \rho \frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} \,\mathrm{d}V = \iint\limits_{S(t)} \sum_{j=1}^{3} \sigma_{ji} n_j \,\mathrm{d}S + \iiint\limits_{V(t)} f_i \,\mathrm{d}V, \qquad i \in \{1,2,3\}.$$

In this integral equation, we wish to express all integrals as volume integrals. Hence, invoking the divergence theorem, we can write

$$\iiint\limits_{V(t)} \rho \frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} \,\mathrm{d}V = \iiint\limits_{V(t)} \sum_{j=1}^{3} \frac{\partial \sigma_{ji}}{\partial x_j} \,\mathrm{d}V + \iiint\limits_{V(t)} f_i \,\mathrm{d}V, \qquad i \in \{1,2,3\}.$$

Using the linearity of the integral operator, we can rewrite this equation as

$$\iiint\limits_{V(t)} \left(\sum_{j=1}^{3} \frac{\partial \sigma_{ji}}{\partial x_j} + f_i - \rho \frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} \right) \mathrm{d}V = 0, \qquad i \in \{1,2,3\}, \tag{2.32}$$

which states the balance of linear momentum, as long as the portion of the continuum contained in volume $V(t)$ remains the same.

To derive Cauchy's equations of motion, consider equation (2.32). For this integral equation to be satisfied for an arbitrary volume that contains the same portion of the continuum, the integrand must be identically zero. Thus, we require

$$\sum_{j=1}^{3} \frac{\partial \sigma_{ji}}{\partial x_j} + f_i = \rho \frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2}, \qquad i \in \{1,2,3\}. \tag{2.33}$$

In view of equation (2.16), $\mathrm{d}^2/\mathrm{d}t^2$ refers to the material time-derivative operator. However, in this book, as discussed in Section 1.2.3, we use the linearized formulation and we can rewrite equations (2.33) as[7]

$$\sum_{j=1}^{3}\frac{\partial\sigma_{ji}}{\partial x_j}+f_i=\rho\frac{\partial^2 u_i}{\partial t^2},\qquad i\in\{1,2,3\}. \tag{2.34}$$

These are Cauchy's equations of motion. As shown in Exercise 2.4, the *SI* units of Cauchy's equations of motion are N/m^3.[8]

Cauchy's equations of motion relate two vectorial quantities, namely, the surface force — which corresponds to the summation term defining the divergence of tensor σ_{ji} — and the body force, to the acceleration vector. In other words, Cauchy's equations of motion state that the acceleration of an element within a continuum results from the application of surface and body forces.

If the acceleration term vanishes in equations of motion (2.34), we obtain the equations of static equilibrium,

$$\sum_{j=1}^{3}\frac{\partial\sigma_{ji}}{\partial x_j}+f_i=0,\qquad i\in\{1,2,3\}. \tag{2.35}$$

These equations describe the equilibrium state of an element of the continuum arising from the application of forces whose resultant is zero. Equations (2.35) are used to illustrate the symmetry of the stress tensor, as shown in Exercise 2.5. Equations (2.35) are also valid for rectilinear, constant-velocity motion.

Consider a system composed of equation of continuity (2.8) and Cauchy's equations of motion (2.34) in a three-dimensional continuum. This system contains four equations and sixteen unknowns, namely, mass density, ρ, stress-tensor components, σ_{11}, σ_{12}, σ_{13}, σ_{21}, σ_{22}, σ_{23}, σ_{31}, σ_{32}, σ_{33}, body-force components, f_1, f_2, f_3, and displacement-vector components, u_1, u_2, u_3.

Note that if we consider conservative systems, the three body-force components are derived from a single scalar function. In other words, $\mathbf{f}=\nabla U(\mathbf{x})$.

In our subsequent studies, we will reduce the discrepancy between the number of equations and the number of unknowns. In Section 2.7, we will

[7] Readers interested in this approximation might refer to Grant, F.S., and West, G.F., (1965) Interpretation theory in applied geophysics: McGraw-Hill Inc., pp. 28 – 29, and to Graff, K.F., (1975/1991) Wave motion in elastic solids: Dover, pp. 586 – 587.

[8] Readers interested in formulation of Cauchy's equations of motion as a generalization of equations of motion for a perfect fluid might refer to Synge, J.L., and Schild, A., (1949/1978) Tensor calculus: Dover, p. 208.

show that the stress tensor is symmetric, which results in only six independent stress-tensor components. Also, we will not consider the body force, $\mathbf{f} = [f_1, f_2, f_3]$.

Where we consider an infinitesimal element of the continuum, as we do in Part I, the body force is irrelevant. We can see this in view of the tetrahedron argument, discussed in Section 2.5.2, in particular, by examining the step between equations (2.27) and (2.28).

Where we study waves and rays, as we do in Part II and Part III, we invoke Cauchy's equations of motion (2.34), which, in general, contain both surface forces and body forces. However, if we consider sufficiently high frequencies, which is the case in applied seismology, the effects of the body forces are negligible as compared to the effects of the surface forces. In other words, the effects of gravitation are negligible as compared to the effects of elasticity.[9]

2.6.2 Example: Surface-forces formulation

To gain insight into the equations of motion without body forces, we rederive equations (2.34) without using the divergence theorem, which relates surface and volume integrals.

Consider the force acting in the positive direction of the x_1-axis on each coordinate plane. In view of definition (2.12), we can write the force acting along the x_1-axis as

$$T_1^{(\mathbf{e}_1)}\,\mathrm{d}x_2\mathrm{d}x_3 + T_1^{(\mathbf{e}_2)}\,\mathrm{d}x_1\mathrm{d}x_3 + T_1^{(\mathbf{e}_3)}\,\mathrm{d}x_1\mathrm{d}x_2, \tag{2.36}$$

where $\mathbf{e}_i$ denotes the unit normal to the coordinate plane on which $T_1^{(\mathbf{e}_i)}$ is acting, and $\mathrm{d}x_j\mathrm{d}x_k$ is the surface area of this planar element. Following expression (2.18), expression (2.36) can be rewritten as

$$\sigma_{11}\,\mathrm{d}x_2\mathrm{d}x_3 + \sigma_{21}\,\mathrm{d}x_1\mathrm{d}x_3 + \sigma_{31}\,\mathrm{d}x_1\mathrm{d}x_2. \tag{2.37}$$

Now, consider a small rectangular box subjected to stresses. Let the rectangular box be spanned by $\mathrm{d}x_1$, $\mathrm{d}x_2$ and $\mathrm{d}x_3$, with its sides being parallel to the orthonormal coordinate axes.

Consider the force acting in the positive direction of the x_1-axis on each face of the rectangular box. The resultant force along the x_1-axis is a sum of forces acting on the three sets of the parallel faces of the rectangular

[9]Readers interested in the effect of gravity on seismic wave propagation might refer to Udías, A., (1999) Principles of seismology: Cambridge University Press, pp. 39 – 40.

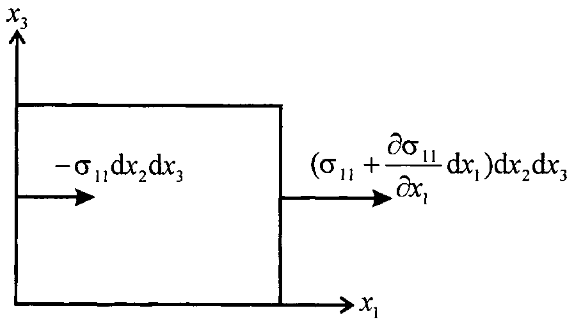


Figure 2.3: Two forces acting along the x_1-axis on faces that are parallel to it.

box. Within each set, the two parallel faces are separated by a distance dx_i. By convention, stated in Section 2.5.1, a stress component is positive if it acts in the positive direction of the coordinate axis and on the plane whose outward normal points in the positive coordinate direction. For each set of the two parallel faces of the aforementioned rectangular box, one face exhibits an outward normal that points in the positive coordinate direction while the other face exhibits an outward normal that points in the negative coordinate direction. Thus, in view of expression (2.37), we can write the resultant force along the x_1-axis as

$$\begin{aligned} dF_1 &= \left[\left(\sigma_{11}+\frac{\partial\sigma_{11}}{\partial x_1}dx_1\right)dx_2dx_3+\left(-\sigma_{11}\,dx_2dx_3\right)\right] \\ &+\left[\left(\sigma_{21}+\frac{\partial\sigma_{21}}{\partial x_2}dx_2\right)dx_1dx_3+\left(-\sigma_{21}\,dx_1dx_3\right)\right] \\ &+\left[\left(\sigma_{31}+\frac{\partial\sigma_{31}}{\partial x_3}dx_3\right)dx_1dx_2+\left(-\sigma_{31}\,dx_1dx_2\right)\right], \end{aligned} \tag{2.38}$$

which, for a given direction, contains all six separate forces acting on all the faces of the rectangular box. The expressions in brackets correspond to the sum of the two forces along the x_1-axis acting on faces orthogonal to the x_1-axis, the x_2-axis, and the x_3-axis, respectively. In other words, the first bracket denotes a sum of the two forces acting along the x_1-axis on the faces normal to it, as shown in Figure 2.3, while the second and the third brackets denote the sums of forces acting along the x_1-axis on the faces parallel to it.

Note that, in view of terms $\sigma_{i1}+(\partial\sigma_{i1}/\partial x_i)\,dx_i$, expression (2.38) is a first-order approximation. This approximation is consistent with our study in the context of linearized theory.

Expression (2.38) immediately simplifies to

$$\mathrm{d}F_1 = \left(\frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{21}}{\partial x_2} + \frac{\partial \sigma_{31}}{\partial x_3} \right) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3. \tag{2.39}$$

Invoking Newton's second law of motion in the form given by

$$\mathrm{d}F_1 = \rho\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \frac{\mathrm{d}^2 u_1}{\mathrm{d}t^2},$$

where ρ is the mass density of the small rectangular box and u_1 is the displacement in the x_1-direction, we can write expression (2.39) as

$$\rho \frac{\mathrm{d}^2 u_1}{\mathrm{d}t^2} = \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{21}}{\partial x_2} + \frac{\partial \sigma_{31}}{\partial x_3}. \tag{2.40}$$

Analogously, for the displacement-vector component along the x_2-axis and the displacement-vector component along the x_3-axis, we can write

$$\rho \frac{\mathrm{d}^2 u_2}{\mathrm{d}t^2} = \frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{32}}{\partial x_3} \tag{2.41}$$

and

$$\rho \frac{\mathrm{d}^2 u_3}{\mathrm{d}t^2} = \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{23}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3}, \tag{2.42}$$

respectively.

In view of linearization discussed in Section 1.2.3, total derivatives with respect to time are equivalent to partial derivatives. Consequently, expressions (2.40), (2.41) and (2.42) are equivalent to Cauchy's equations of motion (2.34) with no body forces.

2.7 Balance of angular momentum

Introductory comments

Motion within a continuum must also obey the balance of angular momentum. In deriving the differential equation to express the balance of angular momentum, we use the conservation of mass — by invoking time derivative of volume integral — and the balance of linear momentum. Hence, the constraints imposed by the balance of angular momentum do not add another independent differential equation. They do, however, reduce the number of unknowns, as shown below.

2.7.1 Integral equation

The balance of angular momentum can be stated as

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)}\left(\mathbf{x}\times\rho\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\right)\mathrm{d}V=\iint\limits_{S(t)}(\mathbf{x}\times\mathbf{T})\,\mathrm{d}S+\iiint\limits_{V(t)}(\mathbf{x}\times\mathbf{f})\,\mathrm{d}V, \tag{2.43}$$

where $V(t)$ is a volume that moves while always containing the same portion of the continuum and $S(t)$ is the surface containing this volume. The integrand on the left-hand side is the angular momentum, namely, the vector product of the distance between a reference point and the element of the continuum with the linear-momentum density $\rho\mathrm{d}\mathbf{u}/\mathrm{d}t$. The first integrand on the right-hand side is the vector product of this distance and force per unit area associated with this element, while the second integrand on the right-hand side is the vector product of that distance and force per unit volume associated with this element.

2.7.2 Symmetry of stress tensor

Since our formulation must obey the conservation of mass and the balance of linear momentum as well as the balance of angular momentum, we obtain an important consequence of these laws — the stress tensor is symmetric.

Theorem 2.1 *Consider a linearized formulation in a three-dimensional continuum. Let the principles of the conservation of mass and the balance of linear momentum hold. The balance of angular momentum holds if and only if*

$$\sigma_{ij}=\sigma_{ji},$$

where $i,j\in\{1,2,3\}$. *In other words, the stress tensor is symmetric.*

Notation 2.1 *The repeated-index summation notation is used in this proof. Any term in which an index appears twice stands for the sum of all such terms as the index assumes all the values between* 1 *and* 3.

Proof. We can rewrite expression (2.43) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)}\left[\rho\left(\mathbf{x}\times\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\right)\right]\mathrm{d}V=\iint\limits_{S(t)}(\mathbf{x}\times\mathbf{T})\,\mathrm{d}S+\iiint\limits_{V(t)}(\mathbf{x}\times\mathbf{f})\,\mathrm{d}V. \tag{2.44}$$

Invoking the time derivative of a moving-volume integral, given by expression (2.10) and letting $\mathcal{A} = \mathbf{x}\times \mathrm{d}\mathbf{u}/\mathrm{d}t$, we can restate expression (2.44) as

$$\iiint\limits_{V(t)} \rho\left(\mathbf{x}\times\frac{\mathrm{d}^2\mathbf{u}}{\mathrm{d}t^2}\right)\mathrm{d}V = \iint\limits_{S(t)} (\mathbf{x}\times\mathbf{T})\,\mathrm{d}S + \iiint\limits_{V(t)} (\mathbf{x}\times\mathbf{f})\,\mathrm{d}V. \tag{2.45}$$

In view of the linearized formulation discussed in Section 1.2.3, we can rewrite the balance of angular momentum, given in expression (2.45), as

$$\iiint\limits_{V(t)} \rho\left(\mathbf{x}\times\frac{\partial^2\mathbf{u}}{\partial t^2}\right)\mathrm{d}V = \iint\limits_{S(t)} (\mathbf{x}\times\mathbf{T})\,\mathrm{d}S + \iiint\limits_{V(t)} (\mathbf{x}\times\mathbf{f})\,\mathrm{d}V. \tag{2.46}$$

Using the stress tensor, invoking the divergence theorem, and in view of the validity of expression (2.46) for an arbitrary integration volume that consistently contains the same portion of the continuum, we obtain the differential equation given by

$$\rho\left(\mathbf{x}\times\frac{\partial^2\mathbf{u}}{\partial t^2}\right) = \nabla\cdot(\mathbf{x}\times\sigma) + \mathbf{x}\times\mathbf{f}. \tag{2.47}$$

Consider the first term on the right-hand side in equation (2.47). The expression in parentheses is a second-rank tensor whose ilth component can be written as

$$(\mathbf{x}\times\sigma)_{il} = \epsilon_{ijk}x_j\sigma_{kl}, \qquad i,l\in\{1,2,3\},$$

where ϵ_{ijk} is the permutation symbol. Taking the ith component of the divergence and using the product rule, we obtain

$$\begin{aligned}[\nabla\cdot(\mathbf{x}\times\sigma)]_i &= \frac{\partial}{\partial x_l}\left(\epsilon_{ijk}x_j\sigma_{kl}\right)\\ &= (\mathbf{x}\times\nabla\cdot\sigma)_i + \epsilon_{ijk}\delta_{jl}\sigma_{kl}, \qquad i\in\{1,2,3\}.\end{aligned} \tag{2.48}$$

Substituting expression (2.48) into equation (2.47), we obtain

$$\rho\left(\mathbf{x}\times\frac{\partial^2\mathbf{u}}{\partial t^2}\right)_i = (\mathbf{x}\times\nabla\cdot\sigma)_i + \epsilon_{ijk}\delta_{jl}\sigma_{kl} + (\mathbf{x}\times\mathbf{f})_i, \qquad i\in\{1,2,3\}.$$

Using the linearity of the cross-product operator, we can write

$$\begin{aligned}\left[\mathbf{x}\times\left(\rho\frac{\partial^2\mathbf{u}}{\partial t^2}\right)\right]_i - (\mathbf{x}\times\nabla\cdot\sigma)_i - (\mathbf{x}\times\mathbf{f})_i &= \left[\mathbf{x}\times\left(\rho\frac{\partial^2\mathbf{u}}{\partial t^2} - \nabla\cdot\sigma - \mathbf{f}\right)\right]_i \\ &= \epsilon_{ijk}\delta_{jl}\sigma_{kl}, \qquad i\in\{1,2,3\}.\end{aligned} \tag{2.49}$$

Invoking Cauchy's equations of motion (2.34), which can be written as

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} - \nabla \cdot \sigma - \mathbf{f} = 0,$$

the term in brackets in equation (2.49) vanishes. This implies

$$\epsilon_{ijk}\delta_{jl}\sigma_{kl} = 0, \qquad i \in \{1,2,3\}.$$

Using the properties of Kronecker's delta, we can rewrite this equation as

$$\epsilon_{ijk}\sigma_{kj} = 0, \qquad i \in \{1,2,3\},$$

which, in view of the properties of the permutation symbol, represents the equation given by

$$\sigma_{jk} = \sigma_{kj},$$

as required. ■

Remark 2.1 $\epsilon_{ijk}\sigma_{jk} = 0$ *is a summation of terms for a given* i*. For instance, for* $i = 1$*, the summation can be written as*

$$\sum_{j=1}^{3}\sum_{k=1}^{3} \epsilon_{1jk}\sigma_{jk} = 0.$$

By the properties of the permutation symbol, ϵ_{ijk}*, only two terms are nonzero; they are* σ_{23} *and* σ_{32}*. Also by the properties of the permutation symbol, these terms exhibit opposite signs. Thus we obtain*

$$\sigma_{23} - \sigma_{32} = 0.$$

Thus, in view of the balance of angular momentum, we see that the stress tensor is symmetric. Hence, the stress tensor has only six independent components and, in view of Section 2.5.2, these components are sufficient to determine the state of stress at any given point within a continuum.

Consequently, considering the system of fundamental equations and not including the body forces, we have four equations and ten unknowns.

2.8 Fundamental equations

The conservation of mass, the balance of linear momentum and the balance of angular momentum are the only three fundamental principles that relate the unknowns in our system. No other balance principles furnish us with additional constraints. For instance, the balance of energy, which deals with thermodynamic processes, does not add another fundamental equation or reduce the number of unknowns since we assume that the heat generated by the deformation is negligible and does not affect the process of deformation. The balance of energy does, however, play a key role in the formulation of the constitutive equations, which are discussed in Chapters 3 and 4.

Let us summarize the fundamental equations that describe the motion within a continuum.

In view of the symmetry of the stress tensor, the system of equations formed by expressions (2.8), (2.40), (2.41) and (2.42) consists of four equations, namely, the equation of continuity,

$$\frac{\partial \rho}{\partial t}+\sum_{i=1}^{3}\frac{\partial}{\partial x_i}\left(\rho\frac{\partial u_i}{\partial t}\right)=0,$$

and Cauchy's equations of motion with no body forces,

$$\sum_{j=1}^{3}\frac{\partial \sigma_{ij}}{\partial x_j}=\rho\frac{\partial^2 u_i}{\partial t^2}, \qquad i\in\{1,2,3\}, \tag{2.50}$$

where $\sigma_{ij}=\sigma_{ji}$, with $i,j\in\{1,2,3\}$. Explicitly, these equations can be written as

$$\frac{\partial \rho}{\partial t}+\frac{\partial}{\partial x_1}\left(\rho\frac{\partial u_1}{\partial t}\right)+\frac{\partial}{\partial x_2}\left(\rho\frac{\partial u_2}{\partial t}\right)+\frac{\partial}{\partial x_3}\left(\rho\frac{\partial u_3}{\partial t}\right)=0, \tag{2.51}$$

and

$$\frac{\partial \sigma_{11}}{\partial x_1}+\frac{\partial \sigma_{12}}{\partial x_2}+\frac{\partial \sigma_{13}}{\partial x_3}-\rho\frac{\partial^2 u_1}{\partial t^2}=0, \tag{2.52}$$

$$\frac{\partial \sigma_{12}}{\partial x_1}+\frac{\partial \sigma_{22}}{\partial x_2}+\frac{\partial \sigma_{23}}{\partial x_3}-\rho\frac{\partial^2 u_2}{\partial t^2}=0, \tag{2.53}$$

$$\frac{\partial \sigma_{13}}{\partial x_1}+\frac{\partial \sigma_{23}}{\partial x_2}+\frac{\partial \sigma_{33}}{\partial x_3}-\rho\frac{\partial^2 u_3}{\partial t^2}=0. \tag{2.54}$$

The resulting system of four equations contains ten unknowns, namely, ρ, u_1, u_2, u_3, σ_{11}, σ_{12}, σ_{13}, σ_{22}, σ_{23} and σ_{33}. This system of equations is underdetermined; there are not enough equations to uniquely determine the behaviour of the continuum. To render the system determined, we turn to constitutive equations, discussed in Chapter 3.

Closing remarks

In Chapter 3, in order to complete the system of equations, we associate Cauchy's equations of motion and the equation of continuity with the constitutive equations describing the relation between stress and strain in an elastic continuum. These constitutive equations also allow us to associate the fundamental equations with the specific properties of elastic materials. Notably, the wave equation and the eikonal equation, used extensively throughout the book, are rooted in Cauchy's equations of motion and the constitutive equations for elastic continua.

Exercises

Exercise 2.1 [10]*Several physical laws discussed in this book are stated as the vanishing of a definite integral which is tantamount to the vanishing of the integrand. Justify this equivalence using a one-dimensional case.*

Solution 2.1 *Consider an integral equation given by*

$$\int_A^B f(x)\,\mathrm{d}x = 0.$$

Let $f(x)$ be a continuous function in the interval $[A, B]$ and let $f(x_0) \neq 0$, for $x_0 \in [A, B]$. Because of the continuity, $f(x) \neq 0$ in the neighbourhood of x_0, and, hence, the integral taken over this neighbourhood does not vanish. Since we require $\int_A^B f(x)\,dx = 0$ for arbitrary limits of integration, we must require that $f(x) = 0$, for all $x \in [A, B]$.

Exercise 2.2 *Using expressions (2.17) and (2.29), obtain the components of the traction vector acting on the plane whose normal is parallel to the x_1-axis. Compare the results to expression (2.30).*

Solution 2.2 *Following expression (2.17), we can immediately write the components of the traction vectors acting on the plane whose normal is parallel to the x_1-axis as $\left[T_1^{(\mathbf{e}_1)}, T_2^{(\mathbf{e}_1)}, T_3^{(\mathbf{e}_1)}\right]$. In view of definition (2.18), we*

[10]See also Section 2.1.1.

can rewrite these components as $[\sigma_{11}, \sigma_{12}, \sigma_{13}]$. *Using equation (2.29), the traction vector acting on the plane whose unit normal is parallel to the* x_1-*axis is given by*

$$\begin{bmatrix} T_1^{(\mathbf{e}_1)} & T_1^{(\mathbf{e}_2)} & T_1^{(\mathbf{e}_3)} \\ T_2^{(\mathbf{e}_1)} & T_2^{(\mathbf{e}_2)} & T_2^{(\mathbf{e}_3)} \\ T_3^{(\mathbf{e}_1)} & T_3^{(\mathbf{e}_2)} & T_3^{(\mathbf{e}_3)} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} T_1^{(\mathbf{e}_1)} \\ T_2^{(\mathbf{e}_1)} \\ T_3^{(\mathbf{e}_1)} \end{bmatrix}.$$

In view of definition (2.18), we can rewrite these components as

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{13} \end{bmatrix},$$

as expected from the property stated in expression (2.30).

Exercise 2.3 [11] *Using the stress tensor, prove the particular case of the following theorem.*

Theorem 2.2 *If an mth-rank tensor is linearly related to an nth-rank tensor through a quantity that possesses* $n+m$ *indices, then this quantity is an* $(n+m)$*rank tensor.*

Solution 2.3 *Consider the stress tensor that relates two vectors, namely, the traction and the unit normal vector. Thus, two first-rank tensors are linearly related by a second-rank tensor.*

Notation 2.2 *The repeated-index summation notation is used in this proof. Any term in which an index appears twice stands for the sum of all such terms as the index assumes all the values between* 1 *and* 3.

Proof. The relationship between the components of the traction, $\mathbf{T}$, in two coordinate systems can be stated as

$$\hat{T}_i = a_{ik} T_k, \qquad i \in \{1, 2, 3\},$$

where a_{ij} are the entries of the transformation matrix. Also, the components of the traction, $\mathbf{T}$, are related to the components of the normal vector, $\mathbf{n}$, by the quantity σ, as

$$T_k = \sigma_{kj} n_j, \qquad k \in \{1, 2, 3\}.$$

[11] See also Sections 2.5.2 and 5.2.1.

Combining both expressions, we can write

$$\hat{T}_i = a_{ik}\sigma_{kj}n_j, \qquad i \in \{1,2,3\}.$$

Since $\mathbf{n}$ is a vector, it obeys the inverse transformation laws, namely,

$$n_j = a_{mj}\hat{n}_m, \qquad j \in \{1,2,3\}.$$

Thus, we can write

$$\hat{T}_i = a_{ik}\sigma_{kj}a_{mj}\hat{n}_m, \qquad i \in \{1,2,3\}.$$

Since the relationship between the components of the traction, $\mathbf{T}$, and the components of the normal vector, $\mathbf{n}$, are valid for all coordinate systems, we can formally write

$$\hat{T}_i = \hat{\sigma}_{im}\hat{n}_m, \qquad i \in \{1,2,3\}.$$

Subtracting the two equations for T_i^* from one another, we obtain

$$\left(a_{ik}\sigma_{kj}a_{mj} - \hat{\sigma}_{im}\right)\hat{n}_m = 0, \qquad i \in \{1,2,3\}.$$

Since the result must hold for any orientation of the vector $\mathbf{n}$, as required by the physical argument discussed in this chapter, we get

$$a_{ik}\sigma_{kj}a_{mj} - \hat{\sigma}_{im} = 0, \qquad i,m \in \{1,2,3\},$$

and we can restate it as

$$\hat{\sigma}_{im} = a_{ik}a_{mj}\sigma_{kj}, \qquad i,m \in \{1,2,3\}. \tag{2.55}$$

The last expression shows that σ obeys standard transformation rules for a second-rank tensor. Consequently, σ, which linearly relates two vectors, is a second-rank tensor. ■

Remark 2.2 *The quotient rule, stated in this theorem, is also exemplified by the stress-strain equations (3.1), where two second-rank tensors are linearly related by a fourth-rank elasticity tensor, namely,*

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl}, \qquad i,j \in \{1,2,3\},$$

with ε_{kl} denoting the strain tensor, and where the repeated index assumes all the values between 1 *and* 3.

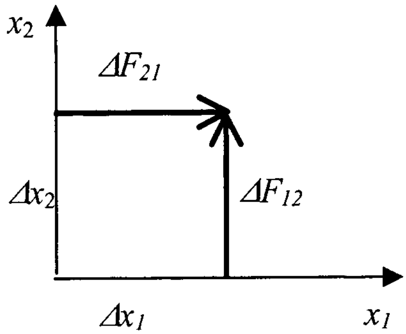


Figure 2.4: An x_1x_2-cross-section of a rectangular box, $\Delta x_1 \Delta x_2 \Delta x_3$, with the moment-producing forces, F_{12} and F_{21}. The directions of the two forces are perpendicular to one another while their magnitudes are equal.

Exercise 2.4 [12]*Find the physical SI units of equations of motion (2.34). Show that these units are consistent for all terms involved.*

Solution 2.4

$$\sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j} + f_i = \rho \frac{\partial^2 u_i}{\partial t^2}, \qquad i \in \{1, 2, 3\}$$

Following the definition of stress as force per unit area, the units of stress tensor are $[N/m^2]$. *Consequently, the units of the first term of the left-hand side are* $[N/m^3]$. *In view of* f_i *being the components of force per unit volume, the units are also* $[N/m^3]$. *On the right-hand side, the units of mass density are* $[kg/m^3]$, *while the units of acceleration are* $[m/s^2]$, *resulting in* $[kg/(m^2 s^2)]$. *Since* $[N] = [kgm/s^2]$, *the units of the right-hand side are also* $[N/m^3]$, *as expected. Thus, the physical units of equations (2.34) are* $[N/m^3]$.

Exercise 2.5 [13]*Using Figure 2.4, prove the following theorem.*

Theorem 2.3 *The stress tensor is symmetric, namely,* $\sigma_{ij} = \sigma_{ji}$, *where* $i, j \in \{1, 2, 3\}$.

[12]See also Section 2.6.1.
[13]See also Section 2.7.2.

Solution 2.5 .

Proof. Consider a rectangular box that is an element within a continuum. Let the volume of this box, whose edges are parallel to the coordinate axes, be

$$\Delta V = \Delta x_1 \Delta x_2 \Delta x_3.$$

We require that this element of volume does not rotate within the continuum. This requirement implies that the sum of moments acting on this box must be zero. The sum of moments about the x_3-axis is zero if

$$\Delta F_{12} \Delta x_1 = \Delta F_{21} \Delta x_2. \tag{2.56}$$

Using formulations of traction and the stress-tensor components, we can write $\sigma_{1j} = \Delta F_{1j} / \Delta S_1$. Thus, we have

$$\Delta F_{12} = \sigma_{12} \Delta S_1 = \sigma_{12} \Delta x_2 \Delta x_3, \tag{2.57}$$

and

$$\Delta F_{21} = \sigma_{21} \Delta S_2 = \sigma_{21} \Delta x_1 \Delta x_3. \tag{2.58}$$

Inserting expressions (2.57) and (2.58) into equation (2.56), we obtain

$$\sigma_{12} \Delta x_1 \Delta x_2 \Delta x_3 = \sigma_{21} \Delta x_1 \Delta x_2 \Delta x_3,$$

which implies

$$\sigma_{12} = \sigma_{21}.$$

Hence, together with the equality of the sum of moments about the x_1-axis and the x_2-axis, we can write

$$\sigma_{ij} = \sigma_{ji}, \qquad i, j \in \{1, 2, 3\},$$

as required. ■

Chapter 3

Stress-strain equations

> ... there is a conjecture that two sets of small motions may be superimposed without interfering with each other in a nonlinear fashion. Another conjecture is that the seismic motions set up by some physical source should be uniquely determined by the combined properties of that source and of the medium of wave propagation. These conjectures, and many others that are generally assumed by seismologists to be true, are properties of infinitesimal motion in classical continuum mechanics for an elastic medium with a linear stress-strain relation;
>
> *Keiiti Aki and Paul G. Richards (1980) Quantitative seismology: Theory and methods*

Preliminary remarks

The equations resulting from the fundamental principles discussed in Chapter 2 are valid for any continuum irrespective of its constitution. In other words, they do not explicitly account for distinctive properties of a particular material. Also, these equations constitute a system of differential equations that contains more unknowns than equations.

In order to consider the properties of a particular material and to formulate a determined system of equations that describes the propagation of deformations within that material, we turn our attention to empirical relations that can be expressed as constitutive equations. These equations are based on experimental observations of actual materials. An elastic continuum is defined by the constitutive equations that, in accordance with experimental observations, state that for elastic materials, forces are linearly related to small deformations.

We begin this chapter with the formulation of linear stress-strain equations, which underlie the theory of elasticity used in this book. We then express these equations in both tensorial and matrix forms.

3.1 Formulation of stress-strain equations

Introductory comments

Perhaps the best-known constitutive equation is based on Hooke's law of elasticity discovered by Robert Hooke in the middle of the seventeenth century. This law furnishes us with the physical justification for the mathematical theory of linear elasticity.

Ut tensio sic vis — "as the extension, so the force" is a famous statement from Hooke's work of 1676. He described it in more detail by writing that

> the power of any spring is in the same proportion with the tension thereof: that is, if one power stretch or bend it in one space, two will bend in two, three will bend in three, and so forward. And this is the rule or law of Nature, upon which all manner of restituent or springing motion doth proceed.

In an earlier paper, "De potentia restitutiva", Hooke published the results of his experiments with elastic materials and stated that

> it is very evident that the rule or law of Nature in every springing body is, that the force or power thereof to restore itself to its natural position is always proportional to the distance or space it is removed therefrom....

In the modern terminology of continuum mechanics and in view of Chapters 1 and 2, the linearity of Hooke's law can be stated in the following manner.

> At any point of a continuum, each component of the stress tensor is a linear function of all the components of the strain tensor.

This statement is used to formulate stress-strain equations, which are introduced in this chapter. The restoring force is discussed in Chapter 4.

3.1.1 Tensorial form

At a given point $\mathbf{x}$ of the continuum, Hooke's law, expressing each stress-tensor component, σ_{ij}, as a linear combination of all the strain-tensor components, ε_{kl}, can be written for a three-dimensional continuum as

$$\sigma_{ij} = \sum_{k=1}^{3} \sum_{l=1}^{3} c_{ijkl} \varepsilon_{kl}, \qquad i,j \in \{1,2,3\}, \tag{3.1}$$

where c_{ijkl} are the components of a tensor, known as the elasticity tensor.[1] Since the units of the stress-tensor components are N/m^2, and the strain-tensor components are dimensionless, the units of the elasticity-tensor components are N/m^2.

Note that c_{ijkl} relates two second-rank tensors. Hence, in view of tensor algebra, the elasticity tensor must be a fourth-rank tensor. Consequently, in a three-dimensional continuum, it has $3^4 = 81$ components.

In this book, we study continua that are described by stress-strain equations (3.1).[2] To understand the description of a continuum that is provided by c_{ijkl}, consider these stress-strain equations. In view of Chapters 1 and 2, tensors σ_{ij} and ε_{kl} are direction-dependent. Hence, the values of c_{ijkl} are intrinsically direction-dependent. Consequently, at a given point $\mathbf{x}$ of a continuum, these values determine the anisotropic properties of the continuum at this point. Furthermore, if the values of c_{ijkl} depend on position $\mathbf{x}$, the continuum is inhomogeneous. This is explicitly used in stress-strain equations (7.2).

Note the following distinction between the continuum model and real materials. While studying anisotropy and inhomogeneity in real materials, we observe that anisotropy is rooted in the inhomogeneity of the material. Intrinsically, anisotropy results from the inhomogeneity exhibited by an atomic structure or crystal lattice. In a seismological context, anisotropy results from the arrangement of grains or layers in the materials through which seismic waves propagate. Hence, physically, at some scale, anisotropy is linked to inhomogeneity. In the mathematical context of continuum mechanics, however, anisotropy and inhomogeneity are two distinct properties.

---

[1] Elasticity tensor is also commonly referred to as the stiffness tensor. Our nomenclature is consistent with Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover, pp. 9 – 10, and with Marsden, J.E., and Ratiu, T.S., (1999) Introduction to mechanics and symmetry: A basic exposition of classical mechanical systems (*2nd* edition): Springer-Verlag, p. 113.

[2] Readers interested in detailed formulations of elasticity, hyperelasticity, linear elasticity, etc., might refer to Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover.

3.1.2 Matrix form

Introductory comments

Due to the symmetries of the stress and strain tensors, constitutive equations (3.1) can be conveniently written in a matrix form containing six independent equations. This form, which allows us to express elasticity tensor (3.1) as an elasticity matrix, is often used in this book.

Note that, although, in some particular cases, the components of a tensor can be written as the entries of a matrix, the matrices and the tensors are distinct mathematical entities.

Stress-tensor and strain-tensor symmetries

At every point of a continuum, as shown in Section 2.7, in view of the balance of angular momentum, the stress tensor is symmetric, namely, $\sigma_{ij} = \sigma_{ji}$. Also, the strain tensor is symmetric, namely, $\varepsilon_{kl} = \varepsilon_{lk}$, by its definition (1.15). Consider stress-strain equations (3.1), which describe the states of stress and strain at a given point.

Consider the symmetry of the stress tensor. In view of this symmetry, we can write stress-strain equations (3.1) as

$$\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl} = \sigma_{ij} = \sigma_{ji} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{jikl}\varepsilon_{kl}, \qquad i,j \in \{1,2,3\}. \tag{3.2}$$

In other words, each double-summation term gives the same value of the stress-tensor component at the given point.

Subtracting the first double-summation term from the second one, we can write

$$\sum_{k=1}^{3}\sum_{l=1}^{3} c_{jikl}\varepsilon_{kl} - \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl} = \sum_{k=1}^{3}\sum_{l=1}^{3} \left(c_{ijkl} - c_{jikl}\right)\varepsilon_{kl} = 0,$$

where $i,j \in \{1,2,3\}$. Thus, for this equation to be satisfied for all strain-tensor components, we require

$$c_{ijkl} = c_{jikl}, \qquad i,j,k,l \in \{1,2,3\}. \tag{3.3}$$

Hence, due to the symmetry of the stress tensor, the elasticity tensor is invariant under permutations in the first pair of subscripts.

Consider the symmetry of the strain tensor. The order of k and l has no effect on stress-strain equations (3.1) since they are the summation indices. Hence, we can write

$$\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijlk}\varepsilon_{lk}, \qquad i,j \in \{1,2,3\}.$$

In view of the symmetry of the strain tensor, we can rewrite it as

$$\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijlk}\varepsilon_{kl}, \qquad i,j \in \{1,2,3\},$$

which, we can also state as

$$\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl} - \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijlk}\varepsilon_{kl} = \sum_{k=1}^{3}\sum_{l=1}^{3} \left(c_{ijkl} - c_{ijlk}\right)\varepsilon_{kl} = 0,$$

where $i,j \in \{1,2,3\}$. For this equation to be satisfied for all strain-tensor components, we require

$$c_{ijkl} = c_{ijlk}, \qquad i,j,k,l \in \{1,2,3\}. \tag{3.4}$$

Hence, due to the symmetry of the strain tensor, the elasticity tensor is invariant under permutations in the second pair of subscripts.

Elasticity matrix

In view of equalities (3.3) and (3.4), the number of independent components of the elasticity tensor is thirty-six. These components can be written as entries C_{mn} of a 6×6 elasticity matrix, which relates the six independent stress-tensor components to the six independent strain-tensor components. To construct this matrix, in view of symmetries (3.3) and (3.4), it is enough to consider the pairs of (i,j) and (k,l) for $i \leq j$ and $k \leq l$, respectively.

Consider such pairs (i,j), where $i,j \in \{1,2,3\}$. Let us arrange them in the order given by

$$(1,1),(2,2),(3,3),(2,3),(1,3),(1,2).$$

Now, we can replace each pair by a single number m that gives the position of the pair in this list; thus, $m \in \{1,\ldots,6\}$. In other words, we make the following replacement $(i,j) \to m$:

$$(1,1) \to 1, (2,2) \to 2, (3,3) \to 3,$$
$$(2,3) \to 4, (1,3) \to 5, (1,2) \to 6.$$

We can concisely write this replacement as

$$\left\{\begin{array}{lll} m=i & \text{if} & i=j \\ m=9-(i+j) & \text{if} & i\neq j \end{array}\right., \qquad i,j\in\{1,2,3\}. \tag{3.5}$$

Considering the analogous pairs (k,l), where $k,l\in\{1,2,3\}$, we see that identical replacements can be made. Consequently, we can replace c_{ijkl}, where $i,j,k,l\in\{1,2,3\}$, by C_{mn}, where $m,n\in\{1,\ldots,6\}$.

Thus, equations (3.1) can be restated as

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix}. \tag{3.6}$$

Note that the factors of 2 result from the fact that for a given $k\neq l$, the corresponding strain-tensor component appears twice in the summation on the right-hand side of equations (3.1) — as ε_{kl} and as ε_{lk}. Also note that, due to the symmetry of the stress tensor, it is sufficient to consider only six among the original nine equations stated in expression (3.1).

We could also replace the pairs of subscripts for ε_{kl} and σ_{ij} by single subscripts. However, we keep the original notation of these components in order that their physical meaning remains apparent, as discussed in Sections 1.3.2 and 2.5.1, respectively.

In concise notation, we write stress-strain equations (3.6) as

$$\underline{\sigma} = \mathbf{C}\underline{\varepsilon}, \tag{3.7}$$

where $\underline{\sigma}$ and $\underline{\varepsilon}$ are six-entry, single-column matrices, composed of the stress-tensor and the strain-tensor components, respectively, while $\mathbf{C}$ is a 6×6 matrix.

3.2 Determined system

Stress-strain equations furnish us with six additional equations and no new unknowns for the system discussed in Chapter 2. The system is no longer underdetermined.

Note that the strain-tensor components, in accordance with definition (1.15), may be expressed in terms of the displacement-vector components,

u_i, where $i \in \{1,2,3\}$, which are the unknowns used in the equations of motion and the equation of continuity, as illustrated in Exercise 3.1. Thus, in a three-dimensional continuum, we have a system of ten equations for ten unknowns. These equations are the equation of continuity, given by expression (2.51), the three equations of motion, given by expressions (2.52), (2.53) and (2.54), and six constitutive equations.

Note that the consistency of this system requires the linearized theory that allows us to ignore the fact that, in principle, equations of motion (2.34) refer to the spatial coordinates while definition (1.15), which is used in formulating stress-strain equations, refers to the material coordinates. In other words, we ignore the distinction between the spatial and the material coordinates and use the equations of motion and the stress-strain equations in the same system of equations.

Closing remarks

We use constitutive equations, namely, stress-strain equations (3.1) or, equivalently, equations (3.6) to obtain a determined system of equations that describes the propagation of deformations in elastic continua. For many seismological studies, the linear equations relating the stress-tensor components and the strain-tensor components agree, within sufficient accuracy, with experimental observations involving small deformations.

Stress-strain equations (3.1) or (3.6) link the fundamental principles with the properties of a particular elastic material. Notably, this link allows us to investigate Cauchy's equations of motion in the context of elastic materials, which leads to the wave equation and the eikonal equation, discussed in Chapters 6 and 7, respectively.

Stress-strain equations (3.1) or (3.6) describe the continuum whose deformations are linearly related to loads. For this continuum to represent an elastic material, we require the existence of the restoring force that allows, upon the removal of the load, the return to the undeformed state. In Chapter 4, we investigate the effects of this requirement upon parameters c_{ijkl} and C_{mn}.

Exercises

Exercise 3.1 [3] *Consider a one-dimensional homogeneous continuum. Using stress-strain equations (3.6), equation of continuity (2.8) and Cauchy's equations of motion (2.34) with no body force, write the resulting system of two differential equations.*

Solution 3.1 *Following equations (3.6) and considering a one-dimensional continuum that coincides with the x_1-axis, we can write*

$$\sigma_{11} = C_{11}\varepsilon_{11},$$

which, in view of definition (1.15) can be written as

$$\sigma_{11} = C_{11}\frac{\partial u_1(x,t)}{\partial x_1}, \tag{3.8}$$

where, due to the homogeneity of the continuum, C_{11} is a constant. The corresponding equation of continuity, whose general form is given by expression (2.8), is

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial}{\partial x_1}\left[\rho(x,t)\frac{\partial u_1(x,t)}{\partial t}\right] = 0, \tag{3.9}$$

and Cauchy's equation of motion, whose general form is given by expression (2.34), is

$$\frac{\partial \sigma_{11}}{\partial x_1} = \rho(x,t)\frac{\partial^2 u_1(x,t)}{\partial t^2}. \tag{3.10}$$

Inserting expression (3.8) into equation (3.10), differentiating and rearranging, we obtain

$$\frac{\partial^2 u_1(x,t)}{\partial x_1^2} = \frac{\rho(x,t)}{C_{11}}\frac{\partial^2 u_1(x,t)}{\partial t^2}. \tag{3.11}$$

Equations (3.9) and (3.11) constitute the required system of two differential equations in two unknowns, namely, $u_1(x,t)$ and $\rho(x,t)$, whose variables are x and t.

Remark 3.1 *If the mass-density function, $\rho(x,t)$, is given by a constant, equation (3.11) is a one-dimensional wave equation, discussed in Chapter 6.*

[3] See also Sections 3.2 and 4.4.2.

Chapter 4

Strain energy

> Ce qui fait la beauté d'une œuvre d'art, ce n'est pas la simplicité de ses parties, c'est plutôt une sorte d'harmonie globale qui donne à l'ensemble un aspect d'unité et d'homogénéité malgré la complication parfois très grande des détails. [...] La beauté des théories scientifiques nous paraît essentiellement de la même nature: elle s'impose quand, dominant sans cesse les raisonnements et les calculs, se retrouve partout une même idée centrale qui unifie et vivifie tout le corps de la doctrine.[1]
>
> *Louis de Broglie (1941) Continue et discontinue en physique moderne*

Preliminary remarks

When a material undergoes a deformation, energy is expended to deform it. In view of balance of energy, the energy expended must be converted into another form of energy. Elasticity of an actual material results from the fact that a large part of the expended energy associated with the deformation is converted to potential energy stored within the deformed material. For elastic continua, we assume that all the expended energy is stored within the strained continuum. We refer to this energy as strain energy.

[1] What makes the beauty of a work of art is not the simplicity of its parts, it is rather a kind of global harmony which gives to the whole an aspect of unity and homogeneity in spite of, at times, very large complications of details. [...] The beauty of scientific theories is of the same nature: this beauty is striking when, constantly dominating the reasoning and the calculations, one finds everywhere the same central idea which unifies and inspires the entire body of the formulation.

It is important to emphasize that the existence of strain energy, which allows the strained continuum to regain its initial state upon the removal of the load, is the defining property of an elastic continuum. The mathematical expression of this physical entity is the strain-energy function.

We begin this chapter with the derivation of the strain-energy function. Subsequently, in view of this function, we obtain another symmetry of c_{ijkl}, beyond the ones shown in Chapter 3. Then we derive the physical constraints on c_{ijkl}, which arise from the strain-energy function. This chapter concludes with the system of equations describing the behaviour of elastic continua.

4.1 Strain-energy function

For elastic continua, we assume that all the expended energy is stored in the strained continuum as a potential energy. In other words, we are dealing with a conservative system. We wish to formulate the corresponding potential-energy function.

To motivate our formulation, consider a force, $\mathbf{F}$, acting on a conservative system to increase the potential energy, $U(\mathbf{x})$, of this system. We can write the components of such a force as $\partial U/\partial x_i = F_i$, where $i \in \{1,2,3\}$. By analogy, let us postulate

$$\frac{\partial W(\varepsilon)}{\partial \varepsilon_{ij}} = \sigma_{ij}, \qquad i,j \in \{1,2,3\}, \tag{4.1}$$

where — in the context of elasticity theory — W is the potential-energy function of a conservative system. In other words, we postulate that the stress tensor is derived from this scalar function. To obtain the explicit expression for W, we use stress-strain equations (3.1) to write expression (4.1) as

$$\frac{\partial W(\varepsilon)}{\partial \varepsilon_{ij}} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl}, \qquad i,j \in \{1,2,3\}. \tag{4.2}$$

Integrating both sides of equations (4.2) with respect to ε_{ij}, we obtain

$$W(\varepsilon) = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl}\varepsilon_{ij}, \tag{4.3}$$

where we set the integration constant to zero. The vanishing of this constant results from the convention that, for unstrained continua, $W = 0$.

W in expression (4.3) is the strain-energy function. It is the desired potential-energy function that corresponds to elastic continua subjected to infinitesimal strains.[2]

Note that W has the units of energy per volume.

Examining expression (4.3), we recognize that strain energy is given by a homogeneous function of degree 2 in the strain-tensor components.[3] The fact that the strain-energy function is homogeneous of degree 2 in the ε_{ij}, follows from Definition A.1, which is discussed in Appendix A. This property of the strain-energy function is illustrated in Exercise 4.2. As shown in this exercise, expression (4.3) can be viewed as a second-degree polynomial in the strain-tensor components where both the constant term and the linear term vanish. A mathematical application of the homogeneity of W is illustrated in Exercise 4.3.

4.2 Strain-energy function and elasticity-tensor symmetry

4.2.1 Fundamental considerations

The existence of the strain-energy function, which defines an elastic continuum, implies the invariance of the elasticity tensor, c_{ijkl}, under permutations of pairs of subscripts ij and kl. This can be derived in the following manner.

Let us return to equations (4.2). Differentiating both sides of these equations with respect to ε_{kl}, we obtain

$$\frac{\partial^2 W(\varepsilon)}{\partial\varepsilon_{kl}\partial\varepsilon_{ij}} = c_{ijkl}, \qquad i,j,k,l \in \{1,2,3\}. \tag{4.4}$$

Now, let us invoke the equality of mixed partial derivatives, which states that, if W is a well-behaved function, the order of differentiation is interchangeable.[4] In view of expression (4.4), this implies that

[2]Readers interested in a general formulation for finite strains might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall., pp. 282 - 285.

[3]Both terms "degree" and "order" are commonly used to describe the homogeneity of a function. In this book, we use the former term since it refers to the value of the exponent and, hence, is consistent with other uses of this term, such as "degree of a polynomial".

[4]The equality of mixed partial derivatives is often used in this book. We can state it by the following theorem.

Theorem 4.1 *Let $f = f(x,y)$. Assume that the partial derivatives $\partial f/\partial x$, $\partial f/\partial y$,*

$$c_{ijkl} = c_{klij}, \qquad i,j,k,l \in \{1,2,3\}. \tag{4.5}$$

Hence, we conclude that the elasticity tensor is invariant under permutations of pairs of subscripts ij and kl.

We can also justify symmetry $c_{ijkl} = c_{klij}$ in the following manner.[5] Recalling that σ_{ij} and ε_{ij} are associated with force and displacement, respectively, we can write the element of work as

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\,\mathrm{d}\varepsilon_{ij}.$$

In view of the balance of energy, the element of work equals the total differential of W, namely,

$$\mathrm{d}W = \sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\,\mathrm{d}\varepsilon_{ij}. \tag{4.6}$$

Note that the requirement for the element of work to be a total differential results from the fact that the value of work must be independent of the integration path. The physical justification for this is that the work cannot depend on the path of deformations, but only on the difference between the initial and final states. Otherwise, we could deform the material following one path and let it return to its initial state along a different path. If the amount of energy were not the same for all paths, we could create or destroy energy by this process.

Invoking stress-strain equations (3.1), we can rewrite expression (4.6) as

$$\mathrm{d}W = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}\varepsilon_{kl}\,\mathrm{d}\varepsilon_{ij}. \tag{4.7}$$

$\partial^2 f/\partial x\partial y$ and $\partial^2 f/\partial y\partial x$ exist and are continuous. Then

$$\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x}.$$

Readers interested in the proof of this theorem might refer to Lang, S., (1973) Calculus of several variables: Addison-Wesley Publishing Co., pp. 110 – 111, or to Stewart, J., (1995), Multivariable calculus ($3rd$ edition): Brooks/Cole Publishing Co., p. A2.

[5] Readers interested in this formulation might also refer to Ting, T.C.T., (1996) Anisotropic elasticity: Theory and applications: Oxford University Press, p. 33.

Since this has to be a total differential, we require that[6]

$$\frac{\partial}{\partial \varepsilon_{st}} \sum_{k=1}^{3} \sum_{l=1}^{3} c_{ijkl} \varepsilon_{kl} = \frac{\partial}{\partial \varepsilon_{ij}} \sum_{k=1}^{3} \sum_{l=1}^{3} c_{stkl} \varepsilon_{kl}, \qquad i,j,s,t \in \{1,2,3\},$$

which gives

$$c_{ijst} = c_{stij}, \qquad i,j,s,t \in \{1,2,3\}.$$

Upon renaming the indices, we can write

$$c_{ijkl} = c_{klij}, \qquad i,j,k,l \in \{1,2,3\}, \tag{4.8}$$

which are equations (4.5). Hence, we rederived equations (4.5).

Also, in view of conditions (4.8), we can integrate both sides of equation (4.7) to obtain

$$W = \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \sum_{l=1}^{3} c_{ijkl} \varepsilon_{kl} \varepsilon_{ij}, \tag{4.9}$$

which is expression (4.3), as expected.

4.2.2 Stress-strain equations

Since every elastic continuum must obey equations (4.1), conditions (4.5) provide constraints on the components of the fourth-rank tensor in stress-strain equations (3.1).

As shown in Sections 3.1.1 and 3.1.2, we can express stress-strain equations in either the tensorial or matrix form. Recalling expression (3.5), the equality $c_{ijkl} = c_{klij}$ implies $C_{mn} = C_{nm}$ since switching ij with kl is tantamount to switching m with n. In other words, the elasticity matrix is symmetric, namely,

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}. \tag{4.10}$$

[6]Readers interested in this requirement might refer to Courant, R., and John, F., (1974/1989) Introduction to calculus and analysis: Springer-Verlag, Vol. II, p. 84, or to Zill, D.G., and Cullen, M.R., (1997) Differential equations with boundary-value problems (*4th* edition): Brooks/Cole Publishing Company, p. 39.

Thus, we see that the existence of the strain-energy function reduces the number of independent entries of **C** from thirty-six, used in equations (3.6), to twenty-one, stated in matrix (4.10).

In view of conditions (4.5) and resulting matrix (4.10), stress-strain equations (3.6) become

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix}. \tag{4.11}$$

Equations (4.11) are the matrix form of the stress-strain equations for a general elastic continuum that obeys Hooke's law. The twenty-one independent entries of the symmetric matrix that relates the stress-tensor components to the strain-tensor components are the elasticity parameters, which — together with mass density — fully describe a given elastic continuum.

4.2.3 Coordinate transformations

The strain energy, W, is a scalar quantity and, hence, its value is invariant under coordinate transformations. To achieve this invariance, in general, the values of the parameters c_{ijkl} or C_{mn} depend on the orientation of the coordinate system. In other words, the values of these parameters ensure that W is invariant under coordinate transformations. Hence, for an elastic continuum, expression (4.9) that contains a given set of elasticity parameters holds only for one orientation of the coordinate system.[7] If, for a particular continuum, expression (4.9) with a given set of elasticity parameters holds for several orientations of the coordinate system, this continuum possesses particular symmetries, which lead to further simplifications of matrix (4.10). Such symmetries are discussed in Chapter 5.

4.3 Stability conditions

4.3.1 Physical justification

Strain-energy function, W, given in expression (4.3), is formulated in terms of parameters c_{ijkl}, where $i, j, k, l \in \{1, 2, 3\}$. This function provides the sole

[7] Interested readers might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, p. 285.

fundamental constraints on these parameters. These constraints are called stability conditions since they constitute a mathematical statement of the fact that it is necessary to expend energy in order to deform a material. In other words, if energy is not expended, the material remains stable in its undeformed state.

In general, energy is a positive quantity. By convention, the strain energy of an undeformed continuum is zero. Thus, the strain-energy function must be a positive quantity that vanishes only in the undeformed state.[8]

4.3.2 Mathematical formulation

Mathematically, the stability conditions are equivalent to the positive-definiteness of the elasticity matrix. This can be derived in the following manner.

In view of expression (4.3) and by equivalence of stress-strain equations (3.1) and (3.6), we can write the strain-energy function as

$$W = \frac{1}{2}\left(\mathbf{C}\underline{\varepsilon}\right) \cdot \underline{\varepsilon}, \tag{4.12}$$

where $\mathbf{C}$ is matrix (4.10), and $\underline{\varepsilon}$ is the strain matrix, shown explicitly in equation (3.6). In view of Section 4.3.1, we require that

$$\left(\mathbf{C}\underline{\varepsilon}\right) \cdot \underline{\varepsilon} \geq 0, \tag{4.13}$$

where the equality sign corresponds to the case where $\underline{\varepsilon} = \mathbf{0}$. Expression (4.13) states the positive-definiteness of matrix $\mathbf{C}$. In other words, matrix $\mathbf{C}$ is positive-definite if and only if $\left(\mathbf{C}\underline{\varepsilon}\right) \cdot \underline{\varepsilon} > 0$, for all $\underline{\varepsilon}$, such that $\underline{\varepsilon} \neq \mathbf{0}$.

4.3.3 Constraints on elasticity parameters

To formulate the conditions of positive-definiteness of the elasticity matrix, we can use either of the following theorems of linear algebra.[9]

---

[8] Readers interested in a more detailed description might refer to Musgrave, M.J.P., (1990) On the constraints of positive-definite strain energy in anisotropic media: Q.J.Mech.appl.Math., **43**, Part 4, 605 – 621.

[9] Readers interested in proofs of Theorem 4.2 and Theorem 4.3 might refer to Ayres, F., (1962) Matrices: Schaum's Outlines, McGraw-Hill, Inc., p. 142, and to Morse P.M., and Feshbach H., (1953) Methods of theoretical physics: McGraw-Hill, Inc., pp. 771 – 774, respectively.

Note that — in view of the fact that every symmetric matrix can be diagonalized — Theorem 4.3 follows from Theorem 4.2.

Theorem 4.2 *A real symmetric matrix is positive-definite if and only if the determinants of all its leading principal minors, including the determinant of the matrix itself, are positive.*

Theorem 4.3 *A real symmetric matrix is positive-definite if and only if all its eigenvalues are positive.*

[10]Since matrix (4.10) is symmetric, the stability conditions can be conveniently formulated based on Theorems 4.2 and 4.3, as shown in Exercises 5.3 and 5.12, respectively. Among these conditions, we find that

$$C_{mm} > 0, \qquad m \in \{1, \ldots, 6\}, \tag{4.14}$$

which implies that all the main-diagonal entries of the elasticity matrix must be positive, as shown in Exercise 4.5.

Stability conditions cannot be violated. However, as shown in Exercise 5.12, interesting and, perhaps, nonintuitive results are allowable within the stability conditions.

4.4 System of equations for elastic continua

4.4.1 Elastic continua

In order to state a complete system of equations describing behaviour of our continua, we note that linearly elastic continua are defined by stress-strain equations (3.1), namely,

$$\sigma_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl}, \qquad i,j \in \{1,2,3\}, \tag{4.15}$$

where, in view of equations (3.3), (3.4) and (4.5), we require

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{klij}, \qquad i,j,k,l \in \{1,2,3\}. \tag{4.16}$$

We recall that symmetries (4.16) result from definition (1.15), which implies $\varepsilon_{kl} = \varepsilon_{lk}$, as well as from the balance of angular momentum, stated in expression (2.43), namely,

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{V(t)}\left(\mathbf{x}\times\rho\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\right)\mathrm{d}V = \iint\limits_{S(t)}(\mathbf{x}\times\mathbf{T})\,\mathrm{d}S + \iiint\limits_{V(t)}(\mathbf{x}\times\mathbf{f})\,\mathrm{d}V,$$

[10]Readers interested in the expressions for the stability conditions for particular continua may derive them from the corresponding stiffness matrices, shown in Chapter 5, or refer to Fedorov, F.I., (1968) Theory of elastic waves in crystals: Plenum Press, New York, p. 16 and p. 33.

which implies $\sigma_{ij} = \sigma_{ji}$. Symmetries (4.16) also result from the existence of the strain-energy function that is given by expression (4.3), namely,

$$W(\varepsilon) = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl}\varepsilon_{ij}, \tag{4.17}$$

and must satisfy equation (4.1), namely,

$$\frac{\partial W(\varepsilon)}{\partial \varepsilon_{ij}} = \sigma_{ij}, \qquad i,j \in \{1,2,3\}. \tag{4.18}$$

As expected, this formulation is consistent with the following statement from the classic book of Augustus Edward Hugh Love (1892) "A treatise on the mathematical theory of elasticity".

> When a body is slightly strained by gradual application of a load, and the temperature remains constant, the stress components are linear functions of the strain components [equations (4.15)], and they are also partial differential coefficients of a function W of the strain components [equations (4.18)]. The strain-energy function, W, is therefore a homogeneous quadratic function of the strain components [equations (4.17)].

4.4.2 Governing equations

We can now show that the behaviour of the continuum discussed in this book is governed by a system of ten equations and ten unknowns. The unknowns of this system are ρ, u_1, u_2, u_3, σ_{11}, σ_{12}, σ_{13}, σ_{22}, σ_{23}, σ_{33}, while a given continuum is defined in terms of twenty-one elasticity parameters, $C_{mn} = C_{nm}$, where $m,n \in \{1,\ldots,6\}$.

Note that four among ten equations result from the fundamental principles, which are given by the conservation of mass, stated in equation (2.6), namely,

$$\int_V \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v})\right] \mathrm{d}V = 0,$$

and the balance of linear momentum, stated in equation (2.16), namely,

$$\iiint_{V(t)} \rho \frac{\mathrm{d}^2\mathbf{u}}{\mathrm{d}t^2} \mathrm{d}V = \iint_{S(t)} \mathbf{T}\,\mathrm{d}S + \iiint_{V(t)} \mathbf{f}\,\mathrm{d}V.$$

The remaining six equations are constitutive equations, which provide a phenomenological description of actual materials.

Explicitly, we can write this system as

$$\left\{\begin{array}{l}
\frac{\partial}{\partial t}\rho=-\left[\frac{\partial}{\partial x_1}\left(\rho\frac{\partial}{\partial t}u_1\right)+\frac{\partial}{\partial x_2}\left(\rho\frac{\partial}{\partial t}u_2\right)+\frac{\partial}{\partial x_3}\left(\rho\frac{\partial}{\partial t}u_3\right)\right]\\
\rho\frac{\partial^2}{\partial t^2}u_1=\frac{\partial}{\partial x_1}\sigma_{11}+\frac{\partial}{\partial x_2}\sigma_{12}+\frac{\partial}{\partial x_3}\sigma_{13}\\
\rho\frac{\partial^2}{\partial t^2}u_2=\frac{\partial}{\partial x_1}\sigma_{12}+\frac{\partial}{\partial x_2}\sigma_{22}+\frac{\partial}{\partial x_3}\sigma_{23}\\
\rho\frac{\partial^2}{\partial t^2}u_3=\frac{\partial}{\partial x_1}\sigma_{13}+\frac{\partial}{\partial x_2}\sigma_{23}+\frac{\partial}{\partial x_3}\sigma_{33}\\
\sigma_{11}=C_{11}\varepsilon_{11}+C_{12}\varepsilon_{22}+C_{13}\varepsilon_{33}+2C_{14}\varepsilon_{23}+2C_{15}\varepsilon_{13}+2C_{16}\varepsilon_{12}\\
\sigma_{22}=C_{12}\varepsilon_{11}+C_{22}\varepsilon_{22}+C_{23}\varepsilon_{33}+2C_{24}\varepsilon_{23}+2C_{25}\varepsilon_{13}+2C_{26}\varepsilon_{12}\\
\sigma_{33}=C_{13}\varepsilon_{11}+C_{23}\varepsilon_{22}+C_{33}\varepsilon_{33}+2C_{34}\varepsilon_{23}+2C_{35}\varepsilon_{13}+2C_{36}\varepsilon_{12}\\
\sigma_{23}=C_{14}\varepsilon_{11}+C_{24}\varepsilon_{22}+C_{34}\varepsilon_{33}+2C_{44}\varepsilon_{23}+2C_{45}\varepsilon_{13}+2C_{46}\varepsilon_{12}\\
\sigma_{13}=C_{15}\varepsilon_{11}+C_{25}\varepsilon_{22}+C_{35}\varepsilon_{33}+2C_{45}\varepsilon_{23}+2C_{55}\varepsilon_{13}+2C_{56}\varepsilon_{12}\\
\sigma_{12}=C_{16}\varepsilon_{11}+C_{26}\varepsilon_{22}+C_{36}\varepsilon_{33}+2C_{46}\varepsilon_{23}+2C_{56}\varepsilon_{13}+2C_{66}\varepsilon_{12}
\end{array}\right., \tag{4.19}$$

where, by definition (1.15), $\varepsilon_{ij}=\left(\partial u_i/\partial x_j+\partial u_j/\partial x_i\right)/2$. The first equation is equation of continuity (2.51), which results from the conservation of mass. The following three equations are Cauchy's equations of motion (2.52), (2.53) and (2.54), which result from the balance of linear momentum. The last six equations are stress-strain equations (4.11), which contain the elasticity parameters that describe a given continuum.

Note that, invoking system (4.19) to study actual materials, we can consider directly only C_{mn} and ρ as properties of a continuum that represents a given material. Indirectly, the values of C_{mn} and ρ can indicate other properties, such as layering and fractures.

In a properly chosen coordinate system, which we will discuss in Section 5.6.2, different materials exhibit different values of the elasticity parameters. These values are determined experimentally and characterize a given material. Often, the goal of our seismological studies is to determine the values of the elasticity parameters and mass density of the subsurface based on the theoretical formulation and the experimental data.

Note that the last six equations of system (4.19) can be substituted into the second, the third and the fourth equations to obtain a system of four partial differential equations for four unknowns, namely, $\rho(\mathbf{x},t)$, $u_1(\mathbf{x},t)$, $u_2(\mathbf{x},t)$, $u_3(\mathbf{x},t)$ in the position variables, $\mathbf{x}=[x_1,x_2,x_3]$, and the time variable, t. For a one-dimensional case, a system of partial differential equations is exemplified in Exercise 3.1. Also, this substitution of stress-strain equations into Cauchy's equations of motion is used extensively in Chapters 6 and 7 to formulate equations of motions specifically for elastic continua.

Closing remarks

For linearly elastic continua, stress is a linear function of strain, which depends on twenty-one elasticity parameters, as shown in expressions (4.11). Furthermore, these elastic continua possess strain energy, which is expressed as a quadratic function of strain, shown in expression (4.3). Thus, for instance, doubling the strain within the continuum doubles the stress within it, while it quadruples the energy stored within.

In our studies, we assume that the elasticity parameters have no temperature dependence, which is tantamount to our assuming in formulating system (4.19) that the process of deformation is isothermal. In other words, this process occurs at a constant temperature. Due to low thermal conductivity of most subsurface materials, we could argue that seismic wave propagation is better approximated by an adiabatic process, where no heat enters or leaves the element of volume. However, we choose the simplicity of the isothermal approach since the difference in experimental determination of elasticity parameters between the isothermal and adiabatic approaches is only of the order of one percent.[11]

Our formulation of elasticity parameters is rooted in the mathematical concept of a continuum. The continuum formulation of these parameters is also consistent with that of condensed-matter physics where, according to common physical knowledge, materials are composed of nuclei and electrons. Physically, the elasticity parameters are functions of the interactions among the nuclei and electrons within a material. They can be calculated by considering the total energy associated with the changes of volume and shape, which result from forces acting on every atom.[12]

A given elastic continuum can possess particular symmetries, which further reduce the number of independent elasticity parameters required to

[11]Interested readers might refer to Brekhovskikh, L.M., and Goncharov, V., (1982/1994) Mechanics of continua and wave dynamics: Springer-Verlag., pp. 45 – 47, to Fung, Y.C., (1977) A first course in continuum mechanics: Prentice-Hall, Inc., pp. 169 – 170, to Grant, F.S., and West, G.F., (1965) Interpretation theory in applied geophysics: McGraw-Hill Book Co., pp. 30 – 31, and to Timoshenko, S.P., and Goodier, J.N., (1934/1987) Theory of elasticity: McGraw-Hill Publishing Company, p. 244.

[12]Readers interested in certain relationships between the continuum formulations and the atomic scale associated with the study of condensed-matter physics might refer to Aoki, H., Syono, Y., and Hemley, R.J., (editors), (2000) Physics meets mineralogy: Condensed-matter physics in geosciences: Cambridge University Press.

describe it. Such symmetries are discussed in Chapter 5.

Exercises

Exercise 4.1 *Using the one-dimensional case illustrated by a spring constant, k, show that for, elastic continua, strain energy is equal to the area below the graph of $F = kx$.*

Solution 4.1 *Following the definition of work and energy in a conservative system, we can write*

$$W = \int \mathbf{F} \cdot \mathrm{d}\mathbf{x},$$

where $\mathbf{F}$ *denotes force and* $\mathrm{d}\mathbf{x}$ *denotes an element of displacement. The one-dimensional stress-strain equation can be written as $F = kx$, where k is an elasticity parameter, commonly known as the spring constant. Thus,*

$$W = \int_0^{\Delta x} kx \,\mathrm{d}x = \frac{1}{2} k \left(\Delta x\right)^2,$$

where $x = 0$ corresponds to the unstrained state while $x = \Delta x$ corresponds to the strained state. This is equal to the triangular area below the straight line, kx, spanned between $x = 0$ and $x = \Delta x$.

Notation 4.1 *In Exercise 4.2, for convenience, we denote the strain-tensor components using single subscripts.*

Exercise 4.2 [13]*Consider the strain-energy function to be a second-degree polynomial given by*

$$W = C_0 + \sum_{n=1}^{6} C_n \varepsilon_n + \frac{1}{2} \sum_{n=1}^{6} \sum_{m=1}^{6} C_{nm} \varepsilon_n \varepsilon_m, \tag{4.20}$$

where ε_l is an entry of matrix $\underline{\varepsilon}$, given in equation (3.6). Show explicitly that the first two terms vanish and, hence, W is homogeneous of degree 2 in the strain-tensor components. Note that since tensor ε_{ij} is symmetric, we only need one index, $i = 1, 2, \ldots, 6$, to represent all components.

[13] See also Section 4.1.

Solution 4.2 *Expression (4.20) can be explicitly written as*

$$
\begin{aligned}
W &= C_0 \\
&+ C_1\varepsilon_1 + C_2\varepsilon_2 + C_3\varepsilon_3 + C_4\varepsilon_4 + C_5\varepsilon_5 + C_6\varepsilon_6 \\
&+ \frac{1}{2}(C_{11}\varepsilon_1^2 + C_{21}\varepsilon_2\varepsilon_1 + C_{31}\varepsilon_3\varepsilon_1 + C_{41}\varepsilon_4\varepsilon_1 + C_{51}\varepsilon_5\varepsilon_1 + C_{61}\varepsilon_6\varepsilon_1 \\
&+ C_{12}\varepsilon_1\varepsilon_2 + C_{22}\varepsilon_2^2 + C_{32}\varepsilon_3\varepsilon_2 + C_{42}\varepsilon_4\varepsilon_2 + C_{52}\varepsilon_5\varepsilon_2 + C_{62}\varepsilon_6\varepsilon_2 \\
&+ C_{13}\varepsilon_1\varepsilon_3 + C_{23}\varepsilon_2\varepsilon_3 + C_{33}\varepsilon_3^2 + C_{43}\varepsilon_4\varepsilon_3 + C_{35}\varepsilon_5\varepsilon_3 + C_{63}\varepsilon_6\varepsilon_3 \\
&+ C_{14}\varepsilon_1\varepsilon_4 + C_{24}\varepsilon_2\varepsilon_4 + C_{34}\varepsilon_3\varepsilon_4 + C_{44}\varepsilon_4^2 + C_{54}\varepsilon_5\varepsilon_4 + C_{64}\varepsilon_6\varepsilon_4 \\
&+ C_{15}\varepsilon_1\varepsilon_5 + C_{25}\varepsilon_2\varepsilon_5 + C_{35}\varepsilon_3\varepsilon_5 + C_{45}\varepsilon_4\varepsilon_5 + C_{55}\varepsilon_5^2 + C_{65}\varepsilon_6\varepsilon_5 \\
&+ C_{16}\varepsilon_1\varepsilon_6 + C_{26}\varepsilon_2\varepsilon_6 + C_{36}\varepsilon_3\varepsilon_6 + C_{46}\varepsilon_4\varepsilon_6 + C_{56}\varepsilon_5\varepsilon_6 + C_{66}\varepsilon_6^2).
\end{aligned}
$$

We assume that W vanishes for the unstrained state. Thus, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = \varepsilon_5 = \varepsilon_6 = 0$, implies $W = 0$; consequently, $C_0 = 0$. Also, following expression (4.1) we require that $\sigma_m = \partial W / \partial \varepsilon_m$. Consider, for instance, $m = 5$; we can specifically write

$$
\begin{aligned}
\sigma_5 &= \frac{\partial W}{\partial \varepsilon_5} \\
&= C_5 + \frac{1}{2}[(C_{15} + C_{51})\,\varepsilon_1 + (C_{25} + C_{52})\,\varepsilon_2 + (C_{35} + C_{53})\,\varepsilon_3 \\
&+ (C_{45} + C_{54})\,\varepsilon_4 + 2C_{55}\varepsilon_5 + (C_{56} + C_{65})\,\varepsilon_6].
\end{aligned}
$$

No strain implies no stress and, hence, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = \varepsilon_5 = \varepsilon_6 = 0 \Longrightarrow \sigma_5 = 0$. Thus, it follows that $C_5 = 0$. Analogously, considering σ_1, σ_2, σ_3, σ_4 and σ_6, we obtain $C_1 = C_2 = C_3 = C_4 = C_6 = 0$. Thus,

$$
\begin{aligned}
W(\varepsilon_m) &= \frac{1}{2}(C_{11}\varepsilon_1^2 + C_{21}\varepsilon_2\varepsilon_1 + C_{31}\varepsilon_3\varepsilon_1 + C_{41}\varepsilon_4\varepsilon_1 + C_{51}\varepsilon_5\varepsilon_1 + C_{61}\varepsilon_6\varepsilon_1 \\
&+ C_{12}\varepsilon_1\varepsilon_2 + C_{22}\varepsilon_2^2 + C_{32}\varepsilon_3\varepsilon_2 + C_{42}\varepsilon_4\varepsilon_2 + C_{52}\varepsilon_5\varepsilon_2 + C_{62}\varepsilon_6\varepsilon_2 \\
&+ C_{13}\varepsilon_1\varepsilon_3 + C_{23}\varepsilon_2\varepsilon_3 + C_{33}\varepsilon_3^2 + C_{43}\varepsilon_4\varepsilon_3 + C_{35}\varepsilon_5\varepsilon_3 + C_{63}\varepsilon_6\varepsilon_3 \\
&+ C_{14}\varepsilon_1\varepsilon_4 + C_{24}\varepsilon_2\varepsilon_4 + C_{34}\varepsilon_3\varepsilon_4 + C_{44}\varepsilon_4^2 + C_{54}\varepsilon_5\varepsilon_4 + C_{64}\varepsilon_6\varepsilon_4 \\
&+ C_{15}\varepsilon_1\varepsilon_5 + C_{25}\varepsilon_2\varepsilon_5 + C_{35}\varepsilon_3\varepsilon_5 + C_{45}\varepsilon_4\varepsilon_5 + C_{55}\varepsilon_5^2 + C_{65}\varepsilon_6\varepsilon_5 \\
&+ C_{16}\varepsilon_1\varepsilon_6 + C_{26}\varepsilon_2\varepsilon_6 + C_{36}\varepsilon_3\varepsilon_6 + C_{46}\varepsilon_4\varepsilon_6 + C_{56}\varepsilon_5\varepsilon_6 + C_{66}\varepsilon_6^2),
\end{aligned}
\tag{4.21}
$$

which is a homogeneous function of degree 2 in the ε_m, as required.

Remark 4.1 *Examining expression (4.21), we observe that*

$$W(c\varepsilon_m) = c^2 W(\varepsilon_m),$$

where c is a real number. Hence, in view of Definition A.1 stated in Appendix A, W is homogeneous of degree 2 in the ε_m.

Exercise 4.3 [14] *Using the property of the strain-energy function, W, stated in expression (4.1), and in view of W being homogeneous of degree 2 in the* ε_{ij}, *show that*

$$W = \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij} \varepsilon_{ij}. \tag{4.22}$$

Solution 4.3 *Since W is homogeneous of degree 2 in the* ε_{ij}, *by Theorem A.1 stated in Appendix A, we can write*

$$\sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial W}{\partial \varepsilon_{ij}} \varepsilon_{ij} = 2W.$$

In view of expression (4.1), we can rewrite the above expression as

$$\sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij} \varepsilon_{ij} = 2W,$$

which immediately yields expression (4.22), as required.

Exercise 4.4 *Assuming that* $c_{ijkl} = c_{klij}$, *derive expression (4.1), namely,*

$$\frac{\partial W}{\partial \varepsilon_{ij}} = \sigma_{ij}, \qquad i, j \in \{1, 2, 3\},$$

directly from expression (4.22).

Solution 4.4 *Differentiating expression (4.22) with respect to a particular strain-tensor component* ε_{kl}, *and recalling that stress is a function of strain, we obtain*

$$\begin{aligned}
\frac{\partial W}{\partial \varepsilon_{kl}} &= \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} \varepsilon_{ij} + \sigma_{ij} \frac{\partial \varepsilon_{ij}}{\partial \varepsilon_{kl}} \right) \\
&= \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} \varepsilon_{ij} + \sigma_{ij} \delta_{ik} \delta_{jl} \right), \qquad k, l \in \{1, 2, 3\}.
\end{aligned}$$

[14] See also Section 4.1.

Using stress-strain equations (3.1), and recalling that c_{ijkl} are independent of strain, we can write

$$\frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} = c_{ijkl}, \qquad i,j,k,l \in \{1,2,3\}.$$

Consequently, using the fact that $c_{ijkl} = c_{klij}$, we obtain

$$\begin{aligned}\frac{\partial W}{\partial \varepsilon_{kl}} &= \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\left(c_{ijkl}\varepsilon_{ij} + \sigma_{ij}\delta_{ik}\delta_{jl}\right) \\ &= \frac{1}{2}\left(\sum_{i=1}^{3}\sum_{j=1}^{3} c_{ijkl}\varepsilon_{ij} + \sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\delta_{ik}\delta_{jl}\right), \qquad k,l \in \{1,2,3\}.\end{aligned}$$

Again, in view of equations (3.1) and (4.5) as well as using the properties of Kronecker's delta, we obtain

$$\frac{\partial W}{\partial \varepsilon_{kl}} = \frac{1}{2}\left(\sigma_{kl} + \sigma_{kl}\right) = \sigma_{kl}, \qquad k,l \in \{1,2,3\},$$

where, in view of the arbitrariness of the subscript symbol, we obtain expression (4.1), as required.

Exercise 4.5 [15] *Using expression (4.12), justify that the main-diagonal entries of the elasticity matrix must be always positive.*

Solution 4.5 *Consider expression (4.12). In view of equations (4.11), we can write*

$$W = \frac{1}{2}\left(\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix}\right) \cdot \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix}.$$

Let the strain matrix, $\underline{\varepsilon}$, have only one nonzero entry. For instance, we can

[15] See also Section 4.3.3.

write

$$W = \frac{1}{2}\left(\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{1}{2}\left[C_{14}, C_{24}, C_{34}, C_{44}, C_{45}, C_{46}\right] \cdot [0,0,0,1,0,0]^T = \frac{1}{2}C_{44}.$$

Similarly, for all other single, nonzero entries, $W = C_{ii}/2$. *Hence, the positive value of the strain-energy function for all possible nonzero strains requires* $C_{ii} > 0$.

Chapter 5

Material symmetry

> Symmetry is a vast subject, significant in art and nature. Mathematics lies at its root, and it would be hard to find a better one on which to demonstrate the working of the mathematical intellect.
>
> *Hermann Weyl (1952) Symmetry*

Preliminary remarks

Materials can possess certain symmetries. In the context of our studies, this means that we can measure a property of a material in several different orientations of the coordinate system and obtain the same result each time. In other words, we are unable to detect the transformations of the reference coordinate system by mechanical experiments. This invariance to the orientation of the coordinate system is called material symmetry. In a properly chosen coordinate system, the form of the elasticity matrix allows us to recognize the symmetry of this continuum. This symmetry is indicative of the properties exhibited by the material represented by this continuum.

We begin this chapter with the formulation of transformations of the coordinate system and the effect of these transformations on the stress-strain equations. Then we formulate the condition that allows us to obtain the elasticity matrix of a continuum that is invariant under a given transformation of coordinates. We complete this chapter by studying several such continua.

5.1 Orthogonal transformations

5.1.1 Transformation matrix

To study material symmetries, we wish to use transformation of an orthonormal coordinate system in the $x_1x_2x_3$-space. A change of an orthonormal coordinate system in our three-dimensional space is given by

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x}, \tag{5.1}$$

where $\mathbf{x} = [x_1, x_2, x_3]^T$ and $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \hat{x}_3]^T$ are the original and transformed coordinate systems, respectively, and $\mathbf{A}$ is the transformation matrix. Equation (5.1) is the matrix form of equations (1.35), shown in Exercise 1.4.

We are interested specifically in distance-preserving transformations, namely, rotations and reflections. In other words, these transformations allow us to change the orientation of the continuum without deforming it. Such transformations are represented by orthogonal matrices, that is, by square matrices given by

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{5.2}$$

that satisfy the orthogonality condition, namely, $\mathbf{A}\mathbf{A}^T = \mathbf{I}$, which is equivalent to $\mathbf{A}^T = \mathbf{A}^{-1}$.

Note that the determinants of these transformation matrices are the Jacobians of the coordinate transformations. This is illustrated in Exercise 5.1.

5.1.2 Symmetry group

Expressing the elasticity matrix of a given continuum in a conveniently chosen orthonormal coordinate system allows us to recognize the material symmetries of that continuum, as discussed in Sections 5.5 – 5.10. In other words, we can recognize the transformations that belong to the symmetry group of that continuum, which can be stated by the following definition.

Definition 5.1 *The set of all orthogonal transformations given by matrices* $\mathbf{A}$ *to which the elastic properties of a given continuum are invariant is called the symmetry group of that continuum.*

If the elastic properties are invariant under orthogonal transformations given by matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, they are also invariant to product $\mathbf{A}_1\mathbf{A}_2$. Furthermore, if these properties are invariant to $\mathbf{A}$, they are also invariant to $\mathbf{A}^{-1}$. This is the reason for our invoking the notion of a group in Definition 5.1.[1]

5.2 Transformation of coordinates

Introductory comments

Recall that the properties of our continuum are formulated in terms of the stress-strain equations. To investigate the material symmetries of a given continuum, we study the stress-strain equations in the context of the orthogonal transformations of the orthonormal coordinate system.

5.2.1 Transformation of stress-tensor components

The components of the stress tensor expressed as a 3×3 matrix transform according to

$$\hat{\sigma} = \mathbf{A}\sigma\mathbf{A}^T, \tag{5.3}$$

where $\mathbf{A}$ stands for matrix (5.2) and the accent symbolizes the transformed entity. This is the matrix form of transformation (2.55), which is proven in Exercise 2.3.

Following matrix (2.18), the stress-tensor components are given by a square matrix

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}, \tag{5.4}$$

which, in view of Theorem 2.1, is symmetric.

Note that since a second-rank tensor has two indices, it is convenient to write it as a matrix, even though, tensors and matrices are distinct mathematical entities. Herein, we use the fact that under the orthogonal transformations, the entries of a matrix behave like the components of a second-rank tensor.

Recall that stress-strain equations (4.11) involve stress-tensor components as a single-column matrix, $\underline{\sigma}$, namely,

$$\underline{\sigma} = [\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{23}, \sigma_{13}, \sigma_{12}]^T. \tag{5.5}$$

[1]Readers interested in physical aspects of the group theory, which is the study of invariants and symmetries, might refer to Arfken, G.B, and Weber, H.J., (2001) Mathematical methods for physicists (5*th* edition): Harcourt/Academic Press, pp. 237 – 301.

We wish to obtain a transformation equation for $\underline{\sigma}$ that is equivalent to equation (5.3). Transformation (5.3) is linear; hence, it can be rewritten as a multiplication of $\underline{\sigma}$ by a matrix. In other words, we can write

$$\underline{\hat{\sigma}} = \underline{\mathbf{A}}\,\underline{\sigma}, \tag{5.6}$$

where $\underline{\mathbf{A}}$ is a 6×6 transformation matrix. To find the entries of $\underline{\mathbf{A}}$, we substitute the elements of the standard basis of the space of symmetric 3×3 matrices for σ.[2]

Consider the first element of the basis, namely,

$$\sigma = \mathbf{E}_{11},$$

where $\mathbf{E}_{ij}$ denotes the matrix with unity in the position (i, j) and with zeros elsewhere. Thus, in this case,

$$\sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.7}$$

Using Kronecker's delta, we can write the entries of matrix (5.7) as

$$\sigma_{ij} = \delta_{i1}\delta_{j1}, \qquad i, j \in \{1, 2, 3\}.$$

Then using equation (5.3), the entries of $\hat{\sigma}$ can be computed as

$$\begin{aligned} \hat{\sigma}_{kl} &= \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ki}\sigma_{ij}A^T_{jl} = \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ki}\delta_{i1}\delta_{j1}A_{lj} \\ &= A_{k1}A_{l1}, \qquad k, l \in \{1, 2, 3\}. \end{aligned}$$

In view of expression (5.5), taking $(k, l) = (1, 1)$, $(2, 2)$, $(3, 3)$, $(2, 3)$, $(1, 3)$ and $(1, 2)$, we obtain

$$\underline{\hat{\sigma}} = \begin{bmatrix} A_{11}A_{11} \\ A_{21}A_{21} \\ A_{31}A_{31} \\ A_{21}A_{31} \\ A_{11}A_{31} \\ A_{11}A_{21} \end{bmatrix}.$$

[2] Readers interested in the underlying aspects of linear algebra might refer to Anton, H., (1973) Elementary linear algebra: John Wiley & Sons, pp. 237 – 238, and to Ayres, F., (1962) Matrices: Schaum's Outlines, McGraw-Hill, p. 88 and p. 94.

Since, herein,

$$\underline{\sigma} = [1, 0, 0, 0, 0, 0]^T,$$

in view of equation (5.6), $\underline{\hat{\sigma}}$ is the first column of $\underline{\mathbf{A}}$. Analogously, we can compute the second and the third columns of $\underline{\mathbf{A}}$ by considering $\sigma = \mathbf{E}_{22}$ and $\sigma = \mathbf{E}_{33}$, respectively.

To find the fourth column of $\underline{\mathbf{A}}$, we use

$$\sigma = \mathbf{E}_{23} + \mathbf{E}_{32}.$$

Thus, in this case,

$$\sigma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{5.8}$$

Using Kronecker's delta, we can write the entries of matrix (5.8), as

$$\sigma_{ij} = \delta_{i2}\delta_{j3} + \delta_{i3}\delta_{j2}, \qquad i, j \in \{1, 2, 3\}.$$

Then, using equation (5.3), the entries of $\hat{\sigma}$ can be computed as

$$\begin{aligned}
\hat{\sigma}_{kl} &= \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ki}\sigma_{ij}A^T_{jl} = \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ki}\delta_{i2}\delta_{j3}A_{lj} + \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ki}\delta_{i3}\delta_{j2}A_{lj} \\
&= A_{k2}A_{l3} + A_{k3}A_{l2}, \qquad k, l \in \{1, 2, 3\}.
\end{aligned}$$

In view of expression (5.5), taking $(k, l) = (1, 1)$, $(2, 2)$, $(3, 3)$, $(2, 3)$, $(1, 3)$ and $(1, 2)$, we obtain

$$\underline{\hat{\sigma}} = \begin{bmatrix} 2A_{12}A_{13} \\ 2A_{22}A_{23} \\ 2A_{32}A_{33} \\ A_{22}A_{33}+A_{23}A_{32} \\ A_{12}A_{33}+A_{13}A_{32} \\ A_{12}A_{23}+A_{13}A_{22} \end{bmatrix},$$

which is the fourth column of $\underline{\mathbf{A}}$. Analogously, we can compute the fifth and the sixth columns of $\underline{\mathbf{A}}$ by considering $\sigma = \mathbf{E}_{13} + \mathbf{E}_{31}$ and $\sigma = \mathbf{E}_{12} + \mathbf{E}_{21}$, respectively.

Now, putting together the six columns of $\underline{\mathbf{A}}$, we obtain

$$\underline{\mathbf{A}} = \tag{5.9}$$

$$\begin{bmatrix}
A_{11}A_{11} & A_{12}A_{12} & A_{13}A_{13} & 2A_{12}A_{13} & 2A_{11}A_{13} & 2A_{11}A_{12} \\
A_{21}A_{21} & A_{22}A_{22} & A_{23}A_{23} & 2A_{22}A_{23} & 2A_{21}A_{23} & 2A_{21}A_{22} \\
A_{31}A_{31} & A_{32}A_{32} & A_{33}A_{33} & 2A_{32}A_{33} & 2A_{31}A_{33} & 2A_{31}A_{32} \\
A_{21}A_{31} & A_{22}A_{32} & A_{23}A_{33} & A_{22}A_{33}+A_{23}A_{32} & A_{21}A_{33}+A_{23}A_{31} & A_{21}A_{32}+A_{22}A_{31} \\
A_{11}A_{31} & A_{12}A_{32} & A_{13}A_{33} & A_{12}A_{33}+A_{13}A_{32} & A_{11}A_{33}+A_{13}A_{31} & A_{11}A_{32}+A_{12}A_{31} \\
A_{11}A_{21} & A_{12}A_{22} & A_{13}A_{23} & A_{12}A_{23}+A_{13}A_{22} & A_{11}A_{23}+A_{13}A_{21} & A_{11}A_{22}+A_{12}A_{21}
\end{bmatrix},$$

which is the desired transformation matrix for the stress-tensor components given by matrix (5.5). Thus, given transformation matrix (5.2), whose entries are A_{ij}, we can immediately write the corresponding $\underline{\mathbf{A}}$ using matrix (5.9). Matrix (5.9) was also formulated by Bond (1943).

5.2.2 Transformation of strain-tensor components

The components of the strain-tensor expressed as a 3×3 matrix transform according to

$$\hat{\varepsilon} = \mathbf{A}\varepsilon\mathbf{A}^T, \tag{5.10}$$

where $\mathbf{A}$ stands for matrix (5.2) and the accent symbolizes the transformed entity. Equation (5.10) is the matrix form of equation (1.37), which is shown in Exercise 1.4.

In equation (5.10), the strain-tensor components are considered as a square matrix

$$\varepsilon = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{21} \\ \varepsilon_{13} & \varepsilon_{21} & \varepsilon_{33} \end{bmatrix},$$

whose symmetry results from definition (1.15).

We wish to rewrite the strain-tensor components as a single-column matrix in a manner similar to that shown in Section 5.2.1. As shown in stress-strain equations (4.11), the single column matrix, $\underline{\varepsilon}$, is formulated with factors of 2, namely,

$$\underline{\varepsilon} = [\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, 2\varepsilon_{23}, 2\varepsilon_{13}, 2\varepsilon_{12}]^T. \tag{5.11}$$

Hence, the corresponding transformation matrix differs from expression (5.6). To account for the factors of 2, we can write

$$\hat{\underline{\varepsilon}} = \mathbf{F}\underline{\mathbf{A}}\mathbf{F}^{-1}\underline{\varepsilon}, \tag{5.12}$$

where $\underline{\mathbf{A}}$ is matrix (5.9) and

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

Thus, the transformation matrix for the strain-tensor components, given by matrix (5.11), can be explicitly written as

$$\mathbf{M}_{\mathbf{A}} = \mathbf{F}\underline{\mathbf{A}}\,\mathbf{F}^{-1} = \tag{5.13}$$

$$\begin{bmatrix} A_{11}A_{11} & A_{12}A_{12} & A_{13}A_{13} & A_{12}A_{13} & A_{11}A_{13} & A_{11}A_{12} \\ A_{21}A_{21} & A_{22}A_{22} & A_{23}A_{23} & A_{22}A_{23} & A_{21}A_{23} & A_{21}A_{22} \\ A_{31}A_{31} & A_{32}A_{32} & A_{33}A_{33} & A_{32}A_{33} & A_{31}A_{33} & A_{31}A_{32} \\ 2A_{21}A_{31} & 2A_{22}A_{32} & 2A_{23}A_{33} & A_{22}A_{33}+A_{23}A_{32} & A_{21}A_{33}+A_{23}A_{31} & A_{21}A_{32}+A_{22}A_{31} \\ 2A_{11}A_{31} & 2A_{12}A_{32} & 2A_{13}A_{33} & A_{12}A_{33}+A_{13}A_{32} & A_{11}A_{33}+A_{13}A_{31} & A_{11}A_{32}+A_{12}A_{31} \\ 2A_{11}A_{21} & 2A_{12}A_{22} & 2A_{13}A_{23} & A_{12}A_{23}+A_{13}A_{22} & A_{11}A_{23}+A_{13}A_{21} & A_{11}A_{22}+A_{12}A_{21} \end{bmatrix}$$

and expression (5.12) can be restated as

$$\underline{\hat{\varepsilon}} = \mathbf{M}_{\mathbf{A}}\,\underline{\varepsilon}. \tag{5.14}$$

Consequently, given transformation matrix (5.2), whose entries are A_{ij}, we can immediately write the corresponding $\mathbf{M}_{\mathbf{A}}$ using matrix (5.13).

5.2.3 Stress-strain equations in transformed coordinates

Now, having formulated $\underline{\hat{\sigma}}$ and $\underline{\hat{\varepsilon}}$, which are given by expressions (5.6) and (5.14), respectively, we can formally write the stress-strain equations in transformed coordinates as

$$\underline{\hat{\sigma}} = \hat{\mathbf{C}}\underline{\hat{\varepsilon}}.$$

Explicitly, we can write these equations as

$$\begin{bmatrix} \hat{\sigma}_{11} \\ \hat{\sigma}_{22} \\ \hat{\sigma}_{33} \\ \hat{\sigma}_{23} \\ \hat{\sigma}_{13} \\ \hat{\sigma}_{12} \end{bmatrix} = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & \hat{C}_{14} & \hat{C}_{15} & \hat{C}_{16} \\ \hat{C}_{12} & \hat{C}_{22} & \hat{C}_{23} & \hat{C}_{24} & \hat{C}_{25} & \hat{C}_{26} \\ \hat{C}_{13} & \hat{C}_{23} & \hat{C}_{33} & \hat{C}_{34} & \hat{C}_{35} & \hat{C}_{36} \\ \hat{C}_{14} & \hat{C}_{24} & \hat{C}_{34} & \hat{C}_{44} & \hat{C}_{45} & \hat{C}_{46} \\ \hat{C}_{15} & \hat{C}_{25} & \hat{C}_{35} & \hat{C}_{45} & \hat{C}_{55} & \hat{C}_{56} \\ \hat{C}_{16} & \hat{C}_{26} & \hat{C}_{36} & \hat{C}_{46} & \hat{C}_{56} & \hat{C}_{66} \end{bmatrix} \begin{bmatrix} \hat{\varepsilon}_{11} \\ \hat{\varepsilon}_{22} \\ \hat{\varepsilon}_{33} \\ 2\hat{\varepsilon}_{23} \\ 2\hat{\varepsilon}_{13} \\ 2\hat{\varepsilon}_{12} \end{bmatrix}, \tag{5.15}$$

where, as discussed in Chapter 4, the elasticity matrix is symmetric due to the strain-energy function.

Recall that a continuum is formulated in terms of the stress-strain equations. Consequently, an examination of equations (4.11) and (5.15) leads to the following definition.

Definition 5.2 *The elastic properties of a continuum are invariant under an orthogonal transformation if* $\mathbf{C} = \hat{\mathbf{C}}$, *in other words, if the transformed elasticity matrix is identical to the original elasticity matrix.*

Thus, material symmetry is exhibited by a change of the reference coordinate system that is undetectable by any mechanical experiment.

5.3 Condition for material symmetry

In view of Definition 5.2, the invariance to an orthogonal transformation imposes certain conditions on the elasticity matrix. For the transformed and the original matrices to be identical to one another, they must possess a particular form. Herein, we study a method where, given an orthogonal transformation, we can find the elasticity matrix that is invariant under this transformation and, hence, describe the material symmetry exhibited by a particular continuum. This method is stated in the following theorem.

Theorem 5.1 *The elastic properties of a continuum are invariant under an orthogonal transformation, given by matrix* $\mathbf{A}$, *if and only if*

$$\mathbf{C} = \mathbf{M}_{\mathbf{A}}^{T} \mathbf{C} \, \mathbf{M}_{\mathbf{A}}, \tag{5.16}$$

where $\mathbf{C}$ *is the elasticity matrix and* $\mathbf{M}_{\mathbf{A}}$ *is matrix (5.13).*

Proof. Consider stress-strain equations (3.7), namely,

$$\underline{\sigma} = \mathbf{C} \, \underline{\varepsilon}, \tag{5.17}$$

which are expressed in terms of the original coordinate system. In the transformed coordinate system, these equations are written as

$$\hat{\underline{\sigma}} = \hat{\mathbf{C}} \, \hat{\underline{\varepsilon}}. \tag{5.18}$$

Consider equation (5.18). Substituting expressions (5.6) and (5.14) into equation (5.18), we obtain

$$\underline{\mathbf{A}} \, \underline{\sigma} = \hat{\mathbf{C}} \, \mathbf{M}_{\mathbf{A}} \, \underline{\varepsilon}.$$

Multiplying both sides by $\underline{\mathbf{A}}^{-1}$, we get

$$\underline{\sigma} = \underline{\mathbf{A}}^{-1} \hat{\mathbf{C}} \, \mathbf{M}_{\mathbf{A}} \, \underline{\varepsilon}.$$

According to Lemma 5.1 shown below, $\underline{\mathbf{A}}^{-1} = \mathbf{M}_{\mathbf{A}}^{T}$. Hence, we can write

$$\underline{\sigma} = \mathbf{M}_{\mathbf{A}}^{T} \hat{\mathbf{C}} \, \mathbf{M}_{\mathbf{A}} \, \underline{\varepsilon}. \tag{5.19}$$

Examining equations (5.17) and (5.19), we conclude that they both hold for any $\underline{\varepsilon}$, if and only if

$$\mathbf{C} = \mathbf{M}_{\mathbf{A}}^{T} \hat{\mathbf{C}} \, \mathbf{M}_{\mathbf{A}},$$

which is the relation between $\mathbf{C}$ and $\hat{\mathbf{C}}$ under transformation matrix $\mathbf{A}$. In view of Definition 5.2, invariance with respect to $\mathbf{A}$ means that

$$\mathbf{C} = \mathbf{M}_{\mathbf{A}}^{T} \mathbf{C} \, \mathbf{M}_{\mathbf{A}},$$

which is expression (5.16), as required. ∎

Lemma 5.1 *Let $\underline{\mathbf{A}}$ be given by matrix (5.9) and $\mathbf{M}_{\mathbf{A}}$ be given by matrix (5.13). It follows that $\underline{\mathbf{A}}^{-1} = \mathbf{M}_{\mathbf{A}}^{T}$.*

Proof. Recall that $\mathbf{A}$ is an orthogonal matrix. Let $\hat{\mathbf{x}} = \mathbf{Ax}$ be the transformed coordinate system. Consider expression $\sigma\mathbf{x}$, which, in view of expression (5.3), can be stated in the $\hat{\mathbf{x}}$-coordinates as $(\mathbf{A}\sigma\mathbf{A}^{\mathbf{T}})\hat{\mathbf{x}}$. Thus, in terms of the $\hat{\mathbf{x}}$-coordinates, the stress-tensor components become

$$\hat{\sigma} = \mathbf{A}\sigma\mathbf{A}^{T}. \tag{5.20}$$

Let us calculate $\underline{\mathbf{A}}^{-1}$. Since $\mathbf{A}$ is an orthogonal matrix, namely, $\mathbf{A}^{T} = \mathbf{A}^{-1}$, we can rewrite equation (5.20) as

$$\sigma = \mathbf{A}^{T}\hat{\sigma}\mathbf{A}. \tag{5.21}$$

Thus, in a manner analogous to that used to obtain expression (5.6), we can rewrite expression (5.21) in the desired notation, as

$$\underline{\sigma} = \underline{\mathbf{A}^{T}}\,\underline{\hat{\sigma}}, \tag{5.22}$$

where $\underline{\mathbf{A}^{T}}$ is constructed as matrix (5.9), but with the entries $A_{ij}^{T} = A_{ji}$ of $\mathbf{A}^{T}$ used in place of the entries A_{ij} of $\mathbf{A}$. Note that the order of operations matters; namely, $\underline{\mathbf{A}^{T}} \neq \underline{\mathbf{A}}^{T}$. Comparing expression (5.6) with expression (5.22), we see that $\underline{\mathbf{A}^{T}} = \underline{\mathbf{A}}^{-1}$. Hence, we can write the inverse of matrix $\underline{\mathbf{A}}$ explicitly, as

$$\underline{\mathbf{A}}^{-1} = \tag{5.23}$$

$$\begin{bmatrix} A_{11}A_{11} & A_{21}A_{21} & A_{31}A_{31} & 2A_{21}A_{31} & 2A_{11}A_{31} & 2A_{11}A_{21} \\ A_{12}A_{12} & A_{22}A_{22} & A_{32}A_{32} & 2A_{22}A_{32} & 2A_{12}A_{32} & 2A_{12}A_{22} \\ A_{13}A_{13} & A_{23}A_{23} & A_{33}A_{33} & 2A_{23}A_{33} & 2A_{13}A_{33} & 2A_{13}A_{23} \\ A_{12}A_{13} & A_{22}A_{23} & A_{32}A_{33} & A_{22}A_{33}+A_{32}A_{23} & A_{12}A_{33}+A_{23}A_{31} & A_{21}A_{32}+A_{22}A_{31} \\ A_{11}A_{13} & A_{21}A_{23} & A_{31}A_{33} & A_{21}A_{33}+A_{31}A_{23} & A_{11}A_{33}+A_{31}A_{13} & A_{11}A_{23}+A_{21}A_{13} \\ A_{11}A_{12} & A_{21}A_{22} & A_{31}A_{32} & A_{21}A_{32}+A_{31}A_{22} & A_{11}A_{32}+A_{31}A_{12} & A_{11}A_{22}+A_{21}A_{12} \end{bmatrix}.$$

Comparing the entries of matrices (5.23) and (5.13), we notice that the former one is equal to the transpose of the latter, as required. ■

Expression (5.16) is a concise statement of conditions that the entries of the elasticity matrix must obey in order for the continuum described by stress-strain equations (4.11) to be invariant under an orthogonal transformation. Given transformation (5.2), expression (5.16) is convenient to apply since it contains twenty-one linear equations for C_{mn}. Furthermore, considering transformations (5.40) and (5.26), matrix $\mathbf{M}_{A}$ is significantly simplified, since $A_{13} = A_{23} = A_{31} = A_{32} = 0$, while $A_{33} = \pm 1$.

5.4 Point symmetry

Let us illustrate condition (5.16) by describing the material symmetry that is valid for all continua described by stress-strain equations (4.11). In the following theorem, we show that at every point, an elastic continuum is invariant under the reflection through the origin of the coordinate system located at this point. Such a reflection is described by the transformation matrix given by

$$\mathbf{A}_{-\mathbf{I}} := \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -\mathbf{I}. \tag{5.24}$$

Theorem 5.2 *At every point, a continuum given by stress-strain equations (4.11) is invariant under the reflection about the origin of a coordinate system that is located at that point.*

Proof. Consider transformation matrix (5.24). Matrix (5.13) becomes

$$\mathbf{M}_{\mathbf{A}_{-\mathbf{I}}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I}.$$

Hence, condition (5.16) becomes

$$\mathbf{C} = \mathbf{I}^T \mathbf{C}\,\mathbf{I},$$

which is identically satisfied for any $\mathbf{C}$. ■

This means that the symmetry group of every continuum contains $\mathbf{A}_{-\mathbf{I}}$.

5.5 Generally anisotropic continuum

A generally anisotropic continuum is the most general continuum describable by stress-strain equations (4.11). The elasticity matrix of a generally anisotropic continuum is given by

$$\mathbf{C}_{\mathrm{GEN}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}. \tag{5.25}$$

The only symmetry exhibited by a generally anisotropic continuum is point symmetry. Hence, a generally anisotropic continuum is described by an elasticity matrix that contains twenty-one independent entries.

5.6 Monoclinic continuum

5.6.1 Elasticity matrix

A continuum whose symmetry group contains a reflection about a plane through the origin is said to be monoclinic. For convenience, let us choose the coordinate system such that this reflection takes place about the x_1x_2-plane, which means, along the x_3-axis.

Consider the orthogonal transformation that is represented by matrix (5.2) in the form given by

$$\mathbf{A} = \begin{bmatrix} \cos\Theta & \sin\Theta & 0 \\ -\sin\Theta & \cos\Theta & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{5.26}$$

Matrix (5.26), whose determinant is equal to negative unity, corresponds to the composition of two transformations, namely, rotation by angle Θ about the x_3-axis and reflection about the x_1x_2-plane. To consider the reflection alone, we let $\Theta = 0$ to obtain

$$\mathbf{A}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{5.27}$$

Following expression (5.13), the corresponding matrix $\mathbf{M}_A$ is

$$\mathbf{M}_{A_3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Theorem 5.1 requires that the elasticity matrix satisfies condition (5.16). This condition requires that

$$\mathbf{C} = \mathbf{M}_{\mathbf{A}_3}^T \mathbf{C}\, \mathbf{M}_{\mathbf{A}_3},$$

which we can explicitly write as

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}$$
$$= \begin{bmatrix} C_{11} & C_{12} & C_{13} & -C_{14} & -C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & -C_{24} & -C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & -C_{34} & -C_{35} & C_{36} \\ -C_{14} & -C_{24} & -C_{34} & C_{44} & C_{45} & -C_{46} \\ -C_{15} & -C_{25} & -C_{35} & C_{45} & C_{55} & -C_{56} \\ C_{16} & C_{26} & C_{36} & -C_{46} & -C_{56} & C_{66} \end{bmatrix}.$$

The equality of these two matrices implies that

$$C_{14} = C_{15} = C_{24} = C_{25} = C_{34} = C_{35} = C_{46} = C_{56} = 0. \tag{5.28}$$

Thus, the elasticity matrix of a continuum that possesses a reflection symmetry along the x_3-axis is

$$\mathbf{C}_{\mathrm{MONO}x_3} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & C_{16} \\ C_{12} & C_{22} & C_{23} & 0 & 0 & C_{26} \\ C_{13} & C_{23} & C_{33} & 0 & 0 & C_{36} \\ 0 & 0 & 0 & C_{44} & C_{45} & 0 \\ 0 & 0 & 0 & C_{45} & C_{55} & 0 \\ C_{16} & C_{26} & C_{36} & 0 & 0 & C_{66} \end{bmatrix}. \tag{5.29}$$

Hence, a monoclinic continuum is described by an elasticity matrix that contains thirteen independent entries.

5.6.2 Natural coordinate system

In general, in an arbitrary coordinate system, all the entries of an elasticity matrix are nonzero. A natural coordinate system is a particular system within which an elasticity matrix has the fewest possible number of nonzero independent entries.

For any continuum, there exists at least three natural coordinate systems.[3] Hence, in principle, we could also rotate our orthonormal coordinate

[3]Interested readers might refer to Fedorov, F.I., (1968) Theory of elastic waves in

system so as to find a natural coordinate system for a generally anisotropic continuum, discussed above in Section 5.5. This, however, is not a simple task. Yet, in a natural coordinate system, a generally anisotropic continuum can be described by eighteen independent elasticity parameters and three Euler's angles that specify the orientation of this system.

A monoclinic continuum can be conveniently used to illustrate the concept of a natural coordinate system. The coordinate system that is used to formulate matrix (5.29) has the x_3-axis coinciding with the normal to the symmetry plane of the continuum. In other words, the x_1x_2-plane coincides with the symmetry plane. The rotation of the coordinate system about the x_3-axis allows us to further reduce the number of elasticity parameters needed to describe a monoclinic continuum. An appropriate rotation reduces matrix (5.29) to a new matrix that contains only twelve parameters. This orientation of the coordinate system is a natural coordinate system for a monoclinic continuum.

Rotation of the coordinate axes about the x_3-axis by angle Θ, where the angle is given by

$$\tan(2\Theta) = \frac{2C_{45}}{C_{44} - C_{55}}, \tag{5.30}$$

leads to a new set of elasticity parameters, which we denote by $\hat{C}_{mn}$.[4] In the new set, $\hat{C}_{45}$ vanishes and elasticity matrix (5.29) is reduced to

$$\hat{\mathbf{C}}_{\mathrm{MONO}} = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & 0 & 0 & \hat{C}_{16} \\ \hat{C}_{12} & \hat{C}_{22} & \hat{C}_{23} & 0 & 0 & \hat{C}_{26} \\ \hat{C}_{13} & \hat{C}_{23} & \hat{C}_{33} & 0 & 0 & \hat{C}_{36} \\ 0 & 0 & 0 & \hat{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{C}_{55} & 0 \\ \hat{C}_{16} & \hat{C}_{26} & \hat{C}_{36} & 0 & 0 & \hat{C}_{66} \end{bmatrix}. \tag{5.31}$$

Hence, in a natural coordinate system, a monoclinic continuum is described by twelve independent elasticity parameters and the angle Θ that describes the orientation of the coordinate system and corresponds to Euler's angle.

crystals: Plenum Press, New York, p. 25 and pp. 110 – 111, to Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon, pp. 163 – 170, to Lanczos, C., (1949/1986) The variational principles of mechanics: Dover, p. 373, to Schouten, J.A. (1951/1989) Tensor analysis for physicists: Dover, p. 162., and to Winterstein, D.F., (1990) Velocity anisotropy terminology for geophysicists: Geophysics, **55**, 1070 – 1088.

[4]Readers interested in the formulation of expression (5.30), in the context presented herein, might refer to Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon, pp. 82 – 83, 94 – 95 and 110 – 116.

Note that expression (5.30) can be verified by diagonalizing a submatrix, namely,

$$\begin{bmatrix} \hat{C}_{44} & 0 \\ 0 & \hat{C}_{55} \end{bmatrix} = \begin{bmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{bmatrix} \begin{bmatrix} C_{44} & C_{45} \\ C_{45} & C_{55} \end{bmatrix} \begin{bmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{bmatrix}. \tag{5.32}$$

We note that considering matrices (5.29) and (5.31), $C_{mn} \neq \hat{C}_{mn}$, where $m, n \in \{1, \ldots, 6\}$. In other words, the rotation about the x_3-axis by the angle (5.30) results in a new elasticity matrix to describe the same continuum.

In the context of ray theory, natural coordinate systems are associated with pure-mode directions, as discussed in Section 10.2.1. In Section 10.2.1, expression (5.30) is obtained by considering the displacement directions of the three types of waves that propagate along the x_3-axis in a monoclinic continuum.

The orthotropic, tetragonal, transversely isotropic, and isotropic continua, discussed in Sections 5.7 – 5.10 are all, *ab initio*, considered in their natural coordinate systems. Notably, for an isotropic continuum, all orthonormal coordinate systems are natural, while for orthotropic, tetragonal and transversely isotropic continua, we can obtain a natural coordinate system by setting the axes of the coordinate system to coincide with the symmetry axes. This is not the case for generally anisotropic and monoclinic continua, where the orientation of a natural coordinate system is more difficult to find.

5.7 Orthotropic continuum

An orthotropic continuum is a continuum that possesses three orthogonal symmetry planes. For convenience, let us choose the coordinate system such that the symmetry planes coincide with the coordinate planes. This is a natural coordinate system for an orthotropic continuum. Hence, the transformation matrices are given by

$$\mathbf{A}_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.33}$$

$$\mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.34}$$

and $\mathbf{A}_3$, given by matrix (5.27), which correspond to the reflections along the x_1-axis, the x_2-axis and the x_3-axis, respectively.

In view of the properties of the symmetry group, the elasticity matrix of an orthotropic continuum can be obtained using any two of the three symmetry planes. This is shown by the following theorem.

Theorem 5.3 *If a continuum given by stress-strain equations (4.11) is invariant under the reflection about two orthogonal planes, it must also be invariant under the reflection about the third orthogonal plane.*

Proof. Consider a continuum that is invariant under the reflections along the x_1-axis and along the x_3-axis. The corresponding orthogonal transformations are given by matrices (5.33) and (5.27), respectively. Also, following Theorem 5.2, all continua possess point symmetry. In other words, they are invariant under the transformation given by matrix (5.24). Since all these transformations belong to the symmetry group of an orthotropic continuum, their products also belong to this group. Consider

$$(\mathbf{A}_1)(\mathbf{A}_3)(\mathbf{A}_{-\mathbf{I}}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which we recognize to be matrix (5.34) corresponding to the reflections along the x_2-axis. Thus, in view of point symmetry, invariance to the reflections about two orthogonal planes also implies invariance to the reflection about the third orthogonal plane. ■

Therefore, to obtain the elasticity matrix of an orthotropic continuum, let us use matrices (5.27) and (5.34). Matrix (5.27) is also used in Section 5.6 where we obtain the relations given in expression (5.28), namely,

$$C_{14} = C_{15} = C_{24} = C_{25} = C_{34} = C_{35} = C_{46} = C_{56} = 0. \tag{5.35}$$

Using matrix (5.34), condition (5.16) becomes

$$\mathbf{C} = \mathbf{M}_{A_2}^T \mathbf{C} \mathbf{M}_{A_2}. \tag{5.36}$$

Hence, using matrix (5.13), we obtain

$$\mathbf{M}_{A_2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \tag{5.37}$$

Inserting matrix (5.37) into equation (5.36), we get

$$C_{16} = C_{26} = C_{36} = C_{45} = 0. \tag{5.38}$$

Thus, combining relations (5.35) and (5.38), we can write the elasticity matrix for an orthotropic continuum as

$$\mathbf{C}_{\mathrm{ORTHO}x_1x_2x_3} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}. \tag{5.39}$$

Hence, in a natural coordinate system, an orthotropic continuum is described by nine independent elasticity parameters.

5.8 Tetragonal continuum

A tetragonal continuum is a continuum whose symmetry group contains a four-fold rotation and a reflection through the plane that contains the axis of rotation. For convenience, let us choose the coordinate system such that the x_3-axis is the axis of rotation, while the reflection is along the x_2-axis. This is a natural coordinate system for a tetragonal continuum.

Consider the orthogonal transformation that is represented by matrix (5.2) in the form given by

$$\mathbf{A}_{\Theta} = \begin{bmatrix} \cos\Theta & \sin\Theta & 0 \\ -\sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.40}$$

Matrix (5.40), whose determinant is equal to unity, corresponds to rotation by angle Θ about the x_3-axis.

The transformation matrices of a tetragonal continuum are given by

$$\mathbf{A}_{\pi/2} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.41}$$

which is matrix (5.40) with $\Theta = \pi/2$, and by matrix (5.34). These matrices correspond to the rotation about the x_3-axis and to the reflections along the x_2-axis, respectively.

Matrix (5.34) also belongs to the symmetry group of an orthotropic continuum discussed in Section 5.7 where we obtained the relations given in expression (5.38), namely,

$$C_{16} = C_{26} = C_{36} = C_{45} = 0. \tag{5.42}$$

These relations also apply to a tetragonal continuum. The additional relations result from matrix (5.41). Using matrix (5.41), condition (5.16) becomes

$$\mathbf{C} = \mathbf{M}^T_{\mathbf{A}_{\pi/2}} \mathbf{C} \mathbf{M}_{\mathbf{A}_{\pi/2}}$$

and results in relations given by

$$C_{22} = C_{11},\ C_{23} = C_{13},\ C_{55} = C_{44}. \tag{5.43}$$

Combining relations (5.42) and (5.43), we obtain the elasticity matrix of a tetragonal continuum, namely,

$$\mathbf{C}_{\mathrm{TETRA}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}. \tag{5.44}$$

Thus, in a natural coordinate system, only six independent elasticity parameters are needed to describe a tetragonal continuum.

Note that, as expected, matrix (5.44) is a special case of matrix (5.39) with additional relations given by expression (5.43).

5.9 Transversely isotropic continuum

5.9.1 Elasticity matrix

Now we will consider a particularly interesting case. Suppose that a continuum is invariant with respect to a single rotation given by matrix (5.40) where Θ is smaller than $\pi/2$. Let us consider, for example, $\Theta = 2\pi/5$, and, hence, assume that the symmetry group contains

$$\mathbf{A}_{2\pi/5} = \begin{bmatrix} \cos\frac{2\pi}{5} & \sin\frac{2\pi}{5} & 0 \\ -\sin\frac{2\pi}{5} & \cos\frac{2\pi}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.45}$$

Following condition (5.16), the elasticity matrix, $\mathbf{C}$, satisfies the equation given by

$$\mathbf{C} = \mathbf{M}^T_{\mathbf{A}_{2\pi/5}} \mathbf{C} \mathbf{M}_{\mathbf{A}_{2\pi/5}}. \tag{5.46}$$

The entries of matrix $\mathbf{M}_{\mathbf{A}_{2\pi/5}}$ are more complicated than the entries of the transformation matrices used in the previous sections, but equation (5.46) can still be solved directly to give relations among the entries of $\mathbf{C}$. The solution to condition (5.46) is the matrix given by

$$C_{\mathrm{TRANS}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \dfrac{C_{11}-C_{12}}{2} \end{bmatrix}. \tag{5.47}$$

Thus, the requirements that the symmetry group contains $\mathbf{A}_{2\pi/5}$ results in a continuum that is described by only five independent elasticity parameters.

5.9.2 Rotation invariance

A particularly important property of matrix (5.47) is the fact that for any angle Θ, this matrix, without any further simplification, satisfies the equation given by

$$\mathbf{C} = \mathbf{M}^T_{\mathbf{A}_\Theta} \mathbf{C} \mathbf{M}_{\mathbf{A}_\Theta}, \tag{5.48}$$

where $\mathbf{A}_\Theta$ is given by matrix (5.40). This property of matrix (5.47) can be directly verified by substituting $\mathbf{M}_{\mathbf{A}_\Theta}$, without specifying the value of Θ, and $\mathbf{C}_{\mathrm{TRANS}}$ into the right-hand side of equation (5.48). The resulting expression reduces to $\mathbf{C}_{\mathrm{TRANS}}$. Therefore, the invariance of $\mathbf{C}_{\mathrm{TRANS}}$ to the five-fold rotation about a given axis implies invariance to the rotation by any angle about this axis. As stated in Theorem 5.4, below, there is nothing special about $2\pi/5$; we could choose any angle smaller than $\pi/2$ to obtain the same elasticity matrix.

To prove Theorem 5.4 below, and to see the reason behind it, consider the fact that the material symmetry of a continuum is equivalent to the symmetry of the strain-energy function, $W(\varepsilon)$, as discussed in Section 4.2.3.[5]

[5]Readers interested in the treatment of symmetries that is based on the strain-energy function might refer to Carcione, J.M., (2001) Wave fields in real media: wave propagation in anisotropic, anelastic and porous media: Pergamon, pp. 2 – 3, to Epstein, M., and Slawinski, M.A., (1998) On some aspects of the continuum-mechanics context. Revue de

Since strain energy is a scalar, its value must be the same for all orientations of the coordinate system. In view of expression (4.3), namely,

$$W(\varepsilon)=\frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}\varepsilon_{ij}\varepsilon_{kl}, \tag{5.49}$$

this is, in general, achieved by the values of the components of c_{ijkl}, which are different for different orientations of the coordinate system in such a way that the value of W remains the same. If a continuum exhibits a given material symmetry, the same values of the components of c_{ijkl} give the same value of W for more than one orientation of the coordinate system.

We wish to express the effect of an orthogonal transformation, $\mathbf{A}$, on the strain-energy function. Since the value of strain energy must be the same for the original and the transformed coordinate systems, we can write

$$\hat{W}(\hat{\varepsilon})=W(\varepsilon), \tag{5.50}$$

where the transformed strain-tensor components are given by expression (5.10), namely,

$$\hat{\varepsilon}=\mathbf{A}\varepsilon\mathbf{A}^{T}, \tag{5.51}$$

which, for brevity, we denote by $\mathbf{A}\circ\varepsilon$, with $\circ$ standing for the orthogonal-transformation operator.

Note that equation (5.50) is always satisfied. Consequently, this equation provides no information about the material symmetry of the continuum. The material symmetry requires that the strain-energy function be invariant under $\mathbf{A}$, namely,

$$W(\hat{\varepsilon})=W(\varepsilon),$$

which we can write as $W(\varepsilon)=W(\mathbf{A}\circ\varepsilon)$. In other words, material symmetry requires that the strain-energy function be the same for both ε and $\hat{\varepsilon}$.

Herein, we are interested in rotations of the coordinate system; hence, we consider transformation $\mathbf{A}_{\Theta}$, which is given by expression (5.40). In view of expressions (5.40) and (5.51), $\hat{\varepsilon}$ can be regarded as a quadratic trigonometric polynomial in Θ. Hence, $W(\mathbf{A}_{\Theta}\circ\varepsilon)$ is a quartic trigonometric polynomial in Θ.[6]

l'Institut Français du Pétrole, **53**, No. 5, pp. 673 – 674, to Lanczos, C., (1949/1986) The variational principles of mechanics: Dover, pp. 373 – 374, and to Macelwane, J.B., and Sohon, F.W., (1936) Introduction to theoretical seismology, Part I: Geodynamics: John Wiley and Sons, Inc., pp. 77 – 78.

[6]Readers interested in trigonometric polynomials might refer to Courant, R., and Hilbert, D., (1924/1989) Methods of mathematical physics: John Wiley & Sons, Vol. I, p. 69 – 70.

Now, let the material symmetry be the invariance under rotation by $\Theta = 2\pi/n$, where $n \geq 5$. Hence, consider the strain-energy function that is invariant under such rotation. Since the symmetries of a continuum form a group, we conclude that $W(\varepsilon)$ is also invariant under rotations by $2m\pi/n$, where $m \in \{0, 1, \ldots, n-1\}$. In other words, the symmetry group contains rotations given by

$$\mathbf{A}_{2m\pi/n} = \begin{bmatrix} \cos\frac{2m\pi}{n} & \sin\frac{2m\pi}{n} & 0 \\ -\sin\frac{2m\pi}{n} & \cos\frac{2m\pi}{n} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{array}{l} n \geq 5 \\ m \in \{0, 1, \ldots, n-1\} \end{array}. \tag{5.52}$$

So, we can write

$$W(\varepsilon) = W\left(\mathbf{A}_{2m\pi/n} \circ \varepsilon\right), \qquad \begin{array}{l} n \geq 5 \\ m \in \{0, 1, \ldots, n-1\} \end{array}, \tag{5.53}$$

which implies that

$$W(\varepsilon) = \frac{1}{n}\sum_{m=0}^{n-1} W\left(\mathbf{A}_{2m\pi/n} \circ \varepsilon\right), \qquad n \geq 5. \tag{5.54}$$

In other words, since equation (5.53) holds for any allowable value of m, the sum on the right-hand side of equation (5.54) is composed of the identical values of W.

Note that for any $W(\varepsilon)$, not necessarily invariant under transformations (5.52), the right-hand side of equation (5.54) is called the symmetrization of W with respect to the group of these transformations. Hence, equation (5.54) means that if $W(\varepsilon)$ is invariant under rotations (5.52), it is equal to its symmetrization with respect to these rotations.

Now we can introduce the key statement that explains why an elasticity matrix invariant to a five-fold rotation about a given axis is necessarily invariant to any rotation about this axis. Since $W(\mathbf{A}_\Theta \circ \varepsilon)$ is a quartic trigonometric polynomial, it follows that, for $n \geq 5$, we can apply Lemma 5.2, below, and rewrite the right-hand side of equation (5.54) as

$$\frac{1}{n}\sum_{m=0}^{n-1} W\left(\mathbf{A}_{2m\pi/n} \circ \varepsilon\right) = \frac{1}{2\pi}\int_0^{2\pi} W(\mathbf{A}_\Theta \circ \varepsilon)\, \mathrm{d}\Theta, \qquad n \geq 5,$$

which leads to

$$W(\varepsilon) = \frac{1}{2\pi}\int_{0}^{2\pi} W(\mathbf{A}_\Theta \circ \varepsilon)\,\mathrm{d}\Theta. \tag{5.55}$$

Equations (5.55) states that $W(\varepsilon)$ is equal to its symmetrization over all possible rotations. This implies that $W(\varepsilon)$ is invariant under all rotations. Hence, we conclude with the following theorem.

Theorem 5.4 *If $W(\varepsilon)$ is invariant under rotations by angle $2\pi/n$ about a given axis, where $n \geq 5$, it is invariant under any rotation about this axis.*

To complete the proof of this theorem, consider the following lemma.

Lemma 5.2 *If f is a trigonometric polynomial of at most degree $n-1$, then*

$$\frac{1}{n}\sum_{m=0}^{n-1} f\left(\frac{2m\pi}{n}\right) = \frac{1}{2\pi}\int_{0}^{2\pi} f(\Theta)\,\mathrm{d}\Theta. \tag{5.56}$$

Proof. Consider a basis of the space of trigonometric polynomials of at most degree $n-1$, given by

$$f_r(\Theta) = e^{ir\Theta}, \qquad r \in \{-(n-1), \ldots, n-1\}.$$

In view of linearity, to prove equation (5.56) for f, it suffices to prove it for f_r, where $r \in \{-(n-1), \ldots, n-1\}$. Set

$$z = e^{ir\frac{2\pi}{n}}, \qquad r \in \{-(n-1), \ldots, n-1\}. \tag{5.57}$$

Then, we can write

$$\frac{1}{n}\sum_{m=0}^{n-1} f_r\left(\frac{2m\pi}{n}\right) = \frac{1}{n}\sum_{m=0}^{n-1} z^m, \qquad r \in \{-(n-1), \ldots, n-1\}. \tag{5.58}$$

Examining expression (5.57) for the case of $r = 0$, we note that $z = 1$ and, thus, the right-hand side of expression (5.58) is equal to 1. Now, for $r \neq 0$, $z \neq 1$ and we can write the right-hand side of expression (5.58) as

$$\frac{1}{n}\sum_{m=0}^{n-1} z^m = \frac{1}{n}\frac{z^n - 1}{z - 1}. \tag{5.59}$$

Examining expression (5.57), we also note that $z^n = 1$ and, thus, the right-hand side of expression (5.59) is equal to 0. To summarize, we can write the left-hand side of equation (5.56) as

$$\frac{1}{n}\sum_{m=0}^{n-1} f_r\left(\frac{2m\pi}{n}\right) = \begin{cases} 0 & \text{if} \quad r \neq 0 \\ 1 & \text{if} \quad r = 0 \end{cases} . \tag{5.60}$$

Performing the integration on the right-hand side of equation (5.56), we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} e^{ir\Theta}\mathrm{d}\Theta = \begin{cases} 0 & \text{if} \quad r \neq 0 \\ 1 & \text{if} \quad r = 0 \end{cases} . \tag{5.61}$$

Thus, for $r \in \{-(n-1), \ldots, n-1\}$, expressions (5.60) and (5.61) are equal to one another and, hence, equation (5.56) is valid for polynomials of at most degree $n-1$, as required. ■

Note that in the proof of Theorem 5.4 we used the fact that $W(\varepsilon)$ is a quadratic polynomial in the strain-tensor components and, hence, $W(\mathbf{A}_\Theta \circ \varepsilon)$ is a quartic trigonometric polynomial in Θ. This corresponds to the fact that Theorem 5.4 is associated with c_{ijkl}, which is a fourth-rank tensor. This theorem can be extended to the higher-rank tensors if they are subject to similar transformations. In general, such a rotation invariance was given by Herman (1945) and is shown in Exercise 5.4.

Since the symmetry of $W(\varepsilon)$ is tantamount to the symmetry of a continuum, we conclude that a continuum described by matrix (5.47) is transversely isotropic.

5.10 Isotropic continuum

5.10.1 Elasticity matrix

A continuum whose symmetry group contains all orthogonal transformations is said to be isotropic. For an isotropic continuum, all coordinate systems are natural coordinate systems and, hence, no particular orientation is required.

Since the symmetry group of an isotropic continuum contains all orthogonal transformations, it must contain all rotations about the x_3-axis. Thus, the elasticity matrix of an isotropic continuum has, at least, the simplicity of the form shown in matrix (5.47). Consider also the invariance to the transformation that exchanges the x_1 and x_3 coordinates, namely,

$$\mathbf{A}_{x_1x_3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Following condition (5.16), we obtain the equation given by

$$\mathbf{C} = \mathbf{M}^T_{A_{x_1x_3}} \mathbf{C} \mathbf{M}_{A_{x_1x_3}},$$

which imposes the additional relations, namely,

$$C_{11} = C_{33},\ C_{12} = C_{13},\ C_{44} = C_{66}.$$

Incorporating these relations into matrix (5.47), we obtain

$$\mathbf{C}_{\mathrm{ISO}} = \begin{bmatrix} C_{11} & C_{11} - 2C_{44} & C_{11} - 2C_{44} & 0 & 0 & 0 \\ C_{11} - 2C_{44} & C_{11} & C_{11} - 2C_{44} & 0 & 0 & 0 \\ C_{11} - 2C_{44} & C_{11} - 2C_{44} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix}. \tag{5.62}$$

Hence, the elasticity matrix of an isotropic continuum contains only two independent elasticity parameters, namely, C_{11} and C_{44}. Furthermore, as shown in Exercise 5.5, the elasticity matrix of an isotropic continuum is symmetric without invoking the existence of the strain-energy function.

5.10.2 Lamé's parameters

The two independent elasticity parameters that describe an isotropic continuum are often expressed as

$$\begin{cases} \lambda := C_{11} - 2C_{44} \\ \mu := C_{44} \end{cases}. \tag{5.63}$$

The two parameters, λ and μ, are called Lamé's parameters. Their physical meaning is described in Section 5.10.4.

Using the definition of Lamé's parameters (5.63), we can rewrite matrix (5.62) as

$$\mathbf{C}_{\mathrm{LAMÉ}} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}. \tag{5.64}$$

5.10.3 Tensorial formulation

Using matrix (5.64) and in view of equations (3.1), we can write the stress-strain equations for an isotropic continuum as

$$\sigma_{ij} = \lambda\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \qquad i,j \in \{1,2,3\}. \tag{5.65}$$

This formulation is used to derive the wave equation in Chapter 6.

Since, in equations (3.1), the elasticity tensor, c_{ijkl}, is a fourth-rank tensor, the number of elasticity parameters for an isotropic continuum can also be derived directly from the mathematical properties of a fourth-rank tensor and the concept of an isotropic tensor.

Note that an isotropic tensor is a tensor whose components are the same in all coordinate systems.

The general form of an isotropic fourth-rank tensor is

$$a_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \xi\delta_{ik}\delta_{jl} + \eta\delta_{il}\delta_{jk}, \qquad i,j,k,l \in \{1,2,3\}, \tag{5.66}$$

where λ, ξ and η are constants. In other words, an isotropic fourth-rank tensor is stated in terms of three constants that do not depend on the choice of the coordinate system. In elasticity theory, since the strain tensor is symmetric, the most general isotropic elasticity tensor is given by expression (5.65), which contains only two constants, λ and μ, where, as shown in Exercise 5.6, $\mu := (\xi + \eta)/2$.

By examining stress-strain equations (5.65) in the context of tensor analysis, we can see that they correspond to the isotropic formulation since they retain the same form for all orthogonal transformations. To gain insight into this statement, we rewrite these equations using definition (1.15) as

$$\sigma_{ij} = \lambda\delta_{ij}\sum_{k=1}^{3}\frac{\partial u_k}{\partial x_k} + \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad i,j \in \{1,2,3\}. \tag{5.67}$$

Equations (5.67) are invariant under the coordinate transformations. We immediately see that the summation term is $\nabla \cdot \mathbf{u}$, which — being a scalar — is invariant under all coordinate transformations. Using transformation rules for the components of a second-rank tensor, we can also show that, upon the coordinate transformation, the term in parentheses retains the same form. Since both the summation term and the term in parentheses are invariant under the coordinate transformations, stress-strain equations (5.65) correspond to isotropic continua.

5.10.4 Physical meaning of Lamé's parameters

We can obtain the physical meaning of Lamé's parameters, λ and μ, by examining stress-strain equations (5.65).

Lamé's parameter μ is a measure of rigidity. We can see that by setting $\lambda = 0$ and considering ε_{ij} with $i \neq j$. Thus, we can write expressions (5.65) as

$$\sigma_{ij} = 2\mu\varepsilon_{ij}, \qquad \begin{matrix} i \neq j \\ i,j \in \{1,2,3\} \end{matrix},$$

which, using definition (1.15), we can rewrite as

$$\sigma_{ij} = \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad \begin{matrix} i \neq j \\ i,j \in \{1,2,3\} \end{matrix}.$$

In view of Section 1.3.2, we see that μ is a coefficient that relates stress to a change in shape. Thus, Lamé's parameter μ describes the rigidity of the continuum.

The physical meaning of Lamé's parameter λ is less immediate. If we let μ vanish and consider ε_{ij} with $i = j$, equations (5.65) become

$$\lim_{\mu\to 0^+} \sigma_{ij} = \lambda\sum_{k=1}^{3}\varepsilon_{kk} = \lambda\left(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}\right) = \lambda\varphi, \qquad \begin{matrix} i = j \\ i,j \in \{1,2,3\} \end{matrix}, \tag{5.68}$$

where φ is the dilatation defined in expression (1.26). Examination of expression (5.68), which can be viewed as corresponding to a fluid, shows that Lamé's parameter λ is akin to the compressibility, κ. Note however that, in view of the positive-definiteness of elasticity matrix (5.64), as required by the stability conditions for an elastic solid, discussed in Section 4.3, we require $\mu > 0$. Hence, we treat the vanishing of μ as a limit.

To study a proper solid, we consider a finite value of μ. We still consider ε_{ij} with $i = j$ and we further assume $\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{33}$. For convenience, let $\varepsilon_{ii} \equiv \tilde{\varepsilon}/3$, where $i \in \{1,2,3\}$. Thus, we can write stress-strain equations (5.65) as

$$\sigma_{ij} = \lambda\tilde{\varepsilon} + \frac{2}{3}\mu\tilde{\varepsilon} = \left(\lambda + \frac{2}{3}\mu\right)\tilde{\varepsilon}, \qquad \begin{matrix} i = j \\ i,j \in \{1,2,3\} \end{matrix}. \tag{5.69}$$

In view of expression (1.25) and letting $\Delta V := \check{V} - V$, we can write

$$\tilde{\varepsilon} = \frac{\Delta V}{V}.$$

Also, σ_{ii} is equal to $-P$, where P denotes the difference in hydrostatic pressure. In other words, P is a pressure difference between the pressure associated with the deformation and the pressure at the undeformed state. Thus, we can write expression (5.69) as

$$-P = \left(\lambda + \frac{2}{3}\mu\right)\frac{\Delta V}{V}. \tag{5.70}$$

To gain insight into the physical meaning of λ, we use the concept of compressibility, κ, that is defined as the relative decrease of volume produced by unit pressure, namely,

$$\kappa := -\frac{1}{P}\frac{\Delta V}{V}. \tag{5.71}$$

Using expression (5.70), we can rewrite the compressibility as

$$\kappa = \frac{1}{\lambda + \frac{2}{3}\mu}.$$

Solving for λ, we obtain

$$\lambda = \frac{1}{\kappa} - \frac{2}{3}\mu.$$

Thus, in the case of vanishing rigidity, $\mu \to 0^{+}$, Lamé's parameter λ is the reciprocal of the compressibility, while, in general, λ has a more complicated physical significance given in terms of both the rigidity and compressibility.

Closing remarks

In this chapter, by studying the elasticity matrix, we investigated the symmetries of the elasticity tensor that correspond to generally anisotropic, monoclinic, orthotropic, tetragonal, transversely isotropic and isotropic continua. By further investigating transformation properties of the elasticity tensor, we could also show that the only two remaining cases are the trigonal and cubic continua, which are described by six and three independent elasticity parameters, respectively. However, these two continua do not commonly appear in seismological studies.

Note that both the trigonal and tetragonal continua are described by the same number of independent elasticity parameters. However, the elasticity matrix of a trigonal continuum is different than the elasticity matrix of a tetragonal continuum. In the former case, the matrix contains eighteen nonzero entries, while, in the latter case, it contains twelve nonzero entries.[7]

Studying the symmetries of a continuum provides us with information about the material that it represents. For instance, by analyzing seismic data, we can infer information about layering and fractures. Also, knowing the smallest number of independent elasticity parameters that is required to describe a given continuum provides us with a convenient way to study seismic-wave propagation in specific materials. For instance, explicit expressions for wave velocities in a generally anisotropic continuum are complicated. However, if we know that a given material can be adequately described by a continuum that possesses particular symmetries, we reduce the complication of these expressions. Explicit expressions for wave velocities in a transversely isotropic continuum are derived in Chapter 10.

Note that the nomenclature commonly used to describe the material symmetries originates in crystallography. Herein, however, we are studying symmetries of continua. Consequently, intuitive and heuristic descriptions associated with crystal lattices are not appropriate in the context of elastic continua.

Exercises

Exercise 5.1 [8]*Show that the Jacobian that is associated with matrix (5.40) is equal to unity.*

Solution 5.1 *Using matrix (5.40), we can write the transformation of coordinates as*

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \begin{bmatrix} \cos\Theta & \sin\Theta & 0 \\ -\sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \tag{5.72}$$

where $\mathbf{x}$ *and* $\hat{\mathbf{x}}$ *are the original and the transformed coordinates, respectively.*

[7]Readers interested in a complete classification of symmetries for the elasticity tensor might refer to Ting, T.C.T., (1996) Anisotropic elasticity: Theory and applications: Oxford University Press, pp. 40 – 51.

[8]See also Section 5.1.1.

The Jacobian is given by

$$J := \det \begin{bmatrix} \frac{\partial \hat{x}_1}{\partial x_1} & \frac{\partial \hat{x}_1}{\partial x_2} & \frac{\partial \hat{x}_1}{\partial x_3} \\ \frac{\partial \hat{x}_2}{\partial x_1} & \frac{\partial \hat{x}_2}{\partial x_2} & \frac{\partial \hat{x}_2}{\partial x_3} \\ \frac{\partial \hat{x}_3}{\partial x_1} & \frac{\partial \hat{x}_3}{\partial x_2} & \frac{\partial \hat{x}_3}{\partial x_3} \end{bmatrix}. \tag{5.73}$$

Thus, examining equations (5.72) and (5.73), we see that the determinant of matrix (5.40) is the Jacobian. Hence, we immediately obtain

$$J = \det \begin{bmatrix} \cos\Theta & \sin\Theta & 0 \\ -\sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \cos^2\Theta + \sin^2\Theta = 1,$$

as required.

Exercise 5.2 *Consider a continuum whose symmetry group contains the reflection about the x_2x_3-plane. This reflection implies that $\varepsilon_{12} = -\hat{\varepsilon}_{12}$ and $\varepsilon_{13} = -\hat{\varepsilon}_{13}$, as well as $\sigma_{12} = -\hat{\sigma}_{12}$ and $\sigma_{13} = -\hat{\sigma}_{13}$. Using stress-strain equations (4.11), show that*

$$C_{15} = C_{16} = C_{25} = C_{26} = C_{35} = C_{36} = C_{45} = C_{46} = 0,$$

and state the resulting elasticity matrix $\mathbf{C}_{\mathrm{MONO}x_1}$.

Solution 5.2 *Consider the stress-tensor components σ_{12} and $\hat{\sigma}_{12}$. Using stress-strain equations (4.11), we can write*

$$\sigma_{12} = C_{16}\varepsilon_{11} + C_{26}\varepsilon_{22} + C_{36}\varepsilon_{33} + 2C_{46}\varepsilon_{23} + 2C_{56}\varepsilon_{13} + 2C_{66}\varepsilon_{12}, \tag{5.74}$$

and

$$\hat{\sigma}_{12} = C_{16}\hat{\varepsilon}_{11} + C_{26}\hat{\varepsilon}_{22} + C_{36}\hat{\varepsilon}_{33} + 2C_{46}\hat{\varepsilon}_{23} + 2C_{56}\hat{\varepsilon}_{13} + 2C_{66}\hat{\varepsilon}_{12}.$$

The second equation can be expressed in terms of the original strain components as

$$\hat{\sigma}_{12} = C_{16}\varepsilon_{11} + C_{26}\varepsilon_{22} + C_{36}\varepsilon_{33} + 2C_{46}\varepsilon_{23} - 2C_{56}\varepsilon_{13} - 2C_{66}\varepsilon_{12}.$$

In view of relations $\sigma_{12} = -\hat{\sigma}_{12}$, and the equality of the stress-strain equations required in view of the assumed symmetry, we obtain

$$\sigma_{12} = -\hat{\sigma}_{12} = -C_{16}\varepsilon_{11} - C_{26}\varepsilon_{22} - C_{36}\varepsilon_{33} - 2C_{46}\varepsilon_{23} + 2C_{56}\varepsilon_{13} + 2C_{66}\varepsilon_{12}. \tag{5.75}$$

Equality between (5.74) and (5.75) requires

$$C_{16} = C_{26} = C_{36} = C_{46} = 0.$$

Similarly, for $\sigma_{13} = -\hat{\sigma}_{13}$, we require

$$C_{15} = C_{25} = C_{35} = C_{45} = 0.$$

Thus, we obtain

$$\mathbf{C}_{\mathrm{MONO}x_1} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 0 & 0 \\ C_{12} & C_{22} & C_{23} & C_{24} & 0 & 0 \\ C_{13} & C_{23} & C_{33} & C_{34} & 0 & 0 \\ C_{14} & C_{24} & C_{34} & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & C_{56} \\ 0 & 0 & 0 & 0 & C_{56} & C_{66} \end{bmatrix}, \qquad (5.76)$$

as required.

Exercise 5.3 [9]*Find the stability conditions for a transversely isotropic continuum described by matrix (5.47).*

Solution 5.3 *In view of Section 4.3, the stability conditions require that matrix (5.47) be positive-definite. Recalling equations (4.14), we obtain*

$$C_{11} > 0, \qquad (5.77)$$

$$C_{33} > 0, \qquad (5.78)$$

$$C_{44} > 0, \qquad (5.79)$$

and

$$C_{11} > C_{12}. \qquad (5.80)$$

We notice that matrix (5.47) is a direct sum of two submatrices given by

$$\mathbf{C}_1 = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{12} & C_{11} & C_{13} \\ C_{13} & C_{13} & C_{33} \end{bmatrix},$$

and

$$\mathbf{C}_2 = \begin{bmatrix} C_{44} & 0 & 0 \\ 0 & C_{44} & 0 \\ 0 & 0 & \dfrac{C_{11} - C_{12}}{2} \end{bmatrix}.$$

[9]See also Section 4.3.3.

Conditions (5.79) and (5.80) ensure that matrix $\mathbf{C}_2$ *is positive-definite. In view of condition (5.77), the remaining conditions for the positive-definiteness of matrix* $\mathbf{C}_1$ *are*

$$\det\begin{bmatrix} C_{11} & C_{12} \\ C_{12} & C_{11} \end{bmatrix} > 0, \tag{5.81}$$

and

$$\det\begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{12} & C_{11} & C_{13} \\ C_{13} & C_{13} & C_{33} \end{bmatrix} > 0. \tag{5.82}$$

The condition resulting from determinant (5.81) is

$$C_{11} > |C_{12}|, \tag{5.83}$$

while the condition resulting from determinant (5.82) is

$$C_{33}(C_{11} - C_{12})(C_{11} + C_{12}) > 2C_{13}^2(C_{11} - C_{12}).$$

In view of expression (5.80), we can rewrite the latter condition as

$$C_{33}(C_{11} + C_{12}) > 2C_{13}^2. \tag{5.84}$$

Also, in view of condition (5.78), we have $C_{11} + C_{12} > 0$. *Consequently, condition (5.83) follows from conditions (5.78) and (5.84). Thus, all the stability conditions for a transversely isotropic continuum are given by expressions (5.77), (5.78), (5.79), (5.80) and (5.84).*

Remark 5.1 *Note that if matrix* $\mathbf{C}_1$ *is positive-definite, we also have*

$$\det\begin{bmatrix} C_{11} & C_{13} \\ C_{13} & C_{33} \end{bmatrix} > 0, \tag{5.85}$$

which we can write as

$$C_{11}C_{33} > C_{13}^2. \tag{5.86}$$

Herein, we will show that condition (5.86) is a consequence of condition (5.84). Let us rewrite condition (5.84) as

$$C_{33}(C_{11} + C_{12}) - 2C_{13}^2 > 0,$$

which we restate as

$$a + b > 0,$$

where

$$a := C_{11}C_{33} - C_{13}^2,$$

and

$$b := C_{12}C_{33} - C_{13}^2.$$

Using this notation, we can write condition (5.86) as $a > 0$. To show that condition (5.86) is a consequence of condition (5.84), we first show that $a > b$, which is equivalent to showing that

$$C_{11}C_{33} > C_{12}C_{33}. \tag{5.87}$$

Inequality (5.87) is true due to conditions (5.78) and (5.80). Hence, since $a + b > 0$ and $a - b > 0$, by summation we get $2a > 0$, which immediately implies that $a > 0$, as required.

Exercise 5.4 [10] *Using the formula for the change of coordinates for the components of a tensor as well as Lemma 5.3 below, show that if a tensor of rank n, given by $T_{i_1 \ldots i_n}$, is invariant under the $(n+1)$-fold rotation about a given axis, it is invariant under any rotation about this axis.*

Lemma 5.3 *Let $P(\Theta)$ be a trigonometric polynomial of at most degree n. If $P(\Theta)$ has a period of $2\pi/(n+1)$, then $P(\Theta) \equiv const.$*

Proof. Consider a basis of the space of trigonometric polynomials of at most degree n, given by

$$f_r(\Theta) = e^{ir\Theta}, \qquad r \in \{-n, \ldots, n\}. \tag{5.88}$$

We can uniquely write

$$P(\Theta) = \sum_{r=-n}^{n} \alpha_r f_r(\Theta), \tag{5.89}$$

where α_r are complex numbers. In view of expressions (5.88) and (5.89), we can write

$$P\left(\Theta + \frac{2\pi}{n+1}\right) = \sum_{r=-n}^{n} \alpha_r f_r\left(\Theta + \frac{2\pi}{n+1}\right) = \sum_{r=-n}^{n} \alpha_r e^{ir2\pi/(n+1)} f_r(\Theta). \tag{5.90}$$

Since $P(\Theta)$ has a period of $2\pi/(n+1)$, examining equations (5.89) and (5.90), we obtain

$$\alpha_r = \alpha_r e^{ir2\pi/(n+1)}, \qquad r \in \{-n, \ldots, n\}.$$

Observing that $e^{ir2\pi/(n+1)} \neq 1$ for all $r \in \{-n, \ldots, n\}$, except $r = 0$, we conclude that $\alpha_r = 0$, except, possibly, α_0. Hence, $P(\Theta)$ is constant. ∎

[10] See also Section 5.9.2.

Solution 5.4 *Consider transformation matrix (5.40), namely,*

$$\mathbf{A}=\left[\begin{array}{ccc}\cos\Theta & \sin\Theta & 0\\ -\sin\Theta & \cos\Theta & 0\\ 0 & 0 & 1\end{array}\right]. \tag{5.91}$$

The transformed tensor components are given by

$$\hat{T}_{i_1\ldots i_n}=\sum_{j_1=1}^{3}\cdots\sum_{j_n=1}^{3}T_{j_1\ldots j_n}A_{i_1j_1}\ldots A_{i_nj_n},\qquad i_1,\ldots,i_n\in\{1,2,3\}.$$

In view of matrix (5.91), we see that $\hat{T}_{i_1\ldots i_n}=\hat{T}_{i_1\ldots i_n}(\Theta)$ *is a trigonometric polynomial in* Θ *of at most degree* n. *Since tensor* $T_{i_1\ldots i_n}$ *is invariant under the rotation by the angle* $2\pi/(n+1)$, *polynomial* $\hat{T}_{i_1\ldots i_n}(\Theta)$ *has a period of* $2\pi/(n+1)$. *Since* $\hat{T}_{i_1\ldots i_n}(\Theta)$ *is at most of degree* n, *it follows from Lemma 5.3 that this trigonometric polynomial is constant. This means that* $T_{i_1\ldots i_n}$ *is invariant under any rotation, as required.*

Exercise 5.5 [11]*Show that for an isotropic continuum the elasticity matrix is symmetric without invoking the strain-energy function.*

Notation 5.1 *The repeated-index summation notation is used in this solution. Any term in which an index appears twice stands for the sum of all such terms as the index assumes all the values between* 1 *and* 3.

Solution 5.5 *In view of Section 4.2, to show the symmetry of the elasticity matrix,* $C_{mn}=C_{nm}$, *it suffices to show that*

$$c_{ijkl}=c_{klij},\qquad i,j,k,l\in\{1,2,3\}.$$

Recall stress-strain equations (3.1), namely,

$$\sigma_{ij}=c_{ijkl}\varepsilon_{kl},\qquad i,j\in\{1,2,3\},$$

as well as a particular case of these equations that corresponds to isotropic continua and is given by equations (5.65), namely,

$$\sigma_{ij}=\lambda\delta_{ij}\varepsilon_{kk}+2\mu\varepsilon_{ij},\qquad i,j\in\{1,2,3\},$$

where λ *and* μ *are Lamé's parameters. Hence, for isotropic continua, we can write*

$$c_{ijkl}\varepsilon_{kl}-(\lambda\delta_{ij}\varepsilon_{kk}+2\mu\varepsilon_{ij})=0,\qquad i,j\in\{1,2,3\}.$$

[11] See also Section 5.10.1.

Using the properties of Kronecker's delta and the summation convention, we can write, $\varepsilon_{kk} = \delta_{kl}\varepsilon_{kl}$, *and* $\varepsilon_{ij} = \delta_{ik}\delta_{jl}\varepsilon_{kl}$, *where* $i, j \in \{1, 2, 3\}$. *Thus,*

$$\begin{aligned} c_{ijkl}\varepsilon_{kl} - (\lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}) &= [c_{ijkl} - (\lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl})]\,\varepsilon_{kl} \\ &= 0, \qquad i, j \in \{1, 2, 3\}. \end{aligned}$$

Since ε_{kl} *is arbitrary, the expression in brackets must vanish. Hence, we require that*

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl}, \qquad i, j, k, l \in \{1, 2, 3\}.$$

By the commutativity of Kronecker's delta, $\delta_{ij}\delta_{kl} = \delta_{kl}\delta_{ij}$, *while by its symmetry,* $\delta_{ik}\delta_{jl} = \delta_{ki}\delta_{lj}$. *Consequently, we can write*

$$\begin{aligned} c_{ijkl} &= \lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl} \\ &= \lambda\delta_{kl}\delta_{ij} + 2\mu\delta_{ki}\delta_{lj} = c_{klij}, \qquad i, j, k, l \in \{1, 2, 3\}, \end{aligned}$$

as required.

Exercise 5.6 [12] *Using Lemma 5.4, prove Theorem 5.5, stated below.*

Notation 5.2 *Repeated-index summation is used in this exercise. Any term in which an index appears twice stands for the sum of all such terms as the index assumes all the values between* 1 *and* 3.

Lemma 5.4 [13] *The general isotropic fourth-rank tensor is*

$$a_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \xi\delta_{ik}\delta_{jl} + \eta\delta_{il}\delta_{jk}, \qquad i, j, k, l \in \{1, 2, 3\}. \tag{5.92}$$

Theorem 5.5 *Given the symmetry of the strain tensor, defined in expression (1.15), the stress-strain equations for a three-dimensional isotropic continuum are given by expression (5.65), namely,*

$$\sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \qquad i, j \in \{1, 2, 3\},$$

where $2\mu = \xi + \eta$.

Solution 5.6 .

[12] See also Section 5.10.3.

[13] Readers interested in a proof of Lemma 5.4 might refer to Matthews, P.C., (1998) Vector calculus: Springer, pp. 124 – 125.

Proof. Consider stress-strain equations (3.1), namely

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl}, \qquad i,j \in \{1,2,3\}.$$

Inserting expression (5.92) for c_{ijkl}, and using the properties of Kronecker's delta, in view of Lemma 5.4, we can write

$$\begin{aligned}\sigma_{ij} &= (\lambda\delta_{ij}\delta_{kl} + \xi\delta_{ik}\delta_{jl} + \eta\delta_{il}\delta_{jk})\,\varepsilon_{kl} \\ &= \lambda\delta_{ij}\varepsilon_{kk} + \xi\varepsilon_{ij} + \eta\varepsilon_{ji}, \qquad i,j \in \{1,2,3\}.\end{aligned}$$

Since, by its definition, the strain tensor, ε_{kl}, is symmetric, we can write

$$\sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + (\xi + \eta)\,\varepsilon_{ij}, \qquad i,j \in \{1,2,3\},$$

and, hence, there are only two independent constants in the stress-strain equations for an isotropic continuum. Since the constants are arbitrary, we can set $2\mu = \xi + \eta$, and write

$$\sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \qquad i,j \in \{1,2,3\},$$

as required. ■

Remark 5.2 *While studying isotropic materials it is common to express the two elasticity parameters, shown in matrices (5.62) and (5.64), in terms of other parameters that possess an immediate physical meaning. Exercises 5.7 – 5.12 discuss such expressions.*

Remark 5.3 *The relations among Poisson's ratio, Young's modulus and Lamé's parameters are given by*

$$\lambda = \frac{\mathrm{E}\nu}{(1+\nu)(1-2\nu)}, \tag{5.93}$$

and

$$\mu = \frac{\mathrm{E}}{2(1+\nu)}. \tag{5.94}$$

Exercise 5.7 *Consider an isotropic continuum. Defining Poisson's ratio as*

$$\nu := -\frac{\varepsilon_{xx}}{\varepsilon_{zz}} = -\frac{\varepsilon_{yy}}{\varepsilon_{zz}}, \tag{5.95}$$

where we subject the continuum to a uniaxial stress along the z-axis so that $\sigma_{xx} = \sigma_{yy} = \sigma_{xy} = \sigma_{yz} = \sigma_{zx} = 0$, show that Poisson's ratio is given by

$$\nu = \frac{\lambda}{2(\lambda+\mu)}, \tag{5.96}$$

where λ and μ are Lamé's parameters.

Solution 5.7 *Following stress-strain equations (5.65), which describe isotropic continua, we can write*

$$\sigma_{xx} = \lambda\left(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}\right) + 2\mu\varepsilon_{xx} = \left(\lambda + 2\mu\right)\varepsilon_{xx} + \lambda\varepsilon_{yy} + \lambda\varepsilon_{zz} = 0.$$

Dividing both sides by ε_{zz}*, we obtain*

$$\left(\lambda + 2\mu\right)\frac{\varepsilon_{xx}}{\varepsilon_{zz}} + \lambda\frac{\varepsilon_{yy}}{\varepsilon_{zz}} + \lambda = 0.$$

Using the definition of Poisson's ratio, we can rewrite the above expression as

$$-\left(\lambda + 2\mu\right)\nu - \lambda\nu + \lambda = -2\left(\lambda + \mu\right)\nu + \lambda = 0.$$

Hence, solving for ν*, we get*

$$\nu = \frac{\lambda}{2\left(\lambda + \mu\right)},$$

which is expression (5.96), as required.

Exercise 5.8 [14]*Consider an isotropic continuum under a uniaxial stress that leads to small deformations. Using expression (5.96), show that no change in volume implies no resistance to change in shape, as stated by* $\mu = 0$.

Solution 5.8 *Consider a rectangular box with initial dimensions* x_1, x_2, *and* x_3. *Its volume is* $V = x_1x_2x_3$. *Let the dimensions after deformation be* $x_1 + \Delta x_1$, $x_2 + \Delta x_2$, *and* $x_3 + \Delta x_3$, *where, after the deformation, the original rectangular box remains rectangular. Thus, the volume after deformation is*

$$\begin{aligned}\breve{V} &= \left(x_1 + \Delta x_1\right)\left(x_2 + \Delta x_2\right)\left(x_3 + \Delta x_3\right) \\ &\approx x_1x_2x_3 + x_2x_3\Delta x_1 + x_1x_3\Delta x_2 + x_1x_2\Delta x_3,\end{aligned} \tag{5.97}$$

where the approximation stems from the assumption of small deformations and, consequently, from neglecting the second-order and the third-order terms involving Δx_i, *where* $i \in \{1, 2, 3\}$. *No change in volume implies*

$$\breve{V} - V = 0.$$

Using expression (5.97), we can write

$$\breve{V} - V = x_2x_3\Delta x_1 + x_1x_3\Delta x_2 + x_1x_2\Delta x_3 = 0.$$

[14]See also Section 1.3.2.

Dividing both sides by $V = x_1 x_2 x_3$, *we get*

$$\frac{\check{V} - V}{V} = \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} + \frac{\Delta x_3}{x_3} = 0.$$

In view of expression (1.25) and denoting $\varepsilon_{11} := \Delta x_1 / x_1$, $\varepsilon_{22} := \Delta x_2 / x_2$, $\varepsilon_{33} := \Delta x_3 / x_3$, *we obtain*

$$\frac{\check{V} - V}{V} = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} = 0.$$

Dividing both sides by ε_{33} *and invoking the definition of Poisson's ratio, given in expression (5.95), we can write*

$$\frac{\varepsilon_{11}}{\varepsilon_{33}} + \frac{\varepsilon_{22}}{\varepsilon_{33}} + 1 = -\nu - \nu + 1 = 0,$$

which implies that the corresponding Poisson's ratio is $\nu = 1/2$. *Using expression (5.96), we obtain*

$$\mu = \frac{1 - 2\nu}{2\nu} \lambda = 0,$$

as required.

Exercise 5.9 *Using equations (5.65), show that in an isotropic continuum, the strain-tensor components,* ε_{ij}, *can be expressed in terms of the stress-tensor components,* σ_{ij}, *as*

$$\varepsilon_{ij} = \frac{1 + \nu}{\mathrm{E}} \sigma_{ij} - \frac{\nu}{\mathrm{E}} \delta_{ij} \sum_{k=1}^{3} \sigma_{kk}, \qquad i, j \in \{1, 2, 3\}, \tag{5.98}$$

where ν *is Poisson's ratio and* E *is Young's modulus.*

Solution 5.9 *Using expressions (5.93) and (5.94), we can write stress-strain equations (5.65) as*

$$\sigma_{ij} = \frac{\mathrm{E}\nu}{(1 + \nu)(1 - 2\nu)} \delta_{ij} \sum_{k=1}^{3} \varepsilon_{kk} + \frac{\mathrm{E}}{1 + \nu} \varepsilon_{ij}, \qquad i, j \in \{1, 2, 3\}.$$

Solving for ε_{ij}, *we obtain*

$$\varepsilon_{ij} = \frac{1 + \nu}{\mathrm{E}} \sigma_{ij} - \frac{\nu}{1 - 2\nu} \delta_{ij} \sum_{k=1}^{3} \varepsilon_{kk}, \qquad i, j \in \{1, 2, 3\}. \tag{5.99}$$

Now, we seek to express strains $\sum_{k=1}^{3}\varepsilon_{kk}$ in terms of stresses. In view of Kronecker's delta and stress-strain equations (5.65), we can write all stress-tensor components for which $\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk}$ does not vanish. They are

$$\sigma_{ii}=\lambda\sum_{k=1}^{3}\varepsilon_{kk}+2\mu\varepsilon_{ii},\qquad i\in\{1,2,3\}.$$

Writing these three equations explicitly, we get

$$\begin{cases}\sigma_{11}=\lambda\sum\limits_{k=1}^{3}\varepsilon_{kk}+2\mu\varepsilon_{11}\\ \sigma_{22}=\lambda\sum\limits_{k=1}^{3}\varepsilon_{kk}+2\mu\varepsilon_{22}\\ \sigma_{33}=\lambda\sum\limits_{k=1}^{3}\varepsilon_{kk}+2\mu\varepsilon_{33}\end{cases}.$$

Summing these three equations, we obtain

$$\begin{aligned}\sigma_{11}+\sigma_{22}+\sigma_{33}&=3\lambda\sum_{k=1}^{3}\varepsilon_{kk}+2\mu\left(\varepsilon_{11}+\varepsilon_{22}+\varepsilon_{33}\right)\\&=3\lambda\sum_{k=1}^{3}\varepsilon_{kk}+2\mu\sum_{k=1}^{3}\varepsilon_{kk}=(3\lambda+2\mu)\sum_{k=1}^{3}\varepsilon_{kk}.\end{aligned}$$

Expressing the left-hand side as a summation, we can write the sought expression

$$\sum_{k=1}^{3}\varepsilon_{kk}=\frac{\sum\limits_{k=1}^{3}\sigma_{kk}}{3\lambda+2\mu}.\tag{5.100}$$

Using expression (5.100), we can write expression (5.99) as

$$\varepsilon_{ij}=\frac{1+\nu}{\mathrm{E}}\sigma_{ij}-\frac{\nu}{(1-2\nu)(3\lambda+2\mu)}\delta_{ij}\sum_{k=1}^{3}\sigma_{kk},\qquad i,j\in\{1,2,3\}.\tag{5.101}$$

Consider the term in parentheses that contains λ and μ. Using expressions (5.93) and (5.94), we can write this term as

$$3\lambda+2\mu=3\frac{\mathrm{E}\nu}{(1+\nu)(1-2\nu)}+\frac{\mathrm{E}}{1+\nu}=\frac{\mathrm{E}}{1-2\nu}.$$

Hence, expression (5.101) becomes

$$\varepsilon_{ij} = \frac{1+\nu}{\mathrm{E}}\sigma_{ij} - \frac{\nu}{\mathrm{E}}\delta_{ij}\sum_{k=1}^{3}\sigma_{kk}, \qquad i,j \in \{1,2,3\},$$

which is expression (5.98), as required.

Exercise 5.10 *Using expression (4.22), show that for isotropic continua the strain-energy function can be expressed in terms of the strain-tensor components as*

$$W = \frac{\lambda}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\varepsilon_{ii}\varepsilon_{jj} + \mu\sum_{i=1}^{3}\sum_{j=1}^{3}\varepsilon_{ij}\varepsilon_{ij}, \tag{5.102}$$

where λ and μ are Lamé's parameters.

Solution 5.10 *Recall expression (4.22), namely,*

$$W = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\varepsilon_{ij}. \tag{5.103}$$

Also, recall that for an isotropic continuum the stress-strain equations are given by expression (5.65), namely,

$$\sigma_{ij} = \lambda\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \qquad i,j \in \{1,2,3\}. \tag{5.104}$$

Inserting expression (5.104) into expression (5.103), we obtain

$$W = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\left[\left(\lambda\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk} + 2\mu\varepsilon_{ij}\right)\varepsilon_{ij}\right].$$

The properties of Kronecker's delta imply that

$$W = \frac{\lambda}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\varepsilon_{ii}\varepsilon_{jj} + \mu\sum_{i=1}^{3}\sum_{j=1}^{3}\varepsilon_{ij}\varepsilon_{ij},$$

which is expression (5.102), as required.

Exercise 5.11 *Using expression (4.22), show that, for isotropic continua, the strain-energy function can be expressed in terms of the stress-tensor components as*

$$W = \frac{1}{2\mathrm{E}}\left[(1+\nu)\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\sigma_{ij} - \nu\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ii}\sigma_{jj}\right], \tag{5.105}$$

where λ and μ are Lamé's parameters, ν is Poisson's ratio, and E is Young's modulus.

Solution 5.11 *Recall expression (4.22), namely,*

$$W = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\varepsilon_{ij}. \tag{5.106}$$

Also, recall expression (5.98), namely,

$$\varepsilon_{ij} = \frac{1+\nu}{\mathrm{E}}\sigma_{ij} - \frac{\nu}{\mathrm{E}}\delta_{ij}\sum_{k=1}^{3}\sigma_{kk}, \qquad i,j \in \{1,2,3\}, \tag{5.107}$$

where ν and E are Poisson's ratio and Young's modulus, respectively. Inserting expression (5.107) into expression (5.106), we obtain

$$W = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\left[\sigma_{ij}\left(\frac{1+\nu}{\mathrm{E}}\sigma_{ij} - \frac{\nu}{\mathrm{E}}\delta_{ij}\sum_{k=1}^{3}\sigma_{kk}\right)\right].$$

The properties of Kronecker's delta, δ_{ij}, imply

$$W = \frac{1}{2\mathrm{E}}\left[(1+\nu)\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\sigma_{ij} - \nu\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ii}\sigma_{jj}\right],$$

which is expression (5.105), as required.

Exercise 5.12 [15] *Consider elasticity matrix (5.64). Find the range of values for Lamé's parameters that is required by the stability conditions. Express this range in terms of Poisson's ratio. Provide a physical interpretation of this result.*

[15] See also Section 4.3.3.

Solution 5.12 *Stability conditions require the elasticity matrix to be positive-definite. Matrix (5.64) is symmetric. As stated in Theorem 4.3, for the positive-definiteness we require all eigenvalues to be positive. Consider the two submatrices, namely,*

$$\begin{bmatrix} \lambda+2\mu & \lambda & \lambda \\ \lambda & \lambda+2\mu & \lambda \\ \lambda & \lambda & \lambda+2\mu \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix}.$$

We obtain the eigenvalues, Λ_i, by solving

$$\det\left(\begin{bmatrix} \lambda+2\mu & \lambda & \lambda \\ \lambda & \lambda+2\mu & \lambda \\ \lambda & \lambda & \lambda+2\mu \end{bmatrix} - \Lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 0,$$

and

$$\det\left(\mu \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \Lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 0.$$

The eigenvalues are $\Lambda_{1,2} = 2\mu$, $\Lambda_3 = 3\lambda + 2\mu$, and $\Lambda_{4,5,6} = \mu$. Positiveness of the eigenvalues means that $\mu > 0$ and $\lambda > -\frac{2}{3}\mu$. Recalling expression (5.96), we obtain the range of physically acceptable values of Poisson's ratio, namely,

$$\nu \in \left[-1, \frac{1}{2}\right].$$

Physically, for a cylindrical sample and in view of $\nu := -\varepsilon_{xx}/\varepsilon_{zz}$, the negative value of Poisson's ratio implies that the diminishing of the length of the cylinder along the z-axis is accompanied by the shortening of the radius along the x-axis. For most solids, we would expect a more limited range, namely, $\nu \in \left[0, \frac{1}{2}\right]$, where the diminishing of the length is accompanied by the extension of the radius.

Part II

Waves and rays

This Page Intentionally Left Blank

Introduction to Part II

> The solution of the equation of motion for an elastic medium results in the existence of elastic waves in its interior. The wave phenomenon is a way of transporting energy without transport of matter. The propagation of energy is, then, a very important aspect of wave propagation.
>
> *Agustín Udías (1999) Principles of seismology*

In *Part I*, we derived Cauchy's equations of motion, the equation of continuity, and formulated the stress-strain equations for elastic continua. These equations form a determined system, which allows us to describe the behaviour of such continua.

In *Part II*, we combine Cauchy's equations of motion with the stress-strain equations to formulate the equations of motion in elastic continua. In the particular case of isotropic homogeneous continua, these equations are wave equations, which possess analytical solutions. However, in anisotropic inhomogeneous continua, we are unable to formulate equations of motion that possess analytical solutions. Hence, we choose to study these equations in terms of the high-frequency approximation, which results in ray theory. This approach allows us to study rays, wavefronts, traveltimes and amplitudes of signals that propagate within such a continuum. Although the resulting expressions are exact only for the case of an infinitely high frequency of a signal, the experimental results agree well with the theoretical predictions, provided that the properties of the continuum do not change significantly within the wavelength of the signal.

Ray methods form an important theoretical platform for seismological studies. They allow us to formulate problems in the context of such mathematical tools as differential geometry and the calculus of variations. While referring to the ray solution in their volumes on "Quantitative seismology: Theory and methods", Aki and Richards state that

> [it] provides the basis for routine interpretation of most seismic body waves, and it always provides a guide to more sophisticated methods, should they be necessary.

However, in view of this being an approximate solution, we must be aware of its limitations. Grant and West, in their book on "Interpretation theory in applied geophysics", state that

> it is often surprising to observe how uncritically their [ray methods] validity in seismological problems is accepted.

Rays, as a scientific entity, can be traced to the work of Willebrord Snell who, at the turn of the sixteenth and seventeenth century, formulated the law of refraction. The mathematical underpinnings of ray theory were established by William Rowan Hamilton in the first half of the nineteenth century. The formulation of rays in terms of asymptotic series, which is the platform for our studies, is associated with the work of Carl Runge, Arnold Sommerfeld and Peter Debye, at the beginning of twentieth century, as well as Vasiliy M. Babich and Joseph B. Keller in the middle of the twentieth century. Further work, specifically in the context of seismic rays, was done by Vlastislav Červený.

Chapter 6

Equations of motion: Isotropic homogeneous continua

> From the study of nature there arose that class of partial differential equations that is at the present time the most thoroughly investigated and probably the most important in the general structure of human knowledge, namely, the equations of mathematical physics.
>
> *Sergei L. Sobolev and Olga A. Ladyzenskaya (1969) Partial differential equations* in *Mathematics (editors: Aleksandrov, et al.)*

Preliminary remarks

Having formulated system (4.19) — a system of equations to describe the behaviour of an elastic continuum — we wish to write Cauchy's equations of motion explicitly in the context of the stress-strain equations for such a continuum. This way, we commence our study of wave phenomena in an elastic continuum.

We begin by choosing the simplest type of elastic continuum, namely an isotropic homogeneous one, and, hence, we derive the corresponding equations of motion, which lead to the wave equations. In the process of formulating these equations, we learn about the existence of the two types of waves that can propagate in isotropic continua. Furthermore, we obtain the expressions for the speed of these waves as functions of the properties of the continuum.

We begin this chapter by combining Cauchy's equations of motion (2.50) with constitutive equations (5.65). This formulation results in the derivation of the wave equations. To gain insight into these equations, we study them in the context of plane waves and displacement potentials. We also investigate the solutions of the wave equations. We conclude this chapter with examples of extensions of the standard form of the wave equation that take into account aspects of anisotropy and of inhomogeneity.

6.1 Wave equations

6.1.1 Equation of motion

To derive the wave equation, assume that a given three-dimensional continuum is isotropic and homogeneous. Thus, we consider the corresponding stress-strain equations given by expression (5.65), namely,

$$\sigma_{ij} = \lambda\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \qquad i,j \in \{1,2,3\}, \tag{6.1}$$

where λ and μ are constants. We also consider Cauchy's equations of motion (2.50), namely,

$$\sum_{j=1}^{3}\frac{\partial\sigma_{ij}}{\partial x_j} = \rho\frac{\partial^2 u_i}{\partial t^2}, \qquad i \in \{1,2,3\}, \tag{6.2}$$

which do not contain body forces.

We wish to combine stress-strain equations (6.1) with equations of motion (6.2) to get the equations of motion in an isotropic homogeneous continuum. In other words, we substitute expression (6.1) into equations (6.2) to obtain

$$\begin{aligned}\rho\frac{\partial^2 u_i}{\partial t^2} &= \sum_{j=1}^{3}\frac{\partial}{\partial x_j}\left(\lambda\delta_{ij}\sum_{k=1}^{3}\varepsilon_{kk} + 2\mu\varepsilon_{ij}\right) \\ &= \sum_{j=1}^{3}\left(\delta_{ij}\lambda\sum_{k=1}^{3}\frac{\partial\varepsilon_{kk}}{\partial x_j} + 2\mu\frac{\partial\varepsilon_{ij}}{\partial x_j}\right), \qquad i \in \{1,2,3\}.\end{aligned} \tag{6.3}$$

Now, we wish to express the right-hand side of equations (6.3) in terms of the displacement vector, $\mathbf{u}$. Invoking the definition of the strain tensor, given in expression (1.15), we can rewrite equations (6.3) as

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^{3}\left[\delta_{ij}\lambda\sum_{k=1}^{3}\frac{\partial}{\partial x_j}\left(\frac{\partial u_k}{\partial x_k}\right) + \mu\frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right], \qquad i \in \{1,2,3\}.$$

Using the property of Kronecker's delta, we obtain

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \lambda \sum_{k=1}^{3} \frac{\partial}{\partial x_i}\left(\frac{\partial u_k}{\partial x_k}\right) + \mu \sum_{j=1}^{3} \frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad i \in \{1,2,3\}.$$

Using the linearity of the differential operators, we can rewrite these equations as

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \lambda \sum_{j=1}^{3} \frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \sum_{j=1}^{3} \frac{\partial^2 u_i}{\partial x_j^2} + \mu \sum_{j=1}^{3} \frac{\partial^2 u_j}{\partial x_j \partial x_i}, \qquad i \in \{1,2,3\},$$

where, in the first summation, for the summation indices, we let $k = j$. Using the equality of mixed partial derivatives, we obtain

$$\begin{aligned} \rho \frac{\partial^2 u_i}{\partial t^2} &= (\lambda + \mu) \sum_{j=1}^{3} \frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \sum_{j=1}^{3} \frac{\partial^2 u_i}{\partial x_j^2} \\ &= (\lambda + \mu) \frac{\partial}{\partial x_i} \sum_{j=1}^{3} \frac{\partial u_j}{\partial x_j} + \mu \left(\sum_{j=1}^{3} \frac{\partial^2}{\partial x_j^2}\right) u_i, \qquad i \in \{1,2,3\}. \end{aligned} \tag{6.4}$$

We can use vector calculus to concisely state equations (6.4). Consider the right-hand side of these equations. The first summation term is the divergence of $\mathbf{u}$, namely, $\nabla \cdot \mathbf{u}$, while the second summation term is Laplace's operator, namely, ∇^2. Consequently, we can rewrite equations (6.4) as

$$\rho \frac{\partial^2 u_i}{\partial t^2} = (\lambda + \mu) \frac{\partial}{\partial x_i} \nabla \cdot \mathbf{u} + \mu \nabla^2 u_i, \qquad i \in \{1,2,3\}. \tag{6.5}$$

We can explicitly write the three equations stated in expression (6.5) as

$$\rho \frac{\partial^2}{\partial t^2} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = (\lambda + \mu) \begin{bmatrix} \dfrac{\partial (\nabla \cdot \mathbf{u})}{\partial x_1} \\ \dfrac{\partial (\nabla \cdot \mathbf{u})}{\partial x_2} \\ \dfrac{\partial (\nabla \cdot \mathbf{u})}{\partial x_3} \end{bmatrix} + \mu \nabla^2 \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

Noticing that the first matrix on the right-hand side involves the gradient operator, we can concisely state the three equations shown in expression (6.5) as

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u}. \tag{6.6}$$

This is the equation of motion that applies to isotropic homogeneous continua.

We wish to write equation (6.6) in a form that allows us to express it in terms of the dilatation and the rotation vector, in accordance with their definitions stated in Chapter 1. Using the vector identity given by

$$\nabla^2 \mathbf{a} = \nabla (\nabla \cdot \mathbf{a}) - \nabla \times (\nabla \times \mathbf{a}), \tag{6.7}$$

and letting $\mathbf{a} = \mathbf{u}$, we can rewrite equation (6.6) as

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \nabla (\nabla \cdot \mathbf{u}) - \mu \nabla \times (\nabla \times \mathbf{u}). \tag{6.8}$$

Equation (6.8) contains information about the deformations expressed in terms of the divergence and the curl operators. Recalling the definitions of the dilatation and the rotation vector, given by expressions (1.26) and (1.30), respectively, we can immediately write

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \nabla \varphi - \mu \nabla \times \mathbf{\Psi}. \tag{6.9}$$

Equation (6.9) describes the propagation of the deformations in terms of both dilatation and the rotation vector in an isotropic homogeneous continuum. It describes the propagation related to both the change in volume and the change in shape. The divergence operator is associated with the change of volume while the curl operator is associated with the change in shape.

Note that $\nabla^2 \mathbf{u}$ behaves as a vector only with respect to the change of orthonormal coordinates. Consequently, equation (6.6) is valid only for such coordinates. Equation (6.8), however, is valid for curvilinear coordinates.

6.1.2 Wave equation for P waves

To gain insight into the types of waves that propagate in an isotropic homogeneous continuum, we wish to split equation (6.9) into its parts, which are associated with the dilatation and with the rotation vector.

In view of vector-calculus identities, we take the divergence of equation (6.9). Since in a homogeneous continuum λ and μ are constants, we can write

$$\nabla \cdot \left[\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right] = (\lambda + 2\mu) \nabla \cdot \nabla \varphi - \mu \nabla \cdot \nabla \times \mathbf{\Psi}. \tag{6.10}$$

The factor of μ vanishes, since $\nabla \cdot \nabla \times \mathbf{\Psi} = 0$. Also, considering the factor of $\lambda + 2\mu$ and invoking the definition of Laplace's operator, we can write

$$\nabla \cdot \nabla \varphi = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right] \cdot \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right] \varphi = \nabla^2 \varphi.$$

Consequently, equation (6.10) becomes

$$\nabla \cdot \left[\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right] = (\lambda + 2\mu) \nabla^2 \varphi. \tag{6.11}$$

Let us consider the left-hand side of equation (6.11). In a homogeneous continuum, the mass density, ρ, is a constant. Hence — in view on the linearity of the differential operators — we can take ρ outside of such operators. Also — in view of the the equality of mixed partial derivatives — we can interchange time and space derivatives. Thus, we obtain

$$\rho \frac{\partial^2 \varphi}{\partial t^2} = (\lambda + 2\mu) \nabla^2 \varphi,$$

where, on the left-hand side, we again use definition (1.26). Rearranging, we obtain

$$\nabla^2 \varphi = \frac{1}{\frac{\lambda + 2\mu}{\rho}} \frac{\partial^2 \varphi}{\partial t^2}. \tag{6.12}$$

Equation (6.12) is the wave equation for P waves, where the wave function is given by dilatation, $\varphi(\mathbf{x}, t) = \nabla \cdot \mathbf{u}$, with $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$. As shown in Section 6.4,

$$v := \sqrt{\frac{\lambda + 2\mu}{\rho}} \tag{6.13}$$

is the propagation speed. In view of Section 5.10, the presence of both Lamé's parameters in expression (6.13) suggests that P waves subject the continuum to both a change in volume and a change in shape.

In view of definition (1.26), P waves are sometimes referred to as dilatational waves. Also, since the dilatation, φ, is the relative change in volume, they are sometimes referred to as pressure waves. Furthermore, since the speed of P waves is always greater than the speed of S waves, which are discussed below, in earthquake observations, P waves are sometimes referred to as primary waves.

6.1.3 Wave equation for S waves

To obtain the wave equation for S waves, we take the curl of equation (6.9) and get

$$\nabla \times \left[\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right] = (\lambda + 2\mu) \nabla \times \nabla \varphi - \mu \nabla \times \nabla \times \boldsymbol{\Psi}. \tag{6.14}$$

Following vector-calculus identities, the curl of the gradient vanishes and, hence, the first term on the right-hand side disappears. Recalling definition

(1.30) and considering the constancy of the mass density, ρ — in view of the linearity of the differential operators as well as the equality of mixed partial derivatives — we obtain

$$\rho \frac{\partial^2 \boldsymbol{\Psi}}{\partial t^2} = -\mu \nabla \times [\nabla \times \boldsymbol{\Psi}] . \tag{6.15}$$

Invoking vector-calculus identity (6.7) and letting $\mathbf{a} = \boldsymbol{\Psi}$, we can write equation (6.15) as

$$\rho \frac{\partial^2 \boldsymbol{\Psi}}{\partial t^2} = -\mu \left[\nabla (\nabla \cdot \boldsymbol{\Psi}) - \nabla^2 \boldsymbol{\Psi} \right] .$$

In view of definition (1.30) and the vanishing of the divergence of a curl, the first term in brackets disappears. Hence, we obtain

$$\nabla^2 \boldsymbol{\Psi} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \boldsymbol{\Psi}}{\partial t^2} , \tag{6.16}$$

where the wave function is given by the rotation vector, $\boldsymbol{\Psi}(\mathbf{x}, t) = \nabla \times \mathbf{u}$, with $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$. As shown in Section 6.4,

$$v := \sqrt{\frac{\mu}{\rho}} \tag{6.17}$$

is the propagation speed. In view of Section 5.10, the presence of the single Lamé's parameter, namely, μ, in expression (6.17), suggests that S waves subject the continuum to a change in shape. Also, due to the vanishing of rigidity in fluids, we can conclude that the propagation of S waves is limited to solids.

In view of definition (1.30), S waves are sometimes referred to as rotational waves. Since the rotation vector is given by $\boldsymbol{\Psi} = \nabla \times \mathbf{u}$, we conclude that $\nabla \cdot \boldsymbol{\Psi} = 0$. If the divergence of a vector field vanishes, this vector field is volume-preserving; hence, S waves are sometimes referred to as the equivoluminal waves. In English, the justification for the letter S is due to the fact that these waves are often referred to as shear waves. Also, due to the fact that the speed of S waves is always smaller than the speed of P waves, in earthquake observations, S waves are sometimes referred to as secondary waves.

6.2 Plane waves

In general, equations (6.4) are complicated partial differential equations. This shows that even in isotropic homogeneous continua, the description of wave phenomena constitutes a serious mathematical problem. We can simplify these equations by introducing certain abstract mathematical entities that allow us to describe particular aspects of wave phenomena. While studying wave propagation in homogeneous media, we can consider plane waves, namely, waves whose displacements are functions of a single direction of propagation. Notably, in Chapter 11, we will use plane waves to study reflection and transmission of waves at an interface separating two anisotropic homogeneous halfspaces.

Herein, to gain insight into equations (6.12) and (6.16), we will study these equations in the context of plane waves. In view of expression (2.15), let the displacement be given by

$$\mathbf{u} = \left[u_1(x_1, t), u_2(x_1, t), u_3(x_1, t)\right].$$

In other words, let the plane waves propagate along the x_1-axis. Since all the partial derivatives of $\mathbf{u}$ with respect to x_2 and x_3 vanish, equations (6.4) become

$$\rho \frac{\partial^2 u_1}{\partial t^2} = (\lambda + 2\mu) \frac{\partial^2 u_1}{\partial x_1^2},$$

$$\rho \frac{\partial^2 u_2}{\partial t^2} = \mu \frac{\partial^2 u_2}{\partial x_1^2},$$

and

$$\rho \frac{\partial^2 u_3}{\partial t^2} = \mu \frac{\partial^2 u_3}{\partial x_1^2}.$$

After algebraic manipulations, we can write

$$\frac{\partial^2 u_1}{\partial x_1^2} = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 u_1}{\partial t^2}, \tag{6.18}$$

$$\frac{\partial^2 u_2}{\partial x_1^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 u_2}{\partial t^2}, \tag{6.19}$$

and

$$\frac{\partial^2 u_3}{\partial x_1^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 u_3}{\partial t^2}. \tag{6.20}$$

Consider equation (6.18). Recall expression (1.26), which in this case becomes

$$\varphi = \frac{\partial u_1}{\partial x_1}. \tag{6.21}$$

Taking the derivative of equation (6.18) with respect to x_1, we obtain

$$\frac{\partial^3 u_1}{\partial x_1^3} = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^3 u_1}{\partial x_1 \, \partial t^2}. \tag{6.22}$$

Using expression (6.21) in equation (6.22), we obtain

$$\frac{\partial^2 \varphi}{\partial x_1^2} = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 \varphi}{\partial t^2}, \tag{6.23}$$

which is a one-dimensional form of equation (6.12). Examining equations (6.21) and (6.23), we recognize that the displacement and the direction of propagation are parallel to one another, which is the key property of P waves in isotropic continua. This property is also shown in Exercise 10.4.

Now, consider equations (6.19) and (6.20). Recall expression (1.30), which in this case becomes

$$\mathbf{\Psi} = \left[0, -\frac{\partial u_3}{\partial x_1}, \frac{\partial u_2}{\partial x_1}\right]^T . \tag{6.24}$$

Taking the derivative of equations (6.20) and (6.19) with respect to x_1, and writing them as a vector, we obtain

$$\begin{bmatrix} 0 \\ -\dfrac{\partial^3 u_3}{\partial x_1^3} \\ \dfrac{\partial^3 u_2}{\partial x_1^3} \end{bmatrix} = \frac{1}{\frac{\mu}{\rho}} \begin{bmatrix} 0 \\ -\dfrac{\partial^3 u_3}{\partial x_1 \, \partial t^2} \\ \dfrac{\partial^3 u_2}{\partial x_1 \, \partial t^2} \end{bmatrix} .$$

Using the equality of mixed partial derivatives and expression (6.24), we obtain

$$\frac{\partial^2 \mathbf{\Psi}}{\partial x_1^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \mathbf{\Psi}}{\partial t^2}, \tag{6.25}$$

which is a one-dimensional form of equation (6.16). Examining equations (6.24) and (6.25), we recognize that the displacements and the direction of propagation are orthogonal to each other, which is the key property of S waves in isotropic continua. This property is also shown in Exercise 10.6.

Plane waves are an approximation that allows us to study, in homogeneous media, a wavefield that results from a distant source. For close sources, we can construct a wavefield as a superposition of plane waves. In such an approach, there is a constructive interference in the regions where the plane waves coincide and a destructive interference outside of these regions.

While studying inhomogeneous media, the behaviour of seismic waves cannot be conveniently described using plane waves and their superposition. For such studies, we will introduce in Section 6.6.3 another abstract mathematical entity — a seismic ray, which belongs to the realm of asymptotic methods and provides us with a different perspective to study seismic wavefields.

6.3 Displacement potentials

6.3.1 Helmholtz's decomposition

In Sections 6.1.2 and 6.1.3, we derived the wave equations for P and S waves, respectively. We can also obtain equations that correspond to P and S waves by following Helmholtz's method of separating a vector function into its scalar and vector potentials.[1]

A differentiable function $\mathbf{u}(\mathbf{x},t)$ can be decomposed into

$$\mathbf{u}(\mathbf{x},t) = \nabla\mathcal{P} + \nabla\times\mathbf{S}, \tag{6.26}$$

where $\mathcal{P}$ and $\mathbf{S} = [S_1, S_2, S_3]$ are the scalar and vector potentials, respectively. In the context of our study, expression (6.26) means that — upon the passage of a wave — the displacement of an element of the continuum can be written in terms of a scalar potential function, $\mathcal{P}$, and a vector potential function, $\mathbf{S}$.

Expression (6.26) can be explicitly written as

$$\begin{aligned}
u_1(\mathbf{x},t) &= \frac{\partial\mathcal{P}}{\partial x_1} + \frac{\partial S_3}{\partial x_2} - \frac{\partial S_2}{\partial x_3}, \\
u_2(\mathbf{x},t) &= \frac{\partial\mathcal{P}}{\partial x_2} + \frac{\partial S_1}{\partial x_3} - \frac{\partial S_3}{\partial x_1}, \\
u_3(\mathbf{x},t) &= \frac{\partial\mathcal{P}}{\partial x_3} + \frac{\partial S_2}{\partial x_1} - \frac{\partial S_1}{\partial x_2},
\end{aligned}$$

which constitute three equations for four unknowns, namely, $\mathcal{P}$, S_1, S_2 and S_3. In this formulation, to obtain a determined system of equations, we also require

$$\nabla\cdot\mathbf{S} = 0. \tag{6.27}$$

[1]Readers interested in Helmholtz's theorem might refer to Arfken, G.B, and Weber, H.J., (2001) Mathematical methods for physicists (5th edition): Harcourt/Academic Press, pp. 96 - 101.

To justify equation (6.27), consider two vector identities. The first identity, namely,

$$\nabla^2 \mathbf{A} = \nabla (\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}),$$

was introduced in expression (6.7), while the second identity is

$$\nabla \cdot (\nabla \times \mathbf{A}) = 0,$$

where $\mathbf{A}(\mathbf{x})$ is a vector function. Let $\mathbf{w} = \nabla^2 \mathbf{A}$, $\mathcal{P} = \nabla \cdot \mathbf{A}$ and $\mathbf{S} = -\nabla \times \mathbf{A}$. We can rewrite these two vector identities as

$$\mathbf{w} = \nabla \mathcal{P} + \nabla \times \mathbf{S},$$

and

$$\nabla \cdot \mathbf{S} = 0,$$

which are expressions (6.26) and (6.27), respectively, with $\mathbf{w} \equiv \mathbf{u}$. We can always set $\nabla \cdot \mathbf{S} = 0$ since $\mathbf{S}$ is arbitrary up to a gradient ∇f, where $f(\mathbf{x})$ is an arbitrary function. To see that, let

$$\tilde{\mathbf{S}} = \mathbf{S} + \nabla f. \tag{6.28}$$

Taking the curl of both sides of equation (6.28), using the linearity of the differential operator and the vanishing of the curl of a gradient, we obtain

$$\nabla \times \tilde{\mathbf{S}} = \nabla \times (\mathbf{S} + \nabla f) = \nabla \times \mathbf{S}.$$

In other words, the same $\mathbf{u}$ is obtained using either $\mathbf{S}$ or $\mathbf{S} + \nabla f$. Now, let us take the divergence of both sides of equation (6.28). We obtain

$$\nabla \cdot \tilde{\mathbf{S}} = \nabla \cdot (\mathbf{S} + \nabla f(\mathbf{x})) = \nabla \cdot \mathbf{S} + \nabla^2 f. \tag{6.29}$$

Examining equation (6.29), we see that we can always choose an arbitrary function $f(\mathbf{x})$ such that $\nabla^2 f = -\nabla \cdot \mathbf{S}$, which results in $\nabla \cdot \tilde{\mathbf{S}} = 0$. Thus, if both $\mathcal{P}$ and $\mathbf{S}$ are derived from a common vector function, $\mathbf{A}(\mathbf{x})$, we can write a determined system of equations consisting of equations (6.26) and (6.27).

6.3.2 Equation of motion

To implement displacement potentials in the equations of motion, we wish to write these equations in terms of $\mathbf{u}$ given by expression (6.26).

Inserting expression (6.26) into equation (6.6), we can write

$$\rho\frac{\partial^2\left(\nabla\mathcal{P}+\nabla\times\mathbf{S}\right)}{\partial t^2}=(\lambda+\mu)\,\nabla\left[\nabla\cdot\left(\nabla\mathcal{P}+\nabla\times\mathbf{S}\right)\right]+\mu\nabla^2\left(\nabla\mathcal{P}+\nabla\times\mathbf{S}\right).$$

Using the vanishing of the divergence of a curl and the definition of Laplace's operator, we obtain

$$\begin{aligned}\rho\frac{\partial^2\left(\nabla\mathcal{P}+\nabla\times\mathbf{S}\right)}{\partial t^2}&=(\lambda+\mu)\,\nabla\left(\nabla\cdot\nabla\mathcal{P}\right)+\mu\nabla^2\left(\nabla\mathcal{P}+\nabla\times\mathbf{S}\right)\\&=(\lambda+\mu)\,\nabla\left(\nabla^2\mathcal{P}\right)+\mu\nabla^2\left(\nabla\mathcal{P}+\nabla\times\mathbf{S}\right).\end{aligned}$$

Using the linearity of the differential operators and the fact that in a homogeneous continuum ρ, λ and μ are constants, as well as using the equality of mixed partial derivatives, we can rewrite this equation as

$$\begin{aligned}\nabla\left(\rho\frac{\partial^2\mathcal{P}}{\partial t^2}\right)+\nabla\times\left(\rho\frac{\partial^2\mathbf{S}}{\partial t^2}\right)&=\nabla\left[(\lambda+\mu)\,\nabla^2\mathcal{P}\right]+\nabla\left(\mu\nabla^2\mathcal{P}\right)+\nabla\times\left(\mu\nabla^2\mathbf{S}\right)\\&=\nabla\left[(\lambda+2\mu)\,\nabla^2\mathcal{P}\right]+\nabla\times\left(\mu\nabla^2\mathbf{S}\right).\end{aligned}$$

Rearranging, we obtain

$$\nabla\left[(\lambda+2\mu)\,\nabla^2\mathcal{P}-\rho\frac{\partial^2\mathcal{P}}{\partial t^2}\right]+\nabla\times\left[\mu\nabla^2\mathbf{S}-\rho\frac{\partial^2\mathbf{S}}{\partial t^2}\right]=\mathbf{0},\tag{6.30}$$

which is a form of the equation of motion in terms of the scalar and vector potentials.

6.3.3 P and S waves

We wish to study the relation of the scalar and vector potential to the two types of waves that propagate in an isotropic homogeneous continuum.

As in Section 6.1.2, let us take the divergence of equation (6.30). Using the vanishing of the divergence of a curl and the definition of Laplace's operator, we obtain

$$\nabla^2\left[(\lambda+2\mu)\,\nabla^2\mathcal{P}-\rho\frac{\partial^2\mathcal{P}}{\partial t^2}\right]=0.\tag{6.31}$$

Using the linearity of the differential operator and the equality of mixed partial derivatives, where $\left(\nabla^2\right)\left(\partial^2/\partial t^2\right)=\left(\partial^2/\partial t^2\right)\left(\nabla^2\right)$, we can rewrite equation (6.31) as

$$(\lambda+2\mu)\,\nabla^2\left(\nabla^2\mathcal{P}\right)-\rho\frac{\partial^2\left(\nabla^2\mathcal{P}\right)}{\partial t^2}=0.\tag{6.32}$$

To relate expression (6.26) to the dilatation, φ, let us take the divergence of expression (6.26). Using the vanishing of the divergence of a curl and recalling definition (1.26) as well as the definition of Laplace's operator, we obtain

$$\varphi := \nabla \cdot \mathbf{u} = \nabla \cdot \nabla \mathcal{P} \equiv \nabla^2 \mathcal{P}. \tag{6.33}$$

In other words, the dilatation is equal to the Laplacian of the scalar potential. Using expression (6.33), we can rewrite equation (6.32) as

$$\nabla^2 \varphi = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 \varphi}{\partial t^2}, \tag{6.34}$$

which is equation (6.12). Thus, we conclude that the Laplacian of the scalar potential, $\mathcal{P}$, satisfies the wave equation for P waves.

As in Section 6.1.3, let us take the curl of equation (6.30). Using the vanishing of the curl of a gradient, we get

$$\nabla \times \nabla \times \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right) = \mathbf{0}.$$

Following identity (6.7) and letting $\mathbf{a}$ denote the term in parentheses, we can rewrite this equation as

$$\nabla \left[\nabla \cdot \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right) \right] = \nabla^2 \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right).$$

Using the linearity of the differential operators and the equality of mixed partial derivatives, we can rewrite the left-hand side of this equation to obtain

$$\nabla \left[\mu \nabla^2 (\nabla \cdot \mathbf{S}) - \rho \frac{\partial^2 (\nabla \cdot \mathbf{S})}{\partial t^2} \right] = \nabla^2 \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right). \tag{6.35}$$

In view of equation (6.27), equation (6.35) becomes

$$\nabla^2 \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right) = 0. \tag{6.36}$$

Using the linearity of the differential operator, ∇^2, and the equality of mixed partial derivatives, we can write equation (6.36) as

$$\mu \nabla^2 \left(\nabla^2 \mathbf{S} \right) - \rho \frac{\partial^2 \left(\nabla^2 \mathbf{S} \right)}{\partial t^2} = 0. \tag{6.37}$$

To relate expression (6.26) to the rotation vector, $\boldsymbol{\Psi}$, let us take the curl of expression (6.26). Using the vanishing of the curl of a gradient and recalling definition (1.30), we obtain

$$\boldsymbol{\Psi} := \nabla \times \mathbf{u} = \nabla \times \nabla \times \mathbf{S}.$$

Recalling identity (6.7) and letting $\mathbf{a} = \mathbf{S}$, we get

$$\boldsymbol{\Psi} = \nabla \times (\nabla \times \mathbf{S}) = \nabla (\nabla \cdot \mathbf{S}) - \nabla^2 \mathbf{S}.$$

In view of equation (6.27), we obtain

$$\boldsymbol{\Psi} = -\nabla^2 \mathbf{S}. \tag{6.38}$$

In other words, the rotation vector is equal to negative of the Laplacian of the vector potential. Using expression (6.38), we can rewrite equation (6.37) as

$$\nabla^2 \boldsymbol{\Psi} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \boldsymbol{\Psi}}{\partial t^2}, \tag{6.39}$$

which is equation (6.16). Thus, we conclude that under condition (6.27) the Laplacian of the vector potential, $\mathbf{S}$, satisfies the wave equation for S waves.

6.4 Solutions of one-dimensional wave equation

To gain further insights into the physical meaning of equations (6.12) and (6.16), we study the solution of their generic form, where we do not specify if the wave function corresponds to P waves or to S waves.[2] Consider the initial-value problem given by

$$\frac{\partial^2 u(x,t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u(x,t)}{\partial t^2} = 0, \tag{6.40}$$

where $u = u(x,t)$ is the wave function and v is a constant. Let the initial conditions be stated by

$$\left\{ \begin{array}{l} u(x,t)\big|_{t=0} = \gamma(x) \\ \\ \left.\dfrac{\partial u(x,t)}{\partial t}\right|_{t=0} = \eta(x) \end{array} \right. . \tag{6.41}$$

[2]Applying Newton's second law of motion, we can derive equation (6.40) for either longitudinal waves or transverse waves, which correspond to P waves or S waves, respectively. Readers interested in such a derivation of the one-dimensional wave equation might refer to Hanna, J.R., (1982) Fourier series and integrals of boundary value problems: John Wiley and Sons, pp. 109 – 111 and pp. 121 – 122.

The following method of solving the wave equation was introduced in 1746 by d'Alembert and further elaborated upon by Euler, with important contributions from Daniel Bernoulli and Lagrange.[3] It is based on the following two lemmas.

Lemma 6.1 *Equation*

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0,$$

is equivalent to

$$\frac{\partial^2 u\,(y,z)}{\partial y \partial z} = 0,$$

where the new coordinates are

$$\left\{ \begin{array}{l} y = x + vt \\ z = x - vt \end{array} \right. . \tag{6.42}$$

Details of the derivation of Lemma 6.1 are shown in Exercise 6.1.

The form $\partial^2 u\,(y,z)\,/\partial y \partial z = 0$ is a normal form of the hyperbolic differential equation, where y and z are referred to as the natural coordinates, and the straight lines, $y = x + vt$ and $z = x - vt$, in the xt-plane are known as the characteristics.[4]

Lemma 6.2 *For equation*

$$\frac{\partial^2 u\,(y,z)}{\partial y \partial z} = 0,$$

the only form of the solution is

$$u\,(y,z) = f\,(y) + g\,(z)\,, \tag{6.43}$$

where f and g are arbitrary functions.

[3] Readers interested in the history of deriving the wave equation including disagreements among d'Alembert, Euler, Bernoulli and Lagrange in accommodating the initial conditions might refer to Kline, M., (1972) Mathematical thought from ancient to modern times: Oxford University Press, Vol. II, pp. 503 - 514.

[4] Readers interested in normal forms of the hyperbolic equations and its association with characteristics might refer to Morse P.M., and Feshbach H., (1953) Methods of theoretical physics: McGraw-Hill, Inc., Part I, pp. 682 - 683.

Readers interested in the characteristics and their significance in wave theory might refer to Musgrave, M.J.P., (1970) Crystal acoustics: Introduction to the study of elastic waves and vibrations in crystals: Holden-Day, pp. 68 - 76.

Details of the derivation of Lemma 6.2 are shown in Exercise 6.2.

Combining Lemma 6.1 and Lemma 6.2, we can state the following corollary.

Corollary 6.1 *Following Lemma 6.1 and Lemma 6.2, and using coordinates (6.42), we can write the solution of equation (6.40) as*

$$u(x,t) = f(x+vt) + g(x-vt), \tag{6.44}$$

where f and g are arbitrary functions.

Solution (6.44) allows arbitrary functions f and g. Further constraints must be imposed on functions f and g if we wish to obtain a particular solution.

Herein, we wish to obtain a particular form of solution (6.44) that satisfies the constraints provided by initial conditions (6.41). Inserting expression (6.44) into system of equations (6.41), we can write

$$\left\{ \begin{array}{l} f(x) + g(x) = \gamma(x) \\ vf'(x) - vg'(x) = \eta(x) \end{array} \right. ,$$

where, in view of $t=0$, $f'(x) \equiv \mathrm{d}f(x)/\mathrm{d}x$ and $g'(x) \equiv \mathrm{d}g(x)/\mathrm{d}x$. This system of equations can be solved explicitly for $f(x)$ and $g(x)$. Integrating both sides of the second equation of this system, we obtain

$$\left\{ \begin{array}{l} f(x) + g(x) = \gamma(x) \\ \\ f(x) - g(x) = \frac{1}{v} \int\limits_{x_0}^{x} \eta(\zeta)\,\mathrm{d}\zeta \end{array} \right. .$$

Adding the two equations together, we get

$$f(x) = \frac{1}{2}\left[\gamma(x) + \frac{1}{v}\int\limits_{x_0}^{x} \eta(\zeta)\,\mathrm{d}\zeta\right], \tag{6.45}$$

while subtracting the second equation from the first one gives us

$$g(x) = \frac{1}{2}\left[\gamma(x) - \frac{1}{v}\int\limits_{x_0}^{x} \eta(\zeta)\,\mathrm{d}\zeta\right]. \tag{6.46}$$

Inserting expressions (6.45) and (6.46) into solution (6.44), we obtain

$$u(x,t)=\frac{1}{2}\left[\gamma(x+vt)+\gamma(x-vt)+\frac{1}{v}\int\limits_{x-vt}^{x+vt}\eta(\zeta)\,\mathrm{d}\zeta\right], \tag{6.47}$$

where we use the fact that reversing the limits of integration changes the sign of the integral.

Expression (6.47) is the solution of the original initial-value problem given by equations (6.40) and (6.41). Viewing x as the position variable and t as the time variable, solution (6.47) corresponds to the propagation of the shape given by function γ in a one-dimensional x-space.

To illustrate the process of propagation, consider expression (6.47) and let the x-axis be horizontal. At a time $t=0$, $\gamma=\gamma(x)$. At a later time, $\gamma=\gamma(x\pm vt)$. This means that γ has moved along the x-axis by a distance vt. This motion is both to the left and right of the original point x. Since t stands for time, the constant v in equation (6.40) stands for the speed of propagation. This illustration is the reason why the equations of the form $v^2\nabla^2 u=\partial^2 u/\partial t^2$ are called wave equations.

6.5 Reduced wave equation

Since the wave equation is a partial differential equation, to solve it we often assume a trial solution. For instance, while studying three-dimensional continua, it is common to assume a harmonic plane-wave solution, which we use in Section 6.6.3. However, we might also require a more complicated position dependence of the solution.

To illustrate a formulation that allows us to study such a position dependence, we consider equation (6.40), namely,

$$\frac{\partial^2 u(x,t)}{\partial x^2}-\frac{1}{v^2}\frac{\partial^2 u(x,t)}{\partial t^2}=0.$$

If we assume an oscillatory motion, we can write a trial solution that is given by

$$u(x,t)=\acute{u}(x)\exp(-i\omega t), \tag{6.48}$$

where ω stands for the angular frequency.[5] Thus, it is assumed that the time dependence of the displacement function $u(x,t)$ is satisfied by $\exp(-i\omega t)=\cos(\omega t)-i\sin(\omega t)$. In other words, while studying the position dependence

[5] In this book, $\exp(\cdot)$ and $e^{(\cdot)}$ are used as synonymous notations.

of the solution of the wave equation, we assume that this solution is sinusoidal in time.

Inserting solution (6.48) into equation (6.40), as shown in Exercise 6.3, we obtain

$$\frac{\mathrm{d}^2 \acute{u}(x)}{\mathrm{d}x^2} + \left(\frac{\omega}{v}\right)^2 \acute{u}(x) = 0, \tag{6.49}$$

which is referred to as the reduced wave equation.

Considering a three-dimensional continuum, we can write trial solution (6.48) as

$$u(\mathbf{x}, t) = \acute{u}(\mathbf{x}) \exp(-i\omega t),$$

where $\mathbf{x} = [x_1, x_2, x_3]$. In view of equation (6.49), we can write

$$\frac{\partial^2 \acute{u}(\mathbf{x})}{\partial x_1^2} + \frac{\partial^2 \acute{u}(\mathbf{x})}{\partial x_2^2} + \frac{\partial^2 \acute{u}(\mathbf{x})}{\partial x_3^2} + \left(\frac{\omega}{v}\right)^2 \acute{u}(\mathbf{x}) = 0,$$

which, using Laplace's operator, we can concisely state as

$$\nabla^2 \acute{u}(\mathbf{x}) + \left(\frac{\omega}{v}\right)^2 \acute{u}(\mathbf{x}) = 0. \tag{6.50}$$

Thus, following the assumption of oscillatory motion, the wave equation, which belongs to the class of hyperbolic partial differential equations, is transformed into equation (6.50), which belongs to the class of elliptic partial differential equations. Since for elliptical partial differential equations there is no time dependence, equation (6.50) allows us study complicated position dependences without dealing with the time dependence.

6.6 Extensions of wave equation

Introductory comments

In Chapter 7, we will derive equations of motion in anisotropic inhomogeneous continua. This is accomplished by combining Cauchy's equations of motion with stress-strain equations for generally anisotropic continua and allowing the elasticity parameters to be functions of position. The fundamental derivation shown in Chapter 7 lies at the root of ray theory, which is subsequently studied in this book.

There are, however, certain cases where the standard wave equation, which is derived for isotropic homogeneous continua, can be extended to account for anisotropy and for inhomogeneity. An investigation of such cases is undertaken in this section.

6.6.1 Standard wave equation

In multidimensional continua, wave equation (6.40), may be written as

$$\nabla^2 u(\mathbf{x},t) - \frac{1}{v^2}\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} = 0, \tag{6.51}$$

which is a partial differential equation with constant coefficients, where, as shown in Section 6.4, constant v is the magnitude of the velocity of the solution. In equation (6.51), $\mathbf{x}$ are the position coordinates. Hence, this equation describes wave propagation in continua characterized by constant speed at all positions $\mathbf{x}$ and in all directions determined by the coordinates. Consequently, this wave equation is valid for isotropic homogeneous continua.

We wish to extend equation (6.51) to the anisotropic case. In certain cases, by transforming the coordinates, we can formulate a wave equation that in homogeneous continua associates different velocities with different directions. An example of such an extension, which results in a wave equation for elliptical velocity dependence, is illustrated in Section 6.6.2.

We also wish to extend equation (6.51) to the inhomogeneous case. By considering the position dependence $v = v(\mathbf{x})$ and assuming that function $v(\mathbf{x})$ varies slowly with $\mathbf{x}$, we can use an approximation that allows us to describe wave propagation in weakly inhomogeneous continua. This extension of equation (6.51) to account for weak inhomogeneity is illustrated in Section 6.6.3 and belongs to the high-frequency approximation.

6.6.2 Wave equation and elliptical velocity dependence

Wave equation

To study an extension of the wave equation to anisotropic cases, consider equation (6.51). For convenience, let v be equal to unity. Hence, we can write

$$\nabla^2 u(\mathbf{x},t) = \frac{\partial^2 u(\mathbf{x},t)}{\partial t^2}. \tag{6.52}$$

Consider a two-dimensional continuum that is contained in the xz-plane. For $\mathbf{x} = [x,z]$, equation (6.52) can be explicitly written as

$$\frac{\partial^2 u(x,z,t)}{\partial x^2} + \frac{\partial^2 u(x,z,t)}{\partial z^2} = \frac{\partial^2 u(x,z,t)}{\partial t^2}. \tag{6.53}$$

Let the linear transformation of the position coordinates be such that

$$\acute{u}(x,z,t) = u\left(\frac{x}{v_x}, \frac{z}{v_z}, t\right), \tag{6.54}$$

where v_x and v_z are constants. Using the chain rule, as shown in Exercise 6.4, we can write equation (6.53) as

$$v_x^2 \frac{\partial^2 \acute{u}(x,z,t)}{\partial x^2} + v_z^2 \frac{\partial^2 \acute{u}(x,z,t)}{\partial z^2} = \frac{\partial^2 \acute{u}(x,z,t)}{\partial t^2}. \tag{6.55}$$

Thus, function $\acute{u}$ is the solution of equation (6.55).

To illustrate the meaning of constants v_x and v_z, consider transformation (6.54) and let

$$u\left(\frac{x}{v_x}, \frac{z}{v_z}, t\right) := u(\xi, \varsigma, t).$$

If point (ξ, ς) is moving in the $\xi\varsigma$-plane at the unit speed, namely,

$$\frac{\mathrm{d}}{\mathrm{d}t}\sqrt{\xi^2 + \varsigma^2} = 1,$$

the solutions $u(\xi, \varsigma, t)$, at different times t, are concentric circles. It follows that, in the xz-plane,

$$\frac{\mathrm{d}}{\mathrm{d}t}\sqrt{\frac{x^2}{v_x^2} + \frac{z^2}{v_z^2}} = 1,$$

and, hence, the solutions $\acute{u}(x,z,t)$, at different times t, are ellipses.

Equation (6.55) is the wave equation that describes the wavefront propagation in a two-dimensional homogeneous continuum where the wave is subjected to an elliptical velocity dependence with direction. The semiaxes of the elliptical wavefronts coincide with the coordinate axes and the magnitudes of the wavefront velocities along these axes are given by v_x and v_z, respectively.

Phase velocity

Knowing that v_x and v_z are the magnitudes of the wavefront velocities along the x-axis and the z-axis, respectively, we wish to find the expression for the wavefront velocity in an arbitrary direction. Since the wavefronts are loci of constant phase, the wavefront velocity is referred to as phase velocity.

To solve equation (6.55), consider the trial solution given by

$$\acute{u}(x,z,t) = \exp\left[i\omega\left(p_x x + p_z z - t\right)\right], \tag{6.56}$$

where the right-hand side is called the phase factor. Hence, assuming monochromatic waves, where a given value of ω is constant, loci of constant phase are given by the constancy of the term in parentheses. Thus, wavefronts at time t are straight lines $p_x x + p_z z = t$, where p_x and p_z are the components of vector $\mathbf{p}$ that is normal to the a given wavefront. Since x and z have units of distance while t is time, it follows that the units of the components of $\mathbf{p}$ are the units of slowness. In other words, $\mathbf{p}$ is the phase-slowness vector, which describes the slowness with which the wavefront propagates. The envelope of all straight lines $p_x x + p_z z = t$ at time t is an elliptical wavefront. Hence, $\mathbf{p}$ describes the slowness with which the line tangent to the elliptical wavefront propagates.

To examine trial solution (6.56), we substitute it into wave equation (6.55). We obtain

$$v_x^2 \omega^2 p_x^2 \exp\left[i\omega\left(xp_x + zp_z - \omega t\right)\right] + v_z^2 \omega^2 p_z^2 \exp\left[i\omega\left(xp_x + zp_z - \omega t\right)\right] \\ = \omega^2 \exp\left[i\omega\left(xk_x + zk_z - t\right)\right].$$

Dividing by ω^2 and by the exponential term, we can write this equation as

$$v_x^2 p_x^2 + v_z^2 p_z^2 = 1, \tag{6.57}$$

where v_x and v_z are the magnitude of the phase velocity along the horizontal and vertical axes, respectively, while p_x and p_z are the components of $\mathbf{p}$ at a given point on the wavefront.

To state expression (6.57) as a function of the orientation of the wavefront, we can express the phase-slowness vector as

$$\mathbf{p} = \left[p_x, p_z\right] = \left[p\left(\vartheta\right)\sin\vartheta, p\left(\vartheta\right)\cos\vartheta\right], \tag{6.58}$$

where $p\left(\vartheta\right)$ stands for the magnitude of the phase-slowness vector in a given direction ϑ, which is measured between the wavefront normal and the z-axis, and is referred to as the phase angle. Using expression (6.58), we can rewrite expression (6.57) as

$$\left[p\left(\vartheta\right)\right]^2 \left(v_x^2 \sin^2\vartheta + v_z^2 \cos^2\vartheta\right) = 1. \tag{6.59}$$

Since the magnitude of phase slowness is the reciprocal of the magnitude of phase velocity, expression (6.59) can be restated as

$$v\left(\vartheta\right) = \frac{1}{p\left(\vartheta\right)} = \sqrt{v_x^2 \sin^2\vartheta + v_z^2 \cos^2\vartheta}. \tag{6.60}$$

Expression (6.60) gives the magnitude of phase velocity as a function of phase angle for the case of elliptical velocity dependence. As shown in

Exercise 10.8, SH waves in transversely isotropic continua are characterized by elliptical velocity dependence.

Thus, by a linear transformation of the coordinate axes, we obtained an exact formulation of a wave equation for the elliptical velocity dependence. A more sophisticated manipulation of coordinates might allow us to consider wave equations to study complicated anisotropic behaviours in homogeneous continua. In this book, however, we will not pursue this approach. Rather, in Chapter 7, we will formulate an approximation to the wave equation that is valid for generally anisotropic continua.

6.6.3 Wave equation and weak inhomogeneity

Weak-inhomogeneity formulation

To study an extension of the wave equation to the inhomogeneous case, consider equation (6.51), namely,

$$\nabla^2 u(\mathbf{x}, t) - \frac{1}{v^2}\frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0, \tag{6.61}$$

which is valid for homogeneous continua with v being a constant denoting the speed of propagation. In order to extend this equation to inhomogeneous continua, we wish to express v as a function of the position coordinates, $\mathbf{x}$. Consequently, we wish to consider the equation given by

$$\nabla^2 u(\mathbf{x}, t) - \frac{1}{[v(\mathbf{x})]^2}\frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0. \tag{6.62}$$

Since equation (6.62) is a differential equation, it corresponds to local properties of the continuum and can be locally solved for a given $\mathbf{x}$. We can also obtain an approximate global solution to equation (6.62) if we assume that function $v(\mathbf{x})$ varies slowly, which means that the inhomogeneity of a continuum is weak. In the seismological context, weak inhomogeneity means that the changes of properties within a single wavelength are negligible.

To formulate a trial solution of equation (6.62), consider the fact that we can write a trial solution of equation (6.61) as

$$u(\mathbf{x}, t) = A \exp[i\omega(\mathbf{p} \cdot \mathbf{x} - t)], \tag{6.63}$$

where A is the amplitude of the displacement. As stated in Section 6.6.2, $\exp[\cdot]$ is the phase factor, which is constant for a wavefront at time t. In three-dimensional continua, trial solution (6.63) is called the plane-wave solution since, for a given time t, $\mathbf{p} \cdot \mathbf{x} = t$ is a plane that corresponds to a moving wavefront. Vector $\mathbf{p}$ is normal to this plane and, as shown in Section 6.6.2, $\mathbf{p}$ is the phase-slowness vector.

If the properties of a three-dimensional continuum vary with position, a planar wavefront is distorted during propagation through this continuum. Consequently, a trial solution of equation (6.62) must account for these changes of shape of the wavefront, which also cause changes of amplitude along the wavefront. Using a form analogous to expression (6.63), we can write

$$u(\mathbf{x}, t) = A(\mathbf{x}) \exp\{i\omega [\psi(\mathbf{x}) - t]\}, \tag{6.64}$$

where $A(\mathbf{x})$ denotes the amplitude of the displacement — which is allowed to vary along the wavefront — and $\psi(\mathbf{x})$ — referred to as the eikonal function — accounts for the distortions in the shape of the wavefront. Herein, both $A(\mathbf{x})$ and $\psi(\mathbf{x})$ are smooth scalar functions of position coordinates.

Note that expression (6.64) is a zeroth-order term of the asymptotic series given by[6]

$$u(\mathbf{x}, t) \sim \sum_{n=0}^{N} \frac{u_n(\mathbf{x})}{(i\omega)^n} \exp\{i\omega [\psi(\mathbf{x}) - t]\},$$

where $A(\mathbf{x}) \equiv u_0(\mathbf{x})$.[7] Hence, the following results belong to the realm of asymptotic methods, which play an important role in seismology. In this book, however, we do not explicitly discuss ray theory in the context of asymptotic methods.

Examining the phase factor of trial solution (6.64) in the context of solutions (6.56) and (6.63), we see that equation $\psi(\mathbf{x}) = t$ represents the moving wavefront. In other words, the level sets of function $\psi(\mathbf{x})$ are the

[6]Readers interested in the motivation for choosing this form of the trial solution might refer to Babich, V.M., and Buldyrev, V.S., (1991) Short-wavelength diffraction theory: Asymptotic methods: Springer-Verlag, pp. 10 – 13, to Bleistein, N., Cohen, J.K., and Stockwell, J.W., (2001) Mathematics of multidimensional seismic imaging, migration, and inversion: Springer-Verlag, pp. 436 – 437, and to Kennett, B.L.N., (2001) The seismic wavefield, Vol. I: Introduction and theoretical development: Cambridge University Press, pp. 153 – 154 and 166 – 167.

[7]For a description of the nature of asymptotic expansions, as well as the ways of obtaining them by the method of steepest descent and the method of stationary phase, readers might refer to Jeffreys, H., and Jeffreys, B., (1946/1999) Methods of Mathematical Physics: Cambridge University Press, pp. 498 – 507.

For a discussion on an application of asymptotic series to ray theory, readers might refer to Kravtsov, Y.A., and Orlov, Y.I., (1990) Geometrical optics of inhomogeneous media: Springer-Verlag, pp. 7 – 9.

wavefronts. Since $\mathbf{p}$ is normal to the wavefront, using properties of the gradient, we obtain an important expression, namely,

$$\mathbf{p} = \nabla\psi. \tag{6.65}$$

In other words, the phase-slowness vector is the gradient of the eikonal function.

Now we insert trial solution (6.64) into equation (6.62). Considering the x_1 component of Laplace's operator and substituting the corresponding form of trial solution (6.64), we obtain

$$\begin{aligned}\frac{\partial^2}{\partial x_1^2} A(x_1) \exp\{i\omega[\psi(x_1) - t]\} = \exp\{i\omega[\psi(x_1) - t]\} \\ \left\{\frac{\partial^2 A}{\partial x_1^2} + i\omega\left[2\frac{\partial A}{\partial x_1}\frac{\partial \psi}{\partial x_1} + A\frac{\partial^2 \psi}{\partial x_1^2}\right]\right. \\ \left. -\omega^2 A\frac{\partial \psi}{\partial x_1}\frac{\partial \psi}{\partial x_1}\right\}.\end{aligned}$$

Considering the second derivative with respect to time and substituting the same form of the trial solution, we get

$$\frac{\partial^2}{\partial t^2} A(x_1) \exp\{i\omega[\psi(x_1) - t]\} = -A\omega^2 \exp\{i\omega[\psi(x_1) - t]\}.$$

Consequently, given the fact that the exponential term is never zero, the corresponding form of equation (6.62) becomes

$$\frac{\partial^2 A}{\partial x_1^2} + A\omega^2\left(\frac{1}{v^2} - \frac{\partial \psi}{\partial x_1}\frac{\partial \psi}{\partial x_1}\right) + i\omega\left[2\frac{\partial A}{\partial x_1}\frac{\partial \psi}{\partial x_1} + A\frac{\partial^2 \psi}{\partial x_1^2}\right] = 0, \tag{6.66}$$

which is a complex-valued function of real variables.

The vanishing of expression (6.66), where both A and ψ are assumed to be real, implies the vanishing of both real and imaginary parts. Assuming $\omega \neq 0$, we obtain a system of two equations,

$$\left\{\begin{array}{l}\dfrac{\partial^2 A}{\partial x_1^2} + A\omega^2\left(\dfrac{1}{v^2} - \dfrac{\partial \psi}{\partial x_1}\dfrac{\partial \psi}{\partial x_1}\right) = 0 \\[2ex] 2\dfrac{\partial A}{\partial x_1}\dfrac{\partial \psi}{\partial x_1} + A\dfrac{\partial^2 \psi}{\partial x_1^2} = 0\end{array}\right. . \tag{6.67}$$

Considering three-dimensional continua, and following the definitions of the gradient operator and Laplace's operator, we can write system (6.67) as

$$\begin{cases} \nabla^2 A + A\omega^2 \left[\dfrac{1}{v^2(\mathbf{x})} - (\nabla\psi)^2\right] = 0 \\ \\ 2\nabla A \cdot \nabla\psi + A\nabla^2\psi = 0 \end{cases}, \tag{6.68}$$

where $(\nabla\psi)^2 := (\partial\psi/\partial x_1)^2 + (\partial\psi/\partial x_2)^2 + (\partial\psi/\partial x_3)^2$. System (6.68) corresponds to equation (6.62), in the context of trial solution (6.64).

System (6.68) is not simpler than equation (6.62). However, further analysis of the first equation of this system leads to a simplification and results in the eikonal equation. The second equation of system (6.68) is called the transport equation.

Eikonal equation

Considering the first equation of system (6.68) and assuming that both ω and A are nonzero, we can write it as

$$\frac{\nabla^2 A}{A\omega^2} + \left[\frac{1}{v^2(\mathbf{x})} - (\nabla\psi)^2\right] = 0. \tag{6.69}$$

If we assume the inhomogeneity of the continuum to be weak, this assumption is tantamount to viewing the wavelength as being short and, hence, the frequency as being high. In the limit, we let $\omega \to \infty$, and equation (6.69) becomes

$$[\nabla\psi(\mathbf{x})]^2 = \frac{1}{v^2(\mathbf{x})}. \tag{6.70}$$

In view of expression (6.65), we can write equation (6.70) as

$$p^2 = \frac{1}{v^2(\mathbf{x})}, \tag{6.71}$$

where $p^2 = \mathbf{p} \cdot \mathbf{p}$. Equation (6.71) is the eikonal equation for isotropic weakly inhomogeneous continua. It can be viewed as an approximation to wave equation (6.62).

In Chapter 7, we will derive the eikonal equation for anisotropic inhomogeneous continua. However, since, in general, the eikonal equation is based on the high-frequency approximation, this equation is always limited to weak inhomogeneity.[8]

Recall that equation (6.62) does not explicitly refer to either P or S waves. Consequently, equation (6.71) does not explicitly correspond to either wave. However, in view of expression (6.13) and (6.17), if $v(\mathbf{x})$ is a smooth function given by

$$v(\mathbf{x}) = \sqrt{\frac{\lambda(\mathbf{x}) + 2\mu(\mathbf{x})}{\rho(\mathbf{x})}}, \tag{6.72}$$

equation (6.71) can be viewed as corresponding to P waves, and if $v(\mathbf{x})$ is a smooth function given by

$$v(\mathbf{x}) = \sqrt{\frac{\mu(\mathbf{x})}{\rho(\mathbf{x})}}, \tag{6.73}$$

equation (6.71) can be viewed as corresponding to S waves. In general, for inhomogeneous continua, equations (6.4) cannot be split into two wave equations analogous to equations (6.12) and (6.16). In other words, the dilatational and rotational waves are coupled due to the inhomogeneity of the medium, as illustrated in Exercise 6.7. However, assuming sufficiently high frequency, there are two distinct wavefronts that propagate in an inhomogeneous continuum with speeds given by expressions (6.72) and (6.73).

The eikonal equation is a nonlinear partial differential equation. Specifically, it is a first-order and second-degree partial differential equation. In other words, the derivatives are of the first order, while the degree of the exponent is equal to 2. In general, the solution of the eikonal equation requires numerical methods. If the velocity function, v, is constant, the solution of the eikonal equation is also the solution of the corresponding wave equation, as shown in Exercises 6.5 and 7.1. Otherwise, in the cases where $v = v(\mathbf{x})$, the solution of the eikonal equation is not, in general, the solution of the wave equation, and equation (6.62) is only an approximation of the wave equation.

[8] Readers interested in high-frequency approximation might refer to Bleistein, N., Cohen, J.K., and Stockwell, J.W., (2001) Mathematics of multidimensional seismic imaging, migration, and inversion: Springer-Verlag, pp. 5 – 7. Therein, the authors state that

> "high frequency" does not refer to absolute values of the frequency content of the waves. What must be considered is the relationship between the *wavelengths* [...] and the natural *length scales* of the medium.

Transport equation

The second equation of system (6.68), namely,

$$2\nabla A \cdot \nabla\psi + A\nabla^2\psi = 0, \tag{6.74}$$

is the transport equation. For a given eikonal function, ψ, the transport equation describes the amplitude along the wavefront.

Vertically inhomogeneous continua

In seismology, we are often interested in studying layered media where the properties vary along only one axis.

Consider a three-dimensional continuum, $\mathbf{x} = [x, y, z]$, and assume that its properties vary slowly along the z-axis, while remaining the same along the other two axes. It can be shown that, if $v(\mathbf{x}) = v(z)$ varies slowly, equation (6.62) is approximately satisfied by the displacements associated with the SH waves for all directions of propagation. For the case of P and SV waves, equation (6.62) provides a good approximation only for the displacements of waves propagating near the direction of the z-axis.[9]

Note, however, that eikonal equation (6.70), which is derived from equation (6.62), provides — within the conditions of this derivation — a good approximation for signal trajectories in all directions of propagation.

Closing remarks

In this chapter, to study wave phenomena, we formulated wave equations. These equations are formulated as special cases of Cauchy's equations of motion for isotropic homogeneous continua. From these equations, we identify two distinct types of waves, namely P and S waves, which propagate with two distinct speeds. In Chapter 7, we will formulate Cauchy's equations of motion in the context of anisotropic inhomogeneous continua. Therein, we show the existence of three types of waves.

All waves discussed in this book propagate within the body of a continuum. Consequently, they correspond to the so-called body waves, as opposed to the surface and interface waves, which we do not discuss.

The derivation of the wave equation shown in this chapter is rooted in the balance of linear momentum. This derivation formulates wave propagation

[9]Readers interested in wave propagation in slowly varying vertically nonuniform continua might refer to Krebes, E.S., (1987) Seismic theory and methods (Lecture notes): The University of Calgary, pp. 5-9 – 5-12.

as a result of a continuum conserving the linear momentum within itself. The wave equation can also be derived by invoking other physical principles. For instance, in Chapter 13, its derivation is based on Hamilton's principle, which formulates wave propagation as a result of a continuum restoring itself to the state of equilibrium through the process governed by the principle of stationary action.

The study of solutions for the wave equation motivated several recent developments in mathematics. As a result of these developments, the theory of generalized functions — in particular, the theory of distributions — extends the solutions for the wave equation to include nondifferentiable functions. Also, studies of wave propagation in elastic media have played an important role in the theory of integral equations.[10]

Exercises

Exercise 6.1 *Show the details of the derivation of Lemma 6.1.*

Solution 6.1 *For the first term of the wave equation, consider*

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial x}.$$

Since, following expression (6.42), $\partial y/\partial x = \partial z/\partial x = 1$, *we obtain*

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z}.$$

[10]Interested readers might refer to Aleksandrov, A.D., Kolmogorov, A.N., Lavrentev, M.A., (editors), (1969/1999) Mathematics: Its content, methods and meaning: Dover, Vol. II, pp. 48 – 54 and Vol. III, pp. 245 – 250, to Bleistein, N., Cohen, J.K., and Stockwell, J.W., (2001) Mathematics of multidimensional seismic imaging, migration, and inversion: Springer-Verlag, pp. 389 – 408, and to Demidov, A.S., (2001) Generalized functions in mathematical physics: Main ideas and concepts: Nova Science Publishers, Inc., pp. 41 – 53.

Consequently,

$$\begin{aligned}
\frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}\right)\\
&= \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}\right)\frac{\partial y}{\partial x}+\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}\right)\frac{\partial z}{\partial x}\\
&= \frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial y\partial z}+\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2}\\
&= \frac{\partial^2 u}{\partial y^2}+2\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2},
\end{aligned} \tag{6.75}$$

where, again, we use the equality given by $\partial y/\partial x = \partial z/\partial x = 1$*, and the equality of mixed partial derivatives. Similarly, for the second term of the wave equation, consider*

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial u}{\partial y}\frac{\partial y}{\partial t}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial t}\\
&= v\frac{\partial u}{\partial y}-v\frac{\partial u}{\partial z}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2} &= \frac{\partial}{\partial t}\left(v\frac{\partial u}{\partial y}-v\frac{\partial u}{\partial z}\right)\\
&= \frac{\partial}{\partial y}\left(v\frac{\partial u}{\partial y}-v\frac{\partial u}{\partial z}\right)\frac{\partial y}{\partial t}+\frac{\partial}{\partial z}\left(v\frac{\partial u}{\partial y}-v\frac{\partial u}{\partial z}\right)\frac{\partial z}{\partial t}\\
&= \left(v\frac{\partial^2 u}{\partial y^2}-v\frac{\partial^2 u}{\partial y\partial z}\right)v+\left(v\frac{\partial^2 u}{\partial z\partial y}-v\frac{\partial^2 u}{\partial z^2}\right)(-v)\\
&= v^2\left(\frac{\partial^2 u}{\partial y^2}-\frac{\partial^2 u}{\partial y\partial z}-\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2}\right)\\
&= v^2\left(\frac{\partial^2 u}{\partial y^2}-2\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2}\right).
\end{aligned} \tag{6.76}$$

where the equality of mixed partial derivatives is used. Inserting expressions

(6.75) and (6.76) into equation (6.40), we obtain

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2}-\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2}&=\frac{\partial^2 u}{\partial y^2}+2\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2}-\frac{1}{v^2}v^2\left(\frac{\partial^2 u}{\partial y^2}-2\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2}\right)\\&=\frac{\partial^2 u}{\partial y^2}+2\frac{\partial^2 u}{\partial z\partial y}+\frac{\partial^2 u}{\partial z^2}-\frac{\partial^2 u}{\partial y^2}+2\frac{\partial^2 u}{\partial z\partial y}-\frac{\partial^2 u}{\partial z^2}\\&=4\frac{\partial^2 u}{\partial y\partial z}\\&=0,\end{aligned}$$

where the equality of mixed partial derivatives is used. Hence, we conclude that

$$\frac{\partial^2 u}{\partial y\partial z}=0,$$

as required.

Exercise 6.2 *Show the details of the derivation of Lemma 6.2.*

Solution 6.2 *Considering the equality of mixed partial derivatives, we can write*

$$\begin{aligned}\frac{\partial}{\partial y}\left[\frac{\partial u\,(y,z)}{\partial z}\right]&=\frac{\partial}{\partial z}\left[\frac{\partial u\,(y,z)}{\partial y}\right].\\&=0.\end{aligned}$$

Consequently, for the second partial derivative to vanish, we require that

$$\left[\frac{\partial u\,(y,z)}{\partial z}\right]=G\,(z)\,,$$

on the left-hand side, and

$$\left[\frac{\partial u\,(y,z)}{\partial y}\right]=F\,(y)\,,$$

on the right-hand side. In other words, we require that G be a function of z only, while F be a function of y only. Hence, integrating, we obtain

$$\begin{aligned}u\,(y,z)&=\int F\,(y)\;\mathrm{d}y\\&=f\,(y)+a\,(z)\,,\end{aligned}$$

where $a(z)$ is the integration constant with respect to dy, and

$$u(y,z) = \int G(z)\,\mathrm{d}z$$
$$= g(z) + b(y),$$

where $b(y)$ is the integration constant with respect to dz. In view of the arbitrariness of the integration constants, we can denote $a(z) \equiv g(z)$ and $b(y) \equiv f(y)$. Thus, we obtain

$$u(y,z) = f(y) + g(z),$$

as required.

Exercise 6.3 *Consider wave equation (6.40). Using solution (6.48), obtain equation (6.49).*

Solution 6.3 *In view of solution (6.48), namely, $u(x,t) = \acute{u}(x)\exp(-i\omega t)$, consider the position derivatives, namely,*

$$\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{\partial^2 \acute{u}(x)}{\partial x^2}\exp(-i\omega t), \tag{6.77}$$

and the time derivatives, namely,

$$\frac{\partial^2 u(x,t)}{\partial t^2} = -\omega^2 \acute{u}(x)\exp(-i\omega t). \tag{6.78}$$

Substituting expressions (6.77) and (6.78) into equation (6.40), and dividing by the exponential factor, we obtain a function of a single variable,

$$\frac{\mathrm{d}^2 \acute{u}(x)}{\mathrm{d}x^2} + \left(\frac{\omega}{v}\right)^2 \acute{u}(x) = 0,$$

which is equation (6.49), as required.

Exercise 6.4 *Consider equation (6.53). In view of transformation (6.54), let*

$$\acute{u}(x,z,t) \equiv u(\xi,\varsigma,t), \tag{6.79}$$

where $\xi := x/v_x$ and $\varsigma := z/v_z$. Using the chain rule, show that equation (6.53) is equivalent to equation (6.55).

Solution 6.4 *Taking the derivative of both sides of equation (6.79) with respect to x, we obtain*

$$\frac{\partial \acute{u}(x,z,t)}{\partial x}=\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x}=\frac{1}{v_x}\frac{\partial u(\xi,\varsigma,t)}{\partial \xi},$$

and

$$\frac{\partial^2 \acute{u}(x,z,t)}{\partial^2 x}=\frac{1}{v_x^2}\frac{\partial^2 u(\xi,\varsigma,t)}{\partial^2 \xi}. \tag{6.80}$$

Similarly, taking the derivative of both sides of equation (6.79) with respect to z, we obtain

$$\frac{\partial^2 \acute{u}(x,z,t)}{\partial^2 z}=\frac{1}{v_z^2}\frac{\partial^2 u(\xi,\varsigma,t)}{\partial^2 \varsigma}, \tag{6.81}$$

while, taking the derivative of both sides of equation (6.79) with respect to t, we get

$$\frac{\partial^2 \acute{u}(x,z,t)}{\partial^2 t}=\frac{\partial^2 u(\xi,\varsigma,t)}{\partial^2 t}. \tag{6.82}$$

We can always write equation (6.53) as

$$\frac{\partial^2 u(\xi,\varsigma,t)}{\partial \xi^2}+\frac{\partial^2 u(\xi,\varsigma,t)}{\partial \varsigma^2}=\frac{\partial^2 u(\xi,\varsigma,t)}{\partial t^2}, \tag{6.83}$$

where ξ and ς are the variables of differentiation. Substituting expressions from equations (6.80), (6.81) and (6.82) into (6.83), we obtain equation (6.55), as required.

Exercise 6.5 [11] *Consider a three-dimensional scalar wave equation given by*

$$\nabla^2 u \equiv \frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}=\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2},$$

where v is the velocity of propagation and t is time. Let the plane-wave solution be $u(\mathbf{x},t)=f(\eta)$, where $\eta=n_1x_1+n_2x_2+n_3x_3-vt$, with n_i being the components of the unit vector that is normal to the wavefront. Show that the plane-wave solution of the wave equation is also a solution of its characteristic equation, given by

$$(\nabla u)^2 \equiv \left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2+\left(\frac{\partial u}{\partial x_3}\right)^2=\frac{1}{v^2}\left(\frac{\partial u}{\partial t}\right)^2. \tag{6.84}$$

[11] See also Section 6.6.3 and Exercise 7.1.

Solution 6.5 *Considering the plane-wave solution, we obtain*

$$\frac{\partial u}{\partial x_i}=\frac{\partial f}{\partial \eta}\frac{\partial \eta}{\partial x_i}=\frac{\partial f}{\partial \eta}n_i,$$

and

$$\frac{\partial u}{\partial t}=\frac{\partial f}{\partial \eta}\frac{\partial \eta}{\partial t}=-\frac{\partial f}{\partial \eta}v.$$

Substituting $\partial u/\partial x_i$ *and* $\partial u/\partial t$ *into equation (6.84), we can write*

$$\left(\frac{\partial f}{\partial \eta}n_1\right)^2+\left(\frac{\partial f}{\partial \eta}n_2\right)^2+\left(\frac{\partial f}{\partial \eta}n_3\right)^2=\frac{1}{v^2}\left(-\frac{\partial f}{\partial \eta}v\right)^2,$$

which yields

$$\left(\frac{\partial f}{\partial \eta}\right)^2\left(n_1^2+n_2^2+n_3^2\right)=\left(\frac{\partial f}{\partial \eta}\right)^2.$$

This equality is justified since for the unit vector, $\mathbf{n}$, *we have* $n_1^2+n_2^2+n_3^2=1$.

Exercise 6.6 *In view of Section 6.5, considering the reduced form of equation (6.62) in a single spatial dimension and using the trial solution given by* $u(\mathbf{x})=A(\mathbf{x})\exp[i\omega\psi(\mathbf{x})]$, *obtain set (6.67).*

Solution 6.6 *Considering a single spatial dimension, the reduced form of equation (6.62) is*

$$\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2}+\left[\frac{\omega}{v(x)}\right]^2u(x)=0.$$

Inserting a one-dimensional form of the given trial solution into this equation, performing the differentiation and dividing both sides of the resulting equation by the exponential term, we obtain

$$\frac{\mathrm{d}^2A}{\mathrm{d}x^2}+A\omega^2\left(\frac{1}{v^2}-\frac{\mathrm{d}\psi}{\mathrm{d}x}\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)+i\omega\left[2\frac{\mathrm{d}A}{\mathrm{d}x}\frac{\mathrm{d}\psi}{\mathrm{d}x}+A\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\right]=0,$$

which is analogous to equation (6.66) and, hence, leads to set (6.67), as required.

Exercise 6.7 [12] *Using stress-strain equations (6.1) and Cauchy's equations of motion (6.2), obtain equations of motion for an isotropic inhomogeneous continuum. Discuss these equations in the context of equations (6.4).*

[12] See also Sections 6.1.1 and 6.6.3.

Solution 6.7 *Considering an inhomogeneous continuum, where Lamé's parameters are functions of position, and in view of definition (1.15), we can write equations (6.1) as*

$$\sigma_{ij} = \lambda(\mathbf{x})\,\delta_{ij} \sum_{k=1}^{3} \frac{\partial u_k}{\partial x_k} + \mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \qquad i,j \in \{1,2,3\}, \quad (6.85)$$

which are stress-strain equations for an isotropic inhomogeneous continuum. Considering an inhomogeneous continuum, where mass density is a function of position, we can write equations (6.2) as

$$\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j}, \qquad i \in \{1,2,3\}, \quad (6.86)$$

which are equations of motion for an isotropic inhomogeneous continuum. Using equations (6.85), we can write equations (6.86) as

$$\begin{aligned}
\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} &= \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left[\lambda(\mathbf{x})\,\delta_{ij} \sum_{k=1}^{3} \frac{\partial u_k}{\partial x_k} + \mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] \\
&= \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left[\lambda(\mathbf{x})\,\delta_{ij} \sum_{k=1}^{3} \frac{\partial u_k}{\partial x_k} \right] + \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left[\mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right],
\end{aligned}$$

where $i \in \{1,2,3\}$. Using the property of Kronecker's delta, we obtain

$$\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i} \left[\lambda(\mathbf{x}) \sum_{k=1}^{3} \frac{\partial u_k}{\partial x_k} \right] + \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left[\mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right],$$

where $i \in \{1,2,3\}$. Letting $k = j$ for the summation index, we can write

$$\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i} \left[\lambda(\mathbf{x}) \sum_{j=1}^{3} \frac{\partial u_j}{\partial x_j} \right] + \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left[\mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right],$$

where $i \in \{1,2,3\}$. Using the product rule and the linearity of differential operators, we obtain

$$\begin{aligned}
\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} &= \frac{\partial \lambda}{\partial x_i} \sum_{j=1}^{3} \frac{\partial u_j}{\partial x_j} + \lambda \frac{\partial}{\partial x_i} \sum_{j=1}^{3} \frac{\partial u_j}{\partial x_j} \\
&\quad + \sum_{j=1}^{3} \frac{\partial \mu}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \mu \sum_{j=1}^{3} \left(\frac{\partial}{\partial x_j} \frac{\partial u_i}{\partial x_j} + \frac{\partial}{\partial x_j} \frac{\partial u_j}{\partial x_i} \right),
\end{aligned}$$

where $i \in \{1,2,3\}$. Differentiating and using the equality of mixed partial derivatives, we obtain

$$
\begin{aligned}
\rho(\mathbf{x})\frac{\partial^2 u_i}{\partial t^2} &= \frac{\partial \lambda}{\partial x_i}\sum_{j=1}^{3}\frac{\partial u_j}{\partial x_j} + \lambda\sum_{j=1}^{3}\frac{\partial^2 u_j}{\partial x_i \partial x_j} \\
&\quad + \sum_{j=1}^{3}\frac{\partial \mu}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \mu\sum_{j=1}^{3}\left(\frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial^2 u_j}{\partial x_i \partial x_j}\right),
\end{aligned}
$$

where $i \in \{1,2,3\}$. Simplifying and rearranging, we get

$$
\begin{aligned}
\rho(\mathbf{x})\frac{\partial^2 u_i}{\partial t^2} &= (\lambda+\mu)\frac{\partial}{\partial x_i}\sum_{j=1}^{3}\frac{\partial u_j}{\partial x_j} + \mu\left(\sum_{j=1}^{3}\frac{\partial^2}{\partial x_j^2}\right)u_i \\
&\quad + \frac{\partial \lambda}{\partial x_i}\sum_{j=1}^{3}\frac{\partial u_j}{\partial x_j} + \sum_{j=1}^{3}\frac{\partial \mu}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad (6.87)
\end{aligned}
$$

where $i \in \{1,2,3\}$. These are equations of motion for an isotropic inhomogeneous continuum.[13]

Examining equations (6.87), we notice that if ρ, λ and μ are constants — as is the case for homogeneous continua — equations (6.87) reduce to equations (6.4), as expected. Also we notice that invoking definitions (1.26) and (1.30) as well as identity (6.7) we can express the displacement vector, $\mathbf{u}$, in the first three terms on the right-hand side, using the dilatation, φ, and the rotation vector, $\boldsymbol{\Psi}$. Investigating the last term on the right-hand side, however, it can be shown that we cannot express the displacement vector on the right-hand side of equations (6.86) using only φ and $\boldsymbol{\Psi}$. Consequently, we cannot split equations (6.87) into two parts that are associated with the dilatation alone and with the rotation vector alone, respectively, as we did in Sections 6.1.2 and 6.1.3 in the case of isotropic homogeneous continua. In other words, the dilatational and rotational waves are coupled due to the inhomogeneity of the continuum.

[13]Readers interested in a solution to these equations might refer to Karal, F.C., and Keller, J.B., (1959) Elastic wave propagation in homogeneous and inhomogeneous media: J. Acoust. Soc. Am., **31** (6), 694 – 705.

Chapter 7

Equations of motion: Anisotropic inhomogeneous continua

> ... an exact solution to a problem in wave phenomena is not an end in itself. Rather, it is the asymptotic solution that provides means of interpretation and a basis for understanding. The exact solution, then, only provides a point of departure for obtaining a meaningful solution.
>
> *Norman Bleistein (1984) Mathematical methods for wave phenomena*

Preliminary remarks

In Chapter 6, to study wave phenomena in an isotropic homogeneous continuum, we obtained the equations of motion by invoking Cauchy's equations of motion and using stress-strain equations that correspond to such a continuum. In this chapter, we will study wave phenomena in an anisotropic inhomogeneous continuum by following a strategy analogous to that used in Chapter 6. In this study, we learn about the existence of three types of waves that can propagate in anisotropic continua.

We begin this chapter with the derivation of the equations of motion in an anisotropic inhomogeneous continuum. We obtain these equations by combining Cauchy's equations of motion with the stress-strain equations for an anisotropic inhomogeneous continuum. To solve the resulting equations, we use a trial solution. Subsequently, we derive the eikonal equation for anisotropic inhomogeneous continua, which is the fundamental equation of ray theory, to be studied in the subsequent chapters of *Part II* and *Part III.*

7.1 Formulation of equations

In Chapter 6, the wave equation is derived by considering Cauchy's equations of motion (2.50) and expressing the stress-tensor components therein in terms of stress-strain equations (5.65), which describe an isotropic homogeneous continuum. In the present chapter, we will derive the equations of motion for an anisotropic inhomogeneous continuum by considering Cauchy's equations of motion (2.50) and expressing the stress-tensor components therein in terms of the stress-strain equations that describe an anisotropic inhomogeneous continuum.

In view of equations of motion (2.50), consider equations

$$\rho(\mathbf{x})\frac{\partial^2 u_i}{\partial t^2}=\sum_{j=1}^{3}\frac{\partial \sigma_{ij}}{\partial x_j}, \qquad i\in\{1,2,3\}. \tag{7.1}$$

Note that in equations (7.1), due to the inhomogeneity of the continuum, mass density, $\rho(\mathbf{x})$, is a function of position. Also note that, in general, displacement vector, $\mathbf{u}=[u_1,u_2,u_3]$, and stress-tensor components, σ_{ij}, are also functions of position. However, for clarity of notation, we explicitly state the $\mathbf{x}$-dependence only for the mass density, ρ, and for the elasticity tensor, c_{ijkl}, which describe a given anisotropic inhomogeneous continuum.

The stress-strain equations that account for the anisotropy as well as the inhomogeneity of the continuum are expressed by writing equations (3.1) with the elasticity tensor being functions of position, namely,

$$\sigma_{ij}=\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\,\varepsilon_{kl}, \qquad i,j\in\{1,2,3\}. \tag{7.2}$$

Using definition (1.15), we can rewrite stress-strain equations (7.2) as

$$\sigma_{ij}=\frac{1}{2}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\left(\frac{\partial u_k}{\partial x_l}+\frac{\partial u_l}{\partial x_k}\right), \qquad i,j\in\{1,2,3\}. \tag{7.3}$$

We wish to combine the equations of motion and the stress-strain equations to obtain the equations of motion in an anisotropic inhomogeneous continuum. Inserting stress-strain equations (7.3) into equations of motion (7.1), we obtain

$$\rho(\mathbf{x})\frac{\partial^2 u_i}{\partial t^2}=\sum_{j=1}^{3}\frac{\partial}{\partial x_j}\left[\frac{1}{2}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\left(\frac{\partial u_k}{\partial x_l}+\frac{\partial u_l}{\partial x_k}\right)\right] \tag{7.4}$$
$$=\frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\left(\frac{\partial u_k}{\partial x_l}+\frac{\partial u_l}{\partial x_k}\right)$$
$$+\frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\left(\frac{\partial^2 u_k}{\partial x_j\partial x_l}+\frac{\partial^2 u_l}{\partial x_j\partial x_k}\right),$$

where $i\in\{1,2,3\}$. Equations (7.4) are equations of motion in anisotropic inhomogeneous continua. For isotropic continua, these equations reduce to equations (6.87), and for isotropic homogeneous continua, they further reduce to equations (6.4). Equations (7.4) are complicated differential equations and, in general, we are unable to find their solutions analytically.

7.2 Formulation of solutions

To investigate equations (7.4), let us consider the trial solution that is a function of position, $\mathbf{x}$, and time, t, given by

$$\mathbf{u}(\mathbf{x},t)=\mathbf{A}(\mathbf{x})f(\eta), \tag{7.5}$$

where $\mathbf{A}$ is a vector function of position, $\mathbf{x}$, and f is a scalar function whose argument is given by

$$\eta=v_0[\psi(\mathbf{x})-t], \tag{7.6}$$

with v_0 being a constant with units of velocity. Function $\psi:\mathbb{R}^3\to\mathbb{R}$, shown in expression (7.6), is referred to as the eikonal function.

To see the physical meaning of this trial solution, consider function $f(\eta)$ in the context of trial solutions (6.56), (6.63) and (6.64). We see that f corresponds to the phase factor. Since along the level sets of $\psi(\mathbf{x})$, function f is constant, these level sets correspond to wavefronts. In other words, $\psi(\mathbf{x})=t$ describes a moving wavefront. Function f gives the waveform as a function of time with $\mathbf{A}$ being the spatially variable amplitude of this waveform.

Inserting trial solution (7.5) into equations (7.4), while using the symmetries of the elasticity tensor, c_{ijkl}, and the equality of mixed partial derivatives, we obtain

$$\begin{aligned}\rho(\mathbf{x})\, v_0^2 A_i(\mathbf{x}) \frac{\mathrm{d}^2 f}{\mathrm{d}\eta^2} = \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} \Bigg\{ & f\left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x})\frac{\partial^2 A_k}{\partial x_j \partial x_l}\right] \\ & + v_0 \frac{\mathrm{d}f}{\mathrm{d}\eta}\left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} A_l \frac{\partial \psi}{\partial x_k}\right. \\ & \left. + c_{ijkl}(\mathbf{x})\left(\frac{\partial A_l}{\partial x_j}\frac{\partial \psi}{\partial x_k} + \frac{\partial A_k}{\partial x_l}\frac{\partial \psi}{\partial x_j} + A_l \frac{\partial^2 \psi}{\partial x_j \partial x_k}\right)\right] \\ & + v_0^2 \frac{\mathrm{d}^2 f}{\mathrm{d}\eta^2}\left[c_{ijkl}(\mathbf{x}) A_k \frac{\partial \psi}{\partial x_j}\frac{\partial \psi}{\partial x_l}\right]\Bigg\},\end{aligned}$$

where $i \in \{1,2,3\}$. Rearranging and using the product rule in the coefficient of $\mathrm{d}f/\mathrm{d}\eta$, we get

$$\begin{aligned} 0 = \sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} \Bigg\{ & f\left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x})\frac{\partial^2 A_k}{\partial x_j \partial x_l}\right] \\ & + v_0 \frac{\mathrm{d}f}{\mathrm{d}\eta}\left[\frac{\partial}{\partial x_j}\left(c_{ijkl}(\mathbf{x}) A_l \frac{\partial \psi}{\partial x_k}\right) + c_{ijkl}(\mathbf{x})\frac{\partial A_k}{\partial x_l}\frac{\partial \psi}{\partial x_j}\right] \\ & + v_0^2 \frac{\mathrm{d}^2 f}{\mathrm{d}\eta^2}\left[c_{ijkl}(\mathbf{x}) A_k \frac{\partial \psi}{\partial x_j}\frac{\partial \psi}{\partial x_l}\right]\Bigg\} - \frac{\mathrm{d}^2 f}{\mathrm{d}\eta^2}\rho(\mathbf{x}) A_i v_0^2, \end{aligned} \tag{7.7}$$

where $i \in \{1,2,3\}$.

Concisely, equation (7.7) can be written as

$$a(\mathbf{x}) f + b(\mathbf{x}) f' + c(\mathbf{x}) f'' = 0,$$

where $f' := \mathrm{d}f/\mathrm{d}\eta$ and $f'' := \mathrm{d}^2 f/\mathrm{d}\eta^2$. For equation (7.7) to be satisfied for an arbitrary f, each of the coefficients — $a(\mathbf{x})$, $b(\mathbf{x})$ and $c(\mathbf{x})$ — must vanish.[1]

Note that, in view of trial solution (7.5), we require the arbitrariness of f in order to allow any function to describe the waveform.

Assuming $v_0 \neq 0$, we obtain three systems of equations, namely,

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left(\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j}\frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x})\frac{\partial^2 A_k}{\partial x_j \partial x_l}\right) = 0, \tag{7.8}$$

[1] Readers interested in an analogous formulation of the three vanishing terms of equation (7.7) might refer to Červený, V., (2001) Seismic ray theory: Cambridge University Press, p. 55, pp. 57 – 58 and pp. 62 – 63.

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left[\frac{\partial}{\partial x_j}\left(c_{ijkl}(\mathbf{x})\,A_l\frac{\partial\psi}{\partial x_k}\right)+c_{ijkl}(\mathbf{x})\,\frac{\partial A_k}{\partial x_l}\frac{\partial\psi}{\partial x_j}\right]=0, \tag{7.9}$$

and

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\,A_k\frac{\partial\psi}{\partial x_j}\frac{\partial\psi}{\partial x_l}-\rho(\mathbf{x})\,A_i=0, \tag{7.10}$$

where $i\in\{1,2,3\}$, which correspond to $a(\mathbf{x})$, $b(\mathbf{x})$ and $c(\mathbf{x})$, respectively. Equations (7.8), (7.9) and (7.10) constitute a system of equations for $\psi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$.

7.3 Eikonal equation

In order to obtain $\psi(\mathbf{x})$, we turn our attention to equations (7.10), from which we can factor out the components of vector $\mathbf{A}(\mathbf{x})$. Hence, we can write

$$\sum_{k=1}^{3}\left(\sum_{j=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\,\frac{\partial\psi}{\partial x_j}\frac{\partial\psi}{\partial x_l}-\rho(\mathbf{x})\,\delta_{ik}\right)A_k(\mathbf{x})=\mathbf{0},\qquad i\in\{1,2,3\}. \tag{7.11}$$

In view of expression (6.65), let us denote

$$p_j:=\frac{\partial\psi}{\partial x_j},\qquad j\in\{1,2,3\}, \tag{7.12}$$

where $\mathbf{p}$ is the phase-slowness vector, which describes the slowness of the propagation of the wavefront.

Note that the meaning of $\mathbf{p}$ can be seen by examining expression (7.6) and considering a three-dimensional continuum. Therein, ψ is a function relating position variables, x_1, x_2 and x_3, to the traveltime, t. Thus, since ψ has units of time, $p_j:=\partial\psi/\partial x_j$ has units of slowness and the level sets of $\psi(\mathbf{x})$ can be viewed as wavefronts at a given time t. Consequently, in view of properties of the gradient operator, $\mathbf{p}=\nabla\psi(\mathbf{x})$ is a vector whose direction corresponds to the wavefront normal and whose magnitude corresponds to the wavefront slowness.

In view of notation (7.12), we can write equations (7.11) as

$$\sum_{k=1}^{3}\left(\sum_{j=1}^{3}\sum_{l=1}^{3}c_{ijkl}(\mathbf{x})\,p_jp_l-\rho(\mathbf{x})\,\delta_{ik}\right)A_k(\mathbf{x})=\mathbf{0},\qquad i\in\{1,2,3\}. \tag{7.13}$$

Equations (7.13) are referred to as Christoffel's equations.

In Chapter 10, we discuss equations (7.13) in the context of the particular symmetries of continua, which were introduced in Chapter 5. Therein, we also show that the eigenvalues resulting from these equations are associated with the velocity of the wavefront while the corresponding eigenvectors are the displacement directions. Herein, we study the general form of equations (7.13).

We know from linear algebra that equations (7.13) have nontrivial solutions if and only if

$$\det\left[\sum_{j=1}^{3}\sum_{l=1}^{3} c_{ijkl}(\mathbf{x})\, p_j p_l - \rho(\mathbf{x})\,\delta_{ik}\right] = 0, \qquad i,k \in \{1,2,3\}. \tag{7.14}$$

Assuming that $p^2 \neq 0$, we can write determinant (7.14) as

$$\left(p^2\right)^3 \det\left[\sum_{j=1}^{3}\sum_{l=1}^{3} c_{ijkl}(\mathbf{x})\,\frac{p_j p_l}{p^2} - \frac{\rho(\mathbf{x})}{p^2}\delta_{ik}\right] = 0, \qquad i,k \in \{1,2,3\}. \tag{7.15}$$

Note that $p^2 = 0$ would mean that the slowness of the propagation of the wavefront is zero. This would imply the velocity to be infinite, which is a nonphysical situation. Also, in view of determinant (7.14), $p^2 = 0$ would result in $\det[\rho(\mathbf{x})\,\delta_{ik}] = 0$, which would imply $\rho(\mathbf{x}) = 0$.

Expression (7.15) is a polynomial of degree 3 in p^2, where the coefficients depend on the direction of the phase-slowness vector, $\mathbf{p}$. Any such polynomial can be factored out as

$$\left[p^2 - \frac{1}{v_1^2(\mathbf{x},\mathbf{p})}\right]\left[p^2 - \frac{1}{v_2^2(\mathbf{x},\mathbf{p})}\right]\left[p^2 - \frac{1}{v_3^2(\mathbf{x},\mathbf{p})}\right] = 0, \tag{7.16}$$

where $1/v_i^2$ are the roots of polynomial (7.15).

The conditions imposed on the c_{ijkl} by the stability conditions — discussed in Section 4.3 — imply that the three roots of polynomial (7.15) are real and positive. These properties are further discussed in Section 10.1. The existence of three roots implies the existence of three types of waves, which can propagate in anisotropic continua.

Now, let us consider a given root of equation (7.16). Each root is the eikonal equation for a given type of wave, namely,

$$p^2 = \frac{1}{v_i^2(\mathbf{x},\mathbf{p})}, \qquad i \in \{1,2,3\}. \tag{7.17}$$

Let us examine the meaning of this equation.[2]

Since $p^2 = \mathbf{p} \cdot \mathbf{p}$ is the squared magnitude of the slowness vector, which is normal to the wavefront, then — in view of the wavefronts being the loci of constant phase — v is the function describing phase velocity. This velocity is a function of position, $\mathbf{x}$, and the direction of $\mathbf{p}$. Hence, equation (7.17) applies to anisotropic inhomogeneous continua and can be viewed as an extension of equation (6.71), which is valid for isotropic inhomogeneous continua.

Considering two adjacent wavefronts, we can view equation (7.17) as an infinitesimal formulation of Huygens' principle.[3]

Note that function v is homogeneous of degree 0 in the p_i. In other words, the orientation of a wavefront is described by the direction of $\mathbf{p}$ and is independent of the length of $\mathbf{p}$. Hence, in equation (7.17) we could also write $v_i = v_i(\mathbf{x}, \mathbf{n})$, where $\mathbf{n}$ is a unit vector in the direction of $\mathbf{p}$. Notably, we will use this notation in Chapter 10.

Furthermore, as shown explicitly in Chapter 10, the phase-velocity function can be expressed in terms of the properties of the continuum, namely, its mass density and elasticity parameters. Thus, the eikonal equation relates the magnitude of the slowness with which the wavefront propagates to the properties of the continuum through which it propagates.

In the mathematical context, the eikonal equation is a differential equation. Recalling expressions (7.12), we can rewrite equation (7.17) as

$$[\nabla \psi(\mathbf{x})]^2 = \frac{1}{v^2(\mathbf{x}, \mathbf{p})}. \tag{7.18}$$

In general, the eikonal equation is a nonlinear, first-order, partial differential equation in $\mathbf{x}$ to be solved for the eikonal function, $\psi(\mathbf{x})$. It belongs to the Hamilton-Jacobi class of differential equations.[4]

Equation (7.9) is the transport equation. This transport equation possesses a vectorial form that is valid for anisotropic inhomogeneous continua.

---

[2] Readers interested in the mathematical formulation of the conditions under which the eikonal equation provides a good approximation to the wave equation might refer to Officer, C.B., (1974) Introduction to theoretical geophysics: Springer-Verlag, pp. 204 – 205.

[3] Readers interested in a formulation relating the eikonal equation to Huygens' principle might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2*nd* edition): Springer-Verlag, pp. 248 – 252, and to Lanczos, C., (1949/1986) The variational principles in mechanics: Dover, pp. 269 – 270.

[4] Readers interested in a mathematical study of the eikonal and transport equations might refer to Taylor, M.E., (1996) Partial differential equations; Basic theory: Springer-Verlag, pp. 79 – 84 and pp. 440 – 447.

It is analogous to the scalar transport equation (6.74), which is valid for isotropic inhomogeneous continua.

Closing remarks

In this chapter, while seeking to study the propagation of waves in anisotropic inhomogeneous continua, we follow a strategy analogous to that used in Chapter 6. However, having obtained the equations of motion, we find that we are unable to investigate them analytically. Thus, we utilize a trial solution that leads us to the eikonal equation, which relates the slowness of propagation of the wavefront to the properties of the continuum through which it propagates. In Chapter 8, we will continue our study of wave propagation in anisotropic inhomogeneous continua by solving the eikonal equation.

Note that equations (7.8), (7.9) and (7.10) constitute an overdetermined system that results from inserting trial solution (7.5) into equations (7.4). Since the system is overdetermined, we can obtain a unique solution without using all the equations that compose this system. If we obtain functions $\psi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$ from equations (7.9) and (7.10), we can investigate how well these functions accommodate equations (7.4) by studying their effect on equation (7.8).

Exercises

Exercise 7.1 *In view of Exercise 6.5, consider a more general form of the solution that is given by $u(\mathbf{x},t) = f(\eta)$, where $\eta = v_0[\psi(\mathbf{x}) - t]$. Show that the necessary condition for characteristic equation (6.84) to be satisfied is the eikonal equation, given by*

$$(\nabla\psi)^2 = \frac{1}{v^2}.$$

Solution 7.1 *Considering the argument of f given by $\eta = v_0[\psi(\mathbf{x}) - t]$, we obtain*

$$\frac{\partial u}{\partial x_i} = \frac{\partial f}{\partial \eta}\frac{\partial \eta}{\partial x_i} = v_0 \frac{\partial f}{\partial \eta}\frac{\partial \psi}{\partial x_i}, \qquad i \in \{1,2,3\},$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial f}{\partial \eta}\frac{\partial \eta}{\partial t} = -v_0\frac{\partial f}{\partial \eta}.$$

Substituting $\partial u/\partial x_i$ *and* $\partial u/\partial t$ *into characteristic equation (6.84), we can write*

$$\left(v_0\frac{\partial f}{\partial \eta}\frac{\partial \psi}{\partial x_1}\right)^2 + \left(v_0\frac{\partial f}{\partial \eta}\frac{\partial \psi}{\partial x_2}\right)^2 + \left(v_0\frac{\partial f}{\partial \eta}\frac{\partial \psi}{\partial x_3}\right)^2 = \frac{1}{v^2}\left(-\frac{\partial f}{\partial \eta}v_0\right)^2,$$

which yields

$$v_0^2\left(\frac{\partial f}{\partial \eta}\right)^2\left[\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 + \left(\frac{\partial \psi}{\partial x_3}\right)^2\right] = \left(\frac{v_0}{v}\right)^2\left(\frac{\partial f}{\partial \eta}\right)^2.$$

Since, in general, $v_0 \neq 0$ *and* $\partial f/\partial \eta \neq 0$, *we can write*

$$\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \left(\frac{\partial \psi}{\partial x_2}\right)^2 + \left(\frac{\partial \psi}{\partial x_3}\right)^2 = \frac{1}{v^2},$$

which is the required eikonal equation.

Remark 7.1 *If* v *is constant, Exercise 7.1 is reduced to Exercise 6.5.*

This Page Intentionally Left Blank

Chapter 8

Hamilton's ray equations

> It is a common physical knowledge that wavefields, rather than rays, are physical reality. None the less, the traditions to endow rays with certain physical properties, traced back to Descartes times, have been deeply rooted in natural sciences. Rays are discussed as if they were real objects.
>
> *Yuri A. Kravtsov and Yuri I. Orlov (1999) Caustics, catastrophes and wavefields*

Preliminary remarks

In Chapter 7, we obtained the eikonal equation that gives us the magnitude of phase slowness as a function of the properties of an anisotropic inhomogeneous continuum through which the wavefront propagates. In this chapter, we will focus our attention on the solution of the eikonal equation.

We begin this chapter by using the method of characteristics to solve the eikonal equation, which is a first-order nonlinear partial differential equation. The solution leads to a system of ordinary first-order differential equations that describe the curves that form the solution surface in the **xp**-space.[1] These are the characteristic equations. Parametrizing the characteristic equations in terms of time, we obtain Hamilton's ray equations, whose solutions give the trajectory of a signal propagating through an anisotropic inhomogeneous continuum. Hamilton's ray equations, which are the key equations of ray theory.

[1] In classical mechanics, the **xp**-space corresponds to the momentum phase space. In this book, however, to avoid the confusion with the term "phase" that we use in the specific context of wave phenomena, we do not use this nomenclature.

Readers who are not familiar with Euler's homogeneous-function theorem might find it useful to study this chapter together with Appendix A.

8.1 Method of characteristics

8.1.1 Level-set functions

The eikonal equation is a first-order nonlinear partial differential equation. It is possible to transform this equation into a system of first-order ordinary differential equations by using the method of characteristics. Then, the solutions of the ordinary differential equations are given as the characteristic curves, which compose the solution surface of the original partial differential equation.

Consider eikonal equation (7.17), namely,

$$p^2 = \frac{1}{v^2(\mathbf{x}, \mathbf{p})}, \tag{8.1}$$

where, $p^2 = \mathbf{p} \cdot \mathbf{p}$, and, in view of definition (7.12),

$$p_i := \frac{\partial \psi}{\partial x_i}, \qquad i \in \{1, 2, 3\}. \tag{8.2}$$

We wish to solve this equation for $\mathbf{p}(\mathbf{x})$.

The solution of the eikonal equation is a surface in the $\mathbf{xp}$-space. We have a choice of several implicit descriptions of this surface as level sets of function $F(\mathbf{x}, \mathbf{p})$. The two obvious choices for function F are

$$F(\mathbf{x}, \mathbf{p}) = p^2 - \frac{1}{v^2(\mathbf{x}, \mathbf{p})}, \tag{8.3}$$

and

$$F(\mathbf{x}, \mathbf{p}) = p^2 v^2(\mathbf{x}, \mathbf{p}). \tag{8.4}$$

This way, in view of eikonal equation (8.1), the surfaces are the level sets of functions (8.3) or (8.4), given by

$$F(\mathbf{x}, \mathbf{p}) = 0, \tag{8.5}$$

and

$$F(\mathbf{x}, \mathbf{p}) = 1, \tag{8.6}$$

respectively. Since each formulation has different advantages, both are used in various sections of this book.

8.1.2 Characteristic equations

[2]We seek to construct a solution $\mathbf{p} = \mathbf{p}(\mathbf{x})$ such that equation (8.5) or equation (8.6) is satisfied.[3] In both cases, since $F(\mathbf{x}, \mathbf{p}(\mathbf{x}))$ is constant, it follows that $\mathrm{d}F = 0$, where F is treated as a function of $\mathbf{x}$ only. We can state the differential of F as

$$\mathrm{d}F[\mathbf{x}, \mathbf{p}(\mathbf{x})] = \sum_{i=1}^{3} \frac{\partial F}{\partial x_i} \mathrm{d}x_i + \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial F}{\partial p_j} \frac{\partial p_j}{\partial x_i} \mathrm{d}x_i = 0.$$

Using definition (8.2), we can express it in terms of the eikonal function, ψ, as

$$\mathrm{d}F[\mathbf{x}, \mathbf{p}(\mathbf{x})] = \sum_{i=1}^{3} \frac{\partial F}{\partial x_i} \mathrm{d}x_i + \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial F}{\partial p_j} \frac{\partial}{\partial x_i} \frac{\partial \psi}{\partial x_j} \mathrm{d}x_i = 0.$$

Since $\mathrm{d}x_i \neq 0$ and using the equality of mixed partial derivatives, for each i, we can write

$$\frac{\partial F}{\partial x_i} + \sum_{j=1}^{3} \frac{\partial F}{\partial p_j} \frac{\partial}{\partial x_j} \frac{\partial \psi}{\partial x_i} = 0, \qquad i \in \{1, 2, 3\},$$

which are second-order partial differential equations. Again, using definition (8.2), we can rewrite these equations as

$$\frac{\partial F}{\partial x_i} + \sum_{j=1}^{3} \frac{\partial F}{\partial p_j} \frac{\partial p_i}{\partial x_j} = 0, \qquad i \in \{1, 2, 3\}. \tag{8.7}$$

For each $i \in \{1, 2, 3\}$, we wish to find curves $[\mathbf{x}(s), p_i(s)]$ in the solution surface $p_i = p_i(\mathbf{x})$. This way we get a parametrization of the solution surface by these curves, which are commonly referred to as characteristics. For this purpose, we use geometrical properties of vectors tangent and normal to the solution surface.

---

[2]Characteristic equations discussed in this chapter are different mathematical entities than the equations discussed in Exercises 6.5 and 7.1, which bear the same name.

[3]Readers interested in the method of characteristics for solving first-order partial differential equations might refer to Bleistein, N., (1984) Mathematical methods for wave phenomena: Academic Press, pp. 1 – 27, to Courant, R., and Hilbert, D., (1989) Methods of mathematical physics: John Wiley & Sons., Vol. II, Chapter II, to McOwen, R.C., (1996) Partial differential equations: Methods and applications: Prentice-Hall, Inc., pp. 29 – 38, and to Spivak, M., (1970/1999) A comprehensive introduction to differential geometry: Publish or Perish, Inc., pp. 3 – 28.

Let us consider a given $i \in \{1,2,3\}$. The corresponding equation among equations (8.7) can be written as a scalar product of two vectors given by

$$\left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \frac{\partial F}{\partial p_3}, -\frac{\partial F}{\partial x_i}\right] \cdot \left[\frac{\partial p_i}{\partial x_1}, \frac{\partial p_i}{\partial x_2}, \frac{\partial p_i}{\partial x_3}, -1\right] = 0. \tag{8.8}$$

Following the properties of the scalar product, we conclude that these two vectors are orthogonal to one another in the four-dimensional $x_1x_2x_3p_i$-space.

Also — for a given $i \in \{1,2,3\}$ — the solution surface $p_i = p_i(\mathbf{x})$ can be written as a zero-level surface of function

$$g_i(\mathbf{x}, p_i) = p_i(\mathbf{x}) - p_i. \tag{8.9}$$

Furthermore — for a given $i \in \{1,2,3\}$ — we can obtain vector $\mathbf{n}_i$ normal to the solution surface $p_i = p_i(\mathbf{x})$ as the gradient of function g_i, namely,

$$\mathbf{n}_i = \nabla g_i = \left[\frac{\partial g_i}{\partial x_1}, \frac{\partial g_i}{\partial x_2}, \frac{\partial g_i}{\partial x_3}, \frac{\partial g_i}{\partial p_i}\right]. \tag{8.10}$$

Inserting $g_i(\mathbf{x}, p_i)$, given in expression (8.9), into expression (8.10), we can write

$$\mathbf{n}_i = \left[\frac{\partial (p_i(\mathbf{x}) - p_i)}{\partial x_1}, \frac{\partial (p_i(\mathbf{x}) - p_i)}{\partial x_2}, \frac{\partial (p_i(\mathbf{x}) - p_i)}{\partial x_3}, \frac{\partial (p_i(\mathbf{x}) - p_i)}{\partial p_i}\right],$$

to obtain

$$\mathbf{n}_i = \left[\frac{\partial p_i}{\partial x_1}, \frac{\partial p_i}{\partial x_2}, \frac{\partial p_i}{\partial x_3}, -1\right]. \tag{8.11}$$

Examining equations (8.8) and (8.11), we realize that

$$\mathbf{n}_i \perp \left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \frac{\partial F}{\partial p_3}, -\frac{\partial F}{\partial x_i}\right].$$

Thus, for a given $i \in \{1,2,3\}$, vector $[\partial F/\partial p_1, \partial F/\partial p_2, \partial F/\partial p_3, -\partial F/\partial x_i]$ is tangent to the solution surface $p_i = p_i(\mathbf{x})$. We denote this vector by $\mathbf{t}_i$. This way, for a given $i \in \{1,2,3\}$, we obtain vectors tangent to the solution surface.

To obtain curve $[x_1(s), x_2(s), x_3(s), p_i(s)]$, which is in the solution surface and whose tangent vector is $\mathbf{t}_i$, we solve equations

$$\left\{\begin{aligned} \frac{\mathrm{d}x_1(s)}{\mathrm{d}s} &= \zeta\frac{\partial F}{\partial p_1} \\ \frac{\mathrm{d}x_2(s)}{\mathrm{d}s} &= \zeta\frac{\partial F}{\partial p_2} \\ \frac{\mathrm{d}x_3(s)}{\mathrm{d}s} &= \zeta\frac{\partial F}{\partial p_3} \\ \frac{\mathrm{d}p_i(s)}{\mathrm{d}s} &= -\zeta\frac{\partial F}{\partial x_i} \end{aligned}\right. ,$$

which we can concisely write as

$$\left\{\begin{aligned} \frac{\mathrm{d}x_j(s)}{\mathrm{d}s} &= \zeta\frac{\partial F}{\partial p_j} \\ \frac{\mathrm{d}p_i(s)}{\mathrm{d}s} &= -\zeta\frac{\partial F}{\partial x_i} \end{aligned}\right. , \qquad j \in \{1,2,3\}, \tag{8.12}$$

where ζ is a scaling factor and s is the parameter along the curve. The choice of ζ determines the parametrization, which we will use in Section 8.2.

Such curves are the characteristics of eikonal equation (8.1). We note that the solutions of system (8.12) depend on the initial conditions, which we can write as $x_j(0) = x_j^0$ and $p_i(0) = p_i^0$, where $i, j \in \{1,2,3\}$. The initial point must lie in the solution surface; hence, $p_i^0 = p_i(\mathbf{x}^0)$ for $i \in \{1,2,3\}$.

Since this derivation, which is shown for a given i, must hold for each $i \in \{1,2,3\}$, we can write equations (8.12) as

$$\left\{\begin{aligned} \frac{\mathrm{d}x_j}{\mathrm{d}s} &= \zeta\frac{\partial F}{\partial p_j} \\ \frac{\mathrm{d}p_i}{\mathrm{d}s} &= -\zeta\frac{\partial F}{\partial x_i} \end{aligned}\right. , \qquad i, j \in \{1,2,3\},$$

which, in view of i and j being the summation indices, we can restate as

$$\left\{\begin{aligned} \frac{\mathrm{d}x_i}{\mathrm{d}s} &= \zeta\frac{\partial F}{\partial p_i} \\ \frac{\mathrm{d}p_i}{\mathrm{d}s} &= -\zeta\frac{\partial F}{\partial x_i} \end{aligned}\right. , \qquad i \in \{1,2,3\}. \tag{8.13}$$

Hence, three second-order partial differential equations (8.7) become six first-order ordinary differential equations (8.13). These are the characteristic equations, which compose the solution surface of the eikonal equation.

Note that F is constant along $x_i(s), p_i(s)$ independently of parameter s. To see it, consider

$$\frac{\mathrm{d}F(\mathbf{x},\mathbf{p})}{\mathrm{d}s}=\sum_{i=1}^{3}\left(\frac{\partial F}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}s}+\frac{\partial F}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}s}\right),$$

which, in view of equations (8.13), we can rewrite as

$$\frac{\mathrm{d}F(\mathbf{x},\mathbf{p})}{\mathrm{d}s}=\sum_{i=1}^{3}\left(\zeta\frac{\partial F}{\partial x_i}\frac{\partial F}{\partial p_i}-\zeta\frac{\partial F}{\partial p_i}\frac{\partial F}{\partial x_i}\right)=0,$$

as required.

8.1.3 Consistency of formulation

As stated in Section 8.1.1, there are two obvious forms of function F. Functions (8.3) and (8.4) differ in certain aspects, such as their homogeneity with respect to the variables p_i. However, as stated by the following lemma, they both result in the same characteristic equations and, hence, the same characteristic curves.

Lemma 8.1 *Both formulations of the function given by expressions (8.3) and (8.4) result in the same characteristic curves.*

Proof. Consider characteristic equations (8.13), namely,

$$\left\{\begin{array}{l}\dfrac{\mathrm{d}x_i}{\mathrm{d}s}=\zeta\dfrac{\partial F}{\partial p_i}\\[2ex]\dfrac{\mathrm{d}p_i}{\mathrm{d}s}=-\zeta\dfrac{\partial F}{\partial x_i}\end{array}\right.,\qquad i\in\{1,2,3\}.$$

Letting $F=p^2v^2(\mathbf{x},\mathbf{p})$ and setting $\zeta=1$, we note that equations (8.13) become

$$\left\{\begin{array}{l}\dfrac{\mathrm{d}x_i}{\mathrm{d}s}=2\left(p_iv^2+p^2v\dfrac{\partial v}{\partial p_i}\right)\\[2ex]\dfrac{\mathrm{d}p_i}{\mathrm{d}s}=-2p^2v\dfrac{\partial v}{\partial x_i}\end{array}\right.,\qquad i\in\{1,2,3\}.$$

We let $F = p^2 - 1/v^2(\mathbf{x}, \mathbf{p})$ and equations (8.13) become

$$\begin{cases} \dfrac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta\left(p_i + \dfrac{1}{v^3}\dfrac{\partial v}{\partial p_i}\right) \\ \\ \dfrac{\mathrm{d}p_i}{\mathrm{d}s} = -2\zeta\dfrac{1}{v^3}\dfrac{\partial v}{\partial x_i} \end{cases}, \qquad i \in \{1,2,3\}.$$

Equating the second equations of each set, we can write

$$p^2 v\frac{\partial v}{\partial x_i} = \zeta\frac{1}{v^3}\frac{\partial v}{\partial x_i}, \qquad i \in \{1,2,3\}.$$

Solving for ζ, we obtain

$$\zeta = \frac{p^2 v\dfrac{\partial v}{\partial x_i}}{\dfrac{1}{v^3}\dfrac{\partial v}{\partial x_i}} = p^2 v^4, \qquad i \in \{1,2,3\}.$$

Substituting $\zeta = p^2 v^4$ into the first equation of the second set, we obtain

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2p^2v^4\left(p_i + \frac{1}{v^3}\frac{\partial v}{\partial p_i}\right) = 2\left(p_i p^2 v^4 + p^2 v\frac{\partial v}{\partial p_i}\right), \qquad i \in \{1,2,3\},$$

which is equivalent to the first equation of the first set along $p^2v^2 = 1$. ∎

Thus, following equations (8.13), both expressions (8.5) and (8.6), yield the same characteristic curves, given that $\zeta = v^2$ and $\zeta = 1$, respectively.

8.2 Time parametrization of characteristic equations

8.2.1 General formulation

Different choices of ζ result in different parametrization of the solution curves for the characteristic equations. For seismological studies, it is often convenient to parametrize characteristic equations (8.13) in terms of time. Since, in view of trial solution (7.5), the values of $\psi(\mathbf{x})$ are expressed in terms of time, the differential of the eikonal function can be written as

$$\mathrm{d}\psi(\mathbf{x}) = \mathrm{d}t,$$

where t denotes time. Differentiating with respect to s, we obtain

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s}.$$

This equation governs the propagation of $\psi(\mathbf{x})$ along the characteristic curves. The physical interpretation of parameter s depends on the choice of scaling factor ζ in system (8.13). If the parameter s is to be equivalent to time, t, we require that

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = 1.$$

We can restate the above condition as

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = \sum_{i=1}^{3} \frac{\partial\psi}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}s} = 1. \tag{8.14}$$

Using definition (8.2), we rewrite condition (8.14) as

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = \sum_{i=1}^{3} p_i\frac{\mathrm{d}x_i}{\mathrm{d}s} = 1, \tag{8.15}$$

which is a condition for the time parametrization of characteristic equations (8.13).

8.2.2 Equations with variable scaling factor

In order to obtain the time parametrization of system (8.13) in the context of function (8.3), we can write $\mathrm{d}x_i/\mathrm{d}s = \zeta\partial F/\partial p_i$, where $F = p^2 - 1/v^2$, as

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta\left(p_i + \frac{1}{v^3}\frac{\partial v}{\partial p_i}\right), \qquad i \in \{1,2,3\}.$$

In view of condition (8.15), we require that

$$\sum_{i=1}^{3} p_i\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta\sum_{i=1}^{3} p_i\left(p_i + \frac{1}{v^3}\frac{\partial v}{\partial p_i}\right) = 2\zeta\left(p^2 + \frac{1}{v^3}\sum_{i=1}^{3} p_i\frac{\partial v}{\partial p_i}\right) = 1.$$

Since v is homogeneous of degree 0 in the p_i, the summation on the right-hand side vanishes by Theorem A.1. Thus, we obtain

$$\sum_{i=1}^{3} p_i\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta p^2 = 1,$$

and solving for ζ, we immediately get $\zeta = 1/\left(2p^2\right)$.

Consequently, given function (8.3), the system of characteristic equations (8.13) that is parametrized in terms of time becomes

$$\left\{\begin{array}{l} \dot{x}_i = \dfrac{1}{2p^2}\dfrac{\partial F}{\partial p_i} \\ \\ \dot{p}_i = -\dfrac{1}{2p^2}\dfrac{\partial F}{\partial x_i} \end{array}\right., \qquad i \in \{1,2,3\}, \tag{8.16}$$

where $\dot{x}_i := \mathrm{d}x_i/\mathrm{d}t$ and $\dot{p}_i := \mathrm{d}p_i/\mathrm{d}t$. Equations (8.16) are characteristic equations (8.13) whose scaling factor is a function of the p_i. In view of eikonal equation (8.1), we can also state this scaling factor as $v^2(\mathbf{x},\mathbf{p})/2$.

An implication of this parametrization is shown in Exercise 8.4. An implication of another parametrization of characteristic equations (8.13) in the context of function (8.3) is shown in Exercise 8.3.

8.2.3 Equations with constant scaling factor

In order to obtain the time parametrization of system (8.13) in the context of function (8.4), we can write $\mathrm{d}x_i/\mathrm{d}s = \zeta\partial F/\partial p_i$, where $F = p^2v^2$, as

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta\left(p_iv^2 + p^2v\frac{\partial v}{\partial p_i}\right), \qquad i \in \{1,2,3\}.$$

In view of condition (8.15), we require

$$\sum_{i=1}^{3} p_i\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta\sum_{i=1}^{3} p_i\left(p_iv^2 + p^2v\frac{\partial v}{\partial p_i}\right) = 2\zeta\left(p^2v^2 + p^2v\sum_{i=1}^{3} p_i\frac{\partial v}{\partial p_i}\right) = 1.$$

Following the eikonal equation, the first product in parentheses on the right-hand side is equal to unity. Since v is homogeneous of degree 0 in the p_i, the summation on the right-hand side vanishes by Theorem A.1. Thus, we obtain

$$\sum_{i=1}^{3} p_i\frac{\mathrm{d}x_i}{\mathrm{d}s} = 2\zeta = 1,$$

and solving for ζ, we immediately get $\zeta = 1/2$.

Consequently, given function (8.4), system (8.13) is parametrized in terms of time if

$$\left\{\begin{aligned} \dot{x}_i &= \frac{1}{2}\frac{\partial F}{\partial p_i} \\ \dot{p}_i &= -\frac{1}{2}\frac{\partial F}{\partial x_i} \end{aligned}\right., \qquad i \in \{1,2,3\}, \tag{8.17}$$

where $\dot{x}_i := \mathrm{d}x_i/\mathrm{d}t$ and $\dot{p}_i := \mathrm{d}p_i/\mathrm{d}t$. Equations (8.17) are characteristic equations (8.13) whose scaling factor is the constant equal to $1/2$.

In view of functions (8.3) and (8.4), the corresponding scaling factors, $\zeta = v^2/2$ and $\zeta = 1/2$, are consistent with one another. This can be seen by examining the proof of Lemma 8.1.

8.2.4 Formulation of Hamilton's ray equations

We now examine systems (8.16) and (8.17), and choose to proceed with the latter one since, therein, ζ is given by a constant. This constant can be brought inside the differential operator and we can write system (8.17) as

$$\left\{\begin{aligned} \dot{x}_i &= \frac{\partial}{\partial p_i}\left(\frac{F}{2}\right) \\ \dot{p}_i &= -\frac{\partial}{\partial x_i}\left(\frac{F}{2}\right) \end{aligned}\right., \qquad i \in \{1,2,3\}. \tag{8.18}$$

Defining $\mathcal{H} := F/2$, where $\mathcal{H}$ is referred to as the ray-theory Hamiltonian[4], we can write equations (8.18) as

$$\left\{\begin{aligned} \dot{x}_i &= \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial \mathcal{H}}{\partial x_i} \end{aligned}\right., \qquad i \in \{1,2,3\}. \tag{8.19}$$

Equations (8.19) constitute a system of first-order ordinary differential equations in t for $\mathbf{x}(t)$ and $\mathbf{p}(t)$. These equations are Hamilton's ray equations.

[4]In this book we use two distinct Hamiltonians denoted by $\mathcal{H}$ and $\mathbb{H}$. Consequently, in the text, we avoid a generic reference to "the Hamiltonian", unless it is clear from the context which one of the two is considered.

System (8.19) governs the signal trajectories in the **xp**-space spanned by the position vectors, $\mathbf{x}$, and the phase-slowness vectors, $\mathbf{p}$.

The first set of equations of system (8.19) corresponds to the components of vectors tangent to the so-called base characteristic curves, namely, $\mathbf{x}(t)$. These curves belong to the physical $\mathbf{x}$-space. They are trajectories along which signals propagate in the physical space and, in the context of ray theory, they are rays.

The second set of equations of system (8.19) describes the rate of change of the phase slowness. If $\mathcal{H}$ is not explicitly a function of a given x_i, we obtain $\dot{p}_i = 0$, which implies that p_i is constant along the ray. Hence, in such a case, p_i is a conserved quantity, known as the ray parameter, which is discussed in Chapter 14. Physically, this means that $v(\mathbf{x}, \mathbf{p})$ is not explicitly a function of x_i and, hence, the continuum is homogeneous along that component.

Note that, in the context of Legendre's transformation, discussed in Appendix B, the first set of equations can be viewed as a definition of a variable, while the essence of the physical formulation is contained in the second set of equations.

The ray-theory Hamiltonian, $\mathcal{H}$, resulting from function (8.4), can be explicitly stated as

$$\mathcal{H} = \frac{1}{2}p^2 v^2 (\mathbf{x}, \mathbf{p}) . \tag{8.20}$$

It is a dimensionless quantity, unlike the classical-mechanics Hamiltonian, discussed in Chapter 13, which has units of energy. In view of eikonal equation (7.17), which states that $p^2v^2 = 1$, and expression (8.20), we require that $\mathcal{H}(\mathbf{x}, \mathbf{p}) = 1/2$, along a ray.

8.3 Example: Ray equations in isotropic inhomogeneous continua

8.3.1 Parametric form

System (8.19) allows us to study ray theory in the context of anisotropic inhomogeneous continua. To gain familiarity with such a system, consider a formulation for isotropic inhomogeneous continua, where eikonal equation (8.1) reduces to

$$p^2 = \frac{1}{v^2(\mathbf{x})}, \tag{8.21}$$

which is eikonal equation (6.70).

To study ray equations in isotropic inhomogeneous continua, let us choose function (8.3), which becomes

$$F(\mathbf{x}) = p^2 - \frac{1}{v^2(\mathbf{x})}. \tag{8.22}$$

Using system (8.13), we can write the corresponding characteristic equations as

$$\left\{ \begin{aligned} \frac{\mathrm{d}x_i}{\mathrm{d}s} &= 2\zeta p_i \\ \\ \frac{\mathrm{d}p_i}{\mathrm{d}s} &= -2\zeta \frac{1}{v^3}\frac{\partial v}{\partial x_i} \end{aligned} \right. , \qquad i \in \{1,2,3\}. \tag{8.23}$$

Also, let us choose scaling factor ζ so that s is the arclength parameter. As shown in Exercise 8.2, we obtain the arclength parametrization of system (8.23) by letting $\zeta = v/2$. Furthermore, as shown in Exercise 8.3, system (8.23) can be restated as a single expression

$$\frac{\mathrm{d}}{\mathrm{d}s}\left[\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\right] = -\frac{\nabla v(\mathbf{x})}{v^2(\mathbf{x})}, \tag{8.24}$$

where $\mathbf{x} = [x_1, x_2, x_3]$.

Equation (8.24) relates the properties of the continuum, which are given by the phase-velocity function $v(\mathbf{x})$, to the ray $\mathbf{x}(s)$, which is described by arclength parameter s.

8.3.2 Explicit form

Consider a three-dimensional isotropic inhomogeneous continuum where $\mathbf{x} = [x, y, z]$. Expression (8.24) can be explicitly written as three parametric equations for $x(s)$, $y(s)$ and $z(s)$, namely,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}x}{\mathrm{d}s}\right) &= -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial x}, \\ \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}y}{\mathrm{d}s}\right) &= -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial y}, \\ \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}z}{\mathrm{d}s}\right) &= -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial z}, \end{aligned} \tag{8.25}$$

where s is the arclength parameter along the ray. Consequently, all three equations are related by $\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2}$, where x, y and z

are the orthonormal coordinates. Consequently, as shown in Exercise 8.6, instead of using the parametric form, under certain conditions related to the behaviour of the curve $\mathbf{x}(s)$, we can write equations (8.25) as two equations for $x(z)$ and $y(z)$, namely,

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{1}{v(\mathbf{x})}\frac{\frac{\mathrm{d}x}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}}\right] = -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial x}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}, \quad (8.26)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{1}{v(\mathbf{x})}\frac{\frac{\mathrm{d}y}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}}\right] = -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial y}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}, \quad (8.27)$$

which form a system of explicit equations for isotropic inhomogeneous continua.

If the continuum exhibits only vertical inhomogeneity, $v = v(z)$, the right-hand sides of equations (8.26) and (8.27) vanish and, for the resulting equations, we can obtain an analytical solution, as shown in Exercise 8.7. If, however, the properties of the medium vary along the x-axis and the y-axis, we must often resort to numerical methods to obtain a solution.[5]

Closing remarks

By solving the eikonal equation using the method of characteristics, we obtain Hamilton's ray equations whose solutions give rays. Hamilton's ray

[5]Readers interested in numerical techniques to solve these differential equations might refer to Červený, V., and Ravindra, R., (1971) Theory of seismic head waves: University of Toronto Press, pp. 25 – 26.

equations are rooted in the high-frequency approximation and the trial solutions discussed in Chapters 6 and 7, and the resulting rays are given by the function $\mathbf{x} = \mathbf{x}(t)$. We can study the entire ray theory in the context of Hamilton's ray equations, which is the most rigorous method for studying seismic rays, traveltimes, wavefronts and amplitudes.

In Chapter 9, however, we will explore another formulation of rays using an approach that transforms Hamilton's six first-order equations into Lagrange's three second-order equations. Also, this Lagrangian formulation coincides with the variational approach to the study of ray theory, which we will discuss in Part III. By investigating both of these approaches, we gain additional physical insight into ray theory, as well as additional knowledge of useful mathematical tools.

In general, ray theory is related to the Wentzel-Kramers-Brillouin-Jeffreys (WKBJ) method for solving differential equations. The WKBJ method is also used in other physical theories, for instance, in quantum mechanics.[6] Ray theory is an approximation of wave theory as classical mechanics is an approximation of quantum mechanics. The high-frequency approximation is analogous to assuming the action, discussed in Section 13.2.1, to be infinitely divisible, as is the case in classical mechanics. This is not the case in quantum mechanics due to the existence of Planck's constant, which is the fundamental unit, or quantum, of action.[7]

Exercises

Exercise 8.1 *Consider a three-dimensional isotropic inhomogeneous continuum. Using Hamilton's ray equations (8.19), show that, in isotropic continua, rays are orthogonal to wavefronts.*

Solution 8.1 *Following expression (8.20), we can explicitly write Hamiltonian $\mathcal{H}(\mathbf{x}, \mathbf{p})$, in a three-dimensional isotropic inhomogeneous continuum,*

[6] Readers interested in the WKBJ method might refer to Aki, K. and Richards, P.G., (2002) Quantitative seismology (2*nd* edition): University Science Books, pp. 434 – 437, (Box 9.6), and to Woodhouse, N.M.J., (1992) Geometric quantization (2*nd* edition): Oxford Science Publications, pp. 197 – 201 and pp. 236 – 249.

[7] Readers interested in the association of the geometrical optics and quantum mechanics might refer to Goldstein, H., (1950/1980) Classical mechanics: Addison-Wesley Publishing Co., pp. 484 – 492.

as

$$\mathcal{H}(\mathbf{x},\mathbf{p}) = \frac{1}{2}p^2v^2(\mathbf{x}) = \frac{1}{2}[p_1,p_2,p_3]\cdot[p_1,p_2,p_3]\,v^2(x_1,x_2,x_3)\,. \tag{8.28}$$

The corresponding Hamilton's ray equations (8.19) are

$$\begin{cases} \dot{x}_1 = p_1v^2 \\ \dot{x}_2 = p_2v^2 \\ \dot{x}_3 = p_3v^2 \\ \dot{p}_1 = -p^2v\dfrac{\partial v}{\partial x_1} \\ \dot{p}_2 = -p^2v\dfrac{\partial v}{\partial x_2} \\ \dot{p}_3 = -p^2v\dfrac{\partial v}{\partial x_3} \end{cases}. \tag{8.29}$$

Recalling definition (7.12), we can write the first three equations of system (8.29) as

$$[\dot{x}_1,\dot{x}_2,\dot{x}_3] = v^2\left[\frac{\partial\psi}{\partial x_1},\frac{\partial\psi}{\partial x_2},\frac{\partial\psi}{\partial x_3}\right]. \tag{8.30}$$

The left-hand side of equation (8.30) is a vector tangent to the curve $\mathbf{x}(t)$*, while the right-hand side is the gradient of function* $\psi(\mathbf{x})$*, scaled by* v^2*. For a given point of the continuum, we can write equation (8.30) as*

$$\mathbf{t}|_{x_1,x_2,x_3} = v^2(\nabla\psi)\big|_{x_1,x_2,x_3}\,.$$

This means that vector $\mathbf{t}$*, which is tangent to curve* $\mathbf{x}(t)$*, is parallel to the gradient of the eikonal function,* $\nabla\psi(\mathbf{x})$*. Since curve* $\mathbf{x}(t)$ *corresponds to the ray and the level sets of the eikonal function correspond to the wavefronts, by the properties of the gradient operator, the rays in an isotropic inhomogeneous continuum are orthogonal to the wavefronts.*

Remark 8.1 *Characteristic equations (8.13) can be parametrized by choosing various expressions for scaling factor* ζ*. Two typical examples are shown in Exercises 8.3 and 8.4, below. In both cases, we invoke function (8.3) and consider isotropic inhomogeneous continua. Hence, characteristic equations (8.13) become equations (8.23).*

Exercise 8.2 [8]*Show that the arclength parametrization of system (8.23) requires* $\zeta = v/2$*.*

[8] See also Section 8.3.1.

Solution 8.2 *In general, if* $\mathbf{x} = \mathbf{x}(s)$, *using definition (8.2), we can write*

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = \sum_{i=1}^{3} \frac{\partial \psi}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}s} = \sum_{i=1}^{3} p_i \frac{\mathrm{d}x_i}{\mathrm{d}s},$$

which, in view of characteristic equations (8.23), we can rewrite as

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = \sum_{i=1}^{3} 2\zeta p_i p_i,$$

which we can immediately restate as

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s} = 2\zeta p^2. \tag{8.31}$$

If s *is the arclength parameter, then*

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2},$$

and, hence,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}, \tag{8.32}$$

where t *stands for traveltime. Combining expressions (8.31) and (8.32), we obtain*

$$\frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}t} = \frac{\mathrm{d}\psi(\mathbf{x})}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t} = 2\zeta p^2 \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}. \tag{8.33}$$

In view of condition (8.14) and since the square root gives the magnitude of velocity, we can rewrite equation (8.33) as

$$1 = 2\zeta p^2 v.$$

Solving for ζ, *where — in view of equation (8.21) — we use* $p^2 v^2 = 1$, *we get*

$$\zeta = \frac{1}{2p^2 v} = \frac{v}{2p^2 v^2} = \frac{v}{2},$$

as required.

Exercise 8.3 [9]*Letting $\zeta = v/2$, show that characteristic equations (8.23) can be reduced to equation (8.24), namely,*

$$\frac{\mathrm{d}}{\mathrm{d}s}\left[\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\right] = -\frac{\nabla v(\mathbf{x})}{v^2(\mathbf{x})}.$$

Solution 8.3 *If $\zeta = v/2$, characteristic equations (8.23) become*

$$\left\{\begin{array}{l} \dfrac{\mathrm{d}x_i}{\mathrm{d}s} = vp_i \\ \\ \dfrac{\mathrm{d}p_i}{\mathrm{d}s} = -\dfrac{1}{v^2}\dfrac{\partial v}{\partial x_i} \end{array}\right. , \qquad i \in \{1,2,3\}.$$

The first equation of this system can be rewritten as

$$p_i = \frac{1}{v}\frac{\mathrm{d}x_i}{\mathrm{d}s}, \qquad i \in \{1,2,3\}.$$

Hence, the second equation can be stated as

$$\frac{\mathrm{d}p_i}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v}\frac{\mathrm{d}x_i}{\mathrm{d}s}\right) = -\frac{1}{v^2}\frac{\partial v}{\partial x_i}, \qquad i \in \{1,2,3\}.$$

Thus, the system of characteristic equations can be written as a single expression

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\right) = -\frac{\nabla v(\mathbf{x})}{v^2(\mathbf{x})},$$

where $\mathbf{x} = [x_1, x_2, x_3]$, which is equation (8.24), as required.

Exercise 8.4 [10]*Letting $\zeta = v^2/2$, show that characteristic equations (8.23) can be written as a system of equations given by*

$$\left\{\begin{array}{l} \dot{x}_i = v^2 p_i \\ \\ \dot{p}_i = -\dfrac{\partial}{\partial x_i}\ln v \end{array}\right. , \qquad i \in \{1,2,3\}.$$

Solution 8.4 *As shown in Section 8.2, using function (8.3) and letting $\zeta = v^2/2$ results in the time parametrization of characteristic equations (8.13).*

[9]See also Section 8.3.1.

[10]See also Section 8.2.2 and Exercise 13.3.

Hence, characteristic equations (8.23) can be written as

$$\left\{\begin{array}{l} \dot{x}_i := \dfrac{\mathrm{d}x_i}{\mathrm{d}t} = v^2 p_i \\ \\ \dot{p}_i := \dfrac{\mathrm{d}p_i}{\mathrm{d}t} = -\dfrac{1}{v}\dfrac{\partial v}{\partial x_i} \end{array}\right. , \qquad i \in \{1,2,3\}.$$

Following the chain rule, we can restate the second equation of this system to obtain

$$\left\{\begin{array}{l} \dot{x}_i = v^2 p_i \\ \\ \dot{p}_i = -\dfrac{\partial}{\partial x_i}\ln v \end{array}\right. , \qquad i \in \{1,2,3\},$$

as required.

Remark 8.2 *Lemma 8.1 shows that both functions (8.3) and (8.4) yield the same characteristics. Thus, in a seismological context, both functions result in the same rays. In view of Exercise 8.4, Exercise 8.5 illustrates this property for isotropic inhomogeneous continua.*

Exercise 8.5 [11] *Using characteristic equations (8.17) and considering functions (8.4), show that, for isotropic inhomogeneous continua, we obtain the system of equations*

$$\left\{\begin{array}{l} \dot{x}_i = v^2 p_i \\ \\ \dot{p}_i = -\dfrac{\partial}{\partial x_i}\ln v \end{array}\right. , \qquad i \in \{1,2,3\}.$$

Solution 8.5 *Considering functions (8.4) for isotropic inhomogeneous continua, characteristic equations (8.17), which are parametrized in terms of time, become*

$$\left\{\begin{array}{l} \dot{x}_i = v^2 p_i \\ \\ \dot{p}_i = -p^2 v \dfrac{\partial v}{\partial x_i} \end{array}\right. , \qquad i \in \{1,2,3\}.$$

Since $p^2 v^2 = 1$, we can write

$$\left\{\begin{array}{l} \dot{x}_i = v^2 p_i \\ \\ \dot{p}_i = -\dfrac{1}{v}\dfrac{\partial v}{\partial x_i} = -\dfrac{\partial}{\partial x_i}\ln v \end{array}\right. , \qquad i \in \{1,2,3\},$$

[11] See also Exercise 13.3.

which is also the solution of Exercise 8.4.

Exercise 8.6 *Formally, show the steps leading from set (8.25) to equations (8.26) and (8.27).*

Solution 8.6 *The first two equations can be written as*

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}x}{\mathrm{d}s}\right) = -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial x},$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}y}{\mathrm{d}s}\right) = -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial y},$$

which leads to

$$\frac{\mathrm{d}s}{\mathrm{d}z}\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}x}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}s}\right) = -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial x}\frac{\mathrm{d}s}{\mathrm{d}z},$$

$$\frac{\mathrm{d}s}{\mathrm{d}z}\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}y}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}s}\right) = -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial y}\frac{\mathrm{d}s}{\mathrm{d}z},$$

where we multiplied both sides of the equations by ds/dz, and we multiplied the factors inside the parentheses by unity in the form $(\mathrm{d}s/\mathrm{d}z)(\mathrm{d}z/\mathrm{d}s)$. *The two equations can be immediately restated as*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}x}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}s}\right) &= -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial x}\frac{\mathrm{d}s}{\mathrm{d}z},\\ \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{v(\mathbf{x})}\frac{\mathrm{d}y}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}s}\right) &= -\frac{1}{v^2(\mathbf{x})}\frac{\partial v(\mathbf{x})}{\partial y}\frac{\mathrm{d}s}{\mathrm{d}z}.\end{aligned} \tag{8.34}$$

We assume the invertibilty of function $z = z(s)$, *which allows us to write* $s = s(z)$. *Furthermore, we assume that the behaviour of the space curve* $[x(s), y(s), z(s)]$ *allows us to express it as* $[x(z), y(z)]$. *Consequently, from formal operations, we get*

$$\frac{\mathrm{d}s}{\mathrm{d}z} \equiv \frac{\mathrm{d}s(x(z), y(z), z)}{\mathrm{d}z} = \frac{\sqrt{[\mathrm{d}x(z)]^2 + [\mathrm{d}y(z)]^2 + [\mathrm{d}z(z)]^2}}{\mathrm{d}z}$$

$$= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}z}\right)^2}.$$

Thus, since $\mathrm{d}s/\mathrm{d}z = 1/\left(\mathrm{d}z/\mathrm{d}s\right)$*, equations (8.34) can be stated as*

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{v\left(\mathbf{x}\right)}\frac{\frac{\mathrm{d}x}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}}\right) = -\frac{1}{v^2\left(\mathbf{x}\right)}\frac{\partial v\left(\mathbf{x}\right)}{\partial x}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1},$$

and

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{v\left(\mathbf{x}\right)}\frac{\frac{\mathrm{d}y}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}}\right) = -\frac{1}{v^2\left(\mathbf{x}\right)}\frac{\partial v\left(\mathbf{x}\right)}{\partial y}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1},$$

which — as required — are equations (8.26) and (8.27), respectively.

Exercise 8.7 *Solve ray equations (8.26) and (8.27) for a vertically inhomogeneous continuum, where* $v = v\left(z\right)$*.*

Solution 8.7 *Since* $v = v\left(z\right)$*, the right-hand sides of equations (8.26) and (8.27) vanish. Consequently, we obtain*

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{1}{v\left(z\right)}\frac{\frac{\mathrm{d}x}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}}\right] = 0, \qquad (8.35)$$

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{1}{v\left(z\right)}\frac{\frac{\mathrm{d}y}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2+1}}\right] = 0.$$

Since the velocity gradient is present only along the z*-axis, the ray is contained in a single vertical plane. Thus, with no loss of generality, we can* *assume that a given ray is contained in the xz-plane and, hence, consider only equation (8.35). In view of the vanishing of the total derivative, equation (8.35) can be restated as*

$$\frac{1}{v(z)}\frac{\frac{\mathrm{d}x}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+1}}=\mathfrak{p}, \tag{8.36}$$

where $\mathfrak{p}$ *is a constant. Equation (8.36), can be rewritten as*

$$\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2=\mathfrak{p}^2v^2\left[\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+1\right].$$

Solving for $\mathrm{d}x/\mathrm{d}z$, *we obtain*

$$\frac{\mathrm{d}x}{\mathrm{d}z}=\frac{\mathfrak{p}v}{\sqrt{1-\mathfrak{p}^2v^2}}, \tag{8.37}$$

and, hence we can state the solution as

$$x(z)=\int\limits_{z_0}^{z}\frac{\mathfrak{p}v(\xi)}{\sqrt{1-\mathfrak{p}^2v^2(\xi)}}\,\mathrm{d}\xi,$$

where ξ *is the integration variable. This is a standard expression for a ray in vertically inhomogeneous continua, where, as shown in Exercise 8.8,* $\mathfrak{p}=\sin\theta/v(z)$.

Exercise 8.8 *Consider equation (8.36). Show that* $\mathfrak{p}=\sin\theta/v(z)$.

Solution 8.8 *Since* $\mathrm{d}x/\mathrm{d}z=\tan\theta$, *following standard trigonometric identities, we can write equation (8.36) as*

$$\mathfrak{p}=\frac{1}{v(z)}\frac{\frac{\mathrm{d}x}{\mathrm{d}z}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2+1}}=\frac{1}{v(z)}\frac{\tan\theta}{\sqrt{\tan^2\theta+1}}=\frac{\frac{\sin\theta}{\cos\theta}}{v(z)\sec\theta}=\frac{\sin\theta}{v(z)}.$$

Exercise 8.9 *Consider a one-dimensional homogeneous continuum. Show that solution* $x(t)$ *of Hamilton's ray equations (8.19) corresponds to coordinates (6.42), which can be written as*

$$x(t)=x_0\pm vt.$$

Solution 8.9 *For a one-dimensional case, letting* $x_1 \equiv x$ *and* $p_1 \equiv p$*, we can write Hamilton's ray equations (8.19) as*

$$\begin{cases} \dot{x} = pv^2 + p^2 v \dfrac{\partial v}{\partial p} \\ \\ \dot{p} = -p^2 v \dfrac{\partial v}{\partial x} \end{cases},$$

To study solution $x(t)$*, we consider the first equation. In elasticity theory, a one-dimensional continuum must be isotropic, hence,* $\partial v/\partial p = 0$*. Thus, we obtain*

$$\dot{x} = pv^2.$$

Since, in the one-dimensional case, p *is the magnitude of the phase-slowness vector, we can write*

$$\dot{x} = \frac{1}{p} p^2 v^2.$$

In view of eikonal equation (7.17) and since $v = \pm 1/p$*, we can write*

$$\dot{x} := \frac{\mathrm{d}x}{\mathrm{d}t} = \pm v.$$

Solving for $\mathrm{d}x$*, we obtain*

$$\mathrm{d}x = \pm v \mathrm{d}t.$$

Integrating both sides, we obtain

$$x(t) = x_0 \pm vt,$$

as required and where x_0 *is the integration constant.*

Exercise 8.10 *Following ray equation (8.24), show that rays are straight lines in homogeneous continua.*

Solution 8.10 *For homogeneous continua,* v *is constant and, hence, the right-hand side of ray equation (8.24) vanishes. Thus, we obtain*

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{v}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\right) = \mathbf{0}.$$

The vanishing of the total derivative implies that the term in parentheses can be written as

$$\frac{1}{v}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = \mathbf{C},$$

where $\mathbf{C}$ *denotes a constant vector. Rearranging and integrating gives*

$$\mathbf{x} = \mathbf{a}s + \mathbf{b},$$

which is an equation of a straight line, where $\mathbf{a} := \mathbf{C}v$.

This Page Intentionally Left Blank

Chapter 9

Lagrange's ray equations

The ancient Greeks had a hard time defining objects like "curves" and "surfaces" in a general way since their algebra was not well developed and always remained on a rather modest level. In fact, some historians think that the final stagnation of Greek mathematics was caused by the Greeks' failure to develop algebra and to apply it to geometry.

Stephan Hildebrandt and Anthony Tromba (1996) The parsimonious universe

Preliminary remarks

In Chapter 8, we obtained Hamilton's ray equations, which allow us to study seismic signals in an anisotropic inhomogeneous continuum. In a three-dimensional continuum, Hamilton's equations constitute a system of six ordinary first-order differential equations, which are expressed in terms of Hamiltonian $\mathcal{H}$ and exist in the $\mathbf{xp}$-space. This system can be also expressed as a system consisting of three ordinary second-order differential equations, which are expressed in terms of Lagrangian $\mathcal{L}$, where function $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ is Legendre's transformation of function $\mathcal{H}(\mathbf{x}, \mathbf{p})$. These second-order differential equations are Lagrange's ray equations, which exist in the $\mathbf{x}\dot{\mathbf{x}}$-space.[1]

We begin this chapter by transforming Hamilton's ray equations into Lagrange's ray equations. Subsequently, we relate the orientations and mag-

[1]In classical mechanics, the $\mathbf{x}\dot{\mathbf{x}}$-space corresponds to the velocity phase space. In this book, however, to avoid the confusion with the term "phase" that we use in the specific context of wave phenomena, we do not use this nomenclature.

nitudes of vectors $\mathbf{p}$ and $\dot{\mathbf{x}}$, which result in expressions relating phase and ray angles as well as phase and ray velocities.

Readers who are not familiar with Legendre's transformation might find it useful to study this chapter together with Appendix B.

9.1 Transformation of Hamilton's ray equations

9.1.1 Formulation of Lagrange's ray equations

To obtain Lagrange's ray equations, consider Hamilton's ray equations (8.19), namely,

$$\left\{\begin{array}{l} \dot{x}_i = \dfrac{\partial \mathcal{H}}{\partial p_i} \\ \\ \dot{p}_i = -\dfrac{\partial \mathcal{H}}{\partial x_i} \end{array}\right. , \qquad i \in \{1,2,3\}. \tag{9.1}$$

System (9.1) is composed of six first-order ordinary differential equations in t to be solved for $\mathbf{x}(t)$ and $\mathbf{p}(t)$. Legendre's transformation, discussed in Appendix B, allows us to express this system as three second-order ordinary differential equations in t to be solved for $\mathbf{x}(t)$.

In view of expression (B.13) in Appendix B, consider a function given by

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \sum_{j=1}^{3} p_j(\mathbf{x}, \dot{\mathbf{x}})\, \dot{x}_j - \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}})), \tag{9.2}$$

where $\mathcal{L}$ is referred to as the ray-theory Lagrangian[2] corresponding to a given $\mathcal{H}$, and $p_j(\mathbf{x}, \dot{\mathbf{x}})$ is a solution of

$$\dot{x}_j = \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial p_j}, \qquad j \in \{1,2,3\}, \tag{9.3}$$

which is equation (B.12). Hence, in view of Appendix B, $\mathcal{L}$ is Legendre's transformation of $\mathcal{H}$. Now, we wish to rewrite Hamilton's ray equations (9.1) in terms of Lagrangian (9.2).

[2] In this book we use four distinct Lagrangians, which are denoted by $\mathcal{L}$, $\mathcal{F}$, $\mathbb{L}$ and F. Consequently, in the text, we avoid a generic reference to "the Lagrangian", unless it is clear from the context which one among the four is considered.

Using expression (9.2), consider its derivative with respect to the first and second arguments, namely, x_i and $\dot{x}_i$, where $i \in \{1,2,3\}$. We obtain

$$\frac{\partial \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i} = \frac{\partial}{\partial x_i}\left[\sum_{j=1}^{3} p_j(\mathbf{x},\dot{\mathbf{x}})\,\dot{x}_j - \mathcal{H}(\mathbf{x},\mathbf{p}(\mathbf{x},\dot{\mathbf{x}}))\right]$$
$$= \sum_{j=1}^{3}\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i}\dot{x}_j - \frac{\partial \mathcal{H}(\mathbf{x},\mathbf{p}(\mathbf{x},\dot{\mathbf{x}}))}{\partial x_i} - \sum_{j=1}^{3}\frac{\partial \mathcal{H}(\mathbf{x},\mathbf{p}(\mathbf{x},\dot{\mathbf{x}}))}{\partial p_j}\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i} \tag{9.4}$$

and

$$\frac{\partial \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i}\left[\sum_{j=1}^{3} p_j(\mathbf{x},\dot{\mathbf{x}})\,\dot{x}_j - \mathcal{H}(\mathbf{x},\mathbf{p}(\mathbf{x},\dot{\mathbf{x}}))\right]$$
$$= \sum_{j=1}^{3}\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i}\dot{x}_j + p_i(\mathbf{x},\dot{\mathbf{x}}) - \sum_{j=1}^{3}\frac{\partial \mathcal{H}(\mathbf{x},\mathbf{p}(\mathbf{x},\dot{\mathbf{x}}))}{\partial p_j}\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i}, \tag{9.5}$$

respectively.

Using Hamilton's ray equations (9.1), we can restate expressions (9.4) and (9.5) as

$$\frac{\partial \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i} = \sum_{j=1}^{3}\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i}\dot{x}_j + \dot{p}_i(\mathbf{x},\dot{\mathbf{x}}) - \sum_{j=1}^{3}\dot{x}_j\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i}, \qquad i \in \{1,2,3\}, \tag{9.6}$$

and

$$\frac{\partial \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i} = \sum_{j=1}^{3}\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i}\dot{x}_j + p_i(\mathbf{x},\dot{\mathbf{x}}) - \sum_{j=1}^{3}\dot{x}_j\frac{\partial p_j(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i}, \qquad i \in \{1,2,3\}, \tag{9.7}$$

respectively.

Examining expressions (9.6) and (9.7), we see that $\partial\mathcal{L}/\partial x_i = \dot{p}_i$ and $\partial\mathcal{L}/\partial\dot{x}_i = p_i$. Hence, we conclude that

$$\frac{\partial \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial x_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial \dot{x}_i}\right), \qquad i \in \{1,2,3\},$$

which we can rewrite as

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_i} = 0, \qquad i \in \{1,2,3\}, \tag{9.8}$$

where $\mathcal{L}$ is given in expression (9.2).

Note that when we introduce $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$, in expression (9.2), $\dot{\mathbf{x}}$ denotes a new variable, which, *a priori*, has no association with $\mathbf{x}$. If we consider the solution of system (9.1), which is given by $(\mathbf{x}(t), \mathbf{p}(t))$, then, in view of $\mathbf{p}(t) = \mathbf{p}(\mathbf{x}(t), \dot{\mathbf{x}}(t))$, we also have the corresponding solution $(\mathbf{x}(t), \dot{\mathbf{x}}(t))$. By examining equation (9.3) together with the first equation of system (9.1), we see that $\mathrm{d}\mathbf{x}(t)/\mathrm{d}t = \dot{\mathbf{x}}(t)$. Consequently, at the end, our initial abuse of notation does no harm and, rather, might be viewed as insightful. In other word, depending on the context, $\dot{\mathbf{x}}$ can be viewed as an independent variable or as a function of t.

Equations (9.8) constitute a system of three second-order ordinary differential equations in t to be solved for $\mathbf{x}(t)$, which is the curve corresponding to the ray. We refer to equations of this form as Lagrange's ray equations.

In view of this derivation, system (9.8) is equivalent to system (9.1). Herein, we have obtained Lagrange's ray equations from Hamilton's ray equations. The duality of Legendre's transformation is such that we can also obtain Hamilton's ray equations from Lagrange's ray equations, as shown in Exercise 9.2 and in Appendix B. This leads to the following proposition.

Proposition 9.1 *Rays, parametrized by time, can be obtained either by solving Hamilton's ray equations (9.1) or by solving Lagrange's ray equations (9.8).*

We note that, in view of Legendre's transformation, the derivation of Lagrange's ray equations requires regularity of Hamiltonian $\mathcal{H}$, namely,

$$\det\left[\frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j}\right] \neq 0, \qquad i, j \in \{1, 2, 3\},$$

which is a necessary condition for Legendre's transformation to be a local diffeomorphism. This limitation is discussed in Section 13.1.

9.1.2 Beltrami's identity

For our subsequent work, we notice that we can write all the equations of system (9.8) as a single equation, namely,

$$\frac{\partial \mathcal{L}}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{3} \dot{x}_i \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \mathcal{L}\right) = 0. \tag{9.9}$$

Equation (9.9) is also valid for an n-dimensional case, where $i \in \{1, \ldots, n\}$. The verification of this expression, for the two-dimensional case, is shown in Exercise 9.1.

We refer to equation (9.9) as Beltrami's identity, since this expression was formulated in 1868 by Eugenio Beltrami. Beltrami's identity plays an important role in our raytracing methods, as illustrated in Section 12.3 and in Chapter 14. This importance results from the fact that if $\mathcal{L}$ does not explicitly depend on t, the first term on the left-hand side in equation (9.9) vanishes and, hence, the term in parentheses is equal to a constant. Furthermore, if $\mathcal{L}$ is homogeneous in the $\dot{x}_i$, the Lagrangian is conserved along the solution, $\mathbf{x}(t)$, as shown in Exercise 13.2.

9.2 Relation between $\mathbf{p}$ and $\dot{\mathbf{x}}$

9.2.1 Phase and ray velocities

General formulation

We wish to study the relation between the orientations and the magnitudes of vectors $\mathbf{p}$ and $\dot{\mathbf{x}}$. Mathematically, the components of these vectors are the variables used in Legendre's transformation. Physically, $\mathbf{p}$ is the vector normal to the wavefront and $\dot{\mathbf{x}}$ is the vector tangent to the ray.

Consider a given point $\mathbf{x}$ of the continuum and, therein, the directional dependence of $\mathcal{H}$. The first set of equations of system (9.1) is

$$\dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \qquad i \in \{1, 2, 3\}. \tag{9.10}$$

Inserting expression (8.20), namely,

$$\mathcal{H} = \frac{1}{2} p^2 v^2 (\mathbf{x}, \mathbf{p}),$$

into equations (9.10) and using the equality resulting from the eikonal equation, namely, $p^2 v^2 = 1$, we obtain

$$\dot{x}_i = p_i v^2 + \frac{1}{v} \frac{\partial v}{\partial p_i}, \qquad i \in \{1, 2, 3\}, \tag{9.11}$$

where the phase-velocity function, v, is a function of the orientation of vector $\mathbf{p}$.

Vector $\dot{\mathbf{x}}$ is tangent to the ray $\mathbf{x}(t)$. Since, at a given point, this vector corresponds to the velocity of the signal along the ray at that point, we refer to it as ray velocity.[3] We wish to find the magnitude of this vector, which can be written as

$$V := |\dot{\mathbf{x}}| = \sqrt{\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}} = \sqrt{\sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij} \dot{x}_i \dot{x}_j}, \tag{9.12}$$

where δ_{ij} is Kronecker's delta. Using expression (9.11), we can write each term of the double summation in radicand (9.12) as

$$\begin{aligned}
(\dot{x}_i)^2 &= \left(p_i v^2 + \frac{1}{v}\frac{\partial v}{\partial p_i}\right)^2 \\
&= (p_i)^2 v^4 + 2 p_i v \frac{\partial v}{\partial p_i} + \frac{1}{v^2}\left(\frac{\partial v}{\partial p_i}\right)^2, \qquad i \in \{1, 2, 3\}.
\end{aligned}$$

Performing the summation of the three terms, we obtain

$$\begin{aligned}
\sum_{i=1}^{3}\left[(p_i)^2 v^4 + 2 p_i v \frac{\partial v}{\partial p_i} + \frac{1}{v^2}\left(\frac{\partial v}{\partial p_i}\right)^2\right] &= v^4 \sum_{i=1}^{3} (p_i)^2 + 2v \sum_{i=1}^{3} p_i \frac{\partial v}{\partial p_i} \\
&\quad + \frac{1}{v^2} \sum_{i=1}^{3} \left(\frac{\partial v}{\partial p_i}\right)^2 \\
&= v^4 \sum_{i=1}^{3} (p_i)^2 + \frac{1}{v^2} \sum_{i=1}^{3} \left(\frac{\partial v}{\partial p_i}\right)^2,
\end{aligned}$$

where, since v is homogeneous of degree 0 in the p_i, the summation of $p_i\left(\partial v / \partial p_i\right)$ vanishes due to Theorem A.1.

Thus, in view of equality $p^2 v^2 = 1$, we can write expression (9.12) as

$$V = \sqrt{v^2 + \frac{1}{v^2}\left(\nabla_{\mathbf{p}} v\right)^2},$$

where $\nabla_{\mathbf{p}} v$ denotes the gradient of the phase-velocity function, v, with respect to the components of the phase-slowness vector, $\mathbf{p}$. Using the chain

[3]In seismology, this quantity is often referred to as the group velocity. Our nomenclature is consistent with Synge, J.L., (1937/1962) Geometrical optics: An introduction to Hamilton's methods: Cambridge University Press, p. 12, and with Born, M., and Wolf, E., (1999) Principles of optics (7*th* edition): Cambridge University Press, p. 792 – 795. Also, our nomenclature appears in Winterstein, D.F., (1990) Velocity anisotropy terminology for geophysicists: Geophysics, **55**, 1070 – 1088, and in Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon, p. 12.

rule, we can rewrite this expression as

$$V = \sqrt{v^2 + \left[\nabla_{\mathbf{p}}\left(\ln\frac{1}{v}\right)\right]^2}.$$

Following the properties of logarithms, we obtain

$$V = \sqrt{v^2 + \left[\nabla_{\mathbf{p}}(\ln v)\right]^2}. \tag{9.13}$$

Expression (9.13) gives the magnitude of the signal velocity along the ray $\mathbf{x}(t)$. In expression (9.13), the magnitude of the ray velocity, V, is given in terms of the magnitude of the phase velocity, v, as a function of the orientation of the wavefront, given by the wavefront-normal vector, $\mathbf{p}$.

Two-dimensional case

To illustrate expression (9.13), consider a two-dimensional continuum that is contained in the x_1x_3-plane. At a given point of the continuum, we can express the orientation of the wavefront-normal vector, $\mathbf{p} = [p_1, p_3]$, in terms of a single angle. This is the phase angle, which, in this two-dimensional continuum, is given by

$$\vartheta = \arctan\frac{p_1}{p_3}. \tag{9.14}$$

Hence, using expression (9.13), the magnitude of the ray velocity can be expressed in terms of the phase velocity and the phase angle.

Herein, using expression (9.13), we can write

$$V = \sqrt{v^2 + \left(\frac{\partial \ln v}{\partial p_1}\right)^2 + \left(\frac{\partial \ln v}{\partial p_3}\right)^2}. \tag{9.15}$$

We wish to express differential operators $\partial/\partial p_i$ in terms of the phase angle. Using the chain rule, we can write

$$\frac{\partial}{\partial p_1} = \frac{\partial\vartheta}{\partial p_1}\frac{\partial}{\partial\vartheta} = \frac{\partial\arctan\frac{p_1}{p_3}}{\partial p_1}\frac{\partial}{\partial\vartheta} = \frac{p_3}{p_1^2+p_3^2}\frac{\partial}{\partial\vartheta} = p_3v^2\frac{\partial}{\partial\vartheta}, \tag{9.16}$$

where $p_1^2 + p_3^2 = p^2 = 1/v^2$. Similarly, we obtain

$$\frac{\partial}{\partial p_3} = -p_1v^2\frac{\partial}{\partial\vartheta}. \tag{9.17}$$

Thus, expression (9.15) can be written as

$$\begin{aligned} V &= \sqrt{v^2 + \left(p_3 v^2 \frac{\partial \ln v}{\partial \vartheta}\right)^2 + \left(-p_1 v^2 \frac{\partial \ln v}{\partial \vartheta}\right)^2} \\ &= \sqrt{v^2 + \left(p_3^2 + p_1^2\right) v^4 \left(\frac{\partial \ln v}{\partial \vartheta}\right)^2} = \sqrt{v^2 + p^2 v^4 \left(\frac{\partial \ln v}{\partial \vartheta}\right)^2} \\ &= \sqrt{[v(\vartheta)]^2 + [v(\vartheta)]^2 \left(\frac{\partial \ln v(\vartheta)}{\partial \vartheta}\right)^2}. \end{aligned}$$

Following the chain rule, we obtain

$$V(\vartheta) = \sqrt{[v(\vartheta)]^2 + \left[\frac{\partial v(\vartheta)}{\partial \vartheta}\right]^2}, \tag{9.18}$$

which gives the magnitude of the ray velocity in terms of the phase velocity as a function of the phase angle.

Since, as shown in Chapter 7, phase velocity is a function of the properties of the continuum — namely, its mass density and the elasticity parameters — expression (9.18) gives the magnitude of the ray velocity in terms of these properties and as a function of the phase angle.

9.2.2 Phase and ray angles

To illustrate the relation between the orientations of vectors $\mathbf{p}$ and $\dot{\mathbf{x}}$, consider a two-dimensional continuum that is contained in the x_1x_3-plane. Therein, the phase angle is given by expression (9.14). Analogously, we can express the orientation of the vector tangent to the ray, namely, $\dot{\mathbf{x}} = [\dot{x}_1, \dot{x}_3]$, in terms of a single angle. This is the ray angle, which, in this two-dimensional continuum, is given by

$$\theta = \arctan \frac{\dot{x}_1}{\dot{x}_3}. \tag{9.19}$$

Using expression (9.11), we can write

$$\dot{x}_i = p_i v^2 + \frac{1}{v}\frac{\partial v}{\partial p_i}, \qquad i \in \{1, 3\}. \tag{9.20}$$

Hence, expression (9.19) becomes

$$\tan\theta = \frac{p_1 v^2 + \dfrac{1}{v}\dfrac{\partial v}{\partial p_1}}{p_3 v^2 + \dfrac{1}{v}\dfrac{\partial v}{\partial p_3}}.$$

We wish to express the differential operators $\partial/\partial p_i$ in terms of the phase angle. Recalling expression (9.16) and (9.17), we obtain

$$\tan\theta = \frac{p_1 v^2 + p_3 v \dfrac{\partial v}{\partial \vartheta}}{p_3 v^2 - p_1 v \dfrac{\partial v}{\partial \vartheta}} = \frac{p_1 + p_3 \dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}}{p_3 - p_1 \dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}}.$$

Recalling expression (9.14), we divide both the numerator and the denominator by p_3, to obtain

$$\tan\theta = \frac{\dfrac{p_1}{p_3} + \dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}}{1 - \dfrac{p_1}{p_3}\dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}} = \frac{\tan\vartheta + \dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}}{1 - \dfrac{\tan\vartheta}{v}\dfrac{\partial v}{\partial \vartheta}}. \tag{9.21}$$

Expression (9.21) relates the phase and the ray angles.

Note that, in view of standard formulations in polar coordinates, expression (9.21) gives the angle θ that corresponds to the vector normal to the curve $1/v(\vartheta)$. We refer to this curve as the phase-slowness curve.[4] Furthermore, as shown in Exercise 9.3, ϑ corresponds to the vector normal to the curve $V(\theta)$. This curve is denoted as the ray-velocity curve. Hence, the phase-slowness curve is the polar reciprocal of the ray-velocity curve. In general, the phase-slowness and ray-velocity surfaces are polar reciprocals of one another.[5]

The possibility of solving expression (9.21) explicitly for ϑ depends on function v. To understand this statement, consider the following description. The explicit solution of expression (9.21) for ϑ requires that we can solve expression (9.20) for the p_i in terms of the $\dot{x}_i$. Since expression (9.20) is derived from expression (B.12), we require the solvability of the latter expression for the p_i in terms of the $\dot{x}_i$. It can be shown that expression (9.21) can be explicitly solved for ϑ if and only if function v is quadratic in the p_i. In a seismological context, this implies an elliptical velocity dependence. Consequently, an explicit ray-velocity expression $V = V(\theta)$, where θ is the

[4] Interested readers might refer to Anton, H., (1984) Calculus with analytic geometry: John Wiley & Sons, pp. 730 – 731.

[5] Interested readers might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2nd edition): Springer-Verlag, pp. 248 – 252, where the relation between the direction normal to the wavefront and the ray direction is formulated in terms of Huygens' principle, as well as to Born, M., and Wolf, E., (1999) Principles of optics (7th edition): Cambridge University Press, pp. 803 – 805, and to Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon, pp. 21 – 29, where the geometrical properties of the physical concepts are formulated.

ray angle, is possible only for elliptical velocity dependence. This expression is illustrated in Exercise 9.8.

9.2.3 Geometrical illustration

In general, at a given point, the direction of a wavefront normal and the direction of a ray are different. Also, considering two wavefronts separated by a given time interval, the magnitudes of phase and ray velocities differ due to the fact that the distance along the wavefront normal is different than the distance along the ray over the same time interval.

As shown in Exercise 9.7, the relationship between the magnitudes of the ray velocity, $V = |\dot{\mathbf{x}}|$, and phase velocity, $v = 1/|\mathbf{p}|$, is given by

$$V = \frac{v}{\mathbf{n} \cdot \mathbf{t}}, \tag{9.22}$$

where $\mathbf{n}$ and $\mathbf{t}$ are unit vectors normal to the wavefront and tangent to the ray, respectively.

Note that, in view of vector algebra, expression (9.22) shows that the phase-velocity vector is the projection of the ray-velocity vector onto the wavefront normal.[6] This means that, in general, the magnitude of ray velocity is always greater than, or equal to, the magnitude of the corresponding phase velocity.

Using the definition of the scalar product and the fact that $|\mathbf{n}| = |\mathbf{t}| = 1$, we can rewrite expression (9.22) as

$$V = \frac{v}{\cos(\theta - \vartheta)}. \tag{9.23}$$

Expression (9.23) conveniently involves all four entities discussed in this chapter, namely, ray velocity, V, phase velocity, v, ray angle, θ, and phase angle, ϑ.

Closing remarks

To describe rays in anisotropic inhomogeneous continua, we can use either Hamilton's ray equations or Lagrange's ray equations. These equations con-

[6] Readers interested in this formulation might refer to Auld, B.A., (1973) Acoustic fields and waves in solids: John Wiley and Sons, Vol. I, p. 222 and p. 227, to Born, M., and Wolf, E., (1999) Principles of optics (7th edition): Cambridge University Press, p. 794, and to Epstein, M., and Śniatycki, J. (1992) Fermat's principle in elastodynamics: Journal of Elasticity, **27**, 45 – 56.

stitute dual descriptions of the same theory. Hamilton's ray equations are directly rooted in fundamental physical principles, while Lagrange's ray equations are based on the same principles via Legendre's transformation, which links the two systems. Thus, Lagrange's ray equations are subject to the singularities of this transformation.

Lagrange's formulation belongs to the realm of variational methods and, hence, allows us to introduce the tools of the calculus of variations, which are the subject of *Part III.*

Exercises

Exercise 9.1 [7] *Considering a two-dimensional continuum, verify that, given Lagrangian $\mathcal{L}$ that satisfies Lagrange's ray equations (9.8), Beltrami's identity (9.9) is also satisfied.*

Solution 9.1 *For a two-dimensional continuum, let $x := x_1$ and $z := x_2$. Then, we can write $\mathcal{L} = \mathcal{L}(x, z, \dot{x}, \dot{z}, t)$. Consequently, Beltrami's identity (9.9) can be written as*

$$\frac{\partial \mathcal{L}}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}\frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{z}\frac{\partial \mathcal{L}}{\partial \dot{z}} - \mathcal{L}\right) = 0. \tag{9.24}$$

Differentiating the left-hand side of equation (9.24), we obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}\frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{z}\frac{\partial \mathcal{L}}{\partial \dot{z}} - \mathcal{L}\right) &= \frac{\partial \mathcal{L}}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{z}\frac{\partial \mathcal{L}}{\partial \dot{z}}\right) - \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} \\
&= \frac{\partial \mathcal{L}}{\partial t} + \ddot{x}\frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{x}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) + \ddot{z}\frac{\partial \mathcal{L}}{\partial \dot{z}} + \dot{z}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{z}}\right) \\
&\quad - \left[\frac{\partial \mathcal{L}}{\partial x}\dot{x} + \frac{\partial \mathcal{L}}{\partial z}\dot{z} + \frac{\partial \mathcal{L}}{\partial \dot{x}}\ddot{x} + \frac{\partial \mathcal{L}}{\partial \dot{z}}\ddot{z} + \frac{\partial \mathcal{L}}{\partial t}\right] \\
&= \dot{x}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) - \frac{\partial \mathcal{L}}{\partial x}\right] + \dot{z}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{z}}\right) - \frac{\partial \mathcal{L}}{\partial z}\right] = 0,
\end{aligned}$$

which agrees with the right-hand side, as required. Note that the vanishing of the left-hand side results from the fact that each expression in brackets

[7] See also Section 14.5

corresponds to a ray equation from system (9.8), namely,

$$\left\{\begin{aligned} &\frac{\partial \mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) = 0 \\ &\frac{\partial \mathcal{L}}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{z}}\right) = 0 \end{aligned}\right. .$$

Exercise 9.2 *Assuming that Hamiltonian $\mathcal{H}$ and Lagrangian $\mathcal{L}$ do not explicitly depend on t, following expression (9.2) and using equations (9.8), derive Hamilton's ray equations (8.19).*

Solution 9.2 *Consider $\mathcal{H}(x,p)$ and $\mathcal{L}(x,\dot{x})$. Following expression (9.2), the differential of $\mathcal{H}$ becomes*

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{3} \mathrm{d}p_i \dot{x}_i + \sum_{i=1}^{3} p_i \mathrm{d}\dot{x}_i - \sum_{i=1}^{3} \frac{\partial \mathcal{L}}{\partial x_i}\mathrm{d}x_i - \sum_{i=1}^{3} \frac{\partial \mathcal{L}}{\partial \dot{x}_i}\mathrm{d}\dot{x}_i .$$

In view of expression (B.7), we can write $p_i = \partial \mathcal{L}/\partial \dot{x}_i$. Hence, the second and last summations on the right-hand side cancel one another, and we obtain

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{3} \dot{x}_i \mathrm{d}p_i - \sum_{i=1}^{3} \frac{\partial \mathcal{L}}{\partial x_i}\mathrm{d}x_i . \tag{9.25}$$

Also, the differential of $\mathcal{H}$, can be formally written as

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{3} \frac{\partial \mathcal{H}}{\partial p_i}\mathrm{d}p_i + \sum_{i=1}^{3} \frac{\partial \mathcal{H}}{\partial x_i}\mathrm{d}x_i . \tag{9.26}$$

Equating the corresponding terms of expression (9.25) and its formal statement (9.26), we can write

$$\left\{\begin{aligned} &\dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ &\frac{\partial \mathcal{L}}{\partial x_i} = -\frac{\partial \mathcal{H}}{\partial x_i} \end{aligned}\right. , \qquad i \in \{1,2,3\} .$$

Invoking Lagrange's ray equation (9.8) and recalling expression (B.7), we can write

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{\mathrm{d}p_i}{\mathrm{d}t} \equiv \frac{\partial \mathcal{L}}{\partial x_i} - \dot{p}_i = 0, \qquad i \in \{1,2,3\} .$$

Hence, we obtain

$$\dot{p}_i = \frac{\partial \mathcal{L}}{\partial x_i}, \qquad i \in \{1, 2, 3\}.$$

Thus, we can write

$$\begin{cases} \dot{x}_i = \dfrac{\partial \mathcal{H}}{\partial x_i} \\ \\ \dot{p}_i = -\dfrac{\partial \mathcal{H}}{\partial x_i} \end{cases}, \qquad i \in \{1, 2, 3\},$$

which are Hamilton's ray equations (8.19), as required.

Exercise 9.3 [8] *In view of the polar reciprocity of the phase-slowness curve and the ray-velocity curve, derive the equation that, while analogous to expression (9.21), relates phase angle to both ray velocities and ray angles, namely,*

$$\tan\vartheta = \frac{\tan\theta - \dfrac{1}{V(\theta)}\dfrac{\partial V}{\partial\theta}}{1 + \dfrac{\tan\theta}{V(\theta)}\dfrac{\partial V}{\partial\theta}}.$$

Solution 9.3 *Phase angle is given by*

$$\tan\vartheta = \frac{p_1}{p_3},$$

where, following Legendre's transformation, p_i is the phase-slowness component given by

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i},$$

and $\mathcal{L}$ is the ray-theory Lagrangian. Considering a two-dimensional medium and following the definition of the Lagrangian, we can write

$$\mathcal{L} = \frac{\dot{x}_1^2 + \dot{x}_3^2}{[V(\theta)]^2},$$

where $\theta = \arctan(\dot{x}_1/\dot{x}_3)$. Consider differential operator $\partial/\partial\dot{x}_i$. To express the differential operator in terms of the ray angle, we can write

$$\frac{\partial}{\partial\dot{x}_1} = \frac{\partial\theta}{\partial\dot{x}_1}\frac{\partial}{\partial\theta} = \frac{\partial \arctan\dfrac{\dot{x}_1}{\dot{x}_3}}{\partial\dot{x}_1}\frac{\partial}{\partial\theta} = \frac{\dfrac{1}{\dot{x}_3}}{1 + \left(\dfrac{\dot{x}_1}{\dot{x}_3}\right)^2}\frac{\partial}{\partial\theta} = \frac{\dot{x}_3}{V^2}\frac{\partial}{\partial\theta},$$

[8] See also Section 11.1.2

and

$$\frac{\partial}{\partial \dot{x}_3} = \frac{\partial \theta}{\partial \dot{x}_3}\frac{\partial}{\partial \theta} = \frac{\partial \arctan \dfrac{\dot{x}_1}{\dot{x}_3}}{\partial \dot{x}_3}\frac{\partial}{\partial \theta} = -\frac{\dfrac{\dot{x}_1}{\dot{x}_3^2}}{1+\left(\dfrac{\dot{x}_1}{\dot{x}_3}\right)^2}\frac{\partial}{\partial \theta} = -\frac{\dot{x}_1}{V^2}\frac{\partial}{\partial \theta}.$$

Consider the expression for the phase-slowness components and for the ray-theory Lagrangian. Using the quotient rule, we can write

$$p_1 = \frac{\partial \mathcal{L}}{\partial \dot{x}_1} = \frac{2\dot{x}_1 V^2 - 2\left(\dot{x}_1^2 + \dot{x}_3^2\right) V \dfrac{\partial V}{\partial \dot{x}_1}}{V^4} = \frac{2\dot{x}_1 - 2V\dfrac{\partial V}{\partial \dot{x}_1}}{V^2},$$

where we used the fact that the expression in parentheses is equal to the square of the magnitude of the ray velocity, namely, V^2. *Using, for* $\partial/\partial \dot{x}_1$, *the differential operator derived above, we obtain*

$$p_1 = 2\frac{\dot{x}_1 - \dfrac{\dot{x}_3}{V}\dfrac{\partial V}{\partial \theta}}{V^2}.$$

Similarly, we get

$$p_3 = 2\frac{\dot{x}_3 + \dfrac{\dot{x}_1}{V}\dfrac{\partial V}{\partial \theta}}{V^2}.$$

Thus,

$$\tan\vartheta = \frac{p_1}{p_3} = \frac{\dot{x}_1 - \dfrac{\dot{x}_3}{V}\dfrac{\partial V}{\partial \theta}}{\dot{x}_3 + \dfrac{\dot{x}_1}{V}\dfrac{\partial V}{\partial \theta}} = \frac{\dfrac{\dot{x}_1}{\dot{x}_3} - \dfrac{1}{V}\dfrac{\partial V}{\partial \theta}}{1 + \dfrac{\dfrac{\dot{x}_1}{\dot{x}_3}}{V}\dfrac{\partial V}{\partial \theta}} = \frac{\tan\theta - \dfrac{1}{V}\dfrac{\partial V}{\partial \theta}}{1 + \dfrac{\tan\theta}{V}\dfrac{\partial V}{\partial \theta}}, \tag{9.27}$$

as required, which shows that $1/v(\vartheta)$ *and* $V(\theta)$ *are polar reciprocals of one another.*

Remark 9.1 *Expression (9.27) requires a closed form expression for the ray velocity as a function of the ray angle,* $V(\theta)$. *Such an expression can be formulated only for elliptical velocity dependence. In such a case the ray-velocity curve is an ellipse.*

Exercise 9.4 *Using expressions (9.21) and (9.27) and following standard trigonometric identities, show that*

$$\frac{\partial}{\partial \vartheta}\ln v = \frac{\partial}{\partial \theta}\ln V.$$

Solution 9.4 *Note that expression (9.21) can be written as*

$$\tan\theta = \frac{\tan\vartheta + \frac{\partial}{\partial\vartheta}\ln v}{1-\tan\vartheta\frac{\partial}{\partial\vartheta}\ln v} = \frac{\tan\vartheta + \tan\left[\arctan\left(\frac{\partial}{\partial\vartheta}\ln v\right)\right]}{1-\tan\vartheta\tan\left[\arctan\left(\frac{\partial}{\partial\vartheta}\ln v\right)\right]},$$

which, following trigonometric identities, we can write as

$$\theta = \vartheta + \arctan\left(\frac{\partial}{\partial\vartheta}\ln v\right). \tag{9.28}$$

Similarly, expression (9.27) can be written as

$$\vartheta = \theta - \arctan\left(\frac{\partial}{\partial\theta}\ln V\right). \tag{9.29}$$

Solving expression (9.29) for θ and equating it to expression (9.28), we obtain

$$\frac{\partial}{\partial\vartheta}\ln v = \frac{\partial}{\partial\theta}\ln V,$$

as required.

Exercise 9.5 *Derive expression (9.21) using level-set function (8.4) and characteristic equations (8.13).*

Solution 9.5 *In view of expression (9.19), the ray angle can be stated as*

$$\tan\theta = \frac{\frac{\mathrm{d}x_1}{\mathrm{d}s}}{\frac{\mathrm{d}x_3}{\mathrm{d}s}}, \tag{9.30}$$

where s defines the parametrization of the ray $\mathbf{x}(s)$, and dx_i/ds are the components of the vector tangent to the ray. Since expression (9.30) is given as a ratio, the actual parametrization has no effect on the ray angle. Consider a given point in an anisotropic continuum and the level-set function given by expression (8.4), namely,

$$F(\mathbf{p}) = p^2v^2(\mathbf{p}) = 1. \tag{9.31}$$

At a given point, expression (9.31) is not a function of $\mathbf{x}$, and, hence, $\partial F/\partial x_i = 0$. Thus, characteristic equations (8.13), are reduced to

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = \zeta\frac{\partial F}{\partial p_i} = 2\zeta\left(p_iv^2 + p^2v\frac{\partial v}{\partial p_i}\right), \qquad i\in\{1,2\}.$$

Following expression (9.30), we can write the ray angle as

$$\tan\theta = \frac{2\zeta\left(p_1 v^2 + p^2 v \dfrac{\partial v}{\partial p_1}\right)}{2\zeta\left(p_3 v^2 + p^2 v \dfrac{\partial v}{\partial p_3}\right)} = \frac{p_1 v^2 + p^2 v \dfrac{\partial v}{\partial p_1}}{p_3 v^2 + p^2 v \dfrac{\partial v}{\partial p_3}}. \tag{9.32}$$

We wish to express the quantities on the right-hand side of expression (9.32) in terms of the phase angle, ϑ. Recalling expression (9.14), we can write the differential operator in the numerator as

$$\begin{aligned}\frac{\partial}{\partial p_1} = \frac{\partial\vartheta}{\partial p_1}\frac{\partial}{\partial\vartheta} &= \frac{\partial \arctan\dfrac{p_1}{p_3}}{\partial p_1}\frac{\partial}{\partial\vartheta} \\ &= \frac{\dfrac{1}{p_3}}{1+\left(\dfrac{p_1}{p_3}\right)^2}\frac{\partial}{\partial\vartheta} = \frac{p_3}{p^2}\frac{\partial}{\partial\vartheta}.\end{aligned}$$

Similarly, we obtain the differential operator in the denominator, which is

$$\frac{\partial}{\partial p_3} = -\frac{p_1}{p^2}\frac{\partial}{\partial\vartheta}.$$

Using these differential operators in expression (9.32), we can rewrite it as

$$\tan\theta = \frac{p_1 v^2 + p_3 v \dfrac{\partial v}{\partial\vartheta}}{p_3 v^2 - p_1 v \dfrac{\partial v}{\partial\vartheta}} = \frac{p_1 + p_3 \dfrac{1}{v}\dfrac{\partial v}{\partial\vartheta}}{p_3 - p_1 \dfrac{1}{v}\dfrac{\partial v}{\partial\vartheta}}.$$

Again, recalling expression (9.14), we divide both the numerator and the denominator by p_3 to obtain

$$\tan\theta = \frac{\dfrac{p_1}{p_3} + \dfrac{1}{v}\dfrac{\partial v}{\partial\vartheta}}{1 - \dfrac{p_1}{p_3}\dfrac{1}{v}\dfrac{\partial v}{\partial\vartheta}} = \frac{\tan\vartheta + \dfrac{1}{v}\dfrac{\partial v}{\partial\vartheta}}{1 - \dfrac{\tan\vartheta}{v}\dfrac{\partial v}{\partial\vartheta}},$$

which is expression (9.21), as required.

Exercise 9.6 *Derive expression (9.21) using level-set function (8.3) and characteristic equations (8.13).*

Solution 9.6 *Recall expression (9.30). Consider a given point in an anisotropic continuum and the level-set function that is given by expression (8.3), namely,*

$$F(\mathbf{p}) = p^2 - \frac{1}{v^2(\mathbf{p})} = 0. \tag{9.33}$$

At a given point, expression (9.33) is not a function of $\mathbf{x}$, *and, hence,* $\partial F/\partial x_i = 0$. *Thus, characteristic equations (8.13), are reduced to*

$$\frac{\mathrm{d}x_i}{\mathrm{d}s} = \zeta \frac{\partial F}{\partial p_i} = 2\zeta \left(p_i + \frac{1}{v^3} \frac{\partial v}{\partial p_i} \right), \qquad i \in \{1, 2\}.$$

Following expression (9.30), we can write the ray angle as

$$\tan\theta = \frac{p_1 + \dfrac{1}{v^3}\dfrac{\partial v}{\partial p_1}}{p_3 + \dfrac{1}{v^3}\dfrac{\partial v}{\partial p_3}}. \tag{9.34}$$

We wish to express the quantities on the right-hand side of expression (9.34) in terms of the phase angle, ϑ. *In view of expression (9.14), we consider the differential operator in the numerator, namely,*

$$\frac{\partial}{\partial p_1} = \frac{\partial \vartheta}{\partial p_1}\frac{\partial}{\partial \vartheta} = \frac{\partial \arctan \dfrac{p_1}{p_3}}{\partial p_1}\frac{\partial}{\partial \vartheta}$$

$$= \frac{\dfrac{1}{p_3}}{1 + \left(\dfrac{p_1}{p_3}\right)^2}\frac{\partial}{\partial \vartheta} = \frac{p_3}{p_1^2 + p_3^2}\frac{\partial}{\partial \vartheta}.$$

Considering the phase-slowness vector given by $\mathbf{p} = [p_1, p_3]$, *we can write* $p^2 = \mathbf{p} \cdot \mathbf{p}$. *Hence, the differential operator becomes*

$$\frac{\partial}{\partial p_1} = \frac{p_3}{p^2}\frac{\partial}{\partial \vartheta}.$$

Similarly, we obtain the differential operator in the denominator, which is

$$\frac{\partial}{\partial p_3} = -\frac{p_1}{p^2}\frac{\partial}{\partial \vartheta}.$$

Using these differential operators in expression (9.34), we can rewrite it as

$$\tan\theta = \frac{p_1 + \dfrac{1}{v^3}\dfrac{p_3}{p^2}\dfrac{\partial v}{\partial \vartheta}}{p_3 - \dfrac{1}{v^3}\dfrac{p_1}{p^2}\dfrac{\partial v}{\partial \vartheta}}.$$

Following eikonal equation (7.17), we can state $p^2v^2 = 1$, and, hence, we can write

$$\tan\theta = \frac{p_1 + \dfrac{p_3}{v}\dfrac{\partial v}{\partial \vartheta}}{p_3 - \dfrac{p_1}{v}\dfrac{\partial v}{\partial \vartheta}}.$$

Again, recalling expression (9.14) and dividing both numerator and denominator by p_3, we obtain

$$\tan\theta = \frac{\dfrac{p_1}{p_3} + \dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}}{1 - \dfrac{\dfrac{p_1}{p_3}}{v}\dfrac{\partial v}{\partial \vartheta}} = \frac{\tan\vartheta + \dfrac{1}{v(\vartheta)}\dfrac{\partial v(\vartheta)}{\partial \vartheta}}{1 - \dfrac{\tan\vartheta}{v(\vartheta)}\dfrac{\partial v(\vartheta)}{\partial \vartheta}},$$

which is expression (9.21), as required.

Exercise 9.7 *Using expressions (9.18) and (9.21), derive expression (9.22).*

Solution 9.7 *Using algebraic manipulation, we can write expression (9.21), namely,*

$$\tan\theta = \frac{\tan\vartheta + \dfrac{1}{v}\dfrac{\partial v}{\partial \vartheta}}{1 - \dfrac{\tan\vartheta}{v}\dfrac{\partial v}{\partial \vartheta}},$$

as

$$\frac{\partial v}{\partial \vartheta} = v\frac{\tan\theta - \tan\vartheta}{1 + \tan\theta\tan\vartheta}.$$

Recognizing the trigonometric identity, we can rewrite it as

$$\frac{\partial v}{\partial \vartheta} = v\tan(\theta - \vartheta). \tag{9.35}$$

Consider expression (9.18). In view of expression (9.35), we can write

$$\begin{aligned} V &= \sqrt{v^2 + \left(\frac{\partial v}{\partial \vartheta}\right)^2} \\ &= \sqrt{v^2 + v^2\tan^2(\theta - \vartheta)} \\ &= v\sqrt{1 + \tan^2(\theta - \vartheta)}. \end{aligned}$$

Using trigonometric identities, we obtain

$$V = \frac{v}{\cos(\theta - \vartheta)}, \tag{9.36}$$

which, notably, is expression (9.23). The argument of the cosine function is the angle between the ray-velocity vector, $\mathbf{V}$*, and the phase-velocity vector,* $\mathbf{v}$*. As defined in expression (9.22), let* $\mathbf{t}$ *be the unit vector tangent to the ray, and* $\mathbf{n}$ *be the unit vector normal to the wavefront. Hence,* $\theta - \vartheta$ *is the angle between* $\mathbf{n}$ *and* $\mathbf{t}$*. Thus, we can immediately rewrite expression (9.36) as*

$$V = \frac{v}{\mathbf{n} \cdot \mathbf{t}},$$

which is expression (9.22), as required.

Exercise 9.8 *Derive a particular case of expression (9.21) that corresponds to the elliptical velocity dependence.*

Solution 9.8 *Inserting expression (6.60) into expression (9.18), we can write the magnitude of the ray-velocity vector as*

$$V(\vartheta) = \sqrt{\frac{v_x^4 \tan^2 \vartheta + v_z^4}{v_x^2 \tan^2 \vartheta + v_z^2}}. \tag{9.37}$$

This is the magnitude of ray velocity in terms of the phase velocity as a function of the phase angle for the case of elliptical velocity dependence. Also, inserting expression (6.60) into expression (9.21), we obtain

$$\tan \theta = \left(\frac{v_x}{v_z}\right)^2 \tan \vartheta, \tag{9.38}$$

which is the relation between the phase angle and the ray angle for elliptical velocity dependence. Expression (9.38) is analogous to expression (10.47), which corresponds to SH waves in transversely isotropic continua. Inserting expression (9.38) into expression (9.37), we can write the magnitude of the ray-velocity vector in terms of ray-related quantities, namely,

$$V(\theta) = V_z \sqrt{\frac{\tan^2 \theta + 1}{\left(\frac{V_z}{V_x}\right)^2 \tan^2 \theta + 1}}, \tag{9.39}$$

where V_x *and* V_z *are the magnitudes of the ray-velocity vector along the x-axis and z-axis, respectively. Herein, we use the fact that, along the axes of the ellipse, the magnitudes of the phase velocity and the ray velocity coincide.*

This Page Intentionally Left Blank

Chapter 10

Christoffel's equations

Mathematical applications to physics occur in at least two aspects. Mathematics is of course the principal tool for solving technical analytical problems, but increasingly it is also a principal guide in our understanding of the basic structure and concepts involved.

Theodore Frankel (1997) The geometry of physics

Preliminary remarks

In Chapter 7, where we studied the equations of motion in anisotropic continua, we noted that waves propagate therein with three distinct phase velocities. Throughout Chapters 7 – 9, we denoted each of these velocities by $v = v(\mathbf{x}, \mathbf{p})$, which is a function of both position and direction. Such a formulation allowed us to derive general forms of the equations governing ray theory in anisotropic inhomogeneous continua, namely, the eikonal equation, Hamilton's ray equations and Lagrange's ray equations. In this chapter, we wish to derive explicit expressions for these three velocities in terms of the properties of a given continuum, namely, its mass density and elasticity parameters.

We begin this chapter by writing Christoffel's equations, derived in Chapter 7, explicitly in terms of mass density and elasticity parameters. Based on the solvability of these equations, we are then able to formulate the expressions for the three wave velocities, as well as for the associated displacement directions. Using these expressions, we study two specific cases — the three waves that propagate along the symmetry axis in a monoclinic continuum and the three waves that propagate in an arbitrary direction in a transversely isotropic continuum. The chapter concludes with a discussion of the three corresponding phase-slowness surfaces and their intersections.

10.1 Explicit form of Christoffel's equations

We wish to study Christoffel's equations, shown in expression (7.13), namely,

$$\sum_{k=1}^{3}\left(\sum_{j=1}^{3}\sum_{l=1}^{3}c_{ijkl}\left(\mathbf{x}\right)p_{j}p_{l}-\rho\left(\mathbf{x}\right)\delta_{ik}\right)A_{k}\left(\mathbf{x}\right)=\mathbf{0},\qquad i\in\left\{1,2,3\right\},\tag{10.1}$$

in the context of a specific continuum. In other words, we wish to rewrite equations (10.1) in a way that allows us to conveniently insert the elasticity parameters of a continuum exhibiting a particular symmetry, as discussed in Chapter 5.

Expressing the phase slowness as the reciprocal of the phase velocity, namely,

$$p^{2}=\frac{1}{v^{2}},\tag{10.2}$$

and letting $n_i^2 = p_i^2/p^2$, where $p^2 := \mathbf{p}\cdot\mathbf{p}$, be the squared components of the unit vector normal to the wavefront, we can rewrite equations (10.1) as

$$p^{2}\sum_{k=1}^{3}\left(\sum_{j=1}^{3}\sum_{l=1}^{3}c_{ijkl}\left(\mathbf{x}\right)n_{j}n_{l}-\rho\left(\mathbf{x}\right)v^{2}\delta_{ik}\right)A_{k}\left(\mathbf{x}\right)=\mathbf{0},\qquad i\in\left\{1,2,3\right\}.\tag{10.3}$$

We can state equations (10.3) in matrix notation as

$$p^{2}\left[\boldsymbol{\Gamma}\left(\mathbf{x},\mathbf{n}\right)-\rho\left(\mathbf{x}\right)v^{2}\mathbf{I}\right]\mathbf{A}\left(\mathbf{x}\right)=\mathbf{0},\tag{10.4}$$

where

$$\boldsymbol{\Gamma}\left(\mathbf{x},\mathbf{n}\right)=\left[\begin{array}{ccc}\sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{1j1l}\left(\mathbf{x}\right)n_{j}n_{l} & \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{1j2l}\left(\mathbf{x}\right)n_{j}n_{l} & \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{1j3l}\left(\mathbf{x}\right)n_{j}n_{l}\\ \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{2j1l}\left(\mathbf{x}\right)n_{j}n_{l} & \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{2j2l}\left(\mathbf{x}\right)n_{j}n_{l} & \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{2j3l}\left(\mathbf{x}\right)n_{j}n_{l}\\ \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{3j1l}\left(\mathbf{x}\right)n_{j}n_{l} & \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{3j2l}\left(\mathbf{x}\right)n_{j}n_{l} & \sum\limits_{j=1}^{3}\sum\limits_{l=1}^{3}c_{3j3l}\left(\mathbf{x}\right)n_{j}n_{l}\end{array}\right].\tag{10.5}$$

[1]Using formula (3.5), we can state the entries of matrix $\mathbf{\Gamma}(\mathbf{x}, \mathbf{n})$ in terms of the elasticity parameters $C_{mn}(\mathbf{x})$, to obtain

$$
\begin{aligned}
\Gamma_{11} &= C_{11}n_1^2 + C_{66}n_2^2 + C_{55}n_3^2 + 2\left(C_{16}n_1n_2 + C_{56}n_2n_3 + C_{15}n_1n_3\right), \\
\Gamma_{22} &= C_{66}n_1^2 + C_{22}n_2^2 + C_{44}n_3^2 + 2\left(C_{26}n_1n_2 + C_{24}n_2n_3 + C_{46}n_1n_3\right), \\
\Gamma_{33} &= C_{55}n_1^2 + C_{44}n_2^2 + C_{33}n_3^2 + 2\left(C_{45}n_1n_2 + C_{34}n_2n_3 + C_{35}n_1n_3\right), \\
\Gamma_{12} &= \Gamma_{21} \\
&= C_{16}n_1^2 + C_{26}n_2^2 + C_{45}n_3^2 \\
&+ \left(C_{12} + C_{66}\right)n_1n_2 + \left(C_{25} + C_{46}\right)n_2n_3 + \left(C_{14} + C_{56}\right)n_1n_3, \\
\Gamma_{13} &= \Gamma_{31} \qquad\qquad (10.6) \\
&= C_{15}n_1^2 + C_{46}n_2^2 + C_{35}n_3^2 \\
&+ \left(C_{14} + C_{56}\right)n_1n_2 + \left(C_{36} + C_{45}\right)n_2n_3 + \left(C_{13} + C_{55}\right)n_1n_3, \\
\Gamma_{23} &= \Gamma_{32} \\
&= C_{56}n_1^2 + C_{24}n_2^2 + C_{34}n_3^2 \\
&+ \left(C_{25} + C_{46}\right)n_1n_2 + \left(C_{23} + C_{44}\right)n_2n_3 + \left(C_{36} + C_{45}\right)n_1n_3,
\end{aligned}
$$

where, for convenience of notation, we do not explicitly write $\Gamma_{rs}(\mathbf{x}, \mathbf{n})$ and $C_{mn}(\mathbf{x})$. Thus, using the elasticity matrices formulated in Chapter 5, expressions (10.6) allow us to state Christoffel's equations for a given

[1]Note that it is also common to divide the elasticity parameters by mass density and, hence, to write Christoffel's equations (10.1) as

$$\sum_{k=1}^{3}\left(\sum_{j=1}^{3}\sum_{l=1}^{3}\frac{c_{ijkl}(\mathbf{x})}{\rho(\mathbf{x})}p_jp_l - \delta_{ik}\right)A_k(\mathbf{x}) = \mathbf{0}, \qquad i \in \{1,2,3\},$$

where, as we see in view of Exercise 2.4, the c_{ijkl}/ρ have units of velocity squared. The corresponding solvability condition can be written as

$$\det\left[\Gamma_{ik}(\mathbf{x}, \mathbf{p}) - \delta_{ik}\right] = 0, \qquad i, k \in \{1,2,3\},$$

where the entries of matrix $\mathbf{\Gamma}(\mathbf{x}, \mathbf{p})$ are

$$\Gamma_{ik}(\mathbf{x}, \mathbf{p}) := \sum_{j=1}^{3}\sum_{l=1}^{3}\frac{c_{ijkl}(\mathbf{x})}{\rho(\mathbf{x})}p_jp_l, \qquad i, k \in \{1,2,3\}.$$

Each of the three eigenvalues of $\mathbf{\Gamma}(\mathbf{x}, \mathbf{p})$, namely, $G_i(\mathbf{x}, \mathbf{p})$, where $i \in \{1,2,3\}$, results in an eikonal equation, which we can write as

$$G_i(\mathbf{x}, \mathbf{p}) = 1, \qquad i \in \{1,2,3\},$$

and which is equivalent to equation (7.17).

continuum. Hence, we can conveniently study behaviour of the continuum in terms of its properties, namely, its mass density and elasticity parameters.

In general, $\mathbf{\Gamma}$ is a symmetric matrix due to the symmetry of the elasticity tensor, c_{ijkl}, discussed in Section 4.2. Consequently, in view of equations (10.4), which constitute a homogeneous system of three linear equations, we can invoke the following theorems of linear algebra.[2]

Theorem 10.1 *Since $\mathbf{\Gamma}$ is symmetric, the corresponding eigenvalues are real.*

Theorem 10.2 *Since $\mathbf{\Gamma}$ is symmetric, the corresponding eigenvectors are orthogonal to each other.*

Furthermore, a homogeneous system of linear equations has either only the trivial solution, namely, $\mathbf{A} = \mathbf{0}$, or infinitely many solutions in addition to the trivial solution. A necessary and sufficient condition for a system of n homogeneous equations in n unknowns to have nontrivial solutions is the vanishing of the determinant of the coefficient matrix.

We wish to examine the solvability of system (10.4). Since, for physically meaningful solutions, we require $p^2 \neq 0$, as discussed in Section 7.3, system (10.4) can be written as

$$\left[\mathbf{\Gamma}(\mathbf{x},\mathbf{n}) - \rho(\mathbf{x})\, v^2 \mathbf{I}\right] \mathbf{A}(\mathbf{x}) = \mathbf{0}. \tag{10.7}$$

Hence, we can write the solvability condition of system (10.7) as

$$\det \begin{bmatrix} \Gamma_{11}(\mathbf{x},\mathbf{n}) - \rho(\mathbf{x})\, v^2 & \Gamma_{12}(\mathbf{x},\mathbf{n}) & \Gamma_{13}(\mathbf{x},\mathbf{n}) \\ \Gamma_{12}(\mathbf{x},\mathbf{n}) & \Gamma_{22}(\mathbf{x},\mathbf{n}) - \rho(\mathbf{x})\, v^2 & \Gamma_{23}(\mathbf{x},\mathbf{n}) \\ \Gamma_{13}(\mathbf{x},\mathbf{n}) & \Gamma_{23}(\mathbf{x},\mathbf{n}) & \Gamma_{33}(\mathbf{x},\mathbf{n}) - \rho(\mathbf{x})\, v^2 \end{bmatrix} = 0. \tag{10.8}$$

In view of Theorem 10.1, the determinantal equation, stated in expression (10.8), has three real roots — the eigenvalues ρv_i^2, where $i = 1, 2, 3$. Furthermore, in view of Theorem 10.2, the three corresponding eigenvectors are orthogonal to each other.

To recognize the physical meaning of the eigenvalues and eigenvectors of system (10.7), consider trial solution (7.5), which leads to Christoffel's equations and can be written as

$$\mathbf{u}(\mathbf{x},t) = \mathbf{A}(\mathbf{x})\, f\left\{v_0\left[\psi(\mathbf{x}) - t\right]\right\}. \tag{10.9}$$

[2] For proofs of Theorem 10.1 and Theorem 10.2, interested readers might refer to Anton, H., (1973) Elementary linear algebra: John Wiley & Sons, p. 289 and p. 399, and pp. 286 - 287, respectively.

Examining expression (10.9) and in view of definition (7.12), namely, $p_j := \partial\psi/\partial x_j$, and expression (10.2), we see that the three eigenvalues correspond to three distinct phase velocities, which are measured normal to the wavefront of a given wave. In view of Theorem 10.1, these velocities are real.

Also, as stated in trial solution (7.5), $\mathbf{A}(\mathbf{x})$ is the displacement vector. Hence, each eigenvector corresponds to the displacements of the continuum associated with the propagation of a given wave. In view of Theorem 10.2, each wave exhibits the displacement vector that is orthogonal to the displacement vectors of the other two waves.

For seismological studies, the three displacement vectors are orthogonal to each other at a given point, if all three corresponding wavefronts exhibit the same direction. If we place a receiver in an inhomogeneous continuum at a certain distance from the source — where, in general, the three wavefront normals do not coincide — the three recorded displacement directions are not orthogonal to each other since each displacement vector corresponds to a wavefront that exhibits a different orientation than the two other wavefronts.

Examining matrix (10.5), we can also conclude that, for a given wave in a continuum defined by stress-strain equations (7.2), the magnitude of the phase velocity, at a given point, depends only on the elasticity parameters and mass density at that point and is a function of the direction of propagation. Hence, given the properties of the continuum, at each point, we can uniquely determine the magnitude of phase velocity for every direction.

The corresponding displacement direction depends on the same quantities and can be also uniquely determined at a given point of an anisotropic continuum. This is not the case in isotropic continua, where the displacement direction of S waves, although contained in the plane orthogonal to the phase-slowness vector, $\mathbf{p}$, cannot be uniquely determined, as shown in Exercise 10.1.

For the remainder of this chapter, we focus our attention on a given point of the continuum. Hence, for convenience of notation, we write $\rho(\mathbf{x}) \equiv \rho$ and $C_{mn}(\mathbf{x}) \equiv C_{mn}$.

10.2 Christoffel's equations and anisotropic continua

Introductory comments

We wish to study equation (10.8), which provides us with the phase velocities of the three waves within an anisotropic continuum, as well as examine the eigenvectors of the corresponding matrix Γ, which are the displacement vectors.

Explicit expressions for these velocities in a generally anisotropic continuum can be obtained by inserting entries (10.6) into equation (10.8). Thus, we obtain three phase velocities, which are functions of both the properties of the continuum — given by its mass density, ρ, and the elasticity parameters, C_{mn} — and the orientation of the wavefront — given by its unit normal, $\mathbf{n}$. Once the phase velocities are obtained, we can find the displacement directions that correspond to each of the three waves by using system (10.7).

Note that in the formulation discussed in Chapters 1 and 2, we assumed the displacements of material points associated with the propagation of the waves to be infinitesimal. This is justified by the fact that these displacements are many orders of magnitude smaller than the size of the continuum under investigation, as well as, several orders of magnitude smaller than the wavelength of a given wave. Nevertheless, seismic receivers measure the direction and the amplitude of these displacements, thereby providing us with important information for our study of the properties of the materials through which waves propagate. These measurements are discussed in this chapter and in Chapter 11, respectively.

To illustrate explicit expressions for phase velocities and displacement directions, we consider two particular cases. In the case of a monoclinic continuum, we investigate velocities and displacements for the three waves that are associated with the propagation along the symmetry axis. Notably, this formulation also allows us to illustrate the condition of the natural coordinate system, discussed in Section 5.1. In the case of a transversely isotropic continuum, we investigate velocities and displacements for the three waves for an arbitrary direction of propagation. Notably, this formulation allows us to show that, in general, for anisotropic continua, the displacement direction is neither parallel nor orthogonal to the direction of propagation, as is the case for isotropic continua.

10.2.1 Monoclinic continua

Christoffel's equations along symmetry axis

Consider a monoclinic continuum and let the x_3-axis coincide with the normal to the symmetry plane. In other words, let the x_3-axis be the symmetry axis. Such a continuum is described by elasticity matrix (5.29).

Consider a propagation along the x_3-axis. Hence, $n_1 = n_2 = 0$, and the unit vector normal to the wavefront is $\mathbf{n} = [0, 0, 1]$. Following entries (10.6) and in view of elasticity matrix (5.29), we note that system (10.7) becomes

$$\begin{bmatrix} C_{55} - \rho v^2 & C_{45} & 0 \\ C_{45} & C_{44} - \rho v^2 & 0 \\ 0 & 0 & C_{33} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{10.10}$$

System (10.10) can be rewritten as

$$\begin{bmatrix} C_{55} - \rho v^2 & C_{45} \\ C_{45} & C_{44} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{10.11}$$

and

$$\left[C_{33} - \rho v^2\right] A_3 = 0. \tag{10.12}$$

The displacement vectors associated with equations (10.11) are contained in the x_1x_2-plane. The displacement vector associated with equation (10.12) coincides with the x_3-axis. Hence, the displacement directions associated with equations (10.11) are orthogonal to the direction of propagation, while the displacement direction associated with equation (10.12) is parallel to the direction of propagation. We refer to the waves whose displacement directions are either orthogonal or parallel to the direction of propagation as the pure-mode waves, and denote them by S or P, respectively.

Note that this monoclinic example illustrates the fact that, along the symmetry axes, all waves propagate as pure-mode waves.

Phase velocities along symmetry axis

In order to obtain the phase velocity along the symmetry axis, consider equations (10.11). The solvability condition is

$$\det \begin{bmatrix} C_{55} - \rho v^2 & C_{45} \\ C_{45} & C_{44} - \rho v^2 \end{bmatrix} = 0.$$

Thus, we obtain the determinantal equation, namely,

$$\rho^2 \left(v^2\right)^2 - \left[\left(C_{44} + C_{55}\right) \rho\right] \left(v^2\right) - \left(C_{45}^2 - C_{44}C_{55}\right) = 0,$$

and, hence, the velocities of the S waves are

$$v_{S_1} = \sqrt{\frac{\left(C_{44} + C_{55}\right) + \sqrt{\left(C_{44} - C_{55}\right)^2 + 4C_{45}^2}}{2\rho}}, \tag{10.13}$$

and

$$v_{S_2} = \sqrt{\frac{(C_{44}+C_{55}) - \sqrt{(C_{44}-C_{55})^2 + 4C_{45}^2}}{2\rho}}.$$

Also, consider equation (10.12). A nontrivial solution requires that $A_3 \neq 0$. Thus, the velocity of the P wave is

$$v_P = \sqrt{\frac{C_{33}}{\rho}}.$$

Displacement directions along symmetry axis

In view of equation (10.12), the P-wave displacement vector is parallel to the x_3-axis. Considering a three dimensional continuum, we can write this displacement vector as

$$\mathbf{A}_P = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = a \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

where a is a nonzero constant.

Now, we wish to find the orientations of the displacement vectors of the S waves. In view of equations (10.11), these vectors are contained in the x_1x_2-plane. Inserting eigenvalue (10.13) into equations (10.11), we obtain

$$\begin{bmatrix} \frac{C_{55}-C_{44}-\sqrt{(C_{44}-C_{55})^2+4C_{45}^2}}{2} & C_{45} \\ C_{45} & \frac{C_{44}-C_{55}-\sqrt{(C_{44}-C_{55})^2+4C_{45}^2}}{2} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{10.14}$$

In view of a three-dimensional continuum, we can write the nontrivial solution of system (10.14) as the displacement vector given by

$$\mathbf{A}_{S_1} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = b \begin{bmatrix} \frac{C_{55}-C_{44}+\sqrt{(C_{44}-C_{55})^2+4C_{45}^2}}{2C_{45}} \\ 1 \\ 0 \end{bmatrix}, \tag{10.15}$$

where b is a nonzero constant. Hence, the angle that this vector makes with a coordinate axis in the x_1x_2-plane is

$$\tan\Theta = \frac{A_1}{A_2} = \frac{C_{55}-C_{44}+\sqrt{(C_{44}-C_{55})^2+4C_{45}^2}}{2C_{45}}. \tag{10.16}$$

We can also find, in an analogous manner, the displacement vector that corresponds to the other S wave. It is given by

$$\mathbf{A}_{S_2} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = c \begin{bmatrix} -1 \\ \dfrac{C_{55} - C_{44} + \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2C_{45}} \\ 0 \end{bmatrix},$$

where c is a nonzero constant.

We recognize that eigenvectors $\mathbf{A}_P$, $\mathbf{A}_{S_1}$ and $\mathbf{A}_{S_2}$are linearly independent. Thus, as expected, by Theorem 10.2, the three displacement directions are orthogonal to each other, since

$$\mathbf{A}_P \cdot \mathbf{A}_{S_1} = \mathbf{A}_P \cdot \mathbf{A}_{S_2} = \mathbf{A}_{S_1} \cdot \mathbf{A}_{S_2} = 0.$$

Furthermore, in this particular case of the waves propagating along the symmetry axis, the displacement vectors are either parallel or orthogonal to the wavefront normal, $\mathbf{n}$.

In general, in anisotropic continua, the wavefront normal, $\mathbf{n}$, is neither parallel nor orthogonal to the displacement vector. However, in any anisotropic continuum, there exist at least three directions of propagation where the wavefront normal is either parallel or orthogonal to the displacement direction.[3] Such directions are called the pure-mode directions. As illustrated herein, symmetry axes are pure-mode directions.

Natural coordinate systems

In Section 5.6.2, we use the natural coordinate system to describe a monoclinic continuum using the smallest number of nonzero elasticity parameters. The relation between the natural coordinate system and pure-mode directions is stated by the following proposition.

Proposition 10.1 *Given a propagation along a pure-mode direction, the coordinate system whose axes coincide with the displacement directions of the three waves is a natural coordinate system.*

To elucidate Proposition 10.1, consider expression (10.16). Invoking the trigonometric identity given by

$$\tan(2\Theta) = \frac{2\tan\Theta}{1 - \tan^2\Theta},$$

[3]Readers interested in further description and additional references might refer to Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon, p. 166.

we can restate expression (10.16) as

$$\tan(2\Theta) = \frac{2C_{45}}{C_{44} - C_{55}}. \tag{10.17}$$

Expression (10.17) is precisely expression (5.30), which allows us to express elasticity matrix (5.29) in a natural coordinate system to obtain matrix (5.31). To further illustrate this result, we notice that, using elasticity matrix (5.31), equations (10.10) become

$$\begin{bmatrix} \hat{C}_{55} - \rho v^2 & 0 & 0 \\ 0 & \hat{C}_{44} - \rho v^2 & 0 \\ 0 & 0 & \hat{C}_{33} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{10.18}$$

where all three displacement directions are along the axes of the natural coordinate system, as expected.

Square submatrix

$$\begin{bmatrix} \hat{C}_{55} - \rho v^2 & 0 \\ 0 & \hat{C}_{44} - \rho v^2 \end{bmatrix},$$

in equation (10.18), is the diagonal form of the square matrix shown in equation (10.11). In terms of a natural coordinate system, such a diagonalization is also obtained using equation (5.32), which in the present case, we can write as

$$\begin{bmatrix} \hat{C}_{55} - \rho v^2 & 0 \\ 0 & \hat{C}_{44} - \rho v^2 \end{bmatrix} = \begin{bmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{bmatrix}$$
$$\begin{bmatrix} C_{55} - \rho v^2 & C_{45} \\ C_{45} & C_{44} - \rho v^2 \end{bmatrix} \begin{bmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{bmatrix},$$

where C_{44} and C_{55} are entries of matrix (5.29), while $\hat{C}_{44}$ and $\hat{C}_{55}$ are entries of matrix (5.31).

Also, examining systems (10.10) and (10.18), we see that the third equation remains unchanged; hence, $C_{33} = \hat{C}_{33}$. This results from the fact that to obtain a natural coordinate system, the original coordinate system is rotated by angle Θ about the x_3-axis, whose orientation remains unchanged.

10.2.2 Transversely isotropic continua

Christoffel's equations

In seismological studies, transverse isotropy plays an important role. For instance, transverse isotropy can be conveniently used to describe layered media.

Consider a transversely isotropic continuum and let the x_3-axis coincide with the normal to the plane of transverse isotropy. In other words, let the x_3-axis be the rotation-symmetry axis. Such a continuum is described by elasticity matrix (5.47). For notational convenience, letting

$$\frac{C_{11} - C_{12}}{2} = C_{66}, \tag{10.19}$$

in matrix (5.47), we can write the entries of matrix $\mathbf{\Gamma}$, given by expressions (10.6), as

$$\begin{aligned}
\Gamma_{11} &= n_1^2 C_{11} + n_2^2 C_{66} + n_3^2 C_{44}\,, \\
\Gamma_{22} &= n_1^2 C_{66} + n_2^2 C_{11} + n_3^2 C_{44}\,, \\
\Gamma_{33} &= \left(n_1^2 + n_2^2\right) C_{44} + n_3^2 C_{33}\,, \\
\Gamma_{12} &= \Gamma_{21} = n_1 n_2 \left(C_{11} - C_{66}\right), \\
\Gamma_{13} &= \Gamma_{31} = n_1 n_3 \left(C_{13} + C_{44}\right), \\
\Gamma_{23} &= \Gamma_{32} = n_2 n_3 \left(C_{13} + C_{44}\right).
\end{aligned} \tag{10.20}$$

Note that, in view of expression (10.19), $C_{12} = C_{11} - 2C_{66}$. Thus, we could also write $\Gamma_{12} = \Gamma_{21} = n_1 n_2 \left(C_{12} + C_{66}\right)$, which is consistent with the pattern of the last two lines of set (10.20). However, in this chapter, we choose to describe a transversely isotropic continuum using C_{11}, C_{13}, C_{33}, C_{44} and C_{66}.

Thus, Christoffel's equations for a transversely isotropic continuum are given by system (10.7) with entries (10.20). Note that, in view of transverse isotropy, with no loss of generality, we can set either $n_1 = 0$ or $n_2 = 0$.

Phase velocities in transverse-isotropy plane

In this section, we wish to obtain three distinct phase-velocity expressions for the pure-mode waves in a transversely isotropic continuum in order to conveniently identify the general expressions, which are derived in the following section. All waves that propagate along the rotation-symmetry axis, as well as the waves that propagate within the plane of transverse isotropy, are pure-mode waves. However, along the rotation-symmetry axis, the displacement directions of the S waves are subject to the same elastic properties, and, hence, their phase-velocity expressions are not distinct. Consequently, to obtain three distinct velocities, we consider the propagation in the plane of transverse isotropy, where $n_3 = 0$.

Furthermore, in view of transverse isotropy, we can consider the propagation in any direction in this plane. We choose the propagation along the x_1-axis and, hence, we set $n_2 = 0$. Consequently, $n_1^2 = 1$. Thus, entries (10.20) become $\Gamma_{11} = C_{11}$, $\Gamma_{22} = C_{66}$, $\Gamma_{33} = C_{44}$ and $\Gamma_{12} = \Gamma_{13} = \Gamma_{23} = 0$. Hence, for the propagation along the x_1-axis, system (10.7) becomes

$$\begin{bmatrix} C_{11} - \rho v^2 & 0 & 0 \\ 0 & C_{66} - \rho v^2 & 0 \\ 0 & 0 & C_{44} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{10.21}$$

By examining system (10.21), we recognize that all equations are independent of each other and, as expected, all three waves propagate as pure-mode waves. To consider a P wave propagating along the x_1-axis, we set the displacement amplitude along the x_1-axis to unity. Hence, the corresponding vector is $\mathbf{A}_P = [1, 0, 0]^T$. This immediately results in the expression for the P-wave velocity along the x_1-axis, namely,

$$v_P = \pm\sqrt{\frac{C_{11}}{\rho}}. \tag{10.22}$$

To consider an S wave propagating along the x_1-axis, we set to unity the displacement amplitude along the axis orthogonal to the x_1-axis and contained in the x_1x_2-plane. We view the x_1x_2-plane as a horizontal plane and, therefore, we refer to this wave as an SH wave. Hence, the corresponding vector is $\mathbf{A}_{SH} = [0, 1, 0]^T$. This immediately results in the expression for the SH-wave velocity along the x_1-axis, namely,

$$v_{SH} = \pm\sqrt{\frac{C_{66}}{\rho}}. \tag{10.23}$$

To consider the other S wave propagating along the x_1-axis, we set to unity the displacement amplitude along the axis orthogonal to the x_1x_2-plane. We refer to this wave as an SV wave. Hence, the corresponding vector is $\mathbf{A}_{SV} = [0, 0, 1]^T$. This immediately results in the expression for the SV-wave velocity along the x_1-axis, namely,

$$v_{SV} = \pm\sqrt{\frac{C_{44}}{\rho}}. \tag{10.24}$$

Expressions (10.22), (10.23) and (10.24) are distinct from each other. Hence, we can use these expressions to identify general expressions for wave velocities, which are derived below.

Phase velocities in arbitrary directions

We wish to obtain general phase-velocity expressions for the three waves propagating in arbitrary directions. Using entries (10.20), we can write expression (10.8) as

$$\begin{aligned}\det\left[\mathbf{\Gamma}-\rho v^{2}\mathbf{I}\right] = {} & \left[C_{66}\left(n_1^2+n_2^2\right)+C_{44}n_3^2-\rho v^2\right] \\ & \{-C_{13}^2\left(n_1^2+n_2^2\right)n_3^2-2C_{13}C_{44}\left(n_1^2+n_2^2\right)n_3^2 \\ & +C_{33}C_{44}n_3^4-C_{44}\left(n_1^2+n_2^2\right)\rho v^2-C_{33}n_3^2\rho v^2-C_{44}n_3^2\rho v^2 \\ & +C_{11}\left(n_1^2+n_2^2\right)\left[C_{44}\left(n_1^2+n_2^2\right)+C_{33}n_3^2-\rho v^2\right]+\rho^2v^4\}.\end{aligned}$$

Examining the above expression and using the properties of the components of the unit vector, namely, $n_1^2+n_2^2=1-n_3^2$, we can write this determinant as a function of a single component, namely, n_3.

Rearranging the determinantal expression, we can write it as a product of the quadratic expression in v multiplied by the biquadratic expression in v, namely,

$$\begin{aligned}\det\left[\mathbf{\Gamma}-\rho v^{2}\mathbf{I}\right] = {} & \left[C_{66}\left(1-n_3^2\right)+C_{44}n_3^2-\rho v^2\right] \\ & \left\{\left[C_{33}C_{44}n_3^4-\left[2C_{13}C_{44}-C_{11}C_{33}+C_{13}^2\right]n_3^2\left(1-n_3^2\right)\right.\right. \\ & \left.+C_{11}C_{44}\left(1-n_3^2\right)^2\right] \\ & \left.+\left[\left(C_{11}-C_{33}\right)n_3^2-\left(C_{11}+C_{44}\right)\right]\rho v^2+\rho^2v^4\right\}.\end{aligned} \tag{10.25}$$

Note that determinant (10.25) is independent of n_1 and n_2. It depends only on n_3, namely, the orientation of the wavefront normal, $\mathbf{n}$, with respect to the x_3-axis, which is the rotation-symmetry axis. The absence of n_1 and n_2 illustrates the fact that to study the properties of a transversely isotropic continuum, we can use an arbitrary plane that contains the rotation-symmetry axis.

Following equation (10.8) and, hence, setting expression (10.25) to zero, we immediately obtain the equation to be solved for the three velocities.

Solving the quadratic equation, shown in brackets in expression (10.25), and considering only the positive root, we obtain

$$v\left(\mathbf{n}\right)=\sqrt{\frac{C_{66}\left(1-n_3^2\right)+C_{44}n_3^2}{\rho}}. \tag{10.26}$$

Setting $n_3=0$ and comparing to expressions (10.22), (10.23) and (10.24), we recognize expression (10.26) as corresponding to expression (10.23). Thus, we denote it as

$$v_{SH}(\mathbf{n}) = \sqrt{\frac{C_{66}\left(1-n_3^2\right)+C_{44}n_3^2}{\rho}}. \tag{10.27}$$

Solving the biquadratic equation, shown in braces in expression (10.25), we obtain two solutions. Again, setting $n_3 = 0$, we recognize them as corresponding to expressions (10.22) and (10.24). We denote them as v_{qP} and v_{qSV}, respectively. Following algebraic simplifications and considering only the positive roots, we can write these two solutions as

$$v_{qP}(\mathbf{n}) = \sqrt{\frac{(C_{33}-C_{11})\,n_3^2+C_{11}+C_{44}+\sqrt{\Delta}}{2\rho}}, \tag{10.28}$$

and

$$v_{qSV}(\mathbf{n}) = \sqrt{\frac{(C_{33}-C_{11})\,n_3^2+C_{11}+C_{44}-\sqrt{\Delta}}{2\rho}}, \tag{10.29}$$

where the discriminant, Δ, is

$$\begin{aligned}\Delta \equiv {} & \left[(C_{11}-C_{33})\,n_3^2-C_{11}-C_{44}\right]^2 \\ & -4\left[C_{33}C_{44}n_3^4-\left[2C_{13}C_{44}-C_{11}C_{33}+C_{13}^2\right]n_3^2\left(1-n_3^2\right)\right. \\ & \left. +C_{11}C_{44}\left(1-n_3^2\right)^2\right].\end{aligned}$$

Note that since the n_3 component can be written as

$$n_3 = \cos\vartheta, \tag{10.30}$$

where ϑ is the phase angle, velocity expressions, (10.27), (10.28) and (10.29), can be immediately stated in terms of the phase angle.

Displacement directions

To find the displacement directions of waves propagating in a transversely isotropic continuum, we consider, with no loss of generality, any plane that contains the rotation-symmetry axis. Letting this plane coincide with the x_1x_3-plane, we set $n_2 = 0$, and, hence, using entries (10.20), we can write the coefficient matrix of system (10.7) as

$$\begin{bmatrix} n_1^2C_{11}+n_3^2C_{44}-\rho v^2 & 0 & n_1n_3\,(C_{13}+C_{44}) \\ 0 & n_1^2C_{66}+n_3^2C_{44}-\rho v^2 & 0 \\ n_1n_3\,(C_{13}+C_{44}) & 0 & n_1^2C_{44}+n_3^2C_{33}-\rho v^2 \end{bmatrix}. \tag{10.31}$$

Considering equations (10.7) and in view of the coefficient matrix (10.31), we see that the second equation is not coupled with the remaining two. Hence, we can rewrite system (10.7) as

$$\left[n_1^2 C_{66} + n_3^2 C_{44} - \rho v^2 (\mathbf{n})\right] A_2 = 0, \tag{10.32}$$

and

$$\begin{bmatrix} n_1^2 C_{11} + n_3^2 C_{44} - \rho v^2 (\mathbf{n}) & n_1 n_3 (C_{13} + C_{44}) \\ n_1 n_3 (C_{13} + C_{44}) & n_1^2 C_{44} + n_3^2 C_{33} - \rho v^2 (\mathbf{n}) \end{bmatrix} \begin{bmatrix} A_1 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{10.33}$$

Note the decoupling of a 3×3 matrix into a 1×1 matrix and a 2×2 matrix, where, the former corresponds to the SH waves while the latter corresponds to the qP and the qSV waves. This decoupling of mathematical entities has a physical reason. The displacement vector associated with equation (10.32) is parallel to the x_2-axis, while the displacement vectors associated with equations (10.33) are contained in the x_1x_3-plane. Since the two sets of displacement vectors are orthogonal to one another and, hence, do not share any components, they do not affect one another.

Let us investigate the displacement vector associated with equation (10.32). The trivial solution is $A_2 = 0$. To find a nontrivial solution, we consider a nonzero vector. This displacement vector, $\mathbf{A} = [0, A_2, 0]$, is parallel to the x_2-axis and, hence, it is orthogonal to the propagation plane. Such a displacement must result from the propagation of a pure SH wave.

We can verify that expression (10.27), which can be written as

$$\rho v_{SH}^2 (\mathbf{n}) = n_1^2 C_{66} + n_3^2 C_{44}, \tag{10.34}$$

corresponds to the SH wave. Inserting expression (10.34) into equation (10.32), we notice that the term in brackets vanishes, as expected. In accordance with the theory of linear equations, this also means that any value of A_2 satisfies equation (10.32). In other words, this equation constrains the orientation, but not the magnitude, of the displacement vector.

Now, we focus our attention on the displacement vectors associated with the remaining two equations, which are stated in system (10.33) and correspond to the qP and qSV waves. The trivial solution is $A_1 = A_3 = 0$. To find a nontrivial solution, we consider a nonzero displacement vector, $\mathbf{A} = [A_1, 0, A_3]$, which is contained in the x_1x_3-plane.

System (10.33) allows us to show that, in general, in anisotropic continua, the displacement direction is neither parallel nor orthogonal to the direction of propagation. To do so, we find the angle that the displacement vector makes with the x_3-axis. This angle is given by

$$\phi = \arctan \frac{A_1}{A_3}. \tag{10.35}$$

Using the second equation of system (10.33), we obtain

$$\frac{A_1}{A_3} = \frac{\rho v^2 (\mathbf{n}) - n_1^2 C_{44} - n_3^2 C_{33}}{n_1 n_3 (C_{13} + C_{44})}. \tag{10.36}$$

Note that the same value of the displacement angle is obtained if we use the first equation of system (10.33), as illustrated in Exercise 10.5.

Since $\mathbf{n}^2 = n_1^2 + n_3^2 = 1$ and n_3 is given by expression (10.30), we can write expression (10.35) as

$$\phi = \arctan \frac{\rho v^2 (\vartheta) - C_{44} \sin^2 \vartheta - C_{33} \cos^2 \vartheta}{(C_{13} + C_{44}) \sin \vartheta \cos \vartheta}, \tag{10.37}$$

where $v(\vartheta)$ is given by expressions (10.28) or (10.29), together with expression (10.30), for the qP or qSV waves, respectively. In other words, if we wish to find the displacement direction associated with the qP wave, we insert expressions (10.28) and (10.30) into expression (10.37). If we wish to find the displacement direction associated with the qSV wave, we insert expressions (10.29) and (10.30) into expression (10.37).

Examining expression (10.37), we see that, in general, ϕ and ϑ are neither equal to one another nor differ by precisely $\pi/2$; this is shown in Figure 10.1. Hence, in general, in anisotropic continua, waves do not propagate as pure-mode waves. However, in many geological materials, the angle between ϕ and ϑ is not much different from 0 or $\pi/2$; this is the reason for our referring to these waves as *quasiP* or *quasiS*, respectively.

Note that, as expected from the theory of linear equations, in spite of having determined the orientation of the displacement vector, we still have infinitely many nontrivial solutions given by $A_1 = s$ and $A_3 = ms$, where s is a nonzero parameter and m is the right-hand side of equation (10.36). In other words, we find the orientation but not the magnitude of the displacement vectors.

10.3 Phase-slowness surfaces

Introductory comments

Let us consider a point within a continuum and the phase-slowness vectors emanating, in every direction, from this point. The phase-slowness surface is a surface that contains the endpoints of these phase-slowness vectors. In general, in view of three distinct velocities, there are three distinct sheets of the phase-slowness surface.

Phase-slowness surfaces are used in formulating and applying seismic theory associated with anisotropic continua, as shown in Chapter 11. They possess important topological properties. For the elasticity parameters used to describe geological materials, the two outer sheets of the phase-slowness surface intersect. In other words, the magnitudes of the phase velocity of the two slower waves coincide for certain propagation directions.

10.3.1 Convexity of innermost sheet

In general, in elastic continua, the phase-slowness surface — which, for the transversely isotropic case, results from the bicubic equation, given by expression (10.25) — is of degree 6. Consequently, any straight line can intersect the surface at, at most, six points. Since the line intersecting the innermost sheet of the phase-slowness surface must intersect the two outer sheets twice, the innermost sheet can be intersected at, at most, two points. This results in the following theorem.[4]

Theorem 10.3 *In elastic continua, the innermost sheet of the phase-slowness surface is convex.*

In other words, the phase-slowness surface of the wave exhibiting the greatest velocity cannot have any inflection points.

10.3.2 Intersection points

In general, S_1 and S_2 waves are the two slower waves. As stated above, along certain directions, the velocities of these waves must be the same. We wish to find these directions for the S waves propagating in transversely isotropic continua, namely, the intersections of the SH and qSV phase-slowness surfaces.

In a transversely isotropic continuum, discussed herein, we consider a cross-section of the phase-slowness surface in the x_1x_3-plane. In view of phase-velocity expressions (10.27), (10.28) and (10.29), and using expression (10.30), the corresponding phase-slowness curves can be generated as a polar plot with the radius given by the reciprocal of the phase-velocity magnitude.

[4]Interested readers might refer to Musgrave, M.J.P., (1970) Crystal acoustics: Introduction to the study of elastic waves and vibrations in crystals: Holden-Day, pp. 91 – 92.

Note that the intersection points of the phase-slowness curves in the x_1x_3-plane correspond to intersection lines of the phase-slowness surfaces in the $x_1x_2x_3$-space. In view of the rotation symmetry about the x_3-axis, these lines are circles that are parallel to the x_1x_2-plane.

Consider determinant (10.25). In view of the fact that the quadratic expression in v^2 contains SH waves while the biquadratic expression in v^2 contains qSV waves, at the intersection points the solution of the quadratic equation must satisfy the biquadratic equation for values of $n_3 \in [-1, 1]$. Thus, inserting v^2 — given by expression (10.27) — into the biquadratic part of equation (10.25) — shown in braces — and simplifying, we obtain

$$\left(n_3^2 - 1\right)\{\left(C_{66} - C_{11}\right)\left(C_{44} - C_{66}\right) + \left[\left(C_{13} + C_{44}\right)^2 - \left(C_{11} - C_{66}\right)\left(C_{33} - 2C_{44} + C_{66}\right)\right] n_3^2\} = 0,$$

which is an expression of the form

$$\left(n_3^2 - 1\right)\left(A + Bn_3^2\right) = 0, \tag{10.38}$$

where

$$A := \left(C_{66} - C_{11}\right)\left(C_{44} - C_{66}\right),$$

and

$$B := \left(C_{13} + C_{44}\right)^2 - \left(C_{11} - C_{66}\right)\left(C_{33} - 2C_{44} + C_{66}\right).$$

Hence, immediate solutions of equation (10.38) are given by

$$n_3 = \pm 1,$$

which correspond to the propagation along the rotation-symmetry axis. Setting $n_3 = \pm 1$ in expressions (10.27) and (10.29), we can verify that these are the velocities of SH and qSV waves that are equal to one another. This equality is consistent with the physical consequences of transverse isotropy, where, for the propagation along the rotation-symmetry axis, all displacements orthogonal to this axis are subject to the same elastic properties.

The remaining solutions of equation (10.38) depend on the values of A and B, namely, on the properties of a given continuum given by its elasticity parameters, C_{mn}. In general, we get four distinct cases, namely,

- if $B = 0$, and $A \neq 0$, there are no additional solutions and the magnitudes of the velocity coincide only for the propagation along the rotation-symmetry axis.

- if $B \neq 0$, and $A/B > 0$, there are no additional solutions and the magnitudes of the velocity coincide only for the propagation along the rotation-symmetry axis. Also, except at those two points, the qSV-wave velocity is greater than the SH-wave velocity.

- if $B \neq 0$, and $A/B \leq 0$, there is an additional solution given by

$$n_3 = \pm\sqrt{\frac{(C_{11} - C_{66})(C_{44} - C_{66})}{(C_{13} + C_{44})^2 - (C_{11} - C_{66})(C_{33} - 2C_{44} + C_{66})}}. \tag{10.39}$$

- if $A = B = 0$, all values of n_3 are the solutions and, hence, the magnitudes of the SH-wave and the qSV-wave velocities coincide for all directions. This is the case for isotropic continua.

In a seismological context, expression (10.39) is of particular interest, because, in connection with expression (10.30), it gives the value of the phase angle at which the intersection points occur, as shown in Exercise 10.2. The equality of the two shear-wave phase velocities results from the equality of two eigenvalues. Consequently, the two corresponding eigenvectors, and, hence, the displacement-vector directions, are not uniquely determined. This is also the case for S waves in isotropic continua, as stated in Remark 10.1, which follows Exercise 10.1.

Closing remarks

Explicit velocity and displacement-angle expressions allow us to study wave phenomena in the context of specific materials. In particular, these expressions can be used in formulating inverse problems where the elasticity parameters are calculated based on the traveltime and displacement-angle information, which are obtained from experimental measurements.

Studying anisotropic materials, we need to consider three types of angles, namely, the phase angles, discussed in Chapters 6 and 7, the ray angles, introduced in Chapter 9, as well as the displacement angles, discussed herein. As illustrated in Exercise 10.11, all three angles are related by analytical expressions. However, each angle plays a distinct role in theoretical formulations and the analysis of experimental measurements.

Exercises

Exercise 10.1 [5]*Formulate and solve equation (10.8) for isotropic continua.*

Solution 10.1 *Since isotropy implies directional invariance, with no loss of generality, consider propagation along the x_3-axis and, hence, let $n_1 = n_2 = 0$ and $n_3 = 1$. Considering elasticity matrix (5.64) and following entries (10.6), we can write equation (10.8) as*

$$\det\begin{bmatrix} \mu-\rho v^2 & 0 & 0 \\ 0 & \mu-\rho v^2 & 0 \\ 0 & 0 & \lambda+2\mu-\rho v^2 \end{bmatrix} = 0, \tag{10.40}$$

to obtain

$$\left(\rho v^2-\mu\right)^2\left[\rho v^2-(\lambda+2\mu)\right] = 0. \tag{10.41}$$

Hence, the solutions are $v_1 = v_2 = \sqrt{\mu/\rho}$ and $v_3 = \sqrt{(\lambda+2\mu)/\rho}$, as expected in view of equations (6.17) and (6.13), respectively.

Remark 10.1 *The first two solutions in Exercise 10.1 correspond to the S waves since we can write the corresponding displacement directions as vectors $\mathbf{A} = [1,0,0]^T$ and $\mathbf{A} = [0,1,0]^T$, which are orthogonal to the direction of propagation, $\mathbf{n} = [0,0,1]^T$. The third solution corresponds to the P waves since we can write the corresponding displacement direction as vector $\mathbf{A} = [0,0,1]^T$, which is parallel to the direction of propagation. In view of the double root in equation (10.41), there are only two eigenspaces associated with matrix $\mathbf{\Gamma}$ for an isotropic case, unlike for the anisotropic case, where there are three eigenspaces. Exercise 10.1 shows that in isotropic continua the displacement directions of S waves are contained in the plane that is orthogonal to the direction of propagation. However, these displacement directions cannot be determined uniquely, as is the case for anisotropic continua.*

Exercise 10.2 *Given the values of the elasticity parameters of the Green-river shale*[6]*, namely,*

$$\left\{\begin{array}{l} C_{11} = 3.13\times 10^{10} N/m^2 \\ C_{13} = 0.34\times 10^{10} N/m^2 \\ C_{33} = 2.25\times 10^{10} N/m^2 \\ C_{44} = 0.65\times 10^{10} N/m^2 \\ C_{66} = 0.88\times 10^{10} N/m^2 \end{array}\right. , \tag{10.42}$$

[5] See also Section 10.3.2

[6] These values are stated by Thomsen, L., (1986) Weak elastic anisotropy: Geophysics, **51**, 1954 – 1966.

and using expression (10.30) and (10.39), find the intersection points for the SH and qSV waves.

Solution 10.2 *Invoking expression (10.30) and (10.39), we obtain*

$$\begin{aligned}\vartheta &= \arccos\sqrt{\frac{(C_{11}-C_{66})(C_{44}-C_{66})}{(C_{13}+C_{44})^2-(C_{11}-C_{66})(C_{33}-2C_{44}+C_{66})}}\\ &\approx 66^0.\end{aligned}$$

Exercise 10.3 *In view of Section 10.3.2, show that for isotropic continua, SH-wave velocity and SV-wave velocity coincide for all directions.*

Solution 10.3 *As shown in elasticity matrix (5.64), for an isotropic continuum, we have*

$$\begin{aligned}C_{11} &= C_{22} = C_{33} = \lambda + 2\mu,\\ C_{13} &= \lambda,\\ C_{44} &= C_{66} = \mu,\end{aligned}$$

where λ and μ are Lamé's parameters. Thus,

$$\begin{aligned}A &= (C_{11}-C_{66})(C_{44}-C_{66}) = (\lambda+\mu)(\mu-\mu)\\ &= 0,\end{aligned}$$

and

$$\begin{aligned}B &= (C_{13}+C_{44})^2-(C_{11}-C_{66})(C_{33}-2C_{44}+C_{66}) = (\lambda+\mu)^2-(\lambda+\mu)^2\\ &= 0.\end{aligned}$$

As stated in Section 10.3.2, if $A = B = 0$, the phase-slowness curves coincide for all directions.

Exercise 10.4 *Using expression (10.37), namely,*

$$\phi = \arctan\frac{\rho v^2(\vartheta) - C_{44}\sin^2\vartheta - C_{33}\cos^2\vartheta}{(C_{13}+C_{44})\sin\vartheta\cos\vartheta}, \tag{10.43}$$

show that, for P waves in isotropic continua, the phase angle, ϑ, and the displacement angle, ϕ, coincide.

Solution 10.4 *Considering the elasticity matrix for an isotropic continuum, namely, matrix (5.62), we see that*

$$C_{13} = C_{11} - 2C_{44},$$

and

$$C_{11} = C_{33}.$$

Considering elasticity matrix (5.64) and expression (6.13), we can express the velocity of a P wave in an isotropic continuum as

$$v_P = \sqrt{\frac{C_{33}}{\rho}}.$$

Hence, expression (10.43) can be rewritten as

$$\phi = \arctan \frac{C_{11} - C_{44}\sin^2\vartheta - C_{11}\cos^2\vartheta}{(C_{11} - C_{44})\sin\vartheta\cos\vartheta}.$$

Rearranging and using standard trigonometric identities, we obtain

$$\phi = \arctan \frac{(C_{11} - C_{44})\sin^2\vartheta}{(C_{11} - C_{44})\sin\vartheta\cos\vartheta} = \arctan(\tan\vartheta).$$

Hence, $\phi = \vartheta$, as required and as expected from our discussion in Section 6.2.

Exercise 10.5 *Expression (10.43) is obtained using the second equation of system (10.33). Verify that using the first equation of this system to obtain A_1/A_3, we get the same result as shown in Exercise 10.4.*

Solution 10.5 *Using the first equation of system (10.33), we can write expression (10.35) as*

$$\phi = \arctan \frac{A_1}{A_3} = \arctan \frac{(C_{13} + C_{44})\sin\vartheta\cos\vartheta}{\rho v^2(\vartheta) - C_{11}\sin^2\vartheta - C_{44}\cos^2\vartheta}. \tag{10.44}$$

In view of the isotropic-case expressions, stated in Exercise 10.4, we can rewrite expression (10.44) as

$$\phi = \arctan \frac{(C_{11} - C_{44})\sin\vartheta\cos\vartheta}{C_{11} - C_{11}\sin^2\vartheta - C_{44}\cos^2\vartheta} = \arctan(\tan\vartheta).$$

Hence, $\phi = \vartheta$, as required.

Exercise 10.6 *Using expression (10.43), show that, for S waves in isotropic continua, the phase angle, ϑ, and the displacement angle, ϕ, differ by $\pi/2$, which implies that the propagation and displacement directions are orthogonal to one another.*

Solution 10.6 *Considering the elasticity matrix for an isotropic continuum, namely, matrix (5.62), we see that*

$$C_{13} = C_{11} - 2C_{44},$$

and

$$C_{11} = C_{33}.$$

Considering elasticity matrix (5.64) and expression (6.17), we can express the velocity of an S wave in an isotropic continuum as

$$v_S = \sqrt{\frac{C_{44}}{\rho}}.$$

Hence, in a manner analogous to the one used to obtain the solution of Exercise 10.4, expression (10.43) becomes

$$\phi = \arctan\left(-\cot\vartheta\right) = -\arctan\left(\cot\vartheta\right).$$

Using properties of the inverse trigonometric functions, we can rewrite this expression as

$$\phi = \arctan\left(\tan\vartheta\right) - \frac{\pi}{2} = \vartheta - \frac{\pi}{2},$$

as required and as expected from our discussion in Section 6.2.

Exercise 10.7 *Using determinant (10.25) obtain expressions (10.22), (10.23) and (10.24).*

Solution 10.7 *Consider the determinantal expression (10.25), namely,*

$$\begin{aligned}\det\left[\boldsymbol{\Gamma} - \rho\left(\mathbf{x}\right)v^2\mathbf{I}\right] = {} & \left[C_{66}\left(1-n_3^2\right) + C_{44}n_3^2 - \rho v^2\right] \\ & \left\{\left[C_{33}C_{44}n_3^4 - \left[2C_{13}C_{44} - C_{11}C_{33} + C_{13}^2\right]n_3^2\left(1-n_3^2\right)\right.\right. \\ & \left. + C_{11}C_{44}\left(1-n_3^2\right)^2\right] \\ & \left. + \left[\left(C_{11}-C_{33}\right)n_3^2 - \left(C_{11}+C_{44}\right)\right]\rho v^2 + \rho^2 v^4\right\}.\end{aligned}$$

To consider propagation in the plane of transverse isotropy, we let $n_3 = 0$ *to obtain*

$$\det\left[\mathbf{\Gamma} - \rho\left(\mathbf{x}\right)v^2\mathbf{I}\right] = \left(C_{66} - \rho v^2\right)\left[\rho^2 v^4 - \left(C_{11} + C_{44}\right)\rho v^2 + C_{11}C_{44}\right]. \tag{10.45}$$

Setting expression (10.45) to zero, we obtain expressions (10.22), (10.23) and (10.24), as required.

Exercise 10.8 [7] *Show that SH waves in transversely isotropic continua exhibit elliptical velocity dependence.*

Solution 10.8 *Consider expression (10.27). Recalling expression (10.30) and using trigonometric identities, we can write*

$$v_{SH}\left(\vartheta\right) = \sqrt{\frac{C_{66}}{\rho}\sin^2\vartheta + \frac{C_{44}}{\rho}\cos^2\vartheta}.$$

Setting $\vartheta = 0$*, we get* $v_{SH}\left(0\right) = \sqrt{C_{44}/\rho}$*, while setting* $\vartheta = \pi/2$*, we get* $v_{SH}\left(\pi/2\right) = \sqrt{C_{66}/\rho}$*, which can be denoted as* v_z *and* v_x*, respectively. Thus, we can write*

$$v_{SH}\left(\vartheta\right) = \sqrt{v_x^2\sin^2\vartheta + v_z^2\cos^2\vartheta},$$

which is expression (6.60), giving the magnitude of phase velocity for the case of elliptical velocity dependence.

Exercise 10.9 *Formulate Hamiltonian* $\mathcal{H}$ *that corresponds to SH waves in a transversely isotropic continuum.*

Solution 10.9 *In view of expression (8.20) and considering a given point of the continuum, we can write the corresponding ray-theory Hamiltonian as*

$$\mathcal{H}\left(\mathbf{p}\right) = \frac{1}{2}p^2 v^2\left(\mathbf{p}\right).$$

Considering the SH-wave velocity given by expression (10.27), namely,

$$v_{SH}^2\left(\mathbf{n}\right) = \frac{C_{66}\left(1 - n_3^2\right) + C_{44}n_3^2}{\rho},$$

and since $n_i^2 = p_i^2/p^2$ *and* $n_1^2 = 1 - n_3^2$*, we can write*

$$v_{SH}^2\left(\mathbf{p}\right) = \frac{C_{66}\dfrac{p_1^2}{p^2} + C_{44}\dfrac{p_3^2}{p^2}}{\rho} = \frac{1}{p^2}\frac{C_{66}p_1^2 + C_{44}p_3^2}{\rho}.$$

[7] See also Section 6.6.2

Hence, we can write

$$\mathcal{H}_{SH}(\mathbf{p}) = \frac{1}{2}\frac{C_{66}p_1^2 + C_{44}p_3^2}{\rho}. \tag{10.46}$$

Exercise 10.10 *Using Legendre's transformation and expression (10.46), find the corresponding relation between the phase and the ray angles for SH waves in a transversely isotropic continuum.*

Solution 10.10 *As shown in expression (9.30), the ray angle is given by*

$$\tan\theta = \frac{\frac{\mathrm{d}x_1}{\mathrm{d}s}}{\frac{\mathrm{d}x_3}{\mathrm{d}s}}.$$

Using time parametrization, we can immediately restate this expression as

$$\tan\theta = \frac{\frac{\mathrm{d}x_1}{\mathrm{d}t}}{\frac{\mathrm{d}x_3}{\mathrm{d}t}} \equiv \frac{\dot{x}_1}{\dot{x}_3},$$

where t denotes time. In view of transformation (B.12), we can write

$$\dot{x}_i = \frac{\partial\mathcal{H}}{\partial p_i}.$$

Thus, using expression (10.46), we obtain

$$\dot{x}_1 = \frac{1}{2}\frac{\partial}{\partial p_1}\frac{C_{66}p_1^2 + C_{44}p_3^2}{\rho} = \frac{C_{66}}{\rho}p_1,$$

and

$$\dot{x}_3 = \frac{1}{2}\frac{\partial}{\partial p_3}\frac{C_{66}p_1^2 + C_{44}p_3^2}{\rho} = \frac{C_{44}}{\rho}p_3.$$

Hence, we can write

$$\tan\theta = \frac{\dot{x}_1}{\dot{x}_3} = \frac{C_{66}}{C_{44}}\frac{p_1}{p_3}.$$

Recalling expression (9.14), we can restate the above expression in terms of the phase angle, as

$$\tan\theta = \frac{C_{66}}{C_{44}}\tan\vartheta. \tag{10.47}$$

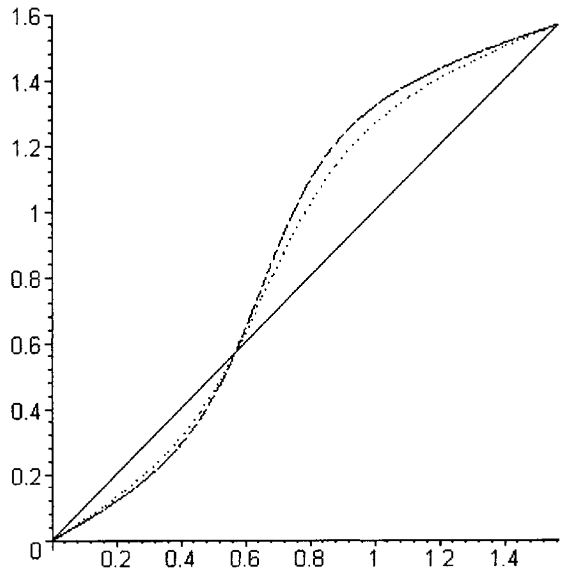


Figure 10.1: Solution of Exercise 10.11. The ray angle (dashed line) and the displacement angle (dotted line) are plotted as functions of the phase angle. The units of both axes are displayed in radians.

Remark 10.2 *Expression (10.47) allows us to explicitly express the phase angle as a function of the ray angle and vice versa, in terms of the properties of the continuum given by its elasticity parameters. An explicit, closed-form expression of the phase angle in terms of the ray angle is possible only for elliptical velocity dependence.*

Exercise 10.11 *Using expression (10.37), namely,*

$$\phi = \arctan \frac{\rho v^2(\vartheta) - C_{44}\sin^2\vartheta - C_{33}\cos^2\vartheta}{(C_{13}+C_{44})\sin\vartheta\cos\vartheta}, \tag{10.48}$$

and expression (9.21), which can be rewritten as

$$\theta = \arctan \frac{\tan\vartheta + \dfrac{1}{v(\vartheta)}\dfrac{\partial v(\vartheta)}{\partial\vartheta}}{1 - \dfrac{\tan\vartheta}{v(\vartheta)}\dfrac{\partial v(\vartheta)}{\partial\vartheta}}, \tag{10.49}$$

as well as the elasticity parameters of the Green-river shale, shown in expressions (10.42), and its mass density, given by $\rho = 2310\ kg/m^3$*, plot the* *displacement angles, ϕ, and the ray angle, θ, as a function of the phase angle, ϑ, for qP waves.*

Solution 10.11 *Inserting phase-velocity expression (10.28) and expression (10.30), into expressions (10.48) and (10.49), we generate the plot of the displacement and the ray angles, respectively. This plot is shown in Figure 10.1.*

Remark 10.3 *Figure 10.1 shows that, in general, the phase angles, the ray angles, and the displacement angles are distinct. For qP waves, the three angles coincide along the pure-mode directions, where qP waves are reduced to P waves. As illustrated using the elasticity parameters of the Green-river shale, the pure-mode directions occur at $\vartheta = 0$ and $\vartheta = \pi/2$, as well as — in view of expressions (10.48) and (10.49) — at the phase angle satisfying equation*

$$\frac{\rho v_{qP}^2(\vartheta) - C_{44}\sin^2\vartheta - C_{33}\cos^2\vartheta}{(C_{13}+C_{44})\sin\vartheta\cos\vartheta} = \frac{\tan\vartheta + \dfrac{1}{v(\vartheta)}\dfrac{\partial v(\vartheta)}{\partial\vartheta}}{1 - \dfrac{\tan\vartheta}{v(\vartheta)}\dfrac{\partial v(\vartheta)}{\partial\vartheta}}.$$

Examining Figure 10.1, we see that the values of the displacement angle are closer to the values of the ray angle than to the values of the phase angle.[8]

[8]Readers interested in relations among the phase angle, the ray angle and the displacement angle might also refer to Tsvankin, I., (2001) Seismic signatures and analysis of reflection data in anisotropic media: Pergamon, pp. 34 – 36.

This Page Intentionally Left Blank

Chapter 11

Reflection and transmission

A "perfect" scientific theory may be described as one which proceeds logically from a few simple hypotheses to conclusions which are in complete agreement with observation, to within the limits of accuracy of observation. [...] As accuracy of observation increases, a theory ceases to be perfect.

John Lighton Synge (1937) Geometrical optics: An introduction to Hamilton's method

Preliminary remarks

Discussing ray theory in Chapter 7, we assumed the smoothness of functions describing mass density and elasticity parameters. Hence, the velocity function was smooth with respect to both position and direction. In other words, we assumed that the continuum was not separated by interfaces.

Certain seismic techniques do not require any *a priori* treatment of interfaces and, hence, smooth velocity functions suffice. For instance, for imaging seismic data, we might only need a background velocity field, which can be given by a smooth function. Other seismological studies, however, require an explicit treatment of interfaces. In particular, we need to consider interfaces to study the phenomena of reflection and transmission. To study these phenomena, we invoke the principles of the continuity of phase, the equality of the sum of displacements and the equality of the traction components across the interface.

We begin this chapter with the derivation of relations among the incidence, reflection, and transmission angles for interfaces between two anisotropic continua. A specific case of elliptical velocity dependence is used to illustrate the general formulation. Then, we consider the amplitudes of the reflected and transmitted signals as functions of the angle of incidence. For a mathematical convenience, the explicit expressions are derived only for the case of SH waves in transversely isotropic continua.

11.1 Angles at interface

11.1.1 Phase angles

Consider a three-dimensional continuum that is composed of parallel homogeneous layers of finite thickness. Let each layer be parallel to the x_1x_2-plane. We choose to view the $x_1x_2x_3$-coordinate system in such a way that we refer to the x_3-axis as the vertical axis. In other words, herein, we study phenomena associated with horizontal layers.

Recall Hamilton's ray equations (8.19), namely,

$$\left\{\begin{array}{l}\dot{x}_i = \dfrac{\partial \mathcal{H}}{\partial p_i}\\[2em] \dot{p}_i = -\dfrac{\partial \mathcal{H}}{\partial x_i}\end{array}\right., \qquad i \in \{1,2,3\}, \tag{11.1}$$

where Hamiltonian $\mathcal{H}$ is given by expression (8.20), namely,

$$\mathcal{H} = \frac{1}{2}p^2 v^2 (\mathbf{x}, \mathbf{p}). \tag{11.2}$$

Examining equations (11.1) and expression (11.2), in view of the horizontal layering, where the elastic properties remain unchanged along the x_1-axis and the x_2-axis, we see that

$$\dot{p}_i \equiv \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial x_i} = 0, \qquad i \in \{1,2\}.$$

Consequently, p_1 and p_2 are constant for a given solution curve $\mathbf{x}(t)$. In other words, the phase-slowness vector components that are parallel to the interfaces are conserved across these interfaces. We refer to this property as the continuity of phase.

The continuity of phase can be justified by a physical argument. The continuity of phase is tantamount to the continuity of wavefronts, which are the loci of constant phase. Equality of p_i, where $i \in \{1,2\}$, across the interface implies that although the orientation of vector $\mathbf{p}$ might change, its horizontal components must remain the same. In other words, the wavefronts are continuous across the interface. We can see this requirement as resulting from Huygens' principle and from the associated causality.

Let us consider propagation in the x_1x_3-plane. In other words, let $\mathbf{p} = [p_1, 0, p_3]$. We can write the horizontal component of the phase-slowness vector as

$$p_1 = |\mathbf{p}|\, n_1,$$

where $|\mathbf{p}|$ is the magnitude of the phase-slowness vector and n_1 is the horizontal component of the unit vector normal to the wavefront. Recalling expression (10.30) and using the fact that, in the x_1x_3-plane, $n_1 = \sqrt{1 - n_3^2}$, we obtain

$$p_1 = |\mathbf{p}| \sin\vartheta,$$

where ϑ is the phase angle, which is measured between the wavefront normal and the vertical axis.

Since p_1 is conserved across the interfaces separating homogeneous horizontal layers, we denote this constant by $\mathfrak{p}$. Now, since the magnitudes of phase slowness and phase velocity are the reciprocals of one another, we can write conserved quantity $\mathfrak{p}$ as

$$\mathfrak{p} = \frac{\sin\vartheta}{v(\vartheta)}, \tag{11.3}$$

where $v(\vartheta)$ gives the magnitude of phase velocity as a function of the phase angle.

Expression (11.3) is a general statement of Snell's law in the context of phase angle and phase velocity. It is valid across interfaces between generally anisotropic continua. Since $\mathfrak{p}$ is a conserved quantity for a given solution curve $\mathbf{x}(t)$, which corresponds to a ray, we refer to $\mathfrak{p}$ as ray parameter. We will discuss it further, in the context of Hamilton's and Lagrange's ray equations, in Section 14.6.

The continuity of the horizontal phase-slowness components provides us with a convenient formulation to relate the angles of incidence, reflection and transmission.[1]

11.1.2 Ray angles

We wish to use the continuity of the phase-slowness components to derive the relation between the ray angles across the interface.

[1]Readers interested in a geometrical formulation of the relation among the incidence, reflection and transmission angles might refer to Auld, B.A., (1973) Acoustic fields and waves in solids: John Wiley and Sons, Vol. II, pp. 1 – 14.

Using Legendre's transformation, an analytic relation between the phase angles and the ray angles was derived in Section 9.2.2 and given by expression (9.21). This expression also states the polar reciprocity between the phase-slowness curve and the ray-velocity curve, which is explicitly shown in Exercise 9.3, and means that, at a given point of the phase-slowness curve, the corresponding ray direction is always normal to the phase-slowness curve. Herein, we will use this geometrical property to formulate expressions relating incidence, reflection and transmission ray angles across an interface between two anisotropic continua.

Note that, while expression (11.3) is generally true for $\vartheta \in (-\pi, \pi)$, obtaining analytical expressions in terms of ray angles and ray velocities is not always possible. If we wish to obtain such expressions, we must restrict our studies to particular symmetries or use convenient approximations.[2] This is a consequence of restrictions imposed by Legendre's transformation. Nevertheless, the geometrical construction relating the phase and ray angles, which results from polar reciprocity, is possible at any given point of the phase-slowness surface of a generally anisotropic continuum.

In the following section, we consider a particular symmetry due to elliptical velocity dependence. Therein, we derive analytical expressions between the ray angles of incidence and transmission.

11.1.3 Example: Elliptical velocity dependence

Phase-slowness curves

Consider a two-dimensional continuum that is contained in the xz-plane. Let this continuum consist of two halfspaces, and let the interface coincide with the x-axis.

We wish to characterize each layer by the phase-slowness curve, which is expressed in terms of the horizontal and vertical velocities.

The two phase-slowness curves can be stated as

$$\left\{ \begin{array}{l} f(p_x, p_z) = ({}_1v_x p_x)^2 + ({}_1v_z p_z)^2 = 1 \\ g(p_x, p_z) = ({}_2v_x p_x)^2 + ({}_2v_z p_z)^2 = 1 \end{array} \right. , \tag{11.4}$$

for the medium of incidence and transmission, respectively, where v_x and v_z specify the horizontal and vertical phase velocities, respectively.

[2]Readers interested in formulations using expressions based on the weak-anisotropy approximation might refer to Slawinski, M.A., Slawinski, R.A, Brown, R.J., and Parkin, J.M., (2000), A generalized form of Snell's law in anisotropic media. Geophysics, **65**, No. 2, 632 – 637.

Note that either expression of set (11.4) is the equation of an ellipse in the $p_x p_z$-plane, given by

$$\frac{p_x^2}{\left(\frac{1}{{}_m v_x}\right)^2} + \frac{p_z^2}{\left(\frac{1}{{}_m v_z}\right)^2} = 1, \qquad m \in \{1,2\},$$

where $m = 1$ corresponds to the medium of incidence, while $m = 2$ corresponds to the medium of transmission.

Conserved quantity in terms of phase angles and phase velocities

Since the continuum is homogeneous along the x-axis, we wish to obtain the quantity that is conserved across the interface in terms of the horizontal and vertical velocities.

In view of expression (11.3), we can write

$$\mathfrak{p} = p_x = \frac{\sin\vartheta}{v(\vartheta)} = \frac{\sin\vartheta_m}{\sqrt{{}_m v_x^2 \sin^2\vartheta_m + {}_m v_z^2 \cos^2\vartheta_m}}, \qquad m \in \{1,2\}, \quad (11.5)$$

where, in view of elliptical velocity dependence, $v(\vartheta)$ is given by expression (6.60).

Conserved quantity in terms of ray angles and ray velocities

We wish to express conserved quantity (11.5) in terms of the ray angle and the ray velocity.

In view of the symmetry of the ellipse, the values of the horizontal phase velocity and vertical phase velocity are equal to the corresponding values of the ray velocities, namely, $v_x = V_x$ and $v_z = V_z$. Hence, set (11.4) can be restated as

$$\left\{ \begin{array}{l} f(p_x, p_z) = ({}_1V_x p_x)^2 + ({}_1V_z p_z)^2 = 1 \\ g(p_x, p_z) = ({}_2V_x p_x)^2 + ({}_2V_z p_z)^2 = 1 \end{array} \right. . \quad (11.6)$$

To find the angle of a normal to a phase-slowness curve, we can consider the phase-slowness curves as the level curves of functions f and g, and use the fact that the ray directions are normal to the phase-slowness curves. In view of the properties of the gradient operator and using, for instance, function f, we can write the unit vector normal to the phase-slowness curve as $\nabla_{\mathbf{p}} f / |\nabla_{\mathbf{p}} f|$, where $\nabla_{\mathbf{p}}$ is the gradient operator given by $[\partial/\partial p_x, \partial/\partial p_z]$. Now, using the scalar product, we obtain the angle between the vector normal to the phase-slowness surface and the vertical axis. This angle, which is the ray angle, is given by

$$\cos\theta_1 = \mathbf{e}_z \cdot \frac{\nabla_{\mathbf{p}} f}{|\nabla_{\mathbf{p}} f|} = \frac{\frac{\partial f}{\partial p_z}}{|\nabla_{\mathbf{p}} f|}$$

evaluated at (p_x, p_z), where $\mathbf{e}_z$ is the unit vector along the vertical axis. Thus, using expressions for f and g stated in set (11.6), we get the corresponding expression for a ray angle in elliptical velocity dependence, namely,

$$\cos\theta_m = \frac{{}_mV_z^2 p_z}{\sqrt{\left({}_mV_x^2 p_x\right)^2 + \left({}_mV_z^2 p_z\right)^2}}, \qquad m \in \{1,2\}. \tag{11.7}$$

To invoke the conserved quantity, $\mathfrak{p} = p_x$, we would like to explicitly solve equations (11.7) for p_x.

Using expressions of set (11.6), we can write

$$p_z = \frac{\sqrt{1 - \left({}_mV_x p_x\right)^2}}{{}_mV_z}, \qquad m \in \{1,2\}, \tag{11.8}$$

and, hence, inserting expressions (11.8) into equations (11.7), we get

$$\cos\theta_m = \frac{{}_mV_z\sqrt{1 - \left({}_mV_x p_x\right)^2}}{\sqrt{{}_mV_x^4 p_x^2 + {}_mV_z^2\left[1 - \left({}_mV_x p_x\right)^2\right]}}, \qquad m \in \{1,2\}. \tag{11.9}$$

Solving equations (11.9) for p_x, we obtain

$$p_x^2 = \frac{{}_mV_z^2 \sin^2\theta_m}{{}_mV_x^2\left({}_mV_z^2 \sin^2\theta_m + {}_mV_x^2\cos^2\theta_m\right)}, \qquad m \in \{1,2\}.$$

Simplifying, we can write

$$p_x^2 = \frac{1}{{}_mV_x^2\left[\left(\frac{{}_mV_x}{{}_mV_z}\right)^2 \cot^2\theta_m + 1\right]}, \qquad m \in \{1,2\}.$$

Consequently, the conserved quantity, $\mathfrak{p} = p_x$, can be written as

$$\mathfrak{p} = \frac{1}{{}_mV_x\sqrt{\left(\frac{{}_mV_x}{{}_mV_z}\right)^2 \cot^2\theta_m + 1}}, \qquad m \in \{1,2\}, \tag{11.10}$$

which is conserved quantity (11.5) stated in terms of ray angles and ray velocities.

Note that we can write expression (11.5) as

$$\mathfrak{p} = \frac{1}{{}_m v_x \sqrt{\left(\frac{{}_m v_z}{{}_m v_x}\right)^2 \cot^2 \vartheta_m + 1}}, \qquad m \in \{1, 2\}, \tag{11.11}$$

which allows us to see the similarity of form between expressions (11.5) and (11.10). Notice, however, that in expression (11.10), we have V_x/V_z, while, in expression (11.11), we have v_z/v_x.

In general, as shown in Exercise 11.1, expressions (11.5) and (11.10) are equivalent to one another. For the isotropic case, as shown in Exercise 11.2, expressions (11.5) and (11.10) become identical.

Following expression (11.10) and in view of set (11.6), we can write

$${}_1V_x^2 \left[\left(\frac{{}_1V_x}{{}_1V_z}\right)^2 \cot^2 \theta_1 + 1\right] = {}_2V_x^2 \left[\left(\frac{{}_2V_x}{{}_2V_z}\right)^2 \cot^2 \theta_2 + 1\right], \tag{11.12}$$

where the subscripts 1 and 2 correspond to the medium of incidence and transmission, respectively. Equation (11.12) can be viewed as a statement of Snell's law for elliptical velocity dependence, expressed in terms of ray angles and ray velocities.

11.2 Amplitudes at interface

11.2.1 Kinematic and dynamic boundary conditions

Introductory comments

In Section 11.1, we related the directions of waves across the interface. For this purpose, we used the continuity of phase. Herein, we will relate the amplitudes of waves across the interface. For this purpose, we will use the equality of the sum of displacements and the equality of the traction components across the interface, which we refer to as the kinematic and the dynamic boundary conditions, respectively.

In general, when a wave encounters an interface, it generates both reflected and transmitted waves. In this process, the energy of the incident wave is partially reflected and partially transmitted. The fractions of the incident-wave energy that are reflected and transmitted are functions of the direction of the incident wave and the material properties on either side of the interface. Since energy carried by a wave is directly proportional to the square of the amplitude of the displacement, which can be measured by a seismic receiver, we discuss reflection and transmission amplitudes.

The formulation presented in this section deals specifically with amplitudes of plane SH waves in the context of a plane interface between two transversely isotropic continua whose rotation-symmetry axes are normal to the interface. Also, these two continua are assumed to be in a welded contact, which implies that they cannot slip with respect to one another. SH waves are used because their elliptical velocity dependence lends itself to a convenient illustration of the physical concepts involved.

Displacement vectors

In a three-dimensional, transversely isotropic continuum, where the rotation-symmetry axis is assumed to coincide with the x_3-axis, we consider an SH wave whose phase-slowness vector, $\mathbf{p}$, is contained in the x_1x_3-plane. Hence, this SH wave exhibits a displacement in the x_2-direction only, and, consequently, we can write its displacement vector as

$$\mathbf{u} = [0, u_2, 0]\,. \tag{11.13}$$

Considering the oscillatory nature of waves and in view of expression (6.63), we can write the nonzero component of displacement as

$$u_2 = A \exp\left[i\omega\left(\mathbf{p}\cdot\mathbf{x} - t\right)\right], \tag{11.14}$$

where A denotes the amplitude of the displacement and $\exp\left[\cdot\right]$ is the phase factor.

Kinematic boundary conditions

Our kinematic boundary conditions require the equality of the sum of displacements on either side of the interface. This equality has the following physical meanings. The equality of displacements parallel to the interface implies that the materials cannot slip with respect to one another. The equality of the displacement normal to the interface implies that the materials cannot separate from one another or penetrate one another. These equalities are tantamount to the assumption of a welded contact.

In view of expression (11.14) and setting the amplitude of the incident signal to unity, we can write the kinematic boundary condition as

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \exp\left[i\omega\left(\mathbf{p}^{i}\cdot\mathbf{x}-t\right)\right]+\begin{bmatrix} 0 \\ A_r \\ 0 \end{bmatrix} \exp\left[i\omega\left(\mathbf{p}^{r}\cdot\mathbf{x}-t\right)\right]$$
$$=\begin{bmatrix} 0 \\ A_t \\ 0 \end{bmatrix} \exp\left[i\omega\left(\mathbf{p}^{t}\cdot\mathbf{x}-t\right)\right],$$

where i, r and t, as superscripts or subscripts, refer to the incident, reflected and transmitted waves, respectively. We can immediately rewrite this kinematic boundary condition as

$$\exp\left[i\omega\left(\mathbf{p}^{i}\cdot\mathbf{x}-t\right)\right]+A_r\exp\left[i\omega\left(\mathbf{p}^{r}\cdot\mathbf{x}-t\right)\right]=A_t\exp\left[i\omega\left(\mathbf{p}^{t}\cdot\mathbf{x}-t\right)\right]. \tag{11.15}$$

Dynamic boundary conditions

The dynamic boundary conditions require the equality of the traction components across the interface.

Note that the inequality of the traction components would imply a finite net force acting on a massless element of the interface. This would lead, in view of Newton's second law of motion, to a physically inacceptable concept of an infinite acceleration.

Recalling the traction, given by expression (2.31), and the symmetry of the stress tensor, stated in Theorem 2.1, we can write

$$T_i=\sum_{j=1}^{3}\sigma_{ij}n_j, \qquad i\in\{1,2,3\},$$

where $\mathbf{n}$ is the unit vector normal to the surface upon which the traction is acting. Considering an interface coinciding with an x_1x_2-plane, where the interface normal is $\mathbf{n}=[0,0,1]$, we see that that the traction components that contain n_1 and n_2 vanish identically. Consequently, their equality is trivially satisfied. Now, we can write the equality of the nonzero components of traction as

$$\begin{bmatrix} \sigma_{13}^{I}n_3 \\ \sigma_{23}^{I}n_3 \\ \sigma_{33}^{I}n_3 \end{bmatrix}=\begin{bmatrix} \sigma_{13}^{II}n_3 \\ \sigma_{23}^{II}n_3 \\ \sigma_{33}^{II}n_3 \end{bmatrix},$$

where the superscript I indicates the medium of incidence and reflection, while the superscript II indicates the medium of transmission. This equation immediately implies the equality of the stress-tensor components, namely,

$$\begin{bmatrix} \sigma_{13}^{I} \\ \sigma_{23}^{I} \\ \sigma_{33}^{I} \end{bmatrix} = \begin{bmatrix} \sigma_{13}^{II} \\ \sigma_{23}^{II} \\ \sigma_{33}^{II} \end{bmatrix}. \tag{11.16}$$

To study the amplitudes of reflected and transmitted waves in terms of the properties of the continua on either side of the interface, we wish to rewrite conditions (11.16) in terms of elasticity parameters and mass density. Recalling definition (1.15) and considering stress-strain equations (4.11) with the elasticity matrix for a transversely isotropic continuum, given by matrix (5.47), we can write

$$\sigma_{13} = C_{44}\left(\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1}\right),$$

$$\sigma_{23} = C_{44}\left(\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2}\right), \tag{11.17}$$

and

$$\sigma_{33} = C_{33}\frac{\partial u_3}{\partial x_3}.$$

In view of displacement vector (11.13), the only nonzero stress-tensor component is given by expression (11.17), which we can rewrite as

$$\sigma_{23} = C_{44}\frac{\partial u_2}{\partial x_3}. \tag{11.18}$$

All other components vanish identically and, thus, their equalities are trivially satisfied. Hence, dynamic boundary conditions (11.16) are reduced to

$$\sigma_{23}^{I} = \sigma_{23}^{II}. \tag{11.19}$$

Considering the displacement-vector components for the incident, reflected and transmitted waves and in view of expression (11.18), we can write boundary condition (11.19) as

$$C_{44}^{I}\left(\frac{\partial u_2^i}{\partial x_3} + \frac{\partial u_2^r}{\partial x_3}\right) = C_{44}^{II}\frac{\partial u_2^t}{\partial x_3}. \tag{11.20}$$

Note that the left-hand side of equation (11.20) contains the contributions to stress-tensor component σ_{23}^{I} of both the incident and reflected waves.

Invoking expression (11.14), we can write equation (11.20) as

$$C_{44}^{I}\left(\omega p_3^i \exp\left[i\omega\left(\mathbf{p}^i\cdot\mathbf{x}-t\right)\right]+\omega p_3^r A_r \exp\left[i\omega\left(\mathbf{p}^r\cdot\mathbf{x}-t\right)\right]\right)$$
$$=\omega p_3^t C_{44}^{II} A_t \exp\left[i\omega\left(\mathbf{p}^t\cdot\mathbf{x}-t\right)\right]. \quad (11.21)$$

Thus, equations (11.15) and (11.21) are the kinematic and dynamic boundary conditions, respectively, for SH waves propagating across the interface separating two transversely isotropic continua in welded contact. These equations form a system of equations to be solved for the reflection and transmission amplitudes.

Note that mass density is implicitly present in condition (11.21) since it is contained in expressions for the phase-slowness vectors.

11.2.2 Reflection and transmission amplitudes

Derivation of expressions

We wish to obtain the values of the reflection amplitude, A_r, and the transmission amplitude, A_t. Thus, we need to solve the system composed of equations (11.15) and (11.21). Since these equations relate to a point on the interface, in view of the previous assumptions, we can make certain simplifications without further affecting the generality of the formulation.

Since we are considering the interface that coincides with the x_1x_2-plane, we set $x_3 = 0$. In view of the transversely isotropic continuum with the x_3-axis corresponding to the rotation-symmetry axis and our choice of the propagation in the x_1x_3-plane, the corresponding phase-slowness vector is $\mathbf{p} = [p_1, 0, p_3]$. Furthermore, the homogeneity of the continuum along the x_1-axis and the x_2-axis allows us to conveniently choose any incidence point on the interface; hence, we choose $(0,0,0)$. Also, at the instant of incidence, the incident, reflected and transmitted waves are considered at the boundary at the same time t. Moreover, considering monochromatic waves, the value of frequency, ω, is the same for the incident, reflected and transmitted waves. Thus, equations (11.15) and (11.21) simplify to

$$1 + A_r = A_t, \quad (11.22)$$

and

$$C_{44}^{I}\left(p_3^i + A_r p_3^r\right) = C_{44}^{II} p_3^t A_t, \quad (11.23)$$

respectively.

We can further simplify condition (11.23). In view of the phase-slowness curve being symmetric about the x_3-axis, the equality of the p_1 components for the incident and reflected waves implies that $p_3^i = -p_3^r$. In other words, the vertical components of the phase-slowness vectors for the incident and reflected waves exhibit the same magnitudes and opposite directions.

Hence, dynamic boundary condition (11.23) becomes

$$C_{44}^I p_3^i (1 - A_r) = C_{44}^{II} p_3^t A_t. \tag{11.24}$$

Now, it is convenient to explicitly include mass density in condition (11.24). Since p_3 is a vertical component of the phase-slowness vector, recalling expression (10.27), we can write

$$C_{44}^I \frac{\cos\vartheta_i}{\sqrt{\frac{C_{66}^I \sin^2\vartheta_i + C_{44}^I \cos^2\vartheta_i}{\rho_1}}} (1 - A_r) = C_{44}^{II} \frac{\cos\vartheta_t}{\sqrt{\frac{C_{66}^{II} \sin^2\vartheta_t + C_{44}^{II} \cos^2\vartheta_t}{\rho_2}}} A_t. \tag{11.25}$$

Equations (11.22) and (11.25) form a system of two equations to be solved for the two unknowns, namely, the reflection and transmission amplitudes. These solutions are

$$A_r(\vartheta) = \frac{\frac{\sqrt{\rho_1} C_{44}^I \cos\vartheta_i}{\sqrt{C_{66}^I \sin^2\vartheta_i + C_{44}^I \cos^2\vartheta_i}} - \frac{\sqrt{\rho_2} C_{44}^{II} \cos\vartheta_t}{\sqrt{C_{66}^{II} \sin^2\vartheta_t + C_{44}^{II} \cos^2\vartheta_t}}}{\frac{\sqrt{\rho_1} C_{44}^I \cos\vartheta_i}{\sqrt{C_{66}^I \sin^2\vartheta_i + C_{44}^I \cos^2\vartheta_i}} + \frac{\sqrt{\rho_2} C_{44}^{II} \cos\vartheta_t}{\sqrt{C_{66}^{II} \sin^2\vartheta_t + C_{44}^{II} \cos^2\vartheta_t}}}, \tag{11.26}$$

and

$$A_t(\vartheta) = \frac{2\frac{\sqrt{\rho_1} C_{44}^I \cos\vartheta_i}{\sqrt{C_{66}^I \sin^2\vartheta_i + C_{44}^I \cos^2\vartheta_i}}}{\frac{\sqrt{\rho_1} C_{44}^I \cos\vartheta_i}{\sqrt{C_{66}^I \sin^2\vartheta_i + C_{44}^I \cos^2\vartheta_i}} + \frac{\sqrt{\rho_2} C_{44}^{II} \cos\vartheta_t}{\sqrt{C_{66}^{II} \sin^2\vartheta_t + C_{44}^{II} \cos^2\vartheta_t}}}. \tag{11.27}$$

Expressions (11.26) and (11.27) give the reflection amplitude and the transmission amplitude, respectively, for SH waves in transversely isotropic continua with the rotation-symmetry axes normal to the interface. The reflection and transmission amplitudes depend on the values of the elasticity parameters and mass density on either side of the interface, and are functions of the phase angles of incidence and transmission.

Interpretation of expressions

Examining expressions (11.26) and (11.27), we learn about the behaviour of the seismic signal in the context of its being transmitted through, or reflected from, the interface.

Depending on the values of elasticity parameters, mass densities and the incidence angle, the value of expression (11.26) can be either positive or negative. The positive sign implies that the direction of the displacement vectors for both the incident wave and the reflected wave is the same. The negative sign implies the reversal of the direction of the displacement vector. Also, while the amplitude of the incident wave is set to unity, the amplitude of the transmitted wave can be greater than unity. This is in agreement with balance of energy, as shown in Exercise 11.5.

If the values of elasticity parameters and mass densities are such that the magnitude of the velocity that is parallel to the interface is greater in the medium of transmission than in the medium of incidence, by examining expression (11.5), we conclude that once ϑ_i is large enough, $\sin\vartheta_t$ is greater than unity and, consequently, $\cos\vartheta_t = \sqrt{1-\sin^2\vartheta_t}$ is purely imaginary. Furthermore, examining expressions (11.26) and (11.27), we conclude that, in such a case, A_r and A_t are complex numbers.

Let us consider the transmitted wave. Returning to expression (11.14), we can write it as

$$\begin{aligned} u_2^t &= A_t \exp\left[i\omega\left(\left|\mathbf{p}^t\right|\cos\vartheta_t z + \left|\mathbf{p}^t\right|\sin\vartheta_t x - t\right)\right] \\ &= A_t \exp\left[i\omega\left(\left|\mathbf{p}^t\right| i\left|\cos\vartheta_t\right| z + \left|\mathbf{p}^t\right|\sin\vartheta_t x - t\right)\right] \\ &= A_t \exp\left(-\omega\left|\mathbf{p}^t\right|\left|\cos\vartheta_t\right| z\right)\exp\left[i\left(\left|\mathbf{p}^t\right|\sin\vartheta_t x - t\right)\right]. \end{aligned} \tag{11.28}$$

Expression (11.28) describes a wave that propagates in the positive x-direction and decays exponentially in the positive z-direction. Such a wave is called evanescent. In such a case there is no energy transmitted across the interface. Also, in such a case, the corresponding magnitude of A_r is equal to unity, as shown in Exercise 11.6.

Since for evanescent waves there is no energy transmitted across the interface, let us focus our attention on the reflected wave. For evanescent waves, A_r is a complex number that we can write as

$$A_r(\vartheta) = |A_r| \exp(i\varkappa), \tag{11.29}$$

where $|A_r|$ is the magnitude and $\varkappa$ is the angle in the complex plane. In view of expressions (11.14) and (11.29), as well as using the fact that $|A_r| = 1$, we can write the nonzero component of displacement of the reflected wave as

$$u_2^r = \exp(i\varkappa)\exp[i\omega(\mathbf{p}\cdot\mathbf{x}-t)] = \exp\{i[\varkappa+\omega(\mathbf{p}\cdot\mathbf{x}-t)]\}, \qquad (11.30)$$

where $\exp\{\cdot\}$ is the phase factor. Consequently, examining expression (11.30) and following the sign convention used for the phase factor in expression (11.14), we see that if $\varkappa > 0$, the reflected wave is phase-delayed relative to the incident wave. This is the consequence of the fact that positive $\varkappa$ results in the phase factor being evaluated at an earlier time. In other words, $\exp[i(\varkappa-\omega t)]$ lags $\exp(-i\omega t)$ in time. Similarly, if $\varkappa < 0$, the reflected wave is phase-advanced.[3]

Expressions in terms of incidence phase angle

As shown in Section 11.1, the incidence and transmission angles can be expressed in terms of one another. Consequently, we wish to state expressions (11.26) and (11.27) in terms of the phase angle of incidence only.

Recall conserved quantity (11.3). Let the phase velocity be given by expression (10.27), and the phase angle be stated by expression (10.30). Thus, we can write

$$\frac{\sin\vartheta_i}{\sqrt{\dfrac{C_{66}^I\sin^2\vartheta_i+C_{44}^I\cos^2\vartheta_i}{\rho_1}}} = \frac{\sin\vartheta_t}{\sqrt{\dfrac{C_{66}^{II}\sin^2\vartheta_t+C_{44}^{II}\cos^2\vartheta_t}{\rho_2}}}. \qquad (11.31)$$

Solving equation (11.31) for the angle of transmission, yields

$$\vartheta_t = \arcsin\sqrt{\frac{\rho_1 C_{44}^{II}\sin^2\vartheta_i}{\left[\rho_2\left(C_{66}^I-C_{44}^I\right)-\rho_1\left(C_{66}^{II}-C_{44}^{II}\right)\right]\sin^2\vartheta_i+\rho_2 C_{44}^I}}. \qquad (11.32)$$

Hence, by inserting expression (11.32) into expressions (11.26) and (11.27), we can state the latter expressions in terms of the phase angle of incidence only.

Expressions in terms of incidence ray angle

It is often convenient to state expressions (11.26) and (11.27) in terms of the ray angle of incidence, rather than the phase angle of incidence. Following

[3]Readers interested in phase shifts might refer to Aki, K., and Richards, P.G., (2002) Quantitative seismology (2*nd* edition): University Science Books, pp. 149 – 157.

equation (10.47), we can express the phase angle in terms of the ray angle as

$$\vartheta_i = \arctan\left(\frac{C_{44}^I}{C_{66}^I}\tan\theta_i\right). \tag{11.33}$$

Consequently, by inserting expression (11.33) into expression (11.32), and inserting the resulting expression into expressions (11.26) and (11.27), we can state the latter expressions in terms of the ray angle of incidence only.

Closing remarks

The reflection-angle and transmission-angle expressions derived in this chapter result from the continuity of phase across the interface. Analogous expressions, resulting from the conserved quantity associated with Fermat's principle of stationary traveltime, are discussed in Chapter 14.

Herein, the reflection-amplitude and transmission-amplitude expressions are derived for SH waves in transversely isotropic continua. This formulation provides a convenient illustration of the derivation process resulting from the boundary conditions that imply the equality of the sum of displacements and the equality of the traction components across the interface. Such a formulation can also be used in more general cases.

Note, however, that the illustration using SH waves does not address the fact that, in general, in anisotropic continua, displacement direction is neither parallel nor orthogonal to the wavefront normal, as discussed in Section 10.2.2. This property would introduce additional complications that are not addressed in this chapter.

Our formulation of the reflection and transmission amplitudes is based on the plane-wave assumption. Considering a point source, the plane-wave assumption provides a good approximation to a general formulation for distant sources. Moreover, other wavefront shapes can be considered as a composition of plane waves. In other words, any wavefront can be decomposed into plane waves.[4]

If we wish to derive a more general formulation, numerous assumptions must be investigated. For instance, considering ray methods in transversely isotropic continua, SH waves are decoupled from the qP and qSV waves. In

[4] Readers interested in evaluation of the applicability of the plane-wave assumption and its extensions might refer to Grant, F.S., and West, G.F., (1965) Interpretation theory in applied geophysics: McGraw-Hill Book Co., Chapter 6.

general, in continua exhibiting different symmetries, all three waves are coupled. Also, for the interface considered in this chapter, the two transversely isotropic continua are oriented in such a way that their rotation-symmetry axes are normal to the interface. Furthermore, the boundary conditions used in this chapter are based on the assumption of the welded contact at the interface. Many of the above concerns are addressed in the existing literature.[5]

Exercises

Exercise 11.1 [6]*Show that expressions (11.5) and (11.10) are equivalent to one another.*

Solution 11.1 *Consider expression (11.10). In view of the symmetry of an ellipse, we know that $V_x = v_x$ and $V_z = v_z$. Hence, we can write*

$$p = \frac{1}{v_x\sqrt{\left(\frac{v_x}{v_z}\right)^2 \cot^2\theta + 1}}.$$

Recalling expression (9.38), we express the ray angle in terms of the phase

[5] Readers interested in the formulation of reflection and transmission coefficients for P, SV and SH waves at different boundary conditions might refer to Aki, K. and Richards, P.G., (2002) Quantitative seismology (2*nd* edition): University Science Books, pp. 128 – 149, and to Červený, V., (2001) Seismic ray theory: Cambridge University Press, pp. 477 – 505. The former reference also contains a convenient weak-inhomogeneity approximation.

Readers interested in a formulation involving qP and qSV waves in transversely isotropic continua might refer to Mavko, G., Mukerji, T., and Dvorkin, J., (1998) The rock physics handbook: Cambridge University Press, pp 65 – 70.

Readers interested in a formulation of reflection and transmission coefficients in anelastic continua might refer to Le, L.H.T., Krebes, E.S., and Quiroga-Goode, G.E., (1994) Synthetic seismograms for SH waves in anelastic transversely isotropic media: Geophys. J. Int, **116**, 598 – 604.

Readers interested in a formulation accounting for phenomena resulting from slip interfaces, including interfaces between two identical continua, might refer to Schoenberg, M., (1980) Elastic wave behaviour across linear slip interfaces: J. Acoust. Soc. Am., **68** (5), 1516 – 1521.

[6] Also see Section 14.6

angle to obtain

$$\mathfrak{p} = \frac{1}{v_x\sqrt{\left(\frac{v_x}{v_z}\right)^2 \frac{1}{\tan^2\left\{\arctan\left[\left(\frac{v_x}{v_z}\right)^2 \tan\vartheta\right]\right\}} + 1}}.$$

Using trigonometric identities, we get

$$\mathfrak{p} = \frac{1}{v_x\sqrt{\left(\frac{v_x}{v_z}\right)^2 \frac{1}{\left(\frac{v_x}{v_z}\right)^4 \tan^2\vartheta} + 1}} = \frac{1}{\sqrt{v_z^2\cot^2\vartheta + v_x^2}}.$$

Multiplying both numerator and denominator by $\sin\vartheta$*, we obtain*

$$\mathfrak{p} = \frac{\sin\vartheta}{\sqrt{v_z^2\cos^2\vartheta + v_x^2\sin^2\vartheta}}.$$

In view of expression (6.60), we can immediately write

$$\mathfrak{p} = \frac{\sin\vartheta}{v(\vartheta)},$$

which is expression (11.5), as required.

Exercise 11.2 *Show that in isotropic continua, expressions (11.5) and (11.10) are identical to one another.*

Solution 11.2 *Consider expression (11.10). In isotropic continua,* $V := V_x = V_z$*. Hence, we can write*

$$\mathfrak{p} = \frac{1}{V\sqrt{\cot^2\theta + 1}}.$$

For isotropic continua, the magnitudes of the phase and ray velocities coincide, namely, $V = v$*. Also, the phase and ray angles coincide, namely,* $\theta = \vartheta$*. Thus, invoking trigonometric identities, we obtain*

$$\mathfrak{p} = \frac{\sin\vartheta}{v},$$

which is the isotropic form of expression (11.5), as required.

Exercise 11.3 *Following expression (11.26), state the expressions for the reflection and transmission amplitudes for isotropic continua in terms of mass density, ρ, and velocity, v.*

Solution 11.3 *In view of matrices (5.62) and (5.64), we let $\mu := C_{44} = C_{66}$ and write*

$$A_r(\vartheta) = \frac{\sqrt{\rho_1\mu_1}\cos\vartheta_i - \sqrt{\rho_2\mu_2}\cos\vartheta_t}{\sqrt{\rho_1\mu_1}\cos\vartheta_i + \sqrt{\rho_2\mu_2}\cos\vartheta_t},$$

and

$$A_t(\vartheta) = \frac{2\sqrt{\rho_1\mu_1}\cos\vartheta_i}{\sqrt{\rho_1\mu_1}\cos\vartheta_i + \sqrt{\rho_2\mu_2}\cos\vartheta_t},$$

for the reflection and transmission amplitudes, respectively. In view of $v = \sqrt{\mu/\rho}$, we can restate these expressions as

$$A_r(\vartheta) = \frac{\rho_1\sqrt{\frac{\mu_1}{\rho_1}}\cos\vartheta_i - \rho_2\sqrt{\frac{\mu_2}{\rho_2}}\cos\vartheta_t}{\rho_1\sqrt{\frac{\mu_1}{\rho_1}}\cos\vartheta_i + \rho_2\sqrt{\frac{\mu_2}{\rho_2}}\cos\vartheta_t} = \frac{\rho_1 v_1\cos\vartheta_i - \rho_2 v_2\cos\vartheta_t}{\rho_1 v_1\cos\vartheta_i + \rho_2 v_2\cos\vartheta_t},$$

and

$$A_t(\vartheta) = \frac{2\rho_1\sqrt{\frac{\mu_1}{\rho_1}}\cos\vartheta_i}{\rho_1\sqrt{\frac{\mu_1}{\rho_1}}\cos\vartheta_i + \rho_2\sqrt{\frac{\mu_2}{\rho_2}}\cos\vartheta_t} = \frac{2\rho_1 v_1\cos\vartheta_i}{\rho_1 v_1\cos\vartheta_i + \rho_2 v_2\cos\vartheta_t}.$$

Following Snell's law, namely, $\vartheta_t = \arcsin\left[(v_2/v_1)\sin\vartheta_i\right]$, we can express both A_r and A_t in terms of the angle of incidence, ϑ_i.

Exercise 11.4 *Using expressions (11.26) and (11.27), state the expressions for the reflection and transmission amplitudes for normal incidence in terms of mass density, ρ, and velocity, $v_{SH}(0)$.*

Solution 11.4 *Consider expressions (11.26) and (11.27). Letting $\vartheta_i = \vartheta_t = 0$, we obtain*

$$A_r(0) = \frac{\sqrt{\rho_1 C_{44}^I} - \sqrt{\rho_2 C_{44}^{II}}}{\sqrt{\rho_1 C_{44}^I} + \sqrt{\rho_2 C_{44}^{II}}},$$

and

$$A_t(0) = \frac{2\sqrt{\rho_1 C_{44}^I}}{\sqrt{\rho_1 C_{44}^I} + \sqrt{\rho_2 C_{44}^{II}}},$$

for the reflection and transmission amplitudes, respectively. In view of expressions (10.27) and (10.30), we obtain $v := v_{SH}(0) = \sqrt{C_{44}/\rho}$ *and, hence, we can restate the above expressions as*

$$A_r(0) = \frac{\rho_1\sqrt{\dfrac{C_{44}^I}{\rho_1}} - \rho_2\sqrt{\dfrac{C_{44}^{II}}{\rho_2}}}{\rho_1\sqrt{\dfrac{C_{44}^I}{\rho_1}} + \rho_2\sqrt{\dfrac{C_{44}^{II}}{\rho_2}}} = \frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2}, \tag{11.34}$$

and

$$A_t(0) = \frac{2\rho_1\sqrt{\dfrac{C_{44}^I}{\rho_1}}}{\rho_1\sqrt{\dfrac{C_{44}^I}{\rho_1}} + \rho_2\sqrt{\dfrac{C_{44}^{II}}{\rho_2}}} = \frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2}, \tag{11.35}$$

where $v_1 \equiv v_{SH}(0)$ *in the medium of incidence and* $v_2 \equiv v_{SH}(0)$ *in the medium of transmission.*

Exercise 11.5 *Consider expression*

$$\langle \mathfrak{E} \rangle = \frac{1}{2}\rho v \omega^2 A^2, \tag{11.36}$$

where $\langle \mathfrak{E} \rangle$ *is the average energy density carried by the wave and* ω *is its angular frequency. Using the expressions for the normal-incidence reflection and transmission amplitudes, derived in Exercise 11.4, show that the energy is conserved.*

Solution 11.5 *The balance of energy states that the energy carried by the incident wave must be equal to the sum of the energies carried by the reflected and transmitted waves, namely,*

$$\langle \mathfrak{E}_i \rangle = \langle \mathfrak{E}_r \rangle + \langle \mathfrak{E}_t \rangle .$$

Considering monochromatic waves and normalizing incident-wave amplitude to unity, in accordance with expression (11.36), we obtain

$$\frac{1}{2}\rho_1 v_1 = \frac{1}{2}\rho_1 v_1 A_r^2 + \frac{1}{2}\rho_2 v_2 A_t^2,$$

which can be rewritten as

$$1 = A_r^2 + \frac{\rho_2 v_2}{\rho_1 v_1} A_t^2. \tag{11.37}$$

Inserting expressions (11.34) and (11.35) into expression (11.37), we get

$$1 = \left(\frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2}\right)^2 + \frac{\rho_2 v_2}{\rho_1 v_1}\left(\frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2}\right)^2$$
$$= \left(\frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2}\right)^2 + \frac{4\rho_1 v_1 \rho_2 v_2}{(\rho_1 v_1 + \rho_2 v_2)^2} = 1,$$

as required.

Exercise 11.6 [7]*Show that if* $\sin\vartheta_t > 1$, *the magnitude of* A_r *is equal to unity.*

Solution 11.6 *If* $\sin\vartheta_t > 1$, *then* $\cos\vartheta_t$ *is a pure imaginary number. In that case, expression (11.26) is of the form*

$$A_r = \frac{a - bi}{a + bi}.$$

The magnitude is given by

$$|A_r| = \sqrt{A_r A_r^*},$$

where

$$A_r^* := \frac{a + bi}{a - bi}$$

is the complex conjugate. Therefore,

$$|A_r| = \sqrt{\frac{a - bi}{a + bi}\frac{a + bi}{a - bi}} = 1,$$

as required.

[7]See also Section 11.2.2

Part III

Variational formulation of rays

This Page Intentionally Left Blank

Introduction to Part III

> What you do is to invent various curves, and calculate on each curve a certain quantity. If you calculate this quantity for one route, and then for another, you will get a different number for each route. There is one route which gives the least possible number, however, and that is the route that the particle in nature actually takes. We are now describing the actual motion by saying something about the whole curve. We have lost the idea of causality, that the particle feels the pull and moves in accordance with it. Instead of that, in some grand fashion it smells all the curves, all the possibilities, and decides which one to take by choosing that for which our quantity is least.[8]
>
> *Richard Feynman (1967) The Character of Physical Law*

The fundamental formulation of ray theory was presented in *Part II*. This theory is based on the high-frequency approximation to Cauchy's equations of motion in anisotropic inhomogeneous continua and results in Hamilton's ray equations. Also, in *Part II*, we used Legendre's transformation of Hamilton's ray equations to obtain Lagrange's ray equations. Thus, within the limitations of this transformation, we have two equivalent forms of the ray equations.

In *Part III*, we will study ray theory in the context of Lagrange's ray equations. We will show that they are the stationarity conditions of the calculus of variations. Hence, we will show that rays, wavefronts and traveltimes can be studied by invoking the concept of stationary traveltime.

We will use the calculus of variations in search of the stationarity condition for a definite integral that describes the traveltime of the signal between

[8]Readers interested in the philosophical aspects of this statement, in the context of analytical mechanics, might refer to Toretti, R., (1999) The philosophy of physics: Cambridge University Press, p. 92.

a source and a receiver. Since, in the variational approach to ray theory, either time or distance constitutes the single variable, the stationarity conditions are a system of ordinary differential equations. Consequently, the variational formulation is an elegant method to describe rays, wavefronts and traveltimes. Also, an intuitive concept of stationarity is a fruitful starting point for many investigations.

The first scientific statement of a variational principle was formulated in optics by Pierre de Fermat in 1657.[9] In its original formulation, this principle was referred to as the principle of least time. Following Fermat's principle, the principle of least action in mechanics was proposed in the first half of the eighteenth century by Pierre-Louis Moreau de Maupertuis and, then, rigorously stated by William Rowan Hamilton in 1835.[10]

The theory of the calculus of variations originated with the statement of Johannes Bernoulli, who, in 1696, posed the problem to determine the shape of a wire along which a bead might slide in the shortest possible time. While this problem might have initially appeared quite particular, it led to an important general theory. In 1900, David Hilbert delivered a talk on "Mathematical Problems" during which he made the following statement.

> The mathematicians of past centuries were accustomed to devote themselves to the solution of difficult individual problems with passionate zeal. They knew the value of difficult problems. I remind you only of the 'problem of the line of quickest descent', proposed by Johannes Bernoulli. [...] It is an error to believe that rigour in the proof is the enemy of simplicity. On the contrary, we find it confirmed by numerous examples that the rigorous method is at the same time simpler and the more easily comprehended. [...] the most striking example of my statement is the calculus of variations.

[9] Interested readers might refer to Born, M., and Wolf, E., (1999) Principles of optics (*7th* edition): Cambridge University Press, p. xxvi.

[10] Readers interested in formal relations between the classical-mechanics principle of stationary action and the ray-theory principle of stationary traveltime as well as their relation to quantuum mechanics might refer to Goldstein, H., (1950/1980) Classical mechanics: Addison-Wesley Publishing Co., pp. 365 – 371 and pp. 484 – 492.

Chapter 12

Euler's equations

> For since the shape of the whole universe is most perfect and, in fact, designed by the wisest creator, nothing at all takes place in the universe in which a rule of maximum or minimum does not appear.
>
> *Leonhard Euler (1744) Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti*[1]

Preliminary remarks

In Chapter 9, we derived Lagrange's ray equations. These equations are variational equations and, hence, allow us to consider ray theory in the context of the calculus of variations.

We begin this chapter with a brief discussion of stationarity of a definite integral and the derivation of the stationarity condition of the calculus of variations, namely, Euler's equation. This is followed by formulations of the generalized and special forms of Euler's equations, which are again used in Chapters 13 and 14. We conclude this chapter by relating Euler's equations to Lagrange's ray equations.

This chapter is intended to give a brief introduction to the calculus of variations for readers who are not familiar with this subject. Otherwise, it can be omitted without affecting the study of subsequent chapters.

[1]Method of finding curved lines enjoying the maximum and minimum property; or the solution of the isoperimetric problem understood in the broadest sense

12.1 Mathematical background

The calculus of variations is the study of methods to obtain stationary values of definite integrals. These values depend on functions that compose a given integrand. In other words, the domain of a definite integral is a set of functions. An integral operates on a set of functions and we seek a particular function that gives a stationary value of this integral. Analogously, in differential calculus, a function operates on a set of points and we seek a particular point that gives a stationary value of this function.

In differential calculus, the condition for stationarity of a function is the vanishing of its first derivative. We wish to formulate an analogous condition for stationarity of a definite integral.[2]

Herein, we focus our study on two-dimensional problems that are contained in the xz-plane. In this study, we require stationary values of an integral expressed as

$$I = \int_a^b F\left(z(x), z'(x); x\right) \mathrm{d}x. \tag{12.1}$$

Thus, we seek function $z(x)$ that makes integral (12.1) stationary. Assuming that $z(x)$ is continuous and smooth, we can view it as a curve in the xz-plane.

Integrand F contains three arguments, namely, $z(x)$, $z'(x) \equiv \mathrm{d}z/\mathrm{d}x$ and x. In formulating the condition of stationarity, we consider these three arguments as independent.

Note that to avoid any confusion, we could choose to write

$$F\left(z(x), z'(x); x\right) \equiv F(\xi_1, \xi_2, \xi_3). \tag{12.2}$$

However, we will not introduce these additional symbols.

We need, however, a new operator symbol. In the search for stationarity, Lagrange introduced a special symbol denoted by δ, which refers to the variations of curve $z(x)$. In other words, among all the variations of $z(x)$ between the fixed end-points a and b, we search for a curve that renders the value of a given integral stationary. This curve is a solution of the variational problem. Hence, the problem of looking for such a curve is symbolically stated as $\delta \int_a^b F \mathrm{d}x = 0$.

[2]Readers interested in a definition of stationarity of a definite integral might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2*nd* edition): Springer-Verlag, p. 57.

Note the distinction between the variational and differential operators. Symbol $\delta z(x)$ refers to a variation from curve to curve for a given x, whereas symbol $\mathrm{d}z(x)$ refers to a differential change along a given curve for a change in x.[3]

Note that, in this chapter and in Chapter 14, we restrict our study to curves in the form $z = z(x)$, rather than in the parametric form, $\mathbf{x}(t)$, used to formulate Hamilton's and Lagrange's ray equations in Chapters 8 and 9, respectively.[4]

The condition of stationarity of integral (12.1) was derived by Euler in 1744.[5] This condition is discussed in the next section.

12.2 Formulation of Euler's equation

In this section, we derive the stationarity condition for integral (12.1). In other words, among all continuously differentiable functions $z(x)$ that satisfy the boundary conditions at $z(a)$ and $z(b)$, we establish the condition to choose a function that renders integral (12.1) stationary. This stationarity condition is stated by the following theorem.

Theorem 12.1 *Function $z(x)$ with the continuous first derivative on interval $[a,b]$ yields a stationary value of integral (12.1), namely,*

$$I = \int_a^b F\left(z(x), z'(x); x\right) \mathrm{d}x, \tag{12.3}$$

in the class of functions with boundary conditions $z(a) = z_a$ and $z(b) = z_b$, if equation

$$\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) = 0. \tag{12.4}$$

[3]Readers interested in the δ operator might refer to Lanczos, C., (1949/1986) The variational principles of mechanics: Dover, pp. 38 – 40, and to Ewing, M.G., (1969/1985) Calculus of variations with applications: Dover, pp. 86 – 88.

[4]Readers interested in the relation between the explicit and parametic formulations of the Euler equations might refer to Ewing, M.G., (1969/1985) Calculus of variations with applications: Dover, pp. 140 – 141, to Gelfand, I.M., and Fomin, S.V., (1963/2000) Calculus of variations: Dover, pp 38 – 42, to Sagan, H., (1969/1992) Introduction to the calculus of variations: Dover, pp. 197 – 202, or to Weinstock, R., (1952/1974) Calculus of variations with applications to physics and engineering: Dover, pp. 34 – 36.

[5]It is also common to refer to this equation as the Euler-Lagrange equation. Readers interested in the history of this equation might refer to Marsden, J.E., and Ratiu, T.S., (1999) Introduction to mechanics and symmetry: A basic exposition of classical mechanical systems (2*nd* edition): Springer-Verlag, pp. 231 – 234.

is satisfied.[6]

We refer to equation (12.4) as Euler's equation. Euler's equation (12.4) is a second-order ordinary differential equation.[7]

To see the connection between integral (12.3) and its stationarity condition, given by Euler's equation (12.4), consider the following heuristic argument.

Replace the integral by a finite sum of subdivisions given by x_0, x_1,..., x_{n-1}, x_n, where the interval of integration $[a, b]$ is $[x_0, x_n]$. The subdivisions are assumed to be equally spaced and we denote this spacing by $\Delta x = (b-a)/n$. A discrete expression approximating integral (12.3) can be written as

$$S_n = \sum_{i=0}^{n-1} F\left(z_{i+1}, z'_{i+1}; x_{i+1}\right) \Delta x,$$

where

$$z'_{i+1} := \frac{z_{i+1} - z_i}{\Delta x}. \tag{12.5}$$

Here S_n is viewed as a function of the $n-1$ variables, $z_1, \ldots, z_{n-1}$.

Note that z_0 and z_n are not included as variables because they are fixed by the boundary conditions, namely, $z_0 = z_a$ and $z_n = z_b$.

To find the stationary value of S_n, we find the stationary points for $n-1$ variables. This is equivalent to setting to zero all partial derivatives of S_n with respect to z_i. In other words, the stationarity condition is

$$\frac{\partial S_n}{\partial z_i} = 0, \qquad i \in \{1, \ldots, n-1\}. \tag{12.6}$$

In view of expression (12.5), in the sum S_n, for any given $i \in (1, \ldots, n-1)$, there are only two consecutive terms that explicitly contain a given z_i, namely,

$$F\left(z_i, z'_i; x_i\right) \Delta x + F\left(z_{i+1}, z'_{i+1}; x_{i+1}\right) \Delta x. \tag{12.7}$$

[6]Readers interested in a rigorous proof of Theorem 12.1 might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2*nd* edition): Springer-Verlag, pp. 57 - 58.

[7]Readers interested in a thorough study of Euler's equations might refer to Courant, R., and Hilbert, D., (1924/1989) Methods of mathematical physics: John Wiley & Sons, Vol. I, pp. 183 - 206, and to Morse P.M., and Feshbach H., (1953) Methods of theoretical physics: McGraw-Hill, Inc., Part I, pp. 276 - 280.

Applying stationarity condition (12.6), we take the derivative of expression (12.7) with respect to z_i and obtain

$$\left[\frac{\partial F}{\partial z}\left(z_i, z_i'; x_i\right) + \frac{\partial F}{\partial z'}\left(z_i, z_i'; x_i\right)\frac{\partial z_i'}{\partial z_i}\right]\Delta x + \left[\frac{\partial F}{\partial z'}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\frac{\partial z_{i+1}'}{\partial z_i}\right]\Delta x = 0, \quad (12.8)$$

where $i \in \{1, \ldots, n-1\}$.

Note that an analogous approach can be followed by viewing S_n as a function of x_i and, hence, by setting all partial derivatives with respect to x_i to zero. As shown in Exercise 12.2, by following this approach, we obtain Beltrami's identity (12.10).

In view of equation (12.8) and recalling expression (12.5), we have

$$\frac{\partial z_i'}{\partial z_i} = \frac{\partial}{\partial z_i}\frac{z_i - z_{i-1}}{\Delta x} = \frac{1}{\Delta x},$$

and

$$\frac{\partial z_{i+1}'}{\partial z_i} = \frac{\partial}{\partial z_i}\frac{z_{i+1} - z_i}{\Delta x} = -\frac{1}{\Delta x}.$$

Hence, equation (12.8) becomes

$$\left[\frac{\partial F}{\partial z}\left(z_i, z_i'; x_i\right) + \frac{\partial F}{\partial z'}\left(z_i, z_i'; x_i\right)\frac{1}{\Delta x}\right]\Delta x - \left[\frac{\partial F}{\partial z'}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\frac{1}{\Delta x}\right]\Delta x = 0,$$

where $i \in \{1, \ldots, n-1\}$. This equation can be rearranged to give

$$\frac{\partial F}{\partial z}\left(z_i, z_i'; x_i\right) - \frac{1}{\Delta x}\left[\frac{\partial F}{\partial z'}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right) - \frac{\partial F}{\partial z'}\left(z_i, z_i'; x_i\right)\right] = 0, \quad (12.9)$$

where $i \in \{1, \ldots, n-1\}$.

We now assume that as $\Delta x \to 0$ and $x_i \to x \in [a, b]$, z_i approaches $z(x)$ and $z_i' = (z_i - z_{i-1})/\Delta x$ approaches $z'(x)$. Then, equation (12.9) becomes

$$\frac{\partial F}{\partial z}\left(z(x), z'(x); x\right) - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial z'}\left(z(x), z'(x); x\right)\right] = 0,$$

which is Euler's equation (12.4), as required.

12.3 Beltrami's identity

A convenient form of Euler's equation (12.4) is Beltrami's identity, discussed in Section 9.1.2. In the two-dimensional case, where we look for the $z(x)$ that is a solution of Euler's equation (12.4), Beltrami's identity is equivalent to that equation. Hence, we can write

$$\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) = 0 = \frac{\partial F}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x}\left(z'\frac{\partial F}{\partial z'} - F\right). \tag{12.10}$$

In general, Beltrami's identity is not equivalent to the corresponding Euler's equations or Lagrange's ray equations. In Chapter 9, for instance, a single expression of Beltrami's identity (9.9) refers to three equations of system (9.8) and, hence, by itself, cannot give a unique solution of system (9.8). A verification of Beltrami's identity (12.10) and its derivation are shown in Exercises 12.1 and 12.2, respectively.

Beltrami's identity is particularly useful when the integrand does not explicitly depend on x, namely, $F = F(z, z')$. In such a case, the first term on the right-hand side of equation (12.10) vanishes. Important consequences of this simplification are discussed in Chapter 14 in the context of ray parameters.

12.4 Generalizations of Euler's equation

Introductory comments

Integral (12.1) depends on a single variable, x, on a single function, $z(x)$, and on its first derivative, $z'(x)$. In mathematical considerations of physically motivated problems, a given integral whose stationary value we seek can also depend on several variables, on several functions and on higher-order derivatives. Such formulations result in stationarity conditions that are second-order partial differential equations, systems of second-order ordinary differential equations and higher-order ordinary differential equations, respectively.

Although a given problem can depend on all of the above quantities, each of the three cases is described separately below.

12.4.1 Case of several variables

Let us consider an integral that contains a single function of several variables. To begin, we consider an integral whose integrand contains a function of two variables, namely,

$$I = \int_{a_y}^{b_y} \int_{a_x}^{b_x} F\left(z\left(x, y\right), z_x, z_y; x, y\right) \mathrm{d}x \, \mathrm{d}y,$$

where $z_x := \partial z / \partial x$ and $z_y := \partial z / \partial y$. Thus, within given constraints on a boundary, we look for a smooth surface, $z\left(x, y\right)$, that renders I stationary. In this case, Euler's equation becomes

$$\frac{\partial F}{\partial z} - \left[\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial z_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial z_y}\right)\right] = 0, \tag{12.11}$$

which is a second-order partial differential equation. The generalization for n variables follows the same pattern, thereby giving

$$\frac{\partial F}{\partial z} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial F}{\partial z_{x_i}}\right) = 0, \tag{12.12}$$

where $z = z\left(x_1, \ldots, x_n\right)$ and $z_{x_i} := \partial z / \partial x_i$, with $i \in \{1, \ldots, n\}$.

Problems involving multiple integrals were considered by Lagrange in his papers dating from 1760 – 1762. A physical example of a double integral is discussed in Section 13.2.3.

12.4.2 Case of several functions

Let us consider an integral that contains several single-variable functions and their first derivatives. To begin, we consider an integral whose integrand contains two functions, namely,

$$I = \int_a^b F\left(y\left(x\right), y'\left(x\right), z\left(x\right), z'\left(x\right); x\right) \, \mathrm{d}x.$$

Thus, we look for smooth curves $y\left(x\right)$ and $z\left(x\right)$ that render I stationary, subject to constraints

$$\begin{cases} y\left(a\right) = a_1 \\ z\left(a\right) = a_2 \\ y\left(b\right) = b_1 \\ z\left(b\right) = b_2 \end{cases}, \tag{12.13}$$

where a_i and b_i are constants. In this case, Euler's equations become a system of second-order ordinary differential equations

$$\left\{\begin{array}{l}\dfrac{\partial F}{\partial y}-\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{\partial F}{\partial y'}\right)=0\\[2ex]\dfrac{\partial F}{\partial z}-\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{\partial F}{\partial z'}\right)=0\end{array}\right. .$$

The generalization for n functions follows the same pattern thereby giving a system of n equations,

$$\frac{\partial F}{\partial \zeta_i}-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial \zeta_i'}\right)=0, \qquad i\in\{1,\ldots,n\}, \tag{12.14}$$

where $\zeta_i=\zeta_i(x)$ and $\zeta_i'=\mathrm{d}\zeta_i(x)/\mathrm{d}x$.

12.4.3 Higher-order derivatives

Let us consider an integral whose integrand contains higher-order derivatives of a single-variable function. To begin, we consider an integral whose integrand contains both the first and second derivatives, namely,

$$I=\int_a^b F\left(z(x),z'(x),z''(x);x\right)\mathrm{d}x.$$

Thus, we look for a smooth curve $z(x)$ that renders I stationary, subject to constraints

$$\left\{\begin{array}{l}z(a)=a_1\\z'(a)=a_2\\z(b)=b_1\\z'(b)=b_2\end{array}\right. , \tag{12.15}$$

where a_i and b_i are constants. In this case, Euler's equation becomes

$$\frac{\partial F}{\partial z}-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right)+\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial F}{\partial z''}\right)=0.$$

This is a fourth-order ordinary differential equation. The generalization for nth-order derivatives follows the same pattern to yield

$$\frac{\partial F}{\partial z}-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right)+\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial F}{\partial z''}\right)+\ldots+(-1)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\partial F}{\partial z^{(n)}}\right)=0,$$

which is an ordinary differential equation of order $2n$.

12.5 Special cases of Euler's equation

Introductory comments

There are cases where, due to the explicit absence of certain arguments or to the particular form of integral (12.1), Euler's equation (12.4) becomes a simpler equation.

Note that in evaluating partial derivatives, only explicit appearances of the variable of differentiation are taken into account. For instance, if we differentiate $F(z(x))$ with respect to z, namely, $\partial F/\partial z$, no allowance is made for the fact that a change in x also results in a change of z. Following expression (12.2), we could choose to write such a differentiation as $\partial F(\xi_1)/\partial \xi_1$ and, thus, at the expense of introducing an additional symbol, avoid any confusion.

12.5.1 Independence of z

Let us consider an integrand that is explicitly independent of z, namely, $F = F(z'; x)$. We see that Euler's equation (12.4) is reduced to

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) = 0.$$

The vanishing of the total derivative implies that

$$\frac{\partial F}{\partial z'} = C_1, \tag{12.16}$$

where C_1 denotes a constant. Thus, $z(x)$ is obtained as a solution of first-order ordinary differential equation (12.16).

12.5.2 Independence of x and z

Let us consider an integrand that is explicitly independent of both x and z. In other words, it is only dependent on z', namely, $F = F(z')$. Since z' is the only variable, we can immediately rewrite equation (12.16) as

$$\frac{\mathrm{d}F(z')}{\mathrm{d}z'} = C_1. \tag{12.17}$$

Denoting $\mathrm{d}F(z')/\mathrm{d}z'$ as $\mathrm{f}(z')$, we can write equation (12.17) as

$$\mathrm{f}(z') = C_1.$$

Assuming that $\mathrm{d}f/\mathrm{d}z' \neq 0$, we can consider inverse function f^{-1}. Thus, we can write

$$z' = \mathrm{f}^{-1}(C_1).$$

Recalling that $z' \equiv \mathrm{d}z/\mathrm{d}x$ and denoting $\mathrm{f}^{-1}(C_1) = C_2$, we can write

$$\frac{\mathrm{d}z}{\mathrm{d}x} = C_2. \tag{12.18}$$

This is a first-order ordinary differential equation, whose solution,

$$z = C_2 x + C_3,$$

is obtained directly by integration.

Thus, finding the curve which gives a stationary value of $\int_a^b F(z')\mathrm{d}x$ consists of writing the equation of a straight line passing through points $[a, z(a)]$ and $[b, z(b)]$.

In a seismological context, this implies that in homogeneous continua, whether the continua be isotropic or anisotropic, if the properties do not depend on position, rays are straight.

12.5.3 Independence of x

Let us consider an integrand that is explicitly independent of x, namely, $F = F(z, z')$. Using Beltrami's identity (12.10), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(z'\frac{\partial F}{\partial z'} - F\right) = 0.$$

The vanishing of the total derivative implies that

$$z'\frac{\partial F}{\partial z'} - F = C, \tag{12.19}$$

where C denotes a constant. Thus, $z(x)$ is obtained as a solution of first-order ordinary differential equation (12.19).

In a seismological context, the case where the traveltime integral is independent of x implies that the continuum is homogeneous along the x-axis — a case commonly encountered in layered media. In such media, the constant in expression (12.19) is a ray parameter, discussed in Chapter 14.

12.5.4 Total derivative

Let integrand $F(x, z, z')$ be a total derivative of function $f(x, z)$ with respect to x, namely,

$$F\left(z, z'; x\right) = \frac{\mathrm{d}f(x, z)}{\mathrm{d}x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} z'. \tag{12.20}$$

Consider the left-hand side of Euler's equation (12.4). Inserting function (12.20), we obtain

$$\begin{aligned} \frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) &= \frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2} z' - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial z}\right) \\ &= \frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2} z' - \left(\frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2} z'\right) \\ &= 0. \end{aligned}$$

Thus, equation (12.4) is identically satisfied. Consequently, if F is a total derivative, Euler's equation (12.4) is satisfied by any $z(x)$. In other words, if a variational problem involves the integral of a total differential, namely,

$$\delta \int_a^b \mathrm{d}f(x, z) = 0,$$

the value of the integral is independent of the integration path and depends only on the limits of integration.

Note that, considering a fixed-ends variational problem, we can add to the integrand a term that is a total derivative without changing the solution of Euler's equations, as shown in Exercise 12.3.[8] Considering such cases, we note that, although a solution curve is not affected by this addition, the value of the integral is changed. For instance, identical rays can result in distinct traveltimes, depending on the properties of the continuum.

12.5.5 Function of x and z

Euler's equation

In physically motivated problems, we often encounter an integral given by $\int_a^b h(\mathbf{x})\mathrm{d}s$, which is an integral of function h, whose value depends on position $\mathbf{x}$ along the arclength element $\mathrm{d}s$. Such an integral represents a certain quantity measured along a trajectory that connects points a and b.

[8] In electromagnetic theory, this property is associated with the gauge invariance. Interested readers might refer to Morse P.M., and Feshbach H., (1953) Methods of theoretical physics: McGraw-Hill, Inc., pp. 210–212.

Considering the two-dimensional case and assuming that the trajectory can be expressed as $z = z(x)$, we can write such an integral as

$$\int_{a_x}^{b_x} h(x,z)\sqrt{1+\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2}\,\mathrm{d}x. \tag{12.21}$$

Thus, Euler's equation (12.4) becomes

$$\frac{\partial}{\partial z}\left[h(x,z)\sqrt{1+(z')^2}\right] - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial}{\partial z'}\left(h(x,z)\sqrt{1+(z')^2}\right)\right] = 0,$$

where $z' \equiv \mathrm{d}z/\mathrm{d}x$. Performing partial derivatives, we obtain

$$\frac{\partial h(x,z)}{\partial z}\sqrt{1+(z')^2} - \frac{\mathrm{d}}{\mathrm{d}x}\left[h(x,z)\frac{z'}{\sqrt{1+(z')^2}}\right] = 0.$$

Then, by the product rule, we get

$$\frac{\partial h(x,z)}{\partial z}\sqrt{1+(z')^2} - \left[\frac{\mathrm{d}h(x,z)}{\mathrm{d}x}\frac{z'}{\sqrt{1+(z')^2}} + h(x,z)\frac{\mathrm{d}}{\mathrm{d}x}\frac{z'}{\sqrt{1+(z')^2}}\right] = 0.$$

Letting $h := h(x,z)$ and using the quotient and chain rules, we obtain

$$\frac{\partial h}{\partial z}\sqrt{1+(z')^2} - \left[\left(\frac{\partial h}{\partial x}+\frac{\partial h}{\partial z}z'\right)\frac{z'}{\sqrt{1+(z')^2}} + h\frac{z''\sqrt{1+(z')^2} - z'\frac{z'z''}{\sqrt{1+(z')^2}}}{1+(z')^2}\right] = 0.$$

An algebraic simplification leads to

$$\frac{\partial h}{\partial z}\sqrt{1+(z')^2} - \frac{\partial h}{\partial x}\frac{z'}{\sqrt{1+(z')^2}} - \frac{\partial h}{\partial z}\frac{(z')^2}{\sqrt{1+(z')^2}} - h\frac{z''}{\left[1+(z')^2\right]^{\frac{3}{2}}} = 0.$$

Rearranging the common factor, we obtain

$$\frac{1}{\sqrt{1+(z')^2}}\left[\frac{\partial h}{\partial z} + \frac{\partial h}{\partial z}(z')^2 - \frac{\partial h}{\partial x}z' - \frac{\partial h}{\partial z}(z')^2 - h\frac{z''}{1+(z')^2}\right] = 0.$$

The cancellation of identical terms results in

$$\frac{1}{\sqrt{1+(z')^2}}\left[\frac{\partial h}{\partial z}-\frac{\partial h}{\partial x}z'-h\frac{z''}{1+(z')^2}\right]=0.$$

Since the factor in front of the brackets is never zero, Euler's equation becomes

$$\frac{\partial h}{\partial z}-\frac{\partial h}{\partial x}\frac{\mathrm{d}z}{\mathrm{d}x}-h\frac{\frac{\mathrm{d}^2 z}{\mathrm{d}x^2}}{1+\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2}=0. \tag{12.22}$$

To study equation (12.22) in the context of ray theory, let function $h(x,z)$ describe slowness in an isotropic inhomogeneous continuum. Hence, letting the velocity function be $v(x,z)=1/h(x,z)$ and rearranging equation (12.22), we obtain

$$v\frac{\mathrm{d}^2 z}{\mathrm{d}x^2}-\frac{\partial v}{\partial x}\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^3+\frac{\partial v}{\partial z}\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2-\frac{\partial v}{\partial x}\frac{\mathrm{d}z}{\mathrm{d}x}+\frac{\partial v}{\partial z}=0, \tag{12.23}$$

where we assume $v(x,z)\neq 0$. In such a case, integral (12.21), which can be rewritten as

$$\int\limits_{a_x}^{b_x}\frac{\sqrt{1+(z')^2}}{v(x,z)}\,\mathrm{d}x, \tag{12.24}$$

represents the traveltime between two points. Thus, a solution of equation (12.23) is a ray, $z(x)$, satisfying Fermat's principle of stationary traveltime.

If $v(x,z)$ is given by a constant, integral (12.24) is explicitly independent of x and z, and it corresponds to the case discussed in Section 12.5.2. In such a case, equation (12.23) reduces to $\mathrm{d}^2z/\mathrm{d}x^2=0$, whose solutions are $z=C_2x+C_3$, where C_2 and C_3 are constants that depend on the limits of integration in integral (12.24).

Geometrical interpretation and physical meaning

Integral (12.21) has a simple geometrical interpretation. Let function h be a smooth and continuous function whose values are positive. Consider an orthonormal coordinate system, where $h(x,z)$ can be represented as a surface above the xz-plane. Let $z(x)$ be a smooth and continuous curve in the xz-plane that connects points a and b. Integral (12.21) is the surface area of a strip that is orthogonal to the xz-plane and whose edges are given by

$z(x)$ and the corresponding values of $h(x, z)$. This strip can be viewed as a fence that follows curve $z(x)$, and whose height, at any point, is determined by function $h(x, z)$.

A solution of equation (12.22), namely, $z = z(x)$, is the curve along which the area of the corresponding strip is stationary. Herein, given the geometry of the variational problem, the area of the strip that results from equation (12.22) is minimum, as illustrated in Exercise 12.4.

Traveltime is the product of slowness and distance travelled. If function h represents the slowness in an isotropic inhomogeneous continuum, the area of the strip represents the traveltime between the two points, which are given by the limits of integration. Hence, a solution of equation (12.23) is a trajectory along which the traveltime is stationary.

12.6 First integrals

Special cases of Euler's equation, which result from the absence of particular arguments in the integrand function, are called first integrals. This name originates in the period of mathematical history when many differential equations were solved by integration. The description shown in Section 12.5.2, where the integrand is explicitly independent of both x and z, exemplifies such an approach.

The term "first integral" implies that the order of the differential equation has been reduced by one, which is equivalent to the integration process. Formally, the meaning of first integral is described in the following definition.

Definition 12.1 *If an nth-order differential equation*

$$f\left(x, z, z', \ldots, z^{(n)}\right) = 0, \tag{12.25}$$

can be transformed to the equivalent form

$$\frac{\mathrm{d}}{\mathrm{d}x} g\left(x, z, z', \ldots, z^{(n-1)}\right) = 0,$$

we see that

$$g\left(x, z, z', \ldots, z^{(n-1)}\right) = C, \tag{12.26}$$

where C is a constant. Expression (12.26) is a "first integral" of equation (12.25).

Note that the fact that the integrand of a variational problem does not explicitly depend on a particular argument is equivalent to saying that this problem is invariant with respect to that argument. This invariance and the associated first integral are contained in Noether's theorem, published in Göttingen in 1918 in her paper entitled "Invariante Variationsprobleme".[9]

In the context of ray theory, we use the property that a first integral of a differential equation is a function that has a constant value along a solution curve. This constant is a ray parameter, which is discussed in Chapter 14.

12.7 Lagrange's ray equations as Euler's equations

To use the calculus of variations in the study of ray theory, we wish to show that Lagrange's ray equations (9.8) belong to the realm of Euler's equations. The parametric form of Euler's equation (12.4) corresponds to a system of two Euler's equations, namely,

$$\left\{\begin{array}{l} \dfrac{\partial G}{\partial x}-\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial G}{\partial \dot{x}}\right)=0 \\ \\ \dfrac{\partial G}{\partial z}-\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial G}{\partial \dot{z}}\right)=0 \end{array}\right., \tag{12.27}$$

where $G=G(x,z,\dot{x},\dot{z})$ with $\dot{x}:=\mathrm{d}x/\mathrm{d}t$ and $\dot{z}:=\mathrm{d}z/\mathrm{d}t$. A solution of system (12.27) is a curve in the xz-plane given by $[x(t),z(t)]$ that corresponds to variational problem

$$\delta\int G\mathrm{d}t=0.$$

To see the relation between G and F, which is stated in integral (12.1), we can write $\mathrm{d}t=\mathrm{d}x/\dot{x}$ and $z':=\mathrm{d}z/\mathrm{d}x$. Hence, $G(x,z,\dot{x},\dot{z})=F(z,\dot{z}/\dot{x},x)\,\dot{x}$.

This parametric formulation allows us to use Euler's equations for an n-dimensional space. In general, we can write a system of n Euler's equations, namely,

$$\frac{\partial G}{\partial x_i}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial \dot{x}_i}=0, \qquad i\in\{1,\ldots,n\}, \tag{12.28}$$

[9]Readers interested in rigorous derivations and proofs might refer to Gelfand, I.M., and Fomin, S.V., (1963/2000) Calculus of variations: Dover, pp. 79 – 83 and pp. 176 – 179, and to Goldstein, H., (1950/1980) Classical mechanics: Addison-Wesley Publishing Co., pp. 588 – 596.

Readers interested in variational aspects of Noether's theorem might refer to Lanczos, C., (1949/1986) The variational principles of mechanics: Dover, pp. 401 – 405.

whose solution is a curve in the $\mathbf{x}$-space given by $\mathbf{x}(t)$. Examining systems (9.8) and (12.28), we conclude that Lagrange's ray equations (9.8) possess the form of Euler's equations (12.28).

The fact that Euler's and Lagrange's equations have equivalent forms is the reason why equations of the form (9.8) and (12.28) are often referred to as the Euler-Lagrange equations. In this book, we use the term Euler's equations to refer to the mathematical condition of stationarity while we reserve the term Lagrange's equations to refer to those among Euler's equations that are endowed with physical meaning associated with ray theory or classical mechanics.

Closing remarks

The fact that Lagrange's ray equations are also Euler's equations implies that rays can be obtained as solutions of a variational problem. This fact allows us to use the tools of the calculus of variations in our investigations of ray theory.

In the calculus of variations, a stationary curve is given by Euler's equation. The conditions to specify that this curve results in a minimum or a maximum value of a given integral are difficult to formulate mathematically and are not addressed in this book.[10] Yet, in physically motivated problems the minimum or maximum nature of the stationary curve is often obvious from the physical context.

In Chapter 13, we will study Fermat's variational principle of stationary traveltime. We will show that the search for a ray is equivalent to the search for a curve along which the traveltime is stationary. In Chapter 14, we will show that first integrals, which correspond to conserved quantities along these rays, can be used in raytracing techniques.

Exercises

Exercise 12.1 *In view of Euler's equation (12.4), verify Beltrami's identity (12.10).*

[10] Readers interested in geodesic fields and its implication to minima and maxima might refer to Kreyszig, E., (1959/1991) Differential geometry: Dover, pp. 162 – 168.

Solution 12.1 *Consider* $F = F(x, z, z')$ *and Beltrami's identity (12.10). We can write*

$$\frac{\partial F}{\partial x}+\frac{\mathrm{d}}{\mathrm{d}x}\left(z'\frac{\partial F}{\partial z'}-F\right)=\frac{\partial F(x,z,z')}{\partial x}+\frac{\mathrm{d}}{\mathrm{d}x}\left(z'\frac{\partial F(x,z,z')}{\partial z'}\right)-\frac{\mathrm{d}F(x,z,z')}{\mathrm{d}x}$$

$$=\frac{\partial F}{\partial x}+z''\frac{\partial F}{\partial z'}+z'\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right)-\left(\frac{\partial F}{\partial x}+\frac{\partial F}{\partial z}z'+\frac{\partial F}{\partial z'}z''\right)$$

$$=z'\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right)-\frac{\partial F}{\partial z}z'=-z'\left[\frac{\partial F}{\partial z}-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right)\right],$$

where the terms in brackets is Euler's equation (12.4). Thus,

$$\frac{\partial F}{\partial x}+\frac{\mathrm{d}}{\mathrm{d}x}\left(z'\frac{\partial F}{\partial z'}-F\right)=0,$$

as required.

Exercise 12.2 *Following the argument used to justify Theorem 12.1, derive the explicit form of Beltrami's identity (12.10).*

Solution 12.2 *To obtain Beltrami's identity, consider term (12.7), namely,*

$$F\left(z_i, z_i'; x_i\right)\Delta x + F\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\Delta x,$$

where $i \in \{1, \ldots, n-1\}$. *Differentiating with respect to* x, *we obtain*

$$\left[\frac{\partial F}{\partial x}\left(z_i, z_i'; x_i\right)+\frac{\partial F}{\partial z'}\left(z_i, z_i'; x_i\right)\frac{\partial z_i'}{\partial x_i}\right]\Delta x+F\left(z_i, z_i'; x_i\right)\frac{\partial \Delta x}{\partial x_i}$$
$$+\frac{\partial F}{\partial z'}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\frac{\partial z_{i+1}'}{\partial x_i}\Delta x+F\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\frac{\partial \Delta x}{\partial x_i}. \tag{12.29}$$

Recalling expression (12.5) and the appropriate expression for Δx, *we can* *write*

$$\frac{\partial z_i'}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{z_i - z_{i-1}}{x_i - x_{i-1}}\right) = -\frac{z_i - z_{i-1}}{\left(x_i - x_{i-1}\right)^2} = -\frac{z_i - z_{i-1}}{\left(\Delta x\right)^2},$$

$$\frac{\partial \Delta x}{\partial x_i}\left(z_i, z_i'; x_i\right) = \frac{\partial}{\partial x_i}\left(x_i - x_{i-1}\right) = 1,$$

$$\frac{\partial z_{i+1}'}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{z_{i+1} - z_i}{x_{i+1} - x_i}\right) = \frac{z_{i+1} - z_i}{\left(x_{i+1} - x_i\right)^2} = \frac{z_{i+1} - z_i}{\left(\Delta x\right)^2},$$

$$\frac{\partial \Delta x}{\partial x_i}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right) = \frac{\partial}{\partial x_i}\left(x_{i+1} - x_i\right) = -1.$$

Hence, expression (12.29) becomes

$$\left[\frac{\partial F}{\partial x}\left(z_i, z_i'; x_i\right) - \frac{\partial F}{\partial z'}\left(z_i, z_i'; x_i\right)\frac{z_i - z_{i-1}}{\left(\Delta x\right)^2}\right]\Delta x + F\left(z_i, z_i'; x_i\right)$$

$$+ \frac{\partial F}{\partial z'}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\frac{z_{i+1} - z_i}{\left(\Delta x\right)^2}\Delta x - F\left(z_{i+1}, z_{i+1}'; x_{i+1}\right).$$

As in stationarity condition (12.6), we have a system of $n-1$ *equations, namely,*

$$\begin{aligned}\frac{\partial S_n}{\partial x_i} &= 0\\ &= F\left(z_i, z_i'; x_i\right) - F\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\\ &\quad - \frac{\partial F}{\partial z'}\left(z_i, z_i'; x_i\right)\frac{z_i - z_{i-1}}{\Delta x} + \frac{\partial F}{\partial z'}\left(z_{i+1}, z_{i+1}'; x_{i+1}\right)\frac{z_{i+1} - z_i}{\Delta x}\\ &\quad + \frac{\partial F}{\partial x}\left(z_i, z_i'; x_i\right)\Delta x,\end{aligned}$$

where $i \in \{1, \ldots, n-1\}$. *Dividing both sides of each equation by* Δx *and* *using the appropriate definition of Δx, we can write*

$$0 = \frac{F(z_i, z_i'; x_i) - F(z_{i+1}, z_{i+1}'; x_{i+1})}{\Delta x}$$

$$+ \frac{\frac{\partial F}{\partial z'}(z_{i+1}, z_{i+1}'; x_{i+1}) \frac{z_{i+1} - z_i}{\Delta x} - \frac{\partial F}{\partial z'}(z_i, z_i'; x_i) \frac{z_i - z_{i-1}}{\Delta x}}{\Delta x}$$

$$+ \frac{\partial F}{\partial x}(z_i, z_i'; x_i),$$

where $i \in \{1, \ldots, n-1\}$. Letting $\Delta x \to 0$, we see that z_i approaches $z(x)$ so that $z_i' = (z_i - z_{i-1})/\Delta x$ approaches $z'(x)$. Recognizing in the resulting statement the definitions of the derivatives, we obtain a single equation

$$0 = -\frac{\mathrm{d}F}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'} z'\right) + \frac{\partial F}{\partial x}.$$

Rearranging and using the linearity of the differential operator, we get

$$\frac{\partial F}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x}\left[z' \frac{\partial F}{\partial z'} - F\right] = 0,$$

which is Beltrami's identity (12.10), as required.

Exercise 12.3 [11]*Prove the following lemma.*

Notation 12.1 *To state Lemma 12.1, below, we use the parametric form of the variational problems, rather than the explicit form used in this chapter.*

Lemma 12.1 *Both variational problems*

$$\delta \int F(\mathbf{x}, \dot{\mathbf{x}})\, \mathrm{d}t = 0, \tag{12.30}$$

and

$$\delta \int \left[cF(\mathbf{x}, \dot{\mathbf{x}}) + \sum_{j=1}^{n} f_j(\mathbf{x})\, \dot{x}_j\right] \mathrm{d}t = 0, \tag{12.31}$$

have the same function $\mathbf{x}(t)$ that renders the corresponding integrals stationary, if $f_i(\mathbf{x})$ are the components of a gradient of a function and c is a nonzero constant.

[11] See also Section 12.5.4

Solution 12.3 *We show two different proofs of Lemma 12.1. Proof A invokes the properties of a variational fixed-ends problem, while Proof B utilizes standard properties of differential calculus in the context of Euler's equations.*

Proof A. To prove that $\mathbf{x}(t)$ is the same for variational problems (12.30) and (12.31), we reduce problem (12.31) to problem (12.30). Consider the integral of variational problem (12.31). In view of the linearity of the integral operator, we can write

$$\begin{aligned}\int\left[cF(\mathbf{x},\dot{\mathbf{x}})+\sum_{j=1}^{n}f_j(\mathbf{x})\dot{x}_j\right]\mathrm{d}t &= c\int F(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t+\int\sum_{j=1}^{n}f_j(\mathbf{x})\dot{x}_j\mathrm{d}t\\ &= c\int F(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t+\int\sum_{j=1}^{n}f_j(\mathbf{x})\,\mathrm{d}x_j.\end{aligned}$$

Consider the integral that involves the summation. Since the $f_j(\mathbf{x})$ are the components of ∇g, for some function $g(\mathbf{x})$, we can restate this integral as

$$\int\sum_{j=1}^{n}f_j(\mathbf{x})\,\mathrm{d}x_j=\int\sum_{j=1}^{n}\frac{\partial g(\mathbf{x})}{\partial x_j}\mathrm{d}x_j. \tag{12.32}$$

Since integral (12.32) is the integral of total differential

$$\frac{\partial g(\mathbf{x})}{\partial x_1}\mathrm{d}x_1+\ldots+\frac{\partial g(\mathbf{x})}{\partial x_n}\mathrm{d}x_n=\mathrm{d}g(\mathbf{x}),$$

the value of integral (12.32) is independent of the integration path. Hence, term $\sum f_j(\mathbf{x})\dot{x}_j$ has no effect on the choice of function $\mathbf{x}(t)$. Recalling that $c\neq 0$, we have reduced variational problem (12.31) to variational problem (12.30) and, hence, the proof is complete. ∎

Proof B. Consider variational problem (12.30). The corresponding Euler's equations are

$$\frac{\partial F}{\partial x_i}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F}{\partial \dot{x}_i}\right)=0,\qquad i\in\{1,\ldots,n\}. \tag{12.33}$$

Now, consider variational problem (12.31). The corresponding Euler's equations are

$$\frac{\partial\left[cF+\sum\limits_{j=1}^{n}f_j(\mathbf{x})\dot{x}_j\right]}{\partial x_i}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\left[cF+\sum\limits_{j=1}^{n}f_j(\mathbf{x})\dot{x}_j\right]}{\partial \dot{x}_i}\right)=0,\qquad i\in\{1,\ldots,n\},$$

which we can write as

$$c\left[\frac{\partial F}{\partial x_i}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F}{\partial \dot{x}_i}\right)\right]+\frac{\partial\left[\sum\limits_{j=1}^{n} f_j\left(\mathbf{x}\right)\dot{x}_j\right]}{\partial x_i}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\left[\sum\limits_{j=1}^{n} f_j\left(\mathbf{x}\right)\dot{x}_j\right]}{\partial \dot{x}_i}\right)=0, \tag{12.34}$$

where $i\in\{1,\ldots,n\}$. To prove that solution $\mathbf{x}(t)$ is the same for equations (12.33) and (12.34), we prove that these two systems of equations are equivalent to one another. Recalling that $c\neq 0$, to prove that equations (12.33) and (12.34) are equivalent, we need to show that

$$\frac{\partial\left[\sum\limits_{j=1}^{n} f_j\left(\mathbf{x}\right)\dot{x}_j\right]}{\partial x_i}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\left[\sum\limits_{j=1}^{n} f_j\left(\mathbf{x}\right)\dot{x}_j\right]}{\partial \dot{x}_i}\right),\qquad i\in\{1,\ldots,n\}. \tag{12.35}$$

Consider the left-hand side of equation (12.35). Using the linearity of the differential operator, we can write

$$\frac{\partial\left[\sum\limits_{j=1}^{n} f_j\left(\mathbf{x}\right)\dot{x}_j\right]}{\partial x_i}=\sum_{j=1}^{n}\frac{\partial f_j\left(\mathbf{x}\right)}{\partial x_i}\dot{x}_j,\qquad i\in\{1,\ldots,n\}. \tag{12.36}$$

Consider the right-hand side of equation (12.35). Using the linearity of the differential operator and taking into account the fact that the only term of $\sum f_j\left(\mathbf{x}\right)\dot{x}_j$ that is dependent on $\dot{x}_i$ is the term where $j=i$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\left[\sum\limits_{j=1}^{n} f_j\left(\mathbf{x}\right)\dot{x}_j\right]}{\partial \dot{x}_i}\right)=\frac{\mathrm{d}f_i\left(\mathbf{x}\right)}{\mathrm{d}t}=\sum_{j=1}^{n}\frac{\partial f_i\left(\mathbf{x}\right)}{\partial x_j}\dot{x}_j,\qquad i\in\{1,\ldots,n\}. \tag{12.37}$$

Examining the coefficients of $\dot{x}_j$ in expressions (12.36) and (12.37), we see that we need to show the equality given by

$$\frac{\partial f_j\left(\mathbf{x}\right)}{\partial x_i}=\frac{\partial f_i\left(\mathbf{x}\right)}{\partial x_j},\qquad i,j\in\{1,\ldots,n\}. \tag{12.38}$$

Recall that $[f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})]$ are the components of ∇g, for some function $g(\mathbf{x})$, namely, $[\partial g/\partial x_1, \ldots, \partial g/\partial x_n]$. Thus, we can write the left-hand side of equation (12.38) as

$$\frac{\partial f_j(\mathbf{x})}{\partial x_i} = \frac{\partial}{\partial x_i}\left[\frac{\partial g(\mathbf{x})}{\partial x_j}\right] = \frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j}, \qquad i, j \in \{1, \ldots, n\}.$$

Analogously, we can write the right-hand side of equation (12.38) as

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\frac{\partial g(\mathbf{x})}{\partial x_i}\right] = \frac{\partial^2 g(\mathbf{x})}{\partial x_j \partial x_i}, \qquad i, j \in \{1, \ldots, n\}.$$

Hence, due to the equality of mixed partial derivatives — which we can write as $\partial^2 g/\partial x_i \partial x_j = \partial^2 g/\partial x_j \partial x_i$ — the proof is complete. ∎

Exercise 12.4 *Consider a variational problem given by integral (12.21). Let $f(x, z) = 1$, and let the endpoints be $(0, 0)$ and $(1, 1)$. Find function $z(x)$ that renders this integral stationary and calculate the value of the integral along this function. Choose another function that connects the endpoints and show that the resulting value of the integral is greater than the one corresponding to the extremizing function. In view of Section 12.5.5, provide a geometrical illustration.*

Solution 12.4 *The variational problem in question is*

$$\delta \int_0^1 \sqrt{1 + \left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2}\, \mathrm{d}x = 0. \tag{12.39}$$

Since integral (12.39) depends only on z', in view of Section 12.5.2, the extremizing function is a straight line given by $z(x) = x$. Inserting $z = x$ into integral (12.39), we obtain the distance along the extremizing function, namely,

$$\int_0^1 \sqrt{2}\, \mathrm{d}x = \sqrt{2},$$

as also expected from Pythagoras' theorem. Now, let us calculate the distance along another curve, for instance, $z(x) = x^2$. Integral (12.39) becomes

$$\int_0^1 \sqrt{1 + (2x)^2}\, \mathrm{d}x = \frac{1}{4}\left[2x\sqrt{1 + (2x)^2} + \operatorname{Arcsinh}(2x)\right]_0^1 \approx 1.48,$$

which is greater than $\sqrt{2}$*, as expected.*
In view of Section 12.5.5, integral (12.39) is the surface area of a strip whose width is equal to unity, due to $f(x,z) = 1$*, and whose length corresponds to the curve* $z(x)$*, between* $x = 0$ *and* $x = 1$*. Since the width of the strip is constant, the least surface area corresponds to the shortest curve connecting the two points. Hence, the extremizing function is* $z(x) = x$*, which is a straight line.*

Exercise 12.5 *Express Euler's equation (12.4) as the corresponding Hamilton's equations in* dz/dx *and* dp/dx*.*

Solution 12.5 *Consider integrand (12.1), namely,* $F(z, z'; x)$*. In view of Legendre's transformation, discussed in Appendix B, let the variable of transformation be denoted by*

$$p := \frac{\partial F}{\partial z'}, \tag{12.40}$$

and the new function be

$$H(z, p; x) = pz' - F(z, z'; x), \tag{12.41}$$

which is the Hamiltonian corresponding to F*. Hence, by the duality of Legendre's transformation, we can write*

$$z' = \frac{\partial H}{\partial p}. \tag{12.42}$$

Invoking Euler's equation (12.4) and in view of expression (12.40), we can write

$$\frac{\partial F}{\partial z} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) = \frac{\mathrm{d}p}{\mathrm{d}x} \equiv p'. \tag{12.43}$$

Hence, using expression (12.41) to express function F *in terms of function* H*, we obtain*

$$p' = \frac{\partial F}{\partial z} = \frac{\partial}{\partial z}\left[pz' - H(z, p; x)\right] = -\frac{\partial H}{\partial z}. \tag{12.44}$$

Thus, using equations (12.42) and (12.44), we can write a system of first-order ordinary differential equations in dz/dx *and* dp/dx*, namely,*

$$\left\{\begin{array}{l} z' = \dfrac{\partial H}{\partial p} \\ \\ p' = -\dfrac{\partial H}{\partial z} \end{array}\right. , \tag{12.45}$$

which are the required Hamilton's equations.

Exercise 12.6 *In view of Exercise 12.5, prove the following theorem.*

Theorem 12.2 *For an integral given by expression (12.1), namely,*

$$\int_a^b F\left[z(x), z'(x); x\right] \mathrm{d}x,$$

if F does not explicitly depend on x, the corresponding Hamiltonian, H, is the first integral of equation (12.4).

Solution 12.6 .

Proof. We can formally write

$$\frac{\mathrm{d}H(p,z;x)}{\mathrm{d}x} = \frac{\partial H}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\partial H}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}x} + \frac{\partial H}{\partial x}. \tag{12.46}$$

Invoking equations (12.45), the first two terms on the right-hand side of equation (12.46) vanish and, hence, this equation is reduced to

$$\frac{\mathrm{d}H(p,z;x)}{\mathrm{d}x} = \frac{\partial H}{\partial x}.$$

In view of expression (12.41), H does not depend on x explicitly if and only if F does not depend on x explicitly. In such a case, we obtain

$$\frac{\mathrm{d}H}{\mathrm{d}x} = 0,$$

and, hence, H is constant. Using expressions (12.40) and (12.41), we can write

$$H = pz' - F = \frac{\partial F}{\partial z'}z' - F = C,$$

where C denotes a constant. This is the first integral of equation (12.4), given by expression (12.19). ■

Chapter 13

Fermat's principle

> There is hardly any other branch of mathematical sciences in which abstract mathematical speculations and concrete physical evidence go so beautifully together and complement each other so perfectly. [...] In spite of all differences in the interpretation, the variational principles of mechanics continue to hold their ground in the description of all the phenomena of nature.
>
> *Cornelius Lanczos (1949) The variational principles of mechanics*

Preliminary remarks

In Chapter 9, we derived Lagrange's ray equations, which, as shown in Chapter 12, are the stationarity conditions for a definite integral. In this chapter, we will show that this definite integral corresponds to the traveltime of a signal between two points in an anisotropic inhomogeneous continuum. Consequently, we can study ray theory in terms of Fermat's variational principle of stationary traveltime.

In general, physical applications of the calculus of variations are based on the fact that the behaviours of physical systems appear to coincide with the extremals of certain integrals. For instance, while in ray theory this integral corresponds to the traveltime, in classical mechanics this integral is given in terms of the kinetic and potential energies.

We begin this chapter with the statement of Fermat's principle as a theorem dealing with rays. Proof of this theorem is rooted in Hamilton's ray equations, where the mathematical concept of a ray originates. Hence, we investigate several properties of the ray-theory Hamiltonian and the resulting Lagrangian and, using these properties, obtain a proof of this theorem. We also discuss another variational principle that is pertinent to our studies, namely, Hamilton's principle.

13.1 Formulation of Fermat's principle

Notation 13.1 *In this section, to show the generality of the formulation, all expressions are derived for an n-dimensional space.*

13.1.1 Statement of Fermat's principle

In 1657, Pierre de Fermat formulated his variational principle for the propagation of light. He stated that light travels along a curve that renders the traveltime minimum. In modern notation, a generic form of this principle, to which we refer as the principle of stationary traveltime, can be restated by the following theorem.

Theorem 13.1 [1]*Rays are the solutions of the variational problem*

$$\delta \int_A^B \frac{\mathrm{d}s}{V(\mathbf{x},\mathbf{n})} = 0, \tag{13.1}$$

where $\mathrm{d}s$ *is an arclength element and* $V(\mathbf{x},\mathbf{n})$ *is the ray velocity in direction* $\mathbf{n} = \mathrm{d}\mathbf{x}/\mathrm{d}s$ *at point* $\mathbf{x}$. A *and* B *are the fixed endpoints of this variational problem.*

Note that, in expression (13.1) and throughout Section 13.1, $\mathbf{n}$ denotes a vector tangent to the ray and not a vector normal to the wavefront, as is the case in other sections of this book.

13.1.2 Properties of Hamiltonian $\mathcal{H}$

In order to prove Theorem 13.1, we must show that the solution of variational problem (13.1) is equivalent to the solution of Hamilton's ray equations

[1] The proof of the theorem shown in this section is based on Bóna, A., and Slawinski, M.A., (2003) Fermat's principle for seismic rays in elastic continua: Journal of Applied Geophysics.

(8.19), namely,

$$\left\{\begin{array}{l}\dot{x}_i = \dfrac{\partial \mathcal{H}}{\partial p_i} \\ \\ \dot{p}_i = -\dfrac{\partial \mathcal{H}}{\partial x_i}\end{array}\right., \qquad i \in \{1, \ldots, n\}. \tag{13.2}$$

Let us investigate the properties of the Hamiltonian that is given by expression (8.20), namely,

$$\mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} p^2 v^2 (\mathbf{x}, \mathbf{p}),$$

and which, in view of v being homogeneous of degree 0 in the p_i, can also be stated as

$$\mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} p^2 v^2 \left(\mathbf{x}, \frac{\mathbf{p}}{|\mathbf{p}|}\right), \tag{13.3}$$

where $|\mathbf{p}|$ is the magnitude of the phase-slowness vector.

By examining expression (13.3), we note the following properties of this Hamiltonian. $\mathcal{H}$ is homogeneous of degree 2 in the p_i. Also, since $\mathcal{H}$ does not explicitly depend on time, its value is conserved along the ray. The latter property can be stated by the following lemma.

Lemma 13.1 *Hamiltonian $\mathcal{H}(\mathbf{x}, \mathbf{p})$, given by expression (13.3), is conserved along the ray.*

Proof. Using system (13.2) and the fact that $\mathcal{H}$ does not explicitly depend on time, we can write

$$\begin{aligned}\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} &= \sum_{i=1}^{n} \frac{\partial \mathcal{H}}{\partial x_i} \dot{x}_i + \sum_{i=1}^{n} \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i + \frac{\partial \mathcal{H}}{\partial t} \\ &= -\sum_{i=1}^{n} \dot{p}_i \dot{x}_i + \sum_{i=1}^{n} \dot{x}_i \dot{p}_i = 0.\end{aligned}$$

■

Moreover, the value of the Hamiltonian, which is conserved along the ray, is equal to 1/2. This results from the fact that the eikonal equation, which is shown in equation (7.17), must be satisfied along the rays. Hence, in view of this equation, which states that $p^2 v^2 = 1$, and expression (8.20), we require that

$$\mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} \tag{13.4}$$

along a ray.

13.1.3 Variational equivalent of Hamilton's ray equations

To show that rays obtained from Hamilton's ray equations (13.2) are solutions of variational problem (13.1), we express these equations in the context of the calculus of variations. As stated in Section 12.7, Lagrange's ray equations (9.8), namely,

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_i} = 0, \qquad i \in \{1, \ldots, n\}, \tag{13.5}$$

possess the form of Euler's equations. Consequently, in view of Chapter 12, we can state the following proposition.

Proposition 13.1 *Rays are the solutions of the variational problem*

$$\delta \int \mathcal{L} \mathrm{d}t = 0, \tag{13.6}$$

where the ray-theory Lagrangian $\mathcal{L}$ is given by expression (9.2), namely

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \sum_{j=1}^{n} p_j (\mathbf{x}, \dot{\mathbf{x}}) \dot{x}_j - \mathcal{H}(\mathbf{x}, \mathbf{p}). \tag{13.7}$$

13.1.4 Properties of Lagrangian $\mathcal{L}$

To examine variational formulations (13.1) and (13.6), we must study the properties of Lagrangian $\mathcal{L}$, given by expression (13.7), in terms of the corresponding Hamiltonian, $\mathcal{H}$. We begin by stating the following lemma.

Lemma 13.2 *If $\mathcal{H}(\mathbf{x}, \mathbf{p})$ is homogeneous of degree 2 in the p_i, then*

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}(\mathbf{x}, \mathbf{p})) = \mathcal{H}(\mathbf{x}, \mathbf{p}),$$

where, by Legendre's transformation, $\dot{x}_i = \partial \mathcal{H} / \partial p_i$.

Proof. Consider Lagrangian

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}(\mathbf{x}, \mathbf{p})) = \sum_{i=1}^{n} p_i \dot{x}_i - \mathcal{H}.$$

In view of Legendre's transformation, we can write

$$\mathcal{L}(\mathbf{x}(t), \dot{\mathbf{x}}(\mathbf{x}(t), \mathbf{p}(t))) = \sum_{i=1}^{n} p_i \frac{\partial \mathcal{H}}{\partial p_i} - \mathcal{H}.$$

If $\mathcal{H}$ is homogeneous of degree 2 in the p_i, by Theorem A.1, stated in Appendix A, we obtain

$$\mathcal{L}\left(\mathbf{x}\left(t\right),\dot{\mathbf{x}}\left(\mathbf{x}\left(t\right),\mathbf{p}\left(t\right)\right)\right)=2\mathcal{H}-\mathcal{H}=\mathcal{H},$$

which completes the proof. ∎

In view of the conserved value of Hamiltonian $\mathcal{H}$, as shown in Lemma 13.1, and following expression (13.4), we obtain the following corollary of Lemma 13.2.

Corollary 13.1 *Along each ray, Lagrangian $\mathcal{L}$ is equal to* $1/2$.

In view of $\mathcal{H}$ being homogeneous of degree 2 in the p_i, the analogous property of $\mathcal{L}$ is shown in the following lemma.

Lemma 13.3 *If Hamiltonian $\mathcal{H}\left(\mathbf{x},\mathbf{p}\right)$ is homogeneous of degree* 2 *in the p_i, then Lagrangian $\mathcal{L}\left(\mathbf{x},\dot{\mathbf{x}}\right)$ is homogeneous of degree* 2 *in the $\dot{x}_i$.*

Proof. By Lemma 13.2, $\mathcal{H}\left(\mathbf{x},\mathbf{p}\right)=\mathcal{L}\left(\mathbf{x},\dot{\mathbf{x}}\left(\mathbf{x},\mathbf{p}\right)\right)$, where $\dot{\mathbf{x}}$ and $\mathbf{p}$ are related by Legendre's transformation $\dot{x}_i=\partial\mathcal{H}/\partial p_i$. Let $\mathbf{p}':=a\mathbf{p}$, where a is a constant. The corresponding Hamilton's equations are

$$\dot{x}_i'=\frac{\partial\mathcal{H}\left(\mathbf{x},\mathbf{p}'\right)}{\partial p_i'}=\frac{\partial\mathcal{H}\left(\mathbf{x},a\mathbf{p}\right)}{\partial\left(ap_i\right)},\qquad i\in\left\{1,\ldots,n\right\}.$$

By the homogeneity of $\mathcal{H}$ and the property of the differential operator, we can write

$$\dot{x}_i'=\frac{\partial\mathcal{H}\left(\mathbf{x},a\mathbf{p}\right)}{\partial\left(ap_i\right)}=\frac{a^2\dfrac{\partial\mathcal{H}\left(\mathbf{x},\mathbf{p}\right)}{\partial p_i}}{\dfrac{\partial\left(ap_i\right)}{\partial p_i}}=\frac{a^2\dfrac{\partial\mathcal{H}\left(\mathbf{x},\mathbf{p}\right)}{\partial p_i}}{a},\qquad i\in\left\{1,\ldots,n\right\}.$$

Hence,

$$\dot{x}_i'=a\frac{\partial\mathcal{H}\left(\mathbf{x},\mathbf{p}\right)}{\partial p_i},\qquad i\in\left\{1,\ldots,n\right\},$$

which, in view of Hamilton's ray equations, given by system (13.2), can be stated as

$$\dot{x}_i'=a\dot{x}_i,\qquad i\in\left\{1,\ldots,n\right\}.$$

Consequently, we can write

$$\mathcal{L}\left(\mathbf{x},a\dot{\mathbf{x}}\right)=\mathcal{L}\left(\mathbf{x},\dot{\mathbf{x}}'\right),$$

which, by Lemma 13.2, yields

$$\mathcal{L}(\mathbf{x}, \mathbf{a}\dot{\mathbf{x}}) = \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}') = \mathcal{H}(\mathbf{x}, \mathbf{p}') = \mathcal{H}(\mathbf{x}, a\mathbf{p}) = a^2\mathcal{H}(\mathbf{x}, \mathbf{p}) = a^2\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}),$$

where the expression in the middle results from the homogeneity of $\mathcal{H}$. This means that Lagrangian $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ is homogeneous of degree 2 in the $\dot{x}_i$. ∎

Lemma 13.3 implies that variational problem (13.6) has a fixed parametrization since $\mathcal{L}$ is homogeneous of degree 2 in the $\dot{x}_i$. For a variational problem to be independent of parametrization, the integrand must be homogeneous of degree 1 in the $\dot{x}_i$, as shown in Exercise 13.1.

Note that, as shown in Section 8.2, the solutions of Hamilton's ray equations (13.2) are parametrized by time; hence, the solutions of system (13.5) are also parametrized by time. Also note that, in view of the homogeneity of the Lagrangian and its not being explicitly dependent on t, Beltrami's identity together with Euler's homogeneous-function theorem imply that $\mathcal{L}$ is conserved along any ray, as shown in Exercise 13.2. As expected, this result is consistent with Corollary 13.1.

13.1.5 Parameter-independent Lagrange's ray equations

Parametrization independence is necessary to state Fermat's principle since its generic form, shown in expression (13.1), is parametrization independent. This results from the fact that the integrand in expression (13.1) is homogeneous of degree 1 in the $\dot{x}_i$.

Let us consider a Lagrangian given by

$$\mathcal{F} = \sqrt{2\mathcal{L}}, \tag{13.8}$$

where $\mathcal{L}$ is given by (13.7). Note that, following Definition A.1, stated in Appendix A, $\mathcal{F}$ is absolute-value homogeneous of degree 1 in the $\dot{x}_i$. Under certain conditions, which are satisfied in our case, the solutions of Lagrange's ray equations (13.5) are also the solutions of the equations given by

$$\frac{\partial \mathcal{F}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{F}}{\partial \dot{x}_i}\right) = 0, \qquad i \in \{1, \ldots, n\}. \tag{13.9}$$

This is stated by the following lemma.

Lemma 13.4 *A solution of equations (13.5) that satisfies the condition given in Corollary 13.1, where $\mathcal{L}$ is given by expression (13.7), is also a solution of equations (13.9), where $\mathcal{F} = \sqrt{2\mathcal{L}}$.*

Proof. Inserting $\mathcal{L} = \mathcal{F}^2/2$ into equations (13.5), we obtain

$$\begin{aligned}\frac{\partial}{\partial x_i}\left(\frac{\mathcal{F}^2}{2}\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial \dot{x}_i}\left(\frac{\mathcal{F}^2}{2}\right)\right] &= \mathcal{F}\frac{\partial \mathcal{F}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left[\mathcal{F}\frac{\partial \mathcal{F}}{\partial \dot{x}_i}\right] \\ &= \mathcal{F}\left[\frac{\partial \mathcal{F}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{F}}{\partial \dot{x}_i}\right)\right] - \frac{\mathrm{d}\mathcal{F}}{\mathrm{d}t}\frac{\partial \mathcal{F}}{\partial \dot{x}_i} \\ &= 0, \qquad i \in \{1, \ldots, n\}.\end{aligned}$$

Since $\mathcal{L} = 1/2$ along a ray, as shown in Corollary 13.1, then $\mathcal{F} = 1$ and, hence, $\mathrm{d}\mathcal{F}/\mathrm{d}t = 0$ along the solutions of equations (13.5). Thus, equations (13.5) become equations (13.9), as required. ■

Consequently, equations (13.9) can also be viewed as Lagrange's ray equations.

If we can show that

$$\mathcal{F} = \frac{|\dot{\mathbf{x}}|}{V\left(\mathbf{x}, \frac{\dot{\mathbf{x}}}{|\dot{\mathbf{x}}|}\right)}, \tag{13.10}$$

where $|\dot{\mathbf{x}}| = \mathrm{d}s/\mathrm{d}t$ and $\dot{\mathbf{x}}/|\dot{\mathbf{x}}| = \mathbf{n}$, then we prove Theorem 13.1, since the right-hand side of equation (13.10) is the integrand of equation (13.1).

13.1.6 Ray velocity

In order to show that the right-hand side of equation (13.10) is the integrand of equation (13.1), we must formulate ray velocity in a variational context. Since, as shown in Lemma 13.3, Lagrangian $\mathcal{L}$ is homogenous of degree 2 in the $\dot{x}_i$, we can write

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \mathcal{L}(\mathbf{x}, |\dot{\mathbf{x}}|\,\mathbf{n}) = |\dot{\mathbf{x}}|^2\,\mathcal{L}(\mathbf{x}, \mathbf{n}),$$

where $\mathbf{n} = \dot{\mathbf{x}}/|\dot{\mathbf{x}}|$ is a unit vector tangent to the ray. Since, as stated in Corollary 13.1, the value of Lagrangian $\mathcal{L}$ along a ray is $1/2$, we can write

$$\frac{1}{2} = |\dot{\mathbf{x}}|^2\,\mathcal{L}(\mathbf{x}, \mathbf{n}).$$

Since this expression is valid along any ray, the ray velocity V, given by $|\dot{\mathbf{x}}|$, can be expressed as

$$V(\mathbf{x}, \mathbf{n}) := |\dot{\mathbf{x}}| = \frac{1}{\sqrt{2\mathcal{L}(\mathbf{x}, \mathbf{n})}}, \tag{13.11}$$

which is consistent with expression (9.12).

Now, we are ready to complete the proof of Theorem 13.1.

13.1.7 Proof of Fermat's principle

Proof. By Lemma 13.4, rays are the solutions of Euler's equations stated in system (13.9). Consequently, rays are the solutions of variational problem

$$\delta \int \mathcal{F}\,\mathrm{d}t = 0. \tag{13.12}$$

In view of expression (13.8), we can restate this variational problem as

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}})\ \mathrm{d}t = \delta \int \sqrt{2\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}\,\mathrm{d}t = 0.$$

Since, as stated in Lemma 13.3, $\mathcal{L}$ is homogeneous of degree 2 in the $\dot{x}_i$, we can write

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}})\ \mathrm{d}t = \delta \int \sqrt{2\left|\dot{\mathbf{x}}\right|^2 \mathcal{L}(\mathbf{x}, \mathbf{n})}\,\mathrm{d}t = \delta \int \left|\dot{\mathbf{x}}\right| \sqrt{2\mathcal{L}(\mathbf{x}, \mathbf{n})}\,\mathrm{d}t = 0.$$

In view of expressions (13.11) and since $\left|\dot{\mathbf{x}}\right|\ \mathrm{d}t = \mathrm{d}s$, we conclude that

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}})\,\mathrm{d}t = \delta \int \frac{\mathrm{d}s}{V(\mathbf{x}, \mathbf{n})} = 0.$$

Hence, the solutions of Hamilton's ray equations that correspond to rays are the solutions of variational problem (13.1). ∎

Theorem 13.1 states that seismic rays in anisotropic inhomogeneous continua obey Fermat's principle of stationary traveltime. Since our proof relies on Legendre's transformation, discussed in Appendix B, it is valid only if the Hamiltonian, $\mathcal{H}$, is regular, namely,

$$\det\left[\frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j}\right] \neq 0, \qquad i, j \in \{1, \ldots, n\}.$$

In other words, we are unable to prove Theorem 13.1 at the inflection points of the phase-slowness surface. As stated in Theorem 10.3, for an elastic continuum defined by constitutive equations (3.1), the innermost phase-slowness surface is always convex and, hence, the Hamiltonian associated with the fastest wave is always regular. For the slower waves, however, there are points where the Hamiltonian is irregular. This does not mean that Fermat's principle does not hold in general; however, the proof of Theorem 13.1 in the context of a phase-velocity function giving an irregular Hamiltonian remains an open problem.

Heuristically, the principle of stationary traveltime can be justified by the fact that among all signals of finite duration, the signals arriving at the receiver at the same instant constructively interfere and, consequently, contribute to the recorded observation, while the contribution of a multitude of signals arriving at different times is negligible.

13.2 Illustration of Hamilton's principle

Introductory comments

Born and Wolf, in their classic book entitled "Principles of optics", make the following statement.

> Variational considerations are of considerable importance as they often reveal analogies between different branches of physics. In particular there is a close analogy between geometrical optics and the mechanics of a moving particle; this was brought out very clearly by the celebrated investigations of Sir W.R. Hamilton, whose approach became of great value in modern physics, especially in applications to de Broglie's wave mechanics.

In this book, we focus on the variational formulation of geometrical optics without explicitly studying the analogies among different branches of physics. In this section, however, we will illustrate the analogy with classical mechanics by deriving the wave equation using Hamilton's variational principle.

13.2.1 Action

Fermat's principle, discussed in Section 13.1, plays an important role in ray theory. Another variational principle, which is pertinent to wave phenomena in elastic continua, is that of Hamilton. As stated by Arnold, in "Mathematical methods of classical mechanics",

> the fundamental notions of classical mechanics arose by the transforming of several very simple and natural notions of geometrical optics, guided by a particular variational principle — that of Fermat, into general variational principles.

In this section, we will illustrate Hamilton's principle in a simple context where the resulting Lagrange's equations of motion can be viewed as a restatement of Newton's second law of motion. Consequently, using the particular case of Hamilton's variational principle, we derive the one-dimensional wave equation, which corresponds to homogeneous continua.

While Newton proposed to measure motion by the rate of change of momentum, Leibniz suggested another quantity, the *vis viva*[2]. In the stan-

[2] *living force.* Readers interested in the origin of this entity might refer to Toretti, R., (1999) The philosophy of physics: Cambridge University Press, pp. 33 – 36.

dard formulation of classical mechanics, *vis viva* can be viewed as twice the kinetic energy. *Vis viva* underlies the concept of action.

The commonly accepted definition of action is

$$\mathbb{A} := \int_{t_1}^{t_2} \mathbb{L}\, \mathrm{d}t, \tag{13.13}$$

where $\mathbb{L}$ is the classical-mechanics Lagrangian that is defined by

$$\mathbb{L} := T - U, \tag{13.14}$$

with T and U denoting the kinetic energy and the potential energy, respectively.

In classical mechanics, the principle of least action was proposed by de Maupertuis who, in 1744 in a document appropriately entitled *Accord des différentes lois de la nature qui avait jusqu'ici paru incompatibles*[3], stated that

> l'action est proportionnelle au produit de la masse par la vitesse et par l'espace. Maintenant, voici ce principe si digne de l'Être suprême: Lorsqu'il arrive quelque changement dans la Nature, la quantité d'action employée pour ce changement est toujours la plus petite qu'il soit possible.[4]

However, careful analysis of the variational methods led to the formulation of the principle of stationary action rather than the principle of least action. The stationary-action principle was rigorously stated by Hamilton who wrote that

> although the law of least action has thus attained a rank among the highest theorems of physics, yet its pretensions to a cosmological necessity, on the grounds of economy in the universe, are now generally rejected. And the rejection appears just, for this, among other reasons, that the quantity pretended to be economized is in fact often lavishly expended.

[3] Agreement of various laws of nature which until now appeared incompatible

[4] Action is proportional to the product of mass, velocity, and displacement. Consequently, the principle so worthy of the Supreme Being: When there is a change in Nature, the value of action used for this change is the smallest possible.

In other words, action may be either a minimum or maximum. As a result, in classical mechanics, the principle of stationary action proposed by Hamilton states that

> if the positions of a conservative system are given at two instants, t_1 and t_2, the value of the time integral of Lagrangian $\mathbb{L}$ is stationary for the path actually described by this system, as compared to any other path that connects the two positions and obeys the constraints of the system.

In other words, in view of definition (13.13), finding a stationary value of action is equivalent to variational problem

$$\delta \mathbb{A} = \delta \int_{t_1}^{t_2} \mathbb{L}\,\mathrm{d}t = 0. \tag{13.15}$$

From the variational principle of action, it is possible to derive many equations of mathematical physics. In particular, a variational derivation of the wave equation is shown in Section 13.2.3. In the context of this illustration, the potential energy is assumed to be a function of position alone, while the kinetic energy is assumed to be a function of velocity alone. In other words, for this illustration of Hamilton's principle, we confine our interests to homogeneous continua.

13.2.2 Lagrange's equations of motion

In this section, we introduce Lagrange's equations of motion using the concepts of particle mechanics in order to familiarize the reader with this classical formulation. In the context of seismic wave propagation, the reader can omit this section and proceed directly to Section 13.2.3.

Considering Hamilton's principle, stated in equation (13.15), and in view of the stationarity conditions, discussed in Chapter 12, the motion of a particle must satisfy Euler's equations. The parametric form of Euler's equations can be written as

$$\frac{\partial \mathbb{L}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbb{L}}{\partial \dot{x}_i}\right) = 0, \qquad i \in \{1,2,3\}, \tag{13.16}$$

where t denotes time, x_i is the position coordinate and, hence, $\dot{x}_i$ is a component of the velocity vector tangent to the trajectory of this particle. Equations (13.16) are Lagrange's equations of motion. In the context of this section, since the kinetic energy does not depend on position, Lagrange's equations of motion (13.16) are just a restatement of Newton's second law of motion. To justify this equivalence, consider the following description.

Considering the first term of Lagrange's equations of motion (13.16) and recalling that T is assumed to be a function of velocity alone, we obtain

$$\frac{\partial \mathbb{L}}{\partial x_i} = \frac{\partial \left[T\left(\dot{\mathbf{x}}\right) - U\left(\mathbf{x}\right)\right]}{\partial x_i} = -\frac{\partial U}{\partial x_i} := F_i, \qquad i \in \{1,2,3\},$$

which is the expression for a component of force in a conservative field. Considering the expression in parentheses in the second term of Lagrange's equations of motion (13.16) and recalling that U is assumed to be a function of position alone, we obtain

$$\frac{\partial \mathbb{L}}{\partial \dot{x}_i} = \frac{\partial \left[T\left(\dot{\mathbf{x}}\right) - U\left(\mathbf{x}\right)\right]}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} := p_i, \qquad i \in \{1,2,3\}, \tag{13.17}$$

which is the expression for a component of momentum.

Since the first term of equations (13.16) is the component of force, while the second term is the rate of change of the corresponding component of momentum, Lagrange's equations of motion (13.16) are equivalent to Newton's second law of motion, namely,

$$F_i - \frac{\mathrm{d}p_i}{\mathrm{d}t} = 0, \qquad i \in \{1,2,3\}.$$

Also, as shown in Exercises 13.6 and 13.7, we can derive Hamilton's equations of motion from Newton's laws of motion.

Lagrange's equations of motion (13.16) apply to discrete systems, where the Lagrangian depends on the position of each particle. However, as shown in the following section, we can use the principle of stationary action in the context of continua, where the motion is defined by coordinates that are functions of both time and position variables.[5]

13.2.3 Wave equation

Continuous systems and Lagrangian density

A seismological application of stationary-action principle (13.15) and, consequently, of Lagrange's equations of motion for elastic continua, is exem-

[5]Readers interested in the energy propagation in the seismological context of the continuum using the Lagrange equations of motion and the generalized coordinates, might refer to Udías, A., (1999) Principles of seismology: Cambridge University Press, pp. 36 – 38.

plified by the derivation of the wave equation.[6] The coordinates of a three-dimensional continuous system are given by three position variables, x_1, x_2, x_3, and the time variable, t. Consequently, the displacement is given as a function of four independent variables, namely, $u = u(x_1, x_2, x_3, t)$. Hence, for a three-dimensional continuum, Lagrangian $\mathbb{L}$ is associated with an element of volume and is given by

$$\mathbb{L} = \iiint \mathfrak{L}\, \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3, \tag{13.18}$$

where $\mathfrak{L}$ is the Lagrangian density

$$\mathfrak{L} = \mathfrak{L}\left(u, \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial t}, x_i, t\right), \qquad i = \{1, 2, 3\}.$$

Variational derivation of wave equations

Consider oscillations of a finite-length string with fixed ends. Let the string itself be massless, have a length l, and contain n equal masses, m_i, spaced at equal intervals, Δx. Let the longitudinal displacements of masses be $u_0, \ldots, u_{n+1}$, with $u_0 = u_{n+1} = 0$ being the boundary conditions corresponding to fixed ends. Assume the force, F, required to stretch a length Δx of the string by amount u, to be

$$F = \frac{k}{\Delta x} u, \tag{13.19}$$

where k denotes a constant.

Note that the term $k/\Delta x$ has the units of $[N/m]$ and expression (13.19) can be viewed as a one-dimensional statement of Hooke's law.

The potential energy, U, is associated with the elasticity of the string and is given by the strain-energy function, discussed in Chapter 4. The potential energy of a segment of the string is

$$\mathrm{d}U = \frac{1}{2} \frac{k}{\Delta x} (\Delta u)^2,$$

where $\Delta u \equiv u_i - u_{i-1}$. Summing all the segments, the potential energy along the entire string containing n discrete mass points is

$$U = \frac{1}{2} k \sum_{i=1}^{n} \frac{(u_i - u_{i-1})^2}{\Delta x}. \tag{13.20}$$

[6]Readers interested in further descriptions of the Lagrangian formulation for continuous systems might refer to Goldstein, H., (1950/1980) Classical mechanics: Addison-Wesley Publishing Co., pp. 548 – 555.

The kinetic energy, T, for the entire string containing n discrete mass points, each of which has a mass m, is

$$T = \frac{1}{2} m \sum_{i=1}^{n} \left[\frac{\partial u (x_i, t)}{\partial t} \right]^2 . \tag{13.21}$$

If $n \to \infty$ and $\Delta x \to 0$ in such a way that $(n+1) \Delta x = l$, the potential energy, U, can be written as

$$U = \frac{1}{2} k \sum_{i=1}^{\infty} \left[\frac{u (x_i, t) - u (x_{i-1}, t)}{x_i - x_{i-1}} \right]^2 (x_i - x_{i-1}) .$$

Thus, in the limit, the term in brackets represents a partial derivative with respect to x, while the summation results in integration. Hence, we can write

$$U = \frac{1}{2} k \int_0^l \left[\frac{\partial u (x, t)}{\partial x} \right]^2 \mathrm{d}x. \tag{13.22}$$

Similarly, for the kinetic energy, rearranging and using the limit, we can write

$$T = \frac{1}{2} \frac{m}{\Delta x} \sum_{i=1}^{\infty} \left[\frac{\partial u (x_i, t)}{\partial t} \right]^2 \Delta x = \frac{1}{2} \rho \int_0^l \left[\frac{\partial u (x, t)}{\partial t} \right]^2 \mathrm{d}x, \tag{13.23}$$

where $\rho := \lim_{\Delta x \to 0} m / \Delta x$, is the mass density of the one-dimensional continuum.

Since the kinetic energy, given in expression (13.23), is not a function of position, we can invoke the classical-mechanics Lagrangian, given by expression (13.14). Thus, using expressions (13.22) and (13.23), we can write

$$\begin{aligned} \mathbb{L} (x, t) &= T (\dot{x}) - U (x) \\ &= \frac{1}{2} \rho \int_0^l \left[\frac{\partial u (x, t)}{\partial t} \right]^2 \mathrm{d}x - \frac{1}{2} k \int_0^l \left[\frac{\partial u (x, t)}{\partial x} \right]^2 \mathrm{d}x \\ &= \int_0^l \left\{ \frac{\rho}{2} \left[\frac{\partial u (x, t)}{\partial t} \right]^2 - \frac{k}{2} \left[\frac{\partial u (x, t)}{\partial x} \right]^2 \right\} \mathrm{d}x. \end{aligned} \tag{13.24}$$

Since we are presently dealing with a one-dimensional continuum, considering expression (13.18), we can write

$$\mathbb{L} = \int_0^l \mathfrak{L}\,\mathrm{d}x, \tag{13.25}$$

where, in view of integral (13.24), $\mathfrak{L}$ is the Lagrangian density given by

$$\mathfrak{L} \equiv \frac{\rho}{2}\left(\frac{\partial u}{\partial t}\right)^2 - \frac{k}{2}\left(\frac{\partial u}{\partial x}\right)^2. \tag{13.26}$$

To invoke a variational formulation, in view of expression (13.25) and following equation (13.15), we can write

$$\delta\int_0^t \mathbb{L}\,\mathrm{d}t = \delta\int_0^t\int_0^l \mathfrak{L}\,\mathrm{d}x\,\mathrm{d}t = 0.$$

Thus, we seek the stationary value of a definite integral that depends on two variables. In view of the corresponding Euler's equation, namely, equation (12.11), we can write the stationarity condition as

$$\frac{\partial\mathfrak{L}}{\partial u} - \left[\frac{\partial}{\partial t}\left(\frac{\partial\mathfrak{L}}{\partial u_t}\right) + \frac{\partial}{\partial x}\left(\frac{\partial\mathfrak{L}}{\partial u_x}\right)\right] = 0, \tag{13.27}$$

where $u_t := \partial u/\partial t$ and $u_x := \partial u/\partial x$.

Equation (13.27) is Lagrange's equation of motion for a one-dimensional continuum.

Inserting the Lagrangian density, stated in expression (13.26), into equation (13.27) and considering ρ and k as constants, we obtain

$$-\left[\frac{\partial}{\partial t}\left(\rho\frac{\partial u}{\partial t}\right) - \frac{\partial}{\partial x}\left(k\frac{\partial u}{\partial x}\right)\right] = -\rho\frac{\partial^2 u}{\partial t^2} + k\frac{\partial^2 u}{\partial x^2} = 0.$$

Rearranging, we can write

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{\frac{k}{\rho}}\frac{\partial^2 u}{\partial t^2},$$

which is a one-dimensional wave equation for longitudinal waves in elastic continua, where $\sqrt{k/\rho}$ denotes the speed of propagation with the units of speed resulting from $[k] = \left[kgm/s^2\right]$ and $[\rho] = [kg/m]$.

Note that the solution of the one-dimensional wave equation is surface $u(x,t)$ — in the xt-space — that renders $\iint \mathfrak{L}\mathrm{d}x\mathrm{d}t$ stationary. This illustrates the fact that a solution of Euler's equation involving two variables is a surface, as stated in Section 12.4.1.

The variational approach to the one-dimensional wave equation for transverse waves is shown in Exercise 13.9.

Closing remarks

As shown in this chapter, rays — originally formulated in terms of Hamilton's ray equations (8.19) — coincide with the curves exhibiting stationary traveltime. This property allows us to invoke Fermat's principle and, hence, to study ray theory using the tools of the calculus of variations. In Chapter 14, we will use the stationarity of traveltime to study raytracing techniques.

Variational formulations are equivalent to Hamilton's ray equations provided we can, using Legendre's transformation, write a given ray-theory Hamiltonian as the corresponding ray-theory Lagrangian. This requirement is satisfied for all convex phase-slowness surfaces. As stated in Theorem 10.3, the phase-slowness surface of the fastest wave is convex. Consequently, we can always use Fermat's principle to study the qP wave. When dealing with the qS wave, we must be aware of the inflection points of its phase-slowness surface. The study of such points belongs to the realm of singularity theory, which is not considered in this book.[7]

Exercises

Exercise 13.1 *Consider a traveltime integral in an anisotropic inhomogeneous continuum, namely,*

$$\check{C}=\int\limits_a^b \mathcal{F}(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t. \tag{13.28}$$

[7]Interested readers might refer to Hanyga, A., and Slawinski, M.A., (2001) Caustics in qSV rayfields of transversely isotropic and vertically inhomogeneous media: Anisotropy 2000: Fractures, converted waves, and case studies: SEG (Special Issue), 409 – 418.

Show that if $\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})$ is homogeneous of degree 1 in $\dot{\mathbf{x}}$, the integral is independent of parametrization.

Solution 13.1 *Let $s = f(t)$ be an arbitrary parametrization. Hence,*

$$\mathrm{d}s = \frac{\mathrm{d}f}{\mathrm{d}t}\mathrm{d}t =: \dot{f}\mathrm{d}t,$$

and

$$\dot{f} = \frac{\mathrm{d}s}{\mathrm{d}t}. \tag{13.29}$$

Consider $\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})$, where

$$\dot{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t},$$

which, in view of expression (13.29), can be written as

$$\dot{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\frac{\mathrm{d}f}{\mathrm{d}t} =: \mathbf{x}'\dot{f}.$$

For the value of the integral (13.28) to be independent of parametrization, we require

$$\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t = \mathcal{F}(\mathbf{x},\mathbf{x}')\,\mathrm{d}s. \tag{13.30}$$

Consider the left-hand side of equation (13.30). Since $\dot{\mathbf{x}} = \mathbf{x}'\dot{f}$ and $\mathrm{d}t = \mathrm{d}s/\dot{f}$, we can write it as

$$\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t = \mathcal{F}\left(\mathbf{x},\mathbf{x}'\dot{f}\right)\frac{\mathrm{d}s}{\dot{f}}.$$

If $\mathcal{F}$ is homogeneous of degree 1 in $\dot{\mathbf{x}}$, following Definition A.1, stated in Appendix A, we obtain

$$\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t = \dot{f}\mathcal{F}(\mathbf{x},\mathbf{x}')\frac{\mathrm{d}s}{\dot{f}} = \mathcal{F}(\mathbf{x},\mathbf{x}')\,\mathrm{d}s,$$

which is equation (13.30), as required.

Remark 13.1 *Exercise 13.1 shows that the general statement of Fermat's principle, namely, $\delta\int\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t = 0$, is independent of parametrization. This is the case since $\mathcal{F}(\mathbf{x},\dot{\mathbf{x}})\,\mathrm{d}t = \mathrm{d}s/V$ is homogeneous of degree 1 in $\dot{\mathbf{x}}$. Note that $\mathrm{d}s$ is homogeneous of degree 1 in $\dot{\mathbf{x}}$, while V is homogeneous of degree 0 in $\dot{\mathbf{x}}$.*

Exercise 13.2 [8]*In view of Lemma 13.3, use Beltrami's identity (9.9), namely,*

$$\frac{\partial \mathcal{L}}{\partial t}+\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{n}\dot{x}_i\frac{\partial \mathcal{L}}{\partial \dot{x}_i}-\mathcal{L}\right)=0, \tag{13.31}$$

to show that Lagrangian $\mathcal{L}$ is conserved along the ray.

Solution 13.2 *Since $\mathcal{L}$, given by expression (13.7), does not explicitly depend on t, equation (13.31) becomes $\mathrm{d}\left(\sum_{i=1}^{n}\dot{x}_i\partial\mathcal{L}/\partial\dot{x}_i-\mathcal{L}\right)/\mathrm{d}t=0$, which implies that $\sum_{i=1}^{n}\dot{x}_i\partial\mathcal{L}/\partial\dot{x}_i-\mathcal{L}=C$, where C denotes a constant. Since, by Lemma 13.3, $\mathcal{L}$ is homogeneous of degree 2 in the $\dot{x}_i$, in view of Theorem A.1, we obtain $2\mathcal{L}-\mathcal{L}=C$. Thus, Lagrangian $\mathcal{L}$ is equal to a constant and, hence, it is conserved along the ray.*

Exercise 13.3 *Consider the system of six characteristic equations for an isotropic inhomogeneous continuum, derived in Exercises 8.4 and 8.5, namely,*

$$\left\{\begin{array}{l}\dfrac{\mathrm{d}x_i}{\mathrm{d}t}=v^2p_i\\[3ex] \dfrac{\mathrm{d}p_i}{\mathrm{d}t}=-\dfrac{\partial\ln v}{\partial x_i}\end{array}\right., \qquad i\in\{1,2,3\}. \tag{13.32}$$

Express system (13.32) as three second-order equations.

Solution 13.3 *Solving the first equation of system (13.32) for the components of the phase-slowness vector, we get*

$$p_i=\frac{1}{v^2(\mathbf{x})}\frac{\mathrm{d}x_i}{\mathrm{d}t}, \qquad i\in\{1,2,3\}.$$

Differentiating with respect to t, we can write

$$\frac{\mathrm{d}p_i}{\mathrm{d}t}=\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{v^2(\mathbf{x})}\frac{\mathrm{d}x_i}{\mathrm{d}t}\right], \qquad i\in\{1,2,3\},$$

which we can equate to the second equation of set (13.32) to obtain

$$\frac{\partial\ln v}{\partial x_i}+\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{v^2(\mathbf{x})}\frac{\mathrm{d}x_i}{\mathrm{d}t}\right]=0, \qquad i\in\{1,2,3\}, \tag{13.33}$$

as required.

[8]See also Section 9.1.2

Exercise 13.4 *Consider the traveltime integral in an isotropic inhomogeneous continuum. Show that equations (13.33) are equivalent to a parametric form of Euler's equations.*

Solution 13.4 *Let the integrand of the traveltime integral in an isotropic inhomogeneous continuum be written as*

$$\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}; t) = \frac{\sqrt{\sum_{i=1}^{3} \dot{x}_i \dot{x}_i}}{V(\mathbf{x})},$$

where $\dot{\mathbf{x}} \equiv \mathrm{d}\mathbf{x}/\mathrm{d}t$. *We invoke equations (13.9), namely,*

$$\frac{\partial \mathcal{F}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{F}}{\partial \dot{x}_i}\right) = 0, \qquad i \in \{1, 2, 3\}, \tag{13.34}$$

which are a parametric form of Euler's equations. Considering integrand $\mathcal{F}$ *and the first term of equations (13.34), we obtain*

$$\frac{\partial \mathcal{F}}{\partial x_i} = -\frac{1}{V^2(\mathbf{x})} \frac{\partial V(\mathbf{x})}{\partial x_i} \sqrt{\sum_{i=1}^{3} \dot{x}_i \dot{x}_i}, \qquad i \in \{1, 2, 3\}.$$

Using the fact that t *denotes time, and, hence, as shown in expression (9.12),*

$$\sqrt{\sum_{i=1}^{3} \dot{x}_i \dot{x}_i} =: V(\mathbf{x}), \tag{13.35}$$

where V *is the magnitude of ray velocity, we can write*

$$\frac{\partial \mathcal{F}}{\partial x_i} = -\frac{1}{V(\mathbf{x})} \frac{\partial V(\mathbf{x})}{\partial x_i}, \qquad i \in \{1, 2, 3\}.$$

Using the chain rule, we can rewrite this expression as

$$\frac{\partial \mathcal{F}}{\partial x_i} = -\frac{\partial}{\partial x_i} \ln V(\mathbf{x}), \qquad i \in \{1, 2, 3\}. \tag{13.36}$$

Considering integrand $\mathcal{F}$ *and the second term of equations (13.34), we obtain*

$$\frac{\partial \mathcal{F}}{\partial \dot{x}_i} = \frac{1}{V(\mathbf{x})} \frac{\dot{x}_i}{\sqrt{\sum_{i=1}^{3} \dot{x}_i \dot{x}_i}}, \qquad i \in \{1, 2, 3\},$$

which, using expression (13.35), we can write as

$$\frac{\partial \mathcal{F}}{\partial \dot{x}_i} = \frac{1}{V^2(\mathbf{x})} \frac{\mathrm{d}x_i}{\mathrm{d}t}, \qquad i \in \{1,2,3\}. \tag{13.37}$$

Consequently, using expressions (13.36) and (13.37), we can write Euler's equations (13.34) as

$$\frac{\partial \ln V}{\partial x_i} + \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{V^2(\mathbf{x})} \frac{\mathrm{d}x_i}{\mathrm{d}t}\right] = 0, \qquad i \in \{1,2,3\}. \tag{13.38}$$

Since in isotropic continua, phase and ray velocities coincide, namely, $V \equiv v$, equations (13.38) are equivalent to equations (13.33), as required.

Remark 13.2 *Exercises 13.3 and 13.4 show that the characteristic equations that are the solutions of the eikonal equation in isotropic inhomogeneous continua are tantamount to Euler's equations that provide the stationarity condition for the traveltime of the signal in such continua. In other words, these exercises verify Fermat's principle in isotropic inhomogeneous continua.*[9]

Exercise 13.5 *Recall the classical-mechanics Lagrangian given in expression (13.14), namely,*

$$\mathbb{L} := T - U, \tag{13.39}$$

where T and U are the kinetic and potential energies. Let the classical-mechanics Hamiltonian be

$$\mathbb{H} := T + U. \tag{13.40}$$

Using the standard expression for kinetic energy and letting p_i be a component of linear momentum, verify Legendre's transformation between $\mathbb{L}$ and $\mathbb{H}$.

Solution 13.5 *In view of Legendre's transformation, discussed in Appendix B, we can write*

$$\mathbb{H} = \sum_{i=1}^{n} p_i \dot{x}_i - \mathbb{L},$$

[9]Readers interested in a formulation linking rays and Fermat's principle in isotropic inhomogeneous continua might also refer to Elmore, W.C., and Heald, M.A., (1969/1985) Physics of waves: Dover, pp. 320 – 322.

where $p_i = mv_i$, *with* v_i *being a component of velocity given by* $v_i = \mathrm{d}x_i/\mathrm{d}t \equiv \dot{x}_i$. *Hence, we can write*

$$\mathbb{H} = m\sum_{i=1}^{n}\dot{x}_i^2 - \mathbb{L} = mv^2 - \mathbb{L}, \tag{13.41}$$

where v *stands for the magnitude of velocity. Recalling definitions (13.39) and (13.40), we can write expression (13.41) as*

$$T + U = mv^2 - (T - U),$$

where T *and* U *are the kinetic and potential energies, respectively. Simplifying, we obtain*

$$T = \frac{1}{2}mv^2,$$

which is the standard expression for kinetic energy.

Exercise 13.6 *Given Newton's second law of motion, stated as a single second-order ordinary differential equation, namely,*

$$m\frac{\mathrm{d}^2x_i}{\mathrm{d}t^2} = -\frac{\partial U(\mathbf{x})}{\partial x_i}, \qquad i \in \{1,2,3\}, \tag{13.42}$$

where $U(\mathbf{x})$ *denotes the scalar potential, write the corresponding two first-order ordinary differential equations in* t *to be solved for the* x_i *and the* p_i, *where* p_i *is a component of the linear momentum.*

Solution 13.6 *We can denote the components of the momentum vector as*

$$p_i := m\frac{\mathrm{d}x_i}{\mathrm{d}t}, \qquad i \in \{1,2,3\}. \tag{13.43}$$

Differentiating both sides of equations (13.43) with respect to t, *we obtain*

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = m\frac{\mathrm{d}^2x_i}{\mathrm{d}t^2}, \qquad i \in \{1,2,3\},$$

which are equations (13.42). Hence, Newton's second law of motion can be written as a set of two first-order differential equations,

$$\begin{cases} \dfrac{\mathrm{d}x_i}{\mathrm{d}t} = \dfrac{p_i}{m} \\[2ex] \dfrac{\mathrm{d}p_i}{\mathrm{d}t} = -\dfrac{\partial U}{\partial x_i} \end{cases}, \qquad i \in \{1,2,3\}. \tag{13.44}$$

Exercise 13.7 [10] *Using expression (13.40), show that equations (13.44), obtained in Exercise 13.6, correspond to Hamilton's equations of motion that are given by*

$$\left\{\begin{array}{l}\dfrac{\mathrm{d}x_i}{\mathrm{d}t}=\dfrac{\partial \mathbb{H}}{\partial p_i}\\[2ex]\dfrac{\mathrm{d}p_i}{\mathrm{d}t}=-\dfrac{\partial \mathbb{H}}{\partial x_i}\end{array}\right., \qquad i\in\{1,2,3\}.$$

Solution 13.7 *Consider expression (13.40). Using the standard expression for the kinetic energy, as well as the definition of linear momentum, we can write this expression as*

$$\mathbb{H}=T+U=\frac{1}{2}m\left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\right)^2+U=\frac{1}{2m}p_i^2+U, \qquad i\in\{1,2,3\}. \tag{13.45}$$

Differentiating equations (13.45) with respect to both the p_i and the x_i, we obtain

$$\left\{\begin{array}{l}\dfrac{\partial \mathbb{H}}{\partial p_i}=\dfrac{p_i}{m}\\[2ex]\dfrac{\partial \mathbb{H}}{\partial x_i}=\dfrac{\partial U}{\partial x_i}\end{array}\right., \qquad i\in\{1,2,3\}.$$

Using Newton's second law of motion, which is stated in expression (13.44), we obtain

$$\left\{\begin{array}{l}\dfrac{\mathrm{d}x_i}{\mathrm{d}t}=\dfrac{\partial \mathbb{H}}{\partial p_i}\\[2ex]\dfrac{\mathrm{d}p_i}{\mathrm{d}t}=-\dfrac{\partial \mathbb{H}}{\partial x_i}\end{array}\right., \qquad i\in\{1,2,3\},$$

which are Hamilton's equations of motion, as required.

Exercise 13.8 *Considering a free-falling body in the vacuum, show that Hamilton's principle is consistent with Newton's concept of acceleration due to gravity.*

Solution 13.8 *Let $T=mv^2/2$ and $U=mgz$, where m is mass, v is velocity, g is acceleration due to gravity and z denotes height. Since $v=\mathrm{d}z/\mathrm{d}t$, we can write the classical-mechanics Lagrangian as*

$$\mathbb{L}=T-U=\frac{1}{2}m\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2-mgz.$$

[10] See also Section 13.2.2

Consider the action integral

$$\mathbb{A} = \int_{t_0}^{t_1} \mathbb{L}\,\mathrm{d}t.$$

Hamilton's principle implies

$$\delta \int_{t_0}^{t_1} (T - U)\;\mathrm{d}t = \delta \int_{t_0}^{t_1} \left(\frac{1}{2} m\,(\dot{z})^2 - mgz\right)\,\mathrm{d}t = 0,$$

where $\dot{z} := \mathrm{d}z/\mathrm{d}t$ *. Invoking Euler's equation, which corresponds to Lagrange's equations of motion (13.16), we obtain*

$$\frac{\partial \mathbb{L}}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbb{L}}{\partial \dot{z}}\right) = -mg - \frac{\mathrm{d}}{\mathrm{d}t}\,(m\dot{z}) = -mg - m\ddot{z} = 0,$$

which can be written as

$$\frac{\mathrm{d}^2 z\,(t)}{\mathrm{d}t^2} = -g,$$

where g *is the free-fall acceleration, as required.*

Exercise 13.9 *Following Section 13.2.3, derive a one-dimensional wave equation for transverse waves.*

Solution 13.9 *Let the transverse displacements of masses be* $u_0, \ldots, u_{n+1}$*, with* $u_0 = u_{n+1} = 0$*, which are boundary conditions corresponding to fixed ends. The potential energy,* U*, is associated with the tension,* μ*, of the string. The potential energy per segment is*

$$\mathrm{d}U = \mu\left[\sqrt{(\Delta x)^2 + (u_i - u_{i-1})^2} - \Delta x\right], \tag{13.46}$$

where the term in parentheses constitutes the extension of the segment Δx*, which is the difference between its original length,* Δx*, and its strained length,* $\sqrt{(\Delta x)^2 + (u_i - u_{i-1})^2}$*. We can rewrite expression (13.46) as*

$$\mathrm{d}U = \mu \Delta x\left[\sqrt{1 + \left(\frac{u_i - u_{i-1}}{\Delta x}\right)^2} - 1\right].$$

Expanding the square root as a power series gives us

$$\mathrm{d}U = \mu\Delta x\left[\frac{1}{2}\left(\frac{u_i - u_{i-1}}{\Delta x}\right)^2 - \frac{1}{8}\left(\frac{u_i - u_{i-1}}{\Delta x}\right)^4 + \ldots\right].$$

Assuming that the term in parentheses is much smaller than unity, which implies that the transverse displacement is much smaller than the length of a segment, we obtain

$$\mathrm{d}U \approx \frac{\mu}{2}\Delta x\left(\frac{u_i - u_{i-1}}{\Delta x}\right)^2.$$

Thus, the potential energy for the entire string is

$$U \approx \frac{1}{2}\mu\sum_{i=1}^{n}\left(\frac{u_i - u_{i-1}}{\Delta x}\right)^2\Delta x.$$

Letting $n \longrightarrow \infty$ and $\Delta x \longrightarrow 0$ in such a way that $n\Delta x = l$, where l is the length of the string, and noticing that the term in parentheses represents a partial derivative with respect to x, we can write

$$U = \frac{1}{2}\mu\sum_{i=1}^{\infty}\left[\frac{\partial u\left(x_i, t\right)}{\partial x}\right]^2\Delta x.$$

Thus, in the limit, we obtain

$$U = \frac{1}{2}\mu\int\limits_0^l\left[\frac{\partial u\left(x, t\right)}{\partial x}\right]^2\mathrm{d}x.$$

The kinetic energy is given by expression (13.23), namely,

$$T = \frac{1}{2}\rho\int\limits_0^l\left[\frac{\partial u\left(x, t\right)}{\partial t}\right]^2\mathrm{d}x.$$

Thus, using $\mathbb{L}$, given in expression (13.14), we can write

$$\mathbb{L} = T - U = \int\limits_0^l\left[\frac{\rho}{2}\left(\frac{\partial u}{\partial t}\right)^2 - \frac{\mu}{2}\left(\frac{\partial u}{\partial x}\right)^2\right]\mathrm{d}x.$$

Invoking Hamilton's principle, stated in expression (13.15), we obtain

$$\delta \int_0^t \mathbb{L}\,\mathrm{d}t = \delta \int_0^t \int_0^l \left[\frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2 - \frac{\mu}{2} \left(\frac{\partial u}{\partial x} \right)^2 \right] \mathrm{d}x\,\mathrm{d}t = 0.$$

Using the corresponding Euler's equation (12.11), we get

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 u}{\partial t^2}. \tag{13.47}$$

Equation (13.47) is a one-dimensional wave equation for transverse waves where the transverse displacement, u, is assumed to be much smaller than the length of the string, l.

This Page Intentionally Left Blank

Chapter 14

Ray parameters

> En général la conservation des forces vives donne toujours une intégrale première des différentes équations différentielles de chaque problème; ce qui est d'une grande utilité dans plusieurs occasions.[1]
>
> *Joseph-Louis Lagrange (1788) Mécanique Analytique*

Preliminary remarks

In the context of ray theory, the trajectories of seismic signals as well as their traveltimes can be obtained by solving Hamilton's ray equations or Lagrange's ray equations, discussed in Chapters 8 and 9, respectively. In certain cases, particular properties of the continuum result in simplifications of these equations, thereby allowing us to obtain their solutions more easily, as well as to gain further insight into these solutions.

In Chapter 13, we showed that rays are the solutions of the variational problem of stationary traveltime and, hence, they are the solutions of the corresponding Euler's equations. For the continua that exhibit particular homogeneities, Euler's equations can be simplified by obtaining the corresponding first integrals, which were introduced in Section 12.6. First integrals are the conserved quantities. In ray theory, these quantities, which are constant along a given ray, are called ray parameters.

We begin this chapter, in which we study only two-dimensional continua, with the formulation of the ray parameter for an anisotropic continuum that

[1] In general, the conservation of living forces yields always a first integral of various differential equations of each problem; this is of great utility on numerous occasions.

is homogeneous along one axis. By integrating the ray-parameter expression, we obtain the expression for the ray. Also, using the ray parameter, we obtain the expression for the traveltime. Then we briefly discuss a case in which ray equations do not possess corresponding ray parameters. We conclude this chapter by discussing the conserved quantities in the context of Hamilton's ray equations.

14.1 Traveltime integrals

Let us consider a two-dimensional continuum that is contained in the xz-plane. The traveltime between two points A and B within this continuum can be stated as

$$\check{C} = \int_A^B \frac{\sqrt{1+(z')^2}}{V(x,z,z')}\,\mathrm{d}x \equiv \int_A^B \mathrm{F}\,\mathrm{d}x, \tag{14.1}$$

where $z' := \mathrm{d}z/\mathrm{d}x$. Since $\mathrm{d}z/\mathrm{d}x = \cot\theta$, where θ is the ray angle, we see that the ray velocity, V, is a function of position, (x,z), and direction, z'. In other words, integral (14.1) allows us to study traveltimes in anisotropic inhomogeneous continua.

Also, let us view the x-axis and the z-axis as the horizontal and vertical axes, respectively, where the vertical axis corresponds to depth within a geological model.

In view of Fermat's principle, discussed in Chapter 13, rays correspond to curves along which the traveltime is stationary. Since integral (14.1) is of the type given by integral (12.1), in general, we can obtain such a curve using Euler's equation (12.4). Consequently, F is a ray-theory Lagrangian.

As discussed in Section 12.5, a particular form of the integral, whose stationary value we seek, may result in simplifications of Euler's equation. Herein, we wish to study special cases of traveltime integral (14.1) that are pertinent to seismic investigations.

14.2 Ray parameters as first integrals

[2]In this section, we will study horizontally layered media. In such a case, where the ray velocity, V, may vary with depth, z, and direction, z', travel-

[2]This section is based on Slawinski, M.A., and Webster, P.S., (1999) On generalized ray parameters for vertically inhomogeneous and anisotropic media, Canadian Journal of Exploration Geophysics, **35**, No. 1/2, 28 – 31.

time integral (14.1) becomes

$$\check{C}=\int_A^B \frac{\sqrt{1+(z')^2}}{V(z,z')}\,\mathrm{d}x \equiv \int_A^B \mathrm{F}(z,z')\,\mathrm{d}x. \tag{14.2}$$

Since traveltime integral (14.2) does not exhibit an explicit dependence on x, to obtain the ray, we use Beltrami's identity (12.10), namely,

$$\frac{\partial \mathrm{F}}{\partial x}+\frac{\mathrm{d}}{\mathrm{d}x}\left(z'\frac{\partial \mathrm{F}}{\partial z'}-\mathrm{F}\right)=0,$$

which immediately leads to

$$z'\frac{\partial \mathrm{F}}{\partial z'}-\mathrm{F}=C, \tag{14.3}$$

where C is a constant. Expression (14.3) is first integral (12.19) and C is a conserved quantity along the ray.

We wish to study this conserved quantity. Inserting integrand F, given in integral (14.2), into expression (14.3), we obtain

$$z'\frac{\partial\left(\frac{\sqrt{1+(z')^2}}{V(z,z')}\right)}{\partial z'}-\frac{\sqrt{1+(z')^2}}{V(z,z')}=-\frac{V'z'\sqrt{1+(z')^2}}{V^2}-\frac{1}{V\sqrt{1+(z')^2}}=C, \tag{14.4}$$

where, for convenience, we denote $V:=V(z,z')$ and $V':=\partial V/\partial z'$. The chain rule implies $-V'/V^2=\partial\left[1/V(z,z')\right]/\partial z'$ and, hence, we get

$$\frac{\partial}{\partial z'}\left(\frac{1}{V}\right)z'\sqrt{1+(z')^2}-\frac{1}{V\sqrt{1+(z')^2}}=C. \tag{14.5}$$

Expression (14.5) is a first integral of Euler's equation (12.4) for traveltime integral (14.2).

In order to express the first integral in terms of the ray angle, we use $z'\equiv \mathrm{d}z/\mathrm{d}x=\cot\theta$, where θ is the ray angle measured from the z-axis. Also, the differential operator in expression (14.5) can be restated as $\partial/\partial z'=(\partial\theta/\partial z')\,\partial/\partial\theta$. Hence, using trigonometric identities, we obtain another form of expression (14.5) given by

$$\mathfrak{p}=\cos\theta\frac{\partial}{\partial\theta}\left[\frac{1}{V(z,\theta)}\right]+\frac{\sin\theta}{V(z,\theta)}, \tag{14.6}$$

where $\mathfrak{p} = -C$ and where V and θ denote ray velocity and ray angle, respectively. Since $\mathfrak{p}$ is conserved along a given ray, $z(x)$, we refer to this conserved quantity as ray parameter. Expression (14.6) is the ray parameter for anisotropic vertically inhomogeneous continua.

For expression (14.6) to be valid, the ray velocity may vary along the z-axis but not along the x-axis. The directional dependence of velocity, however, need not exhibit any particular symmetry. In other words, the angular velocity dependence is arbitrary.

Note that in the context of elasticity theory, the availability of exact and explicit ray-velocity expressions $V(\theta)$ is limited due to the requirements of Legendre's transformation. An explicit, closed-form expression for $V(\theta)$ is only possible for the case of elliptical velocity dependence.

14.3 Example: Ellipticity and linearity

Introductory comments

In this section, we study a particular case of wave propagation that is associated with both an elliptical velocity dependence with direction and a linear velocity dependence with depth. This assumption allows us to obtain analytic expressions for rays and traveltimes.

Since Euler's equation (12.4), or its Beltrami's identity (12.10), is a second-order ordinary differential equation, in view of Definition 12.1, first integral (14.5) and ray parameter (14.6) are first-order ordinary differential equations. If the integration of the ray parameter is possible, this integration results in a solution of Euler's equation and its Beltrami's identity, which can be given by $z(x)$ or $x(z)$. In other words, the expressions for ray velocity, V, that result in integrable expression (14.6) allow us to obtain rays by integration.[3] Ray velocity that results in a conveniently integrable ray parameter is provided by the case of elliptical velocity dependence with direction and linear velocity dependence with depth.

[3] For certain cases with applications to continua exhibiting folded layers, readers might refer to Epstein, M., and Slawinski, M.A., (1999) On rays and ray parameters in inhomogeneous isotropic media. Canadian Journal of Exploration Geophysics. **35**, No. 1/2, 7 – 19.

14.3.1 Rays

Derivation

To obtain an analytic expression for a ray, we wish to use an exact ray-velocity expression to be inserted into expression (14.6). For this purpose, we consider expression (9.39), namely,

$$V(\theta) = V_z \sqrt{\frac{1+\tan^2\theta}{1+\left(\frac{V_z}{V_x}\right)^2 \tan^2\theta}}, \tag{14.7}$$

which gives the magnitude of the ray velocity as a function of the ray angle for the case of elliptical velocity dependence. For convenience, let the measure of ellipticity be given by

$$\chi := \frac{V_x^2 - V_z^2}{2V_z^2}, \tag{14.8}$$

where V_x and V_z stand for the magnitude of the horizontal and the vertical ray velocities, respectively.

Using χ, we can write expression (14.7) as

$$V(\theta) = V_z \sqrt{\frac{1+2\chi}{1+2\chi\cos^2\theta}}.$$

Let us assume that the ray velocity varies along the z-axis in such a way that χ remains constant. This implies that the ratio of magnitudes of horizontal and vertical ray velocities remains constant. In such a case, we can write

$$V(\theta, z) = V_z(z) \sqrt{\frac{1+2\chi}{1+2\chi\cos^2\theta}}. \tag{14.9}$$

Furthermore, we assume that the magnitude of the ray velocity increases linearly along the z-axis.[4] In such a case, we can write expression (14.9) as

$$V(\theta, z) = (a + bz) \sqrt{\frac{1+2\chi}{1+2\chi\cos^2\theta}}, \tag{14.10}$$

[4]Readers interested in a seismological formulation of linearly increasing velocity might refer to Epstein, M., and Slawinski, M.A., (1999) On raytracing in constant velocity-gradient media: Geometrical approach, Canadian Journal of Exploration Geophysics. **35**, No. 1/2, 1 – 6, and to Slawinski, R.A., and Slawinski, M.A., (1999) On raytracing in constant velocity-gradient media: Calculus approach, Canadian Journal of Exploration Geophysics. **35**, No. 1/2, 24 – 27.

where a and b are positive constants.

Inserting expression (14.10) into expression (14.6), we obtain

$$\begin{aligned}\mathfrak{p} &= \cos\theta\frac{\partial}{\partial\theta}\left[\frac{1}{(a+bz)\sqrt{\frac{1+2\chi}{1+2\chi\cos^2\theta}}}\right]+\frac{\sin\theta}{(a+bz)\sqrt{\frac{1+2\chi}{1+2\chi\cos^2\theta}}}\\ &= \frac{1}{(a+bz)\sqrt{1+2\chi}}\left(\cos\theta\frac{\partial}{\partial\theta}\sqrt{1+2\chi\cos^2\theta}+\sin\theta\sqrt{1+2\chi\cos^2\theta}\right)\\ &= \frac{\sin\theta}{(a+bz)\sqrt{1+2\chi}\sqrt{1+2\chi\cos^2\theta}}.\end{aligned}\tag{14.11}$$

To obtain an expression for a ray, we wish to state ray parameter (14.11) in terms of position variables x and z. Dividing both the numerator and the denominator by $\sin\theta$, we rewrite expression (14.11) as

$$\mathfrak{p} = \frac{1}{(a+bz)\sqrt{1+2\chi}\sqrt{1+\frac{2\chi+1}{\tan^2\theta}}}.$$

Squaring both sides and rearranging, we obtain

$$\frac{1}{\tan^2\theta} = \frac{1-\mathfrak{p}^2(a+bz)^2(1+2\chi)}{\mathfrak{p}^2(a+bz)^2(1+2\chi)^2}.\tag{14.12}$$

Since $1/\tan^2\theta = (\mathrm{d}z/\mathrm{d}x)^2$, we have a first-order ordinary differential equation, which is a special case of first integral (12.19). We can write equation (14.12) as

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\sqrt{1-\mathfrak{p}^2(a+bz)^2(1+2\chi)}}{\mathfrak{p}(a+bz)(1+2\chi)},\tag{14.13}$$

which can be restated as

$$\mathrm{d}x = \frac{\mathfrak{p}(a+bz)(1+2\chi)}{\sqrt{1-\mathfrak{p}^2(a+bz)^2(1+2\chi)}}\mathrm{d}z.\tag{14.14}$$

To integrate equation (14.14), we set the initial conditions in such a way that $z(0) = 0$. In other words, the source is located at the origin of the coordinate system. Hence, integrating both sides, namely,

$$\int_0^x \mathrm{d}\xi = \int_0^z \frac{\mathfrak{p}(a+b\zeta)(1+2\chi)}{\sqrt{1-\mathfrak{p}^2(a+b\zeta)^2(1+2\chi)}}\mathrm{d}\zeta,$$

where ξ and ζ are the integration variables, we obtain

$$x = \frac{1}{\mathfrak{p}b}\left[\sqrt{1-\mathfrak{p}^2a^2(1+2\chi)} - \sqrt{1-\mathfrak{p}^2(a+bz)^2(1+2\chi)}\right], \tag{14.15}$$

which is the ray given by $x(z)$.

Interpretation

To obtain a geometrical interpretation of equation (14.15), we rearrange it and write

$$\frac{\left(x - \dfrac{\sqrt{1-\mathfrak{p}^2a^2(1+2\chi)}}{\mathfrak{p}b}\right)^2}{\left(\dfrac{1}{\mathfrak{p}b}\right)^2} + \frac{\left(z+\dfrac{a}{b}\right)^2}{\left(\dfrac{1}{\mathfrak{p}b\sqrt{1+2\chi}}\right)^2} = 1. \tag{14.16}$$

This is the equation of an ellipse whose axes are parallel to the axes of the coordinate system with the origin of this system located at the source. The centre of the ellipse is located at

$$\left[\frac{\sqrt{1-\mathfrak{p}^2a^2(1+2\chi)}}{\mathfrak{p}b}, -\frac{a}{b}\right]. \tag{14.17}$$

In a seismological context, where the vertical z-axis points downwards, the centre of the ellipse is located on the horizontal line positioned $-a/b$ units above the x-axis. Ellipse (14.16) passes through the origin, as can be verified by setting $z = 0$ in equation (14.15). The portion of the ellipse that is below the x-axis corresponds to the ray.

The greater the distance between the source and the centre of the ellipse, the smaller the curvature of the ray. For constant-velocity fields, where $b = 0$, the centre of the ellipse is located infinitely far from the source. Consequently, the ray is a straight line. For a signal propagating along the z-axis, $\theta = 0$ and, following expression (14.11), $\mathfrak{p} = 0$. Hence, in view of expression (14.17), the x-coordinate of the centre of the ellipse is located infinitely far from the source. Consequently, the ray is a vertical straight line. For the isotropic case, where $\chi = 0$, equation (14.16) reduces to the expression for a circle. Consequently the rays are circular arcs.

14.3.2 Traveltimes

We can also use ray parameter (14.6) to obtain the traveltime along the corresponding ray. For this purpose, we wish to rewrite integral (14.2) to include the ray parameter for a given source-receiver pair.

Integral (14.2) can be viewed as $\int \mathrm{d}s/V$, where $\mathrm{d}s$ is the arclength element along the ray. In the xz-plane, the arclength element can be written as $\mathrm{d}s = \mathrm{d}z/\cos\theta$, where θ is the ray angle. Hence, traveltime integral (14.2) between the source at $(0,0)$ and the receiver at (X,Z) is

$$\check{C} = \int\limits_0^Z \frac{\mathrm{d}z}{V(z,\theta)\cos\theta}, \tag{14.18}$$

where $V(z,\theta)$ is given by expression (14.10). Hence, we can explicitly write

$$\check{C} = \int\limits_0^Z \frac{\mathrm{d}z}{(a+bz)\sqrt{\dfrac{1+2\chi}{1+2\chi\cos^2\theta}}\cos\theta}. \tag{14.19}$$

To integrate, we must express $\cos\theta$ in terms of constants, a, b, χ, $\mathfrak{p}$, and integration variable z. Using expression (14.11) and trigonometric identities, we obtain

$$\cos\theta = \sqrt{\frac{1-\mathfrak{p}^2(a+bz)^2(1+2\chi)}{1+2\chi\mathfrak{p}^2(a+bz)^2(1+2\chi)}}. \tag{14.20}$$

Inserting expression (14.20) into integral (14.19), after algebraic manipulation, we obtain

$$\check{C} = \int\limits_0^Z \frac{\mathrm{d}z}{(a+bz)\sqrt{1-\mathfrak{p}^2(a+bZ)^2(1+2\chi)}}. \tag{14.21}$$

Integrating between $z=0$ and $z=Z$, while treating $\mathfrak{p}$ as a constant, we obtain the expression for the value of the traveltime, namely,

$$\check{C} = \frac{1}{b}\ln\left[\frac{a+bZ}{a}\,\frac{1+\sqrt{1-\mathfrak{p}^2a^2(1+2\chi)}}{1+\sqrt{1-\mathfrak{p}^2(a+bZ)^2(1+2\chi)}}\right]. \tag{14.22}$$

Note that we can treat $\mathfrak{p}$ as a constant since, for a given source-receiver pair in a laterally homogeneous continuum, $\mathfrak{p}$ is a conserved quantity along the ray.

To find the expression for $\mathfrak{p}$ that corresponds to the source at $(0,0)$ and the receiver at (X,Z), we can write expression (14.15) as

$$X=\frac{1}{\mathfrak{p}b}\left[\sqrt{1-\mathfrak{p}^2a^2(1+2\chi)}-\sqrt{1-\mathfrak{p}^2(a+bZ)^2(1+2\chi)}\right]. \tag{14.23}$$

Solving expression (14.23) for $\mathfrak{p}$, we obtain

$$\mathfrak{p}=\frac{2X}{\sqrt{\left[X^2+(1+2\chi)Z^2\right]\left[(2a+bZ)^2(1+2\chi)+b^2X^2\right]}}. \tag{14.24}$$

Thus, studying the properties of the continuum in terms of a, b and χ, we can use expression (14.22), with $\mathfrak{p}$ given by expression (14.24), to obtain the traveltime between the source and the receiver. These expressions can be conveniently used for inverse problems that are based on traveltime measurements.

14.3.3 Isotropic extension

As shown in Section 6.6.2, by using a linear transformation of coordinates, we can treat elliptical velocity dependence as an isotropic case. Consequently, we can also obtain expressions (14.22) and (14.24) by the following method.

Consider an elliptical velocity dependence with magnitudes of the horizontal and vertical velocities given by

$$v_x=a\sqrt{1+2\chi}, \tag{14.25}$$

where χ is given by expression (14.8), and

$$v_z=a, \tag{14.26}$$

respectively.

Note that, since v_x and v_z are the magnitudes of velocities along the symmetry axes, expressions (14.25) and (14.26) are the same for both phase and ray velocities.

The wavefronts resulting from a point source are ellipses with axes tv_x and tv_z, where t is the traveltime. We can write such a wavefront as

$$\frac{x^2}{t^2v_x^2}+\frac{z^2}{t^2v_z^2}=1,$$

which, using expressions (14.25) and (14.26), we can rewrite as

$$\frac{x^2}{1+2\chi}+z^2=t^2a^2. \tag{14.27}$$

Since v_x and v_z are the magnitudes of velocities along the x-axis and the z-axis, respectively, we can scale the z-axis by a factor of $\sqrt{1+2\chi}$ to obtain circular wavefronts, which correspond to an isotropic case. In other words, we transform the xz-plane into the $x\zeta$-plane, where

$$\zeta = z\sqrt{1+2\chi}. \tag{14.28}$$

Thus, in view of expression (14.28), we let $z = \zeta/\sqrt{1+2\chi}$ to write expression (14.27) as

$$x^2+\zeta^2 = t^2\alpha^2,$$

where

$$\alpha = a\sqrt{1+2\chi} \tag{14.29}$$

is the velocity in the $x\zeta$-plane. Let us also assume that the magnitude of velocity increases linearly along the ζ-axis, namely, $v(\zeta) = \alpha + b\zeta$.

Note that the units of b are $[1/s]$. Consequently, its value does not depend on the scaling of position coordinates.

Dealing with an isotropic case in the $x\zeta$-plane, we can derive the traveltime expression between the source at $(0,0)$ and the receiver at (X,Ξ), to obtain[5]

$$\breve{C} = \frac{1}{b}\ln\left[\frac{\alpha+b\Xi}{\alpha}\frac{1+\sqrt{1-\mathfrak{p}^2\alpha^2}}{1+\sqrt{1-\mathfrak{p}^2(\alpha+b\Xi)^2}}\right], \tag{14.30}$$

where

$$\mathfrak{p} = \frac{2X}{\sqrt{(X^2+\Xi^2)\left[(2\alpha+b\Xi)^2+b^2X^2\right]}}. \tag{14.31}$$

Substituting expression (14.29) into expressions (14.30) and (14.31), as well as — in view of expression (14.28) — letting $\Xi = Z\sqrt{1+2\chi}$, we obtain expressions (14.22) and (14.24), as expected.

14.4 Rays in isotropic continua

In Sections 14.2 and 14.3, we studied ray equations in two-dimensional anisotropic continua and obtained analytical expressions for rays and traveltimes. The availability of analytical expressions resulted from the assumption of homogeneity along the x-axis and, hence, from the existence of a first

[5]Readers interested in details of this derivation might refer to Slotnick, M.M., (1959) Lessons in seismic computing: Society of Exploration Geophysicists, Lesson 37.

integral. In this section, to emphasize the convenience of first integrals, we look briefly at ray equations in a two-dimensional isotropic continuum that is contained in the xz-plane. The traveltime between two points A and B within this continuum can be stated as

$$\check{C} = \int_A^B \frac{\sqrt{1 + (z')^2}}{V(x, z)} \, \mathrm{d}x. \tag{14.32}$$

Since the continuum is isotropic, V is not a function of z'. However, the integrand is an explicit function of x, z and z', and, hence, the corresponding Euler's equation does not have a first integral.

In view of the stationarity of traveltime and Section 12.5.5, the corresponding ray equation, which results from Euler's equation (12.4), is given by equation (12.23), namely,

$$V\frac{\mathrm{d}^2 z}{\mathrm{d}x^2} - \frac{\partial V}{\partial x}\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^3 + \frac{\partial V}{\partial z}\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right)^2 - \frac{\partial V}{\partial x}\frac{\mathrm{d}z}{\mathrm{d}x} + \frac{\partial V}{\partial z} = 0, \tag{14.33}$$

where, due to the isotropy of the continuum, phase and ray velocities coincide, namely, $v \equiv V$. Equation (14.33) is a nonlinear ordinary differential equation, which requires numerical methods to obtain rays and corresponding traveltimes.

14.5 Lagrange's ray equations in xz-plane

In this chapter, as well as in Chapter 12, Euler's equations and Lagrange's ray equations are formulated in the context of explicit functions. Such a formulation is convenient for many raytracing applications. It rules out, however, complicated rays that are given by multiple-valued functions. To generalize the formulation so as to allow such rays, we can formulate our problem in a parametric form.

Consider traveltime integral (14.1). An analogous parametric representation can be given in terms of $x(t)$, $z(t)$, $\dot{x} := \mathrm{d}x/\mathrm{d}t$ and $\dot{z} := \mathrm{d}z/\mathrm{d}t$. Then, the traveltime integral is

$$\check{C} = \int \frac{\mathrm{d}s}{V} = \int \frac{\sqrt{\dot{x}^2 + \dot{z}^2}}{V(x, z, \dot{x}, \dot{z})} \, \mathrm{d}t := \int \mathcal{F} \, \mathrm{d}t, \tag{14.34}$$

where $\mathcal{F}$ is a two-dimensional form of expression (13.10).

In view of the principle of stationary traveltime, we can use Lagrange's ray equations (13.9). In the two-dimensional case, discussed herein, these equations constitute the system

$$\left\{\begin{array}{l}\dfrac{\partial\mathcal{F}}{\partial x}-\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial\mathcal{F}}{\partial\dot{x}}\right)=0\\[2ex]\dfrac{\partial\mathcal{F}}{\partial z}-\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial\mathcal{F}}{\partial\dot{z}}\right)=0\end{array}\right., \tag{14.35}$$

where $\mathcal{F}$ denotes the integrand of the traveltime integral.

Also, the equations of system (14.35) are related by Beltrami's identity, namely,

$$\frac{\partial\mathcal{F}}{\partial t}+\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}\frac{\partial\mathcal{F}}{\partial\dot{x}}+\dot{z}\frac{\partial\mathcal{F}}{\partial\dot{z}}-\mathcal{F}\right)=0. \tag{14.36}$$

The justification for this form of Beltrami's identity is shown in Exercise 9.1.

In view of Theorem A.1, stated in Appendix A, $\mathcal{F}$ cannot depend explicitly on parameter t. Mathematically, we can justify this corollary in the following way.

Since ray-velocity function, V, is homogeneous of degree 0 in the variables $\dot{x}$ and $\dot{z}$ and $\sqrt{\dot{x}^2+\dot{z}^2}$ is absolute-value homogeneous of degree 1 in the same variables, the integrand of the traveltime integral is absolute-value homogeneous of degree 1 in these variables. Thus, since $\mathcal{F}$ is absolute-value homogeneous of degree 1, it follows from Theorem A.1 that

$$\mathcal{F}=\dot{x}\frac{\partial\mathcal{F}}{\partial\dot{x}}+\dot{z}\frac{\partial\mathcal{F}}{\partial\dot{z}}. \tag{14.37}$$

Consequently, the term in parentheses of Beltrami's identity (14.36) vanishes and equation (14.36) implies that $\mathcal{F}$ cannot depend explicitly on t, and, hence, V does not explicitly depend on t, which justifies our corollary.

Physically, this independence means that the ray-velocity function does not change with time. In other words, the properties of the continuum are time-invariant.

Also, the parametric formulation of the traveltime integral conveniently allows us to obtain ray parameters. Consider system (14.35). If $\mathcal{F}$ is not explicitly dependent on x, the first equation becomes $\partial\mathcal{F}/\partial\dot{x}=\mathfrak{p}$, where $\mathfrak{p}$ is a conserved quantity. This conserved quantity is equivalent to ray-parameter expression (14.6), as shown in Exercise 14.3.

14.6 Conserved quantities and Hamilton's ray equations

In this chapter, we study the conserved quantities along the ray in the context of the calculus of variations. In other words, we study these quantities using the Lagrangian formulation of the ray theory. In view of the fact that we can study ray theory in terms of both the Hamiltonian and Lagrangian formulations, let us briefly look at the conserved quantities in terms of Hamilton's ray equations (8.19), namely,

$$\left\{\begin{array}{l} \dot{x}_i = \dfrac{\partial \mathcal{H}}{\partial p_i} \\ \\ \dot{p}_i = -\dfrac{\partial \mathcal{H}}{\partial x_i} \end{array}\right., \qquad i \in \{1,2,3\}.$$

Examining the second equation, we see that if $\mathcal{H}$ does not explicitly depend on the x_i, the corresponding p_i is constant along solution curve $\mathbf{x}(t)$, since $\mathrm{d}p_i/\mathrm{d}t = 0$. To elucidate the consequences of this statement, recall expression (8.20), namely,

$$\mathcal{H} = \frac{1}{2}p^2v^2(\mathbf{x},\mathbf{p}).$$

We see that $\mathcal{H}$ does not explicitly depend on the x_i if and only if function v does not depend on the x_i coordinate. Since phase velocity, v, is a function of the properties of the continuum, we conclude that $\mathcal{H}$ does not depend on the x_i if and only if the continuum is homogeneous along the x_i-axis.

Also, in view of Lagrange's ray equations (9.8), namely,

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right) = 0, \qquad i \in \{1,2,3\}, \tag{14.38}$$

if $\mathcal{L}$ does not explicitly depend on the x_i, the equation of system (14.38) that corresponds to the given subscript i is reduced to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right) = 0, \tag{14.39}$$

which implies that the term in parentheses of expression (14.39) is constant. In view of Legendre's transformation, following expressions (B.7), shown in Appendix B, we can write

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i}, \tag{14.40}$$

where $\mathfrak{p} = p_i$ is the conserved quantity along the solution curve $\mathbf{x}(t)$.

Expressions given by $\mathrm{d}p_i/\mathrm{d}t = 0$ and $\mathrm{d}(\partial\mathcal{L}/\partial\dot{x}_i)/\mathrm{d}t = 0$, formulated in terms of Hamilton's and Lagrange's ray equations, respectively, result from the homogeneity of the continuum along the x_i-axis. These are different formulations of the same conserved quantity. Fundamentally, this quantity results from Noether's theorem, which relates the conserved quantities to the symmetries.

The fact that the same ray parameter can be obtained from both Hamilton's ray equations and Lagrange's ray equations allows us to use both phase velocities, v, and ray velocities, V, as well as phase angles, ϑ, and ray angles, θ, to study rayfields in the context of conserved quantities. For instance, considering anisotropic vertically inhomogeneous continua, we can write

$$\mathfrak{p} = \frac{\sin\vartheta}{v(z,\vartheta)} = \cos\theta\frac{\partial}{\partial\theta}\left[\frac{1}{V(z,\theta)}\right] + \frac{\sin\theta}{V(z,\theta)},$$

where the relation between the magnitudes of phase and ray velocities is given by expressions (9.18), while the relation between the phase and ray angles is given by expression (9.21). An example illustrating this equivalence is shown in Exercise 11.1.

Closing remarks

In this chapter, we used ray parameters, which are first integrals of ray equations, to obtain expressions for rays and traveltimes. In a general inhomogeneous continuum, there are no ray parameters since the integrand of the traveltime integral is an explicit function of all the coordinates. In other words, the inhomogeneity of the continuum does not possess any convenient symmetry that would allow us to formulate expressions for conserved quantities. In such cases, we can still solve Hamilton's or Lagrange's ray equations to obtain rays, even though these equations may be analytically and numerically involved.

Exercises

Exercise 14.1 *Using polar coordinates, formulate the conserved quantity for radially inhomogeneous continua, where the traveltime integral is given* *by*

$$\int_a^b \frac{\sqrt{r^2 + (r')^2}}{V(r)}\, \mathrm{d}\xi.$$

Explain the physical context of the conserved quantity.

Solution 14.1 *Consider Beltrami's identity given by expression*

$$\frac{\partial \mathrm{F}}{\partial \xi} + \frac{\mathrm{d}}{\mathrm{d}\xi}\left(r'\frac{\partial \mathrm{F}}{\partial r'} - \mathrm{F}\right) = 0,$$

where F *denotes the integrand of the traveltime integral and* $r' := dr/\,d\xi$. *Due to the explicit absence of the latitude angle,* ξ, *we obtain*

$$r'\frac{\partial \mathrm{F}}{\partial r'} - \mathrm{F} = C,$$

where C *is a constant. Thus, performing the partial differentiation, we obtain*

$$C = -\frac{r^2}{V(r)\sqrt{r^2 + (r')^2}},$$

which is the expression for the conserved quantity. The conserved quantity results from the traveltime integral's invariance to the latitude angle. In other words, the velocity field consists of concentric circles.

Remark 14.1 *Noticing that*

$$\frac{r}{\sqrt{r^2 + (r')^2}} = \sin\theta,$$

where θ *is the ray angle measured between the ray and the radial direction, we can write*

$$C = -\frac{r\sin\theta}{V(r)}, \tag{14.41}$$

which is a standard form of the ray parameter for radially inhomogeneous continua.[6] *Note that ray parameter (14.41) has different units than ray parameter (14.6).*

[6]Readers interested in traveltime expressions for rays whose ray parameter is given by expression (14.41) might refer to Kennett, B.L.N., (2001) The seismic wavefield, Vol. I: Introduction and theoretical development: Cambridge University Press, pp. 171 – 174.

Exercise 14.2 *Given expression (14.11), determine how the value of the anisotropy parameter χ affects the maximum depth of a ray.*

Solution 14.2 *Solving equation (14.11) for z, we obtain*

$$z = \frac{1}{\mathfrak{p}b}\left(\frac{\sin\theta}{\sqrt{1+2\chi}\sqrt{2\chi\cos^2\theta+1}} - \mathfrak{p}a\right).$$

Consequently, the maximum depth is given by setting $\theta = \pi/2$ to obtain

$$z_{\max} = \frac{1}{\mathfrak{p}b}\left(\frac{1}{\sqrt{1+2\chi}} - \mathfrak{p}a\right). \tag{14.42}$$

To state expression (14.42) in terms of the initial ray angle, we set $z = 0$ and let the corresponding $\theta := \theta_0$, in expression (14.11). Hence, expression (14.11) becomes

$$\mathfrak{p} = \frac{\sin\theta_0}{a\sqrt{1+2\chi}\sqrt{2\chi\cos^2\theta_0+1}}. \tag{14.43}$$

Inserting expression (14.43) into expression (14.42), we obtain

$$z_{\max} = \frac{a}{b}\left(\frac{\sqrt{2\chi\cos^2\theta_0+1}}{\sin\theta_0} - 1\right),$$

which gives the maximum depth for a given initial angle. This expression shows that for $\chi \in (-0.5, 0)$, the maximum depth reached is less than that for the isotropic case, $\chi = 0$. Conversely, for $\chi \in (0, \infty)$, the maximum depth reached is greater than that for the isotropic case, $\chi = 0$. In other words, negative values of parameter χ increase the curvature of the ray while positive values decrease it.

Remark 14.2 *In most seismological studies of sedimentary layers, χ is positive. Hence, as shown in Exercise 14.2, the presence of anisotropy in a vertically inhomogeneous medium tends to straighten the rays and, hence, increase the maximum depth they reach.*

Exercise 14.3 *Show that the parametric form of the ray parameter, given by $\mathfrak{p} = \partial\mathcal{F}/\partial\dot{x}$ and discussed in Section 14.5, is equivalent to ray parameter (14.6).*

Solution 14.3 *Using the first equation of system (14.35) and considering the case where the argument x is not explicitly present in the integrand, we obtain a conserved quantity*

$$\mathfrak{p} = \frac{\partial \mathcal{F}}{\partial \dot{x}}. \tag{14.44}$$

In view of $\mathcal{F}$ given in expression (14.34), we obtain

$$\mathfrak{p} = \frac{1}{V}\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{z}^2}} + \sqrt{\dot{x}^2+\dot{z}^2}\frac{\partial}{\partial \dot{x}}\left(\frac{1}{V}\right). \tag{14.45}$$

In order to state expression (14.45) in terms of the ray angle, θ, we can write the differential operator as

$$\frac{\partial}{\partial \dot{x}} = \frac{\partial \theta}{\partial \dot{x}}\frac{\partial}{\partial \theta} + \frac{\partial \dot{z}}{\partial \dot{x}}\frac{\partial}{\partial \dot{z}} = \frac{\partial \theta}{\partial \dot{x}}\frac{\partial}{\partial \theta} = \frac{1}{\frac{\partial \dot{x}}{\partial \theta}}\frac{\partial}{\partial \theta}.$$

Since $\dot{x} = \dot{z}\tan\theta$, we obtain

$$\frac{\partial}{\partial \dot{x}} = \frac{1}{\dot{z}\frac{1}{\cos^2\theta}}\frac{\partial}{\partial \theta} = \frac{\cos^2\theta}{\dot{z}}\frac{\partial}{\partial \theta}.$$

Thus, returning to expression (14.45), we can write

$$\begin{aligned}
\mathfrak{p} &= \frac{\sin\theta}{V} + \sqrt{\dot{x}^2+\dot{z}^2}\frac{\cos^2\theta}{\dot{z}}\frac{\partial}{\partial \theta}\left(\frac{1}{V}\right) \\
&= \frac{\sin\theta}{V} + \sqrt{\tan^2\theta+1}\cos^2\theta\frac{\partial}{\partial \theta}\left(\frac{1}{V}\right) \\
&= \frac{\sin\theta}{V} + \cos\theta\frac{\partial}{\partial \theta}\left(\frac{1}{V}\right),
\end{aligned}$$

which is identical to expression (14.6), as required.

Exercise 14.4 *In view of Lemma 12.1, show that if $V(z,\theta) = A(z)B(\theta)$, where $B(\theta) = 1/(1+C\cos\theta)$, then the ray in an anisotropic inhomogeneous continuum, $V(z,\theta)$ is the same as the ray in an isotropic inhomogeneous continuum, $A(z)$.*

Solution 14.4 *To express $B(\theta)$, where θ is measured from the z-axis, in terms of z', we invoke trigonometric identity $\cos\theta = \cot\theta/\sqrt{1+\cot^2\theta}$. Noting that $\cot\theta = dz/dx := z'$, we obtain $\cos\theta = z'/\sqrt{1+(z')^2}$. Consequently,*

$$B(z') = \frac{1}{1 + C\frac{z'}{\sqrt{1+(z')^2}}},$$

and,

$$V\left(z,z'\right) = A\left(z\right)B\left(z'\right) = \frac{A\left(z\right)}{1 + C\dfrac{z'}{\sqrt{1+\left(z'\right)^2}}}. \tag{14.46}$$

Consider traveltime integral

$$\check{C} = \int\limits_{x_1}^{x_2} \frac{\sqrt{1+\left(z'\right)^2}}{V\left(z,z'\right)}\,\mathrm{d}x = \int\limits_{x_1}^{x_2} \frac{\sqrt{1+\left(z'\right)^2}}{\dfrac{A\left(z\right)}{1 + C\dfrac{z'}{\sqrt{1+\left(z'\right)^2}}}}\,\mathrm{d}x.$$

Upon algebraic manipulations, we obtain

$$\check{C} = \int\limits_{x_1}^{x_2} \frac{\sqrt{1+\left(z'\right)^2} + Cz'}{A\left(z\right)}\,\mathrm{d}x \equiv \int\limits_{x_1}^{x_2} \mathrm{F}\,\mathrm{d}x.$$

To find the ray, we invoke Euler's equation (12.4) to obtain

$$\begin{aligned} &\frac{\partial \mathrm{F}}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial \mathrm{F}}{\partial z'}\right) \\ &\quad = -\left(\sqrt{1+\left(z'\right)^2} + Cz'\right)\frac{\dfrac{\partial A}{\partial z}}{A^2} - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{1}{A}\left(\frac{z'}{\sqrt{1+\left(z'\right)^2}} + C\right)\right] = 0. \end{aligned} \tag{14.47}$$

Considering only factors which contain C, gives us

$$-Cz'\frac{\dfrac{\partial A}{\partial z}}{A^2} + Cz'\frac{\dfrac{\partial A}{\partial z}}{A^2} = 0.$$

Thus, Euler's equation is independent of C. In view of expression (14.46), the ray resulting from equation (14.47) is the same for both $V\left(z,z'\right)$ and $A\left(z\right)$.

Part IV

Appendices

This Page Intentionally Left Blank

Introduction to Part IV

> Physics is the science upon which all other sciences rest, since it attempts to explain the nature of the universe of things. [...] Mathematics is a language, which enables us to express certain kinds of ideas (e.g., order, patterns) much more precisely than whatever everyday language we speak. [...] Mathematical physics can therefore be regarded as the 'dialect' of mathematics spoken by physicists when they wish to express and use the 'laws' or theories of physics clearly and unambiguously.
>
> *Michael G. Rochester (1997) Lecture notes on mathematical physics*

In the presentation of this book, we assume that the reader is familiar with several mathematical subjects typically taught in undergraduate studies in the faculty of science. These subjects consist of linear algebra, differential and integral calculus, vector and tensor calculus, as well as ordinary and partial differential equations. Another subject that plays an important role in this book — but is not commonly included in an undergraduate curriculum — consists of the calculus of variations. Chapter 12 is devoted to the aspects of this subject that are pertinent to this book.

In *Part IV*, we describe two additional mathematical concepts that are used in the book and with which the reader might not be familiar, namely, Euler's homogeneous-function theorem and Legendre's transformations. Notably, in the context of this book, the applications of these two concepts are often associated with one another. In view of Euler's theorem, different degrees of homogeneity exhibited by several functions formulated in this book give us insight into their physical meanings and allow us to manipulate them. Legendre's transformation is the tool that allows us to transform Hamilton's ray equations into Lagrange's ray equations. Consequently, this transformation links the concepts discussed in *Part II* with those discussed in *Part III*.

Throughout the book, the meaning of a given symbol used in an equation is stated in the proximity of the pertinent equation to avoid ambiguity among several meanings that can be associated with the same symbol. To facilitate clarity, certain symbols are uniquely associated with a particular mathematical or physical meaning. These symbols, together with their meanings, are listed in *List of symbols.*

Appendix A

Euler's homogeneous-function theorem

> Mathematicians can pursue many conflicting directions to derive new results. In the absence of internal criteria that favour or justify one direction rather than another, a choice must be based on external considerations. Of these, certainly the most important is the traditional and still most justifiable reason for the creation and development of mathematics, its value to the sciences.
>
> *Morris Kline (1980) Mathematics: The loss of certainty*

Preliminary remarks

In this book, seismological quantities are expressed in terms of mathematical entities. In accordance with physical principles, we require that these entities possess certain mathematical properties. Using these properties, we can study these mathematical formulations to obtain further insight into their physical meaning. The homogeneity of a function and Euler's homogeneous-function theorem are of particular use in our work.

We begin this appendix by stating the definition of a homogeneous function. Then, we state and prove Euler's homogeneous-function theorem.

A.1 Homogeneous functions

Several functions that play an important role in this book are homogeneous. Notably, the Hamiltonian, stated in expression (8.20), namely,

$$\mathcal{H}(\mathbf{x},\mathbf{p}) = \frac{1}{2}p^2 v^2(\mathbf{x},\mathbf{p}), \tag{A.1}$$

where $p^2 \equiv \mathbf{p}\cdot\mathbf{p}$, is homogeneous of degree 2 in $\mathbf{p}$. To see this property, consider Definition A.1.

Definition A.1 *A real function $f(x_1,\ldots,x_n)$ is homogeneous of degree r in the variables $x_1,\ldots,x_n$ if*

$$f(cx_1,\ldots,cx_n) = c^r f(x_1,\ldots,x_n),$$

for every real number c. If $f(cx_i) = |c|^r f(x_i)$, where $i \in \{1,\ldots,n\}$, we say that f is absolute-value homogeneous of degree r in the x_i.

Remark A.1 *Both terms "degree" and "order" are commonly used to describe the homogeneity of a function. In this book, we use the former term since it refers to the value of the exponent and, hence, is consistent with other uses of this term, such as "degree of a polynomial".*

Now, consider the fact that v is the phase-velocity function that depends on position $\mathbf{x}$ and direction, which is given by the vector normal to the wavefront, namely, $\mathbf{p}$. Since the orientation of the wavefront, indicated by $\mathbf{p}$, does not depend on the magnitude of $\mathbf{p}$, we can rewrite expression (A.1) as

$$\mathcal{H}(\mathbf{x},\mathbf{p}) = \frac{1}{2}p^2 v^2\left(\mathbf{x},\frac{\mathbf{p}}{|\mathbf{p}|}\right), \tag{A.2}$$

where $\mathbf{p}/|\mathbf{p}|$ is a unit vector normal to the wavefront. Hence, we see that v is homogeneous of degree 0 in $\mathbf{p}$. In other words, we can multiply $\mathbf{p}$ by any number and v remains unchanged. In view of Definition A.1, we can write

$$v\left(\mathbf{x},\frac{c\mathbf{p}}{|c\mathbf{p}|}\right) = c^0 v\left(\mathbf{x},\frac{\mathbf{p}}{|\mathbf{p}|}\right) = v\left(\mathbf{x},\frac{\mathbf{p}}{|\mathbf{p}|}\right).$$

This immediately implies that function (A.2) is homogeneous of degree 2 in $\mathbf{p}$, since

$$\begin{aligned}\mathcal{H}(\mathbf{x},c\mathbf{p}) &= \frac{1}{2}\left[(c\mathbf{p})\cdot(c\mathbf{p})\right]v^2\left(\mathbf{x},\frac{c\mathbf{p}}{|c\mathbf{p}|}\right) = \frac{1}{2}c^2p^2v^2\left(\mathbf{x},\frac{c\mathbf{p}}{|c\mathbf{p}|}\right)\\ &= \frac{c^2}{2}p^2v^2\left(\mathbf{x},\frac{\mathbf{p}}{|\mathbf{p}|}\right) = c^2\mathcal{H}(\mathbf{x},\mathbf{p}).\end{aligned} \tag{A.3}$$

We can also illustrate Definition A.1 by the following straightforward example.

Example A.1 *Consider the function*

$$f(x_1, x_2, x_3) = x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2. \tag{A.4}$$

Let

$$\begin{aligned} f(cx, cx_2, cx_3) &= cx_1 cx_2 cx_3 + cx_1 (cx_2)^2 + cx_2 (cx_3)^2 \\ &= c^3 \left(x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2\right) \\ &= c^3 f(x_1, x_2, x_3). \end{aligned}$$

Thus, in view of Definition A.1, f is homogeneous of degree 3 in the x_i.

Homogeneity of a function allows us to use Euler's homogeneous-function theorem, stated in Theorem A.1. This theorem plays an important role in the formulations described in this book. It allows us to simplify numerous expressions and gain insight into their physical meaning.

A.2 Homogeneous-function theorem

[1]Euler's homogeneous-function theorem can be stated in the following way.

Theorem A.1 *If function $f(x_1, \ldots, x_n)$ is homogeneous of degree r in $x_1, \ldots, x_n$, then*

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n)\, x_i = r f(x_1, \ldots, x_n). \tag{A.5}$$

Proof. In view of Definition A.1, we can write

$$f(cx_1, \ldots, cx_n) = c^r f(x_1, \ldots, x_n). \tag{A.6}$$

Differentiating both sides of equation (A.6) with respect to c, we obtain

$$\sum_{i=1}^{n} f_i(cx_1, \ldots, cx_n) \frac{\partial (cx_i)}{\partial c} = r c^{r-1} f(x_1, \ldots, x_n), \tag{A.7}$$

[1]Interested readers might refer to Olmsted, J.M.H., (1961) Advanced calculus: Prentice-Hall, Inc., p. 272.

where f_i denotes the derivative of function f with respect to its ith argument. To obtain the expression stated in Theorem A.1, we consider a particular case where $c = 1$. Letting $c = 1$, we can rewrite equation (A.7) as

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n)\, x_i = r f(x_1, \ldots, x_n), \tag{A.8}$$

which is equation (A.5), as required. ■

To illustrate Theorem A.1, we can study function (A.4), as shown in the following example.

Example A.2 *Using function (A.4), namely,*

$$f(x_1, x_2, x_3) = x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2,$$

we can write the left-hand side of equation (A.5) as

$$\begin{aligned}\sum_{i=1}^{3} \frac{\partial f}{\partial x_i} x_i &= \left(x_2 x_3 + x_2^2\right) x_1 + \left(x_1 x_3 + 2 x_1 x_2 + x_3^2\right) x_2 + \left(x_1 x_2 + 2 x_2 x_3\right) x_3 \\ &= 3\left(x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2\right) \\ &= 3 f(x_1, x_2, x_3).\end{aligned} \tag{A.9}$$

Expression (A.9) is the right-hand side of equation (A.5) for a function that is homogeneous of degree 3 *in the* x_i*, as expected from Theorem A.1.*

Equation (A.5) is often invoked in this book. For instance, in the proof of Lemma 13.2 — knowing that $\mathcal{H}$ is homogeneous of degree 2 in $\mathbf{p}$, as shown in expression (A.3) — we can write

$$\sum_{i=1}^{n} \frac{\partial \mathcal{H}}{\partial p_i} p_i = 2\mathcal{H},$$

which allows us to complete that proof.

The following example illustrates equation (A.5) in the context of physics.

Example A.3 *Following the standard classical-mechanics formulation, let the kinetic energy be*

$$T(v) = \frac{1}{2} m v^2,$$

where m and v denote mass and velocity, respectively. In view of Definition A.1, T is homogeneous of degree 2 in v since

$$T(cv) = \frac{1}{2}m(cv)^2 = \frac{c^2}{2}mv^2 = c^2 T(v),$$

where c denotes a constant. Thus, following Theorem A.1, we can write

$$\frac{\partial T}{\partial v}v = 2T.$$

We can directly verify this result, namely,

$$\frac{\partial T}{\partial v}v = \left[\frac{\partial}{\partial v}\left(\frac{mv^2}{2}\right)\right]v = mv^2 = 2T.$$

Closing remarks

Note that a multivariable function can be homogeneous in a particular set of variables. In this book, $\mathcal{H}(\mathbf{x}, \mathbf{p})$, given in expression (8.20), is homogeneous of degree 2 in $\mathbf{p}$. $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$, given in expression (9.2), is homogeneous of degree 2 in $\dot{\mathbf{x}}$. $\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}})$, given in expression (13.8), is absolute-value homogeneous of degree 1 in $\dot{\mathbf{x}}$. None of these functions is homogeneous in $\mathbf{x}$. The properties of homogeneity of these functions allow us to prove Theorem 13.1, which is the statement of Fermat's principle.

Certain functions used in our studies exhibit no homogeneity. For instance, traveltime integrand $\mathrm{F}(z, z')$, given in expression (14.2), is not homogeneous in either variable.

Euler's homogeneous-function theorem is explicitly used in Chapters 8, 9 and 13.

This Page Intentionally Left Blank

Appendix B

Legendre's transformation

> To penetrate into symplectic geometry while bypassing the long historical route, it is simplest to use the axiomatic method, which has, as Bertrand Russell observed, many advantages, similar to the advantages of stealing over honest work.
>
> *Vladimir Igorevitch Arnold (1992) Catastrophe theory*

Preliminary remarks

Legendre's transformation is a transformation in which we replace a function by a new function that depends on partial derivatives of the original function with respect to original independent variables. In the context of this book, we replace the ray-theory Hamiltonian, $\mathcal{H}(\mathbf{x}, \mathbf{p})$, by the ray-theory Lagrangian, $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$, which depends on the $\dot{x}_i = \partial\mathcal{H}/\partial p_i$, where $i \in \{1, 2, 3\}$.

We begin this appendix with the derivation of Legendre's transformation in a geometrical context, where we consider functions of single variables. Then we proceed to multivariable functions and formulate Legendre's transformation between $\mathcal{H}(\mathbf{x}, \mathbf{p})$ and $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$. We conclude by using Legendre's transformations of these functions to derive the corresponding ray equations.

B.1 Geometrical context

B.1.1 Surface and its tangent planes

[1]Legendre's transformation can be illustrated in a geometrical context. Let an n-dimensional surface in the $(n+1)$-dimensional space be given by equation

$$y = f(x_1, \ldots, x_n). \tag{B.1}$$

Consider the set of all possible n-dimensional planes that are tangent to this surface. The envelope of these planes is the original surface. We wish to derive the equation that describes these tangent planes.

A general form of the equation of an n-dimensional plane is $y = u_1 x_1 + \ldots + u_n x_n - v$, where $u_1, \ldots, u_n$ and v are real numbers that define the plane uniquely and, hence, can be viewed as coordinates of the plane. The transformation from the equation of a surface, given by equation (B.1), to the equation that describes all of its tangent planes, given by

$$v = g(u_1, \ldots, u_n),$$

is Legendre's transformation.

Note that Legendre's transformation is possible if the surface is differentiable and if there are no tangent planes to this surface that are parallel to each other. Otherwise, for the same set $(u_1, \ldots, u_n)$, we have different values of v. In other words, v is not a single-valued function of $(u_1, \ldots, u_n)$.

B.1.2 Single-variable case

To illustrate the geometrical context, consider a smooth curve in the xy-plane. We can describe this curve as a set of points in the plane, where the y-coordinate is determined by the function of one variable, namely, $y = f(x)$. Also, this curve can be regarded as the envelope of its tangent lines. We wish to derive equation $v = g(u)$ that describes all the lines $y = ux - v$, in the xy-plane, that are tangent to the original curve.

The line $y = ux - v$ is tangent to the curve $y = f(x)$, at some point x, if and only if the line passes through the point $(x, f(x))$ and has the same slope as the curve at this point. In other words,

$$v = ux - f(x), \tag{B.2}$$

[1]Readers interested in the geometrical motivation of the Legendre transformation might refer to Courant, R., and Hilbert, D., (1924/1989) Methods of mathematical physics: John Wiley & Sons., Vol. II, pp. 32 – 39.

and

$$u = \frac{\mathrm{d}f}{\mathrm{d}x}, \tag{B.3}$$

respectively.

To complete the derivation of function $g(u)$, we would like to express x in terms of u. This is not always possible since we might not be able to uniquely solve equation (B.3) for x. Our ability to express x in terms of u depends on the form of function $f(x)$.

Assuming that we can obtain $x = x(u)$, the set of all tangent lines is described by $v = g(u)$ where

$$g(u) = ux(u) - f(x(u)). \tag{B.4}$$

Thus, $g(u)$ is Legendre's transformation of $f(x)$. This construction is illustrated by the following example.

Example B.1 *Let $f(x) = x^2$. Then, following equation (B.3), we can write*

$$u = \frac{\mathrm{d}f}{\mathrm{d}x} = 2x.$$

Hence, we obtain

$$x = \frac{u}{2}.$$

Consequently, in view of equation (B.4), we get

$$g(u) = ux(u) - f(x(u)) = \frac{u^2}{2} - \left(\frac{u}{2}\right)^2 = \frac{u^2}{4}.$$

Therefore, $v = u^2/4$ is Legendre's transformation of $y = x^2$.

We can also view Legendre's transformation in a different way. Consider a curve $y = f(x)$ and a straight line $y = ux$, where u is a real number. For a given x-coordinate, we can view $h(x) = ux - f(x)$ as the distance between a point on the curve and a point on the straight line. We wish to find point $x(u)$ that maximizes that distance. Therefore, we set

$$\frac{\mathrm{d}h}{\mathrm{d}x} = u - \frac{\mathrm{d}f}{\mathrm{d}x} = 0,$$

which gives

$$u = \frac{\mathrm{d}f}{\mathrm{d}x}.$$

If we can solve this equation for x, namely, $x = x(u)$, then $g(u) = h(x(u))$ is Legendre's transformation of $f(x)$.

B.2 Duality of transformation

Legendre's transformation is often referred to as a dual transformation since if transformation of f leads to g, then, transformation of g must lead to f.[2] We can illustrate this property by inverting the transformation shown in Example B.1.

Example B.2 *Let $g(u) = u^2/4$. Consider a new function given by*

$$\begin{aligned} f(x) &= ux - g(u) \\ &= ux - \frac{u^2}{4}, \end{aligned} \tag{B.5}$$

where the new independent variable is

$$x = \frac{\mathrm{d}g}{\mathrm{d}u} = \frac{u}{2}.$$

In view of the new independent variable, we can uniquely express u in terms of x, namely, $u = 2x$. Hence, we can write function (B.5) as

$$f(x) = 2x^2 - \frac{(2x)^2}{4} = x^2,$$

as expected from Example B.1. Therefore, $y = x^2$ is Legendre's transformation of $v = u^2/4$.

B.3 Transformation between $\mathcal{L}$ and $\mathcal{H}$

Notation B.1 *In this appendix, to familiarize the reader with the fact that the phase slowness is a covector, $\mathbf{p}$, while the ray velocity is a vector, $\dot{\mathbf{x}}$, following the standard convention, their components appear as subscripts and superscripts, respectively. This distinction is not used in the text of the book.*

Notation B.2 *In this appendix, to show the generality of the formulation, all expressions are derived for an n-dimensional space.*

Remark B.1 *Throughout this book, we formulate our expressions in terms of orthonormal coordinates. The distinction between vectors and covectors becomes important if curvilinear coordinates are used.*

[2]For a proof of this duality, readers might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2nd edition): Springer-Verlag, p. 63.

[3]In the context of this book, Legendre's transformation relates the ray-theory Lagrangian, $\mathcal{L}$, to the ray-theory Hamiltonian, $\mathcal{H}$. The transformation between functions $\mathcal{L}(\mathbf{x},\cdot)$ and $\mathcal{H}(\mathbf{x},\cdot)$ is analogous to the transformation between functions $f(\cdot)$ and $g(\cdot)$, discussed above, where $\cdot$ stands for the variables of transformation. Note that $\mathbf{x}$, while specifying the point in the continuum where the transformation is performed, plays no role in this transformation. At a given point $\mathbf{x}$, $\dot{\mathbf{x}}$ and $\mathbf{p}$ are the active variables of transformation.

Let $\mathcal{L} = \mathcal{L}(\mathbf{x},\dot{\mathbf{x}})$. In view of expression (B.4), consider a new function given by

$$\mathcal{H}(\mathbf{x},\mathbf{p}) = \sum_{i=1}^{n} \dot{x}^i p_i - \mathcal{L}(\mathbf{x},\dot{\mathbf{x}}), \tag{B.6}$$

where, in view of expression (B.3), the new variables are

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}^i}, \qquad i \in \{1,\ldots,n\}. \tag{B.7}$$

Following expression (B.6), we can write the differential of $\mathcal{H}(\mathbf{x},\mathbf{p})$ as

$$\begin{aligned} \mathrm{d}\mathcal{H} &= \sum_{i=1}^{n} \left(p_i \mathrm{d}\dot{x}^i + \dot{x}^i \mathrm{d}p_i\right) - \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial x^i}\mathrm{d}x^i + \frac{\partial \mathcal{L}}{\partial \dot{x}^i}\mathrm{d}\dot{x}^i\right) \\ &= \sum_{i=1}^{n} \left(p_i \mathrm{d}\dot{x}^i + \dot{x}^i \mathrm{d}p_i - \frac{\partial \mathcal{L}}{\partial x^i}\mathrm{d}x^i - \frac{\partial \mathcal{L}}{\partial \dot{x}^i}\mathrm{d}\dot{x}^i\right). \end{aligned} \tag{B.8}$$

In view of expression (B.7), the first and the last term in expression (B.8) cancel one another. Thus, we obtain

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{n} \left(\dot{x}^i \mathrm{d}p_i - \frac{\partial \mathcal{L}}{\partial x^i}\mathrm{d}x^i\right). \tag{B.9}$$

Also, we can formally write the differential of $\mathcal{H}(\mathbf{x},\mathbf{p})$ as

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{n} \left(\frac{\partial \mathcal{H}}{\partial x^i}\mathrm{d}x^i + \frac{\partial \mathcal{H}}{\partial p_i}\mathrm{d}p_i\right), \tag{B.10}$$

which is a statement of the chain rule.

[3]Readers interested in an insightful description of Legendre's transformation, including the duality of the transformation and the application of the transformation to $\mathcal{L}$ and $\mathcal{H}$, might refer to Lanczos, C., (1949/1986) The variational principles in mechanics: Dover, pp. 161 – 172.

Equating the right-hand sides of equations (B.9) and (B.10), we can write

$$\sum_{i=1}^{n}\left(\dot{x}^{i}\mathrm{d}p_{i}-\frac{\partial\mathcal{L}}{\partial x^{i}}\mathrm{d}x^{i}\right)=\sum_{i=1}^{n}\left(\frac{\partial\mathcal{H}}{\partial x^{i}}\mathrm{d}x^{i}+\frac{\partial\mathcal{H}}{\partial p_{i}}\mathrm{d}p_{i}\right). \tag{B.11}$$

By examining equation (B.11), we conclude that

$$\dot{x}^{i}=\frac{\partial\mathcal{H}}{\partial p_{i}},\qquad i\in\{1,\ldots,n\}. \tag{B.12}$$

Examining expressions (B.7) and (B.12), we recognize the duality of these expressions. Thus, we can write the counterpart of expression (B.6), namely,

$$\mathcal{L}(\mathbf{x},\dot{\mathbf{x}})=\sum_{i=1}^{n}\dot{x}^{i}p_{i}-\mathcal{H}(\mathbf{x},\mathbf{p}), \tag{B.13}$$

where the active variables are given by expression (B.12).

B.4 Transformation and ray equations

Knowing that $\mathcal{H}(\mathbf{x},\mathbf{p})$ is Legendre's transformation of $\mathcal{L}(\mathbf{x},\dot{\mathbf{x}})$, and vice-versa, we wish to consider the effect of this transformation on the corresponding ray equations. Herein, in view of the duality of the transformation, we derive Hamilton's ray equations from Lagrange's ray equations. This process is the inverse of the transformation used in Chapter 9.

Recall Lagrange's ray equations (9.8), which, in general, can be written as

$$\frac{\partial\mathcal{L}}{\partial x^{i}}=\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}^{i}},\qquad i\in\{1,\ldots,n\}. \tag{B.14}$$

By examining equation (B.11), we conclude that

$$-\frac{\partial\mathcal{L}(\mathbf{x},\dot{\mathbf{x}})}{\partial x^{i}}=\frac{\partial\mathcal{H}(\mathbf{x},\mathbf{p})}{\partial x^{i}},\qquad i\in\{1,\ldots,n\}, \tag{B.15}$$

where relations between $\dot{\mathbf{x}}$ and $\mathbf{p}$ are given by expressions (B.7) and (B.12). Hence, using expressions (B.7) and (B.15), we can write equation (B.14) as

$$-\frac{\partial\mathcal{H}}{\partial x^{i}}=\frac{\mathrm{d}p_{i}}{\mathrm{d}t},\qquad i\in\{1,\ldots,n\},$$

which can be immediately restated as

$$\dot{p}_{i}=-\frac{\partial\mathcal{H}}{\partial x^{i}},\qquad i\in\{1,\ldots,n\}. \tag{B.16}$$

Thus, we conclude that using the new function, given in expression (B.6), and the new variables, given in expression (B.7), we obtain expression (B.12), while, invoking Lagrange's ray equations (9.8), we obtain expression (B.16). We notice that the system composed of equations (B.12) and (B.16) are Hamilton's ray equations (8.19), namely,

$$\begin{cases} \dot{x}^i = \dfrac{\partial \mathcal{H}}{\partial p_i} \\ \\ \dot{p}_i = -\dfrac{\partial \mathcal{H}}{\partial x^i} \end{cases}, \qquad i \in \{1, \ldots, n\}. \tag{B.17}$$

Hence, Legendre's transformation of $\mathcal{L}$, which leads to $\mathcal{H}$, allows us to derive Hamilton's ray equations from Lagrange's ray equations. In view of this derivation, we recognize that the first equation of system (B.17) is the definition of a variable for Legendre's transformation, while the second equation is endowed with the physical content since it results from Lagrange's ray equations.

Closing remarks

In the context of our work, the fundamental physical principles are directly contained in Hamilton's ray equations, which originate in Cauchy's equations of motion. The fundamental justification of Lagrange's ray equations relies on Legendre's transformations and, hence, it is subject to the singularities of this transformation.[4] Furthermore, if we wish to express the governing equations explicitly in terms of Lagrangian $\mathcal{L}$, we need to solve equations (B.12) for the p_i, in a closed form, which is not always possible.

In the context of elastic continua, the desired transformation is possible for any convex phase-slowness surface. Furthermore, if $\mathcal{H}$ is a quadratic function in the p_i — in other words, the phase-slowness surface is elliptical — we can always obtain explicit, closed-form expressions for the ray velocity and the ray angle.

Legendre's transformation links the concepts of *Part II* with those of *Part III* in this book, and is explicitly used in Chapters 9 and 13.

[4]In general, depending on the context, we can view either the Hamiltonian or the Lagrangian formulation as being more fundamental. Readers interested in this question might refer to Marsden, J.E., and Ratiu, T.S., (1999) Introduction to mechanics and symmetry: A basic exposition of classical mechanical systems (2*nd* edition): Springer-Verlag, pp. 1 – 6.

This Page Intentionally Left Blank

Appendix C

List of symbols

> Our symbolic mechanism is eminently useful and powerful, but the danger is ever-present that we become drowned in a language which has its well-defined grammatical rules but evidently loses all content and becomes a nebulous sham.
>
> *Cornelius Lanczos (1961) Linear differential operators*

Remark C.1 *Symbols listed herein correspond to the given meaning throughout the entire book.*

C.1 Mathematical relations and operations

$=$	equality
$\approx$	approximation
$\equiv$	identity
$:=$	definition
$\sim$	asymptotic relation
$\circ$	orthogonal-transformation operator
$\cdot$	scalar product
$\times$	vector product
∇	gradient
$\nabla\cdot$	divergence
$\nabla\times$	curl
d	total derivative
∂	partial derivative
δ	variation
δ_{ij}	Kronecker's delta
ϵ_{ijk}	permutation

$\in$	"belongs to a set"
$\longrightarrow$	"maps to" or "tends to"
Θ	coordinate-rotation angle
J	Jacobian
$\mathbb{R}$	real numbers
$\mathbb{R}^n$	n-dimensional space of real numbers
$\mathbf{e}_i$	unit vector along the x_i-axis
$f\left(\cdot\right)\rvert_a$	function $f\left(\cdot\right)$ evaluated at $\cdot = a$

C.2 Physical quantities

C.2.1 Greek letters

ε_{kl}	strain tensor see expression (1.15)
ϑ	phase angle see expressions (10.30) and (9.14)
θ	ray angle see expressions (9.30) and (9.21)
κ	compressibility see expressions (5.71)
λ	Lamé's parameter see expressions (5.63)
μ	Lamé's parameter, also known as rigidity modulus see expressions (5.63)
ν	Poisson's ratio see expressions (5.95) and (5.96)
ξ_{ij}	rotation tensor see expression (1.29)
$\mathbf{\Psi}$	rotation vector see expression (1.30)
ρ	mass density see expression (2.1)
σ_{ij}	stress tensor see expression (2.18) and Figure 2.1
φ	dilatation see expression (1.26)
ϕ	displacement angle see expression (10.35)
ω	angular frequency see expression (6.48)

C.2.2 Roman letters

c_{ijkl} elasticity tensor
see expression (3.1)

C_{mn} elasticity-matrix entries, also known as elasticity parameters
see expression (4.11)

E Young's modulus
see Remark 5.3

$\mathcal{F}$ ray-theory Lagrangian,
absolute-value homogeneous of degree 1 in the $\dot{x}_i$
see expression (13.8)

F ray-theory Lagrangian, inhomogeneous
see integral (14.2)

$\mathcal{H}$ ray-theory Hamiltonian, homogeneous of degree 2 in the p_i
see expression (8.20)

$\mathbb{H}$ classical-mechanics Hamiltonian
see Exercise 13.5

$\mathcal{L}$ ray-theory Lagrangian, homogeneous of degree 2 in the $\dot{x}_i$
see expression (9.2)

$\mathbb{L}$ classical-mechanics Lagrangian
see expression (13.14)

W strain energy
see expressions (4.3) and (4.1)

Bibliography

This list contains works that served the author as a source of information and inspiration in writing this book. Many concepts discussed in this book result from the author's studying the works included in this list.

This bibliography contains numerous republications of previously published works. In certain cases, there are two dates that refer to the original publication and to the edition, translation or facsimile used herein.

Achenbach, J.D., (1973) Wave propagation in elastic solids: North Holland

Achenbach, J.D., Gausten, A.K., and McMaken, H., (1982) Ray methods for waves in elastic solids: with applications to scattering by cracks: Pitman Publishing Inc.

Ait-Haddou, R. (1996) Courbes à hodographe Pythagorien en géométrie de Minkowski et modélisation géométric (Ph.D. thesis): Université Joseph Fourier – Grenoble I

Ait-Haddou, R., Biard, L., and Slawinski, M.A., (2000) Minkowski isoperimetric-hodograph curves. Comp. Aided Geom. Design. **17**, Issues 9, 835 – 861.

Aki, K., and Richards, P.G., (1980) Quantitative seismology: Theory and methods: W.H. Freeman & Co.

Aki, K., and Richards, P.G., (2002) Quantitative seismology (2*nd* edition): University Science Books

Akivis, M.A., and Goldberg, V.V., (1972) An introduction to linear algebra and tensors: Dover

Aoki, H., Syono, Y., and Hemley, R.J., (editors), (2000) Physics meets mineralogy: Condensed-matter physics in geosciences: Cambridge University Press

Aleksandrov, A.D., Kolmogorov, A.N., Lavrentev, M.A., (editors), (1969/1999) Mathematics: Its content, methods and meaning: Dover

Antman, S.S., (1995) Nonlinear problems of elasticity: Springer-Verlag

Anton, H., (1973) Elementary linear algebra: John Wiley & Sons

Anton, H., (1984) Calculus with analytic geometry: John Wiley & Sons

Antonelli, P.L., Ingarden, R.S., and Matsumoto, M., (1993) The theory of sprays and Finsler spaces with applications in physics and biology: Kluwer

Appell, P., (1905) Cours de mécanique à l'usage des élèves de la classe de mathématique spéciale: Gauthier-Villars, Paris

Arfken, G.B., (1985) Mathematical methods for physicists ($3rd$ edition): Academic Press, Inc.

Arfken, G.B, and Weber, H.J., (2001) Mathematical methods for physicists ($5th$ edition): Harcourt/Academic Press

Arnold, V.I. (1984) Ordinary differential equations ($3rd$ edition): Springer-Verlag

Arnold, V.I., (1989) Mathematical methods of classical mechanics ($2nd$ edition): Springer-Verlag

Arnold, V.I., (1991) The theory of singularities and its applications: Cambridge University Press

Arnold, V.I., (1992) Catastrophe theory ($3rd$ edition): Springer-Verlag

Auld, B.A., (1973) Acoustic fields and waves in solids: John Wiley and Sons

Ayres, F., (1962) Matrices: Schaum's Outlines, McGraw-Hill, Inc.

Babich, V.M., (1961) Ray method of calculating the intensity of wavefronts in the case of a heterogeneous, anisotropic, elastic medium: In *Problems of the dynamic theory of propagation of seismic waves* (in Russian), **5**, ed. G.I. Petrashen, 36 – 46: Leningrad University Press. (English translation: Geophys. J. Int., 1994, **118**, 379 – 383.)

Babich, V.M., and Buldyrev, V.S., (1972/1991) Short-wavelength diffraction theory: Asymptotic methods: Springer-Verlag

Backus, G., (1970) A geometrical picture of anisotropic elastic tensors: Rev. Geophys., **8** (3), 633 – 671.

Bao, D., Chern, S.-S., and Shen, Z., (Editors), (1996) Finsler geometry: American Mathematical Society

Bao, D., Chern, S.-S., and Shen, Z., (2000) An introduction to Riemann-Finsler geometry: Springer-Verlag

Båth, M., and Berkhout, A.J., (1984) Mathematical aspects of seismology: Geophysical Press

Ben-Menahem, A., and Singh, S.J., (1981/2000) Seismic waves and sources: Dover

Bérest, P., (1997) Calcul des variations, Application à la mécanique et à la physique: Ellipses

Beyer, W.H., (editor), (1984) CRC Standard mathematical tables (27*th* edition): CRC Press, Inc.

Bleistein, N., (1984) Mathematical methods for wave phenomena: Academic Press

Bleistein, N., Cohen, J.K., and Stockwell, J.W., (2001) Mathematics of multidimensional seismic imaging, migration, and inversion: Springer-Verlag

Bleistein, N., and Handelsman, R.A., (1975/1986) Asymptotic expansions of integrals: Dover

Bóna, A., and Slawinski, M.A., (2003) Fermat's principle for seismic rays in elastic continua: Journal of Applied Geophysics (*in press*)

Bond, W.L., (1943) The mathematics of the physical properties of crystals: The Bell System Technical Journal, **22**, No. 1, 1 – 72.

Born, M., and Wolf, E., (1999) Principles of optics (7*th* edition): Cambridge University Press

Brekhovskikh, L.M., and Goncharov, V., (1982/1994) Mechanics of continua and wave dynamics: Springer-Verlag.

Bronson, R., (1973) Modern introductory differential equations: Schaum's Outlines, McGraw-Hill, Inc.

Buchanan, G.R., (1995) Finite element analysis: Schaum's Outlines, McGraw-Hill, Inc.

Buchdahl, H.A., (1970/1993) An introduction to Hamiltonian optics: Dover

Bullen, K.E., and Bolt, B.A., (1947/1985) An introduction to the theory of seismology: Cambridge University Press

Carcione, J.M., (2001) Wave fields in real media: wave propagation in anisotropic, anelastic and porous media: Pergamon

Cauchy, A. L., (1827) De la pression ou tension dans un corps solide: Ex. de Math, **2**, 42 – 56.

Červený, V., 2002. Fermat's variational principle for anisotropic inhomogeneous media: Studia geophysica and geodætica, **46**, 567 – 588.

Červený, V., (2001) Seismic ray theory: Cambridge University Press

Červený, V., and Ravindra, R., (1971) Theory of seismic head waves: University of Toronto Press

Červený, V., Molotkov, I.A., and Pšenčík, I., (1977) Ray method in seismology: Univerzita Karlova

Chadwick, P., (1976/1999) Continuum mechanics: Concise theory and problems: Dover

Chew, W.C., (1990) Waves and fields in inhomogeneous media: Van Nostrand Reinhold

Chou, P.C., and Pagano, N.J., (1967/1992) Elasticity: tensor, dyadic, and engineering approaches: Dover

Cox, D., Little, J., and O'Shea, D., (1992) Ideals, varieties and algorithms (2*nd* edition): Springer-Verlag

Collins, R.E., (1968/1999) Mathematical methods for physicists and engineers: Dover

Corben, H.C., and Stehle, P., (1950/1994) Classical mechanics: Dover

Courant, R., and Hilbert, D., (1924/1989) Methods of mathematical physics: John Wiley & Sons

Courant, R., and John, F., (1974/1989) Introduction to calculus and analysis: Springer-Verlag

Courant, R., and Robbins, H., (1941/1996) What is mathematics?: Oxford University Press

de Hoop, A. T., (1995) Textbook of radiation and scattering of waves: Academic Press, Inc.

Dahlen, F.A., and Tromp, J., (1998) Theoretical global seismology: Princeton University Press

Demidov, A.S., (2001) Generalized functions in mathematical physics: Main ideas and concepts: Nova Science Publishers, Inc.

Dennery, P., and Krzywicki, A., (1967) Mathematics for physicists: Harper and Row

Dingle, R.B., (1973) Asymptotic expansions: Their derivation and interpretation: Academic Press

do Carmo, M.P., (1976) Differential geometry of curves and surfaces: Prentice-Hall, Inc.

Duistermaat, J.J., (1996) Fourier Integral operators: Birkhäuser

Eisenhart, L. P., (1925/1997) Riemannian Geometry: Princeton University Press

Elmore, W.C., and Heald, M.A., (1969/1985) Physics of waves: Dover

Elsgolc, L.E., (1962) Calculus of Variations: Pergamon

Epstein, M., and Slawinski, M.A., (1998) On some aspects of the continuum-mechanics context, Revue de l'Institut Français du Pétrole, **53**, No. 5, 669 – 677.

Epstein, M., and Slawinski, M.A., (1999) On raytracing in constant velocity-gradient media: Geometrical approach, Canadian Journal of Exploration Geophysics, **35**, No. 1/2, 1 – 6.

Epstein, M., and Slawinski, M.A., (1999) On rays and ray parameters in inhomogeneous isotropic media. Canadian Journal of Exploration Geophysics, **35**, No. 1/2, 7 – 19.

Epstein, M., and Śniatycki, J. (1992) Fermat's principle in elastodynamics: Journal of Elasticity, **27**, 45 – 56.

Eringen, A.C., (1967) Mechanics of Continua: John Wiley & Sons

Ewing, M.G., (1969/1985) Calculus of variations with applications: Dover

Ewing, W.M., Jardetzky, W.S., and Press, F., (1957) Elastic waves in layered media: McGraw-Hill, Inc.

Farlow, S.J., (1982) Partial differential equations for scientists and engineers: Dover

Fedorov, F.I., (1968) Theory of elastic waves in crystals: Plenum Press

Feynman, R.P., Leighton, R.B., and Sands, M., (1963/1989) Feynman's lectures on physics: Addison-Wesley Publishing Co.

Feynman, R.P., (1967) The character of physical law: The MIT Press

Flanders, H., (1963/1989) Differential forms with applications to the physical sciences: Dover

Frankel, T., (1997) The geometry of physics: an introduction: Cambridge University Press

Franklin, J.N., (1968/2000) Matrix theory: Dover

Fung, Y.C., (1965) Foundations of solid mechanics: Prentice-Hall, Inc.

Fung, Y.C., (1977) A first course in continuum mechanics: Prentice-Hall, Inc.

Gelfand, I.M., and Fomin, S.V., (1963/2000) Calculus of variations: Dover

Gibson, C.G., (2001) Elementary geometry of differentiable curves: An undergraduate introduction: Cambridge University Press

Goldberg, S.I., (1962/1998) Curvature and homology: Dover

Goldstein, H., (1950/1980) Classical mechanics: Addison-Wesley Publishing Co.

Graff, K.F., (1975/1991) Wave motion in elastic solids: Dover

Grant, F.S., and West, G.F., (1965) Interpretation theory in applied geophysics: McGraw-Hill, Inc.

Greenspan, D., (1961/2000) Introduction to partial differential equations: Dover

Greenwood, D.T., (1977/1997) Classical dynamics: Dover

Guggenheimer, H.W., (1963/1977) Differential geometry: Dover

Guillemin, V., and Sternberg, S., (1977) Geometric asymptotics: American Mathematical Society

Hanna, J.R., (1982) Fourier series and integrals of boundary value problems: John Wiley and Sons

Hanyga, A., (1985) Seismic wave propagation in the Earth: Elsevier

Hanyga, A., (1985) Mathematical theory of non-linear elasticity: PWN — Polish Scientific Publishers

Hanyga, A., (2001) Fermat's principle in anisotropic elastic media: Advances in anisotropy: Selected theory, modeling, and case studies: Society of Exploration Geophysicists (Special Issue), 271 – 288.

Hanyga, A., and Slawinski, M.A., (2001), Caustics in qSV rayfields of transversely isotropic and vertically inhomogeneous media: Anisotropy 2000: Fractures, converted waves, and case studies: Society of Exploration Geophysicists (Special Issue), 409 – 418.

Hardy, G.H., (1940/1996) A mathematician's apology: Cambridge University Press

Harper, C., (1976) Introduction to mathematical physics: Prentice-Hall, Inc.

Hausner, M., (1965/1998) A vector space approach to geometry: Dover

Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon

Herman, B., (1945) Some theorems of the theory of anisotropic media: Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS, **48** (2), 89 – 92.

Hildebrand, F.B., (1948) Advanced calculus for applications: Prentice-Hall, Inc.

Hildebrandt, S. and Tromba, A., (1996) The parsimonious universe: shape and form in the natural world: Springer-Verlag

Ihlenburg, F., (1998) Finite element analysis of acoustic scattering: Springer-Verlag

Jeans, J., (1943/1981) Physics and philosophy: Dover

Jeffreys, H., and Jeffreys, B., (1946/1999) Methods of Mathematical Physics: Cambridge University Press

Kac, M., and Ulam, S.M., (1968/1992) Mathematics and logic: Dover

Karal, F.C., and Keller, J.B., (1959) Elastic wave propagation in homogeneous and inhomogeneous media: J. Acoust. Soc. Am., **31** (6), 694 – 705.

Kennett, B.L.N., (2001) The seismic wavefield, Vol. I: Introduction and theoretical development: Cambridge University Press

Kilmister, C.W., and Reeve, J.E., (1966) Rational mechanics: Longmans

Kline, M., (1959/1981) Mathematics and the physical world: Dover

Kline, M., (1972) Mathematical thought from ancient to modern times: Oxford University Press

Kline, M., (1980) Mathematics: The loss of certainty: Oxford University Press

Knobel, R., (2000) An introduction to the mathematical theory of waves: American Mathematical Society, Institute for Advance Studies

Kolmogorov, A.N., and Fomin, S.V., (1954 – 1960/1999) Elements of the theory of functions and functional analysis: Dover

Kolsky, H., (1953/1963) Stress waves in solids: Dover

Kravtsov, Y.A., and Orlov, Y.I., (1990) Geometrical optics of inhomogeneous media: Springer-Verlag

Kravtsov, Y.A., and Orlov, Y.I., (1999) Caustics, catastrophes and wave fields: Springer-Verlag

Krebes, E.S., (1987) Seismic theory and methods (Lecture notes): University of Calgary

Kreyszig, E., (1959/1991) Differential geometry: Dover

Krysicki, W., and Włodarski, L., (1980) Analiza matematyczna w zadaniach: PWN — Polish Scientific Publishers

Lafontaine, J., (1996) Introduction aux variétés différentielles: Presses Universitaires de Grenoble

Lagrange, J-L., (1788/1989) Mécanique analytique: Éditions Jacques Gabay

Lancaster, P., and Šalkauskas, K., (1996) Transform methods in applied mathematics: An introduction: John Wiley and Sons

Lanczos, C., (1949/1986) The variational principles of mechanics: Dover

Lanczos, C., (1956/1988) Applied analysis: Dover

Lanczos, C., (1961/1997) Linear differential operators: Dover

Lang, S., (1973) Calculus of several variables: Addison-Wesley Publishing Co.

Lang, S., (1995) Differential and Riemannian manifolds: Springer-Verlag

Le, L.H.T., Krebes, E.S., and Quiroga-Goode, G.E., (1994) Synthetic seismograms for SH waves in anelastic transversely isotropic media: Geophys. J. Int, **116**, 598 - 604.

Lee, J.M., (2000) Introduction to topological manifolds: Springer-Verlag

Lipschutz, M.M., (1969) Theory and problems of differential geometry: Schaum's Outline Series, McGraw-Hill, Inc.

Love, A.E.H., (1892/1944) A treatise on the mathematical theory of elasticity: Dover

Lovelock, D., and Rund, H., (1975/1989) Tensors, differential forms, and variational principles: Dover

MacBeth, C., (2002) Multi-component VSP analysis for applied seismic anisotropy: Pergamon

Macelwane, J.B., and Sohon, F.W., (1936) Introduction to theoretical seismology, Part I: Geodynamics: John Wiley and Sons

Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall

Marion, J.B., and Thornton, S.T., (1970/1988) Classical dynamics of particles and systems: Harcourt Brace Jovanovich

Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover

Marsden, J.E., and Ratiu, T.S., (1999) Introduction to mechanics and symmetry: A basic exposition of classical mechanical systems (2*nd* edition): Springer-Verlag

Marsden, J.E., and Tromba, A., (1976) Vector calculus: W.H. Freeman & Co.

Martin, D., (2002) Manifold theory: An introduction for mathematical physicists: Horwood Publishing

Mase, G.E., (1970) Theory and problems of continuum mechanics: Schaum's Outlines, McGraw-Hill, Inc.

Matthews, P.C., (1998) Vector calculus: Springer-Verlag

Mavko, G., Mukerji, T., and Dvorkin, J., (1998) The rock physics book: Cambridge University Press

McCuskey, S.W., (1959) An introduction to advanced dynamics: Addison-Wesley Publishing Co.

McOwen, R.C., (1996) Partial differential equations: Methods and applications: Prentice-Hall, Inc.

Meirovitch, L., (1970) Methods of analytical dynamics: McGraw-Hill, Inc.

Menzel, D.H., (1947/1961) Mathematical physics: Dover

Menzel, D.H., (1955/1960) Fundamental formulas of physics: Dover

Morse P.M., and Feshbach H., (1953) Methods of theoretical physics: McGraw-Hill, Inc.

Musgrave, M.J.P., (1970) Crystal acoustics: Introduction to the study of elastic waves and vibrations in crystals: Holden-Day

Muskhelishvili, N.I., (1933/1975) Some basic problems of the mathematical theory of elasticity: Noordhoff International Publishing

Neusel, M.D., and Smith, L., (2002) Invariant theory of finite groups: American Mathematical Society

Novozhilov, V.V., (1948/1999) Foundations of the nonlinear theory of elasticity: Dover

Nye, J.F., (1957/1985) Physical properties of crystals: their representation by tensors and matrices: Oxford University Press

Nye, J.F., (1999) Natural focusing and fine structure of light: Caustics and wave dislocations: Institute of Physics Publishing

Officer, C.B., (1974) Introduction to theoretical geophysics: Springer-Verlag

Olmsted, J.M.H., (1961) Advanced calculus: Prentice-Hall, Inc.

Olver, P.J., (1993) Applications of Lie groups to differential equations (2*nd* edition): Springer-Verlag

Porteous, I.R., (1994) Geometric differentiation for the intelligence of curves and surfaces: Cambridge University Press

Pitt, V.H., (1977) The Penguin's dictionary of physics: Penguin books

Potter, M.C., (1978) Mathematical methods in the physical sciences: Prentice-Hall, Inc.

Poincaré, H., (1902/1968) La science et l'hypothèse: Flammarion

Pšenčík, I., (1994) Introduction to seismic methods (Lecture notes): PPPG/UFBa, Salvador

Radin, S.H. and Folk, R.T., (1982) Physics for scientists and engineers: Prentice-Hall, Inc.

Rasband, S. N., (1990) Chaotic dynamics of nonlinear systems: John Wiley & Sons

Rauch, J., (1991) Partial differential equations: Springer-Verlag

Rayleigh, J.W.S., (1877/1945) The theory of sound: Dover

Riley, K.F., (1974) Mathematical methods for the physical sciences: An informal treatment for students of physics and engineering: Cambridge University Press.

Rochester, M.G., (1986) Continuum mechanics (Lecture notes): Memorial University of Newfoundland

Rogister, Y., (2000) Etude théorique des deformations globales de la terre (Ph.D. thesis): Université de Liège

Ross, S.L., (1989) Introduction to ordinary differential equations (4*th* edition): John Wiley & Sons

Rund, H., (1959) The differential geometry of Finsler spaces: Springer-Verlag

Rund, H., (1966) The Hamilton-Jacobi theory in the calculus of variations: D. Van Nostrand

Rund, H., (1971) Invariant theory of variational problems on subspaces of a Riemannian manifold: Vandenhoeck & Ruprecht

Rymarz, C., (1993) Mechanika ośrodkow ciągłych: PWN — Polish Scientific Publishers

Sagan, H., (1969/1992) Introduction to the calculus of variations: Dover

Schoenberg, M., (1980) Elastic wave behaviour across linear slip interfaces: J. Acoust. Soc. Am., **68** (5), 1516 – 1521.

Schouten, J.A., (1951/1989) Tensor analysis for physicists: Dover

Schrödinger, E., (1996) Nature and the Greeks: Cambridge University Press

Schutz, B., (1980) Geometrical methods of mathematical physics: Cambridge University Press

Schwartz, L., (1966) Mathematics for the physical sciences: Addison-Wesley Publishing Co.

Segel, L.A., (1977/1987) Mathematics applied to continuum mechanics: Dover

Seto, W.W., (1971) Acoustics: Schaum's Outlines, McGraw-Hill, Inc.

Shen, Z., (2001) Differential geometry of spray and Finsler spaces: Kluwer Academic Publishers

Sheriff, R.E., (2002) Encyclopedic dictionary of applied geophysics: Society of Exploration Geophysicists

Sheriff, R.E., and Geldart, L.P., (1982) Exploration seismology: Cambridge University Press

Shilov, G.E., (1974/1996) Elementary functional analysis: Dover

Simmonds, J.G., (1994) A brief on tensor analysis: Springer-Verlag

Slater, J.C., and Frank, N.H., (1947) Mechanics: McGraw-Hill, Inc.

Slawinski, M.A., and Webster, P.S., (1999) On generalized ray parameters for vertically inhomogeneous and anisotropic media, Canadian Journal of Exploration Geophysics, **35**, No. 1/2, 28 – 31.

Slawinski, M.A., Slawinski, R.A, Brown, R.J., and Parkin, J.M., (2000) A generalized form of Snell's law in anisotropic media, Geophysics, **65**, No. 2, 632 – 637.

Slawinski, R.A., and Slawinski, M.A. (1999) On raytracing in constant velocity-gradient media: Calculus approach, Canadian Journal of Exploration Geophysics, **35**, No. 1/2, 24 – 27.

Slotnick, M.M., (1959) Lessons in seismic computing: Society of Exploration Geophysicists

Smith, D.R., (1974/1998) Variational methods in optimization: Dover

Spiegel, M.R., (1967) Theory and problems of theoretical mechanics with an introduction to Lagrange's equations and Hamiltonian theory: Schaum's Outlines, McGraw-Hill, Inc.

Spivak, M., (1970/1999) A comprehensive introduction to differential geometry: Publish or Perish, Inc.

Stewart, J., (1995), Multivariable calculus (*3rd* edition): Brooks/Cole Publishing Co.

Stolk, C.C., (2000) On the modeling and inversion of seismic data (Ph.D. thesis): Universiteit Utrecht

Struik, D.J., (1950/1988) Lectures on classical differential geometry: Dover

Synge, J.L., (1937/1962) Geometrical optics: An introduction to Hamilton's methods: Cambridge University Press

Synge, J.L., and Schild, A., (1949/1978) Tensor calculus: Dover

Talpaert, Y., (2001) Differential geometry with applications to mechanics and physics: Marcel Dekker, Inc.

Taylor, M.E., (1996) Partial differential equations; Basic theory: Springer-Verlag

Telford, W.M., Geldart, L.P., Sheriff, R.E., and Keys, D.A., (1976) Applied geophysics: Cambridge University Press

Ter Haar, D., (1961) Elements of Hamiltonian mechanics: North Holland

Thomsen, L., (1986) Weak elastic anisotropy: Geophysics, **51**, 1954 – 1966.

Thomson, W. (Lord Kelvin), (1890/1856) Mathematical and physical papers, Vol. III, elasticity, heat, electromagnetism: Cambridge University Press

Timoshenko, S.P., and Goodier, J.N., (1934/1987) Theory of elasticity: McGraw-Hill, Inc.

Ting, T.C.T., (1996) Anisotropic elasticity: Theory and applications: Oxford University Press

Toretti, R., (1999) The philosophy of physics: Cambridge University Press

Tsvankin, I., (2001) Seismic signatures and analysis of reflection data in anisotropic media: Pergamon

Tymoczko, T., (editor) (1986/1998) New directions in the philosophy of mathematics: Princeton University Press

Udías, A., (1999) Principles of seismology: Cambridge University Press

Wang, C.C., and Truesdell, C., (1973) Introduction to rational elasticity: Noordhoff International Publishing

Weinstock, R., (1952/1974) Calculus of variations with applications to physics and engineering: Dover

Weisstein, E.W., (1999) CRC concise encyclopedia of mathematics: CRC Press

Weyl, H., (1921/1952) Space-time-matter: Dover

Weyl, H., (1952) Symmetry: Princeton University Press

Wilcox, H.J., and Myers, D.L., (1978/1994) An introduction to Lebesgue integration and Fourier series: Dover

Winterstein, D.F., (1990) Velocity anisotropy terminology for geophysicists: Geophysics, **55**, 1070 – 1088.

Woodhouse, N.M.J., (1992) Geometric quantization (2*nd* edition): Oxford Science Publications

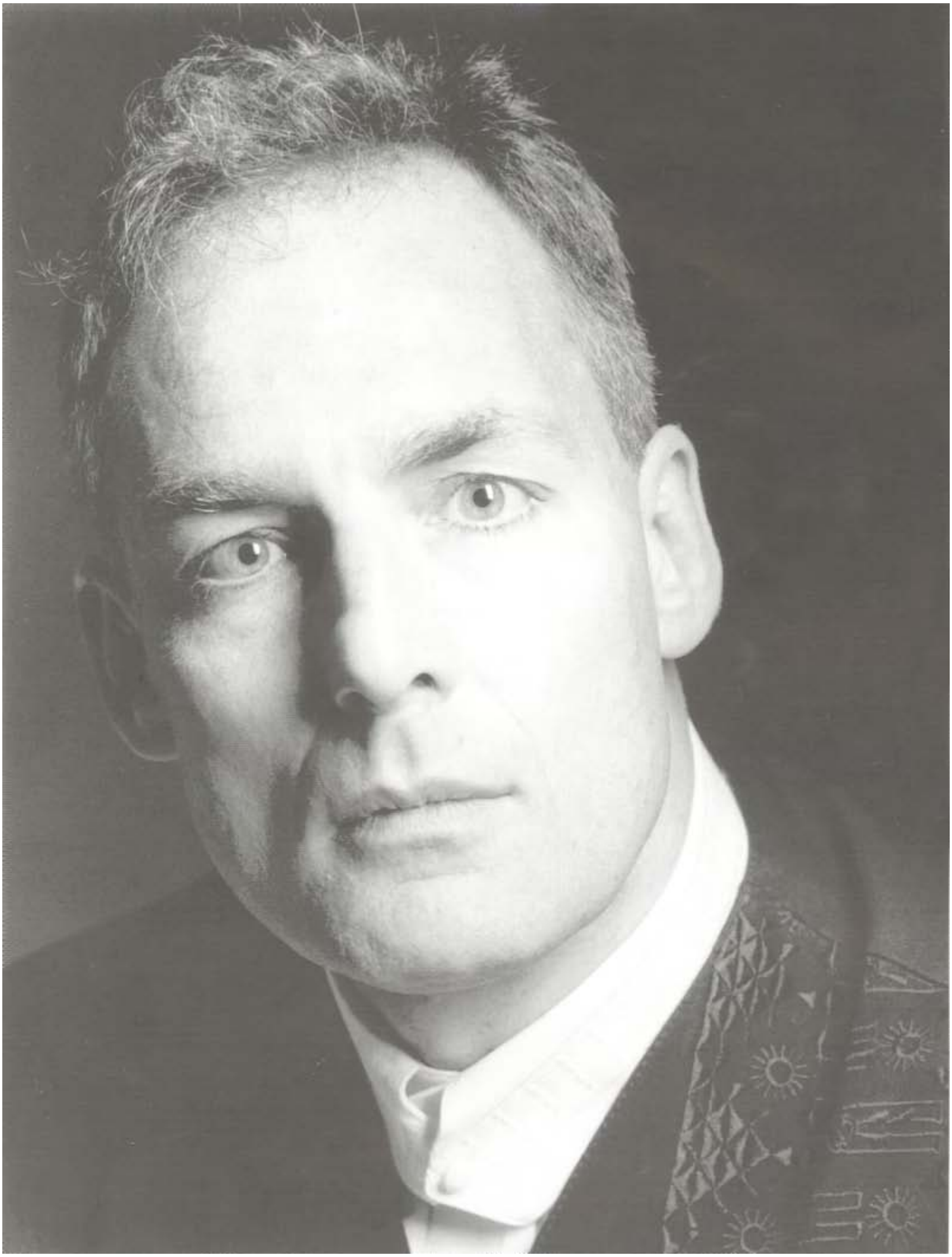
Woodhouse, N.M.J., (1987) Introduction to analytical dynamics: Oxford Science Publications

Zemanian, A.H., (1965/1987) Distribution theory and transform analysis: An introduction to generalized functions, with applications: Dover

Zill, D.G., and Cullen, M.R., (1997) Differential equations with boundary-value problems (4*th* edition): Brooks/Cole Publishing Co.

Zwillinger D., (1992) Book of differential equations (2*nd* edition): Academic Press

Index



This Page Intentionally Left Blank

About the author

Michael Slawinski was born in Warsaw in 1961. He studied at the University of Warsaw, University of Paris and, subsequently, at the University of Calgary. There, in 1988, he obtained his M.Sc. with a thesis entitled "Investigation of inhomogeneous body waves in an elastic/anelastic medium". In 1996, he obtained his Ph.D. from the University of Calgary with a thesis entitled "On elastic-wave propagation in anisotropic media: Reflection/refraction laws, raytracing and traveltime inversion".

Between 1986 and 1997, Slawinski worked for several years as a geophysicist in the petroleum industry. In 1998, he joined the faculty of the Department of Mechanical Engineering at the University of Calgary. In 2001, he assumed the position of Research Chair in Applied Seismology at Memorial University in St. John's, Newfoundland. At the same time, he became an adjunct associate professor in the Department of Mathematics and Statistics at the University of Calgary.

Since 1998, Slawinski has been the director of The Geomechanics Project — a theoretical-research group sponsored by industry — whose research focuses on seismic ray theory in the context of differential geometry and continuum mechanics.

This Page Intentionally Left Blank