

WAVES AND RAYS IN ELASTIC CONTINUA

Second edition

Michael A. Slawinski

Professor
Department of Earth Sciences
Memorial University
30-09-2007

Copyright ©2007 by Michael A. Slawinski

This work is licensed under the Creative Commons Attribution-No Derivative Works 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Contents

List of Figures	xiii
Dedication	xvii
Acknowledgements	xix
Preface	xxi
Changes to this edition	xxv
Part 1. Elastic continua	1
Introduction to Part 1	3
Chapter 1. Deformations	7
Preliminary remarks	7
1.1. Notion of continuum	8
1.2. Rudiments of continuum mechanics	9
1.2.1. Axiomatic format	9
1.2.2. Primitive concepts of continuum mechanics	9
1.3. Material and spatial descriptions	11
1.3.1. Fundamental concepts	11
1.3.2. Material time derivative	13
1.3.3. Conditions of linearized theory	15
1.4. Strain	18
1.4.1. Introductory comments	18
1.4.2. Derivation of strain tensor	18
1.4.3. Physical meaning of strain tensor	21
1.5. Rotation tensor and rotation vector	25
Closing remarks	26
1.6. Exercises	27
Chapter 2. Forces and balance principles	33
Preliminary remarks	33

2.1. Conservation of mass	33
2.1.1. Introductory comments	33
2.1.2. Integral equation	34
2.1.3. Equation of continuity	35
2.2. Time derivative of volume integral	36
2.3. Stress	38
2.3.1. Stress as description of surface forces	38
2.3.2. Traction	38
2.4. Balance of linear momentum	39
2.5. Stress tensor	41
2.5.1. Traction on coordinate planes	41
2.5.2. Traction on arbitrary planes	42
2.6. Cauchy's equations of motion	46
2.6.1. General formulation	46
2.6.2. Example: Surface-forces formulation	48
2.7. Balance of angular momentum	50
2.7.1. Introductory comments	50
2.7.2. Integral equation	51
2.7.3. Symmetry of stress tensor	52
2.8. Fundamental equations	55
Closing remarks	56
2.9. Exercises	57
Chapter 3. Stress-strain equations	63
Preliminary remarks	63
3.1. Rudiments of constitutive equations	63
3.2. Formulation of stress-strain equations: Hookean solid	65
3.2.1. Introductory comments	65
3.2.2. Tensor form	66
3.2.3. Matrix form	68
3.3. Determined system	70
3.4. Anelasticity: Example	70
3.4.1. Introductory comments	70
3.4.2. Viscosity: Stokesian fluid	70
3.4.3. Viscoelasticity: Kelvin-Voigt model	71
Closing remarks	73
3.5. Exercises	74

Chapter 4. Strain energy	79
Preliminary remarks	79
4.1. Strain-energy function	80
4.2. Strain-energy function and elasticity-tensor symmetry	81
4.2.1. Fundamental considerations	81
4.2.2. Elasticity parameters	82
4.2.3. Matrix form of stress-strain equations	83
4.2.4. Coordinate transformations	83
4.3. Stability conditions	84
4.3.1. Physical justification	84
4.3.2. Mathematical formulation	84
4.3.3. Constraints on elasticity parameters	84
4.4. System of equations for elastic continua	85
4.4.1. Elastic continua	85
4.4.2. Governing equations	86
Closing remarks	87
4.5. Exercises	89
Chapter 5. Material symmetry	93
Preliminary remarks	93
5.1. Orthogonal transformations	93
5.1.1. Transformation matrix	93
5.1.2. Symmetry group	94
5.2. Transformation of coordinates	94
5.2.1. Introductory comments	94
5.2.2. Transformation of stress-tensor components	95
5.2.3. Transformation of strain-tensor components	98
5.2.4. Stress-strain equations in transformed coordinates	99
5.3. Condition for material symmetry	100
5.4. Point symmetry	102
5.5. Generally anisotropic continuum	103
5.6. Monoclinic continuum	103
5.6.1. Elasticity matrix	103
5.6.2. Vanishing of tensor components	104
5.6.3. Natural coordinate system	105
5.7. Orthotropic continuum	107
5.8. Trigonal continuum	109

5.8.1. Elasticity matrix	109
5.8.2. Natural coordinate system	110
5.9. Tetragonal continuum	111
5.10. Transversely isotropic continuum	112
5.10.1. Elasticity matrix	112
5.10.2. Rotation invariance	112
5.11. Cubic continuum	116
5.12. Isotropic continuum	118
5.12.1. Elasticity matrix	118
5.12.2. Lamé's parameters	118
5.12.3. Tensor formulation	119
5.12.4. Physical meaning of Lamé's parameters	120
5.13. Relations among symmetry classes	121
Closing remarks	122
5.14. Exercises	124
Part 2. Waves and rays	139
Introduction to Part 2	141
Chapter 6. Equations of motion: Isotropic homogeneous continua	143
Preliminary remarks	143
6.1. Wave equations	143
6.1.1. Equation of motion	143
6.1.2. Wave equation for P waves	146
6.1.3. Wave equation for S waves	147
6.2. Plane waves	148
6.3. Displacement potentials	149
6.3.1. Helmholtz's decomposition	149
6.3.2. Gauge transformation	150
6.3.3. Equation of motion	151
6.3.4. P and S waves	152
6.4. Solutions of wave equation for single spatial dimension	154
6.4.1. d'Alembert's approach	154
6.4.2. Directional derivative	157
6.4.3. Well-posed problem	158
6.4.4. Causality, finite propagation speed and sharpness of signals	161
6.5. Solution of wave equation for two and three spatial dimensions	163

6.5.1. Introductory comments	163
6.5.2. Three spatial dimensions	163
6.5.3. Two spatial dimensions	165
6.6. On evolution equation	166
6.7. Solutions of wave equation for one-dimensional scattering	168
6.8. On weak solutions of wave equation	173
6.8.1. Introductory comments	173
6.8.2. Weak derivatives	174
6.8.3. Weak solution of wave equation	175
6.9. Reduced wave equation	176
6.9.1. Harmonic-wave trial solution	176
6.9.2. Fourier's transform of wave equation	177
6.10. Extensions of wave equation	179
6.10.1. Introductory comments	179
6.10.2. Standard wave equation	179
6.10.3. Wave equation and elliptical velocity dependence	180
6.10.4. Wave equation and weak inhomogeneity	183
Closing remarks	187
6.11. Exercises	189
 Chapter 7. Equations of motion: Anisotropic inhomogeneous continua	 209
Preliminary remarks	209
7.1. Formulation of equations	209
7.2. Formulation of solutions	210
7.2.1. Introductory comments	210
7.2.2. Trial-solution formulation: General wave	210
7.2.3. Trial-solution formulation: Harmonic wave	212
7.2.4. Asymptotic-series formulation	214
7.3. Eikonal equation	218
Closing remarks	220
7.4. Exercises	222
 Chapter 8. Hamilton's ray equations	 229
Preliminary remarks	229
8.1. Method of characteristics	229
8.1.1. Level-set functions	229
8.1.2. Characteristic equations	231
8.1.3. Consistency of formulation	233

8.2. Time parametrization of characteristic equations	234
8.2.1. General formulation	234
8.2.2. Equations with variable scaling factor	235
8.2.3. Equations with constant scaling factor	236
8.2.4. Formulation of Hamilton's ray equations	236
8.3. Physical interpretation of Hamilton's ray equations and solutions	237
8.3.1. Equations	237
8.3.2. Solutions	238
8.4. Relation between \mathbf{p} and $\dot{\mathbf{x}}$	238
8.4.1. General formulation	238
8.4.2. Phase and ray velocities	239
8.4.3. Phase and ray angles	241
8.4.4. Geometrical illustration	242
8.5. Example: Elliptical anisotropy and linear inhomogeneity	243
8.5.1. Introductory comments	243
8.5.2. Eikonal equation	243
8.5.3. Hamilton's ray equations	245
8.5.4. Initial conditions	245
8.5.5. Physical interpretation of equations and conditions	246
8.5.6. Solution of Hamilton's ray equations	247
8.5.7. Solution of eikonal equation	250
8.5.8. Physical interpretation of solutions	251
8.6. Example: Isotropy and inhomogeneity	252
8.6.1. Parametric form	252
8.6.2. Explicit form	252
Closing remarks	253
8.7. Exercises	255
Chapter 9. Christoffel's equations	267
Preliminary remarks	267
9.1. Explicit form of Christoffel's equations	268
9.2. Christoffel's equations and anisotropic continua	271
9.2.1. Introductory comments	271
9.2.2. Monoclinic continua	272
9.2.3. Transversely isotropic continua	275
9.3. Phase-slowness surfaces	281
9.3.1. Introductory comments	281

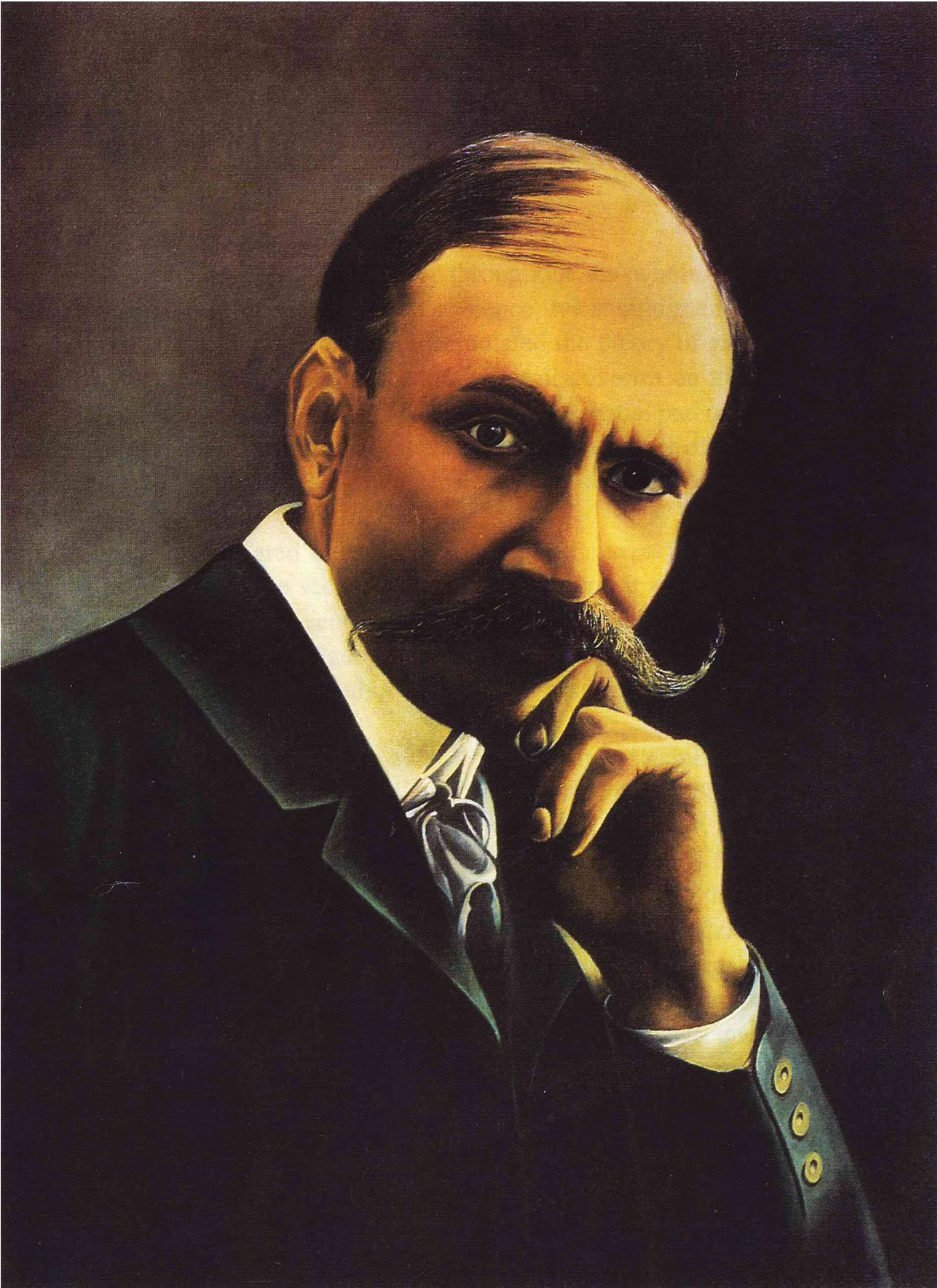
9.3.2. Convexity of innermost sheet	281
9.3.3. Intersection points	282
Closing remarks	284
9.4. Exercises	285
Chapter 10. Reflection and transmission	291
Preliminary remarks	291
10.1. Angles at interface	291
10.1.1. Phase angles	291
10.1.2. Ray angles	293
10.1.3. Example: Elliptical velocity dependence	294
10.2. Amplitudes at interface	296
10.2.1. Kinematic and dynamic boundary conditions	296
10.2.2. Reflection and transmission amplitudes	300
Closing remarks	304
10.3. Exercises	306
Chapter 11. Lagrange's ray equations	311
Preliminary remarks	311
11.1. Legendre's transformation of Hamiltonian	311
11.2. Formulation of Lagrange's ray equations	312
11.3. Beltrami's identity	314
Closing remarks	314
11.4. Exercises	315
Part 3. Variational formulation of rays	319
Introduction to Part 3	321
Chapter 12. Euler's equations	323
Preliminary remarks	323
12.1. Mathematical background	323
12.2. Formulation of Euler's equation	325
12.3. Beltrami's identity	327
12.4. Generalizations of Euler's equation	327
12.4.1. Introductory comments	327
12.4.2. Case of several variables	327
12.4.3. Case of several functions	328
12.4.4. Higher-order derivatives	329

12.5. Special cases of Euler's equation	329
12.5.1. Introductory comments	329
12.5.2. Independence of z	329
12.5.3. Independence of x and z	330
12.5.4. Independence of x	330
12.5.5. Total derivative	331
12.5.6. Function of x and z	331
12.6. First integrals	333
12.7. Lagrange's ray equations as Euler's equations	334
Closing remarks	335
12.8. Exercises	336
Chapter 13. Variational principles	343
Preliminary remarks	343
13.1. Fermat's principle	343
13.1.1. Statement of Fermat's principle	344
13.1.2. Properties of Hamiltonian \mathcal{H}	344
13.1.3. Variational equivalent of Hamilton's ray equations	345
13.1.4. Properties of Lagrangian \mathcal{L}	345
13.1.5. Parameter-independent Lagrange's ray equations	347
13.1.6. Ray velocity	348
13.1.7. Proof of Fermat's principle	348
13.2. Hamilton's principle: Example	349
13.2.1. Introductory comments	349
13.2.2. Action	350
13.2.3. Lagrange's equations of motion	352
13.2.4. Wave equation	353
Closing remarks	356
13.3. Exercises	357
Chapter 14. Ray parameters	365
Preliminary remarks	365
14.1. Traveltime integrals	366
14.2. Ray parameters as first integrals	366
14.3. Example: Elliptical anisotropy and linear inhomogeneity	367
14.3.1. Introductory comments	367
14.3.2. Rays	368
14.3.3. Traveltimes	371

14.4. Rays in isotropic continua	373
14.5. Lagrange's ray equations in xz -plane	373
14.6. Conserved quantities and Hamilton's ray equations	374
Closing remarks	376
14.7. Exercises	377
Part 4. Appendices	387
Introduction to Part 4	389
Appendix A. Euler's homogeneous-function theorem	391
Preliminary remarks	391
A.1. Homogeneous functions	391
A.2. Homogeneous-function theorem	392
Closing remarks	394
Appendix B. Legendre's transformation	395
Preliminary remarks	395
B.1. Geometrical context	395
B.1.1. Surface and its tangent planes	395
B.1.2. Single-variable case	396
B.2. Duality of transformation	397
B.3. Transformation between Lagrangian \mathcal{L} and Hamiltonian \mathcal{H}	398
B.4. Transformation and ray equations	399
Closing remarks	400
Appendix C. List of symbols	401
C.1. Mathematical relations and operations	401
C.2. Physical quantities	403
C.2.1. Greek letters	403
C.2.2. Roman letters	404
Appendix. Index	405
Bibliography	427
	442

List of Figures

1.4.1	Deformation: Uniaxial extension	23
1.4.2	Deformation: Relative change in angles	25
2.5.1	Index convention for stress-tensor components	42
2.5.2	Cauchy's tetrahedron	44
2.6.1	Surfaces forces	49
2.9.1	Angular momentum	61
5.13.1	Relations among symmetry classes	122
6.4.1	Domain of dependence and range of influence	162
9.4.1	Ray, phase and displacement angles	290



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 in the world 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

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. 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 Krebs, Michael Rochester, Yves Rogister, Raphaël Slawinski, Jędrzej Śniatycki and Chad Wheaton.

The second edition of this book benefited from further collaborations with several people stated above, namely, Andrej Bóna, Len Bos, Nelu Bucataru, David Dalton, Klaus Helbig, Misha Kotchetov, Michael Rochester and Yves Rogister, as well as Çağrı Diner, Peter Smith and Jason Soo. Norm Bleistein's review of the first edition provided feedback that benefited the second one. Furthermore, reviews of second-edition manuscripts by Vassily M. Babich, Ed Krebs and Małgorzata Seredyńska — as well as insightful questions from Andreas Atle and Ayiaz Kaderali — have all improved the final version.

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

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. As stated by Bunge in his book on “Philosophy of Science, Vol. I: From problem to theory”,

A nice illustration of the intertwining of empirical and theoretical events in the actual practice of science is offered by seismology, the study of elastic disturbances of Terra. [...] In conclusion, in order to “read” a seismogram so that it may become a set of data regarding an event (e.g., an earthquake) or an evidence relevant to a theory (e.g., about the inner structure of our planet), the seismologist employs elasticity theory and all the theories that may enter the design and interpretation of the seismograph.

The present book emphasizes the interdependence of mathematical formulation and physical meaning in the description of seismic phenomena. Herein, we 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 1*, 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 2*, we use these equations to identify the types of body

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.

To emphasize this statement of Poincaré, let us also consider the following quotation.

As the biggest library if it is in disorder is not as useful as a small but well-arranged one, so you may accumulate a vast amount of knowledge but it will be of far less value to you than a much smaller amount if you have not thought it over for yourself; because only through ordering what you know by comparing every truth with every other truth can you take complete possession of your knowledge and get it into your power.

Arthur Schopenhauer (1851) Parerga and Paralipomena, Volume 2

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 3*, 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 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 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 3* 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. While some exercises extend the topics covered, others are referred to in the main text. Reciprocally, the footnotes attached to these latter exercises refer the reader to the sections in the main text, where a given exercise is mentioned. Often, the exercises referred to in the main text supply steps that are omitted from the exposition in the text. 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 statement² of Marian Smoluchowski, Rudzki's colleague and friend.

Tematyka geofizyczna musiała nęcić Rudzkiego, tak wielkiego, fantastycznego miłośnika przyrody, z drugiej zaś strony ta właśnie nauka odpowiadała najwybitniejszej właściwości umysłu Rudzkiego, jego dążeniu do matematycznej ścisłości w rozumowaniu.³

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

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.

Changes to this edition

Mathematics gives to science the power of abstraction and generalization, and a symbolism that says what it has to say with the greatest possible clarity and economy.

John Lighton Synge and Byron A. Griffith (1949) Principles of mechanics

The conviction expressed by the above quote has inspired both the first and second editions of this book. Significant additions and modifications have been made to this second edition. These changes came about as a result of teaching from this text, readers' enquiries, and research into concepts discussed in the first edition. Summary of these changes and the purpose for their inclusion are given below.

Part 1:

Chapter 1: To familiarize the reader with the fundamentals that underly the seismic theory, we introduce in Section 1.2 the three rudimentary concepts of continuum mechanics, namely, material body, manifold of experience and system of forces.

Chapter 2: To deepen the reader's understanding of the arguments used to formulate the balance principles of continuum mechanics, we include in Sections 2.4 and 2.7, the particle-mechanics motivation of the balance of linear momentum and the balance of angular momentum, respectively. To make the reader aware of possible extensions within continuum mechanics, we discuss in Section 2.7 the distinction between the strong and weak forms of Newton's third law of motion. As a results, we show that the symmetry of the stress tensor is a consequence of the strong form, and not a fundamental physical law.

Chapter 3: To deepen the reader's understanding of the fundamentals that underly the formulation of seismic theory, we introduce in Section 3.1 the three principles of constitutive equations, namely, determinism, local action and objectivity. To provide the background that allows us to emphasize the meaning of elasticity, we introduce in Section 3.4 constitutive equations of anelastic continua.

Chapter 5: Following a demonstration that only eight classes of the elasticity tensor exist,⁴ all these classes are included in this edition. The addition consists of Sections 5.8 and 5.11, where we study trigonal and cubic continua.

⁴Readers interested in the existence of the eight symmetry classes might refer to Bóna, A., Bucataru, I., and Slawinski, M.A., (2005) Characterization of elasticity-tensor symmetries using SU(2). *Journal of Elasticity* **75**(3), 267 – 289, and to Bóna, A., Bucataru, I., and Slawinski, M.A., (2004) Material symmetries of elasticity tensor. *The Quarterly Journal of Mechanics and Applied Mathematics* **57**(4), 583 – 598.

Part 2

Chapter 6: Since this chapter deals with the wave equation, which plays an important role in seismology, we wish to deepen the reader's understanding of different aspects of this equation. In Section 6.4, we discuss well-posedness of the wave equation with its initial conditions. To provide further insight into the properties of the wave equation, we compare it in Section 6.6 to the evolution equation, which is ill-posed. Another significant addition consists of the formulation and study of solutions of the wave equation in two and three spatial dimensions, as presented in Section 6.5. Therein, we discuss also the range of influence and the domain of dependence of these solutions. To familiarize the reader with the concepts of reflection and transmission in the context of the wave equation, we study the solutions of one-dimensional scattering, which is discussed in Section 6.7. To familiarize the reader with the several-century-long debate concerning the applicability of the wave equation, which is a differential equation, in order to study nondifferentiable solutions, we present the theory of distributions and its application to the wave equation. The crux of this addition is Section 6.8. Furthermore, in Section 6.9, we elaborate on the concept of the reduced wave equation and Fourier's transport of the wave equation, which is used in the following chapter to discuss trial solutions of the equations of motion.

Chapter 7: Since this chapter deals with the equations of motion whose trial solutions necessitate asymptotic methods, we wish to familiarize the reader with the concept of the asymptotic series in the context of ray theory, which is commonly referred to as the asymptotic ray theory. This addition is contained in Sections 7.2.3 and 7.2.4.

Chapter 8: Since Hamilton's ray equations are the basis of ray theory, we wish to deepen the reader's understanding of their meaning, as well as the meaning of their solutions. To do so, we present an analytic solution of Hamilton's ray equations, which allows us to gain both mathematical and physical insights into Hamilton's equations. Furthermore, an equivalent solution is presented using Lagrange's ray equations. Hence, this analytic solution allows us to relate the Hamiltonian and Lagrangian formulations of ray theory — two distinct approaches. The crux of this addition consists of Section 8.5.

Chapter 9: The convexity and detachment of the innermost phase-slowness surface are commonly invoked in seismology. In this edition, we show particular cases for which this surface is not detached, and we comment on the fact that detachment is not necessary to prove the convexity of the innermost surface.

Chapter 10: In this edition, we use Newton's third law of motion — rather than the second one used in the first edition — to derive in Section 10.2.1, the dynamic boundary conditions on the interface between two media. Since we are dealing with two media acting on one another, the third law lends itself more naturally than the second one for which we had to treat the interface as a layer with a vanishing thickness.

Part 3

Chapter 14: In this edition, the traveltime expression for a signal in a continuum exhibiting an elliptical velocity dependence with direction and a linear dependence with depth is valid for the entire ray: both the downgoing and upgoing segments. In the first edition, the expression was valid for a downgoing ray only.

In this edition, the chapter on Lagrange's ray equations is found at the end of *Part 2*, rather than immediately after the chapter on Hamilton's ray equations. This placement of the chapter emphasizes the physical motivation; namely, we can study the entire ray theory in the context of Hamilton's ray equations. Lagrange's ray equations, on the other hand, open a new path of study to which we devote *Part 3*.

As a result of teaching from the first edition, the second edition acquired more exercises and figures. There is also a *List of Figures* included.

With the above additions, the title of the book has been modified to emphasize the generality of the approach and the importance of the concept of continuum. Thus, "Seismic waves and rays in elastic media" became "Waves and rays in elastic continua".

Part 1

Elastic continua

Introduction to Part 1

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.⁵

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”.

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.

⁵Readers 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. Readers interested in this statement in the context of variational mechanics, which we will discuss in Section 13.2, might refer to Poincaré, H., (1902/1968) La science et l’hypothèse: Flammarion, pp. 219 – 225.

Also, readers might refer to Einstein, A., and Infeld, L., (1938) Evolution of physics from early concepts to relativity and quanta: Simon & Schuster, p. 125:

During the second half of the nineteenth century new and revolutionary ideas were introduced into physics; they opened the way to a new philosophical view, differing from the mechanical one.

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 that is, however, general enough to be of significance in applied seismology. Our attention focuses on elastic continua. Any 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 — to which we refer to as the strain 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 too literal an 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.

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.² In a seismological context, the Earth is regarded as a continuum that transmits mechanical disturbances. The notion of continuum allows us to describe the deformations and forces experienced by a deformable body in terms of strains and stresses 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

... 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;

²Readers interested in the concept of models and physical understanding might refer to Weinert, F., (2005) The scientist as philosopher: Springer-Verlag, pp. 45 – 47.

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.³

In the context of the philosophy of science, continuum mechanics is associated with the concept of emergence, which is also called methodological holism. In physics, emergence is used to describe properties, laws or phenomena that occur at macroscopic scales but not at microscopic ones, in spite of the fact that a macroscopic system can be viewed as an ensemble of microscopic ones. As an example, let us consider colour. Elementary particles, such as protons or electrons, have no colour. Colour emerges if these particles are arranged in atoms, which absorb or emit specific wavelengths of light and can thus be said to have colour. Emergent concepts in continuum mechanics are elasticity, rigidity, viscosity, friction, and so on. The approach in which we invoke elementary particles to describe properties of matter belongs to the field of the condensed-matter physics, and philosophically is associated also with the concept of reductionism and so-called methodological individualism.

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 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. To gain further insight into our statement, let us consider the following quote from “La science et l’hypothèse” of Poincaré.

³Readers 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.

Dans la plupart des questions, l'analyste suppose, au debut de son calcul, soit que la matière est continue, soit, inversement, qu'elle est formée d'atomes. Il aurait fait le contraire que ses résultats n'en auraient pas été changés; il aurait eu plus de peine à les obtenir, voilà tout.⁴

1.2. Rudiments of continuum mechanics

1.2.1. Axiomatic format. Modern continuum mechanics is a physical theory that adopts an axiomatic format rather than a historical exposition or a heuristic approach. In this format, the structure of the theory is a hypothetico-deductive system. In other words, it is a system that starts from a set of hypotheses and proceeds deductively; hence, conclusions are of no greater generality than premises, and conclusions are as certain as the premises.

To axiomatize a theory, we need to lay down a set of primitive concepts such that there is a sufficient characterization of basic ideas and a platform for the subsequent statements of the theory. To do so, we need to use the language of mathematics and the principles of logic. In other words, an axiomatic theory of physics presupposes both logic and mathematics; it requires them as a formal apparatus of description. However, this apparatus does not suffice to construct a physical theory because it is devoid of intrinsic physical meaning. Herein, we might recall that the purpose of physics is to accurately describe physical phenomena, while the purpose of logic and mathematics is to consistently define abstract concepts. To axiomatize a physical theory is to articulate its explicitness. Indeed, only explicitly articulated theories can be tested.

Among the advantages of an axiomatic format are the recognition of presupposition and assumptions contained within the theory, as well as the rigour and consistency, which provide clarity and foster coherent developments of the theory. Among the inconveniences of an axiomatic approach are the propensity for formalism at the expense of intuition, as well as the propensity for generality at the expense of concreteness. Since the purpose of this book is concrete — namely, the description of wave phenomena in elastic continua — we trust that our presentation might enjoy the advantages of the axiomatic format and avoid its inconveniences.

1.2.2. Primitive concepts of continuum mechanics.

Introductory comments. The primitive basis of a physical theory is a set of formal concepts that are assigned a physical meaning. These concepts cannot be proven within a given theory but are assumed to be true.⁵ The three primitive concepts upon which the continuum mechanics is formed are the material body,

⁴For most problems, the researcher assumes, at the beginning of his calculations, that either the matter is continuous or, otherwise, it is composed of atoms. If he makes the other assumption, his results do not change; they will be just more difficult to obtain.

⁵Note that, as stated on page 4, only abstract statements can be proven. Herein, the primitive concepts, although abstract, cannot be proven within the theory whose foundation they form. If they could be proven, they would be a part of the theory and not its primitive basis.

the manifold of physical experience, and the system of internal forces.⁶ These concepts are consistent with each other since the set they form is free of contradictions. They exhibit a weak deductive completeness since all known statements within the theory can be formulated in the context of the three primitive concepts, but not every statement is entailed by these three concepts — hence, the development of the theory can continue to accommodate new ideas and observations.

Material body. The first primitive concept is the material body, \mathcal{B} . It is a Euclidean three-dimensional smooth manifold composed of material points \mathbf{X} , 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. 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.⁷ The body manifold possesses the following properties. Every sufficiently smooth portion of a body is a body. Also, there is a measure called mass, m , which is a nonnegative scalar quantity such that

$$m(\mathcal{B}_1 \cup \mathcal{B}_2) = m(\mathcal{B}_1) + m(\mathcal{B}_2),$$

where \mathcal{B}_1 and \mathcal{B}_2 are disjoint subsets of \mathcal{B} . In other words, the mass of a body is the sum of the masses of its parts. Furthermore, mass of a material body occupying volume V is

$$m(\mathcal{B}) = \iiint_V \rho \, dV,$$

where ρ is the mass density of the material composing \mathcal{B} .

As an abstract entity, \mathcal{B} is not accessible to direct observations. We encounter its representations at particular times, t , and in particular spatial locations, \mathbf{x} . Points t and \mathbf{x} constitute a manifold of physical experience, which is the second primitive concept of the theory of continuum mechanics.

Manifold of physical experience. The second primitive concept is the manifold of physical experience. It is a Euclidean space-time $\mathbb{E}^3 \times t$, where \mathbb{E}^3 is composed of points \mathbf{x} and is endowed with the Euclidean metric, while t denotes time. Symbol \times stands for the direct product, which is a set of all possible ordered pairs (\mathbf{x}, t) , which we can write explicitly as (x_1, x_2, x_3, t) . We assume that a given time interval is the same for all observers — the time is absolute and universal. Also, we assume that the distance between two given locations is the same for all observers — the space is absolute. These assumptions are tantamount to limiting our study to nonrelativistic continuum mechanics. In the encounters with \mathcal{B} , we consider a sequence of configurations given by

$$(1.2.1) \quad \mathbf{x} = \chi(\mathbf{X}, t),$$

⁶Readers interested in more detail and an elegant exposition might refer to Truesdell, C., (1966) Six lectures on modern natural philosophy: Springer-Verlag, pp. 2 – 3 and pp. 96 – 97, and to Bunge, M., (1967) Foundations of physics: Springer-Verlag, pp. 143 – 157.

⁷According to Coirier, J., (1997) Mécanique des milieux continus: Concepts de base: Cours et exercices corrigés: Dunod, pp. 4 – 5, the size of such an element of volume is of the order of $10^{-15} \, m^3$ and contains of the order of 10^{10} molecules.

where χ is the motion of \mathcal{B} given for a fixed t by an isomorphism: a structure-preserving continuous map between topological spaces that is injective, which means that to every element of one set there corresponds at most one element of the other set, and surjective, which means that to every element of one set there corresponds at least one element of the other set; thus the mapping between \mathbf{X} and \mathbf{x} is one-to-one. Hence, two distinct elements of \mathcal{B} cannot occupy the same point on the manifold of physical experience, and no single element of \mathcal{B} can occupy two distinct points. Since χ is an isomorphism, its inverse is continuous. Thus, we can write

$$\mathbf{X} = \chi^{-1}(\mathbf{x}, t),$$

which allows us to consider the spatial description, as opposed to the material one given by equation (1.2.1). In a material description, the observer identifies the location of points $\mathbf{X} \in \mathcal{B}$ that are immersed in \mathbb{E}^3 ; material point (X_1, X_2, X_3) and time t are the independent variables of this description. In a spatial description, the observer cannot identify particular material points but only spatial locations $\mathbf{x} \in \mathbb{E}^3$; in other words, the observer is unable to follow the displacement of a given material point. In this description, location (x_1, x_2, x_3) and time t are the independent variables. We will discuss these two descriptions in more detail in Section 1.3.

Note that the goal of theoretical sciences is to give the best possible conceptual representation of a given system. The best representation is as close as possible to an isomorphism, which is attainable only in mathematics. Isomorphism is a perfect formal analogy; herein, whatever happens in \mathbf{X} has its isomorphic image in \mathbf{x} , and vice versa. Commonly, the same system can be represented in a variety of inequivalent ways.

System of internal forces. The third primitive concept is the system of internal forces within body \mathcal{B} . These forces are described by establishing Cauchy's stress principle, which can be stated in the following way.

Cauchy's stress principle: The action of the material occupying a portion within the body that is exterior to the closed surface on the material within this surface is represented by vector field \mathbf{T} .

Vector \mathbf{T} is called traction. It acts on a surface whose outward normal unit vector is \mathbf{n} . \mathbf{T} is assumed to depend continuously on \mathbf{n} . Its physical dimensions are force per unit area. The concept of the stress tensor that is invoked by the system of internal forces is central to the theory of continuum mechanics. We will discuss it in detail in Section 2.3.

1.3. Material and spatial descriptions

1.3.1. 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. In continuum mechanics, we can describe such a displacement in at

least two ways; namely, by studying material and spatial descriptions.⁸ 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. As stated in Section 1.2.2, the first approach is called the material description of motion while the second one is called the spatial description of motion. These approaches are also known as the Lagrangian description and the Eulerian description, respectively.⁹

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.

To gain insight into the meaning of the material and spatial descriptions, consider a moving continuum and let the observer focus the 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 1.6.1, which follows Exercise 1.1. At a later time t , the position vector of the material point \mathbf{X} is given by \mathbf{x} . With a certain abuse of notation, we write expression (1.2.1) as $\mathbf{x}(\mathbf{X}, t)$; this mapping gives position, \mathbf{x} , of material point \mathbf{X} at time t . This is the material description, where variable \mathbf{X} identifies the material point. We assumed that, for a given time t , this mapping is one-to-one and 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. Again, with a certain abuse of notation, 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 introduce the material and spatial coordinates, consider an orthonormal coordinate system, where

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

and

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

with the components x_i and X_i being the spatial and material coordinates, respectively.

⁸Material 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.

⁹Readers 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. Readers interested in heuristic descriptions might refer to Snieder, R., (2004) A guided tour of mathematical methods for the physical sciences: Cambridge University Press, pp. 57 – 61.

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

1.3.2. 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

$$(1.3.1) \quad \frac{dF}{dt} = \left. \frac{dF(\mathbf{X}, t)}{dt} \right|_{\mathbf{X}},$$

where symbol $|_{\mathbf{X}}$ means that the derivative is evaluated at \mathbf{X} .

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

$$(1.3.2) \quad \frac{\partial f}{\partial t} = \left. \frac{\partial f(\mathbf{x}, t)}{\partial t} \right|_{\mathbf{x}},$$

where symbol $|_{\mathbf{x}}$ means that the derivative is evaluated at \mathbf{x} .

The material and spatial descriptions are related by the chain rule of differentiation. To see this relation, consider a three-dimensional continuum and explicitly write

$$(1.3.3) \quad F(X_1, X_2, X_3, t) = f(x_1(X_1, X_2, X_3, t), x_2(X_1, X_2, X_3, t), x_3(X_1, X_2, X_3, t), t).$$

Taking the time derivative of both sides, we get

$$\frac{dF}{dt} = \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial t} + \frac{\partial f}{\partial x_3} \frac{\partial x_3}{\partial t} \right] \Big|_{\mathbf{X}}.$$

Since the derivative is evaluated for a given material point, \mathbf{X} , it implies that $\partial x_i / \partial t$ are the components of velocity of this point moving in space; we will write these components as v_i . Thus,

$$(1.3.4) \quad \frac{dF}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x_1} v_1 + \frac{\partial f}{\partial x_2} v_2 + \frac{\partial f}{\partial x_3} v_3.$$

As shown in expression (1.3.1), dF/dt describes the temporal variation of a given quantity for a particular material point within the continuum, and as shown in expression (1.3.2), $\partial f / \partial t$ describes the temporal variation of this quantity at a particular point in space. To discuss the relation between dF/dt and $\partial f / \partial t$, we write the last three terms on the right-hand side of the above equation as a scalar product to get

$$\frac{dF}{dt} = \frac{\partial f}{\partial t} + [v_1, v_2, v_3] \cdot \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right].$$

Recognizing that the second term in brackets is the gradient of function f , and denoting the velocity vector by $\mathbf{v} = [v_1, v_2, v_3]$, we write

$$\frac{dF}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f,$$

which is the relation between the time derivatives of F and f .

We can formally rewrite the right-hand side of the above equation as

$$(1.3.5) \quad \frac{dF(\mathbf{X}, t)}{dt} = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) f(\mathbf{x}, t),$$

where the term in parentheses is an operator acting on function f , and the material and spatial coordinates are related by

$$(1.3.6) \quad \mathbf{x} = \mathbf{x}(\mathbf{X}, t).$$

The term in parentheses on the right-hand side of equation (1.3.5) is called the material time-derivative operator. It can be applied to a scalar, to a vector, or to a tensor function of position and time coordinates. It is common to denote this operator by D/Dt so as to concisely write equation (1.3.5) as

$$\frac{dF(\mathbf{X}, t)}{dt} = \frac{Df(\mathbf{x}, t)}{Dt},$$

where $D/Dt := \partial/\partial t + \mathbf{v} \cdot \nabla$.

Examining the left-hand side of equation (1.3.5) in view of expression (1.3.1), we conclude that the material time derivative is a rate of change associated with a particular element of the continuum. In other words, it is measured by an observer travelling with this element. Mathematically, the material time derivative is the time derivative with material coordinates held constant. To consider spatial coordinates in this context, let us examine the right-hand side of equation (1.3.5). The first term of $(\partial/\partial t + \mathbf{v} \cdot \nabla)$ describes the time rate of change at the location \mathbf{x} , while the second term describes the rate of change associated with the motion of material points; more explicitly, the second term describes the spatial rate of change of material point \mathbf{X} moving with velocity \mathbf{v} . Recalling expression (1.3.3), we can write expression (1.3.4) with a certain abuse of notation as

$$\frac{dF}{dt} = \left. \frac{\partial f}{\partial t} \right|_{\mathbf{x}} + \sum_{i=1}^3 \left. \frac{\partial f}{\partial x_i} \right|_t v_i|_{\mathbf{x}},$$

where $\partial f/\partial t$ is evaluated at location \mathbf{x} , $\partial f/\partial x_i$ is evaluated at instant t for a given material point \mathbf{X} moving with velocity \mathbf{v} .

As we see from the formulation presented above, in general, dF/dt differs from $\partial f/\partial t$ by term $\mathbf{v} \cdot \nabla f$.¹⁰ This term vanishes in the absence of motion, $\mathbf{v} = \mathbf{0}$, or if f does not vary spatially, $\nabla f = \mathbf{0}$. Also, if

¹⁰Readers interested in this term and its meaning as the convective derivative might refer to Sedov, L.I., (1971) A course in continuum mechanics: Wolters-Noordhoff Publishing, Vol. I, pp. 31 – 32.

this term is negligible, we need not distinguish between the material and spatial descriptions. This is an important concept that often allows us simplify our approach; we will discuss it in the next section.

1.3.3. 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.¹¹

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 1.3.2, 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

$$(1.3.7) \quad \mathbf{U}(\mathbf{X}, t) = \mathbf{x}(\mathbf{X}, t) - \mathbf{X},$$

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

$$(1.3.8) \quad \mathbf{u}(\mathbf{x}, t) = \mathbf{x} - \mathbf{X}(\mathbf{x}, t).$$

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

Since the same quantity is given by expressions (1.3.7) and (1.3.8), we can write

$$(1.3.9) \quad \mathbf{U}(\mathbf{X}, t) = \mathbf{u}(\mathbf{x}, t),$$

where the material and spatial coordinates are related by equation (1.3.6).

¹¹Readers 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.

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

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

where $i \in \{1, 2, 3\}$. Assuming that the gradient of the displacement, which is shown in brackets, is vanishingly small, we can consider only the first term of the series. Thus, we can write

$$(1.3.10) \quad \mathbf{U}(\mathbf{X}, t) \approx \mathbf{U}(\mathbf{x}, t).$$

Hence, expression (1.3.7) can be written as

$$(1.3.11) \quad \mathbf{U}(\mathbf{x}, t) \approx \mathbf{x}(\mathbf{X}, t) - \mathbf{X}.$$

Since in expression (1.3.11) \mathbf{U} is a function of \mathbf{x} , we rewrite — using expression (1.3.8) — the displacement as a function of \mathbf{x} to get

$$(1.3.12) \quad \mathbf{U}(\mathbf{x}, t) \approx \mathbf{x} - \mathbf{X}(\mathbf{x}, t).$$

Comparing expressions (1.3.8) and (1.3.12), we see that

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

Thus, in view of expression (1.3.10), we conclude that — for infinitesimal displacements — we can write

$$(1.3.13) \quad \mathbf{U}(\mathbf{X}, t) \approx \mathbf{u}(\mathbf{x}, t).$$

To gain insight into the meaning of this result, we examine equations (1.3.9) and (1.3.13). Equation (1.3.9) states that $\mathbf{U} = \mathbf{u}$, with \mathbf{x} related to \mathbf{X} by equation (1.3.6). Equation (1.3.13) states that $\mathbf{U} \approx \mathbf{u}$, where we can simply replace \mathbf{x} by \mathbf{X} , without invoking equation (1.3.6). 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 (1.3.5) becomes

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

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

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

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 (1.3.5) becomes

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

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

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

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 (2.6.3).

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, $d\mathbf{U}/dt \approx \partial \mathbf{u}/\partial t$ and $d^2 \mathbf{U}/dt^2 \approx \partial^2 \mathbf{u}/\partial t^2$, respectively.

It is important to 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.¹²

In this section, we attempted to justify the linearization by *a priori* arguments. Let us also mention another approach that plays an important role in continuum mechanics and consists of an *a posteriori* justification; in other words, the results justify the assumptions. Herein, the initial approach could be to try the linear formulation by ignoring the nonlinear terms without invoking any physical reason; it is common to begin with a linear formulation and to investigate its applicability. Having developed a theory based on such a formulation, we would compare the predictions of this theory with experimental results. If the agreement between the theory and experiments is satisfactory, we could accept the assumptions. In our context, many seismological measurements agree with theoretical prediction of a linearized theory, thus providing an *a posteriori* justification.

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 2 and 3 of the book, refer to the propagation velocities; namely, 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.

¹²Readers interested in details of this linearization might refer to Achenbach, J.D., (1973) Wave propagation in elastic solids: North Holland, pp. 17 – 21.

1.4. Strain

1.4.1. Introductory comments¹³. 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 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.

1.4.2. Derivation of strain tensor. In this section we show that the strain tensor describes the deformation within the continuum.

The concept of deformation implies that distances among points within the continuum change. To derive the strain tensor in a three-dimensional continuum, consider two infinitesimally close points therein whose coordinates are given by $[x, y, z]$ and $[x + dx, y + dy, z + dz]$. The square of the distance between these points is given by

$$(1.4.1) \quad (ds)^2 = (dx)^2 + (dy)^2 + (dz)^2.$$

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

$$(1.4.2) \quad \left[x + u_x|_{x,y,z}, y + u_y|_{x,y,z}, z + u_z|_{x,y,z} \right],$$

while the coordinates of the second point are given by

$$(1.4.3) \quad \left[\begin{aligned} &x + dx + u_x|_{x+dx,y+dy,z+dz}, \\ &y + dy + u_y|_{x+dx,y+dy,z+dz}, \\ &z + dz + u_z|_{x+dx,y+dy,z+dz} \end{aligned} \right],$$

¹³Readers 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.

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 (1.4.2) from the corresponding components given in expression (1.4.3), we obtain the difference between the corresponding coordinates of the two points after the deformation; namely,

$$(1.4.4) \quad \begin{aligned} & \left[dx + u_x|_{x+dx, y+dy, z+dz} - u_x|_{x, y, z}, \right. \\ & dy + u_y|_{x+dx, y+dy, z+dz} - u_y|_{x, y, z}, \\ & \left. dz + u_z|_{x+dx, y+dy, z+dz} - u_z|_{x, y, z} \right]. \end{aligned}$$

Examining expression (1.4.4), we can gain insight into the physical meaning of the displacement vector. In general, u_x , u_y and u_z are each a function of x , y and z . Thus, the length and orientation of vector $\mathbf{u} = [u_x, u_y, u_z]$ depends on position. Hence, in general, the displacement vector at $x + dx$, $y + dy$, $z + dz$ has different length and orientation than it has at x , y , z ; for instance, $u_x|_{x+dx, y+dy, z+dz}$ differs from $u_x|_{x, y, z}$. The dependence of \mathbf{u} on position results — upon application of \mathbf{u} — in a relative change of positions of points within the continuum. Prior to deformation, the three coordinates of our two points were separated by dx , dy and dz , respectively. After the deformation, they are separated by different amounts, as shown in expression (1.4.4). We can get further insight into \mathbf{u} by considering special cases. If \mathbf{u} is given by constants, then its components are the same at all positions. In such a case, expression (1.4.4) reduces to $[dx, dy, dz]$. This means that there is no change of positions of points within the continuum relative to each other. In such a case, we can view \mathbf{u} as resulting in the translation of the whole medium without deformation. Other special cases are discussed in Section 1.4.3. Let us return to our derivation. In view of infinitesimal displacements, the components of \mathbf{u} that are evaluated in expression (1.4.4) at $x + dx$, $y + dy$, $z + dz$ can be approximated by the first two terms of Taylor's series about (x, y, z) ; namely,

$$\begin{aligned} u_x|_{x+dx, y+dy, z+dz} &\approx u_x|_{x, y, z} + \frac{\partial u_x}{\partial x} \Big|_{x, y, z} dx + \frac{\partial u_x}{\partial y} \Big|_{x, y, z} dy + \frac{\partial u_x}{\partial z} \Big|_{x, y, z} dz, \\ u_y|_{x+dx, y+dy, z+dz} &\approx u_y|_{x, y, z} + \frac{\partial u_y}{\partial x} \Big|_{x, y, z} dx + \frac{\partial u_y}{\partial y} \Big|_{x, y, z} dy + \frac{\partial u_y}{\partial z} \Big|_{x, y, z} dz \end{aligned}$$

and

$$u_z|_{x+dx, y+dy, z+dz} \approx u_z|_{x, y, z} + \frac{\partial u_z}{\partial x} \Big|_{x, y, z} dx + \frac{\partial u_z}{\partial y} \Big|_{x, y, z} dy + \frac{\partial u_z}{\partial z} \Big|_{x, y, z} dz.$$

Inserting these terms into expression (1.4.4) and simplifying, we obtain the approximation for the difference of the corresponding coordinates of the two points after the deformation; namely,

$$\begin{aligned} & \left[dx + \frac{\partial u_x}{\partial x} \Big|_{x, y, z} dx + \frac{\partial u_x}{\partial y} \Big|_{x, y, z} dy + \frac{\partial u_x}{\partial z} \Big|_{x, y, z} dz, \right. \\ & dy + \frac{\partial u_y}{\partial x} \Big|_{x, y, z} dx + \frac{\partial u_y}{\partial y} \Big|_{x, y, z} dy + \frac{\partial u_y}{\partial z} \Big|_{x, y, z} dz, \\ & \left. dz + \frac{\partial u_z}{\partial x} \Big|_{x, y, z} dx + \frac{\partial u_z}{\partial y} \Big|_{x, y, z} dy + \frac{\partial u_z}{\partial z} \Big|_{x, y, z} dz \right]. \end{aligned}$$

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

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

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

$$\begin{aligned}
 (1.4.5) \quad (d\check{s})^2 &\approx (dx)^2 + (dy)^2 + (dz)^2 \\
 &+ 2 \left(\frac{\partial u_x}{\partial x} \Big|_{x,y,z} (dx)^2 + \frac{\partial u_y}{\partial y} \Big|_{x,y,z} (dy)^2 + \frac{\partial u_z}{\partial z} \Big|_{x,y,z} (dz)^2 \right. \\
 &+ \frac{\partial u_x}{\partial y} \Big|_{x,y,z} dx dy + \frac{\partial u_x}{\partial z} \Big|_{x,y,z} dx dz + \frac{\partial u_y}{\partial x} \Big|_{x,y,z} dx dy \\
 &\left. + \frac{\partial u_y}{\partial z} \Big|_{x,y,z} dy dz + \frac{\partial u_z}{\partial x} \Big|_{x,y,z} dx dz + \frac{\partial u_z}{\partial y} \Big|_{x,y,z} dy dz \right),
 \end{aligned}$$

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

Using expressions (1.4.1) and (1.4.5), we obtain the difference in the square of the distance between the two points that results from the deformation; namely,

$$\begin{aligned}
 (d\check{s})^2 - (ds)^2 &\approx 2 \left[\frac{\partial u_x}{\partial x} \Big|_{x,y,z} (dx)^2 + \frac{\partial u_y}{\partial y} \Big|_{x,y,z} (dy)^2 + \frac{\partial u_z}{\partial z} \Big|_{x,y,z} (dz)^2 \right. \\
 &+ \left(\frac{\partial u_x}{\partial y} \Big|_{x,y,z} + \frac{\partial u_y}{\partial x} \Big|_{x,y,z} \right) dx dy + \left(\frac{\partial u_y}{\partial z} \Big|_{x,y,z} + \frac{\partial u_z}{\partial y} \Big|_{x,y,z} \right) dy dz \\
 &\left. + \left(\frac{\partial u_x}{\partial z} \Big|_{x,y,z} + \frac{\partial u_z}{\partial x} \Big|_{x,y,z} \right) dx dz \right].
 \end{aligned}$$

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

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

The left-hand side is a scalar while dx_i and dx_j are components of a vector. The term in parentheses on the right-hand side is a component of a second-rank tensor¹⁴, as shown in Exercise 1.4. In elasticity theory,

¹⁴Both 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.

the term in parentheses is the definition of the strain tensor for infinitesimal displacements; namely,

$$(1.4.6) \quad \varepsilon_{ij} := \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j \in \{1, 2, 3\},$$

¹⁵where $u_i = u_{x_i}$, $u_j = u_{x_j}$ and the partial derivatives are evaluated at $\mathbf{x} = [x_1, x_2, x_3]$.¹⁶ As indicated above, the strain tensor is a second-rank tensor.

Thus, if we suppose that a continuum is deformed in such a way that points are displaced by vector $\mathbf{u}(\mathbf{x})$, then, the strain tensor is defined by expression (1.4.6). Considering infinitesimal displacements, the components of this tensor allow us to describe the deformation associated with any such a displacement. Examining expression (1.4.6), we see that in the particular case discussed on page 19, where \mathbf{u} is given by constants, $\varepsilon_{ij} = 0$ for all i and j . In other words, according to the strain tensor, the continuum is not deformed, as expected in view of our discussion on page 19.

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; namely,

$$\nabla f(x_1, x_2, x_3) := \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right].$$

As shown in the derivation of expression (1.4.6), 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 $i, j \in \{1, 2, 3\}$.

1.4.3. 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.

¹⁵In this book, symbol $:=$ refers to a definition. In particular, “ $:=$ ” is read as “is defined by” and “ $=$ ” as “defines”.

¹⁶Readers 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.

Relative change in length. To illustrate a length change expressed by a strain tensor, we revisit the derivation shown in Section 1.4.2 and consider the one-dimensional case.

Let $\mathbf{x} = [x_1, 0, 0]$ and $\mathbf{x} + d\mathbf{x} = [x_1 + dx_1, 0, 0]$ 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

$$(1.4.7) \quad \check{x}_1 = x_1 + u_1|_{x_1,0,0},$$

and

$$(1.4.8) \quad \check{x}_1 + d\check{x}_1 = x_1 + dx_1 + u_1|_{x_1+dx_1,0,0}.$$

The distance between their components along the x_1 -axis after the deformation is given by the difference between expressions (1.4.7) and (1.4.8); namely,

$$(1.4.9) \quad d\check{x}_1 = dx_1 + u_1|_{x_1+dx_1,0,0} - u_1|_{x_1,0,0}.$$

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

$$u_1|_{x_1+dx_1,0,0} = u_1|_{x_1,0,0} + \left. \frac{\partial u_1}{\partial x_1} \right|_{x_1,0,0} dx_1 + \frac{1}{2} \left. \frac{\partial^2 u_1}{\partial x_1^2} \right|_{x_1,0,0} (dx_1)^2 + \dots$$

Using the approximation consisting of the first two terms, we can write expression (1.4.9) as

$$d\check{x}_1 \approx dx_1 + \left. \frac{\partial u_1}{\partial x_1} \right|_{x_1,0,0} dx_1,$$

which can be restated as

$$d\check{x}_1 \approx \left(1 + \left. \frac{\partial u_1}{\partial x_1} \right|_{x_1,0,0} \right) dx_1.$$

Hence, in view of definition (1.4.6), we can write the distance between the two points after deformation as

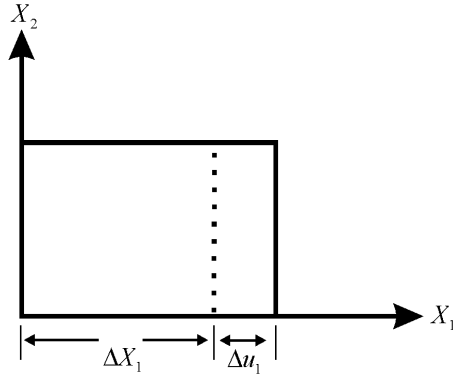
$$(1.4.10) \quad d\check{x}_1 \approx (1 + \varepsilon_{11}) dx_1,$$

where dx_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 1.4.1 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

$$(1.4.11) \quad \frac{\Delta X_1 + \Delta u_1}{\Delta X_1} = 1 + \frac{\Delta u_1}{\Delta X_1}.$$

FIGURE 1.4.1. Deformation: Uniaxial extension in the X_1 -axis direction

Considering infinitesimal gradients of the displacement, discussed in Section 1.3.3, and in view of Exercise 1.5, we can restate expression (1.4.11) as

$$(1.4.12) \quad 1 + \frac{\partial u_1}{\partial x_1}.$$

Expression (1.4.12) is a relative change in length due to deformation. Now, recall equation (1.4.10), which we can restate as

$$(1.4.13) \quad \frac{d\check{x}_1}{dx_1} \approx 1 + \varepsilon_{11},$$

to describe a relative change in length due to deformation. Hence, examining expressions (1.4.12) and (1.4.13), 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

$$(1.4.14) \quad V = \Delta x_1 \Delta x_2 \Delta x_3.$$

After the deformation, following expression (1.4.10), 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

$$(1.4.15) \quad \check{V} = (1 + \varepsilon_{11})(1 + \varepsilon_{22})(1 + \varepsilon_{33}) V.$$

Note that to state \check{V} as written in expression (1.4.15), we require that after the deformation the original rectangular box remains rectangular. In the context of our study, where the deformations are assumed to be small, the errors resulting from departing from this requirement are considered to be negligible. In other words, we use expression (1.4.15) even if Δx_1 , Δx_2 and Δx_3 are no longer parallel to the corresponding axes.

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

$$(1.4.16) \quad \check{V} \approx (1 + \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) V.$$

Thus, using expressions (1.4.14) and (1.4.16), we can state the relative change in volume as

$$(1.4.17) \quad \frac{\check{V} - V}{V} \approx \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} =: \varphi.$$

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 (1.4.6), expression (1.4.17) can be stated as divergence, since we can write

$$(1.4.18) \quad \varphi = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = \nabla \cdot \mathbf{u}.$$

To gain insight into the physical meaning of φ , we can revisit a special case discussed on page 19. If \mathbf{u} is given by constants, there is no deformation and, hence, no change in volume, as expected. Also, if $\mathbf{u} = [u_1(x_2, x_3), u_2(x_1, x_3), u_3(x_1, x_2)]$, then $\varphi = 0$. Such a displacement vector causes an infinitesimal deformation that results in change of shape, as discussed in the next section, but not in change of volume.

The dilatation will appear in stress-strain equations (5.12.4). It will also appear in the wave equation for P waves given in expression (6.1.12). Since dilatation is associated with a change in volume, P waves can be viewed as the propagation of compression within the continuum.

Note that, in terms of tensor algebra, expression (1.4.17) 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 proved 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 5.10.

Change in shape. The strain tensor also describes deformations leading to a change in shape. To gain geometrical insight, consider Figure 1.4.2 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

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

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 1.4.2, we can write

$$(1.4.19) \quad \beta_1 + \beta_2 \approx \frac{\Delta u_2}{\Delta X_1} + \frac{\Delta u_1}{\Delta X_2}.$$

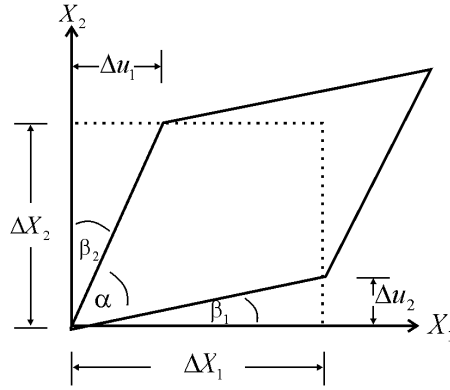


FIGURE 1.4.2. Deformation: Relative change in angles

Considering infinitesimal displacements, discussed in Section 1.3.3, and in view of Exercise 1.5, we can write equation (1.4.19) as

$$(1.4.20) \quad \beta_1 + \beta_2 \approx \frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2} = 2\varepsilon_{21} = 2\varepsilon_{12},$$

where we assume the equivalence between 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.¹⁷

Examining Figure 1.4.2, we see that equation (1.4.20) implies that the original segments are deviated by small angles β_1 and β_2 that can be stated 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.¹⁸

1.5. Rotation tensor and rotation vector

In Section 1.4.3, 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 (1.4.18). 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

$$(1.5.1) \quad \xi_{ij} := \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right), \quad i, j \in \{1, 2, 3\}.$$

In view of definition (1.5.1), $\xi_{11} = \xi_{22} = \xi_{33} = 0$, 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 1.4.3 and illustrated in Figure 1.4.2, the quantities $\partial u_i/\partial x_j$,

¹⁷Readers 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.

¹⁸Readers 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.

where $i \neq j$, are tantamount to the small deviation angles. Following the properties of the curl operator, we can associate tensor (1.5.1) with a vector given by

$$(1.5.2) \quad \Psi = \nabla \times \mathbf{u},$$

as shown in Exercise 1.7. We refer to Ψ as the rotation vector.¹⁹

Rotation vector (1.5.2) will be used in formulating the wave equation involving S waves, as shown in expression (6.1.16). In other words, S waves can be viewed as the propagation of 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 (1.4.6) and (1.5.1), we can write the partial derivative of a component of displacement as

$$(1.5.3) \quad \frac{\partial u_i}{\partial x_j} = \varepsilon_{ij} + \xi_{ij}, \quad i, j \in \{1, 2, 3\}.$$

Equation (1.5.3) 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.



¹⁹Readers 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.

1.6. Exercises

EXERCISE 1.1. ²⁰Given a material description of motion,

$$(1.6.1) \quad \mathbf{x}(\mathbf{X}, t) = \begin{cases} x_1 = X_1 e^t + X_3 (e^t - 1) \\ x_2 = X_2 + X_3 (e^t - e^{-t}) \\ x_3 = X_3 \end{cases},$$

verify that the transformation between the material, \mathbf{X} , and spatial, \mathbf{x} , coordinates exists, and find the spatial description of this motion.²¹

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

$$(1.6.2) \quad J := \det \begin{bmatrix} \frac{\partial x_1}{\partial X_1} & \frac{\partial x_1}{\partial X_2} & \frac{\partial x_1}{\partial X_3} \\ \frac{\partial x_2}{\partial X_1} & \frac{\partial x_2}{\partial X_2} & \frac{\partial x_2}{\partial X_3} \\ \frac{\partial x_3}{\partial X_1} & \frac{\partial x_3}{\partial X_2} & \frac{\partial x_3}{\partial X_3} \end{bmatrix},$$

does not vanish. Using equations (1.6.1), we obtain

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

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 $\mathbf{x} = \mathbf{A}\mathbf{X}$. 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,

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

REMARK 1.6.1. 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} .

²⁰See also Section 1.3.1.

²¹In this book, $e^{(\cdot)}$ and $\exp(\cdot)$ are used as synonymous notations.

EXERCISE 1.2. ²²Consider

$$(1.6.3) \quad F(X) = a \sin \frac{X}{b},$$

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 (1.6.3) 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

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

and

$$\lim_{(u/b) \rightarrow 0} \sin \frac{u(x)}{b} = 0.$$

Thus, we can write

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

as required.

REMARK 1.6.2. 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

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

²²See also Section 1.3.3.

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 1.6.3. 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 linear systems — a commonly used property in mathematical physics.²³

EXERCISE 1.4. ²⁴Using definition (1.4.6) 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 1.6.4. 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 values 1, 2 and 3.

SOLUTION 1.4. Following definition (1.4.6), consider $\partial \hat{u}_i / \partial \hat{x}_j$, where \hat{u}_i are the components of the displacement vector, \mathbf{u} , in the transformed coordinates \hat{x}_j . The transformation rule of the coordinate points is given by

$$(1.6.4) \quad \hat{x}_j = a_{jl} x_l, \quad j \in \{1, 2, 3\},$$

where the entries of matrix \mathbf{a} are the projections between the transformed and original axes. Matrix \mathbf{a} is an orthogonal matrix, which means that its inverse is equal to its transpose. Hence,

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

Consequently, we obtain

$$(1.6.5) \quad \frac{\partial x_j}{\partial \hat{x}_l} = a_{lj}.$$

Since \mathbf{u} is a vector, its components transform according to the rule

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

Thus, we can write

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

which can be restated as

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

²³Readers interested in the epistemological justification of linearity and the principle of superposition might refer to Steiner, M., (1998) The applicability of mathematics as a philosophical problem: Harvard University Press, pp. 30 – 32.

²⁴See also Sections 1.4.2, 5.1.1 and 5.2.3.

Hence, in view of equation (1.6.5), we can write

$$(1.6.6) \quad \frac{\partial \hat{u}_i}{\partial \hat{x}_l} = a_{ik} a_{lj} \frac{\partial u_k}{\partial x_j}, \quad i, l \in \{1, 2, 3\},$$

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. ²⁵Considering the one-dimensional case and assuming infinitesimal displacement gradients, in view of expressions (1.3.7) and (1.3.8), show that

$$(1.6.7) \quad \frac{\partial u}{\partial x} \approx \frac{\partial U}{\partial X}.$$

SOLUTION 1.5. Consider the one-dimensional case of expressions (1.3.7) and (1.3.8), namely

$$\begin{cases} U(X, t) = x(X, t) - X \\ u(x, t) = x - X(x, t) \end{cases}.$$

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

$$(1.6.8) \quad \begin{cases} \frac{\partial U}{\partial X} = \frac{\partial x}{\partial X} - 1 \\ \frac{\partial u}{\partial x} = 1 - \frac{\partial X}{\partial x} \end{cases}.$$

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

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

Hence, we can write expression (1.6.8) as

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

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

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

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

$$(1.6.9) \quad \frac{\partial u}{\partial x} = \frac{\frac{\partial U}{\partial X}}{\frac{\partial U}{\partial X} + 1}.$$

²⁵See also Section 1.4.3.

Examining equation (1.6.9), we notice that for the infinitesimal displacement gradients, namely, $\partial U / \partial X \ll 1$, we can write $\partial u / \partial x \approx \partial U / \partial X$, which is expression (1.6.7), as required.

EXERCISE 1.6. ²⁶Prove the following theorem.

THEOREM 1.6.5. *The sum of diagonal elements of a second-rank tensor is invariant under orthogonal 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 given by

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

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

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

Hence, by orthogonality of \mathbf{b} , we have

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

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 implies that the value of the sum of the diagonal elements is invariant under orthogonal transformations of the coordinate system. The sum of the diagonal elements of a second-rank tensor is a scalar. \square

REMARK 1.6.6. Following Exercise 1.4, we can see that dilatation, φ , defined by expression (1.4.18) 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.

²⁶See also Section 1.4.3.

EXERCISE 1.7. ²⁷In view of the properties of vector operators, show that the components of the second-rank tensor, given by expression (1.5.1), namely,

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

are associated with rotation vector (1.5.2).

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

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

where \mathbf{e}_i denotes a unit vector along the x_i -axis. Following expression (1.5.1), 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 $\Psi = [2\xi_{32}, 2\xi_{13}, 2\xi_{21}]$, we obtain definition (1.5.2).

REMARK 1.6.7. The association between the components of the second-rank tensor ξ_{ij} and the components of vector Ψ is due to the antisymmetry of this tensor that results in only three independent components.

²⁷See also Section 1.5.

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.

In general, the fundamental principles of continuum mechanics consist of the conservation of mass, balance of linear momentum, balance of angular momentum, balance of energy, balance of electric charge, balance of magnetic flux, and the principle of irreversibility: entropy. In this book, in which we work within the theory of linear elasticity, we need to invoke explicitly only a few of these principles.

We begin this chapter with the study of the conservation of mass, 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

2.1.1. 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.

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

$$(2.1.1) \quad m(t) = \iiint_V \rho(\mathbf{x}, t) \, dV,$$

where ρ denotes mass density. The rate of change of mass contained in this volume is given by the differentiation of equation (2.1.1) with respect to time; namely,

$$(2.1.2) \quad \frac{d}{dt} m(t) = \frac{d}{dt} \iiint_V \rho(\mathbf{x}, t) \, dV.$$

Furthermore, for an arbitrary volume that is fixed, V , we can rewrite expression (2.1.2) as

$$(2.1.3) \quad \frac{d}{dt} m(t) = \iiint_V \frac{\partial \rho(\mathbf{x}, t)}{\partial t} \, dV.$$

We can also express dm/dt in a different way. Since we assume that 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

$$(2.1.4) \quad \frac{d}{dt} m(t) = - \iint_S \rho(\mathbf{x}, t) \mathbf{v} \cdot \mathbf{n} \, dS,$$

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. The negative sign results from the fact that the vector normal to the surface points away from volume V . Herein, we assume the element dS to be sufficiently small that it might be considered as a plane and to have the same mass flow across all its points.

Expressions (2.1.3) and (2.1.4) describe the same quantity; namely,

$$\iiint_V \frac{\partial \rho(\mathbf{x}, t)}{\partial t} \, dV = - \iint_S \rho(\mathbf{x}, t) \mathbf{v} \cdot \mathbf{n} \, dS,$$

which means that the rate of change of the amount of material inside a closed surface is equal to the net rate with which this material flows through this surface. To combine expressions (2.1.3) and (2.1.4), we will express the right-hand side of equation (2.1.4) as a volume integral. Invoking the divergence theorem —

according to which the surface integral of vector field $\rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x})$ over a closed surface can be expressed as the volume integral of the divergence of this vector field integrated over the volume enclosed by this surface — we write

$$(2.1.5) \quad \iint_S \rho \mathbf{v} \cdot \mathbf{n} dS = \iiint_V \nabla \cdot (\rho \mathbf{v}) dV.$$

Equating the right-hand sides of equations (2.1.3) and (2.1.4), and using expression (2.1.5) in the latter one, we obtain

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

Combining the two volume integrals, we write

$$(2.1.6) \quad \iiint_V \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right] dV = 0,$$

which is a statement of the conservation of mass for a fixed volume V .

Note that in this derivation of equation (2.1.6), 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 (2.1.6) 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. Recalling our discussion in Section 1.3, we might note that for conservation of mass we used the spatial point of view, while for balances of linear and angular momenta we will use the material point of view.

2.1.3. 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 (2.1.6). For this equation to be true for an arbitrary fixed volume, the integrand must be identically zero. Thus, we require

$$(2.1.7) \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0.$$

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 (2.1.7) as

$$(2.1.8) \quad \frac{\partial \rho}{\partial t} + \nabla \cdot \left(\rho \frac{\partial \mathbf{u}}{\partial t} \right) = 0.$$

This is the equation of continuity.

Note that we could rewrite equation (2.1.8) as

$$\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right] \cdot [\rho, \rho u_1, \rho u_2, \rho u_3] = 0,$$

where the differential operator is a divergence in the four-dimensional space-time. This is a typical form of a fundamental conservation principle.¹

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 will study

$$(2.2.1) \quad \frac{d}{dt} \iiint_{V(t)} \rho(\mathbf{x}, t) \mathcal{A}(\mathbf{x}, t) dV,$$

where ρ is mass density, \mathcal{A} is a scalar, a vector or a tensor that describes a physical quantity of interest, while $V(t)$ is a volume that varies with time but always contains the same portion of the continuum. The temporal variation of \mathcal{A} in $V(t)$ can be also expressed in a different way; namely,

$$(2.2.2) \quad \iiint_{V(t)} \frac{\partial}{\partial t} (\rho \mathcal{A}) dV + \iint_{S(t)} \rho \mathcal{A} \sum_{j=1}^3 v_j n_j dS.$$

Herein, the volume integral describes the change of the amount of \mathcal{A} due to its creation or destruction within the volume, and the surface integral describes the net change of the amount of \mathcal{A} due to its flow across the surface; v_j are the components of the velocity vector with which \mathcal{A} flows across the surface whose normal is $\mathbf{n} = [n_1, n_2, n_3]$. Since expressions (2.2.1) and (2.2.2) describe the same entity, we can

¹Readers interested in fundamentals of conservation principles might refer to Bunge, M., (1967) Foundations of physics: Springer-Verlag, pp. 47 – 51.

equate them to write

$$(2.2.3) \quad \frac{d}{dt} \iiint_{V(t)} \rho(\mathbf{x}, t) \mathcal{A}(\mathbf{x}, t) dV = \iiint_{V(t)} \frac{\partial}{\partial t} (\rho \mathcal{A}) dV + \iint_{S(t)} \rho \mathcal{A} \sum_{j=1}^3 v_j n_j dS,$$

and proceed to express the right-hand side in terms of volume integrals.

Note that the above formulation is analogous to the one presented in Section 2.1. Expressions (2.2.1) and (2.2.2) are analogous to expressions (2.1.2) and (2.1.4), respectively. The key difference between expressions (2.1.4) and (2.2.2) is the appearance of $\iiint_{V(t)} \partial(\rho \mathcal{A}) / \partial t dV$ in the latter one, which is due to the fact that, although mass cannot be created or destroyed in classical physics, no such restriction governs quantity \mathcal{A} .

To express the right-hand side in terms of volume integrals, we invoke the divergence theorem to write equation (2.2.3) as

$$\frac{d}{dt} \iiint_{V(t)} \rho \mathcal{A} dV = \iiint_{V(t)} \frac{\partial}{\partial t} (\rho \mathcal{A}) dV + \iiint_{V(t)} \sum_{j=1}^3 \frac{\partial (\rho \mathcal{A} v_j)}{\partial x_j} dV.$$

Differentiating and rearranging, we get

$$\frac{d}{dt} \iiint_{V(t)} \rho \mathcal{A} dV = \iiint_{V(t)} \left\{ \rho \left(\frac{\partial \mathcal{A}}{\partial t} + \mathbf{v} \cdot \nabla \mathcal{A} \right) + \mathcal{A} \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right] \right\} dV.$$

We note that the term in brackets vanishes due to equation of continuity (2.1.7). Thus, we write

$$\frac{d}{dt} \iiint_{V(t)} \rho \mathcal{A} dV = \iiint_{V(t)} \rho \left(\frac{\partial \mathcal{A}}{\partial t} + \mathbf{v} \cdot \nabla \mathcal{A} \right) dV,$$

which can be restated as

$$(2.2.4) \quad \frac{d}{dt} \iiint_{V(t)} \rho \mathcal{A} dV = \iiint_{V(t)} \rho \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \mathcal{A} dV.$$

In view of expression (1.3.5), we note that the operator in parentheses in expression (2.2.4) is the material time-derivative operator acting on $\mathcal{A}(\mathbf{x}, t)$, namely, $d\mathfrak{A}(\mathbf{X}, t)/dt$, where \mathfrak{A} is the material description of \mathcal{A} . In accordance with the linearization discussed in Section 1.3.3, we let $\mathfrak{A}(\mathbf{X}, t)$ be approximated by $\mathcal{A}(\mathbf{x}, t)$. Hence, we obtain

$$(2.2.5) \quad \frac{d}{dt} \iiint_{V(t)} \rho(\mathbf{x}, t) \mathcal{A}(\mathbf{x}, t) dV = \iiint_{V(t)} \rho(\mathbf{x}, t) \frac{D\mathcal{A}(\mathbf{x}, t)}{Dt} dV,$$

which is the desired result. Symbol D/Dt on the right-hand side denotes the material time-derivative operator. Also, we can view the time derivative of a moving-volume integral as a case of Leibniz's integration rule, as illustrated in Exercise 2.3.

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 concept of stress 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.²

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 enclosed by an imaginary surface interacts with the portion of the continuum outside of this surface, as discussed on page 11. 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 given by

$$(2.3.1) \quad \bar{\mathbf{T}} = \frac{\Delta \mathbf{F}}{\Delta S}.$$

Cauchy's stress principle — the fundamental principle of continuum mechanics stated on page 11 — implies that as $\Delta S \rightarrow 0$, ratio (2.3.1) tends to a finite limit.³ The resulting traction vector is given by

$$(2.3.2) \quad \mathbf{T}^{(\mathbf{n})} = \lim_{\Delta S \rightarrow 0} \frac{\Delta \mathbf{F}}{\Delta S} = \frac{d\mathbf{F}}{dS},$$

where the superscript, \mathbf{n} , specifies the surface element, ΔS , upon which the traction is acting by stating unit vector \mathbf{n} that is normal to this surface element. 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.

Since the value of traction is finite even if the element of the surface area becomes infinitesimal, we can describe forces at any given point within the continuum. Also, since the traction is explicitly dependent on the orientation of the surface element, as indicated by vector \mathbf{n} , we can describe forces in any given direction within the continuum. Consequently, we can study both inhomogeneity and anisotropy of the continuum.

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

³Interested readers might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, p. 70.

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 assumed to behave as an action-at-a-distance, which is the instantaneous action between two bodies in spatial separation. In general, this is not the case, as illustrated by the electromagnetism or general relativity.

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 (2.3.2); namely,

$$\mathbf{T} = \frac{d\mathbf{F}}{dS},$$

and the body forces are given by

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

Consequently, the total force is

$$(2.4.1) \quad \mathbf{F}_T = \iint_S \mathbf{T} dS + \iiint_V \mathbf{f} dV,$$

where S is the surface enclosing volume V .

Note that \mathbf{T} and \mathbf{f} have units of force per area and force per volume, respectively. Hence, both integrals have units of force, as required.

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

$$(2.4.2) \quad \frac{d}{dt} \iiint_{V(t)} \rho \frac{d\mathbf{u}}{dt} dV = \iint_{S(t)} \mathbf{T} dS + \iiint_{V(t)} \mathbf{f} dV,$$

where the displacement,

$$(2.4.3) \quad \mathbf{u} = [u_1(\mathbf{x}, t), u_2(\mathbf{x}, t), u_3(\mathbf{x}, t)],$$

is a function of both space and time. This integral equation states the balance of linear momentum: The rate of change of the linear momentum of an element within the continuum is equal to the sum of the external forces acting upon this element. This statement is analogous to Newton's second law of motion in particle mechanics.

To see this analogy, let us consider a system of n particles whose masses are m_i . For such a system, we can write Newton's second law of motion as

$$(2.4.4) \quad \frac{d}{dt} \sum_{i=1}^n m_i \frac{d\mathbf{x}_i}{dt} = \sum_{i=1}^n \mathbf{F}_i + \sum_{i=1}^n \sum_{j=1}^n \mathbf{F}_{ji}.$$

The sum on the left-hand side is the linear momentum of the system of particles. The first sum on the right-hand side denotes the external forces acting on the particles of the system. The second sum on the right-hand side denotes the internal forces among the particles of the system; \mathbf{F}_{ji} is the force on the i th particle due to the j th particle. Since i and j are summation indices, we can rewrite the second sum on the right-hand side as

$$\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\mathbf{F}_{ji} + \mathbf{F}_{ij}).$$

Now, let us invoke Newton's third law of motion: Forces between particles are equal in magnitude and opposite in direction, $\mathbf{F}_{ji} = -\mathbf{F}_{ij}$. Also, in classical mechanics, it is assumed — though seldom explicitly stated — that a point particle exerts no force on itself, only on other particles; thus, $\mathbf{F}_{ii} = \mathbf{0}$. Consequently, the double sum vanishes and, hence, we can write Newton's second law of motion for a system of particles as

$$\frac{d}{dt} \sum_{i=1}^n m_i \frac{d\mathbf{x}_i}{dt} = \sum_{i=1}^n \mathbf{F}_i.$$

In other words, only the external forces affect the change of linear momentum. This means that the system itself cannot change its own linear momentum. In equation (2.4.2), the external forces are \mathbf{T} and \mathbf{f} ; these forces are external to the portion of a continuum contained in volume V .

Also, in applying Newton's third law of motion, we assume that the two particles remain in constant contact.⁴ This assumption allows us to avoid the physically undesirable concept of action-at-a-distance.

Let us return to equation (2.4.2). Invoking expression (2.2.5) and letting $\mathcal{A} = d\mathbf{u}/dt$, we can rewrite equation (2.4.2) as

$$(2.4.5) \quad \iiint_{V(t)} \rho \frac{d^2 \mathbf{u}}{dt^2} dV = \iint_{S(t)} \mathbf{T} dS + \iiint_{V(t)} \mathbf{f} dV.$$

Herein, d^2/dt^2 is the material time-derivative operator, which is shown in expression (1.3.5). One could rewrite equations (2.4.2) and (2.4.5) using D/Dt rather than d/dt inside the volume integrals. For the purpose of this book, we choose to use the more familiar notation, and to explicitly state the material time-derivative operator every time it is invoked.

In Section 2.5.2, we will use the balance of linear momentum to formulate the stress tensor. In Section 2.6, following the formulation of the stress tensor, we will use equation (2.4.5) to derive Cauchy's equations of motion.

⁴Readers interested in an insightful explanation for the requirement of the constant contact might refer to Schutz, B., (2003) Gravity from the ground up: Cambridge University Press, p. 12.

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 i th 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. At a given point, the three tractions associated with the three mutually orthogonal planes can be explicitly written as three vectors given by

$$(2.5.1) \quad \begin{bmatrix} \mathbf{T}^{(\mathbf{e}_1)} \\ \mathbf{T}^{(\mathbf{e}_2)} \\ \mathbf{T}^{(\mathbf{e}_3)} \end{bmatrix} = \begin{bmatrix} T_1^{(\mathbf{e}_1)} & T_2^{(\mathbf{e}_1)} & T_3^{(\mathbf{e}_1)} \\ T_1^{(\mathbf{e}_2)} & T_2^{(\mathbf{e}_2)} & T_3^{(\mathbf{e}_2)} \\ T_1^{(\mathbf{e}_3)} & T_2^{(\mathbf{e}_3)} & T_3^{(\mathbf{e}_3)} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix}.$$

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 (2.5.1) as

$$(2.5.2) \quad \sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}.$$

By examining expressions (2.5.1) and (2.5.2), we immediately see that σ_{ij} represents the j th 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 2.5.1.⁵ We also wish to distinguish between tension and compression for the traction components normal to a given face, as well as denote the direction 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 2.5.1, we see that all the traction components on each of the six faces illustrated therein

⁵This 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 (2nd 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.

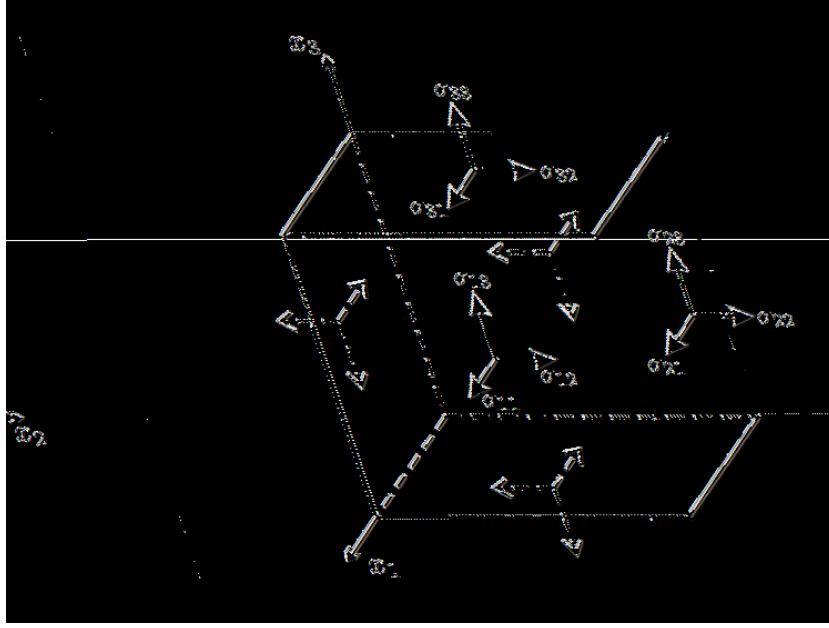


FIGURE 2.5.1. Index convention for σ_{ij} components: Symbol σ_{ij} represents the j th component of the surface force acting on the surface whose normal is parallel to the x_i -axis. All components shown herein are positive.

are positive. In the context of the normal components, our sign convention implies that tension is positive while compression is negative.⁶

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

$$(2.5.3) \quad \mathbf{T}^{(\mathbf{n})} = -\mathbf{T}^{(-\mathbf{n})}$$

is always true; herein, $\mathbf{T}^{(\mathbf{n})}$ is due to the force acting on the plane whose outward normal is \mathbf{n} , and $-\mathbf{T}^{(-\mathbf{n})}$ is due to the force acting on the other side of this plane.

As formulated herein, the entries of matrix (2.5.2) 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; such a construction is also called Cauchy's 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 2.5.2. Thus,

⁶This sign convention is consistent with Aki, K. and Richards, P.G., (2002) Quantitative seismology (2nd edition): University Science Books, p. 15.

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 (2.4.5), for a finite-size tetrahedron, we can write

$$(2.5.4) \quad \Delta \mathbf{F} + \Delta \mathbf{F}^{(\mathbf{e}_1)} + \Delta \mathbf{F}^{(\mathbf{e}_2)} + \Delta \mathbf{F}^{(\mathbf{e}_3)} + \bar{\mathbf{f}} \Delta V = \bar{\rho} \Delta V \frac{d\bar{\mathbf{v}}}{dt},$$

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 (2.5.4) 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 (2.3.1), we can write

$$(2.5.5) \quad \Delta \mathbf{F} = \bar{\mathbf{T}}^{(\mathbf{n})} \Delta S.$$

Using expressions (2.5.3) and (2.5.5), we can rewrite equation (2.5.4) as

$$(2.5.6) \quad \bar{\mathbf{T}}^{(\mathbf{n})} \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_1)} \Delta S_1 - \bar{\mathbf{T}}^{(\mathbf{e}_2)} \Delta S_2 - \bar{\mathbf{T}}^{(\mathbf{e}_3)} \Delta S_3 + \bar{\mathbf{f}} \Delta V = \bar{\rho} \Delta V \frac{d\bar{\mathbf{v}}}{dt},$$

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 (2.5.6), $\bar{\mathbf{T}}^{(\cdot)}$ is a resultant traction that corresponds to a given face. 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 (2.5.6), are illustrated in Figure 2.5.2.

To study equation (2.5.6), 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

$$(2.5.7) \quad \Delta S_i = n_i \Delta S, \quad i \in \{1, 2, 3\},$$

where n_i are the components of the unit vector, \mathbf{n} , which is normal to the oblique face. Using expression (2.5.7), we can rewrite equation (2.5.6) as

$$(2.5.8) \quad \bar{\mathbf{T}}^{(\mathbf{n})} \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_1)} n_1 \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_2)} n_2 \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_3)} n_3 \Delta S + \bar{\mathbf{f}} \Delta V = \bar{\rho} \Delta V \frac{d\bar{\mathbf{v}}}{dt}.$$

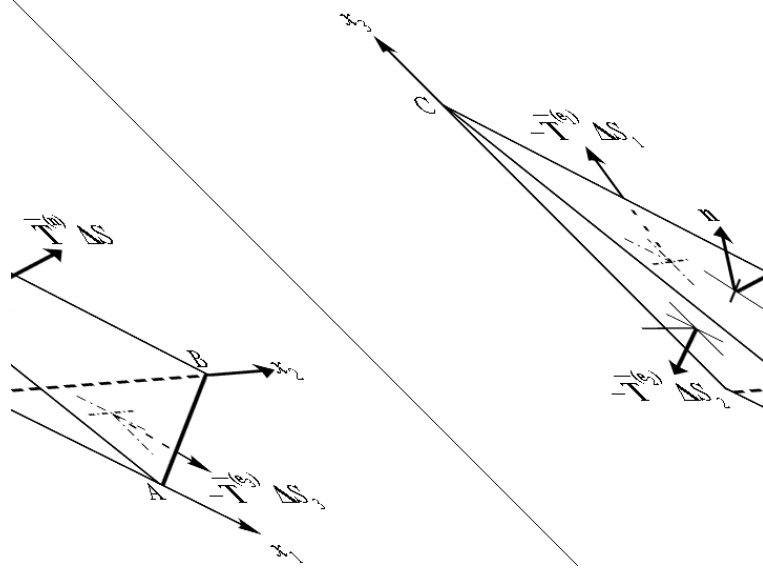


FIGURE 2.5.2. Cauchy's tetrahedron: Tetrahedron used in the formulation of the stress tensor

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

$$(2.5.9) \quad \Delta V = \frac{h}{3} \Delta S,$$

where h is the height of the tetrahedron. Hence, using expression (2.5.9), we can rewrite equation (2.5.8) as

$$(2.5.10) \quad \bar{\mathbf{T}}^{(\mathbf{n})} \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_1)} n_1 \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_2)} n_2 \Delta S - \bar{\mathbf{T}}^{(\mathbf{e}_3)} n_3 \Delta S + \frac{\bar{\mathbf{f}} h}{3} \Delta S = \frac{\bar{\rho} h}{3} \Delta S \frac{d\bar{\mathbf{v}}}{dt}.$$

Dividing both sides of equation (2.5.10) by ΔS , we obtain

$$(2.5.11) \quad \bar{\mathbf{T}}^{(\mathbf{n})} - \bar{\mathbf{T}}^{(\mathbf{e}_1)} n_1 - \bar{\mathbf{T}}^{(\mathbf{e}_2)} n_2 - \bar{\mathbf{T}}^{(\mathbf{e}_3)} n_3 + \frac{\bar{\mathbf{f}} h}{3} = \frac{\bar{\rho} h}{3} \frac{d\bar{\mathbf{v}}}{dt}.$$

To describe the state of stress at a point within the continuum, we let $h \rightarrow 0$ 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

$$(2.5.12) \quad \mathbf{T}^{(\mathbf{n})} = \mathbf{T}^{(\mathbf{e}_1)} n_1 + \mathbf{T}^{(\mathbf{e}_2)} n_2 + \mathbf{T}^{(\mathbf{e}_3)} n_3.$$

Note that in equation (2.5.12), 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 (2.5.12) is valid for any coordinate system.

Equation (2.5.12) 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.⁷

Expressing the orthogonal-face tractions in terms of their components, equation (2.5.12) can be explicitly written as

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

Equation (2.5.13) 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 (2.5.1) and (2.5.13), we conclude that

$$(2.5.14) \quad \mathbf{T}^{(\mathbf{n})} = \sigma^T \mathbf{n},$$

where σ is given in expression (2.5.2) and T denotes transpose. 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 2.4.

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 (2.5.2), we can rewrite equation (2.5.14) 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

$$(2.5.15) \quad T_i^{(\mathbf{n})} = \sum_{j=1}^3 \sigma_{ji} n_j, \quad i \in \{1, 2, 3\}.$$

Expression (2.5.15) 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

⁷Readers 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.

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. The main-diagonal entries describe the components acting along the surface-normal vector. Hence, in view of the sign convention stated on page 41, if σ_{ii} is positive, it is a tensile component; if it is negative, it is a compressional component. The nondiagonal entries describe shear components.⁸

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 (2.5.15), we can write the balance of linear momentum, stated in equation (2.4.5), in terms of components, as

$$\iiint_{V(t)} \rho \frac{d^2 u_i}{dt^2} dV = \iint_{S(t)} \sum_{j=1}^3 \sigma_{ji} n_j dS + \iiint_{V(t)} f_i dV, \quad i \in \{1, 2, 3\}.$$

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

$$\iiint_{V(t)} \rho \frac{d^2 u_i}{dt^2} dV = \iiint_{V(t)} \sum_{j=1}^3 \frac{\partial \sigma_{ji}}{\partial x_j} dV + \iiint_{V(t)} f_i dV, \quad i \in \{1, 2, 3\}.$$

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

$$(2.6.1) \quad \iiint_{V(t)} \left(\sum_{j=1}^3 \frac{\partial \sigma_{ji}}{\partial x_j} + f_i - \rho \frac{d^2 u_i}{dt^2} \right) dV = 0, \quad i \in \{1, 2, 3\},$$

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 (2.6.1). 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

$$(2.6.2) \quad \sum_{j=1}^3 \frac{\partial \sigma_{ji}}{\partial x_j} + f_i = \rho \frac{d^2 u_i}{dt^2}, \quad i \in \{1, 2, 3\}.$$

⁸Readers 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.

In view of equation (2.4.5), d^2/dt^2 refers to the material time-derivative operator. However, in this book, as discussed in Section 1.3.3, we use the linearized formulation and we can rewrite equations (2.6.2) as⁹

$$(2.6.3) \quad \sum_{j=1}^3 \frac{\partial \sigma_{ji}}{\partial x_j} + f_i = \rho \frac{\partial^2 u_i}{\partial t^2}, \quad i \in \{1, 2, 3\}.$$

These are Cauchy's equations of motion. As shown in Exercise 2.6, the *SI* units of Cauchy's equations of motion are N/m^3 .¹⁰ These equations are also known as Cauchy's first law of motion.¹¹

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 (2.6.3), we obtain the equations of static equilibrium,

$$(2.6.4) \quad \sum_{j=1}^3 \frac{\partial \sigma_{ji}}{\partial x_j} + f_i = 0, \quad i \in \{1, 2, 3\}.$$

These equations describe the equilibrium state of an element of the continuum arising from the application of forces whose resultant is zero. Equations (2.6.4) are used to illustrate the symmetry of the stress tensor, as shown in Exercise 2.7. Equations (2.6.4) are also valid for rectilinear, constant-velocity motion.

Consider a system composed of equation of continuity (2.1.8) and Cauchy's equations of motion (2.6.3) in a three-dimensional continuum. This system contains four equations and sixteen unknowns, namely, mass density, ρ , stress-tensor components, $\sigma_{11}, \sigma_{12}, \sigma_{13}, \sigma_{21}, \sigma_{22}, \sigma_{23}, \sigma_{31}, \sigma_{32}, \sigma_{33}$, 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 show that the stress tensor is symmetric, which results in only six independent stress-tensor components. Also, we will ignore the body force, $\mathbf{f} = [f_1, f_2, f_3]$, for the following two reasons.

⁹Readers 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.

¹⁰Readers 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.

¹¹Readers interested in an insightful interpretation of Cauchy's first law of motion might refer to Bunge, M., (1967) Foundations of physics: Springer-Verlag, pp. 152 – 154.

The body force is irrelevant when we consider an infinitesimal element of the continuum, as we do in *Part 1*. We can see this in view of the tetrahedron argument, discussed in Section 2.5.2; in particular, by examining the step between equations (2.5.11) and (2.5.12).

The effects of the body forces are negligible as compared to the effects of the surface forces if we consider sufficiently high frequencies, which is the case of applied seismology discussed in *Part 2* and *Part 3*. In other words, the effects of gravitation are negligible as compared to the effects of elasticity.¹²

2.6.2. Example: Surface-forces formulation. To gain insight into the equations of motion without body forces, we rederive equations (2.6.3) 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 (2.3.2), we can write the force acting along the x_1 -axis as

$$(2.6.5) \quad T_1^{(\mathbf{e}_1)} dx_2 dx_3 + T_1^{(\mathbf{e}_2)} dx_1 dx_3 + T_1^{(\mathbf{e}_3)} dx_1 dx_2,$$

where \mathbf{e}_i denotes the unit normal to the coordinate plane on which $T_1^{(\mathbf{e}_i)}$ is acting, and $dx_j dx_k$ is the surface area of this planar element. Following expression (2.5.2), expression (2.6.5) can be rewritten as

$$(2.6.6) \quad \sigma_{11} dx_2 dx_3 + \sigma_{21} dx_1 dx_3 + \sigma_{31} dx_1 dx_2.$$

Now, consider a small rectangular box subjected to stresses. Let the rectangular box be spanned by dx_1 , dx_2 and dx_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 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 (2.6.6), we can write the resultant force

¹²Readers interested in the effect of gravitation on seismic wave propagation might refer to Udías, A., (1999) Principles of seismology: Cambridge University Press, pp. 39 – 40.

Let us note here that term 'gravity' is often used carelessly. One should use 'gravitation' to refer to the universal attraction between masses, and 'gravity' to refer to the downward acceleration experienced by a mass at rest in a reference frame rotating with the Earth; gravity is the vector resultant of gravitation and centrifugal force.

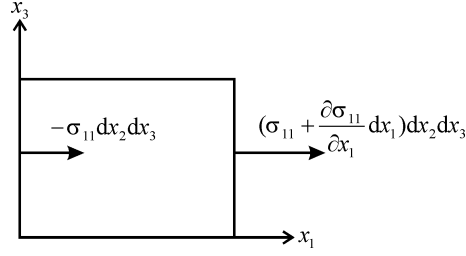


FIGURE 2.6.1. Surface forces: Two forces acting along the x_1 -axis on faces that are normal to it

along the x_1 -axis as

$$(2.6.7) \quad dF_1 = \left[\left(\sigma_{11} + \frac{\partial \sigma_{11}}{\partial x_1} dx_1 \right) dx_2 dx_3 + (-\sigma_{11} dx_2 dx_3) \right] \\ + \left[\left(\sigma_{21} + \frac{\partial \sigma_{21}}{\partial x_2} dx_2 \right) dx_1 dx_3 + (-\sigma_{21} dx_1 dx_3) \right] \\ + \left[\left(\sigma_{31} + \frac{\partial \sigma_{31}}{\partial x_3} dx_3 \right) dx_1 dx_2 + (-\sigma_{31} dx_1 dx_2) \right],$$

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 2.6.1, 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 (2.6.7) is a first-order approximation. This approximation is consistent with our study in the context of linearized theory.

Expression (2.6.7) immediately simplifies to

$$(2.6.8) \quad dF_1 = \left(\frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{21}}{\partial x_2} + \frac{\partial \sigma_{31}}{\partial x_3} \right) dx_1 dx_2 dx_3.$$

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

$$dF_1 = \rho dx_1 dx_2 dx_3 \frac{d^2 u_1}{dt^2},$$

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 (2.6.8) as

$$(2.6.9) \quad \rho \frac{d^2 u_1}{dt^2} = \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{21}}{\partial x_2} + \frac{\partial \sigma_{31}}{\partial x_3}.$$

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

$$(2.6.10) \quad \rho \frac{d^2 u_2}{dt^2} = \frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{32}}{\partial x_3}$$

and

$$(2.6.11) \quad \rho \frac{d^2 u_3}{dt^2} = \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{23}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3},$$

respectively.

In view of linearization discussed in Section 1.3.3, total derivatives with respect to time are equivalent to partial derivatives. Consequently, expressions (2.6.9), (2.6.10) and (2.6.11) are equivalent to Cauchy's equations of motion (2.6.3) with no body forces.

2.7. Balance of angular momentum

2.7.1. Introductory comments. Above we postulated that the continuum obeys the conservation of mass and the balance of linear momentum. In this section, we will formulate the balance of angular momentum in the context of the conservation of mass, the balance of linear momentum and Newton's third law of motion. To obtain the balance of angular momentum, we use the fact that the time rate of change of angular momentum for a given system is equal to the vector sum of the torques due to the external forces acting on that system. This property allows us to study the behaviour of a system without investigating the details of its internal behaviour. In the context of particle mechanics, we can formulate such a system by assuming that each two particles act on one another with forces of equal magnitude that are collinear and have opposite directions; such forces are called the central forces. In view of this assumption, all the internal forces within the system cancel out, which means that the system cannot change its own angular momentum.

To see this formulation in the context of particle mechanics, let us return to the Newton's second law of motion for a system of particles as stated by equation (2.4.4). Multiplying both sides by vector \mathbf{r}_i , which is the vector between a reference point and the i th particle, we get

$$\frac{d}{dt} \sum_{i=1}^n \mathbf{r}_i \times m_i \frac{d\mathbf{x}_i}{dt} = \sum_{i=1}^n \mathbf{r}_i \times \mathbf{F}_i + \sum_{i=1}^n \sum_{j=1}^n \mathbf{r}_i \times \mathbf{F}_{ji}.$$

The sum on the left-hand side is the angular momentum of the system of n particles. The first sum on the right-hand side is the torque on the system due to external forces. The second sum on the right-hand side is the torque due to the forces within the system. In a manner analogous to the one used in Section 2.4, we wish to eliminate the second sum. Since i and j are summation indices, we can rewrite the double sum as

$$\sum_{i=1}^n \sum_{j=1}^n \mathbf{r}_i \times \mathbf{F}_{ji} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\mathbf{r}_i \times \mathbf{F}_{ji} + \mathbf{r}_j \times \mathbf{F}_{ij}).$$

Invoking Newton's third law of motion, $\mathbf{F}_{ji} = -\mathbf{F}_{ij}$, together with $\mathbf{F}_{ii} = \mathbf{0}$, we can rewrite this sum as

$$\sum_{i=1}^n \sum_{j=1}^n \mathbf{r}_i \times \mathbf{F}_{ji} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ji}].$$

Since \mathbf{r}_i is the vector between the reference point and the i th particle while \mathbf{r}_j is the vector between the reference point and the j th particle, $\mathbf{r}_i - \mathbf{r}_j$ is the vector between the two particles. For this sum to vanish, we require that $\mathbf{r}_i - \mathbf{r}_j$ be parallel to \mathbf{F}_{ji} . This means that we require the forces between the particles to be central forces.¹³ This formulation relies on the strong form of Newton's third law of motion: Forces between particles are equal in magnitude and opposite in direction as well as act along the line joining the particles. Notably, the particle-mechanics analogue that we used in Section 2.4 for the balance of linear momentum required only the weak form of Newton's third law of motion; the forces are not required to act along the line joining the particles. Herein, in view of the strong form of Newton's third law of motion, we write the balance of angular momentum for the system of particles as

$$(2.7.1) \quad \frac{d}{dt} \sum_{i=1}^n \mathbf{r}_i \times m_i \frac{d\mathbf{x}_i}{dt} = \sum_{i=1}^n \mathbf{r}_i \times \mathbf{F}_i.$$

We will also use the strong form of Newton's third law of motion for a continuum. This means that for our formulation we will consider only central forces; we will not consider couple stresses within the continuum.¹⁴

2.7.2. Integral equation. Considering only the central forces acting within the continuum, we can state the balance of angular momentum as

$$(2.7.2) \quad \frac{d}{dt} \iiint_{V(t)} \left(\mathbf{x} \times \rho \frac{d\mathbf{u}}{dt} \right) dV = \iint_{S(t)} (\mathbf{x} \times \mathbf{T}) dS + \iiint_{V(t)} (\mathbf{x} \times \mathbf{f}) dV,$$

where $V(t)$ is a volume that moves while always containing the same portion of the continuum and $S(t)$ is the surface enclosing this volume.¹⁵ In other words, the rate of change of the angular momentum of an element within the continuum is equal to the sum of torques acting upon this element. The integrand on the left-hand side is the angular momentum, namely, the vector product of the distance, \mathbf{x} , between a reference point and the element of the continuum with the linear-momentum density, $\rho d\mathbf{u}/dt$. The first

¹³Readers interested in more details of such formulations in the context of particle mechanics might refer to Doran, C., and Lasenby, A., (2003) *Geometric algebra for physicists*, Cambridge University Press, pp. 57 – 61, Feynman, R.P., Leighton, R.B., and Sands, M., (1963/1989) *Feynman's lectures on physics*: Addison-Wesley Publishing Co. Vol. I, p. 18-7, to Goldstein, H., (1950/1980) *Classical mechanics*: Addison-Wesley Publishing Co. pp. 5 – 9, and to Radin, S.H., and Folk, R.T., (1982) *Physics for scientists and engineers*: Prentice-Hall, Inc., pp. 156 – 159, 164.

Readers interested in a formulation that invokes Lagrange's equations of motion, which we will discuss in Section 13.2, might refer to Landau, L.D., and Lifshitz, E.M., (1969) *Course of theoretical physics*, Vol. 1: *Mechanics*: Elsevier, pp. 18 – 19. Therein, the balance of angular momentum is derived as a fundamental consequence of the inertial frame of reference.

¹⁴Readers interested in a description of couple stresses might refer to Malvern, L.E., (1969) *Introduction to the mechanics of a continuous medium*: Prentice-Hall, pp. 215 – 220.

¹⁵Readers interested in the balance of angular momentum that includes couple stresses might refer to Malvern, L.E., (1969) *Introduction to the mechanics of a continuous medium*: Prentice-Hall, pp. 218 – 219.

integrand on the right-hand side is the vector product of this distance and force per unit area, \mathbf{T} , 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, \mathbf{f} , associated with this element.

This integral equation states the balance of angular momentum: The rate of change of the angular momentum of an element within the continuum is equal to the sum of torques due to the external forces acting upon this element. $\mathbf{x} \times \mathbf{T}$ and $\mathbf{x} \times \mathbf{f}$ in equation (2.7.2) are torque densities per unit area and unit volume, respectively; the forces, \mathbf{T} and \mathbf{f} , causing the torques are external to the portion of the continuum contained in volume V . The sum of the two integrals on the right-hand side of equation (2.7.2) is analogous to the sum on the right-hand side in equation (2.7.1).

The balance of angular momentum stated in equation (2.7.2) is valid only for the strong form of Newton's third law of motion.

Note that we have arrived at the balance of angular momentum as a consequence of the conservation of mass, balance of linear momentum and the strong form of Newton's third law of motion. It is worth noting that one could adopt the balance of angular momentum as an independent postulate.

2.7.3. Symmetry of stress tensor. Considering the conservation of mass and the balance of linear momentum — together with the balance of angular momentum under the assumption of only central forces acting within the continuum — we obtain an important consequence of these laws: The stress tensor is symmetric. In deriving the equations resulting from the balance of angular momentum, we will implicitly use the conservation of mass — by invoking time derivative of volume integral — and explicitly use the balance of linear momentum.

THEOREM 2.7.1. *Consider a linearized formulation of displacement and its associated quantities in a three-dimensional continuum. Let the principles of the conservation of mass and the balance of linear momentum hold. Then, the balance of angular momentum stated in expression (2.7.2) holds if*

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

where $i, j \in \{1, 2, 3\}$.

NOTATION 2.7.2. 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 values 1, 2 and 3.

The time-derivative operators, d/dt , inside the volume integrals are the material time-derivative operators.

PROOF. Since ρ is a scalar, we can rewrite expression (2.7.2) as

$$(2.7.3) \quad \frac{d}{dt} \iiint_{V(t)} \left[\rho \left(\mathbf{x} \times \frac{d\mathbf{u}}{dt} \right) \right] dV = \iint_{S(t)} (\mathbf{x} \times \mathbf{T}) dS + \iiint_{V(t)} (\mathbf{x} \times \mathbf{f}) dV.$$

Invoking the time derivative of a moving-volume integral, which is given by expression (2.2.5) and where we let $\mathcal{A} = \mathbf{x} \times d\mathbf{u}/dt$, we can restate expression (2.7.3) as

$$(2.7.4) \quad \iiint_{V(t)} \rho \left(\mathbf{x} \times \frac{d^2 \mathbf{u}}{dt^2} \right) dV = \iint_{S(t)} (\mathbf{x} \times \mathbf{T}) dS + \iiint_{V(t)} (\mathbf{x} \times \mathbf{f}) dV.$$

In view of the linearized formulation discussed in Section 1.3.3, we can rewrite expression (2.7.4) as

$$(2.7.5) \quad \iiint_{V(t)} \rho \left(\mathbf{x} \times \frac{\partial^2 \mathbf{u}}{\partial t^2} \right) dV = \iint_{S(t)} (\mathbf{x} \times \mathbf{T}) dS + \iiint_{V(t)} (\mathbf{x} \times \mathbf{f}) dV.$$

Using the stress tensor given in expression (2.5.14), invoking the divergence theorem, and requiring expression (2.7.5) to be valid for an arbitrary volume that consistently contains the same portion of the continuum, we obtain the differential equation given by

$$(2.7.6) \quad \rho \left(\mathbf{x} \times \frac{\partial^2 \mathbf{u}}{\partial t^2} \right) = \nabla \cdot (\mathbf{x} \times \sigma^T) + \mathbf{x} \times \mathbf{f},$$

where T denotes transpose. Consider the first term on the right-hand side in equation (2.7.6). The expression in parentheses is a second-rank tensor whose il th component can be written as

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

where ϵ_{ijk} is the permutation symbol and where we reversed the subscripts of the stress tensor to consider its transpose. Taking the i th component of the divergence and using the product rule, we obtain

$$(2.7.7) \quad \begin{aligned} [\nabla \cdot (\mathbf{x} \times \sigma^T)]_i &= \frac{\partial}{\partial x_l} (\epsilon_{ijk} x_j \sigma_{lk}) \\ &= (\mathbf{x} \times \nabla \cdot \sigma^T)_i + \epsilon_{ijk} \delta_{jl} \sigma_{lk}, \quad i \in \{1, 2, 3\}. \end{aligned}$$

Substituting expression (2.7.7) into equation (2.7.6), we obtain

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

Using the linearity of the cross-product operator, we can rearrange the above equation and write

$$(2.7.8) \quad \begin{aligned} \left[\mathbf{x} \times \left(\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right) \right]_i - (\mathbf{x} \times \nabla \cdot \sigma^T)_i - (\mathbf{x} \times \mathbf{f})_i &= \left[\mathbf{x} \times \left(\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} - \nabla \cdot \sigma^T - \mathbf{f} \right) \right]_i \\ &= \epsilon_{ijk} \delta_{jl} \sigma_{lk}, \quad i \in \{1, 2, 3\}. \end{aligned}$$

Now, we invoke Cauchy's equations of motion (2.6.3), which can be written as

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

hence, the term in brackets in equation (2.7.8) vanishes. This implies

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

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

$$\epsilon_{ijk}\sigma_{jk} = 0, \quad i \in \{1, 2, 3\},$$

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

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

as required. □

REMARK 2.7.3. $\epsilon_{ijk}\sigma_{jk} = 0$ 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 given by equation (2.7.2), we see that the stress tensor is symmetric. Herein, this symmetry is established without explicitly assuming the equilibrium of an element within the continuum; this assumption is explicitly invoked in Exercise 2.7. We conclude that the symmetry of stress tensor is a general property for any continuum as long as we impose the strong form of Newton's third law of motion; otherwise, the stress tensor is not symmetric.

Herein, due to its symmetry, 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. Thus, the constraints imposed by the balance of angular momentum reduced the number of unknowns.

The symmetry of stress tensor is also known as Cauchy's second law of motion. It is one of the key tenets of classical continuum mechanics. ¹⁶There are generalizations of the classical theory that depart from Cauchy's second law of motion; these generalizations are not discussed in this book. ¹⁷

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

¹⁶Readers interested in an insightful interpretation of Cauchy's second law of motion within classical continuum mechanics might refer to Bunge, M., (1967) Foundations of physics: Springer-Verlag, pp. 153 – 154. Readers interested in the symmetry of stress tensor in the context of classical field theories might refer to Landau, L.D., and Lifshitz, E.M., Course of theoretical physics: Elsevier, Vol. 7 (1986): Theory of elasticity, p. 7, and Vol. 2 (1975): The classical theory of fields, pp. 82 – 85.

¹⁷Readers interested in the generalizations of classical continuum mechanics might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, pp. 217 – 220, to Rymarz, C., (1993) Mechanika ośrodków ciągłych: PWN, pp. 109 – 112, to Truesdell, C., (1966) Six lectures on modern natural philosophy: Springer-Verlag, Lecture II: Polar and oriented media, and study Eringen, A.C., (1999) Microcontinuum field theories I: Foundations and solids: Springer-Verlag, and Forest, S., (2006) Milieux continus généralisés et matériaux hétérogènes: Mines Paris, as well as the books referenced therein.

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 (2.1.8), (2.6.9), (2.6.10) and (2.6.11) 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 the three Cauchy's equations of motion with no body forces,

$$(2.8.1) \quad \sum_{j=1}^3 \frac{\partial \sigma_{ij}}{\partial x_j} = \rho \frac{\partial^2 u_i}{\partial t^2}, \quad i \in \{1, 2, 3\},$$

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

$$(2.8.2) \quad \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_1} \left(\rho \frac{\partial u_1}{\partial t} \right) + \frac{\partial}{\partial x_2} \left(\rho \frac{\partial u_2}{\partial t} \right) + \frac{\partial}{\partial x_3} \left(\rho \frac{\partial u_3}{\partial t} \right) = 0,$$

and

$$(2.8.3) \quad \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} + \frac{\partial \sigma_{13}}{\partial x_3} - \rho \frac{\partial^2 u_1}{\partial t^2} = 0,$$

$$(2.8.4) \quad \frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{23}}{\partial x_3} - \rho \frac{\partial^2 u_2}{\partial t^2} = 0,$$

$$(2.8.5) \quad \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{23}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3} - \rho \frac{\partial^2 u_3}{\partial t^2} = 0.$$

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, which we will discuss in Chapter 3.

To further motivate the formulation of constitutive equation, let us examine again equations (2.8.1). We see that — given σ and ρ — we can obtain $\partial^2 \mathbf{u} / \partial t^2$. However, given the mass density and acceleration, we can obtain only the divergence of the stress tensor and not the stress tensor itself, which would allow us

to describe the internal forces. Notably, the same accelerations associated with two different materials of the same mass density might result in two different responses. Hence, the general principles do not suffice to determine internal forces.

Closing remarks

In Chapter 3, in order to complete the system of equations containing Cauchy's equations of motion and the equation of continuity, we will introduce the constitutive equations describing the relation between stress and strain in an elastic continuum. These constitutive equations will also allow us to associate the fundamental equations formulated in Chapter 2 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.



2.9. Exercises

EXERCISE 2.1. ¹⁸Several 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

$$\int_A^B f(x) \, dx = 0.$$

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. ¹⁹Show that equation of continuity (2.1.7), namely,

$$(2.9.1) \quad \frac{\partial \rho(\mathbf{x}, t)}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0,$$

can be written as

$$(2.9.2) \quad \frac{D\rho(\mathbf{x}, t)}{Dt} + \rho \nabla \cdot \mathbf{v}(\mathbf{x}, t) = 0,$$

where D/Dt stands for the material time-derivative operator, namely,

$$(2.9.3) \quad \frac{D}{Dt} := \frac{\partial}{\partial t} + \mathbf{v}(\mathbf{x}, t) \cdot \nabla$$

SOLUTION 2.2. Taking the divergence on the right-hand side of equation (2.9.1), we write

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \equiv \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v_1, \rho v_2, \rho v_3) = \frac{\partial \rho}{\partial t} + \frac{\partial (\rho v_1)}{\partial x_1} + \frac{\partial (\rho v_2)}{\partial x_2} + \frac{\partial (\rho v_3)}{\partial x_3} = 0.$$

Using the product rule and gathering similar terms, we get

$$\frac{\partial \rho}{\partial t} + \rho \left(\frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3} \right) + [v_1, v_2, v_3] \cdot \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right] \rho = 0.$$

Recognizing the divergence of \mathbf{v} and the gradient operator, we rewrite this expression as

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} + \rho \nabla \cdot \mathbf{v}(\mathbf{x}, t) + \mathbf{v} \cdot \nabla \rho = \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \rho(\mathbf{x}, t) + \rho \nabla \cdot \mathbf{v}(\mathbf{x}, t) = 0.$$

Invoking expression (2.9.3), we obtain

$$\frac{D\rho(\mathbf{x}, t)}{Dt} + \rho \nabla \cdot \mathbf{v}(\mathbf{x}, t) = 0,$$

¹⁸See also Section 2.1.2.

¹⁹See also Section 2.1.3.

which is equation (2.9.2), as required.

EXERCISE 2.3. *Using a single spatial dimension, show that equation (2.2.5) is a case of Leibniz's integral rule; namely,*

$$(2.9.4) \quad \frac{d}{dt} \int_{a(t)}^{b(t)} f(x, t) dx = \int_{a(t)}^{b(t)} \frac{\partial f(x, t)}{\partial t} dx + f(b(t), t) \frac{db}{dt} - f(a(t), t) \frac{da}{dt}.$$

SOLUTION 2.3. Let us write equation (2.2.5) in a single spatial dimension; namely,

$$\frac{d}{dt} \int_{a(t)}^{b(t)} \rho(x, t) \mathcal{A}(x, t) dx = \int_{a(t)}^{b(t)} \rho(x, t) \frac{D\mathcal{A}(x, t)}{Dt} dx,$$

where D/Dt stands for the material time-derivative operator. Invoking its definition (1.3.5) for a single spatial dimension, we write explicitly

$$\begin{aligned} \frac{d}{dt} \int_{a(t)}^{b(t)} \rho(x, t) \mathcal{A}(x, t) dx &= \int_{a(t)}^{b(t)} \rho(x, t) \left(\frac{\partial}{\partial t} + v(x, t) \frac{\partial}{\partial x} \right) \mathcal{A}(x, t) dx \\ &= \int_{a(t)}^{b(t)} \left[\rho(x, t) \frac{\partial \mathcal{A}(x, t)}{\partial t} + \rho(x, t) v(x, t) \frac{\partial \mathcal{A}(x, t)}{\partial x} \right] dx. \end{aligned}$$

In view of expression (2.9.4) and anticipating the substitution given by $f(x, t) = \rho(x, t) \mathcal{A}(x, t)$, let us rewrite the right-hand side of the above equation as

$$\begin{aligned} \int_{a(t)}^{b(t)} \left[\rho(x, t) \frac{\partial \mathcal{A}(x, t)}{\partial t} + \frac{\partial \rho(x, t)}{\partial t} \mathcal{A}(x, t) + \rho(x, t) v(x, t) \frac{\partial \mathcal{A}(x, t)}{\partial x} + \frac{\partial \rho(x, t) v(x, t)}{\partial x} \mathcal{A}(x, t) \right. \\ \left. - \frac{\partial \rho(x, t)}{\partial t} \mathcal{A}(x, t) - \frac{\partial \rho(x, t) v(x, t)}{\partial x} \mathcal{A}(x, t) dx \right]. \end{aligned}$$

Recognizing the product rule of differentiation, we write

$$\int_{a(t)}^{b(t)} \left\{ \frac{\partial}{\partial t} [\rho(x, t) \mathcal{A}(x, t)] + \frac{\partial}{\partial x} [\rho(x, t) v(x, t) \mathcal{A}(x, t)] - \left[\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial \rho(x, t) v(x, t)}{\partial x} \right] \mathcal{A}(x, t) \right\} dx.$$

Recognizing that the third bracketed term is equation of continuity (2.1.7) in a single spatial dimension, we get

$$\int_{a(t)}^{b(t)} \left\{ \frac{\partial}{\partial t} [\rho(x, t) \mathcal{A}(x, t)] + \frac{\partial}{\partial x} [\rho(x, t) v(x, t) \mathcal{A}(x, t)] \right\} dx,$$

which we can write as the sum of two integrals to obtain

$$\begin{aligned} \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} [\rho(x, t) \mathcal{A}(x, t)] dx + \int_{a(t)}^{b(t)} \frac{\partial}{\partial x} [\rho(x, t) v(x, t) \mathcal{A}(x, t)] dx \\ = \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} [\rho(x, t) \mathcal{A}(x, t)] dx + [\rho(x, t) v(x, t) \mathcal{A}(x, t)]_{x=a(t)}^{x=b(t)}. \end{aligned}$$

Using the fact that $v(x, t) = dx/dt$, we write the above result as

$$\int_{a(t)}^{b(t)} \frac{\partial}{\partial t} [\rho(x, t) \mathcal{A}(x, t)] dx + \rho(b(t), t) \frac{db(t)}{dt} \mathcal{A}(b(t), t) - \rho(a(t), t) \frac{da(t)}{dt} \mathcal{A}(a(t), t).$$

Letting $f(x, t) = \rho(x, t) \mathcal{A}(x, t)$, we write

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x, t) dx = \int_{a(t)}^{b(t)} \frac{\partial f(x, t)}{\partial t} dx + f(b(t), t) \frac{db(t)}{dt} - f(a(t), t) \frac{da(t)}{dt},$$

which is equation (2.9.4), as required.

EXERCISE 2.4. ²⁰Using the stress tensor, prove the particular case of the following theorem.

THEOREM 2.9.1. *If an m th-rank tensor is linearly related to an n th-rank tensor through a quantity that possesses $n + m$ indices, then this quantity is an $(n + m)$ th-rank tensor.*

SOLUTION 2.4. 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 2.9.2. 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 values 1, 2 and 3.

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

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

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

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

Combining both expressions, we can write

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

²⁰See also Sections 2.5.2 and 5.2.2.

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

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

Thus, we can write

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

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

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

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

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

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

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

and we can restate it as

$$(2.9.5) \quad \hat{\sigma}_{im} = a_{ik} a_{mj} \sigma_{kj}, \quad i, m \in \{1, 2, 3\}.$$

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. \square

REMARK 2.9.3. The quotient rule, stated in this theorem, is also exemplified by the stress-strain equations (3.2.1), where two second-rank tensors are linearly related by a fourth-rank elasticity tensor, namely,

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

with ε_{kl} denoting the strain tensor, and where the repeated index assumes values 1, 2 and 3.

EXERCISE 2.5. Using expressions (2.5.1) and (2.5.13), 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 (2.5.14).

SOLUTION 2.5. Following expression (2.5.1), we can immediately write the components of the traction vectors acting on the plane whose normal is parallel to the x_1 -axis as $[T_1^{(\mathbf{e}_1)}, T_2^{(\mathbf{e}_1)}, T_3^{(\mathbf{e}_1)}]$. In view of definition (2.5.2), we can rewrite these components as $[\sigma_{11}, \sigma_{12}, \sigma_{13}]$. Using equation (2.5.13), the traction vector acting on the plane whose unit normal is parallel to the x_1 -axis is given by

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

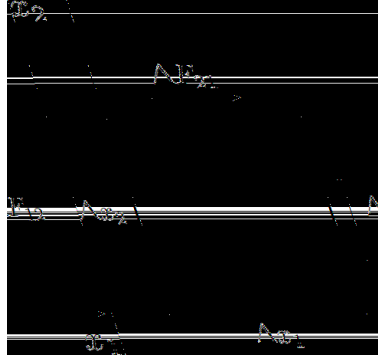


FIGURE 2.9.1. Angular momentum: An x_1x_2 -cross-section of a rectangular box, $\Delta x_1\Delta x_2\Delta x_3$, with the moment-producing forces, F_{12} and F_{21} . The directions of the two forces are perpendicular to one another and their magnitudes are equal.

In view of definition (2.5.2), we can rewrite these components as

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

as expected from the property stated in expression (2.5.14).

EXERCISE 2.6. ²¹Find the physical SI units of equations of motion (2.6.3). Show that these units are consistent for all terms involved.

SOLUTION 2.6. Consider equations (2.6.3); namely,

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

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 $[kg/(m^2s^2)]$. 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 (2.6.3) are $[N/m^3]$.

EXERCISE 2.7. ²²Using Figure 2.9.1, prove the following theorem.

THEOREM 2.9.4. The stress tensor is symmetric; namely, $\sigma_{ij} = \sigma_{ji}$, where $i, j \in \{1, 2, 3\}$.

²¹See also Section 2.6.1.

²²See also Section 2.7.3.

SOLUTION 2.7. PROOF. ²³Consider 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; in other words, we assume this element to be in equilibrium. 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

$$(2.9.6) \quad \Delta F_{12} \Delta x_1 = \Delta F_{21} \Delta x_2.$$

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

$$(2.9.7) \quad \Delta F_{12} = \sigma_{12} \Delta S_1 = \sigma_{12} \Delta x_2 \Delta x_3,$$

and

$$(2.9.8) \quad \Delta F_{21} = \sigma_{21} \Delta S_2 = \sigma_{21} \Delta x_1 \Delta x_3.$$

Inserting expressions (2.9.7) and (2.9.8) into equation (2.9.6), 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

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

as required. □

²³Readers interested in more details of the argument used herein as well as the consequences of this symmetry might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, pp. 77 – 79.

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 a brief discussion of the rudiments of constitutive equations. Subsequently, we formulate the linear stress-strain equations, which underlie the theory of elasticity used in this book. We express these equations in both tensor and matrix forms.

3.1. Rudiments of constitutive equations

Fundamentally, a constitutive equation is a relation between two physical quantities that is specific to a material of a given composition. This relation does not stem from any fundamental physical principle but is not contradictory with such principles. In the context of this book, the two physical quantities being

related by the constitutive equation are the stress and strain tensors. The key objective of constitutive equations is the determination of stress within the continuum and the representation of the variety of materials. Furthermore, we need these equations to verify the theory of continuum mechanics. In other words, for an empirical validation of the theory, we have to enrich the general theory with constitutive equations that allow us to predict experimental results.

¹There are three principles that constitutive equations must obey. First, the principle of determinism requires that the stresses within a body be determined uniquely by the history of its deformation. Secondly, the principle of local action requires that the deformation outside an arbitrarily small neighbourhood of a point may be disregarded in determining stresses at that point. This means that there is no need to invoke action-at-a-distance between stress and strain or to consider a delay due to spatial separation, which would require relativistic continuum mechanics.

These two principles are satisfied trivially in the theory of linear elasticity discussed in this book, where a stress at a given instant does not depend on the history of deformation, and only tractional forces are involved. Thirdly, the principle of objectivity, also known as the principle of material frame indifference, requires that any two observers find the same stress within a given body. In other words, constitutive equations must be invariant under changes of a reference frame that preserves the essential structure of the manifold of physical experience discussed on page 10; the principle of objectivity limits our study to nonrelativistic continuum mechanics. ²For further insight, let us quote an example stated in Truesdell's "Six lectures on modern natural philosophy" that is particularly applicable to the theory of elasticity.

Take a spring, and on one end hang a weight of one pound. The spring lengthens, say by one inch. Now, lay a spring on a horizontal table, fastening one end to the center, and leaving the weight attached to the other end. Spin the table, and adjust the angular speed until the spring again stretches exactly one inch. [...] For an observer standing on the floor as well as for an observer seated upon that table [...] one inch of extension corresponds to one poundal of force.³

Unlike the general principles, the constitutive equations can be, and often are, contradictory among each other. Their limited validity results in specialized theories, as exemplified by elasticity and by fluid mechanics.

¹Readers interested in more detail and elegant exposition might refer to Truesdell, C., (1966) Six lectures on modern natural philosophy: Springer-Verlag, pp. 3 – 6.

²Readers interested in a rigorous formulation of the principle of material frame indifference might refer to Marsden, J.E., and Hughes, T.J.R., (1983/1994) Mathematical foundations of elasticity: Dover, p. 8 and pp. 189 – 199.

³Readers interested in the equality of inertial mass and gravitational mass, which is invoked in the above example, might refer to Bunge, M., (1967) Foundations of physics: Springer-Verlag, pp. 207 – 210.

3.2. Formulation of stress-strain equations: Hookean solid

3.2.1. 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.⁴ 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.

The mathematical justification of Hooke's law and the resulting linear theory stems from the assumption of small deformations since every function is linear on the infinitesimal level. We also assume that the stress tensor is zero for the undeformed state, which corresponds to the strain tensor being zero; there are no constant terms in the stress-strain equations.

⁴Readers interested in a historical comments about Hooke's law might refer to Love, A.E.H., (1892/1944) A treatise on the mathematical theory of elasticity: Dover, p. 2.

3.2.2. Tensor form.

Stress-strain equations. At a given point of a three-dimensional continuum, Hooke's law — which linearly relates each stress-tensor component, σ_{ij} , to all the strain-tensor components, ε_{kl} — can be written as

$$(3.2.1) \quad \sigma_{ij} = \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl}, \quad i, j \in \{1, 2, 3\},$$

where c_{ijkl} are the components of a tensor, known as the elasticity tensor.⁵ Equations (3.2.1) are the constitutive equations of a linearly elastic continuum. 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 .

The essence of the above formulation is the assumption of a functional relationship given by $\sigma_{ij} = \sigma_{ij}(\varepsilon_{ij})$; in other words, stress is only a function of strain, and vice versa. If we develop this function in a power series, set the constant term and all the nonlinear terms to zero, we obtain expression (3.2.1). As stated above, the justification for setting the constant term to zero is the assumption of no initial stress, and the justification of setting the nonlinear terms to zero is the assumption of infinitesimal displacements. Such a formulation of stress-strain equations (3.2.1) is referred to as Cauchy's approach. We could also formulate these equations in a manner stated on page 80, and referred to as Green's approach.

Tensor 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. However, in view of the symmetries of the stress and strain tensors, the number of independent components is thirty-six, as shown below. The physical meaning of the rank of a tensor can be viewed as the number of directions necessary to measure the corresponding property. Tensor c_{ijkl} relates four directions necessary to measure elasticity: two directions of the stress tensor and two directions of the strain tensor, which are the direction of the force together with the normal to the face on which the force is acting and the direction of the displacement vector together with the orientation of the measurement axis, respectively.

The elasticity tensor describes properties of the continuum; hence, it is a property tensor in the study of elasticity in the same manner as a dielectric tensor is a property tensor in the study of electricity. The stress and strain tensors are not property tensors since we can apply various forces to, or cause different deformations of, a given material to study its elasticity. In the same manner, we can apply various electric fields and displacements to study material's ability to store electric charge.⁶

⁵Elasticity 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.

⁶Readers interested in analogies among mechanical, electrical and thermal variables might refer to Heckmann's diagrams, e.g., Newnham, R.E., (2005) *Properties of materials: Anisotropy, symmetry, structure*: Oxford University Press, pp. 1 – 3.

Stress-tensor and strain-tensor symmetries. At every point of a continuum, the stress tensor is symmetric, namely, $\sigma_{ij} = \sigma_{ji}$, and consequently has six independent components, as shown in Section 2.7 in the context of the balance of angular momentum. Also, the strain tensor is symmetric, namely, $\varepsilon_{kl} = \varepsilon_{lk}$, by its definition (1.4.6), and also has only six independent components. To investigate the number of independent components of tensor c_{ijkl} , which relates the stress and strain tensors, let us consider stress-strain equations (3.2.1).

Consider the symmetry of the stress tensor. In view of this symmetry, we can write stress-strain equations (3.2.1) as

$$(3.2.2) \quad \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl} = \sigma_{ij} = \sigma_{ji} = \sum_{k=1}^3 \sum_{l=1}^3 c_{jikl} \varepsilon_{kl}, \quad i, j \in \{1, 2, 3\}.$$

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

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

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

$$(3.2.3) \quad c_{ijkl} = c_{jikl}, \quad i, j, k, l \in \{1, 2, 3\}.$$

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 (3.2.1) since they are the summation indices. Hence, we can write

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

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

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

which, we can also state as

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

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

$$(3.2.4) \quad c_{ijkl} = c_{ijlk}, \quad i, j, k, l \in \{1, 2, 3\}.$$

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

In view of equalities (3.2.3) and (3.2.4), the number of independent components of the elasticity tensor is thirty-six.

Anisotropy and inhomogeneity. In this book, we study continua that are described by stress-strain equations (3.2.1).⁷ 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 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 (7.1.2).

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, layers or fractures 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.

3.2.3. Matrix form.

Introductory comments. Due to the symmetries of the stress and strain tensors, constitutive equations (3.2.1) can be conveniently written in a matrix form containing six independent equations. This form, which allows us to express elasticity tensor (3.2.1) 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.

Elasticity matrix. The thirty-six components of the elasticity tensor that are independent of each other in spite of the symmetries of stress and strain can be written as entries C_{mn} of a 6×6 elasticity matrix, which relates each independent stress-tensor component to the six independent strain-tensor components. To construct this matrix, in view of symmetries (3.2.3) and (3.2.4), 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

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

⁷Readers 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.

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

$$\begin{aligned} (1, 1) &\rightarrow 1, (2, 2) \rightarrow 2, (3, 3) \rightarrow 3, \\ (2, 3) &\rightarrow 4, (1, 3) \rightarrow 5, (1, 2) \rightarrow 6. \end{aligned}$$

We can concisely write this replacement as

$$(3.2.5) \quad \begin{cases} m = i & \text{if } i = j \\ m = 9 - (i + j) & \text{if } i \neq j \end{cases}, \quad i, j \in \{1, 2, 3\}.$$

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 $m, n \in \{1, \dots, 6\}$, to obtain the elasticity matrix given by

$$(3.2.6) \quad \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix},$$

which we use below to write the matrix form of the stress-strain equations.

Stress-strain equations. Using matrix (3.2.6) and in view of the symmetries of the stress and strain tensors, equations (3.2.1) can be restated as

$$(3.2.7) \quad \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix}.$$

The factors of 2 result from the symmetry of the strain tensor: for each $k \neq l$, the corresponding strain-tensor component appears twice in the summation on the right-hand side of equations (3.2.1) — as ε_{kl} and as ε_{lk} — as shown in Exercise 3.1. Also, due to the symmetry of the stress tensor, it is sufficient to consider only six of the nine equations stated in expression (3.2.1). Using formula (3.2.5), we could have replaced 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 1.4.3 and 2.5.1. In a concise notation, we write stress-strain equations (3.2.7) as

$$(3.2.8) \quad \underline{\sigma} = \mathbf{C}\underline{\varepsilon},$$

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.

Writing the stress-strain equations in a matrix notation, as shown in expression (3.2.7), is a convenient way to display the stress-tensor, strain-tensor and elasticity-tensor components. We will use this notation extensively in Chapter 5.

3.3. 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 (1.4.6), 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 3.2. 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 (2.8.2), the three equations of motion, given by expressions (2.8.3), (2.8.4) and (2.8.5), 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 (2.6.3) refer to the spatial coordinates while definition (1.4.6), 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.

3.4. Anelasticity: Example

3.4.1. Introductory comments. For the study presented in this book, we will use constitutive equations (3.2.1). In other words, we will study physical phenomena in the context of linear elasticity. To see this choice in the scope of other possibilities, we will briefly discuss the concept of viscoelasticity.

3.4.2. Viscosity: Stokesian fluid. Constitutive equations (3.2.1) state that deformation is proportional to the applied load; they define a Hookean solid. We could also state that deformation is proportional to the rate of the application of the load. With a certain abuse of notation, let us formally write

$$(3.4.1) \quad \sigma(t) = \eta \frac{d\varepsilon(t)}{dt},$$

where η is the viscosity parameter. Constitutive equation (3.4.1) defines a Stokesian fluid, named in deference to George Gabriel Stokes, an English physicist. Integrating both sides of this equation, we can rewrite it as

$$(3.4.2) \quad \varepsilon(t) = \frac{1}{\eta} \int \sigma(t) dt.$$

If we submit an elastic continuum, $\sigma = c\varepsilon$, to a constant load, σ_0 , it deforms instantaneously to $\varepsilon = \sigma_0/c$, where c is the elasticity parameter. Thereafter, the elastic continuum remains deformed at σ_0/c as long as the load is applied, and immediately returns to its undeformed state upon the removal of the load; this behaviour is shown in Exercise 3.3.

To study the effect of the constant load on a viscous continuum, let us first write such a load as

$$(3.4.3) \quad \sigma(t) = \sigma_0 h(t),$$

where h is Heaviside's function, which is defined by

$$h(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}.$$

To subject a viscous continuum — represented by a Stokesian fluid — to the constant load, we insert expression (3.4.3) into equation (3.4.2) to get

$$(3.4.4) \quad \varepsilon(t) = \frac{\sigma_0}{\eta} \int h(t) dt = \frac{\sigma_0}{\eta} \int_0^t d\tau = \frac{\sigma_0}{\eta} t.$$

Since $(\sigma_0/\eta)t$ is a straight line whose slope is σ_0/η , this result means that the deformation increases linearly as long as the load is applied. If we remove the load at time t_1 , the deformation remains at $(\sigma_0/\eta)t_1$, as shown in Exercise 3.4.

3.4.3. Viscoelasticity: Kelvin-Voigt model. To model real media, it is possible to combine elastic and viscous elements.⁸ A common combination, called the Kelvin-Voigt model, consists of combining in parallel elastic and viscous elements, which can be represented by a spring and dashpot, respectively. Before continuing our study, let us gain a perspective on the meaning of arrangements of springs and dashpots by quoting Coleman and Noll.

We feel that the physicist's confidence in the usefulness of the theory of infinitesimal viscoelasticity does not stem from the belief that the materials to which the theory is applied are really composed of microscopic networks of springs and dashpots, but comes rather from other considerations. First, there is the observation that the theory works for many real materials. But second, and perhaps more important, is the fact that the theory looks plausible because it seems to be a mathematization of little more than certain intuitive prejudices about smoothness in macroscopic phenomena.

⁸Readers interested in a description of several such combinations might refer to Ben-Menahem, A., and Singh, S.J., (1981/2000) *Seismic waves and sources*: Dover, pp. 850 – 865.

Hence, in the spirit of continuum mechanics, we write the constitutive equation of the Kelvin-Voigt model as

$$(3.4.5) \quad \sigma(t) = c\varepsilon(t) + \eta \frac{d\varepsilon(t)}{dt}.$$

Herein, ε is the same for both elements. This would not be the case if we connected these two elements in series, which is known as the Maxwell model. Examining equation (3.4.5), we see that it reduces to $\sigma = c\varepsilon$, if $\eta = 0$, and to $\sigma = \eta d\varepsilon/dt$, if $c = 0$. In other words, the Hookean solid and the Stokesian fluid can be viewed as the two extreme cases of the viscoelastic continuum given by the Kelvin-Voigt model.

Let us examine the behaviour of the Kelvin-Voigt model under load, σ . Given σ , we can solve differential equation (3.4.5) to get

$$(3.4.6) \quad \varepsilon(t) = \frac{\varepsilon(0)}{\exp\left(\frac{c}{\eta}t\right)} + \frac{1}{\eta} \int_0^t \frac{\sigma(\tau)}{\exp\left[\frac{c}{\eta}(t-\tau)\right]} d\tau,$$

as shown in Exercise 3.5.

Let us consider a constant load, σ_0 , applied at t_0 and removed at t_1 . We can express such a load as

$$(3.4.7) \quad \sigma(t) = \sigma_0 [h(t) - h(t - t_1)],$$

where h is Heaviside's function. Inserting this expression into expression (3.4.6) and assuming that $\varepsilon(0) = 0$, which means that there is no deformation prior to $t = 0$, we get

$$(3.4.8) \quad \varepsilon(t) = \frac{\sigma_0}{c} \left[1 - \frac{1}{\exp\left(\frac{c}{\eta}t\right)} \right],$$

for $0 < t < t_1$, which describes the process of deformation, and

$$(3.4.9) \quad \varepsilon(t) = \varepsilon(t_1) \frac{1}{\exp\left[\frac{c}{\eta}(t - t_1)\right]},$$

for $t > t_1$, which describes the process of recovery.

Let us discuss these results. Investigating equation (3.4.8), we see that, if η tends to zero for $0 < t < t_1$, $\varepsilon(t) = \sigma_0/c$. Investigating equation (3.4.9), we see that if η tends to zero for $t > t_1$, $\varepsilon(t) = 0$. These cases correspond to the behaviour exhibited by elasticity. If c tends to zero, equation (3.4.8) reduces to $\varepsilon(t) = (\sigma_0/\eta)t$, as can be seen by invoking de l'Hôpital's rule to differentiate both the numerator and denominator with respect to c . If c tends to zero, equation (3.4.9) reduces to $\varepsilon(t) = \varepsilon(t_1)$. The two latter results correspond to the behaviour exhibited by viscosity, as described in Section 3.4.2.

If we take the derivative of expression (3.4.8) and evaluate it at $t = 0$, we get $d\varepsilon/dt|_{t=0} = \sigma_0/\eta$, which is the rate of deformation of the viscous element, as shown in equation (3.4.4). If we take the limit of the same expression as t tends to infinity, we get $\varepsilon(t) = \sigma_0/c$, which is the deformation of the elastic

element — as t tends to infinity, the deformation asymptotically approaches the state that would be reached instantaneously by the elastic element alone. Thus, initially, the stress is sustained by the viscous element and then it is carried by the elastic element.

If we take the limit of expression (3.4.9) as t tends to infinity, we get $\varepsilon(t) = 0$. This means that upon the removal of the load, the model asymptotically approaches its undeformed state, which would be reached instantaneously by the elastic element alone.

The viscous element slows down the processes of deformation and recovery. In this book, we assume these two processes to be instantaneous by using equations (3.2.1).

Closing remarks

We use constitutive equations, namely, stress-strain equations (3.2.1) or, equivalently, equations (3.2.7) 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 (3.2.1), (3.2.7) or (5.14.7) 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 (3.2.1) or (3.2.7) describe the continuum whose deformations are linearly related to loads. Also, examining these equations, we see that the vanishing of stress in this continuum is accompanied by the disappearance of strain therein, as illustrated in Exercise 3.3. For such a 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} .



3.5. Exercises

EXERCISE 3.1. *Using stress-strain equations (3.2.1) exemplify the presence of the factor 2 in stress-strain equations (3.2.7).*

SOLUTION 3.1. Using stress-strain equations (3.2.1), namely,

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

we explicitly write

$$\sigma_{23} = c_{2311}\varepsilon_{11} + c_{2312}\varepsilon_{12} + c_{2313}\varepsilon_{13} + c_{2321}\varepsilon_{21} + c_{2322}\varepsilon_{22} + c_{2323}\varepsilon_{23} + c_{2331}\varepsilon_{31} + c_{2332}\varepsilon_{32} + c_{2333}\varepsilon_{33}.$$

In view of the symmetry of the strain tensor, $\varepsilon_{ij} = \varepsilon_{ji}$, which implies that $c_{23ij} = c_{23ji}$, we can write

$$\sigma_{23} = c_{2311}\varepsilon_{11} + c_{2322}\varepsilon_{22} + c_{2333}\varepsilon_{33} + 2c_{2323}\varepsilon_{23} + 2c_{2313}\varepsilon_{13} + 2c_{2312}\varepsilon_{12}.$$

Using formula (3.2.5), we restate this equation in the matrix form, namely,

$$\sigma_{23} = \begin{bmatrix} C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix},$$

with the factor 2, as required.

EXERCISE 3.2. ⁹*Consider a one-dimensional homogeneous continuum. Using stress-strain equations (3.2.7), equation of continuity (2.1.8) and Cauchy's equations of motion (2.6.3) with no body force, write the resulting system of two differential equations.*

SOLUTION 3.2. Following equations (3.2.7) 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 (1.4.6) can be written as

$$(3.5.1) \quad \sigma_{11} = C_{11} \frac{\partial u_1(x, t)}{\partial x_1},$$

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 (2.1.8), is

$$(3.5.2) \quad \frac{\partial \rho(x, t)}{\partial t} + \frac{\partial}{\partial x_1} \left[\rho(x, t) \frac{\partial u_1(x, t)}{\partial t} \right] = 0,$$

⁹See also Sections 3.3 and 4.4.2.

and Cauchy's equation of motion, whose general form is given by expression (2.6.3), is

$$(3.5.3) \quad \frac{\partial \sigma_{11}}{\partial x_1} = \rho(x, t) \frac{\partial^2 u_1(x, t)}{\partial t^2}.$$

Inserting expression (3.5.1) into equation (3.5.3), differentiating and rearranging, we obtain

$$(3.5.4) \quad \frac{\partial^2 u_1(x, t)}{\partial x_1^2} = \frac{\rho(x, t)}{C_{11}} \frac{\partial^2 u_1(x, t)}{\partial t^2}.$$

Equations (3.5.2) and (3.5.4) 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 3.5.1. If the mass-density function, $\rho(x, t)$, is given by a constant, equation (3.5.4) is a one-dimensional wave equation, discussed in Chapter 6.

EXERCISE 3.3. Consider the stress-strain equation for an elastic continuum written formally as

$$\sigma(t) = c\varepsilon(t).$$

Examine deformation ε that results from the load given by

$$(3.5.5) \quad \sigma(t) = \sigma_0 [h(t) - h(t - t_1)],$$

where h is Heaviside's function.

REMARK 3.5.2. Expression (3.5.5) describes a load whose magnitude is equal to σ_0 . This load is applied at $t = 0$ and removed at $t = t_1$.

SOLUTION 3.3. Let us write strain as a function of stress; in other words,

$$\varepsilon(t) = \frac{1}{c} \sigma(t).$$

Inserting herein expression (3.5.5), we obtain

$$\varepsilon(t) = \frac{\sigma_0}{c} [h(t) - h(t - t_1)].$$

Examining this result in view of the definition of Heaviside's function, we see that the continuum is deformed by the value of $\varepsilon = \sigma_0/c$ at $t = 0$, and it returns to its undeformed state — namely, $\varepsilon = 0$ — at $t = t_1$. We also see that in the idealized material described by Hooke's law, both the deformation and return to the undeformed state are instantaneous.

EXERCISE 3.4. Show that for a viscous element the deformation remains at $(\sigma_0/\eta)t_1$ upon the removal of load at time t_1 .

SOLUTION 3.4. Inserting expression (3.5.5) into expression (3.4.2), we write

$$\varepsilon(t) = \frac{\sigma_0}{\eta} \int [h(t) - h(t - t_1)] dt.$$

Following the properties of Heaviside's function, we get

$$\varepsilon(t) = \frac{\sigma_0}{\eta} \left(\int_0^t d\tau - \int_{t_1}^t d\tau \right) = \frac{\sigma_0}{\eta} [t - (t - t_1)] = \frac{\sigma_0}{\eta} t_1,$$

as required.

EXERCISE 3.5. Given function σ , solve equation (3.4.5), namely,

$$\sigma(t) = c\varepsilon(t) + \eta \frac{d\varepsilon(t)}{dt},$$

where c and η are constants.

SOLUTION 3.5. This is a linear first-order ordinary differential equation. We would like to integrate $d\varepsilon/dt$ to obtain $\varepsilon(t)$. To achieve that, we multiply both sides by $\varsigma(t)$ to get

$$(3.5.6) \quad \frac{1}{\eta} \sigma(t) \varsigma(t) = \frac{d\varepsilon(t)}{dt} \varsigma(t) + \varepsilon(t) \frac{c}{\eta} \varsigma(t).$$

Then, we consider

$$(3.5.7) \quad \frac{d}{dt} [\varepsilon(t) \varsigma(t)] = \frac{d\varepsilon(t)}{dt} \varsigma(t) + \varepsilon(t) \frac{d\varsigma(t)}{dt}.$$

We wish to find η such that the right-hand sides of these equations are equal to one another. Thus, we require

$$\frac{d\varsigma(t)}{dt} = \frac{c}{\eta} \varsigma(t),$$

which we can rewrite as

$$\frac{\frac{d\varsigma(t)}{dt}}{\varsigma(t)} = \frac{c}{\eta}.$$

Integrating both sides, we get

$$\ln |\varsigma(t)| = \frac{c}{\eta} t + C,$$

where C is the integration constant. Solving for ς , we get

$$(3.5.8) \quad \varsigma(t) = \exp \left(\frac{c}{\eta} t + C \right) = K \exp \left(\frac{c}{\eta} t \right),$$

where $K = \exp C$; for our use of ς , we can let $C = 0$, and hence $K = 1$. Since the right-hand sides of equations (3.5.6) and (3.5.7) are equal to one another, if $\varsigma = \exp [(c/\eta) t]$, we can equate also the left-hand sides of these equations by writing

$$\frac{d}{dt} \left[\varepsilon(t) \exp \left(\frac{c}{\eta} t \right) \right] = \frac{1}{\eta} \exp \left(\frac{c}{\eta} t \right) \sigma(t).$$

Integrating both sides with respect to the integration variable, τ , we get

$$\varepsilon(\tau) \exp \left(\frac{c}{\eta} \tau \right) \Big|_0^t = \frac{1}{\eta} \int_0^t \exp \left(\frac{c}{\eta} \tau \right) \sigma(\tau) d\tau,$$

with the lower limit of integration referring to the onset of load at $t = 0$. Evaluating the left-hand side, we get

$$\varepsilon(t) \exp\left(\frac{c}{\eta}t\right) - \varepsilon(0) = \frac{1}{\eta} \int_0^t \exp\left(\frac{c}{\eta}\tau\right) \sigma(\tau) d\tau.$$

Solving for $\varepsilon(t)$, we obtain

$$\varepsilon(t) = \frac{\varepsilon(0)}{\exp\left(\frac{c}{\eta}t\right)} + \frac{1}{\eta} \int_0^t \frac{\sigma(\tau)}{\exp\left[\frac{c}{\eta}(t-\tau)\right]} d\tau,$$

which is the required solution.

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.

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 the elasticity tensor, beyond the ones shown in Chapter 3. Then we derive the physical constraints on the components of the elasticity tensor; these constraints arise from the strain-energy function. This chapter concludes with the system of equations describing the behaviour of elastic continua.

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 that 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 that unifies and inspires the entire body of the formulation.

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

$$(4.1.1) \quad \frac{\partial W(\varepsilon)}{\partial \varepsilon_{ij}} = \sigma_{ij}, \quad i, j \in \{1, 2, 3\},$$

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 emphasize the fundamental importance of expression (4.1.1), let us comment on stress-strain equations (3.2.1). To obtain these equations in Chapter 3, we used Cauchy's approach, which assumes the existence of a function given by $\sigma_{ij} = \sigma_{ij}(\varepsilon_{ij})$. Also, we could view expression (4.1.1) as another point of departure; such a method is called Green's approach. The essence of such a formulation is the existence of function W , which underlies the concept of elasticity. For Green's approach, we would expand W in a power series of ε_{ij} , and retain only the quadratic terms.²

Herein, to obtain the explicit expression for W , we use stress-strain equations (3.2.1) to write expression (4.1.1) as

$$(4.1.2) \quad \frac{\partial W(\varepsilon)}{\partial \varepsilon_{ij}} = \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl}, \quad i, j \in \{1, 2, 3\}.$$

Integrating both sides of equations (4.1.2) with respect to ε_{ij} , we obtain

$$(4.1.3) \quad W(\varepsilon) = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl} \varepsilon_{ij},$$

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 (4.1.3) is the strain-energy function. It is the desired potential-energy function that corresponds to elastic continua subjected to infinitesimal strains.³

Note that W has the units of energy per volume.

²Readers interested in similarities and distinctions between Cauchy's and Green's approaches to formulate stress-strain equations (3.2.1) might refer to Graff, K.F., (1975/1991) Wave motion in elastic solids: Dover, pp. 590 – 591.

³Readers 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.

Examining expression (4.1.3), we recognize that strain energy is given by a homogeneous function of degree 2 in the strain-tensor components.⁴ The fact that the strain-energy function is homogeneous of degree 2 in the ε_{ij} , follows from Definition A.1.1, 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 (4.1.3) 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 (4.1.2). Differentiating both sides of these equations with respect to ε_{kl} , we obtain

$$(4.2.1) \quad \frac{\partial^2 W(\varepsilon)}{\partial \varepsilon_{kl} \partial \varepsilon_{ij}} = c_{ijkl}, \quad i, j, k, l \in \{1, 2, 3\}.$$

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.⁵ In view of expression (4.2.1), this implies that

$$(4.2.2) \quad c_{ijkl} = c_{klij}, \quad i, j, k, l \in \{1, 2, 3\}.$$

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.⁶ Recalling that σ_{ij} and ε_{ij} are associated with force and displacement, respectively, we can write the element of work as

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

⁴Both 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”.

⁵The equality of mixed partial derivatives is often used in this book. We can state it by the following theorem.

THEOREM 4.2.1. *Let $f = f(x, y)$. Assume that the partial derivatives $\partial f / \partial x$, $\partial f / \partial y$, $\partial^2 f / \partial x \partial y$ 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.

⁶Readers interested in this formulation might also refer to Ting, T.C.T., (1996) *Anisotropic elasticity: Theory and applications*: Oxford University Press, p. 33.

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

$$(4.2.3) \quad dW = \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} d\varepsilon_{ij}.$$

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 (3.2.1), we can rewrite expression (4.2.3) as

$$(4.2.4) \quad dW = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl} d\varepsilon_{ij}.$$

Since this has to be a total differential, we require that⁷

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

which gives

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

Upon renaming the indices, we can write

$$(4.2.5) \quad c_{ijkl} = c_{klij}, \quad i, j, k, l \in \{1, 2, 3\},$$

which are equations (4.2.2). Hence, we rederived equations (4.2.2).

Also, in view of conditions (4.2.5), we can integrate both sides of equation (4.2.4) to obtain

$$(4.2.6) \quad W = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl} \varepsilon_{ij},$$

which is expression (4.1.3), as expected.

4.2.2. Elasticity parameters. In view of the symmetries of the stress and strain tensors, the eighty-one components of the elasticity tensor, c_{ijkl} , consist of only thirty-six independent components, as shown in Section 3.2.2. Furthermore, since every elastic continuum must obey equations (4.1.1), conditions (4.2.2) provide additional constraints on this tensor. Specifically, in view of the strain-energy function and the resulting symmetry stated in expression (4.2.2), the number of independent components is twenty-one.

⁷Readers 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.

Since there are no more reductions in the number of independent components of the elasticity tensor that stem from the fundamental equations, we refer to these remaining twenty-one independent components as the elasticity parameters. These parameters — together with mass density — fully describe a linearly elastic continuum.

4.2.3. Matrix form of stress-strain equations. As shown in Sections 3.2.2 and 3.2.3, we can express stress-strain equations in both the tensor and matrix forms. Recalling expressions (3.2.5) and (4.2.2), we see that $c_{ijkl} = c_{klij}$ 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,

$$(4.2.7) \quad \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}.$$

Thus, the number of independent entries of \mathbf{C} is reduced from thirty-six, used in equations (3.2.7), to twenty-one, stated in matrix (4.2.7).

In view of conditions (4.2.2) and resulting matrix (4.2.7), stress-strain equations (3.2.7) become

$$(4.2.8) \quad \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{bmatrix}.$$

Equations (4.2.8) are the matrix form of the stress-strain equations for a general elastic continuum that obeys Hooke's law.

4.2.4. 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 (4.2.6) that contains a given set of elasticity parameters holds only for one orientation of the coordinate system.⁸ If, for a particular continuum, expression (4.2.6) with a given set of

⁸Interested readers might refer to Malvern, L.E., (1969) Introduction to the mechanics of a continuous medium: Prentice-Hall, p. 285.

elasticity parameters holds for several orientations of the coordinate system, this continuum possesses particular symmetries, which lead to further simplifications of matrix (4.2.7). Such symmetries are discussed in Chapter 5. The invariance of W under coordinate transformations will be explicitly used in Section 5.10.2.

4.3. Stability conditions

4.3.1. Physical justification. Strain-energy function, W , given in expression (4.1.3), is formulated in terms of parameters c_{ijkl} , where $i, j, k, l \in \{1, 2, 3\}$. This function provides the sole 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.⁹

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 (4.1.3) and by equivalence of stress-strain equations (3.2.1) and (3.2.7), we can write the strain-energy function as

$$(4.3.1) \quad W = \frac{1}{2} (\mathbf{C}\underline{\underline{\varepsilon}}) \cdot \underline{\underline{\varepsilon}},$$

where \mathbf{C} is matrix (4.2.7), and $\underline{\underline{\varepsilon}}$ is the strain matrix, shown explicitly in equation (3.2.7). In view of Section 4.3.1, we require that

$$(4.3.2) \quad (\mathbf{C}\underline{\underline{\varepsilon}}) \cdot \underline{\underline{\varepsilon}} \geq 0,$$

where the equality sign corresponds to the case where $\underline{\underline{\varepsilon}} = \mathbf{0}$. Expression (4.3.2) states the positive-definiteness of matrix \mathbf{C} . In other words, matrix \mathbf{C} is positive-definite if and only if $(\mathbf{C}\underline{\underline{\varepsilon}}) \cdot \underline{\underline{\varepsilon}} > 0$, for all $\underline{\underline{\varepsilon}}$, such that $\underline{\underline{\varepsilon}} \neq \mathbf{0}$.

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 algebra¹⁰; namely,

⁹Readers 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.

¹⁰Readers interested in proofs of Theorem 4.3.1 and Theorem 4.3.2 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 4.3.2 follows from Theorem 4.3.1.

THEOREM 4.3.1. *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.*

or

THEOREM 4.3.2. *A real symmetric matrix is positive-definite if and only if all its eigenvalues are positive.*

Since matrix (4.2.7) is symmetric, the stability conditions can be conveniently formulated based on Theorems 4.3.1 and 4.3.2, as shown in Exercises 5.5 and 5.14, respectively. Among these conditions, we find that

$$(4.3.3) \quad C_{mm} > 0, \quad m \in \{1, \dots, 6\},$$

which implies that all the main-diagonal entries of the elasticity matrix must be positive, as shown in Exercise 4.5.¹¹

Stability conditions cannot be violated. However, as shown in Exercise 5.14, 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 (3.2.1), namely,

$$(4.4.1) \quad \sigma_{ij} = \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl}, \quad i, j \in \{1, 2, 3\},$$

where, in view of equations (3.2.3), (3.2.4) and (4.2.2), we require

$$(4.4.2) \quad c_{ijkl} = c_{jikl} = c_{ijlk} = c_{klij}, \quad i, j, k, l \in \{1, 2, 3\}.$$

We recall that symmetries (4.4.2) result from definition (1.4.6), which implies $\varepsilon_{kl} = \varepsilon_{lk}$, as well as from the balance of angular momentum, stated in expression (2.7.2), namely,

$$\frac{d}{dt} \iiint_{V(t)} \left(\mathbf{x} \times \rho \frac{d\mathbf{u}}{dt} \right) dV = \iint_{S(t)} (\mathbf{x} \times \mathbf{T}) dS + \iiint_{V(t)} (\mathbf{x} \times \mathbf{f}) dV,$$

which implies $\sigma_{ij} = \sigma_{ji}$. Symmetries (4.4.2) also result from the existence of the strain-energy function that is given by expression (4.1.3), namely,

$$(4.4.3) \quad W(\varepsilon) = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{kl} \varepsilon_{ij},$$

¹¹Readers 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.

and must satisfy equation (4.1.1); namely,

$$(4.4.4) \quad \frac{\partial W(\varepsilon)}{\partial \varepsilon_{ij}} = \sigma_{ij}, \quad i, j \in \{1, 2, 3\}.$$

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 (4.4.1)], and they are also partial differential coefficients of a function W of the strain components [equations (4.4.4)]. The strain-energy function, W , is therefore a homogeneous quadratic function of the strain components [equations (4.4.3)].

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 $m, n \in \{1, \dots, 6\}$.

Note that four among ten equations result from the fundamental principles, which are given by the conservation of mass, stated in equation (2.1.6), namely,

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

and the balance of linear momentum, stated in equation (2.4.5), namely,

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

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

Explicitly, we can write this system as

$$(4.4.5) \quad \left\{ \begin{array}{l} \frac{\partial}{\partial t} \rho = - \left[\frac{\partial}{\partial x_1} \left(\rho \frac{\partial}{\partial t} u_1 \right) + \frac{\partial}{\partial x_2} \left(\rho \frac{\partial}{\partial t} u_2 \right) + \frac{\partial}{\partial x_3} \left(\rho \frac{\partial}{\partial t} u_3 \right) \right] \\ \rho \frac{\partial^2}{\partial t^2} u_1 = \frac{\partial}{\partial x_1} \sigma_{11} + \frac{\partial}{\partial x_2} \sigma_{12} + \frac{\partial}{\partial x_3} \sigma_{13} \\ \rho \frac{\partial^2}{\partial t^2} u_2 = \frac{\partial}{\partial x_1} \sigma_{12} + \frac{\partial}{\partial x_2} \sigma_{22} + \frac{\partial}{\partial x_3} \sigma_{23} \\ \rho \frac{\partial^2}{\partial t^2} u_3 = \frac{\partial}{\partial x_1} \sigma_{13} + \frac{\partial}{\partial x_2} \sigma_{23} + \frac{\partial}{\partial x_3} \sigma_{33} \\ \sigma_{11} = C_{11} \varepsilon_{11} + C_{12} \varepsilon_{22} + C_{13} \varepsilon_{33} + 2C_{14} \varepsilon_{23} + 2C_{15} \varepsilon_{13} + 2C_{16} \varepsilon_{12} \\ \sigma_{22} = C_{12} \varepsilon_{11} + C_{22} \varepsilon_{22} + C_{23} \varepsilon_{33} + 2C_{24} \varepsilon_{23} + 2C_{25} \varepsilon_{13} + 2C_{26} \varepsilon_{12} \\ \sigma_{33} = C_{13} \varepsilon_{11} + C_{23} \varepsilon_{22} + C_{33} \varepsilon_{33} + 2C_{34} \varepsilon_{23} + 2C_{35} \varepsilon_{13} + 2C_{36} \varepsilon_{12} \\ \sigma_{23} = C_{14} \varepsilon_{11} + C_{24} \varepsilon_{22} + C_{34} \varepsilon_{33} + 2C_{44} \varepsilon_{23} + 2C_{45} \varepsilon_{13} + 2C_{46} \varepsilon_{12} \\ \sigma_{13} = C_{15} \varepsilon_{11} + C_{25} \varepsilon_{22} + C_{35} \varepsilon_{33} + 2C_{45} \varepsilon_{23} + 2C_{55} \varepsilon_{13} + 2C_{56} \varepsilon_{12} \\ \sigma_{12} = C_{16} \varepsilon_{11} + C_{26} \varepsilon_{22} + C_{36} \varepsilon_{33} + 2C_{46} \varepsilon_{23} + 2C_{56} \varepsilon_{13} + 2C_{66} \varepsilon_{12} \end{array} \right. ,$$

where, by definition (1.4.6), $\varepsilon_{ij} = (\partial u_i / \partial x_j + \partial u_j / \partial x_i) / 2$. The first equation is equation of continuity (2.8.2), which results from the conservation of mass. The following three equations are Cauchy's equations of motion (2.8.3), (2.8.4) and (2.8.5), which result from the balance of linear momentum. The last six equations are stress-strain equations (4.2.8), which contain the elasticity parameters that describe a given continuum.

Note that, invoking system (4.4.5) 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 5.6.3, 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 (4.4.5) 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 3.2. 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 (4.2.8). Furthermore, these elastic continua possess strain energy, which is expressed as a quadratic function of strain, shown in expression (4.1.3). 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 (4.4.5) 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.¹²

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.¹³

A given elastic continuum can possess particular symmetries, which further reduce the number of independent elasticity parameters required to describe it. Such symmetries are discussed in Chapter 5.



¹²Interested 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.

¹³Readers 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.

4.5. 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

$$(4.5.1) \quad W = \int \mathbf{F} \cdot d\mathbf{x},$$

where \mathbf{F} denotes force and $d\mathbf{x}$ 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,

$$W = \int_0^{\Delta x} kx \, dx = \frac{1}{2}k(\Delta x)^2,$$

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

EXERCISE 4.2. ¹⁴*Consider the strain-energy function to be a second-degree polynomial given by*

$$(4.5.2) \quad W = C_0 + \sum_{n=1}^6 C_n \varepsilon_n + \frac{1}{2} \sum_{n=1}^6 \sum_{m=1}^6 C_{nm} \varepsilon_n \varepsilon_m,$$

where ε_l is an entry of matrix $\underline{\varepsilon}$, given in equation (3.2.7). 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, $i = 1, 2, \dots, 6$, to represent all components.

SOLUTION 4.2. Expression (4.5.2) can be explicitly written as

$$\begin{aligned} W = & C_0 \\ & + C_1 \varepsilon_1 + C_2 \varepsilon_2 + C_3 \varepsilon_3 + C_4 \varepsilon_4 + C_5 \varepsilon_5 + C_6 \varepsilon_6 \\ & + \frac{1}{2} (C_{11} \varepsilon_1^2 + C_{21} \varepsilon_2 \varepsilon_1 + C_{31} \varepsilon_3 \varepsilon_1 + C_{41} \varepsilon_4 \varepsilon_1 + C_{51} \varepsilon_5 \varepsilon_1 + C_{61} \varepsilon_6 \varepsilon_1 \\ & + C_{12} \varepsilon_1 \varepsilon_2 + C_{22} \varepsilon_2^2 + C_{32} \varepsilon_3 \varepsilon_2 + C_{42} \varepsilon_4 \varepsilon_2 + C_{52} \varepsilon_5 \varepsilon_2 + C_{62} \varepsilon_6 \varepsilon_2 \\ & + C_{13} \varepsilon_1 \varepsilon_3 + C_{23} \varepsilon_2 \varepsilon_3 + C_{33} \varepsilon_3^2 + C_{43} \varepsilon_4 \varepsilon_3 + C_{53} \varepsilon_5 \varepsilon_3 + C_{63} \varepsilon_6 \varepsilon_3 \\ & + C_{14} \varepsilon_1 \varepsilon_4 + C_{24} \varepsilon_2 \varepsilon_4 + C_{34} \varepsilon_3 \varepsilon_4 + C_{44} \varepsilon_4^2 + C_{54} \varepsilon_5 \varepsilon_4 + C_{64} \varepsilon_6 \varepsilon_4 \\ & + C_{15} \varepsilon_1 \varepsilon_5 + C_{25} \varepsilon_2 \varepsilon_5 + C_{35} \varepsilon_3 \varepsilon_5 + C_{45} \varepsilon_4 \varepsilon_5 + C_{55} \varepsilon_5^2 + C_{65} \varepsilon_6 \varepsilon_5 \\ & + C_{16} \varepsilon_1 \varepsilon_6 + C_{26} \varepsilon_2 \varepsilon_6 + C_{36} \varepsilon_3 \varepsilon_6 + C_{46} \varepsilon_4 \varepsilon_6 + C_{56} \varepsilon_5 \varepsilon_6 + C_{66} \varepsilon_6^2). \end{aligned}$$

¹⁴See also Section 4.1.

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 (4.1.1) we require that $\sigma_m = \partial W / \partial \varepsilon_m$. Consider, for instance, $m = 5$; we can specifically write

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

No strain implies no stress and, hence, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = \varepsilon_5 = \varepsilon_6 = 0 \implies \sigma_5 = 0$. Thus, it follows that $C_5 = 0$. Analogously, considering $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ and σ_6 , we obtain $C_1 = C_2 = C_3 = C_4 = C_6 = 0$. Thus,

$$\begin{aligned}W(\varepsilon_m) &= \frac{1}{2}(C_{11}\varepsilon_1^2 + C_{21}\varepsilon_2\varepsilon_1 + C_{31}\varepsilon_3\varepsilon_1 + C_{41}\varepsilon_4\varepsilon_1 + C_{51}\varepsilon_5\varepsilon_1 + C_{61}\varepsilon_6\varepsilon_1 \\ &\quad + C_{12}\varepsilon_1\varepsilon_2 + C_{22}\varepsilon_2^2 + C_{32}\varepsilon_3\varepsilon_2 + C_{42}\varepsilon_4\varepsilon_2 + C_{52}\varepsilon_5\varepsilon_2 + C_{62}\varepsilon_6\varepsilon_2 \\ &\quad + C_{13}\varepsilon_1\varepsilon_3 + C_{23}\varepsilon_2\varepsilon_3 + C_{33}\varepsilon_3^2 + C_{43}\varepsilon_4\varepsilon_3 + C_{53}\varepsilon_5\varepsilon_3 + C_{63}\varepsilon_6\varepsilon_3 \\ &\quad + C_{14}\varepsilon_1\varepsilon_4 + C_{24}\varepsilon_2\varepsilon_4 + C_{34}\varepsilon_3\varepsilon_4 + C_{44}\varepsilon_4^2 + C_{54}\varepsilon_5\varepsilon_4 + C_{64}\varepsilon_6\varepsilon_4 \\ &\quad + C_{15}\varepsilon_1\varepsilon_5 + C_{25}\varepsilon_2\varepsilon_5 + C_{35}\varepsilon_3\varepsilon_5 + C_{45}\varepsilon_4\varepsilon_5 + C_{55}\varepsilon_5^2 + C_{65}\varepsilon_6\varepsilon_5 \\ &\quad + C_{16}\varepsilon_1\varepsilon_6 + C_{26}\varepsilon_2\varepsilon_6 + C_{36}\varepsilon_3\varepsilon_6 + C_{46}\varepsilon_4\varepsilon_6 + C_{56}\varepsilon_5\varepsilon_6 + C_{66}\varepsilon_6^2),\end{aligned}\tag{4.5.3}$$

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

REMARK 4.5.2. Examining expression (4.5.3), we observe that

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

where c is a real number. Hence, in view of Definition A.1.1 stated in Appendix A, W is homogeneous of degree 2 in the ε_m .

EXERCISE 4.3. ¹⁵Using the property of the strain-energy function, W , stated in expression (4.1.1), and in view of W being homogeneous of degree 2 in the ε_{ij} , show that

$$W = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \varepsilon_{ij}.\tag{4.5.4}$$

SOLUTION 4.3. Since W is homogeneous of degree 2 in the ε_{ij} , by Theorem A.2.1 stated in Appendix A, we can write

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

¹⁵See also Section 4.1.

In view of expression (4.1.1), we can rewrite the above expression as

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

which immediately yields expression (4.5.4), as required.

EXERCISE 4.4. Assuming that $c_{ijkl} = c_{klij}$, derive expression (4.1.1), namely,

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

directly from expression (4.5.4).

SOLUTION 4.4. Differentiating expression (4.5.4) with respect to a particular strain-tensor component ε_{kl} , and recalling that stress is a function of strain, we obtain

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

Using stress-strain equations (3.2.1), and recalling that c_{ijkl} are independent of strain, we can write

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

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

$$\frac{\partial W}{\partial \varepsilon_{kl}} = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 (c_{ijkl} \varepsilon_{ij} + \sigma_{ij} \delta_{ik} \delta_{jl}) = \frac{1}{2} \left(\sum_{i=1}^3 \sum_{j=1}^3 c_{klij} \varepsilon_{ij} + \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \delta_{ik} \delta_{jl} \right),$$

where $k, l \in \{1, 2, 3\}$. Again, in view of equations (3.2.1) and (4.2.2) as well as using the properties of Kronecker's delta, we obtain

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

where, in view of the arbitrariness of the subscript symbol, we obtain expression (4.1.1), as required.

EXERCISE 4.5. ¹⁶Using expression (4.3.1), justify that the main-diagonal entries of the elasticity matrix must be always positive.

¹⁶See also Section 4.3.3.

SOLUTION 4.5. Consider expression (4.3.1). In view of equations (4.2.8), 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 write

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

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

¹Materials 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. Subsequently, we study the eight symmetries of an elastic continuum, which are all the possible symmetries of the elasticity tensor.

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

$$(5.1.1) \quad \hat{\mathbf{x}} = \mathbf{A}\mathbf{x},$$

¹This chapter is based on the work that was published by Bos, L.P., Gibson, P.C., Kotchetov, M., and Slawinski, M.A., (2004) Classes of anisotropic media: a tutorial: *Studia geophysica and geodætica*, **48**, pp. 265 – 287.

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 (5.1.1) is the matrix form of equations (1.6.4), 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

$$(5.1.2) \quad \mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

that satisfy the orthogonality condition, namely, $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, which is equivalent to $\mathbf{A}^T = \mathbf{A}^{-1}$, where T denotes transform. This orthogonality condition is discussed in Exercise 5.2.

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 – 5.12. 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 5.1.1. 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.

According to the group theory, 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} . These invariances are the reason for our invoking the notion of a group in Definition 5.1.1.²

5.2. Transformation of coordinates

5.2.1. 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.

²Readers 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 (5th edition)*: Harcourt/Academic Press, pp. 237 – 301.

5.2.2. Transformation of stress-tensor components. The components of the stress tensor expressed as a 3×3 matrix transform according to

$$(5.2.1) \quad \hat{\sigma} = \mathbf{A} \sigma \mathbf{A}^T,$$

where \mathbf{A} stands for matrix (5.1.2) and the accent symbolizes the transformed entity. This is the matrix form of transformation (2.9.5), which is proved in Exercise 2.4.

Following matrix (2.5.2), the stress-tensor components can be written as a square matrix given by

$$(5.2.2) \quad \sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix},$$

which, herein — in view of Theorem 2.7.1 — is symmetric; $\hat{\sigma}$ looks analogous to σ .

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.

In this chapter, we would like to use the elasticity matrix, which is given in expression (4.2.7). The corresponding stress-strain equations are given in expression (4.2.8) and involve stress-tensor components as a single-column matrix, $\underline{\sigma}$, namely,

$$(5.2.3) \quad \underline{\sigma} = [\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{23}, \sigma_{13}, \sigma_{12}]^T.$$

Thus, we wish to obtain a transformation equation for $\underline{\sigma}$ that is equivalent to equation (5.2.1). Since transformation (5.2.1) is linear, it can be rewritten as a multiplication of $\underline{\sigma}$ by a matrix. In other words, we can write

$$(5.2.4) \quad \underline{\hat{\sigma}} = \underline{\mathbf{A}} \underline{\sigma},$$

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 .³

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. In other words, we write

$$(5.2.5) \quad \sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

³Readers 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.

Now, using equation (5.2.1) and matrix (5.1.2), we can compute the corresponding entries of $\hat{\sigma}$ as

$$\begin{aligned}
 \hat{\sigma} &= \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}^T \\
 (5.2.6) \quad &= \begin{bmatrix} A_{11}A_{11} & A_{11}A_{21} & A_{11}A_{31} \\ A_{11}A_{21} & A_{21}A_{21} & A_{21}A_{31} \\ A_{11}A_{31} & A_{21}A_{31} & A_{31}A_{31} \end{bmatrix}.
 \end{aligned}$$

For conciseness, using Kronecker's delta, we can begin by writing the entries of matrix (5.2.5) as

$$\sigma_{ij} = \delta_{i1}\delta_{j1}, \quad i, j \in \{1, 2, 3\},$$

and, hence, write relation (5.2.6) as

$$\begin{aligned}
 \hat{\sigma}_{kl} &= \sum_{i=1}^3 \sum_{j=1}^3 A_{ki} \sigma_{ij} A_{jl}^T = \sum_{i=1}^3 \sum_{j=1}^3 A_{ki} \delta_{i1} \delta_{j1} A_{lj} \\
 &= A_{k1} A_{l1}, \quad k, l \in \{1, 2, 3\}.
 \end{aligned}$$

To consider the matrix form of stress-strain equations, in view of expression (5.2.3) and taking $(k, l) = (1, 1), (2, 2), (3, 3), (2, 3), (1, 3), (1, 2)$, we obtain

$$(5.2.7) \quad \underline{\hat{\sigma}} = \begin{bmatrix} A_{11}A_{11} \\ A_{21}A_{21} \\ A_{31}A_{31} \\ A_{21}A_{31} \\ A_{11}A_{31} \\ A_{11}A_{21} \end{bmatrix}.$$

Column matrix (5.2.7) is the result of unfolding the entries of the rightmost square matrix in expression (5.2.6) in the order given by $(1, 1), (2, 2), (3, 3), (2, 3), (1, 3)$ and $(1, 2)$.

Since, herein, σ is chosen as shown in expression (5.2.5), we write

$$\underline{\sigma} = [1, 0, 0, 0, 0, 0]^T.$$

In view of equation (5.2.4), $\underline{\hat{\sigma}}$ given in expression (5.2.7) 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

$$\sigma = \mathbf{E}_{23} + \mathbf{E}_{32},$$

which is equivalent to

$$(5.2.8) \quad \sigma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Using Kronecker's delta, we can write the entries of matrix (5.2.8), as

$$\sigma_{ij} = \delta_{i2}\delta_{j3} + \delta_{i3}\delta_{j2}, \quad i, j \in \{1, 2, 3\}.$$

Then, using equation (5.2.1), the corresponding entries for $\hat{\sigma}$ can be computed as

$$\begin{aligned} \hat{\sigma}_{kl} &= \sum_{i=1}^3 \sum_{j=1}^3 A_{ki} \sigma_{ij} A_{jl}^T = \sum_{i=1}^3 \sum_{j=1}^3 A_{ki} \delta_{i2}\delta_{j3} A_{lj} + \sum_{i=1}^3 \sum_{j=1}^3 A_{ki} \delta_{i3}\delta_{j2} A_{lj} \\ &= A_{k2}A_{l3} + A_{k3}A_{l2}, \quad k, l \in \{1, 2, 3\}. \end{aligned}$$

To consider the matrix form of stress-strain equations, in view of expression (5.2.3) and taking $(k, l) = (1, 1), (2, 2), (3, 3), (2, 3), (1, 3), (1, 2)$, we obtain

$$\hat{\underline{\sigma}} = \begin{bmatrix} 2A_{12}A_{13} \\ 2A_{22}A_{23} \\ 2A_{32}A_{33} \\ A_{22}A_{33} + A_{23}A_{32} \\ A_{12}A_{33} + A_{13}A_{32} \\ A_{12}A_{23} + A_{13}A_{22} \end{bmatrix},$$

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

$$(5.2.9) \quad \underline{\mathbf{A}} = \begin{bmatrix} A_{11}A_{11} & A_{12}A_{12} & A_{13}A_{13} & 2A_{12}A_{13} & 2A_{11}A_{13} & 2A_{11}A_{12} \\ A_{21}A_{21} & A_{22}A_{22} & A_{23}A_{23} & 2A_{22}A_{23} & 2A_{21}A_{23} & 2A_{21}A_{22} \\ A_{31}A_{31} & A_{32}A_{32} & A_{33}A_{33} & 2A_{32}A_{33} & 2A_{31}A_{33} & 2A_{31}A_{32} \\ A_{21}A_{31} & A_{22}A_{32} & A_{23}A_{33} & A_{22}A_{33} + A_{23}A_{32} & A_{21}A_{33} + A_{23}A_{31} & A_{21}A_{32} + A_{22}A_{31} \\ A_{11}A_{31} & A_{12}A_{32} & A_{13}A_{33} & A_{12}A_{33} + A_{13}A_{32} & A_{11}A_{33} + A_{13}A_{31} & A_{11}A_{32} + A_{12}A_{31} \\ A_{11}A_{21} & A_{12}A_{22} & A_{13}A_{23} & A_{12}A_{23} + A_{13}A_{22} & A_{11}A_{23} + A_{13}A_{21} & A_{11}A_{22} + A_{12}A_{21} \end{bmatrix},$$

which is the desired transformation matrix for the stress-tensor components given by matrix (5.2.3). Thus, given transformation matrix (5.1.2), whose entries are A_{ij} , we can immediately write the corresponding $\underline{\mathbf{A}}$ using matrix (5.2.9). Matrix (5.2.9) was also formulated by Bond (1943).

5.2.3. Transformation of strain-tensor components. The components of the strain-tensor expressed as a 3×3 matrix transform according to

$$(5.2.10) \quad \hat{\varepsilon} = \mathbf{A}\varepsilon\mathbf{A}^T,$$

where \mathbf{A} stands for matrix (5.1.2) and the accent symbolizes the transformed entity. Equation (5.2.10) is the matrix form of equation (1.6.6), which is shown in Exercise 1.4. Herein, the strain-tensor components are a square matrix given by

$$\varepsilon = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix},$$

whose symmetry results from definition (1.4.6); $\hat{\varepsilon}$ looks analogous to ε .

To consider the matrix form of stress-strain equations, we wish to rewrite the strain-tensor components as a single-column matrix in a manner similar to that shown in Section 5.2.2. As shown in stress-strain equations (4.2.8), the single column matrix, $\underline{\varepsilon}$, is formulated with factors of 2; namely,

$$(5.2.11) \quad \underline{\varepsilon} = [\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, 2\varepsilon_{23}, 2\varepsilon_{13}, 2\varepsilon_{12}]^T.$$

Hence, the corresponding transformation matrix differs from expression (5.2.9).⁴ To account for the factors of 2, we can write

$$(5.2.12) \quad \hat{\underline{\varepsilon}} = \mathbf{F}\underline{\mathbf{A}}\mathbf{F}^{-1}\underline{\varepsilon},$$

where $\underline{\mathbf{A}}$ is matrix (5.2.9) and

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

⁴If the stress-strain equations were written as a vector equation in the six-dimensional space, we would not require two distinct transformation matrices. Interested readers might refer to Chapman, C., (2004) Fundamentals of seismic wave propagation: Cambridge University Press, pp. 92 – 94, and to the footnote on page 102.

Thus, the transformation matrix for the strain-tensor components, given by matrix (5.2.11), can be explicitly written as

$$(5.2.13) \quad \mathbf{M}_A = \mathbf{F} \mathbf{A} \mathbf{F}^{-1} = \begin{bmatrix} A_{11}A_{11} & A_{12}A_{12} & A_{13}A_{13} & A_{12}A_{13} & A_{11}A_{13} & A_{11}A_{12} \\ A_{21}A_{21} & A_{22}A_{22} & A_{23}A_{23} & A_{22}A_{23} & A_{21}A_{23} & A_{21}A_{22} \\ A_{31}A_{31} & A_{32}A_{32} & A_{33}A_{33} & A_{32}A_{33} & A_{31}A_{33} & A_{31}A_{32} \\ 2A_{21}A_{31} & 2A_{22}A_{32} & 2A_{23}A_{33} & A_{22}A_{33} + A_{23}A_{32} & A_{21}A_{33} + A_{23}A_{31} & A_{21}A_{32} + A_{22}A_{31} \\ 2A_{11}A_{31} & 2A_{12}A_{32} & 2A_{13}A_{33} & A_{12}A_{33} + A_{13}A_{32} & A_{11}A_{33} + A_{13}A_{31} & A_{11}A_{32} + A_{12}A_{31} \\ 2A_{11}A_{21} & 2A_{12}A_{22} & 2A_{13}A_{23} & A_{12}A_{23} + A_{13}A_{22} & A_{11}A_{23} + A_{13}A_{21} & A_{11}A_{22} + A_{12}A_{21} \end{bmatrix}$$

and expression (5.2.12) can be restated as

$$(5.2.14) \quad \hat{\underline{\underline{\varepsilon}}} = \mathbf{M}_A \underline{\underline{\varepsilon}}.$$

Consequently, given transformation matrix (5.1.2), whose entries are A_{ij} , we can immediately write the corresponding \mathbf{M}_A using matrix (5.2.13).

5.2.4. Stress-strain equations in transformed coordinates. Now, having formulated $\hat{\underline{\underline{\sigma}}}$ and $\hat{\underline{\underline{\varepsilon}}}$, which are given by expressions (5.2.4) and (5.2.14), respectively, we can formally write the stress-strain equations in transformed coordinates as

$$\hat{\underline{\underline{\sigma}}} = \hat{\mathbf{C}} \hat{\underline{\underline{\varepsilon}}}.$$

Explicitly, we can write these equations as

$$(5.2.15) \quad \begin{bmatrix} \hat{\sigma}_{11} \\ \hat{\sigma}_{22} \\ \hat{\sigma}_{33} \\ \hat{\sigma}_{23} \\ \hat{\sigma}_{13} \\ \hat{\sigma}_{12} \end{bmatrix} = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & \hat{C}_{14} & \hat{C}_{15} & \hat{C}_{16} \\ \hat{C}_{12} & \hat{C}_{22} & \hat{C}_{23} & \hat{C}_{24} & \hat{C}_{25} & \hat{C}_{26} \\ \hat{C}_{13} & \hat{C}_{23} & \hat{C}_{33} & \hat{C}_{34} & \hat{C}_{35} & \hat{C}_{36} \\ \hat{C}_{14} & \hat{C}_{24} & \hat{C}_{34} & \hat{C}_{44} & \hat{C}_{45} & \hat{C}_{46} \\ \hat{C}_{15} & \hat{C}_{25} & \hat{C}_{35} & \hat{C}_{45} & \hat{C}_{55} & \hat{C}_{56} \\ \hat{C}_{16} & \hat{C}_{26} & \hat{C}_{36} & \hat{C}_{46} & \hat{C}_{56} & \hat{C}_{66} \end{bmatrix} \begin{bmatrix} \hat{\varepsilon}_{11} \\ \hat{\varepsilon}_{22} \\ \hat{\varepsilon}_{33} \\ 2\hat{\varepsilon}_{23} \\ 2\hat{\varepsilon}_{13} \\ 2\hat{\varepsilon}_{12} \end{bmatrix},$$

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 (4.2.8) and (5.2.15) leads to the following definition.

DEFINITION 5.2.1. 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 5.2.1, 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 5.3.1. *The elastic properties of a continuum are invariant under an orthogonal transformation, given by matrix \mathbf{A} , if and only if*

$$(5.3.1) \quad \mathbf{C} = \mathbf{M}_{\mathbf{A}}^T \mathbf{C} \mathbf{M}_{\mathbf{A}},$$

where \mathbf{C} is the elasticity matrix and $\mathbf{M}_{\mathbf{A}}$ is matrix (5.2.13).

PROOF. Consider stress-strain equations (3.2.8); namely,

$$(5.3.2) \quad \underline{\sigma} = \mathbf{C} \underline{\varepsilon},$$

which are expressed in terms of the original coordinate system. In the transformed coordinate system, these equations are written as

$$(5.3.3) \quad \hat{\underline{\sigma}} = \hat{\mathbf{C}} \hat{\underline{\varepsilon}}.$$

Consider equation (5.3.3). Substituting expressions (5.2.4) and (5.2.14) into equation (5.3.3), 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 5.3.2 shown below, $\underline{\mathbf{A}}^{-1} = \mathbf{M}_{\mathbf{A}}^T$. Hence, we can write

$$(5.3.4) \quad \underline{\sigma} = \mathbf{M}_{\mathbf{A}}^T \hat{\mathbf{C}} \mathbf{M}_{\mathbf{A}} \underline{\varepsilon}.$$

Examining equations (5.3.2) and (5.3.4), 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 5.2.1, invariance with respect to \mathbf{A} means that

$$\mathbf{C} = \mathbf{M}_{\mathbf{A}}^T \mathbf{C} \mathbf{M}_{\mathbf{A}},$$

which is expression (5.3.1), as required. □

LEMMA 5.3.2. *Let $\underline{\mathbf{A}}$ be given by matrix (5.2.9) and $\mathbf{M}_{\mathbf{A}}$ be given by matrix (5.2.13). It follows that $\underline{\mathbf{A}}^{-1} = \mathbf{M}_{\mathbf{A}}^T$.*

PROOF. Recall that \mathbf{A} is an orthogonal matrix. Let $\hat{\mathbf{x}} = \mathbf{A}\mathbf{x}$ be the transformed coordinate system. Consider expression $\sigma\mathbf{x}$, which, in view of expression (5.2.1), can be stated in the $\hat{\mathbf{x}}$ -coordinates as $(\mathbf{A}\sigma\mathbf{A}^T)\hat{\mathbf{x}}$. Thus, in terms of the $\hat{\mathbf{x}}$ -coordinates, the stress-tensor components become

$$(5.3.5) \quad \hat{\sigma} = \mathbf{A}\sigma\mathbf{A}^T.$$

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 (5.3.5) as

$$(5.3.6) \quad \sigma = \mathbf{A}^T \hat{\sigma} \mathbf{A}.$$

Thus, in a manner analogous to that used to obtain expression (5.2.4), we can rewrite expression (5.3.6) in the desired notation, as

$$(5.3.7) \quad \underline{\sigma} = \underline{\mathbf{A}}^T \hat{\sigma},$$

where $\underline{\mathbf{A}}^T$ is constructed as matrix (5.2.9), 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 (5.2.4) with expression (5.3.7), we see that $\underline{\mathbf{A}}^T = \underline{\mathbf{A}}^{-1}$. Hence, we can write the inverse of matrix $\underline{\mathbf{A}}$ explicitly, as

$$(5.3.8) \quad \underline{\mathbf{A}}^{-1} = \begin{bmatrix} A_{11}A_{11} & A_{21}A_{21} & A_{31}A_{31} & 2A_{21}A_{31} & 2A_{11}A_{31} & 2A_{11}A_{21} \\ A_{12}A_{12} & A_{22}A_{22} & A_{32}A_{32} & 2A_{22}A_{32} & 2A_{12}A_{32} & 2A_{12}A_{22} \\ A_{13}A_{13} & A_{23}A_{23} & A_{33}A_{33} & 2A_{23}A_{33} & 2A_{13}A_{33} & 2A_{13}A_{23} \\ A_{12}A_{13} & A_{22}A_{23} & A_{32}A_{33} & A_{22}A_{33} + A_{32}A_{23} & A_{12}A_{33} + A_{23}A_{31} & A_{21}A_{32} + A_{22}A_{31} \\ A_{11}A_{13} & A_{21}A_{23} & A_{31}A_{33} & A_{21}A_{33} + A_{31}A_{23} & A_{11}A_{33} + A_{31}A_{13} & A_{11}A_{23} + A_{21}A_{13} \\ A_{11}A_{12} & A_{21}A_{22} & A_{31}A_{32} & A_{21}A_{32} + A_{31}A_{22} & A_{11}A_{32} + A_{31}A_{12} & A_{11}A_{22} + A_{21}A_{12} \end{bmatrix}.$$

Comparing the entries of matrices (5.3.8) and (5.2.13), we notice that the former one is equal to the transpose of the latter, as required. \square

Expression (5.3.1) 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 (4.2.8) to be invariant under an orthogonal transformation. Given transformation (5.1.2), expression (5.3.1) is convenient to apply since it contains twenty-one linear equations for C_{mn} . Furthermore, considering transformations (5.8.1) and (5.6.1), which are discussed below, matrix $\mathbf{M}_{\mathbf{A}}$ is significantly simplified, since $A_{13} = A_{23} = A_{31} = A_{32} = 0$, while $A_{33} = \pm 1$.

In our formulation of the material-symmetry condition, we could have avoided different transformation matrices for stress and strain, which are given in expressions (5.2.9) and (5.2.13), respectively, if we used

the stress-strain equations formulated in Exercise 5.3, below. Such an approach would have resulted in a different form of matrix \mathbf{M} in expression (5.3.1).⁵ Our choice of equations (4.2.8) was motivated by the simplicity of the elasticity matrix therein, which does not have factors $\sqrt{2}$ and 2, which appear in elasticity matrix (5.14.6) in Exercise 5.3.

5.4. Point symmetry

Let us illustrate condition (5.3.1) by describing the material symmetry that is valid for all continua described by stress-strain equations (4.2.8). 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

$$(5.4.1) \quad \mathbf{A}_{-\mathbf{I}} := \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -\mathbf{I}.$$

THEOREM 5.4.1. *At every point, a continuum given by stress-strain equations (4.2.8) is invariant under the reflection about the origin of a coordinate system that is located at that point.*

PROOF. Consider transformation matrix (5.4.1). Matrix (5.2.13) 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 (5.3.1) 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 $\mathbf{A}_{-\mathbf{I}}$.

⁵Readers interested in the material-symmetry condition in the context of stress-strain equations discussed in Exercise 5.3 might refer to Chapman, C., (2004) Fundamentals of seismic wave propagation: Cambridge University Press, pp. 92 – 94.

5.5. Generally anisotropic continuum

A generally anisotropic continuum is the most general continuum describable by stress-strain equations (4.2.8). The elasticity matrix of a generally anisotropic continuum is given by

$$(5.5.1) \quad \mathbf{C}_{\text{GEN}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}.$$

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 (5.1.2) in the form given by

$$(5.6.1) \quad \mathbf{A} = \begin{bmatrix} \cos \Theta & \sin \Theta & 0 \\ -\sin \Theta & \cos \Theta & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Matrix (5.6.1), 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

$$(5.6.2) \quad \mathbf{A}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Following expression (5.2.13), 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 5.3.1 requires that the elasticity matrix satisfies condition (5.3.1). 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

$$(5.6.3) \quad C_{14} = C_{15} = C_{24} = C_{25} = C_{34} = C_{35} = C_{46} = C_{56} = 0.$$

Thus, the elasticity matrix of a continuum that possesses a reflection symmetry along the x_3 -axis is

$$(5.6.4) \quad \mathbf{C}_{\text{MONO}x_3} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & C_{16} \\ C_{12} & C_{22} & C_{23} & 0 & 0 & C_{26} \\ C_{13} & C_{23} & C_{33} & 0 & 0 & C_{36} \\ 0 & 0 & 0 & C_{44} & C_{45} & 0 \\ 0 & 0 & 0 & C_{45} & C_{55} & 0 \\ C_{16} & C_{26} & C_{36} & 0 & 0 & C_{66} \end{bmatrix}.$$

Hence, a monoclinic continuum is described by an elasticity matrix that contains thirteen independent entries.

5.6.2. Vanishing of tensor components. We can recognize the pattern of the vanishing elasticity parameters in Section 5.6.1 by considering the elasticity tensor rather than the elasticity matrix. Recalling expression (3.2.5), we can write expression (5.6.3) as

$$(5.6.5) \quad c_{1123} = c_{1113} = c_{2223} = c_{2213} = c_{3323} = c_{3313} = c_{2312} = c_{1312} = 0.$$

Examining this result, we see that all the components that disappear as a consequence of the invariance to the reflection along the x_3 -axis have an odd number of subscript 3.

We can exemplify this pattern in view of tensor algebra. In general, the elasticity tensor transforms according to the rule given by

$$\hat{c}_{mnp r} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 A_{mi} A_{nj} A_{pk} A_{rl} c_{ijkl}, \quad m, n, p, r \in \{1, 2, 3\},$$

where \mathbf{A} is a transformation matrix and \hat{c}_{mnpr} denotes the components of the elasticity tensor in the transformed coordinates. In our case, due to the invariance resulting from the symmetry of the elasticity tensor, $\hat{c}_{mnpr} = c_{mnpr}$, as stated in Definition 5.2.1, we write

$$c_{mnpr} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 A_{mi} A_{nj} A_{pk} A_{rl} c_{ijkl}, \quad m, n, p, r \in \{1, 2, 3\},$$

which is the tensor form of condition (5.3.1). To illustrate the vanishing of a component, let us consider, for instance, c_{1113} . We can write

$$c_{1113} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 A_{1i} A_{1j} A_{1k} A_{3l} c_{ijkl}.$$

Since matrix (5.6.1) is a diagonal matrix, let us write only the nonzero entries of \mathbf{A} to get

$$c_{1113} = A_{11} A_{11} A_{11} A_{33} c_{1113}.$$

Specifically, we obtain

$$c_{1113} = (1)(1)(1)(-1)c_{1113} = -c_{1113}.$$

This equality requires that $c_{1113} = 0$, as expected.

5.6.3. 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.⁶ Hence, in principle, we could also rotate our orthonormal coordinate 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 (5.6.4) 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 (5.6.4) to a new matrix that contains only twelve parameters. This orientation of the coordinate system is a natural coordinate system for a monoclinic continuum.

⁶Interested readers might refer to Fedorov, F.I., (1968) Theory of elastic waves in 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.

Rotation of the coordinate axes about the x_3 -axis by angle Θ , where the angle is given by

$$(5.6.6) \quad \tan(2\Theta) = \frac{2C_{45}}{C_{44} - C_{55}},$$

leads to a new set of elasticity parameters, which we denote by \hat{C}_{mn} .⁷ In the new set, \hat{C}_{45} vanishes and elasticity matrix (5.6.4) is reduced to

$$(5.6.7) \quad \hat{\mathbf{C}}_{\text{MONO}} = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & 0 & 0 & \hat{C}_{16} \\ \hat{C}_{12} & \hat{C}_{22} & \hat{C}_{23} & 0 & 0 & \hat{C}_{26} \\ \hat{C}_{13} & \hat{C}_{23} & \hat{C}_{33} & 0 & 0 & \hat{C}_{36} \\ 0 & 0 & 0 & \hat{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{C}_{55} & 0 \\ \hat{C}_{16} & \hat{C}_{26} & \hat{C}_{36} & 0 & 0 & \hat{C}_{66} \end{bmatrix}.$$

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.

We can obtain equation (5.6.6) by using the rotation in the x_1x_2 -plane to diagonalize the submatrix of interest in matrix (5.6.4); namely,

$$(5.6.8) \quad \begin{bmatrix} \hat{C}_{44} & 0 \\ 0 & \hat{C}_{55} \end{bmatrix} = \begin{bmatrix} \cos \Theta & \sin \Theta \\ -\sin \Theta & \cos \Theta \end{bmatrix} \begin{bmatrix} C_{44} & C_{45} \\ C_{45} & C_{55} \end{bmatrix} \begin{bmatrix} \cos \Theta & -\sin \Theta \\ \sin \Theta & \cos \Theta \end{bmatrix}.$$

We note that considering matrices (5.6.4) and (5.6.7), $C_{mn} \neq \hat{C}_{mn}$, where $m, n \in \{1, \dots, 6\}$. In other words, the rotation about the x_3 -axis by the angle (5.6.6) results in a new elasticity matrix to describe the same continuum.

Let us comment on the natural coordinate systems in view of our subsequent work. The orthotropic continuum discussed in Section 5.7 possesses three orthogonal symmetry planes; hence, in accordance with our use of the orthonormal coordinate system, we will consider it, *ab initio*, in its natural coordinate system. While discussing the trigonal continuum, in Section 5.8, we will again search for its natural coordinate system. The tetragonal continuum, discussed in Section 5.9, will be obtained as a special case of the orthotropic one; hence, it will be already in its natural coordinate system. For a transversely isotropic continuum, which we will discuss in Section 5.10, due to its property of rotation invariance, all orthonormal coordinate systems whose one axis coincides with the rotation axis are natural coordinate systems. Hence, our initial formulation will be already set in a natural coordinate system. The cubic continuum, discussed in Section 5.11, can be viewed as a special case of the tetragonal continuum; hence, it will be already in its natural coordinate system. Finally, for an isotropic continuum, discussed in Section 5.12, all orthonormal coordinate systems are natural.

⁷Readers interested in the formulation of expression (5.6.6), 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.

In the context of ray theory, natural coordinate systems are associated with pure-mode directions, as discussed in Section 9.2.2. In particular, in Section 9.2.2, expression (5.6.6) is obtained by considering the displacement directions of the three types of waves that propagate along the x_3 -axis in a monoclinic continuum.

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

$$(5.7.1) \quad \mathbf{A}_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$(5.7.2) \quad \mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and \mathbf{A}_3 , given by matrix (5.6.2), 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 5.7.1. *If a continuum given by stress-strain equations (4.2.8) 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 (5.7.1) and (5.6.2), respectively. Also, following Theorem 5.4.1, all continua possess point symmetry. In other words, they are invariant under the transformation given by matrix (5.4.1). Since all these transformations belong to the symmetry group of an orthotropic continuum, their products also belong to this group. Consider

$$(\mathbf{A}_1)(\mathbf{A}_3)(\mathbf{A}_{-I}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which we recognize to be matrix (5.7.2) 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. \square

Therefore, to obtain the elasticity matrix of an orthotropic continuum, let us use matrices (5.6.2) and (5.7.2). Matrix (5.6.2) is also used in Section 5.6 where we obtain the relations given in expression (5.6.3); namely,

$$(5.7.3) \quad C_{14} = C_{15} = C_{24} = C_{25} = C_{34} = C_{35} = C_{46} = C_{56} = 0.$$

Using matrix (5.7.2), we write condition (5.3.1) as

$$(5.7.4) \quad \mathbf{C} = \mathbf{M}_{A_2}^T \mathbf{C} \mathbf{M}_{A_2},$$

where, using matrix (5.2.13), we explicitly write matrix \mathbf{M}_{A_2} as

$$(5.7.5) \quad \mathbf{M}_{A_2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Let us also use — in this condition — \mathbf{C} given by $\mathbf{C}_{\text{MONO}x_3}$, which is the matrix stated in expression (5.6.4). Thus, solving equation (5.7.4), we get the vanishing entries of \mathbf{C} that distinguish the monoclinic continuum from the orthotropic continuum; they are

$$(5.7.6) \quad C_{16} = C_{26} = C_{36} = C_{45} = 0.$$

Combining relations (5.7.3) and (5.7.6), we can write the elasticity matrix for an orthotropic continuum as

$$(5.7.7) \quad \mathbf{C}_{\text{ORTHO}x_1x_2x_3} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}.$$

Hence, in a natural coordinate system, an orthotropic continuum is described by nine independent elasticity parameters.

To conclude, in view of Section 5.6.2, we can rewrite expression (5.7.6) as

$$(5.7.8) \quad c_{1112} = c_{2221} = c_{3312} = c_{2313} = 0.$$

Examining this result, we see that all the elasticity-tensor components that disappear as a result of the invariance under the reflection along the x_2 -axis have an odd number of subscript 2, as expected. Combining expressions (5.6.5) and (5.7.8), which — in view of Theorem 5.7.1 — correspond to the elasticity tensor that is invariant under reflections along the x_1 -axis, the x_2 -axis and the x_3 -axis, we see that each vanishing component has two occurrences of an odd number of subscript 1, 2 or 3. This is consistent with Theorem 5.7.1, which implies that each component vanishes independently twice; in other words, it vanishes for reflections along two, among the three axes. Also, while examining expressions (5.6.5) and (5.7.8), we can write the components that vanish as a result of the invariance under the reflection along the x_1 -axis; they are: c_{1113} , c_{2213} , c_{3313} , c_{2312} , c_{1112} , c_{2221} , c_{3312} and c_{2313} .

5.8. Trigonal continuum

5.8.1. Elasticity matrix. A trigonal continuum is a continuum whose symmetry group contains rotations about an axis by Θ , where $\Theta = 2\pi/3$ and $\Theta = 4\pi/3$. To obtain the elasticity matrix for this continuum, we consider the orthogonal transformation that is represented by matrix (5.1.2) in the form given by

$$(5.8.1) \quad \mathbf{A}_{x_3\Theta} = \begin{bmatrix} \cos \Theta & \sin \Theta & 0 \\ -\sin \Theta & \cos \Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrix (5.8.1), whose determinant is equal to unity, corresponds to rotation by angle Θ about the x_3 -axis. Therefore, considering $\Theta = 2\pi/3$, we obtain

$$\mathbf{A}_{x_3 2\pi/3} = \begin{bmatrix} \cos \frac{2\pi}{3} & \sin \frac{2\pi}{3} & 0 \\ -\sin \frac{2\pi}{3} & \cos \frac{2\pi}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Following expression (5.2.13), we write the corresponding 6×6 transformation matrix, which is

$$(5.8.2) \quad \mathbf{M}_{\mathbf{A}_{x_3 2\pi/3}} = \begin{bmatrix} \cos^2 \frac{2\pi}{3} & \sin^2 \frac{2\pi}{3} & 0 & 0 & 0 & \cos \frac{2\pi}{3} \sin \frac{2\pi}{3} \\ \sin^2 \frac{2\pi}{3} & \cos^2 \frac{2\pi}{3} & 0 & 0 & 0 & -\sin \frac{2\pi}{3} \cos \frac{2\pi}{3} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos \frac{2\pi}{3} & -\sin \frac{2\pi}{3} & 0 \\ 0 & 0 & 0 & \sin \frac{2\pi}{3} & \cos \frac{2\pi}{3} & 0 \\ -2 \cos \frac{2\pi}{3} \sin \frac{2\pi}{3} & 2 \sin \frac{2\pi}{3} \cos \frac{2\pi}{3} & 0 & 0 & 0 & \cos^2 \frac{2\pi}{3} - \sin^2 \frac{2\pi}{3} \end{bmatrix}.$$

Using matrix (5.8.2) in condition (5.3.1), we obtain

$$(5.8.3) \quad C_{16} = C_{26} = C_{34} = C_{35} = C_{36} = C_{45} = 0$$

and

$$(5.8.4) \quad C_{11} = C_{22}, C_{13} = C_{23}, C_{14} = -C_{56} = -C_{24},$$

$$C_{66} = \frac{1}{2}(C_{22} - C_{12}), C_{15} = C_{46} = -C_{25}, C_{44} = C_{55}.$$

We do not need to repeat this process for $\Theta = 4\pi/3$ since for invariance under a such rotation we would obtain the results identical to the ones shown in expressions (5.8.3) and (5.8.4). This is so, because $\mathbf{A}_{x_3 4\pi/3}$ is a morphism of groups $\mathbf{A}_{x_3 2\pi/3}$ and $\mathbf{A}_{x_3 2\pi/3}$. In other words,

$$\mathbf{M}_{\mathbf{A}_{x_3 4\pi/3}} = \mathbf{M}_{\mathbf{A}_{x_3 2\pi/3}} \mathbf{M}_{\mathbf{A}_{x_3 2\pi/3}},$$

where $\mathbf{M}_{\mathbf{A}_{x_3 2\pi/3}}$ is matrix (5.8.2).

Thus, using expressions (5.8.3) and (5.8.4), we can already write the elasticity matrix for the trigonal continuum as

$$(5.8.5) \quad \mathbf{C}_{\text{TRIGONAL}_{x_3}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & 0 \\ C_{12} & C_{11} & C_{13} & -C_{14} & -C_{15} & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ C_{14} & C_{14} & 0 & C_{44} & 0 & -C_{15} \\ C_{15} & -C_{15} & 0 & 0 & C_{44} & C_{14} \\ 0 & 0 & 0 & -C_{15} & C_{14} & \frac{C_{11}-C_{12}}{2} \end{bmatrix}.$$

Examining matrix (5.8.5), we see that the trigonal continuum is described by an elasticity matrix with seven independent entries.

5.8.2. Natural coordinate system. With an appropriate rotation of the coordinate system about the x_3 -axis, matrix (5.8.5) can be further reduced to a new matrix with only six independent entries. This rotation is given by

$$\tan(3\Theta) = \frac{C_{14}}{C_{15}}.$$

As a result of this rotation, C_{14} vanishes and matrix (5.8.5) becomes

$$\hat{\mathbf{C}}_{\text{TRIGONAL}} = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & 0 & \hat{C}_{15} & 0 \\ \hat{C}_{12} & \hat{C}_{11} & \hat{C}_{13} & 0 & -\hat{C}_{15} & 0 \\ \hat{C}_{13} & \hat{C}_{13} & \hat{C}_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{C}_{44} & 0 & -\hat{C}_{15} \\ \hat{C}_{15} & -\hat{C}_{15} & 0 & 0 & \hat{C}_{44} & 0 \\ 0 & 0 & 0 & -\hat{C}_{15} & 0 & \frac{\hat{C}_{11}-\hat{C}_{12}}{2} \end{bmatrix}.$$

Therefore, in a natural coordinate system, the trigonal continuum is described by six independent elasticity parameters.

5.9. 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.

The transformation matrices of a tetragonal continuum are given by

$$(5.9.1) \quad \mathbf{A}_{x_3\pi/2} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which is matrix (5.8.1) with $\Theta = \pi/2$, and by matrix (5.7.2). These matrices correspond to the rotation about the x_3 -axis and to the reflections along the x_2 -axis, respectively.

Matrix (5.7.2) also belongs to the symmetry group of an orthotropic continuum discussed in Section 5.7, where we obtained the relations given in expression (5.7.6); namely,

$$(5.9.2) \quad C_{16} = C_{26} = C_{36} = C_{45} = 0.$$

These relations also apply to a tetragonal continuum. The additional relations result from matrix (5.9.1). Using matrix (5.9.1), condition (5.3.1) becomes

$$\mathbf{C} = \mathbf{M}_{\mathbf{A}_{x_3\pi/2}}^T \mathbf{C} \mathbf{M}_{\mathbf{A}_{x_3\pi/2}}$$

and results in relations given by

$$(5.9.3) \quad C_{22} = C_{11}, C_{23} = C_{13}, C_{55} = C_{44}.$$

Combining relations (5.9.2) and (5.9.3), we obtain the elasticity matrix of a tetragonal continuum; namely,

$$(5.9.4) \quad \mathbf{C}_{\text{TETRA}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}.$$

Thus, in a natural coordinate system, only six independent elasticity parameters are needed to describe a tetragonal continuum.

Note that, as expected, matrix (5.9.4) is a special case of matrix (5.7.7) with additional relations given by expression (5.9.3).

5.10. Transversely isotropic continuum

5.10.1. 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 (5.8.1) where Θ is smaller than $\pi/2$. Let us consider, for example, $\Theta = 2\pi/5$, and, hence, assume that the symmetry group contains

$$(5.10.1) \quad \mathbf{A}_{x_3 2\pi/5} = \begin{bmatrix} \cos \frac{2\pi}{5} & \sin \frac{2\pi}{5} & 0 \\ -\sin \frac{2\pi}{5} & \cos \frac{2\pi}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Following condition (5.3.1), the elasticity matrix, \mathbf{C} , satisfies the equation given by

$$(5.10.2) \quad \mathbf{C} = \mathbf{M}_{\mathbf{A}_{x_3 2\pi/5}}^T \mathbf{C} \mathbf{M}_{\mathbf{A}_{x_3 2\pi/5}}.$$

The entries of matrix $\mathbf{M}_{\mathbf{A}_{x_3 2\pi/5}}$ are more complicated than the entries of the transformation matrices used in the previous sections, but equation (5.10.2) can still be solved directly to give relations among the entries of \mathbf{C} . The solution to condition (5.10.2) is the matrix given by

$$(5.10.3) \quad \mathbf{C}_{\text{TRANS}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{C_{11}-C_{12}}{2} \end{bmatrix}.$$

Thus, the requirements that the symmetry group contains $\mathbf{A}_{x_3 2\pi/5}$ results in a continuum that is described by only five independent elasticity parameters.

5.10.2. Rotation invariance. A particularly important property of matrix (5.10.3) is the fact that for any angle Θ , this matrix, without any further simplification, satisfies the equation given by

$$(5.10.4) \quad \mathbf{C} = \mathbf{M}_{\mathbf{A}_{x_3 \Theta}}^T \mathbf{C} \mathbf{M}_{\mathbf{A}_{x_3 \Theta}},$$

where $\mathbf{A}_{x_3 \Theta}$ is given by matrix (5.8.1). This property of matrix (5.10.3) can be directly verified by substituting $\mathbf{M}_{\mathbf{A}_{x_3 \Theta}}$, without specifying the value of Θ , and $\mathbf{C}_{\text{TRANS}}$ into the right-hand side of equation (5.10.4). The resulting expression reduces to $\mathbf{C}_{\text{TRANS}}$. Therefore, the invariance of $\mathbf{C}_{\text{TRANS}}$ to the five-fold rotation about a given axis implies invariance to the rotation by any angle about this axis. As stated in Theorem 5.10.1, 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 5.10.1 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 4.2.4.⁸ Since strain energy is a scalar, its value must be the same for all orientations of the coordinate system. In view of expression (4.1.3), namely,

$$(5.10.5) \quad W(\varepsilon) = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{ij} \varepsilon_{kl},$$

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

$$(5.10.6) \quad \hat{W}(\hat{\varepsilon}) = W(\varepsilon),$$

where the transformed strain-tensor components are given by expression (5.2.10), namely,

$$(5.10.7) \quad \hat{\varepsilon} = \mathbf{A} \varepsilon \mathbf{A}^T,$$

which, for brevity, we denote by $\mathbf{A} \circ \varepsilon$, with \circ standing for the orthogonal-transformation operator.

Note that equation (5.10.6) 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 $\mathbf{A}_{x_3\Theta}$, which is given by expression (5.8.1). In view of expressions (5.8.1) and (5.10.7), $\hat{\varepsilon}$ can be regarded as a quadratic trigonometric polynomial in Θ . Hence, $W(\mathbf{A}_{x_3\Theta} \circ \varepsilon)$ is a quartic trigonometric polynomial in Θ .⁹

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

⁸Readers 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 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.

⁹Readers interested in trigonometric polynomials might refer to Courant, R., and Hilbert, D., (1924/1989) Methods of mathematical physics: John Wiley & Sons, Vol. I, pp. 69 – 70.

continuum form a group, we conclude that $W(\varepsilon)$ is also invariant under rotations by $2m\pi/n$, where $m \in \{0, 1, \dots, n-1\}$. In other words, the symmetry group contains rotations given by

$$(5.10.8) \quad \mathbf{A}_{x_3 2m\pi/n} = \begin{bmatrix} \cos \frac{2m\pi}{n} & \sin \frac{2m\pi}{n} & 0 \\ -\sin \frac{2m\pi}{n} & \cos \frac{2m\pi}{n} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{matrix} n \geq 5 \\ m \in \{0, 1, \dots, n-1\} \end{matrix}.$$

So, we can write

$$(5.10.9) \quad W(\varepsilon) = W(\mathbf{A}_{x_3 2m\pi/n} \circ \varepsilon), \quad \begin{matrix} n \geq 5 \\ m \in \{0, 1, \dots, n-1\} \end{matrix},$$

which implies that

$$(5.10.10) \quad W(\varepsilon) = \frac{1}{n} \sum_{m=0}^{n-1} W(\mathbf{A}_{x_3 2m\pi/n} \circ \varepsilon), \quad n \geq 5.$$

In other words, since equation (5.10.9) holds for any allowable value of m , the sum on the right-hand side of equation (5.10.10) is composed of the identical values of W .

Note that for any $W(\varepsilon)$, not necessarily invariant under transformations (5.10.8), the right-hand side of equation (5.10.10) is called the symmetrization of W with respect to the group of these transformations. Hence, equation (5.10.10) means that if $W(\varepsilon)$ is invariant under rotations (5.10.8), 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 $W(\mathbf{A}_{x_3 \Theta} \circ \varepsilon)$ is a quartic trigonometric polynomial, it follows that, for $n \geq 5$, we can apply Lemma 5.10.2, below, and rewrite the right-hand side of equation (5.10.10) as

$$\frac{1}{n} \sum_{m=0}^{n-1} W(\mathbf{A}_{x_3 2m\pi/n} \circ \varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} W(\mathbf{A}_{x_3 \Theta} \circ \varepsilon) d\Theta, \quad n \geq 5,$$

which immediately allows us to write

$$(5.10.11) \quad W(\varepsilon) = \frac{1}{2\pi} \int_0^{2\pi} W(\mathbf{A}_{x_3 \Theta} \circ \varepsilon) d\Theta.$$

Equation (5.10.11) 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 5.10.1. *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 5.10.2. *If f is a trigonometric polynomial of at most degree $n - 1$, then*

$$(5.10.12) \quad \frac{1}{n} \sum_{m=0}^{n-1} f\left(\frac{2m\pi}{n}\right) = \frac{1}{2\pi} \int_0^{2\pi} f(\Theta) d\Theta.$$

PROOF. Consider a basis of the space of trigonometric polynomials of at most degree $n - 1$ given by

$$f_r(\Theta) = e^{ir\Theta}, \quad r \in \{-(n-1), \dots, n-1\}.$$

In view of linearity, to prove equation (5.10.12) for f , it suffices to prove it for f_r , where

$$r \in \{-(n-1), \dots, n-1\}.$$

Set

$$(5.10.13) \quad z = e^{ir\frac{2\pi}{n}}, \quad r \in \{-(n-1), \dots, n-1\}.$$

Then, we can write

$$(5.10.14) \quad \frac{1}{n} \sum_{m=0}^{n-1} f_r\left(\frac{2m\pi}{n}\right) = \frac{1}{n} \sum_{m=0}^{n-1} z^m, \quad r \in \{-(n-1), \dots, n-1\}.$$

Examining expression (5.10.13) for the case of $r = 0$, we note that $z = 1$ and, thus, the right-hand side of expression (5.10.14) is equal to 1. Now, for $r \neq 0$, $z \neq 1$ and we can write the right-hand side of expression (5.10.14) as

$$(5.10.15) \quad \frac{1}{n} \sum_{m=0}^{n-1} z^m = \frac{1}{n} \frac{z^n - 1}{z - 1}.$$

Examining expression (5.10.13), we also note that $z^n = 1$ and, thus, the right-hand side of expression (5.10.15) is equal to 0. To summarize, we can write the left-hand side of equation (5.10.12) as

$$(5.10.16) \quad \frac{1}{n} \sum_{m=0}^{n-1} f_r\left(\frac{2m\pi}{n}\right) = \begin{cases} 0 & \text{if } r \neq 0 \\ 1 & \text{if } r = 0 \end{cases}.$$

Performing the integration on the right-hand side of equation (5.10.12), we obtain

$$(5.10.17) \quad \frac{1}{2\pi} \int_0^{2\pi} e^{ir\Theta} d\Theta = \begin{cases} 0 & \text{if } r \neq 0 \\ 1 & \text{if } r = 0 \end{cases}.$$

Thus, for $r \in \{-(n-1), \dots, n-1\}$, expressions (5.10.16) and (5.10.17) are equal to one another and, hence, equation (5.10.12) is valid for polynomials of at most degree $n - 1$, as required. \square

Note that in the proof of Theorem 5.10.1 we used the fact that $W(\varepsilon)$ is a quadratic polynomial in the strain-tensor components and, hence, $W(\mathbf{A}_{x_3\Theta} \circ \varepsilon)$ is a quartic trigonometric polynomial in Θ . This

corresponds to the fact that Theorem 5.10.1 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 in 1945, and is shown in Exercise 5.6.

Since the symmetry of $W(\varepsilon)$ is tantamount to the symmetry of a continuum, we conclude that a continuum described by matrix (5.10.3) is transversely isotropic.

5.11. Cubic continuum

A cubic continuum is a continuum whose symmetry group contains rotations by $\Theta = \pi/2$ about two axes that are orthogonal to one another. Herein, let us choose the x_3 -axis and x_1 -axis.

To consider the rotation about the x_3 -axis, we set $\Theta = \pi/2$ in matrix (5.8.1) to get

$$\mathbf{A}_{x_3\pi/2} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which is matrix (5.9.1) used already in Section 5.9 to discuss the tetragonal continuum. Following expression (5.2.13), we write the corresponding 6×6 transformation matrix, which is

$$(5.11.1) \quad \mathbf{M}_{\mathbf{A}_{x_3\pi/2}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Using matrix (5.11.1) in condition (5.3.1), we obtain

$$(5.11.2) \quad C_{14} = C_{15} = C_{24} = C_{25} = C_{34} = C_{35} = C_{36} = C_{45} = C_{46} = C_{56} = 0$$

and

$$(5.11.3) \quad C_{11} = C_{22}, C_{13} = C_{23}, C_{16} = -C_{26}, C_{44} = C_{55}.$$

To consider the rotation about the x_1 -axis, we use the transformation matrix given by

$$\mathbf{A}_{x_1\Theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \Theta & -\sin \Theta \\ 0 & \sin \Theta & \cos \Theta \end{bmatrix},$$

which — when evaluated for $\Theta = \pi/2$ — becomes

$$\mathbf{A}_{x_1\pi/2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The corresponding 6×6 transformation matrix is

$$(5.11.4) \quad \mathbf{M}_{\mathbf{A}_{x_1\pi/2}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}.$$

Using matrix (5.11.4) in condition (5.3.1), we obtain

$$(5.11.5) \quad C_{14} = C_{15} = C_{16} = C_{25} = C_{26} = C_{35} = C_{36} = C_{45} = C_{46} = C_{56} = 0$$

and

$$(5.11.6) \quad C_{12} = C_{13}, C_{22} = C_{33}, C_{24} = -C_{34}, C_{55} = C_{66}.$$

Combining expressions (5.11.2), (5.11.3), (5.11.5) and (5.11.6), we obtain the elasticity matrix for the cubic continuum. This matrix is

$$\mathbf{C}_{\text{CUBIC}_{x_1x_3}} = \begin{bmatrix} C_{11} & C_{13} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix},$$

which contains three independent elasticity parameters.

Examining this derivation, we see that cubic symmetry can be viewed as resulting from combining two tetragonal symmetries for the two rotation axes that are orthogonal to one another.

In view of work presented in the sections above, we see that the cubic continuum is a special case of the tetragonal continuum, tetragonal continuum is a special case of the orthotropic continuum, which in turn is a special case of the monoclinic continuum. In other words, the symmetry group of the monoclinic continuum is contained in the symmetry group of the cubic continuum. In Section 5.13, we will study the relations among all symmetry classes.

5.12. Isotropic continuum

5.12.1. 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 (5.10.3). 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 (5.3.1), we obtain the equation given by

$$\mathbf{C} = \mathbf{M}_{A_{x_1x_3}}^T \mathbf{C} \mathbf{M}_{A_{x_1x_3}},$$

which imposes the additional relations; namely,

$$C_{11} = C_{33}, C_{12} = C_{13}, C_{44} = C_{66}.$$

Incorporating these relations into matrix (5.10.3), we obtain

$$(5.12.1) \quad \mathbf{C}_{\text{ISO}} = \begin{bmatrix} C_{11} & C_{11} - 2C_{44} & C_{11} - 2C_{44} & 0 & 0 & 0 \\ C_{11} - 2C_{44} & C_{11} & C_{11} - 2C_{44} & 0 & 0 & 0 \\ C_{11} - 2C_{44} & C_{11} - 2C_{44} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix}.$$

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 5.7, the elasticity matrix of an isotropic continuum is symmetric without invoking the existence of the strain-energy function.

5.12.2. Lamé's parameters. The two independent elasticity parameters that describe an isotropic continuum are often expressed as

$$(5.12.2) \quad \begin{cases} \lambda := C_{11} - 2C_{44} \\ \mu := C_{44} \end{cases}.$$

The two parameters, λ and μ , are called Lamé's parameters. Their physical meaning is described in Section 5.12.4. Mathematically, these two parameters are closely related to the eigenvalues of elasticity

matrix (5.12.1). If we write matrix (5.12.1) in the form discussed in Exercise 5.3, namely,

$$\begin{bmatrix} C_{11} & C_{11} - 2C_{44} & C_{11} - 2C_{44} & 0 & 0 & 0 \\ C_{11} - 2C_{44} & C_{11} & C_{11} - 2C_{44} & 0 & 0 & 0 \\ C_{11} - 2C_{44} & C_{11} - 2C_{44} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2C_{44} \end{bmatrix},$$

its eigenvalues are given by $3C_{11} - 4C_{44}$ and by $2C_{44}$, which is repeated five times. In general, for the anisotropic continua, the eigenvalues of an elasticity matrix play an important role in identifying the symmetry class since — unlike the elasticity parameters — they are invariant under the orthogonal transformations.¹⁰

Using the definition of Lamé's parameters (5.12.2), we can rewrite matrix (5.12.1) as

$$(5.12.3) \quad \mathbf{C}_{\text{LAME}} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}.$$

5.12.3. Tensor formulation. Using matrix (5.12.3) and in view of equations (3.2.1), we can write the stress-strain equations for an isotropic continuum as

$$(5.12.4) \quad \sigma_{ij} = \lambda \delta_{ij} \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{ij}, \quad i, j \in \{1, 2, 3\},$$

where $\sum \varepsilon_{kk}$ is the dilatation defined in expression (1.4.18). This formulation is used to derive the wave equation in Chapter 6.

Since, in equations (3.2.1), 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

$$(5.12.5) \quad a_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \xi \delta_{ik} \delta_{jl} + \eta \delta_{il} \delta_{jk}, \quad i, j, k, l \in \{1, 2, 3\},$$

¹⁰Readers interested in the use of the eigenvalues of the elasticity matrix for identifying the symmetry class might refer to Bóna, A., Bucataru, I., Slawinski, M.A. (2007) Coordinate-free classification of elasticity tensor: *Journal of Elasticity* **87**(2-3), 109–132, and to Rychlewski, J., (1985) On Hooke's law: *Prikl. Matem. Mekhan.*, **48** (3), 420 – 435.

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 (5.12.4), which contains only two constants, λ and μ , where, as shown in Exercise 5.8, $\mu := (\xi + \eta) / 2$.

By examining stress-strain equations (5.12.4) 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 (1.4.6) as

$$(5.12.6) \quad \sigma_{ij} = \lambda \delta_{ij} \sum_{k=1}^3 \frac{\partial u_k}{\partial x_k} + \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j \in \{1, 2, 3\}.$$

Equations (5.12.6) 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 (5.12.4) correspond to isotropic continua.

5.12.4. Physical meaning of Lamé's parameters. We can obtain the physical meaning of Lamé's parameters, λ and μ , by examining stress-strain equations (5.12.4).

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 (5.12.4) as

$$\sigma_{ij} = 2\mu\varepsilon_{ij}, \quad \begin{matrix} i \neq j \\ i, j \in \{1, 2, 3\} \end{matrix},$$

which, using definition (1.4.6), we can rewrite as

$$\sigma_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad \begin{matrix} i \neq j \\ i, j \in \{1, 2, 3\} \end{matrix}.$$

In view of Section 1.4.3, 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 $i = j$, equations (5.12.4) become

$$(5.12.7) \quad \lim_{\mu \rightarrow 0^+} \sigma_{ij} = \lambda \sum_{k=1}^3 \varepsilon_{kk} = \lambda (\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) = \lambda \varphi, \quad \begin{matrix} i = j \\ i, j \in \{1, 2, 3\} \end{matrix},$$

where φ is the dilatation defined in expression (1.4.18). Examination of expression (5.12.7), 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 (5.12.3), 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 (5.12.4) as

$$(5.12.8) \quad \sigma_{ij} = \lambda \tilde{\varepsilon} + \frac{2}{3} \mu \tilde{\varepsilon} = \left(\lambda + \frac{2}{3} \mu \right) \tilde{\varepsilon}, \quad \begin{matrix} i = j \\ i, j \in \{1, 2, 3\} \end{matrix}.$$

In view of expression (1.4.17) 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 (5.12.8) as

$$(5.12.9) \quad -P = \left(\lambda + \frac{2}{3} \mu \right) \frac{\Delta V}{V}.$$

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,

$$(5.12.10) \quad \kappa := -\frac{1}{P} \frac{\Delta V}{V}.$$

Using expression (5.12.9), 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, $\mu \rightarrow 0^+$, 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.

5.13. Relations among symmetry classes

The eight symmetry classes are related to each other. In this section, we will briefly discuss certain aspects of these relations.

The symmetry group of an isotropic continuum contains all orthogonal transformations. Hence it contains, as its subgroups, all other symmetry classes. The symmetry group of the generally anisotropic continuum contains only two orthogonal transformations: the identity and the point symmetry. These two

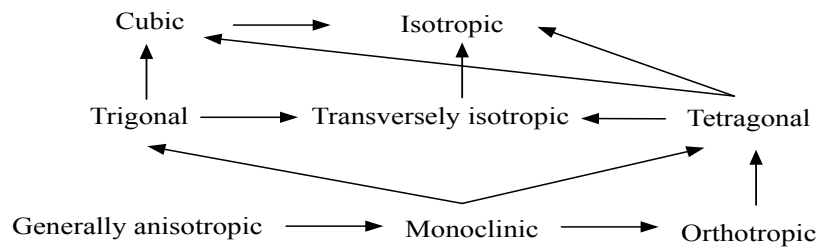


FIGURE 5.13.1. Relations among symmetry classes: The arrows state the subgroup relations; for example, Orthotropic is a subgroup of Tetragonal.

transformations are contained, as a subgroup, in all symmetry classes. However, not all pairs among the classes possess such a subgroup relation; for instance, neither cubic symmetry is a subgroup of transversely isotropic symmetry nor vice versa. The set where not all pairs of elements exhibit the property that one is a subgroup of the other is called partial ordering. The set of the symmetry classes of an elastic continuum, which is shown in Figure 5.13.1, is a partially ordered set. The eight symmetry classes discussed above are related to Curie's symmetry groups, which can be represented geometrically as spheres, cylinders and cones.¹¹

Another important aspect of a given symmetry class is the number of independent elasticity parameters. However, arranging the classes according to this number has its limitations. We can illustrate these limitations with two examples. A transversely isotropic continuum exhibits five elasticity parameters while a cubic one exhibits only three. Yet, as shown above, ordering according to this number cannot be viewed as ordering in terms of the symmetry group. Secondly, both the trigonal and tetragonal continua are described by the same number of independent elasticity parameters, namely six. However, they are distinct continua; their symmetry groups are different from one another and, hence, so are their elasticity matrices.¹²

Closing remarks

In this chapter, by studying the elasticity matrix, we investigated all eight symmetries of the elasticity tensor. In other words, we investigated the elasticity tensors that correspond to generally anisotropic, monoclinic, orthotropic, trigonal, tetragonal, transversely isotropic, cubic and isotropic continua.¹³

¹¹Interested readers might refer to Newnham, R.E., (2005) Properties of materials: Anisotropy, symmetry, structure: Oxford University Press, pp. 26 – 28.

¹²Readers interested in the 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.

¹³Readers interested in the existence of the eight symmetry classes might refer to Bóna, A., Bucataru, I., and Slawinski, M.A., (2005) Characterization of elasticity-tensor symmetries using SU(2). *Journal of Elasticity* **75**(3), 267 – 289, and to Bóna, A., Bucataru, I., and Slawinski, M.A., (2004) Material symmetries of elasticity tensor. *The Quarterly Journal of Mechanics and Applied Mathematics* **57**(4), 583 – 598.

Studying the symmetries of a continuum provides us with information about the material 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 required to describe a given continuum provides us with a convenient way to study 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 9.

The nomenclature commonly used to describe the material symmetries originates in crystallography. However, since we are studying symmetries of continua, intuitive and heuristic descriptions associated with crystal lattices are not always appropriate in this context.



5.14. Exercises

EXERCISE 5.1. ¹⁴Show that the Jacobian that is associated with matrix (5.8.1) is equal to unity.

SOLUTION 5.1. Using matrix (5.8.1), we can write the transformation of coordinates as

$$(5.14.1) \quad \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \begin{bmatrix} \cos \Theta & \sin \Theta & 0 \\ -\sin \Theta & \cos \Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

where \mathbf{x} and $\hat{\mathbf{x}}$ are the original and the transformed coordinates, respectively. The Jacobian is given by

$$(5.14.2) \quad J := \det \begin{bmatrix} \frac{\partial \hat{x}_1}{\partial x_1} & \frac{\partial \hat{x}_1}{\partial x_2} & \frac{\partial \hat{x}_1}{\partial x_3} \\ \frac{\partial \hat{x}_2}{\partial x_1} & \frac{\partial \hat{x}_2}{\partial x_2} & \frac{\partial \hat{x}_2}{\partial x_3} \\ \frac{\partial \hat{x}_3}{\partial x_1} & \frac{\partial \hat{x}_3}{\partial x_2} & \frac{\partial \hat{x}_3}{\partial x_3} \end{bmatrix}.$$

Thus, examining equations (5.14.1) and (5.14.2), we see that the determinant of matrix (5.8.1) 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. ¹⁵Formulate the condition that a transformation matrix must obey in order to preserve the scalar product. Provide a physical interpretation of the result.

SOLUTION 5.2. The scalar product is preserved if, for any two vectors, we can write

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

In other words, the value of the product is the same in the original and transformed coordinate systems. Using expression (5.1.1), we can rewrite this condition as

$$(5.14.3) \quad \mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{u} = \mathbf{v} \cdot \mathbf{u},$$

where \mathbf{A} is a transformation matrix. In terms of components, we write the left-hand side of equation (5.14.3) as

$$\mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{u} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 A_{ij} v_j A_{ik} u_k.$$

¹⁴See also Section 5.1.1.

¹⁵See also Section 5.1.1.

We can change the order of this summation to write

$$\mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{u} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 A_{ij} A_{ik} v_j u_k.$$

To state this summation in terms of matrix multiplication, we write

$$\mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{u} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 (A^T)_{ji} A_{ik} v_j u_k = (\mathbf{A}^T \mathbf{A}) \mathbf{v} \cdot \mathbf{u}.$$

Thus, we can write equation (5.14.3) as

$$\mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{u} = (\mathbf{A}^T \mathbf{A}) \mathbf{v} \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{u}.$$

Since this equality must be valid for any \mathbf{u} , we require that

$$(\mathbf{A}^T \mathbf{A}) \mathbf{v} = \mathbf{v},$$

which implies that $\mathbf{A}^T \mathbf{A} = \mathbf{I}$. This is the condition that a transformation matrix must obey in order to preserve the scalar product; we refer to $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ as the orthogonality condition. Physically, since the distance is expressed in terms of the scalar product, the transformations that obey the orthogonality condition are distance-preserving.

EXERCISE 5.3. Rewrite stress-strain equations (4.2.8) in such a manner that both the matrix containing the stress-tensor components and the matrix containing the strain-tensor components have the same form.

SOLUTION 5.3. Both matrices will have the same form if we multiply $\underline{\sigma}$ and $\underline{\varepsilon}$ by $[\sqrt{2}\mathbf{I}]$ and $[(1/\sqrt{2})\mathbf{I}]$, respectively, where \mathbf{I} is the 6×6 identity matrix. Let us consider equations (4.2.8), which we write as

$$\underline{\sigma} = \mathbf{C}\underline{\varepsilon}.$$

Since we can perform the same operation to both sides of an equation, let us write

$$[\sqrt{2}\mathbf{I}] \underline{\sigma} = [\sqrt{2}\mathbf{I}] \mathbf{C} \underline{\varepsilon}.$$

Since we can multiply either side of this equation by the identity matrix, let us write

$$(5.14.4) \quad [\sqrt{2}\mathbf{I}] \underline{\sigma} = [\sqrt{2}\mathbf{I}] \mathbf{C} \underline{\varepsilon},$$

and let the identity matrix in front of $\underline{\varepsilon}$ have a particular form; namely,

$$\mathbf{I} = [\sqrt{2}\mathbf{I}] [\sqrt{2}\mathbf{I}]^{-1} = [\sqrt{2}\mathbf{I}] \left[\frac{1}{\sqrt{2}} \mathbf{I} \right].$$

Thus, we can write equation (5.14.4) as

$$(5.14.5) \quad [\sqrt{2}\mathbf{I}] \underline{\sigma} = [\sqrt{2}\mathbf{I}] \mathbf{C} [\sqrt{2}\mathbf{I}] \left[\frac{1}{\sqrt{2}} \mathbf{I} \right] \underline{\varepsilon},$$

where, explicitly,

$$\left[\sqrt{2}\mathbf{I} \right] \underline{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\sigma_{23} \\ \sqrt{2}\sigma_{13} \\ \sqrt{2}\sigma_{12} \end{bmatrix}$$

and

$$\left[\frac{1}{\sqrt{2}}\mathbf{I} \right] \underline{\varepsilon} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \sqrt{2}\varepsilon_{23} \\ \sqrt{2}\varepsilon_{13} \\ \sqrt{2}\varepsilon_{12} \end{bmatrix}.$$

In other words, both the matrix containing the stress-tensor components and the matrix containing the strain-tensor components have the same form, as required. Completing the matrix multiplications, we obtain

$$(5.14.6) \quad \left[\sqrt{2}\mathbf{I} \right] \mathbf{C} \left[\sqrt{2}\mathbf{I} \right] = \begin{bmatrix} C_{11} & C_{12} & C_{13} & \sqrt{2}C_{14} & \sqrt{2}C_{15} & \sqrt{2}C_{16} \\ C_{12} & C_{22} & C_{23} & \sqrt{2}C_{24} & \sqrt{2}C_{25} & \sqrt{2}C_{26} \\ C_{13} & C_{23} & C_{33} & \sqrt{2}C_{34} & \sqrt{2}C_{35} & \sqrt{2}C_{36} \\ \sqrt{2}C_{14} & \sqrt{2}C_{24} & \sqrt{2}C_{34} & 2C_{44} & 2C_{45} & 2C_{46} \\ \sqrt{2}C_{15} & \sqrt{2}C_{25} & \sqrt{2}C_{35} & 2C_{45} & 2C_{55} & 2C_{56} \\ \sqrt{2}C_{16} & \sqrt{2}C_{26} & \sqrt{2}C_{36} & 2C_{46} & 2C_{56} & 2C_{66} \end{bmatrix},$$

which is the corresponding elasticity matrix. Thus, another form of stress-strain equations (4.2.8) can be written as

$$(5.14.7) \quad \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\sigma_{23} \\ \sqrt{2}\sigma_{13} \\ \sqrt{2}\sigma_{12} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & \sqrt{2}C_{14} & \sqrt{2}C_{15} & \sqrt{2}C_{16} \\ C_{12} & C_{22} & C_{23} & \sqrt{2}C_{24} & \sqrt{2}C_{25} & \sqrt{2}C_{26} \\ C_{13} & C_{23} & C_{33} & \sqrt{2}C_{34} & \sqrt{2}C_{35} & \sqrt{2}C_{36} \\ \sqrt{2}C_{14} & \sqrt{2}C_{24} & \sqrt{2}C_{34} & 2C_{44} & 2C_{45} & 2C_{46} \\ \sqrt{2}C_{15} & \sqrt{2}C_{25} & \sqrt{2}C_{35} & 2C_{45} & 2C_{55} & 2C_{56} \\ \sqrt{2}C_{16} & \sqrt{2}C_{26} & \sqrt{2}C_{36} & 2C_{46} & 2C_{56} & 2C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \sqrt{2}\varepsilon_{23} \\ \sqrt{2}\varepsilon_{13} \\ \sqrt{2}\varepsilon_{12} \end{bmatrix}.$$

It is convenient to use equations (5.14.7) to investigate eigenvalues of the corresponding elasticity matrix to gain an insight into invariants of the stress-strain equations.¹⁶

¹⁶Interested readers might refer to Bóna, A., Bucataru, I., Slawinski, M.A. (2007) Coordinate-free classification of elasticity tensor: Journal of Elasticity **87**(2-3), 109–132, and to Rychlewski, J., (1985) On Hooke's law: Prikl. Matem. Mekhan., **48** (3), 420 – 435.

EXERCISE 5.4. 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 (4.2.8), 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 $\mathbf{C}_{\text{MONO}x_1}$.

SOLUTION 5.4. Consider the stress-tensor components σ_{12} and $\hat{\sigma}_{12}$. Using stress-strain equations (4.2.8), we can write

$$(5.14.8) \quad \sigma_{12} = C_{16}\varepsilon_{11} + C_{26}\varepsilon_{22} + C_{36}\varepsilon_{33} + 2C_{46}\varepsilon_{23} + 2C_{56}\varepsilon_{13} + 2C_{66}\varepsilon_{12},$$

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

$$(5.14.9) \quad \sigma_{12} = -\hat{\sigma}_{12} = -C_{16}\varepsilon_{11} - C_{26}\varepsilon_{22} - C_{36}\varepsilon_{33} - 2C_{46}\varepsilon_{23} + 2C_{56}\varepsilon_{13} + 2C_{66}\varepsilon_{12}.$$

Equality between (5.14.8) and (5.14.9) 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

$$(5.14.10) \quad \mathbf{C}_{\text{MONO}x_1} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 0 & 0 \\ C_{12} & C_{22} & C_{23} & C_{24} & 0 & 0 \\ C_{13} & C_{23} & C_{33} & C_{34} & 0 & 0 \\ C_{14} & C_{24} & C_{34} & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & C_{56} \\ 0 & 0 & 0 & 0 & C_{56} & C_{66} \end{bmatrix},$$

as required.

EXERCISE 5.5. ¹⁷Find the stability conditions for a transversely isotropic continuum described by matrix (5.10.3).

¹⁷See also Section 4.3.3.

SOLUTION 5.5. In view of Section 4.3, the stability conditions require that matrix (5.10.3) be positive-definite. Recalling equations (4.3.3), we obtain

$$(5.14.11) \quad C_{11} > 0,$$

$$(5.14.12) \quad C_{33} > 0,$$

$$(5.14.13) \quad C_{44} > 0,$$

and

$$(5.14.14) \quad C_{11} > C_{12}.$$

We notice that matrix (5.10.3) 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

$$\mathbf{C}_2 = \begin{bmatrix} C_{44} & 0 & 0 \\ 0 & C_{44} & 0 \\ 0 & 0 & \frac{C_{11}-C_{12}}{2} \end{bmatrix}.$$

Conditions (5.14.13) and (5.14.14) ensure that matrix \mathbf{C}_2 is positive-definite. In view of condition (5.14.11), the remaining conditions for the positive-definiteness of matrix \mathbf{C}_1 are

$$(5.14.15) \quad \det \begin{bmatrix} C_{11} & C_{12} \\ C_{12} & C_{11} \end{bmatrix} > 0,$$

and

$$(5.14.16) \quad \det \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{12} & C_{11} & C_{13} \\ C_{13} & C_{13} & C_{33} \end{bmatrix} > 0.$$

The condition resulting from determinant (5.14.15) is

$$(5.14.17) \quad C_{11} > |C_{12}|,$$

while the condition resulting from determinant (5.14.16) is

$$C_{33} (C_{11} - C_{12}) (C_{11} + C_{12}) > 2C_{13}^2 (C_{11} - C_{12}).$$

In view of expression (5.14.14), we can rewrite the latter condition as

$$(5.14.18) \quad C_{33} (C_{11} + C_{12}) > 2C_{13}^2.$$

Also, in view of condition (5.14.12), we have $C_{11} + C_{12} > 0$. Consequently, condition (5.14.17) follows from conditions (5.14.12) and (5.14.18). Thus, all the stability conditions for a transversely isotropic continuum are given by expressions (5.14.11), (5.14.12), (5.14.13), (5.14.14) and (5.14.18).

REMARK 5.14.1. Note that if matrix \mathbf{C}_1 is positive-definite, we also have

$$(5.14.19) \quad \det \begin{bmatrix} C_{11} & C_{13} \\ C_{13} & C_{33} \end{bmatrix} > 0,$$

which we can write as

$$(5.14.20) \quad C_{11}C_{33} > C_{13}^2.$$

Herein, we will show that condition (5.14.20) is a consequence of condition (5.14.18). Let us rewrite condition (5.14.18) 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 (5.14.20) as $a > 0$. To show that condition (5.14.20) is a consequence of condition (5.14.18), we first show that $a > b$, which is equivalent to showing that

$$(5.14.21) \quad C_{11}C_{33} > C_{12}C_{33}.$$

Inequality (5.14.21) is true due to conditions (5.14.12) and (5.14.14). Hence, since $a + b > 0$ and $a - b > 0$, by summation we get $2a > 0$, which immediately implies that $a > 0$, as required.

EXERCISE 5.6. ¹⁸Using the formula for the change of coordinates for the components of a tensor as well as Lemma 5.14.2 below, show that if a tensor of rank n , given by $T_{i_1 \dots i_n}$, is invariant under the $(n + 1)$ -fold rotation about a given axis, it is invariant under any rotation about this axis.

LEMMA 5.14.2. Let $P(\Theta)$ be a trigonometric polynomial of at most degree n . If $P(\Theta)$ has a period of $2\pi/(n + 1)$, then $P(\Theta) \equiv \text{const.}$

PROOF. Consider a basis of the space of trigonometric polynomials of at most degree n , given by

$$(5.14.22) \quad f_r(\Theta) = e^{ir\Theta}, \quad r \in \{-n, \dots, n\}.$$

¹⁸See also Section 5.10.2.

We can uniquely write

$$(5.14.23) \quad P(\Theta) = \sum_{r=-n}^n \alpha_r f_r(\Theta),$$

where α_r are complex numbers. In view of expressions (5.14.22) and (5.14.23), we can write

$$(5.14.24) \quad P\left(\Theta + \frac{2\pi}{n+1}\right) = \sum_{r=-n}^n \alpha_r f_r\left(\Theta + \frac{2\pi}{n+1}\right) = \sum_{r=-n}^n \alpha_r e^{ir2\pi/(n+1)} f_r(\Theta).$$

Since $P(\Theta)$ has a period of $2\pi/(n+1)$, examining equations (5.14.23) and (5.14.24), we obtain

$$\alpha_r = \alpha_r e^{ir2\pi/(n+1)}, \quad r \in \{-n, \dots, n\}.$$

Observing that $e^{ir2\pi/(n+1)} \neq 1$ for all $r \in \{-n, \dots, n\}$, except $r = 0$, we conclude that $\alpha_r = 0$, except, possibly, α_0 . Hence, $P(\Theta)$ is constant. \square

SOLUTION 5.6. Consider transformation matrix (5.8.1); namely,

$$(5.14.25) \quad \mathbf{A} = \begin{bmatrix} \cos \Theta & \sin \Theta & 0 \\ -\sin \Theta & \cos \Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The transformed tensor components are given by

$$\hat{T}_{i_1 \dots i_n} = \sum_{j_1=1}^3 \dots \sum_{j_n=1}^3 A_{i_1 j_1} \dots A_{i_n j_n} T_{j_1 \dots j_n}, \quad i_1, \dots, i_n \in \{1, 2, 3\}.$$

In view of matrix (5.14.25), we see that $\hat{T}_{i_1 \dots i_n} = \hat{T}_{i_1 \dots i_n}(\Theta)$ is a trigonometric polynomial in Θ of at most degree n . Since tensor $T_{i_1 \dots i_n}$ is invariant under the rotation by the angle $2\pi/(n+1)$, polynomial $\hat{T}_{i_1 \dots i_n}(\Theta)$ has a period of $2\pi/(n+1)$. Since $\hat{T}_{i_1 \dots i_n}(\Theta)$ is at most of degree n , it follows from Lemma 5.14.2 that this trigonometric polynomial is constant. This means that $T_{i_1 \dots i_n}$ is invariant under any rotation, as required.

EXERCISE 5.7. ¹⁹Show that the elasticity matrix of an isotropic continuum is symmetric even without invoking the strain-energy function.

NOTATION 5.14.3. 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 values 1, 2 and 3.

SOLUTION 5.7. In view of Section 4.2, to show the symmetry of the elasticity matrix, $C_{mn} = C_{nm}$, where, $m, n \in \{1, \dots, 6\}$, it suffices to show that

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

¹⁹See also Section 5.12.1.

Recall stress-strain equations (3.2.1); namely,

$$(5.14.26) \quad \sigma_{ij} = c_{ijkl}\varepsilon_{kl}, \quad i, j \in \{1, 2, 3\},$$

as well as a particular case of these equations that corresponds to isotropic continua and is given by equations (5.12.4), namely,

$$(5.14.27) \quad \sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \quad i, j \in \{1, 2, 3\},$$

where λ and μ are Lamé's parameters. Equating the right-hand sides of equations (5.14.26) and (5.14.27), we can write

$$(5.14.28) \quad c_{ijkl}\varepsilon_{kl} - (\lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}) = 0, \quad i, j \in \{1, 2, 3\}.$$

Let us factor out the strain-tensor components, ε_{kl} . Using the properties of Kronecker's delta, 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, we can rewrite equations (5.14.28) as

$$(5.14.29) \quad [c_{ijkl} - (\lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl})]\varepsilon_{kl} = 0, \quad i, j \in \{1, 2, 3\}.$$

In general, ε_{kl} is not zero. Thus, for equations (5.14.29) to be always true, we require the expression in brackets to be zero. In other words,

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl}, \quad i, j, k, l \in \{1, 2, 3\}.$$

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

$$\begin{aligned} c_{ijkl} &= \lambda\delta_{ij}\delta_{kl} + 2\mu\delta_{ik}\delta_{jl} \\ &= \lambda\delta_{kl}\delta_{ij} + 2\mu\delta_{ki}\delta_{lj} = c_{klij}, \quad i, j, k, l \in \{1, 2, 3\}, \end{aligned}$$

as required.

EXERCISE 5.8. ²⁰Using Lemma 5.14.5, prove Theorem 5.14.6, stated below.

NOTATION 5.14.4. 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 values 1, 2 and 3.

LEMMA 5.14.5. ²¹The general isotropic fourth-rank tensor is

$$(5.14.30) \quad a_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \xi\delta_{ik}\delta_{jl} + \eta\delta_{il}\delta_{jk}, \quad i, j, k, l \in \{1, 2, 3\}.$$

THEOREM 5.14.6. Given the symmetry of the strain tensor, defined in expression (1.4.6), the stress-strain equations for a three-dimensional isotropic continuum are given by expression (5.12.4), namely,

$$\sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \quad i, j \in \{1, 2, 3\},$$

²⁰See also Section 5.12.3.

²¹Readers interested in a proof of Lemma 5.14.5 might refer to Matthews, P.C., (1998) Vector calculus: Springer, pp. 124 – 125.

where $2\mu = \xi + \eta$.

SOLUTION 5.8. **PROOF.** Consider stress-strain equations (3.2.1), namely

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

Inserting expression (5.14.30) for c_{ijkl} , and using the properties of Kronecker's delta, in view of Lemma 5.14.5, we can write

$$\begin{aligned} \sigma_{ij} &= (\lambda\delta_{ij}\delta_{kl} + \xi\delta_{ik}\delta_{jl} + \eta\delta_{il}\delta_{jk})\varepsilon_{kl} \\ &= \lambda\delta_{ij}\varepsilon_{kk} + \xi\varepsilon_{ij} + \eta\varepsilon_{ji}, \quad i, j \in \{1, 2, 3\}. \end{aligned}$$

Since, by its definition, the strain tensor, ε_{kl} , is symmetric, we can write

$$\sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + (\xi + \eta)\varepsilon_{ij}, \quad i, j \in \{1, 2, 3\},$$

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

$$\sigma_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \quad i, j \in \{1, 2, 3\},$$

as required. □

REMARK 5.14.7. While studying isotropic materials it is common to express the two elasticity parameters, shown in matrices (5.12.1) and (5.12.3), in terms of quantities that possess an immediate physical meaning. In Exercises 5.9 – 5.14, we will use Poisson's ratio, which is defined as

$$(5.14.31) \quad \nu := -\frac{\varepsilon_{xx}}{\varepsilon_{zz}} = -\frac{\varepsilon_{yy}}{\varepsilon_{zz}},$$

and Young's modulus, which is defined as

$$E := \frac{\sigma_{xx}}{\varepsilon_{xx}} = \frac{\sigma_{yy}}{\varepsilon_{yy}} = \frac{\sigma_{zz}}{\varepsilon_{zz}}.$$

The relations among Poisson's ratio, Young's modulus and Lamé's parameters are given by

$$(5.14.32) \quad \lambda = \frac{E\nu}{(1 + \nu)(1 - 2\nu)},$$

and

$$(5.14.33) \quad \mu = \frac{E}{2(1 + \nu)}.$$

EXERCISE 5.9. Consider an isotropic continuum. Subjecting this 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

$$(5.14.34) \quad \nu = \frac{\lambda}{2(\lambda + \mu)},$$

where λ and μ are Lamé's parameters.

SOLUTION 5.9. Following stress-strain equations (5.12.4), which describe isotropic continua, we can write

$$\sigma_{xx} = \lambda (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) + 2\mu\varepsilon_{xx} = (\lambda + 2\mu) \varepsilon_{xx} + \lambda\varepsilon_{yy} + \lambda\varepsilon_{zz} = 0.$$

Dividing both sides by ε_{zz} , we obtain

$$(\lambda + 2\mu) \frac{\varepsilon_{xx}}{\varepsilon_{zz}} + \lambda \frac{\varepsilon_{yy}}{\varepsilon_{zz}} + \lambda = 0.$$

Invoking the definition of Poisson's ratio given in expression (5.14.31), we can rewrite the above expression as

$$-(\lambda + 2\mu) \nu - \lambda\nu + \lambda = -2(\lambda + \mu) \nu + \lambda = 0.$$

Hence, solving for ν , we get

$$\nu = \frac{\lambda}{2(\lambda + \mu)},$$

which is expression (5.14.34), as required.

EXERCISE 5.10. ²²Consider an isotropic continuum under a uniaxial stress that leads to small deformations. Using expression (5.14.34), show that no change in volume implies no resistance to change in shape, as stated by $\mu = 0$.

SOLUTION 5.10. 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

$$\begin{aligned} \check{V} &= (x_1 + \Delta x_1)(x_2 + \Delta x_2)(x_3 + \Delta x_3) \\ (5.14.35) \quad &\approx x_1x_2x_3 + x_2x_3\Delta x_1 + x_1x_3\Delta x_2 + x_1x_2\Delta x_3, \end{aligned}$$

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

$$\check{V} - V = 0.$$

Using expression (5.14.35), we can write

$$\check{V} - V = x_2x_3\Delta x_1 + x_1x_3\Delta x_2 + x_1x_2\Delta x_3 = 0.$$

Dividing both sides by $V = x_1x_2x_3$, 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 (1.4.17) 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

$$\frac{\check{V} - V}{V} = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} = 0.$$

²²See also Section 1.4.3.

Dividing both sides by ε_{33} and invoking the definition of Poisson's ratio, given in expression (5.14.31), 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 (5.14.34), we obtain

$$\mu = \frac{1 - 2\nu}{2\nu} \lambda = 0,$$

as required.

EXERCISE 5.11. *Using equations (5.12.4), 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*

$$(5.14.36) \quad \varepsilon_{ij} = \frac{1 + \nu}{E} \sigma_{ij} - \frac{\nu}{E} \delta_{ij} \sum_{k=1}^3 \sigma_{kk}, \quad i, j \in \{1, 2, 3\},$$

where ν is Poisson's ratio and E is Young's modulus.

SOLUTION 5.11. Using expressions (5.14.32) and (5.14.33), we can write stress-strain equations (5.12.4) as

$$\sigma_{ij} = \frac{E\nu}{(1 + \nu)(1 - 2\nu)} \delta_{ij} \sum_{k=1}^3 \varepsilon_{kk} + \frac{E}{1 + \nu} \varepsilon_{ij}, \quad i, j \in \{1, 2, 3\}.$$

Solving for ε_{ij} , we obtain

$$(5.14.37) \quad \varepsilon_{ij} = \frac{1 + \nu}{E} \sigma_{ij} - \frac{\nu}{1 - 2\nu} \delta_{ij} \sum_{k=1}^3 \varepsilon_{kk}, \quad i, j \in \{1, 2, 3\}.$$

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 (5.12.4), we can write all stress-tensor components for which $\delta_{ij} \sum_{k=1}^3 \varepsilon_{kk}$ does not vanish. They are

$$\sigma_{ii} = \lambda \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{ii}, \quad i \in \{1, 2, 3\}.$$

Writing these three equations explicitly, we get

$$\begin{cases} \sigma_{11} = \lambda \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{11} \\ \sigma_{22} = \lambda \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{22} \\ \sigma_{33} = \lambda \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{33} \end{cases}.$$

Summing these three equations, we obtain

$$\begin{aligned}\sigma_{11} + \sigma_{22} + \sigma_{33} &= 3\lambda \sum_{k=1}^3 \varepsilon_{kk} + 2\mu (\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) \\ &= 3\lambda \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \sum_{k=1}^3 \varepsilon_{kk} = (3\lambda + 2\mu) \sum_{k=1}^3 \varepsilon_{kk}.\end{aligned}$$

Expressing the left-hand side as a summation, we can write the sought expression

$$(5.14.38) \quad \sum_{k=1}^3 \varepsilon_{kk} = \frac{\sum_{k=1}^3 \sigma_{kk}}{3\lambda + 2\mu}.$$

Using expression (5.14.38), we can write expression (5.14.37) as

$$(5.14.39) \quad \varepsilon_{ij} = \frac{1+\nu}{E} \sigma_{ij} - \frac{\nu}{(1-2\nu)(3\lambda+2\mu)} \delta_{ij} \sum_{k=1}^3 \sigma_{kk}, \quad i, j \in \{1, 2, 3\}.$$

Consider the term in parentheses that contains λ and μ . Using expressions (5.14.32) and (5.14.33), we can write this term as

$$3\lambda + 2\mu = 3 \frac{E\nu}{(1+\nu)(1-2\nu)} + \frac{E}{1+\nu} = \frac{E}{1-2\nu}.$$

Hence, expression (5.14.39) becomes

$$\varepsilon_{ij} = \frac{1+\nu}{E} \sigma_{ij} - \frac{\nu}{E} \delta_{ij} \sum_{k=1}^3 \sigma_{kk}, \quad i, j \in \{1, 2, 3\},$$

which is expression (5.14.36), as required.

EXERCISE 5.12. Using expression (4.5.4), show that for isotropic continua the strain-energy function can be expressed in terms of the strain-tensor components as

$$(5.14.40) \quad W = \frac{\lambda}{2} \sum_{i=1}^3 \sum_{j=1}^3 \varepsilon_{ii} \varepsilon_{jj} + \mu \sum_{i=1}^3 \sum_{j=1}^3 \varepsilon_{ij} \varepsilon_{ij},$$

where λ and μ are Lamé's parameters.

SOLUTION 5.12. Recall expression (4.5.4); namely,

$$(5.14.41) \quad W = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \varepsilon_{ij}.$$

Also, recall that for an isotropic continuum the stress-strain equations are given by expression (5.12.4); namely,

$$(5.14.42) \quad \sigma_{ij} = \lambda \delta_{ij} \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{ij}, \quad i, j \in \{1, 2, 3\}.$$

Inserting expression (5.14.42) into expression (5.14.41), 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 (5.14.40), as required.

EXERCISE 5.13. *Using expression (4.5.4), show that, for isotropic continua, the strain-energy function can be expressed in terms of the stress-tensor components as*

$$(5.14.43) \quad W = \frac{1}{2E} \left[(1 + \nu) \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \sigma_{ij} - \nu \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ii} \sigma_{jj} \right],$$

where λ and μ are Lamé's parameters, ν is Poisson's ratio, and E is Young's modulus.

SOLUTION 5.13. Recall expression (4.5.4); namely,

$$(5.14.44) \quad W = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \varepsilon_{ij}.$$

Also, recall expression (5.14.36); namely,

$$(5.14.45) \quad \varepsilon_{ij} = \frac{1 + \nu}{E} \sigma_{ij} - \frac{\nu}{E} \delta_{ij} \sum_{k=1}^3 \sigma_{kk}, \quad i, j \in \{1, 2, 3\},$$

where ν and E are Poisson's ratio and Young's modulus, respectively. Inserting expression (5.14.45) into expression (5.14.44), we obtain

$$W = \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \left[\sigma_{ij} \left(\frac{1 + \nu}{E} \sigma_{ij} - \frac{\nu}{E} \delta_{ij} \sum_{k=1}^3 \sigma_{kk} \right) \right].$$

The properties of Kronecker's delta, δ_{ij} , imply

$$W = \frac{1}{2E} \left[(1 + \nu) \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ij} \sigma_{ij} - \nu \sum_{i=1}^3 \sum_{j=1}^3 \sigma_{ii} \sigma_{jj} \right],$$

which is expression (5.14.43), as required.

EXERCISE 5.14. ²³Consider elasticity matrix (5.12.3). 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.

²³See also Section 4.3.3.

SOLUTION 5.14. Stability conditions require the elasticity matrix to be positive-definite. Matrix (5.12.3) is symmetric. As stated in Theorem 4.3.2, 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 (5.14.34), 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, $\nu \in [0, 1/2]$, where the diminishing of the length is accompanied by the extension of the radius.

Part 2

Waves and rays

Introduction to Part 2

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 1*, 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 2*, 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 analytic solutions. However, in anisotropic inhomogeneous continua, we are unable to formulate equations of motion that possess analytic 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.²⁴ 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 Pieter Debye, at the beginning of twentieth century, as well as Vassily 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.

²⁴Readers interested in Hamilton's formulation of the ray theory might refer to Hankins, T.L., (1980) Sir William Rowan Hamilton: The Johns Hopkins University Press, pp. 59 – 95.

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 (4.4.5) — 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 (2.8.1) with constitutive equations (5.12.4). 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, including solutions in various spatial dimensions and nondifferentiable solutions. 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 (5.12.4), namely,

$$(6.1.1) \quad \sigma_{ij} = \lambda \delta_{ij} \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{ij}, \quad i, j \in \{1, 2, 3\},$$

where λ and μ are constants. We also consider Cauchy's equations of motion (2.8.1), namely,

$$(6.1.2) \quad \sum_{j=1}^3 \frac{\partial \sigma_{ij}}{\partial x_j} = \rho \frac{\partial^2 u_i}{\partial t^2}, \quad i \in \{1, 2, 3\}.$$

We wish to combine stress-strain equations (6.1.1) with equations of motion (6.1.2) to get the equations of motion in an isotropic homogeneous continuum. In other words, we substitute expression (6.1.1) into equations (6.1.2) to obtain

$$(6.1.3) \quad \begin{aligned} \rho \frac{\partial^2 u_i}{\partial t^2} &= \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left(\lambda \delta_{ij} \sum_{k=1}^3 \varepsilon_{kk} + 2\mu \varepsilon_{ij} \right) \\ &= \sum_{j=1}^3 \left(\delta_{ij} \lambda \sum_{k=1}^3 \frac{\partial \varepsilon_{kk}}{\partial x_j} + 2\mu \frac{\partial \varepsilon_{ij}}{\partial x_j} \right), \quad i \in \{1, 2, 3\}. \end{aligned}$$

Now, we wish to express the right-hand side of equations (6.1.3) in terms of the displacement vector, \mathbf{u} . Invoking the definition of the strain tensor, given in expression (1.4.6), we can rewrite equations (6.1.3) as

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^3 \left[\delta_{ij} \lambda \sum_{k=1}^3 \frac{\partial}{\partial x_j} \left(\frac{\partial u_k}{\partial x_k} \right) + \mu \frac{\partial}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right], \quad i \in \{1, 2, 3\}.$$

Using the property of Kronecker's delta, we obtain

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \lambda \sum_{k=1}^3 \frac{\partial}{\partial x_i} \left(\frac{\partial u_k}{\partial x_k} \right) + \mu \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i \in \{1, 2, 3\}.$$

Using the linearity of the differential operators, we can rewrite these equations as

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \lambda \sum_{j=1}^3 \frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \sum_{j=1}^3 \frac{\partial^2 u_i}{\partial x_j^2} + \mu \sum_{j=1}^3 \frac{\partial^2 u_j}{\partial x_j \partial x_i}, \quad i \in \{1, 2, 3\},$$

where, in the first summation, for the summation indices, we let $k = j$. Using the equality of mixed partial derivatives, we obtain

$$(6.1.4) \quad \begin{aligned} \rho \frac{\partial^2 u_i}{\partial t^2} &= (\lambda + \mu) \sum_{j=1}^3 \frac{\partial^2 u_j}{\partial x_i \partial x_j} + \mu \sum_{j=1}^3 \frac{\partial^2 u_i}{\partial x_j^2} \\ &= (\lambda + \mu) \frac{\partial}{\partial x_i} \sum_{j=1}^3 \frac{\partial u_j}{\partial x_j} + \mu \left(\sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2} \right) u_i, \quad i \in \{1, 2, 3\}. \end{aligned}$$

We can use vector calculus to concisely state equations (6.1.4). 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 (6.1.4) as

$$(6.1.5) \quad \rho \frac{\partial^2 u_i}{\partial t^2} = (\lambda + \mu) \frac{\partial}{\partial x_i} \nabla \cdot \mathbf{u} + \mu \nabla^2 u_i, \quad i \in \{1, 2, 3\}.$$

We can explicitly write the three equations stated in expression (6.1.5) as

$$\rho \frac{\partial^2}{\partial t^2} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = (\lambda + \mu) \begin{bmatrix} \frac{\partial(\nabla \cdot \mathbf{u})}{\partial x_1} \\ \frac{\partial(\nabla \cdot \mathbf{u})}{\partial x_2} \\ \frac{\partial(\nabla \cdot \mathbf{u})}{\partial x_3} \end{bmatrix} + \mu \nabla^2 \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

Noticing that the first matrix on the right-hand side involves the gradient operator, we can concisely state the three equations shown in expression (6.1.5) as

$$(6.1.6) \quad \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u}.$$

This is the equation of motion that applies to isotropic homogeneous continua.

We wish to write equation (6.1.6) 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

$$(6.1.7) \quad \nabla^2 \mathbf{a} = \nabla (\nabla \cdot \mathbf{a}) - \nabla \times (\nabla \times \mathbf{a}),$$

and letting $\mathbf{a} = \mathbf{u}$, we can rewrite equation (6.1.6) as

$$(6.1.8) \quad \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \nabla (\nabla \cdot \mathbf{u}) - \mu \nabla \times (\nabla \times \mathbf{u}).$$

Equation (6.1.8) 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 (1.4.18) and (1.5.2), respectively, we can immediately write

$$(6.1.9) \quad \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \nabla \varphi - \mu \nabla \times \mathbf{\Psi}.$$

Equation (6.1.9) 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. This is due to the fact that, in general, ∇^2 is defined for vectors whose direction is set. For vector fields, however,

this direction changes from point to point. Since $\mathbf{u}(\mathbf{x}, t)$ is a vector field, equation (6.1.6) is valid only for orthonormal coordinates.¹ Equation (6.1.8), 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 (6.1.9) into its parts, which are associated with the dilatation and with the rotation vector.

To obtain the wave equation for P waves, we take the divergence of equation (6.1.9). Since in a homogeneous continuum λ and μ are constants, we can write

$$(6.1.10) \quad \nabla \cdot \left[\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right] = (\lambda + 2\mu) \nabla \cdot \nabla \varphi - \mu \nabla \cdot \nabla \times \mathbf{\Psi}.$$

The factor of μ disappears since $\nabla \cdot \nabla \times \mathbf{\Psi} = 0$ for all $\mathbf{\Psi}$. 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 (6.1.10) becomes

$$(6.1.11) \quad \nabla \cdot \left[\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right] = (\lambda + 2\mu) \nabla^2 \varphi.$$

Let us consider the left-hand side of equation (6.1.11). 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 the divergence. Also — in view of the the equality of mixed partial derivatives — we can interchange time and space derivatives. Thus, we get

$$\rho \frac{\partial^2 \varphi}{\partial t^2} = (\lambda + 2\mu) \nabla^2 \varphi,$$

where, on the left-hand side, we used again definition (1.4.18). Rearranging, we obtain

$$(6.1.12) \quad \nabla^2 \varphi = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 \varphi}{\partial t^2},$$

which is the wave equation whose wave function is given by dilatation, $\varphi(\mathbf{x}, t) = \nabla \cdot \mathbf{u}(\mathbf{x}, t)$.

Equation (6.1.12) is the wave equation wave for P waves. As shown in Section 6.4,

$$(6.1.13) \quad v := \sqrt{\frac{\lambda + 2\mu}{\rho}}$$

is the propagation speed. In view of Section 5.12, the presence of both Lamé's parameters in expression (6.1.13) suggests that P waves subject the continuum to both a change in volume and a change in shape.

¹Readers interested in this statement might refer to Feynman, R.P., Leighton, R.B., and Sands, M., (1963/1989) Feynman's lectures on physics: Addison-Wesley Publishing Co., Vol. II, p. 2-12.

In view of definition (1.4.18), 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 (6.1.9) and write

$$(6.1.14) \quad \nabla \times \left[\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \right] = (\lambda + 2\mu) \nabla \times \nabla \varphi - \mu \nabla \times \nabla \times \Psi.$$

The factor of $\lambda + 2\mu$ disappears since $\nabla \times \nabla \varphi = \mathbf{0}$ for all φ . Recalling definition (1.5.2) 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 get

$$(6.1.15) \quad \rho \frac{\partial^2 \Psi}{\partial t^2} = -\mu \nabla \times [\nabla \times \Psi].$$

Invoking vector-calculus identity (6.1.7) and letting $\mathbf{a} = \Psi$, we can write equation (6.1.15) as

$$\rho \frac{\partial^2 \Psi}{\partial t^2} = -\mu [\nabla (\nabla \cdot \Psi) - \nabla^2 \Psi].$$

In view of definition (1.5.2) and the vanishing of the divergence of a curl, the first term in brackets disappears. Hence, we obtain

$$(6.1.16) \quad \nabla^2 \Psi = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \Psi}{\partial t^2},$$

which is the wave equation whose wave function is given by the rotation vector, $\Psi(\mathbf{x}, t) = \nabla \times \mathbf{u}(\mathbf{x}, t)$.

Equation (6.1.16) is the wave equation wave for S waves. As shown in Section 6.4,

$$(6.1.17) \quad v := \sqrt{\frac{\mu}{\rho}}$$

is the propagation speed. In view of Section 5.12, the presence of the single Lamé's parameter, namely, μ , in expression (6.1.17), 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 (1.5.2), S waves are sometimes referred to as rotational waves. Since the rotation vector is given by $\Psi = \nabla \times \mathbf{u}$, we conclude that $\nabla \cdot \Psi = 0$. 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 (6.1.4) 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. These are the waves for which the components of the displacement vector are functions of the direction of propagation only.

To gain insight into the concept of plane waves, let us revisit equations (6.1.12) and (6.1.16). Let the plane waves propagate along the x_1 -axis. Thus, in view of the properties of plane waves and following expression (2.4.3), we write the displacement vector as

$$\mathbf{u} = [u_1(x_1, t), u_2(x_1, t), u_3(x_1, t)].$$

Since all the partial derivatives of \mathbf{u} with respect to x_2 and x_3 vanish, equations (6.1.4) 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

$$(6.2.1) \quad \frac{\partial^2 u_1}{\partial x_1^2} = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 u_1}{\partial t^2},$$

$$(6.2.2) \quad \frac{\partial^2 u_2}{\partial x_1^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 u_2}{\partial t^2},$$

and

$$(6.2.3) \quad \frac{\partial^2 u_3}{\partial x_1^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 u_3}{\partial t^2}.$$

Consider equation (6.2.1). Recall expression (1.4.18), which in this case becomes

$$(6.2.4) \quad \varphi = \frac{\partial u_1}{\partial x_1}.$$

Taking the derivative of equation (6.2.1) with respect to x_1 , we obtain

$$(6.2.5) \quad \frac{\partial^3 u_1}{\partial x_1^3} = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^3 u_1}{\partial x_1 \partial t^2}.$$

Using expression (6.2.4) in equation (6.2.5), we obtain

$$(6.2.6) \quad \frac{\partial^2 \varphi}{\partial x_1^2} = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 \varphi}{\partial t^2},$$

which is a plane-wave form of equation (6.1.12). Examining equations (6.2.4) and (6.2.6), 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 9.4.

Now, consider equations (6.2.2) and (6.2.3). Recall expression (1.5.2), which in this case becomes

$$(6.2.7) \quad \Psi = \left[0, -\frac{\partial u_3}{\partial x_1}, \frac{\partial u_2}{\partial x_1} \right].$$

Taking the derivative of equations (6.2.3) and (6.2.2) with respect to x_1 , and writing them as a vector, we obtain

$$\begin{bmatrix} 0 \\ -\frac{\partial^3 u_3}{\partial x_1^3} \\ \frac{\partial^3 u_2}{\partial x_1^3} \end{bmatrix} = \frac{1}{\frac{\mu}{\rho}} \begin{bmatrix} 0 \\ -\frac{\partial^3 u_3}{\partial x_1 \partial t^2} \\ \frac{\partial^3 u_2}{\partial x_1 \partial t^2} \end{bmatrix}.$$

Using the equality of mixed partial derivatives and expression (6.2.7), we obtain

$$(6.2.8) \quad \frac{\partial^2 \Psi}{\partial x_1^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \Psi}{\partial t^2},$$

which is a plane-wave form of equation (6.1.16). Examining equations (6.2.7) and (6.2.8), 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 9.6.

Plane waves are an approximation that allows us to study, in homogeneous media, a wavefield that results from a distant source. Notably, in Chapter 10, we will use plane waves to study reflection and transmission of waves at an interface separating two anisotropic homogeneous halfspaces. 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 6.10.4 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. To derive the wave equation for P waves, which we expressed in terms of scalar

function φ , we took the divergence of Cauchy's equations of motion. To derive the wave equation for S waves, which we expressed in terms of vector function Ψ , we took the curl of these equations. Herein, we will obtain equations that correspond to P and S waves by using Helmholtz's method of separating a vector function into its scalar and vector potentials. We will obtain these equations by inserting the potentials into Cauchy's equations of motion.

According to Helmholtz's theorem², a differentiable function $\mathbf{u}(\mathbf{x}, t)$ can be decomposed into

$$(6.3.1) \quad \mathbf{u}(\mathbf{x}, t) = \nabla \mathcal{P}(\mathbf{x}, t) + \nabla \times \mathbf{S}(\mathbf{x}, t),$$

where \mathcal{P} and $\mathbf{S} = [S_1, S_2, S_3]$ are called the scalar and vector potentials, respectively. Following the definitions of the gradient and curl operators, we can explicitly write the components of \mathbf{u} as

$$u_1(\mathbf{x}, t) = \frac{\partial \mathcal{P}}{\partial x_1} + \frac{\partial S_3}{\partial x_2} - \frac{\partial S_2}{\partial x_3},$$

$$u_2(\mathbf{x}, t) = \frac{\partial \mathcal{P}}{\partial x_2} + \frac{\partial S_1}{\partial x_3} - \frac{\partial S_3}{\partial x_1}$$

and

$$u_3(\mathbf{x}, t) = \frac{\partial \mathcal{P}}{\partial x_3} + \frac{\partial S_2}{\partial x_1} - \frac{\partial S_1}{\partial x_2},$$

which constitute a system of differential equations. This system does not have a unique solution. It is common to consider also another equation; namely,

$$(6.3.2) \quad \nabla \cdot \mathbf{S}(\mathbf{x}, t) = 0.$$

since we can always find \mathcal{P} and \mathbf{S} that satisfy the system composed of equations (6.3.1) and (6.3.2).³ Introducing equation (6.3.2) does not result in a unique determination of \mathbf{S} . It only reduces possible choices of this vector, as shown below.

We will use expressions (6.3.1) and (6.3.2) in Sections 6.3.3 and 6.3.4. In the next section, we will justify our introducing equation (6.3.2).

6.3.2. Gauge transformation. Let us consider equation (6.3.2) in the context of equation (6.3.1). We are allowed to set $\nabla \cdot \mathbf{S} = 0$ since, in view of properties of the vector operators, \mathbf{S} used in equation (6.3.1) is determined up to a gradient, ∇f , where $f(\mathbf{x})$ is any differentiable function. In mathematical physics, changing \mathbf{S} by adding ∇f to it is called a gauge transformation. Let

$$(6.3.3) \quad \tilde{\mathbf{S}} = \mathbf{S} + \nabla f.$$

²Readers 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.

³Interested readers might refer to Box 4.2 in Aki, K., and Richards, P.G., (2002) *Quantitative seismology (2nd edition)*: University Science Books.

Taking the curl of both sides of equation (6.3.3), using the linearity of the differential operator and the vanishing of the curl of a gradient, we obtain

$$\nabla \times \tilde{\mathbf{S}} = \nabla \times (\mathbf{S} + \nabla f) = \nabla \times \mathbf{S}.$$

Examining this result in the context of expression (6.3.1), we see that the same \mathbf{u} is obtained using either \mathbf{S} or $\mathbf{S} + \nabla f$. We can use this freedom of choice to set $\nabla \cdot \tilde{\mathbf{S}} = 0$. This is tantamount to finding f such that $\nabla^2 f = -\nabla \cdot \mathbf{S}$. To reach this conclusion, we took the divergence of both sides of equation (6.3.3) to get

$$(6.3.4) \quad \nabla \cdot \tilde{\mathbf{S}} = \nabla \cdot (\mathbf{S} + \nabla f(\mathbf{x})) = \nabla \cdot \mathbf{S} + \nabla^2 f = 0.$$

Since both $\tilde{\mathbf{S}}$ and \mathbf{S} result in the same \mathbf{u} , we have justified our adding equation (6.3.2) to system (6.3.1).

In particular, it is interesting to notice by examining equation (6.3.4), that if f is a harmonic function — in other words, if f is a solution of $\nabla^2 f = 0$ — then, $\nabla \cdot \tilde{\mathbf{S}} = \nabla \cdot \mathbf{S}$. Consequently, in view of equation (6.3.3), we can always add to \mathbf{S} the gradient of a harmonic function.

6.3.3. Equation of motion. To study the equations of motion in terms of the displacement potentials, we insert expression (6.3.1) into equation (6.1.6) and write

$$\rho \frac{\partial^2 (\nabla \mathcal{P} + \nabla \times \mathbf{S})}{\partial t^2} = (\lambda + \mu) \nabla [\nabla \cdot (\nabla \mathcal{P} + \nabla \times \mathbf{S})] + \mu \nabla^2 (\nabla \mathcal{P} + \nabla \times \mathbf{S}).$$

Using the vanishing of the divergence of a curl and the definition of Laplace's operator, we obtain

$$\begin{aligned} \rho \frac{\partial^2 (\nabla \mathcal{P} + \nabla \times \mathbf{S})}{\partial t^2} &= (\lambda + \mu) \nabla (\nabla \cdot \nabla \mathcal{P}) + \mu \nabla^2 (\nabla \mathcal{P} + \nabla \times \mathbf{S}) \\ &= (\lambda + \mu) \nabla (\nabla^2 \mathcal{P}) + \mu \nabla^2 (\nabla \mathcal{P} + \nabla \times \mathbf{S}). \end{aligned}$$

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

$$\begin{aligned} \nabla \left(\rho \frac{\partial^2 \mathcal{P}}{\partial t^2} \right) + \nabla \times \left(\rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right) &= \nabla [(\lambda + \mu) \nabla^2 \mathcal{P}] + \nabla (\mu \nabla^2 \mathcal{P}) + \nabla \times (\mu \nabla^2 \mathbf{S}) \\ &= \nabla [(\lambda + 2\mu) \nabla^2 \mathcal{P}] + \nabla \times (\mu \nabla^2 \mathbf{S}). \end{aligned}$$

Rearranging, we obtain

$$(6.3.5) \quad \nabla \left[(\lambda + 2\mu) \nabla^2 \mathcal{P} - \rho \frac{\partial^2 \mathcal{P}}{\partial t^2} \right] + \nabla \times \left[\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right] = \mathbf{0},$$

which is the equation of motion for isotropic homogeneous continua in terms of the scalar and vector potentials.

6.3.4. P and S waves. Looking at equation (6.3.5), we see that it is satisfied if

$$(6.3.6) \quad \nabla^2 \mathcal{P} - \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 \mathcal{P}}{\partial t^2} = 0$$

and

$$(6.3.7) \quad \nabla^2 \mathbf{S} - \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 \mathbf{S}}{\partial t^2} = \mathbf{0},$$

which, in view of equations (6.1.12) and (6.1.16), appear to be associated with P and S waves, respectively. A rigorous analysis of this result is associated with Lamé's theorem.⁴ Equations (6.3.6) and (6.3.7) are wave equations whose wave functions are the scalar and vector potentials, respectively. Motivated by this observation, we wish to study the relation of the scalar and vector potentials to the two wave equations whose wave functions are given by the dilatation and the rotation vector; namely equations (6.1.12) and (6.1.16), respectively.

As in Section 6.1.2, let us take the divergence of equation (6.3.5). Using the vanishing of the divergence of a curl and the definition of Laplace's operator, we obtain

$$(6.3.8) \quad \nabla^2 \left[(\lambda + 2\mu) \nabla^2 \mathcal{P} - \rho \frac{\partial^2 \mathcal{P}}{\partial t^2} \right] = 0.$$

Using the linearity of the differential operator and the equality of mixed partial derivatives, we can rewrite equation (6.3.8) as

$$(6.3.9) \quad (\lambda + 2\mu) \nabla^2 (\nabla^2 \mathcal{P}) - \rho \frac{\partial^2 (\nabla^2 \mathcal{P})}{\partial t^2} = 0.$$

To relate the scalar potential, \mathcal{P} , to the dilatation, φ , let us take the divergence of expression (6.3.1). Using the vanishing of the divergence of a curl and recalling definition (1.4.18) as well as the definition of Laplace's operator, we obtain

$$(6.3.10) \quad \varphi := \nabla \cdot \mathbf{u} = \nabla \cdot \nabla \mathcal{P} \equiv \nabla^2 \mathcal{P}.$$

In other words, the dilatation is equal to the Laplacian of the scalar potential. Using expression (6.3.10), we can rewrite equation (6.3.9) as

$$(6.3.11) \quad \nabla^2 \varphi = \frac{1}{\frac{\lambda+2\mu}{\rho}} \frac{\partial^2 \varphi}{\partial t^2},$$

which is equation (6.1.12), as expected. Thus, we conclude that the Laplacian of the scalar potential, \mathcal{P} , satisfies the wave equation for P waves.

⁴Readers interested in this theorem might refer to Aki, K., and Richards, P.G., (2002) *Quantitative seismology (2nd edition)*: University Science Books, pp. 67 – 69.

As in Section 6.1.3, let us take the curl of equation (6.3.5). Using the vanishing of the curl of a gradient, we get

$$\nabla \times \nabla \times \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right) = \mathbf{0}.$$

Recalling identity (6.1.7) 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 this equation as

$$(6.3.12) \quad \nabla \left[\mu \nabla^2 (\nabla \cdot \mathbf{S}) - \rho \frac{\partial^2 (\nabla \cdot \mathbf{S})}{\partial t^2} \right] = \nabla^2 \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right).$$

In view of equation (6.3.2), $\nabla \cdot \mathbf{S} = 0$, equation (6.3.12) becomes

$$(6.3.13) \quad \nabla^2 \left(\mu \nabla^2 \mathbf{S} - \rho \frac{\partial^2 \mathbf{S}}{\partial t^2} \right) = \mathbf{0}.$$

Again using the linearity of the differential operator and the equality of mixed partial derivatives, we can write equation (6.3.13) as

$$(6.3.14) \quad \mu \nabla^2 (\nabla^2 \mathbf{S}) - \rho \frac{\partial^2 (\nabla^2 \mathbf{S})}{\partial t^2} = \mathbf{0}.$$

To relate the vector potential, \mathbf{S} , to the rotation vector, $\mathbf{\Psi}$, let us take the curl of expression (6.3.1). Using the vanishing of the curl of a gradient and recalling definition (1.5.2), we obtain

$$\mathbf{\Psi} := \nabla \times \mathbf{u} = \nabla \times \nabla \times \mathbf{S}.$$

Following identity (6.1.7) and letting $\mathbf{a} = \mathbf{S}$, we get

$$\mathbf{\Psi} = \nabla \times (\nabla \times \mathbf{S}) = \nabla (\nabla \cdot \mathbf{S}) - \nabla^2 \mathbf{S}.$$

In view of equation (6.3.2), $\nabla \cdot \mathbf{S} = 0$, we obtain

$$(6.3.15) \quad \mathbf{\Psi} = -\nabla^2 \mathbf{S}.$$

In other words, the rotation vector is equal to the negative Laplacian of the vector potential. Using expression (6.3.15), we can rewrite equation (6.3.14) as

$$(6.3.16) \quad \nabla^2 \mathbf{\Psi} = \frac{1}{\rho} \frac{\partial^2 \mathbf{\Psi}}{\partial t^2},$$

which is equation (6.1.16), as expected. Thus, we conclude that under condition (6.3.2) the Laplacian of the vector potential, \mathbf{S} , satisfies the wave equation for S waves.

This derivation of equations (6.3.11) and (6.3.16) is analogous to the method for obtaining Maxwell's equations in the electromagnetic theory using the vector and scalar potentials.⁵

6.4. Solutions of wave equation for single spatial dimension

6.4.1. d'Alembert's approach. To gain further insights into the physical meaning of equations (6.1.12) and (6.1.16), 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.⁶ Consider the initial-value problem given by

$$(6.4.1) \quad \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u(x, t)}{\partial t^2} = 0,$$

where $u = u(x, t)$ is the wave function and v is a constant. Let the initial conditions be stated by

$$(6.4.2) \quad \begin{cases} u(x, t)|_{t=0} = \gamma(x) \\ \frac{\partial u(x, t)}{\partial t}|_{t=0} = \eta(x) \end{cases}.$$

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.⁷ It is based on the following two lemmas.

LEMMA 6.4.1. *Equation*

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0,$$

is equivalent to

$$(6.4.3) \quad \frac{\partial^2 u(y, z)}{\partial y \partial z} = 0,$$

where the new coordinates are

$$(6.4.4) \quad \begin{cases} y = x + vt \\ z = x - vt \end{cases}.$$

Details of the derivation of Lemma 6.4.1 are shown in Exercise 6.1.

⁵Readers interested in the derivation of Maxwell's equations using the potentials might refer to Feynman, R.P., Leighton, R.B., and Sands, M., (1963/1989) Feynman's lectures on physics: Addison-Wesley Publishing Co., Vol. II, pp. 18-9 – 18-11.

⁶Applying Newton's second law of motion, we can derive equation (6.4.1) for either longitudinal waves or transverse waves, which correspond to P waves or S waves, respectively. Readers interested in such derivations of the one-dimensional wave equation for the P and S waves 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, or to Garrity, T.A., (2001) All the mathematics you missed [but need to know for graduate school]: Cambridge University Press, pp. 274 – 277, where a derivation for the S waves is presented.

⁷Readers 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.

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. For y and z given by constants, $y = x + vt$ and $z = x - vt$ are straight lines in the xt -plane, which are known as the characteristics of the wave equation. Expressions $y(x, t)$ and $z(x, t)$ satisfy the characteristic equation of wave equation (6.4.1), as shown in Exercise 6.17.⁸

LEMMA 6.4.2. *For equation*

$$\frac{\partial^2 u(y, z)}{\partial y \partial z} = 0,$$

the only form of the solution is

$$(6.4.5) \quad u(y, z) = f(y) + g(z),$$

where f and g are twice-differentiable arbitrary functions.

Details of the derivation of Lemma 6.4.2 are shown in Exercise 6.2.

Combining Lemma 6.4.1 and Lemma 6.4.2, we can state the following corollary.

COROLLARY 6.4.3. *Following Lemma 6.4.1 and Lemma 6.4.2, and using coordinates (6.4.4), we can write the only form of the solution of equation (6.4.1) as*

$$(6.4.6) \quad u(x, t) = f(x + vt) + g(x - vt),$$

where f and g are arbitrary twice-differentiable functions.

Solution (6.4.6) allows arbitrary twice-differentiable 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 (6.4.6) that satisfies the constraints provided by initial conditions (6.4.2). Inserting expression (6.4.6) into system of equations (6.4.2), we can write

$$(6.4.7) \quad \begin{cases} f(x) + g(x) = \gamma(x) \\ v f'(x) - v g'(x) = \eta(x) \end{cases},$$

where we used the chain rule with f' and g' denoting the derivatives with respect to arguments $(x + vt)$ and $(x - vt)$, respectively, and evaluated the results at $t = 0$. 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

$$(6.4.8) \quad \begin{cases} f(x) + g(x) = \gamma(x) \\ f(x) - g(x) = \frac{1}{v} \int_{x_0}^x \eta(\zeta) d\zeta \end{cases},$$

⁸Readers 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.

where ζ is the integration variable. Adding the two equations together, we get

$$(6.4.9) \quad f(x) = \frac{1}{2} \left[\gamma(x) + \frac{1}{v} \int_{x_0}^x \eta(\zeta) d\zeta \right],$$

while subtracting the second equation from the first one gives us

$$(6.4.10) \quad g(x) = \frac{1}{2} \left[\gamma(x) - \frac{1}{v} \int_{x_0}^x \eta(\zeta) d\zeta \right].$$

Inserting expressions (6.4.9) and (6.4.10) into solution (6.4.6), we write

$$u(x, t) = \frac{1}{2} \left\{ \gamma(x + vt) + \gamma(x - vt) + \frac{1}{v} \left[\int_{x_0}^{x+vt} \eta(\zeta) d\zeta - \int_{x_0}^{x-vt} \eta(\zeta) d\zeta \right] \right\}.$$

Using the fact that reversing the limits of integration changes the sign of the integral, we obtain

$$(6.4.11) \quad u(x, t) = \frac{1}{2} \left[\gamma(x + vt) + \gamma(x - vt) + \frac{1}{v} \int_{x-vt}^{x+vt} \eta(\zeta) d\zeta \right],$$

which is the solution of the initial-value problem given by equations (6.4.1) and (6.4.2).

Let us interpret the physical meaning of solution (6.4.11). If we view x as the position variable and t as the time variable, solution (6.4.11) describes propagation of u in the one-dimensional x -space. Solution $u(x, t)$ is completely determined by the differential equation and the initial conditions, which describe the solution at the initial time, $u(x, 0) = \gamma(x)$, and the velocity of displacement of u at that instant, $\eta(x)$; in other words, at point x , $\gamma(x)$ displaces with velocity η .

Note that for the initial-value problem discussed herein, the one-dimensional space has an infinite length. To consider finite and semifinite cases, we would also have to consider boundary conditions.⁹

To examine the concept of propagation, let us consider solution (6.4.6) with $g = 0$; namely,

$$(6.4.12) \quad u(x, t) = f(x + vt).$$

We wish to examine the propagation of a given point that belongs to f . Such a point corresponds to a particular value of f . In view of expression (6.4.12), we see that a particular value of $u(x, t)$ remains the same if the value of $x + vt$ stays the same. According to conditions (6.4.2), at time $t = 0$, we have $u(x, 0) = \gamma(x)$. Let us consider location x_0 , at that time. At this location and at that time, we have $\gamma(x_0)$. Let us follow this value of γ . At time t_1 , we have $\gamma(x_1 + vt_1)$. To follow the same value, we require the constancy of the argument; in other words, $x_0 = x_1 + vt_1$. Solving for x_1 , we get $x_1 = x_0 - vt$, which

⁹Readers interested in a finite and semifinite cases might refer to Arnold, V.I., (2004) Lectures on partial differential equations: Springer-Verlag, pp. 29 – 32.

means that the value of γ that was at x_0 moved to $x_0 - vt$; it moved left along the x -axis by distance vt . We conclude that constant v in equation (6.4.1) is the propagation speed.

Returning to solution (6.4.6) and following an argument analogous to the one presented above but with $f = 0$, we conclude that f and g move in opposite directions. In general, f and g are different from one another; they are explicitly stated in terms of the two initial conditions by expressions (6.4.9) and (6.4.10). Examining these expressions, we note that the two functions propagating in the opposite directions are the same if $\eta(x) = 0$; in other words, if there is no initial velocity of displacement.

6.4.2. Directional derivative. To gain further insight into wave equation (6.4.1) and solution (6.4.6), let us rewrite this equation using directional derivatives. We write

$$\left(v^2 \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2} \right) u(x, t) = 0,$$

where the term in parentheses is a differential operator, which we can rewrite as a composition of two differential operators; namely,¹⁰

$$(6.4.13) \quad \left(v \frac{\partial}{\partial x} - \frac{\partial}{\partial t} \right) \left(v \frac{\partial}{\partial x} + \frac{\partial}{\partial t} \right) u(x, t) = 0.$$

Using the scalar product, we can write each operator as

$$(6.4.14) \quad \left([v, -1] \cdot \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial t} \right] \right) \left([v, 1] \cdot \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial t} \right] \right) u(x, t) = 0,$$

where the terms in parentheses have the form of directional derivatives with directions $[v, -1]$ and $[v, 1]$. Equation (6.4.14) means that u is constant in these directions. In other words, u is constant along the lines whose slopes are

$$\frac{dx}{dt} = \pm v.$$

Solving this ordinary differential equation, we get

$$x = \pm vt + C,$$

where C is the integration constant. Solving for C , we write

$$C = x \mp vt.$$

Since a function of a constant is a constant, the solution of equation (6.4.14) can be written as

$$u(x, t) = f(x + vt) + g(x - vt),$$

which is solution (6.4.6); it states that f and g are constant along $x = \pm vt + C$.

¹⁰Readers interested in using expression (6.4.13) to obtain d'Alembert's solution via a first-order partial differential equation might refer to Evans, L.C., (1998) Partial differential equations: AMS, pp. 67 – 68.

In view of equation (6.4.14), we can formulate the solution in a different way. Function u remains unchanged along directions $[v, -1]$ and $[v, 1]$, which correspond to lines that are described by their perpendiculars; namely, $[1, v]$ and $[1, -v]$, respectively. Hence, the arguments of f and g are

$$[x, t] \cdot [1, v] = x + vt$$

and

$$[x, t] \cdot [1, -v] = x - vt,$$

respectively.

6.4.3. Well-posed problem. We are interested in knowing whether or not the solution of the initial-value problem given by equations (6.4.1) and (6.4.2) is unique. Also, since this solution results from the initial conditions, we wish to know whether or not it depends smoothly on the initial data. In other words, we wish to verify that the dependence is such that a small change in input affects the solution by a small amount only. We refer to such a solution as a stable solution. A problem consisting of equations that result in a unique and stable solution is called a well-posed problem. Wave equation (6.4.1) together with conditions (6.4.2) constitute a well-posed problem as we will see below.

It is important to note that many mathematical physics questions do not constitute well-posed problems; they are ill-posed problems. However, this classical nomenclature does not imply that a well-posed problem is physically more realistic than an ill-posed problem. For instance, inverse problems, which are of great interest in seismology, often do not possess unique solutions.

Let us demonstrate the uniqueness of solution (6.4.11). We will demonstrate it explicitly by following the construction of our solution derived in Section 6.4.1.

Consider equation (6.4.1), namely,

$$(6.4.15) \quad \frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0.$$

Following Lemma 6.4.1 and Lemma 6.4.2, we obtained Corollary 6.4.3, according to which

$$(6.4.16) \quad u(x, t) = f(x + vt) + g(x - vt)$$

is the only form of the solution of equation (6.4.15). Using initial conditions (6.4.2), we obtain system (6.4.8), which allows us to uniquely solve for f and g in terms of γ and η . Consequently, we can write solution (6.4.16) in terms of γ and η to obtain expression (6.4.11). Thus, in view of the construction of solution (6.4.11), we conclude that the solution is unique.

Similarly, examining the construction that led to solution (6.4.11), as described in Section 6.4.1, we see that γ and η are linear combinations of the continuous and differentiable functions f and g and of their first derivatives. Hence, functions γ and η , which are the initial conditions, are also continuous. Since the

solution depends linearly on γ and on the integral of η , we conclude that a small change in γ or η results in a small change in u . Hence the solution, $u(x, t)$, is stable.

We can also show the uniqueness of the solution of the wave equation by studying the energy of wave function $u(x, t)$ at a given instant in time. Herein, by analogy to the above section, we will discuss the energy of a wave function in a single spatial dimension, where $u = u(x, t)$. The presented method, however, can be easily extended to higher dimensions, thereby showing that, in general, the wave equations with appropriate initial conditions is a well-posed problem. Furthermore, we could also define the wave-function energy for equations in dissipative and dispersive media; this aspect of wave phenomena, however, is beyond the scope of this book.¹¹

Considering equation (6.4.1), let us define the wave-function energy to be

$$(6.4.17) \quad \mathcal{E}(t) := \frac{1}{2} \int_{-\infty}^{\infty} \left[\left(\frac{\partial u}{\partial t} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] dx.$$

In the study of partial differential equations, it is common to define energy as an integral whose integrand is composed of squares of first derivatives of a function, as we did in expression (6.4.17) for function u .

To interpret the physical meaning of this definition, we first note that, as written, the physical units of \mathcal{E} are the product of velocity squared and distance, which are not units of energy. However, if we multiply \mathcal{E} by unit mass per unit length, $[kg/m]$, the units of \mathcal{E} are the product of velocity squared and mass, which are the units of energy. Also, we can write \mathcal{E} as the sum of two integrals, namely, $\int \left[(\partial u / \partial t)^2 / 2 \right] dx$ and $\int \left[v^2 (\partial u / \partial x)^2 / 2 \right] dx$. We can view the former integral as corresponding to the kinetic energy of displacement. Since $v^2/2$ is a constant, we can view the latter integral as corresponding to the potential energy; in particular, it corresponds to the strain energy that is associated with deformation u .

Differentiating expression (6.4.17) with respect to time and using the fact that limits of integration are fixed, we get

$$\frac{d\mathcal{E}}{dt} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial}{\partial t} \left[\left(\frac{\partial u}{\partial t} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] dx = \int_{-\infty}^{\infty} \left(\frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} + v^2 \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} \right) dx.$$

To study this integral, we rewrite it as two integrals, namely,

$$(6.4.18) \quad \frac{d\mathcal{E}}{dt} = \int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} dx + v^2 \int_{-\infty}^{\infty} \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} dx.$$

¹¹Readers interested in various definitions of energy in the study of partial differential equations might refer to McOwen, R.C., (1996) Partial differential equations: Methods and applications: Prentice-Hall, Inc., pp. 91 – 97.

Solving the second integral by parts, we get

$$\int_{-\infty}^{\infty} \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} dx = \frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \Big|_{x=-\infty}^{x=\infty} - \int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial x^2} dx.$$

We assume that η , which appears in condition (6.4.2), has compact support. This means that $\eta(x)$ is nonzero on a closed and bounded interval, but is zero everywhere else. Thus, $\partial u / \partial t$ vanishes at both positive and negative infinity of the space variable, x . In other words, there is no displacement velocity infinitely far from the neighbourhood of $x = 0$. For the integrated term, $(\partial u / \partial t) (\partial u / \partial x) |_{-\infty}^{\infty}$, to vanish, we must also assume $\partial u / \partial x$ to be finite at both positive and negative infinity of x . If we assume that function γ has compact support, $\partial u / \partial x = 0$ at $x = \pm\infty$ since the propagation speed, v , is finite. This implies that, at an infinite distance, $u(x, t) = 0$, for all $t < \infty$. Using these assumptions, we can write

$$\int_{-\infty}^{\infty} \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} dx = - \int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial x^2} dx.$$

Inserting this expression into expression (6.4.18), we get

$$\frac{d\mathcal{E}}{dt} = \int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t \partial x^2} dx - v^2 \int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial x^2} dx = \int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \left[\frac{\partial^2 u}{\partial t^2} - v^2 \frac{\partial^2 u}{\partial x^2} \right] dx.$$

Examining the integrand in view of wave equation (6.4.1), we see that the term in brackets vanishes. Hence, we obtain

$$(6.4.19) \quad \frac{d\mathcal{E}}{dt} = 0,$$

which implies that \mathcal{E} is constant.

Equation (6.4.19) states that for wave equation (6.4.1), the wave-function energy defined by expression (6.4.17) is conserved.¹² We can state it as the following theorem.

THEOREM 6.4.4. *If $u(x, t)$ is a solution of wave equation (6.4.1) together with initial conditions (6.4.2) that are given by functions with compact support, then energy defined by expression (6.4.17) is conserved. In other words, $\mathcal{E}(t) = \mathcal{E}(0)$ for all t .*

In Exercise 6.3, we prove an analogous theorem for a boundary-value, rather than an initial-value, problem.

We can use Theorem 6.4.4 to show that the solution of the problem given by equations (6.4.1) and (6.4.2) is unique.¹³ Since equation (6.4.1) is linear, a solution can be composed of a difference of two arbitrary

¹²Readers interested in the energy of function $u(\mathbf{x}, t)$ for a three-dimensional wave equation, for a frequency-dispersive wave equation, and for the dissipative wave equation might refer to McOwen, R.C., (1996) Partial differential equations: Methods and applications: Prentice-Hall, Inc., pp. 91 – 92, 95 – 96 and 96 – 97, respectively.

¹³Readers interested in the uniqueness of solution of the wave equation and finiteness of the propagation speed in the context of conservation of energy might refer to Taylor, M.E., (1996) Partial differential equations; Basic theory: Springer-Verlag, pp. 140 – 148.

solutions. Each arbitrary solution must obey the initial conditions. Examining the first of conditions (6.4.2), we see that — at $t = 0$ — each solution is $\gamma(x)$. This means that — at $t = 0$ — the solution composed of a difference of two solutions is zero. If the difference of two arbitrary solutions is zero, these two solutions are equal to one another; in other words, the solution is unique at $t = 0$. Now, using Theorem 6.4.4, we wish to verify this uniqueness for $t > 0$.

Following definition (6.4.17) and using the fact that, at $t = 0$, the solution composed of a difference of two solutions is $u(x, t) = 0$, we see that the corresponding $\mathcal{E}(0) = 0$. Then, Theorem 6.4.4 states that $\mathcal{E}(t) = \mathcal{E}(0) = 0$, for all t . Invoking definition (6.4.17), we can explicitly write

$$\mathcal{E}(t_1) = \frac{1}{2} \int_{-\infty}^{\infty} \left[\left(\frac{\partial u}{\partial t} \Big|_{t=t_1} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t=t_1} \right)^2 \right] dx = 0,$$

where t_1 is an arbitrary time, $t_1 \in (0, \infty)$. Since $(\partial u / \partial t)^2$, v^2 and $(\partial u / \partial x)^2$ are positive for all t and x , for the integral to vanish we require that

$$\frac{\partial u(x, t)}{\partial t} = 0$$

and

$$\frac{\partial u(x, t)}{\partial x} = 0,$$

for any x and for arbitrary t . This implies that

$$\frac{du(x, t)}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} v + \frac{\partial u}{\partial t} = 0,$$

where we used the fact that v is a finite velocity of propagation. Hence, we can say that solution u is constant. Since u is unique at $t = 0$ while being constant for all $t \in (0, \infty)$, it must be unique for all t .

Thus we can conclude with the following corollary of Theorem 6.4.4.

COROLLARY 6.4.5. *Solution $u(x, t)$ of wave equation (6.4.1) together with conditions (6.4.2) is unique.*

6.4.4. Causality, finite propagation speed and sharpness of signals. Let us continue to examine solution (6.4.11), namely,

$$(6.4.20) \quad u(x, t) = \frac{1}{2} \left[\gamma(x + vt) + \gamma(x - vt) + \frac{1}{v} \int_{x-vt}^{x+vt} \eta(\zeta) d\zeta \right].$$

For given (x, t) , the value of solution u depends on the values of γ at points $x \mp vt$ and on the values of η on interval $[x - vt, x + vt]$. Hence, this interval is the domain of dependence of $u(x, t)$, as illustrated in Figure 6.4.1. To examine the effect of a point source, at x_0 , on a location, say x_1 , let us consider this figure. Point x_0 belongs to the domains of dependence of points (x, t) that lie in a triangular region whose apex is at x_0 and whose sides have the slopes of $\mp 1/v$. This region is the range of influence of x_0 on solutions $u(x, t)$. The signal generated at x_0 at $t = 0$ will not reach x_1 until time t_1 . Prior to that instant, the source has no effect at x_1 . Thus, we conclude that the process is causal and the propagation speed of

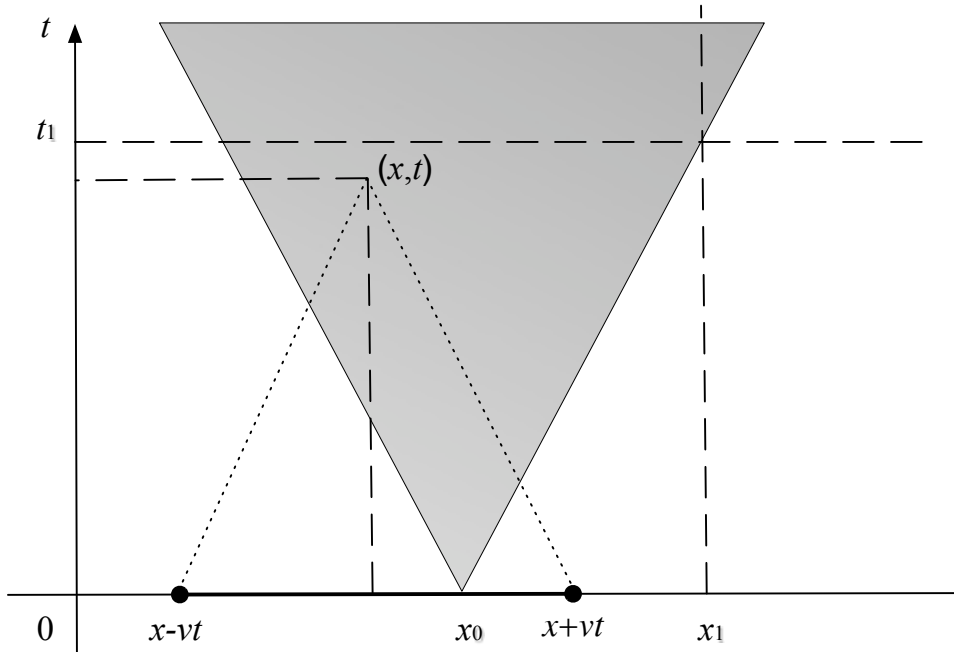


FIGURE 6.4.1. The domain of dependence and the range of influence for the wave equation in a single spatial dimension

the signal is finite; its magnitude is $(x_1 - x_0) / t_1$. There is another important consequence of the range of influence. At x_1 , the effect of the signal is observed not only at t_1 but also afterwards. This means that, in general, a sharp signal generated at x_0 does not propagate as a sharp signal — its effect persists after t_1 . It is interesting to note that for the acoustic case, where u represents pressure, the constant value of u is not audible since our hearing relies on the change in pressure. Thus, we would hear only the initial arrival of the signal — a sharp-signal effect. For the elastic case, where u represents displacement, the constant value is observable.

As stated above, in general, a sharp signal generated at a point does not propagate as a sharp signal. It does so in particular cases, however. By setting the initial displacement velocity, η , to zero, we rewrite solution (6.4.20) as

$$u(x, t) = \frac{1}{2} [\gamma(x + vt) + \gamma(x - vt)];$$

a common example of $\eta = 0$ is the case of a taut string that is pulled, and then released. Herein, the value of the solution depends on points $x \mp vt$ only. Hence, if there is no initial displacement velocity, the effect of the signal does not persist after the signal passes; consequently, sharp signals can propagate.

In Section 6.5, we will see that the solution of the wave equation changes with the spatial dimension of the problem being considered; unlike in the cases of one and two dimensions, three-dimensional media

allow for the propagation of sharp signals. Such a change of behaviour is a particular property of the wave equation; solutions of the heat equation and the steady-state equation, commonly known as Laplace's equation, do not exhibit such changes in physical interpretation due to dimensions.¹⁴

6.5. Solution of wave equation for two and three spatial dimensions

6.5.1. Introductory comments. Having obtained the solution of the wave equation in one spatial dimension in Section 6.4, we wish to investigate the solutions of the wave equation in two and three spatial dimensions. Let us consider

$$(6.5.1) \quad \nabla^2 u(\mathbf{x}, t) - \frac{1}{v^2} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0,$$

where ∇^2 is Laplace's operator, and the corresponding initial conditions; namely,

$$(6.5.2) \quad u(\mathbf{x}, t)|_{t=0} = \gamma(\mathbf{x})$$

and

$$(6.5.3) \quad \left. \frac{\partial u(\mathbf{x}, t)}{\partial t} \right|_{t=0} = \eta(\mathbf{x}).$$

This is a general form of the initial-value problem stated on page 154.

6.5.2. Three spatial dimensions. To consider the case of three spatial dimensions, we write equation (6.5.1) explicitly as

$$(6.5.4) \quad \sum_{i=1}^3 \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial x_i^2} - \frac{1}{v^2} \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial t^2} = 0.$$

To solve this equation, we take its Fourier's transform and the transform of conditions (6.5.2) and (6.5.3) with \mathbf{x} and \mathbf{k} being the transformation variables. As shown in Exercise 6.4, we obtain

$$(6.5.5) \quad \frac{\partial^2 \tilde{u}(\mathbf{k}, t)}{\partial t^2} + v^2 |\mathbf{k}|^2 \tilde{u}(\mathbf{k}, t) = 0.$$

The corresponding initial conditions are

$$(6.5.6) \quad \tilde{u}(\mathbf{k}, t)|_{t=0} = \tilde{\gamma}(\mathbf{k})$$

and

$$(6.5.7) \quad \left. \frac{\partial \tilde{u}(\mathbf{k}, t)}{\partial t} \right|_{t=0} = \tilde{\eta}(\mathbf{k}).$$

¹⁴Readers interested in the concept of the range of influence, including an insightful physical consequences of these ranges for the wave equation, the diffusion equation, and Laplace's equation, might refer to Abbott, M.B., (1966) An introduction to the method of characteristics: Elsevier, pp. 16 – 18 as well as p. 66 and p. 70.

As shown in Exercise 6.5, the general solution of equation (6.5.5) is

$$(6.5.8) \quad \tilde{u}(\mathbf{k}, t) = F(\mathbf{k}) \exp\{iv|\mathbf{k}|t\} + G(\mathbf{k}) \exp\{-iv|\mathbf{k}|t\}.$$

To get expressions for F and G , we use conditions (6.5.6) and (6.5.7). Also, as shown in Exercise 6.5, we get

$$(6.5.9) \quad F(\mathbf{k}) = \frac{1}{2} \left(\tilde{\gamma}(\mathbf{k}) + \frac{1}{iv|\mathbf{k}|} \tilde{\eta}(\mathbf{k}) \right)$$

and

$$(6.5.10) \quad G(\mathbf{k}) = \frac{1}{2} \left(\tilde{\gamma}(\mathbf{k}) - \frac{1}{iv|\mathbf{k}|} \tilde{\eta}(\mathbf{k}) \right).$$

Inserting these expressions into solution (6.5.8), we write

$$(6.5.11) \quad \tilde{u}(\mathbf{k}, t) = \frac{1}{2} \left(\left(\tilde{\gamma}(\mathbf{k}) + \frac{1}{iv|\mathbf{k}|} \tilde{\eta}(\mathbf{k}) \right) \exp\{iv|\mathbf{k}|t\} + \left(\tilde{\gamma}(\mathbf{k}) - \frac{1}{iv|\mathbf{k}|} \tilde{\eta}(\mathbf{k}) \right) \exp\{-iv|\mathbf{k}|t\} \right),$$

which is the solution of the initial-value problem given by equations (6.5.5), (6.5.6) and (6.5.7). To obtain the solution of the initial-value problem given by equations (6.5.4), (6.5.2) and (6.5.3), we need to find the inverse Fourier's transform of solution (6.5.11). To do so, let us rewrite the solution by factoring out the two initial conditions; namely,

$$\tilde{u}(\mathbf{k}, t) = \tilde{\gamma}(\mathbf{k}) \frac{\exp\{iv|\mathbf{k}|t\} + \exp\{-iv|\mathbf{k}|t\}}{2} + \tilde{\eta}(\mathbf{k}) \frac{\exp\{iv|\mathbf{k}|t\} - \exp\{-iv|\mathbf{k}|t\}}{2iv|\mathbf{k}|}.$$

Examining the two fractions, we notice that they are related as follows:

$$\frac{d}{dt} \frac{\exp\{iv|\mathbf{k}|t\} - \exp\{-iv|\mathbf{k}|t\}}{2iv|\mathbf{k}|} = \frac{\exp\{iv|\mathbf{k}|t\} + \exp\{-iv|\mathbf{k}|t\}}{2}.$$

Using this relation, we write

$$(6.5.12) \quad \tilde{u}(\mathbf{k}, t) = \tilde{\gamma}(\mathbf{k}) \frac{d}{dt} \frac{\exp\{iv|\mathbf{k}|t\} - \exp\{-iv|\mathbf{k}|t\}}{2iv|\mathbf{k}|} + \tilde{\eta}(\mathbf{k}) \frac{\exp\{iv|\mathbf{k}|t\} - \exp\{-iv|\mathbf{k}|t\}}{2iv|\mathbf{k}|}.$$

Since each term is a product of two functions, we will invoke the fact that a product in the \mathbf{k} -domain is a convolution in the \mathbf{x} -domain. As shown in Exercise 6.6, the fractions in solution (6.5.12) are the transforms of the distribution given by

$$(6.5.13) \quad \frac{(2\pi)^3}{4\pi v^2 t} \iint_{S(\mathbf{0}, vt)} \tau d\zeta = \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{0}, vt)} \tau d\zeta,$$

where τ is a test function and $d\zeta$ is the surface element on the sphere, $S(\mathbf{0}, vt)$, that is centred at $\mathbf{x} = \mathbf{0}$ and whose radius is vt . In other words, expression (6.5.13) is the inverse transform of the fraction in solution (6.5.12). As formulated in Exercise 6.7, using expression (6.5.13) and the properties of convolution, we

obtain the solution of the initial-value problem given by equations (6.5.4), (6.5.2) and (6.5.3); namely,

$$(6.5.14) \quad u(\mathbf{x}, t) = \frac{d}{dt} \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{x}, vt)} \gamma(\mathbf{y}) d\zeta(\mathbf{y}) + \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{x}, vt)} \eta(\mathbf{y}) d\zeta(\mathbf{y}),$$

where \mathbf{y} are the variables of integration on the sphere centred at \mathbf{x} .

Examining expression (6.5.14), we see that for any point in the $x_1 x_2 x_3 t$ -space, solution $u(\mathbf{x}, t)$ depends on the values of γ and η whose domain is the surface of a sphere in the $x_1 x_2 x_3$ -space that is centred at the \mathbf{x} coordinates of that point, and whose radius is vt . In a manner analogous to the one discussed in Section 6.4.4, we conclude that a point source in the $x_1 x_2 x_3$ -space influences the points in the $x_1 x_2 x_3 t$ -space that are on the surface of the three-dimensional ‘cone’ embedded in four dimensions whose apex is the source and whose slope is $1/v$. Thus, a signal emitted at $t = 0$ will arrive at a point located at a distance d from the source at time $t = d/v$. Prior to that instant, there is no effect of the source, which means that the process is causal and the speed of signal propagation is finite. At that instant, the value of u is finite since both integrals in expression (6.5.14) are bounded. Afterwards, again there is no signal. The signal is confined to the spherical shell of radius vt , which is the propagating wavefront. Thus, sharp signals can propagate in three dimensions, which is in agreement with our experience.

6.5.3. Two spatial dimensions. Mathematically, the solution of the wave equation in two dimensions can be viewed as a particular case of the three-dimensional case. Physically, however, the two results have important distinctions.

To investigate the case of two spatial dimensions, we return to equation (6.5.1) and write it explicitly as

$$(6.5.15) \quad \sum_{i=1}^2 \frac{\partial^2 u(x_1, x_2, t)}{\partial x_i^2} - \frac{1}{v^2} \frac{\partial^2 u(x_1, x_2, t)}{\partial t^2} = 0,$$

where the corresponding initial conditions are given by expressions (6.5.2) and (6.5.3) with $\mathbf{s} = [s_1, s_2, 0]$, where \mathbf{s} stands for coordinates of the three-dimensional space. Using solution (6.5.14), we can write the solution of equation (6.5.15) as

$$u(\mathbf{x}, t) = \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{x}, vt)} \eta(s_1, s_2) d\zeta(\mathbf{s}) + \frac{d}{dt} \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{x}, vt)} \gamma(s_1, s_2) d\zeta(\mathbf{s}).$$

As shown in Exercise 6.8, we can rewrite this solution as

$$(6.5.16) \quad u(x_1, x_2, t) = \frac{d}{dt} \left(\frac{4\pi^2}{v} \iint_{D(\mathbf{x}, vt)} \frac{\gamma(s_1, s_2)}{\sqrt{(vt)^2 - [(s_1 - x_1)^2 + (s_2 - x_2)^2]}} ds_1 ds_2 \right) \\ + \frac{4\pi^2}{v} \iint_{D(\mathbf{x}, vt)} \frac{\eta(s_1, s_2)}{\sqrt{(vt)^2 - [(s_1 - x_1)^2 + (s_2 - x_2)^2]}} ds_1 ds_2,$$

where $D(\mathbf{x}, vt)$ is the disc that is centred at \mathbf{x} and whose radius is vt .

Examining expression (6.5.16), we see that for any point in the x_1x_2t -space, solution $u(\mathbf{x}, t)$ depends on the values of γ and η whose domain is a disc on the x_1x_2 -plane that is centred at the \mathbf{x} coordinates of that point, and whose radius is vt . In a manner analogous to the one discussed in Section 6.4.4, we conclude that a point source on the x_1x_2 -plane influences the points in the x_1x_2t -space that are both on the surface of and contained within the right circular cone whose apex is the point source and whose slope is $1/v$. The projections onto the x_1x_2 -plane of a circular section of this cone is the wavefront at a given time. Thus, a signal emitted at $t = 0$ will arrive at a point located at a distance d from the source at time $t = d/v$. Prior to that instant, there is no effect of the source, which means that the process is causal and the speed of signal propagation is finite. At that instant, the value of u is infinite since the bracketed terms of the radicands of equation (6.5.16) are equal to $(vt)^2$, and hence the two integrals in expression (6.5.16) are unbounded. Afterwards, the value of u diminishes but remains nonzero forever.¹⁵ Hence, sharp signals do not propagate in two dimensions, which we can visualize by picturing a pebble dropped into a pond: the surface is affected by this disturbance after the wavefront has passed.¹⁶

Thus for all the three cases discussed in Sections 6.4 and 6.5, the process of wave propagation described by the solution of the initial-value problem given by expressions (6.5.1), (6.5.2) and (6.5.3) is causal and the speed of propagation is finite. Also, as shown in Section 6.4.3, the solution of this initial-value problem is unique and stable. However, sharp signals propagate only in three dimensions.

6.6. On evolution equation

To gain further insight into the properties of solutions of the wave equation and its initial conditions, let us consider a first-order time-evolution equation in a single spatial dimension, namely,

$$(6.6.1) \quad \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{\partial u(x, t)}{\partial t} = 0,$$

and the initial condition given by $u(x, t)|_{t=0} = \gamma(x)$. As we will see, the solution of this initial-value problem is causal but its speed of propagation is infinite. Furthermore, the mapping $\gamma \rightarrow u$ does not have a unique inverse in the limit as $t \rightarrow \infty$.

¹⁵Readers interested in a quantitative description of fading of the value of u as recorded by the observer in the two-dimensional case might refer to Barton, G., (1989) Elements of Green's functions and propagation: Potentials, diffusion and waves: Oxford Science Publications, pp. 278 – 282.

¹⁶Readers interested in the range of influence for the wave propagation in the two-dimensional and three-dimensional media, including the consequences for the corresponding wave equations on sharpness of the signal as well as on Huygens' principle, might refer to Kline, M., (1972) Mathematical thought from ancient to modern times: Oxford University Press, Vol 2, pp. 690 – 691, to McOwen, R.C., (1996) Partial differential equations: Methods and applications: Prentice-Hall, Inc., pp. 83 – 90 and pp. 130 – 131, and to Renardy, M., and Rogers, R.C., (1993) An introduction to partial differential equations: Springer-Verlag, pp. 158 – 159. Also, an insightful description of the wave equation in one spatial dimension is presented in Spivak, M., (1970/1999) A comprehensive introduction to differential geometry: Publish or Perish, Inc., Vol. V., pp. 68 – 71.

To solve this initial-value problem, we will investigate its Fourier's transform in a manner analogous to the one discussed in Section 6.5.2. We get

$$\frac{\partial \tilde{u}(k, t)}{\partial t} + k^2 \tilde{u}(k, t) = 0,$$

which we can solve as if it were a first-order ordinary differential equation in t , with initial condition $\tilde{u}(k, t)|_{t=0} = \tilde{\gamma}(k)$. Hence, in the transformed domain, the solution is

$$\tilde{u}(k, t) = \tilde{\gamma}(k) \exp \{-k^2 t\}.$$

To obtain the solution of the initial-value problem, we proceed to find the inverse transform. We write

$$u(x, t) = \int_{-\infty}^{\infty} \tilde{\gamma}(k) \exp \{-k^2 t\} \exp \{ikx\} dk.$$

Invoking the definition of Fourier's transform, we write

$$\tilde{\gamma}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma(\zeta) \exp \{-i\zeta k\} d\zeta,$$

where ζ is the integration variable. Inserting this result into the above expression for u , we get

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(\zeta) \exp \{-i\zeta k\} \exp \{-k^2 t\} \exp \{ikx\} d\zeta dk.$$

Combining the exponential terms, we write

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(\zeta) \exp \{i(x - \zeta)k - k^2 t\} d\zeta dk.$$

Exchanging the order of integration and since γ is independent of k , we write

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \exp \{i(x - \zeta)k - k^2 t\} dk \right) \gamma(\zeta) d\zeta.$$

As shown in Exercise 6.9, the inner integral can be written as

$$\int_{-\infty}^{\infty} \exp \{i(x - \zeta)k - k^2 t\} dk = \frac{\exp \left\{ -\frac{(x - \zeta)^2}{4t} \right\}}{\sqrt{t}} \int_{-\infty}^{\infty} e^{-s^2} ds,$$

where s is the variable of integration. Evaluating the definite integral, as shown in Exercise 6.10, we write the inner integral as

$$\int_{-\infty}^{\infty} \exp \{i(x - \zeta)k - k^2 t\} dk = \sqrt{\frac{\pi}{t}} \exp \left\{ -\frac{(x - \zeta)^2}{4t} \right\}.$$

Returning to the above expression for u , we write

$$(6.6.2) \quad u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} \exp \left\{ -\frac{(x - \zeta)^2}{4t} \right\} \gamma(\zeta) d\zeta,$$

which is the solution of the initial-value problem, as stated by the following theorem.

THEOREM 6.6.1. *If $\gamma(x)$ is bounded and continuous on $x \in \mathbb{R}$, then $u(x, t)$ given in expression (6.6.2) is an infinitely differentiable function satisfying — for all $x \in \mathbb{R}$ and $t > 0$ — the initial-value problem of equation (6.6.1), and extends continuously to $t = 0$, where $u(x, t) = \gamma(x)$.*

Let us consider the domain of dependence of solution (6.6.2), and the range of influence of a given point on the x -axis. Examining expression (6.6.2) in view of Theorem 6.6.1, we see that for any (x, t) on the xt -plane where $t > 0$, solution $u(x, t)$ depends on the value of function γ along the entire x -axis. In other words, the domain of dependence is the entire x -axis. By reciprocity, any point on the x -axis influences all the locations on the xt -plane where $t > 0$. In other words, this entire halfplane is the range of influence of such a point. This means that the effect of the initial condition, $\gamma(x)$, upon the solution, $u(x, t)$, is instantaneous; in other words, the solution has an infinite propagation speed. Hence, even if $\gamma(x)$ is confined to a finite domain along the x -axis, its effect for $t > 0$ covers the entire axis. Also, this means that γ affects all points in that halfplane from $t = 0$ to $t = \infty$.

Since, in view of Theorem 6.6.1, solution (6.6.2) is not valid for $t < 0$, the process is causal. Also the forward solution is unique. However, any bounded γ with compact support results in $u \rightarrow 0$ as $t \rightarrow \infty$; hence, we cannot determine the initial state, γ , from u at $t \rightarrow \infty$, which means that the inverse to the mapping $\gamma \rightarrow u$ is not unique at infinity — the inverse problem is ill-posed. We note that the wave equation and its initial conditions constitute well-posed forward and inverse problems.

6.7. Solutions of wave equation for one-dimensional scattering

To introduce wave propagation in inhomogeneous media, let us study a one-dimensional continuum where the wave propagates with speed v_1 at $x \leq 0$ and v_2 at $x > 0$. In view of solution (6.4.6) for the wave equations in the aforementioned continuum, namely,

$$(6.7.1) \quad \begin{cases} \frac{\partial^2 u}{\partial x^2} - \frac{1}{v_1^2} \frac{\partial^2 u}{\partial t^2} = 0, & x \leq 0 \\ \frac{\partial^2 u}{\partial x^2} - \frac{1}{v_2^2} \frac{\partial^2 u}{\partial t^2} = 0, & x > 0 \end{cases},$$

the general solution is

$$(6.7.2) \quad u(x, t) = \begin{cases} f_1(x + v_1 t) + g_1(x - v_1 t) & x \leq 0 \\ f_2(x + v_2 t) + g_2(x - v_2 t) & x > 0 \end{cases}.$$

To examine this problem, let us consider only a wave propagating from the left towards the origin, $x = 0$. The effects associated with $x = 0$ are called scattering since at this point the incident wave is separated

into a reflected wave and a transmitted wave, as we will see below. We can write the expression for the incident wave as

$$(6.7.3) \quad w = w(x - v_1 t),$$

where w is an arbitrary function that we set to zero for $x > 0$; hence, any wave to the right of the origin will be the transmitted wave: the wave that went across $x = 0$.

In order for expression (6.7.3) to represent the initial disturbance, we formulate the initial conditions corresponding to equations (6.7.1) as

$$(6.7.4) \quad u|_{t=0} = w(x)$$

and

$$(6.7.5) \quad \left. \frac{\partial u}{\partial t} \right|_{t=0} = -v_1 w'(x),$$

where w' denotes the derivative of w .

Considering solution (6.7.2) we see that $f_1 + g_1$ describes the incident and reflected waves, and $f_2 + g_2$ describes the transmitted wave. To write it specifically in the context of conditions (6.7.4) and (6.7.5), we restate these conditions as

$$w(y) = u(y, 0) = \begin{cases} f_1(y) + g_1(y), & y \leq 0 \\ f_2(y) + g_2(y), & y > 0 \end{cases}$$

and

$$-v_1 w'(y) = \left. \frac{\partial u}{\partial t} \right|_{t=0} = \begin{cases} v_1 [f_1'(y) - g_1'(y)], & y \leq 0 \\ v_2 [f_2'(y) - g_2'(y)], & y > 0 \end{cases},$$

respectively; to avoid the confusion between the position, x , and the argument of a function, we use y as the argument. Recalling from expression (6.7.3) that at $t = 0$ and to the right of the origin, $w \equiv 0$, we rewrite these conditions as

$$u(y, 0) = \begin{cases} f_1(y) + g_1(y) = w(y), & y \leq 0 \\ f_2(y) + g_2(y) = 0, & y > 0 \end{cases}$$

and

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = \begin{cases} f_1'(y) - g_1'(y) = -w'(y), & y \leq 0 \\ f_2'(y) - g_2'(y) = 0, & y > 0 \end{cases}.$$

To solve these equations, we group them according to the domains of the functions. Thus, we write

$$\begin{cases} f_1(y) + g_1(y) = w(y), \\ f_1'(y) - g_1'(y) = -w'(y), \end{cases} \quad y \leq 0$$

and

$$\begin{cases} f_2(y) + g_2(y) = 0, \\ f_2'(y) - g_2'(y) = 0, \end{cases} \quad y > 0.$$

Integrating the second equations of both sets, we obtain

$$\begin{cases} f_1(y) + g_1(y) = w(y), \\ f_1(y) - g_1(y) = -w(y) + A, \end{cases} \quad y \leq 0$$

and

$$\begin{cases} f_2(y) + g_2(y) = 0, \\ f_2(y) - g_2(y) = B, \end{cases} \quad y > 0,$$

where A and B are integration constants. We have two systems of equations. Each system consists of two equations in two unknowns. We can proceed to solve these systems in their corresponding domains of y .

Let us consider the nonpositive arguments. We add the two equations of the first system to get $f_1(y) = A/2$. Hence, $g_1(y) = w(y) - A/2$. Since f_1 is a constant, its value can be incorporated into g_1 ; in other words, we can let $f_1 = A/2 \equiv 0$, without loss of generality. Thus, $g_1(y) = w(y)$.

Let us consider the positive arguments. We add the two equations of the second system to get $f_2(y) = B/2$. Herein, the constancy of f_2 implies the constancy of g_2 , and we can let $f_2(y) = g_2(y) \equiv 0$, without loss of generality.

Using these results, the fact that $v_1 > 0$, $v_2 > 0$ and $t \geq 0$, and examining solution (6.7.2), we will study the arguments of f_1 , g_1 , f_2 and g_2 to rewrite the general solution. Let us consider $x \leq 0$. The argument of f_1 is $y = x + v_1 t$, which can be either positive or negative. Hence, even though $f_1 \equiv 0$ for negative arguments, we must consider it in the solution. The argument of g_1 is $y = x - v_1 t$, which is negative. Hence, $g_1 = w$. Let us consider $x > 0$. The argument of f_2 is $y = x + v_2 t$, which is positive. Hence, we set $f_2 \equiv 0$. The argument of g_2 is $y = x - v_2 t$, which can be either positive or negative. Hence, even though $g_2 \equiv 0$ for positive arguments, we must consider it in the solution. Thus, we can rewrite expression (6.7.2) as

$$(6.7.6) \quad u(x, t) = \begin{cases} w(x - v_1 t) + f_1(x + v_1 t), & x \leq 0 \\ g_2(x - v_2 t), & x > 0 \end{cases}.$$

Physically, we can interpret expression (6.7.6) in the following way: w is the incident wave, f_1 propagates in the opposite direction along the same domain of x — it is the reflected wave, g_2 propagates in the same direction as w but on the other side of $x = 0$ — it is the transmitted wave; $f_2 \equiv 0$ since there is no wave propagating towards the origin from the right.

In view of the physical interpretation, we could have argued the form of expression (6.7.6) directly from expression (6.7.2) upon introducing expression (6.7.3). However, the rigorous approach presented above provides us with an insight into the relation among the differential equation, its initial conditions and solution.

In expression (6.7.6), f_1 and f_2 are arbitrary functions. However, by invoking physical constraints, we will be able to relate them to w ; in other words, we will obtain relations among the incident, reflected and transmitted waves.

In view of physical considerations, which will be discussed in detail in Section 10.2, we require that both the displacement, u , and the stress be continuous along the entire x -axis. In particular, they must be continuous across $x = 0$. Using expression (6.7.6), we can state the former requirement at $x = 0$ as the equality given by

$$(6.7.7) \quad f_1(v_1 t) + w(-v_1 t) = g_2(-v_2 t).$$

To state the latter requirement, we recall stress-strain equations (4.2.8) with elasticity matrix (5.12.2), which corresponds to an isotropic continuum. These stress-strain equations are given by equations (3.2.1), namely,

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

In our present study of a one-dimensional medium that coincides with the x -axis and considering only longitudinal displacements, we use $\sigma_{xx} := \sigma_{11}$ and $\varepsilon_{xx} := \varepsilon_{11}$ to write the stress-strain equation for this one-dimensional continuum as

$$\sigma_{xx} = (\lambda + 2\mu) \varepsilon_{xx},$$

which, in view of definition (1.4.6), we can restate as

$$\sigma_{xx} = (\lambda + 2\mu) \frac{\partial u}{\partial x}.$$

Furthermore, using expression (6.1.13), we can rewrite it as

$$(6.7.8) \quad \sigma_{xx} = \rho v^2 \frac{\partial u}{\partial x}.$$

Following expression (6.7.8), we can state the continuity of stress across $x = 0$ as the equality given by

$$\rho_1 v_1^2 \left. \frac{\partial u_1}{\partial x} \right|_{x=0} = \rho_2 v_2^2 \left. \frac{\partial u_2}{\partial x} \right|_{x=0},$$

where u_1 and u_2 are the adjacent displacements on either side of $x = 0$. Differentiating the appropriate terms of expression (6.7.6), we can explicitly write this continuity condition as

$$(6.7.9) \quad \rho_1 v_1^2 [f_1'(v_1 t) + w'(-v_1 t)] = \rho_2 v_2^2 g_2'(-v_2 t),$$

where symbol $'$ denotes the derivative of a given function with respect to its argument.

We wish to solve equations (6.7.7) and (6.7.9) for f_1 and g_2 in terms of w . In other words, we wish to express the reflected and transmitted waves in terms of the incident wave. To do so, let us impose the equality of time derivatives of displacement at $x = 0$; in other words, the speed of displacement is continuous across $x = 0$ — just like the displacement itself, as stated in equation (6.7.7). Thus, differentiating

equations in system (6.7.6) with respect to t and equating the results, we obtain

$$(6.7.10) \quad v_1 f_1'(v_1 t) - v_1 w'(-v_1 t) = -v_2 g_2'(-v_2 t).$$

Solving for f_1' , we get

$$f_1'(v_1 t) = w'(-v_1 t) - \frac{v_2}{v_1} g_2'(-v_2 t).$$

Substituting into equation (6.7.9) and simplifying, we get

$$g_2'(-v_2 t) = \frac{2\rho_1 v_1^2}{v_2 (\rho_1 v_1 + \rho_2 v_2)} w'(-v_1 t).$$

Integrating both sides with respect to t and simplifying, we get

$$g_2(-v_2 t) = \frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2} w(-v_1 t),$$

where, in view of the initial conditions, the integration constant is zero. Herein, we set $y = -v_2 t$; hence, $t = -y/v_2$ and the argument of w becomes $(v_1/v_2)y$. Thus, we can write

$$g_2(y) = \frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2} w\left(\frac{v_1}{v_2} y\right),$$

where product $\rho_i v_i$ is called the acoustic impedance. Similarly, as shown in Exercise 6.11, we obtain

$$f_1(y) = \frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2} w(-y),$$

where $y = v_1 t$.

Thus, we can write expression (6.7.6) as

$$(6.7.11) \quad u(x, t) = \begin{cases} w(x - v_1 t) + \frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2} w(-(x + v_1 t)), & x \leq 0 \\ \frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2} w\left(\frac{v_1}{v_2}(x - v_2 t)\right), & x > 0 \end{cases},$$

which is the solution of our scattering problem stated in terms of the incident wave, w , and the properties of the discontinuous medium given by ρ_1 , v_1 and ρ_2 , v_2 .

Having obtained the appropriate form of the solution, let us interpret its physical meaning. For $x \leq 0$, the two terms correspond to the incident and the reflected waves, respectively. For $x > 0$, the solution corresponds to the transmitted wave.

The factors in front of w are the amplitudes of given waves. Thus, the incident wave has amplitude set to unity. Examining expressions

$$(6.7.12) \quad \frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2} =: A_r$$

and

$$(6.7.13) \quad \frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2} =: A_t,$$

which are the amplitudes of the reflected and transmitted waves, respectively, we see that, if $\rho_1 v_1 \neq \rho_2 v_2$, the amplitude of the reflected wave is always smaller than the amplitude of the incident wave, while — in agreement with conservation of energy — the amplitude of the transmitted wave can be either smaller or greater than the amplitude of the incident wave; this is further discussed in Section 10.2.2 and Exercise 10.5. Furthermore, expressions (6.7.12) and (6.7.13) are derived in a different way in Exercise 10.4.

The factor v_1/v_2 in the transmitted wave corresponds to the wavelength: The wavelength of the incident wave is v_1/v_2 times the wavelength of the transmitted wave.

Examining solution (6.7.11), we see that upon transmission from a lower to a higher acoustic impedance, the amplitude of the wave decreases. We also see that explicitly mass density has no effect on the wavelength; implicitly, it is contained in the expressions for the speeds of propagation, v_1 and v_2 . Upon transmission from a lower to a higher speed, the wavelength decreases.

To complete this section, let us consider a particular case of solution (6.7.11). If $\rho_1 v_1 = \rho_2 v_2 = \rho v$, the solution reduces to

$$u(x, t) = w(x - vt),$$

for all x , as expected. In other words, if there is no change in acoustic impedance, the wave travels without any scattering.

6.8. On weak solutions of wave equation

6.8.1. Introductory comments. In a classical approach to differential equations, we expect the differentiability of solutions to, at least, match the order of the equation. Yet, examining solution (6.4.6), without considering equation (6.4.1), we see that the solution itself does not require f and g to be differentiable; f and g could be even discontinuous. Requirements for differentiability of the solution of the wave equation were the subject of long discussions between d'Alembert and Euler in the second half of the eighteenth century. In these discussions, which lasted for almost thirty years, Euler deemed it necessary from physical considerations of wave propagation to admit nondifferentiable functions as solutions, while d'Alembert strictly required differentiability. A rigorous formulation that allows us to consider nondifferentiable solutions was not available until the middle of the twentieth century. In this section, we will briefly study this formulation.

A standard method of verifying that solution (6.4.6) satisfies equation (6.4.1) consists of inserting this solution into the equation. To do so, however, we require that f and g be twice-differentiable. Sergei Sobolev, during his presentation in 1934, stated that

[t]he class of functions that we can consider as solutions to the wave equation from the classical point of view consists of twice-differentiable functions. But in various practical applications it seems convenient to consider functions with well-defined singularities.

In other words, we might wish to investigate waves that cannot be described by twice-differentiable functions. Consequently, we would like to extend the solutions of wave equation (6.4.1) to incorporate functions that are not differentiable. To do so, we consider the theory of generalized functions, commonly referred to as the theory of distributions, which was formulated by Sergei Sobolev in the first half of the twentieth century and thoroughly developed by Laurent Schwartz in the middle of the twentieth century. The key philosophical point of this theory is that a generalized function is not described by itself alone but by its effect on other functions.

We define the effect of $h(x)$ on $\tau(x)$ by the value of $\int_{-\infty}^{\infty} h(x) \tau(x) dx$, where τ is assumed to be infinitely differentiable and compactly supported. In view of this definition, τ is often called the test function. In this formulation, we do not require h to be differentiable. Furthermore, h need not be a function in the classical sense of the term; in general, h is a generalized function, also called a distribution; Dirac's delta is a famous example of such an entity.

Expression $\int_{-\infty}^{\infty} h(x) \tau(x) dx$ generalizes our study of functions. If h is a function in the classical sense, we can describe it on \mathbb{R}^1 by its values, $h(x)$, for appropriate points $x \in \mathbb{R}^1$, as well as by the values of $\int_{-\infty}^{\infty} h\tau dx$ for appropriate functions τ . If h is not a function, we cannot describe it by its values, but we can still describe it by the values of $\int_{-\infty}^{\infty} h\tau dx$, as exemplified by Dirac's delta, where $\delta(x)$ by itself alone does not make sense, while $\int_{-\infty}^{\infty} \delta(x) \tau(x) dx = \tau(0)$ is a well-defined quantity.

6.8.2. Weak derivatives. To study the solutions of the wave equation, which is a partial differential equation, we wish to investigate derivatives of nondifferentiable functions; so called weak derivatives. To do so, let us consider the effect of these derivatives on test functions. Consider $\int_{-\infty}^{\infty} h'(x) \tau(x) dx$. Integrating by parts, we obtain

$$\int_{-\infty}^{\infty} h'(x) \tau(x) dx = h(x) \tau(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} h(x) \tau'(x) dx.$$

Since τ has compact support, $\lim_{x \rightarrow \pm\infty} \tau = 0$; hence, the first term on the right-hand side vanishes, if we consider h to be finite at $\pm\infty$. Thus, we can state the derivative of h as

$$(6.8.1) \quad \int_{-\infty}^{\infty} h'(x) \tau(x) dx = - \int_{-\infty}^{\infty} h(x) \tau'(x) dx.$$

This is the formula for the weak derivative of order one for the case of a single variable. The use of this formula is exemplified in Exercise 6.12.

In order to obtain higher-order derivatives for the case of n variables, we can repeat the process of integration by parts to arrive at the general equation for an m th partial derivative, which is given by

$$(6.8.2) \quad \int_{\mathbb{R}^n} \frac{\partial^m h(\mathbf{x})}{\partial x_i^m} \tau(\mathbf{x}) d\mathbf{x} = (-1)^m \int_{\mathbb{R}^n} h(\mathbf{x}) \frac{\partial^m \tau(\mathbf{x})}{\partial x_i^m} d\mathbf{x},$$

where $\int_{\mathbb{R}^n} dx$ stands for n integrals over n variables from $-\infty$ to $+\infty$.

Examining equation (6.8.2), we notice the crux of this formulation. To study the effect of the derivatives of h , shown on the left-hand side, we study the expression on the right-hand side where no derivatives of h appear. Hence, we do not need h to be differentiable.

6.8.3. Weak solution of wave equation. We wish to extend the solutions of wave equation (6.4.1) to allow nondifferentiable functions. To distinguish the solutions that can be verified directly from the solutions discussed below, we refer to the latter ones as weak solutions.

In view of Section 6.8.2, let us consider the effect of the second derivatives. We write

$$(6.8.3) \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \left(\frac{\partial^2}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \right) [f(x+vt) + g(x-vt)] \right\} \tau(x, t) dx dt,$$

where the term in parentheses is the differential operator of wave equation (6.4.1) and the term in brackets is solution (6.4.6), where we wish to allow f and g to be nondifferentiable.

To verify that $f + g$ is a weak solution of equation (6.4.1), we must show that expression (6.8.3) vanishes for all test functions, τ . Following equation (6.8.2), we write

$$(6.8.4) \quad \begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \left(\frac{\partial^2}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \right) [f(x+vt) + g(x-vt)] \right\} \tau(x, t) dx dt \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x+vt) + g(x-vt)] \left[\left(\frac{\partial^2}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \right) \tau(x, t) \right] dx dt. \end{aligned}$$

Recalling that τ is infinitely differentiable, we let $y = x + vt$ and $z = x - vt$ and use Lemma 6.4.1 to rewrite the right-hand side of the above equation as

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(y) + g(z)] \left[\frac{\partial^2 \tau(y, z)}{\partial y \partial z} \right] dy dz \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y) \frac{\partial^2 \tau(y, z)}{\partial y \partial z} dy dz + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(z) \frac{\partial^2 \tau(y, z)}{\partial y \partial z} dy dz. \end{aligned}$$

Changing the order of integration, we get

$$\int_{-\infty}^{\infty} f(y) \int_{-\infty}^{\infty} \frac{\partial^2 \tau(y, z)}{\partial y \partial z} dz dy + \int_{-\infty}^{\infty} g(z) \int_{-\infty}^{\infty} \frac{\partial^2 \tau(y, z)}{\partial y \partial z} dy dz.$$

In order to show that wave equation (6.4.1) is satisfied — in the weak sense — by $f + g$, we must show that the above expression vanishes. It suffices to show that

$$\int_{-\infty}^{\infty} \frac{\partial^2 \tau(y, z)}{\partial y \partial z} dz = 0$$

and

$$\int_{-\infty}^{\infty} \frac{\partial^2 \tau(y, z)}{\partial y \partial z} dy = 0.$$

Integrating, we obtain

$$\left. \frac{\partial \tau(y, z)}{\partial y} \right|_{z=-\infty}^{z=\infty} = 0$$

and

$$\left. \frac{\partial \tau(y, z)}{\partial z} \right|_{y=-\infty}^{y=\infty} = 0,$$

respectively. Since τ has compact support, all its derivatives — in particular, $\partial \tau / \partial y$ and $\partial \tau / \partial z$ — also have compact support. Thus, $\lim_{x \rightarrow \pm \infty} \partial \tau / \partial y = \lim_{x \rightarrow \pm \infty} \partial \tau / \partial z = 0$, which means that both above equations are satisfied for all test functions, τ , as required. We have verified solution (6.4.6) of wave equation (6.4.1) without requiring differentiability of f or g . In view of expression (6.8.4), we can define the weak solution of the one-dimensional wave equation as $u(x, t)$ that satisfies

$$(6.8.5) \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x, t) \left[\frac{\partial^2 \tau(x, t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 \tau(x, t)}{\partial t^2} \right] dx dt = 0,$$

where τ is the test function.

Any differentiable solution of the wave equation satisfies also equation (6.8.5). In other words, any strong solution is also a weak solution. However, since the opposite is not true, we need equation (6.8.5) to study solutions that are not differentiable.

6.9. Reduced wave equation

6.9.1. Harmonic-wave trial solution. To motivate the form and the name of the reduced wave equation, let us consider a particular trial solution for the 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 plane-wave solution. However, we might wish to study a more complicated position dependence of the solution; hence, we would require another trial solution.

Let us consider equation (6.4.1), namely,

$$(6.9.1) \quad \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u(x, t)}{\partial t^2} = 0.$$

If we study an oscillatory motion, we can write our trial solution as

$$(6.9.2) \quad u(x, t) = \acute{u}(x) \exp(-i\omega t),$$

where ω stands for the angular frequency and $\exp(-i\omega t)$ is called the phase factor.¹⁷ The right-hand side of expression (6.9.2) is a standard form of a complex number whose magnitude is $\acute{u}(x)$ and whose phase is ωt . We can write solution (6.9.2) as

$$(6.9.3) \quad \acute{u}(x) \exp(-i\omega t) = \acute{u}(x) [\cos(\omega t) - i \sin(\omega t)].$$

By using this trial solution, we assume that the time dependence of the displacement function, $u(x, t)$, is satisfied by the term in brackets in expression (6.9.3). In other words, we assume that the solution is sinusoidal in time; such waves are called harmonic waves.

Inserting solution (6.9.2) into equation (6.9.1), we obtain

$$(6.9.4) \quad \frac{d^2 \acute{u}(x)}{dx^2} + \left(\frac{\omega}{v}\right)^2 \acute{u}(x) = 0,$$

as shown in Exercise 6.13. This is the reduced wave equation. Since wave equation (6.9.1) is in a single spatial dimension and since time is not a variable in equation (6.9.4), we obtained an ordinary differential equation. However, to a certain extent, it is a matter of notation: \acute{u} is the solution of equation (6.9.4) for a particular value of ω ; hence, \acute{u} is implicitly a function of ω . Using this notation, we emphasize that equation (6.9.4) is equation (6.9.1) with the temporal dependence given by $\exp(-i\omega t)$. Such a reduced wave equation allows us to study particular problems, such as the steady-state problems or standing-wave problems. Notably, the reduced wave equation was used by Hermann von Helmholtz in the mid-nineteenth century to study oscillations in the organ pipes, and it belongs now to the class of Helmholtz's equations. Gustav Kirchhoff continued the work of Helmholtz on the reduced wave equation to study the solution of the initial-value problem, which resulted in a mathematical statement of Huygens' principle.¹⁸

In the next section, we will study the wave equation using Fourier's transform. Therein, we will obtain a partial differential equation that is almost equivalent to equation (6.9.4).

6.9.2. Fourier's transform of wave equation. Writing equation (6.9.4) as

$$\frac{\partial^2 \acute{u}}{\partial x^2} + \left(\frac{\omega}{v}\right)^2 \acute{u} = 0,$$

we notice that we could view this equation as Fourier's transform of equation (6.9.1) with t and ω being the transformation variables. Let us investigate this property.

¹⁷In this book, $\exp(\cdot)$ and $e^{(\cdot)}$ are used as synonymous notations.

¹⁸Readers interested in the role of the reduced wave equation for studies of solutions of the steady-state equations, as well as in the development of mathematical physics, might refer to Kline, M., (1972) *Mathematical thought from ancient to modern times*: Oxford University Press, Vol. 2, pp. 693 – 696.

Let us take Fourier's transform of both sides of equation (6.9.1) with t and ω being the variables of transformation. Thus, we write

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\frac{\partial^2 u(x, t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u(x, t)}{\partial t^2} \right] \exp(-i\omega t) dt = 0.$$

We write the integral of a difference as a difference of integrals and factor out the constant term, $1/v^2$, to get

$$(6.9.5) \quad \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial^2 u(x, t)}{\partial x^2} \exp(-i\omega t) dt \right] - \frac{1}{v^2} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial^2 u(x, t)}{\partial t^2} \exp(-i\omega t) dt \right] = 0.$$

Let us consider the first bracketed term in equation (6.9.5). We can immediately rewrite it as

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial^2}{\partial x^2} [u(x, t) \exp(-i\omega t)] dt$$

and interchange the integration and differentiation to get

$$\frac{\partial^2}{\partial x^2} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} u(x, t) \exp(-i\omega t) dt \right];$$

to do so, we used the fact that we take the derivative with respect to x while the limits of integration refer to t . Herein, the term in brackets is the definition of Fourier's transform of function $u(x, t)$, which we denote as $\tilde{u}(x, \omega)$. Let us consider the second bracketed term in equation (6.9.5). We evaluate the integral using integration by parts twice to get

$$(i\omega)^2 \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} u(x, t) \exp(-i\omega t) dt \right].$$

We recognize that the term in brackets is Fourier's transform of function u , namely, \tilde{u} . Returning to equation (6.9.5), we rewrite that equation as

$$(6.9.6) \quad \frac{\partial^2 \tilde{u}(x, \omega)}{\partial x^2} + \left(\frac{\omega}{v} \right)^2 \tilde{u}(x, \omega) = 0,$$

which is Fourier's transform of equation (6.9.1).

Equation (6.9.6) exhibits a similarity to equation (6.9.4). Let us compare these two equations. We note that \dot{u} is a function of x alone, while \tilde{u} is a function of both x and ω . As stated in Section 6.9.1, it is a matter of notation: \dot{u} is the solution of equation (6.9.4) for a particular value of ω ; hence, it is implicitly a function of ω — the fact that is explicitly stated for \tilde{u} .

To complete this section, we can state that if we consider a three-dimensional equivalent of equation (6.9.1), we can take its Fourier's transform to get

$$(6.9.7) \quad \nabla^2 \tilde{u}(\mathbf{x}, \omega) + \left(\frac{\omega}{v}\right)^2 \tilde{u}(\mathbf{x}, \omega) = 0.$$

In this section we have shown that the reduced wave equation can be viewed as a particular case of the wave equation subjected to Fourier's transform with t and ω being the transformation variables. In our subsequent work, when invoking the reduced wave equation, we will use the wave equation subjected to Fourier's transform. The formulation of the reduced wave equation in terms of a trial solution has provided us with an insight into both its physical meaning and its nomenclature.

6.10. Extensions of wave equation

6.10.1. 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.

6.10.2. Standard wave equation. In multidimensional continua, wave equation (6.4.1), may be written as

$$(6.10.1) \quad \nabla^2 u(\mathbf{x}, t) - \frac{1}{v^2} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0,$$

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 (6.10.1), \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 (6.10.1) 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 6.10.3.

We also wish to extend equation (6.10.1) 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 (6.10.1) to account for weak inhomogeneity is illustrated in Section 6.10.4 and belongs to the high-frequency approximation.

6.10.3. Wave equation and elliptical velocity dependence.

Wave equation. To study an extension of the wave equation to anisotropic cases, consider equation (6.10.1). For convenience, let v be equal to unity. Hence, we can write

$$(6.10.2) \quad \nabla^2 u(\mathbf{x}, t) = \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2}.$$

Consider a two-dimensional continuum that is contained in the xz -plane. For $\mathbf{x} = [x, z]$, equation (6.10.2) can be explicitly written as

$$(6.10.3) \quad \frac{\partial^2 u(x, z, t)}{\partial x^2} + \frac{\partial^2 u(x, z, t)}{\partial z^2} = \frac{\partial^2 u(x, z, t)}{\partial t^2}.$$

Let the linear transformation of the position coordinates be such that

$$(6.10.4) \quad \acute{u}(x, z, t) = u\left(\frac{x}{v_x}, \frac{z}{v_z}, t\right),$$

where v_x and v_z are constants. Using the chain rule, as shown in Exercise 6.14, we can write equation (6.10.3) as

$$(6.10.5) \quad v_x^2 \frac{\partial^2 \acute{u}(x, z, t)}{\partial x^2} + v_z^2 \frac{\partial^2 \acute{u}(x, z, t)}{\partial z^2} = \frac{\partial^2 \acute{u}(x, z, t)}{\partial t^2}.$$

Thus, function \acute{u} is the solution of equation (6.10.5).

To illustrate the meaning of constants v_x and v_z , consider transformation (6.10.4) 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,

$$\frac{d}{dt} \sqrt{\xi^2 + \varsigma^2} = 1,$$

the solutions $u(\xi, \varsigma, t)$, at different times t , are concentric circles. It follows that, in the xz -plane,

$$\frac{d}{dt} \sqrt{\frac{x^2}{v_x^2} + \frac{z^2}{v_z^2}} = 1,$$

and, hence, the solutions $\acute{u}(x, z, t)$, at different times t , are ellipses.

Equation (6.10.5) 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 (6.10.5), consider the trial solution given by

$$(6.10.6) \quad \dot{u}(x, z, t) = \exp[i\omega(p_x x + p_z z - t)].$$

If we consider 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 (6.10.6), we substitute it into wave equation (6.10.5). We obtain

$$v_x^2 \omega^2 p_x^2 \exp[i\omega(x p_x + z p_z - t)] + v_z^2 \omega^2 p_z^2 \exp[i\omega(x p_x + z p_z - t)] = \omega^2 \exp[i\omega(x k_x + z k_z - t)].$$

Dividing by ω^2 and by the exponential term, we can write this equation as

$$(6.10.7) \quad v_x^2 p_x^2 + v_z^2 p_z^2 = 1,$$

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.

NOTATION 6.10.1. To avoid any confusion, let us clarify the meaning of notation v_x , v_z , p_x and p_z that is used in the entire book. As in expression (6.10.7), v_x and v_z denote constants that define the properties of the velocity field by giving the magnitude of the wavefront velocity along the x -axis and along the z -axis, respectively; v_x and v_z are not the components of a vector. Symbols p_x and p_z , on the other hand, stand for the components of the phase-slowness vector, \mathbf{p} .

In other words, p_x and p_z specify the orientation of a wavefront in the velocity field defined by v_x and v_z .

We also observe that — for the wavefront propagating along the x -axis — the magnitudes of p_x and p_z are $1/v_x$ and 0, respectively, while — for the wavefront propagating along the z -axis — they are 0 and $1/v_z$. In general, for elliptical velocity dependence, $|p_x| \in [0, 1/v_x]$ and $|p_z| \in [0, 1/v_z]$, as we can confirm by examining expression (6.10.11), below.

Let us return to expression (6.10.7). To state this expression as a function of the orientation of the wavefront, we can express the phase-slowness vector as

$$(6.10.8) \quad \mathbf{p} = [p_x, p_z] = [p(\vartheta) \sin \vartheta, p(\vartheta) \cos \vartheta],$$

where $p(\vartheta)$ 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. Thus, using equation (6.10.8), we can also express the phase angle as

$$\frac{p_x}{p_z} = \frac{\sin \vartheta}{\cos \vartheta} = \tan \vartheta,$$

which means that

$$(6.10.9) \quad \vartheta = \arctan \frac{p_x}{p_z}.$$

Using expression (6.10.8), we can rewrite expression (6.10.7) as

$$(6.10.10) \quad [p(\vartheta)]^2 (v_x^2 \sin^2 \vartheta + v_z^2 \cos^2 \vartheta) = 1.$$

Since the magnitude of phase slowness is the reciprocal of the magnitude of phase velocity, expression (6.10.10) can be restated as

$$(6.10.11) \quad v(\vartheta) = \frac{1}{p(\vartheta)} = \sqrt{v_x^2 \sin^2 \vartheta + v_z^2 \cos^2 \vartheta}.$$

Expression (6.10.11) gives the magnitude of phase velocity as a function of phase angle for the case of elliptical velocity dependence. As shown in Exercise 9.8, 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.

Prior to completing this section, we wish to discuss the meaning of the elliptical velocity dependence with direction. According to wave equation (6.10.5) — with v_x and v_z being the magnitudes of the wavefront velocities along the x -axis and the z -axis, respectively — the wavefronts originating from a point source are elliptical, as shown in discussion that follows expression (6.10.6). This ellipticity is exhibited by the dependence of the magnitude of the phase-slowness vector, \mathbf{p} , with direction. This can be illustrated by the polar plot of expression (6.10.11) written as

$$p(\vartheta) = \frac{1}{\sqrt{v_x^2 \sin^2 \vartheta + v_z^2 \cos^2 \vartheta}},$$

which is an ellipse. In other words, phase slowness, p , exhibits elliptical dependence on the phase angle, ϑ . We note that phase velocity does not exhibit elliptical dependence on the phase angle. This can be illustrated by the polar plot of expression (6.10.11) for $v(\vartheta)$, which is not an ellipse, and — for large enough difference between v_x and v_z — is not even a convex curve. Geometrically, the ellipticity of the wavefronts and of the phase slowness is a result of polar reciprocity to be discussed in Section 8.4.3.

Analytically, this ellipticity results from the fact that \mathbf{p} and $\dot{\mathbf{x}}$, which are associated with the phase slowness and ray velocity, respectively, are the variables of Legendre's transformation, which will be also discussed in Section 8.4. As will be shown in Chapter 8 in the context of Hamilton's ray equations, it is \mathbf{p} and $\dot{\mathbf{x}}$, rather than phase velocity, that are fundamental entities of seismic theory. Thus, we conclude that the term elliptical velocity dependence with direction — which we introduced to refer to elliptical wavefronts generated by a point source and governed by wave equation (6.10.5) — is tantamount to the dependence of phase slowness on phase angle, as shown above, and to the dependence of ray velocity on ray angle, as will be discussed in Section 8.4.

6.10.4. Wave equation and weak inhomogeneity.

Weak inhomogeneity: Formulation of equation. To study an extension of the wave equation to the inhomogeneous case, consider equation (6.10.1), namely,

$$(6.10.12) \quad \nabla^2 u(\mathbf{x}, t) - \frac{1}{v^2} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0,$$

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

$$(6.10.13) \quad \nabla^2 u(\mathbf{x}, t) - \frac{1}{[v(\mathbf{x})]^2} \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} = 0.$$

Since equation (6.10.13) 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 (6.10.13) 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.

Replacing v with $v(\mathbf{x})$ is an arbitrary replacement: it is not justified by a derivation from the fundamentals. There is no guarantee that in a weakly inhomogeneous medium, u satisfies equation (6.10.13). Strictly speaking, such an equation should be derived in a manner to be discussed in Section 7.1.

In seismology, we are interested often in studying layered media where the properties vary along only one axis. Considering a three-dimensional continuum, where $\mathbf{x} = [x, y, z]$, we assume often 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 (6.10.13) 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 (6.10.13) provides a good approximation only for the displacements of waves propagating near the direction of

the z -axis.¹⁹ However, the eikonal equation, which we will derive from equation (6.10.13), provides — within the conditions of this derivation — a good approximation for signal trajectories in all directions of propagation.

Weak inhomogeneity: Formulation of solution. To formulate a trial solution of equation (6.10.13), consider the fact that we can write a trial solution of equation (6.10.12) as

$$(6.10.14) \quad u(\mathbf{x}, t) = A \exp [i\omega (\mathbf{p} \cdot \mathbf{x} - t)],$$

where A is the amplitude of the displacement that varies sinusoidally in space and time as described by $\exp [i\omega (\mathbf{p} \cdot \mathbf{x} - t)]$. As stated in Section 6.10.3, $\exp [\cdot]$ is the phase factor, which is constant for a wavefront at time t . In three-dimensional continua, trial solution (6.10.14) 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 6.10.3, \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 (6.10.13) 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 (6.10.14), we write

$$(6.10.15) \quad u(\mathbf{x}, t) = A(\mathbf{x}) \exp \{i\omega [\psi(\mathbf{x}) - t]\},$$

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, which 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. Examining the phase factor of trial solution (6.10.15) in the context of solutions (6.10.6) and (6.10.14), 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 wavefronts. Since \mathbf{p} is normal to the wavefront, using properties of the gradient, we obtain an important expression, namely,

$$(6.10.16) \quad \mathbf{p} = \nabla \psi.$$

In other words, the phase-slowness vector is the gradient of the eikonal function.

We could write the exponential term as the product of $\exp (i\omega \psi)$ and $\exp (-i\omega t)$; the latter term transforms equation (6.10.13) into the reduced wave equation in a manner shown in Section 6.9.1. This property is illustrated in Exercise 6.18. Following the discussion in Section 6.9 — in particular referring to equation (6.9.7) — we take Fourier's transform of equation (6.10.13) with t and ω being the transformation variables

¹⁹Readers interested in wave propagation in slowly varying vertically nonuniform continua might refer to Grant, F.S., and West, G.F., (1965) Interpretation theory in applied geophysics: McGraw-Hill, Inc., pp. 43 – 47, 49 – 50, and to Krebs, E.S., (2004) Seismic theory and methods (Lecture notes): The University of Calgary, pp. 5-8 – 5-11.

to get

$$(6.10.17) \quad \nabla^2 \tilde{u}(\mathbf{x}, \omega) + \left[\frac{\omega}{v(\mathbf{x})} \right]^2 \tilde{u}(\mathbf{x}, \omega) = 0.$$

Let us use the trial solution analogous to solution (6.10.15) and given by

$$(6.10.18) \quad \tilde{u}(\mathbf{x}, \omega) = A(\mathbf{x}) \exp[i\omega\psi(\mathbf{x})].$$

We insert trial solution (6.10.18) into equation (6.10.17). Firstly, considering the x_i term of Laplace's operator, where $i \in \{1, 2, 3\}$, we obtain

$$\frac{\partial^2}{\partial x_i^2} A(\mathbf{x}) \exp[i\omega\psi(\mathbf{x})] = \exp[i\omega\psi(\mathbf{x})] \left[\frac{\partial^2 A}{\partial x_i^2} + i\omega \left(2 \frac{\partial A}{\partial x_i} \frac{\partial \psi}{\partial x_i} + A \frac{\partial^2 \psi}{\partial x_i^2} \right) - \omega^2 A \frac{\partial \psi}{\partial x_i} \frac{\partial \psi}{\partial x_i} \right].$$

Secondly, considering the second term on the left-hand side of equation (6.10.17), we write

$$\left[\frac{\omega}{v(\mathbf{x})} \right]^2 \tilde{u}(\mathbf{x}, \omega) = \left[\frac{\omega}{v(\mathbf{x})} \right]^2 A(\mathbf{x}) \exp\{i\omega[\psi(\mathbf{x})]\}.$$

Consequently, since the exponential term is never zero, equation (6.10.17) becomes

$$(6.10.19) \quad \sum_{i=1}^3 \frac{\partial^2 A}{\partial x_i^2} + A\omega^2 \left(\frac{1}{v^2} - \sum_{i=1}^3 \frac{\partial \psi}{\partial x_i} \frac{\partial \psi}{\partial x_i} \right) + i\omega \sum_{i=1}^3 \left(2 \frac{\partial A}{\partial x_i} \frac{\partial \psi}{\partial x_i} + A \frac{\partial^2 \psi}{\partial x_i^2} \right) = 0,$$

which is a complex-valued function of real variables.

The vanishing of expression (6.10.19), where both A and ψ are assumed to be real, implies the vanishing of both real and imaginary parts. Assuming $\omega \neq 0$ and following the definitions of the gradient operator and Laplace's operator, we obtain

$$(6.10.20) \quad \begin{cases} \nabla^2 A + A\omega^2 \left[\frac{1}{v^2(\mathbf{x})} - (\nabla\psi)^2 \right] = 0 \\ 2\nabla A \cdot \nabla\psi + A\nabla^2\psi = 0 \end{cases},$$

where $(\nabla\psi)^2 := (\partial\psi/\partial x_1)^2 + (\partial\psi/\partial x_2)^2 + (\partial\psi/\partial x_3)^2$. System (6.10.20) corresponds to equation (6.10.17), in the context of trial solution (6.10.18).

Initially, system (6.10.20) might appear not simpler than equation (6.10.17). However, further analysis of the first equation of this system leads to an important simplification.

Eikonal equation. Considering the first equation of system (6.10.20) and assuming that both ω and A are nonzero, we can write it as

$$(6.10.21) \quad \frac{\nabla^2 A}{A\omega^2} + \left[\frac{1}{v^2(\mathbf{x})} - (\nabla\psi)^2 \right] = 0.$$

If we assume the inhomogeneity of the continuum to be weak, this assumption is tantamount to viewing the wavelength as being short with respect to the characteristic distance over which the properties of the

continuum change significantly; this is analogous to the frequency being high.²⁰ In the limit, we let ω tend to infinity, and equation (6.10.21) becomes

$$(6.10.22) \quad [\nabla\psi(\mathbf{x})]^2 = \frac{1}{v^2(\mathbf{x})}.$$

In view of expression (6.10.16), we can write equation (6.10.22) as

$$(6.10.23) \quad p^2 = \frac{1}{v^2(\mathbf{x})},$$

where $p^2 = \mathbf{p} \cdot \mathbf{p}$.

Equation (6.10.23) is the eikonal equation for isotropic weakly inhomogeneous continua. It can be viewed as an approximation to wave equation (6.10.13): the eikonal equation is the high-frequency approximation; hence, for physical applications it is limited to weak inhomogeneity.²¹ In Chapter 7, we will derive the eikonal equation for anisotropic inhomogeneous continua.

Recall that equation (6.10.13) does not explicitly refer to either P or S waves. Consequently, equation (6.10.23) does not explicitly correspond to either wave. Moreover, in view of the comment on page 183 where we state that replacing v with $v(\mathbf{x})$ in equation (6.10.13) is an arbitrary replacement, our formulation herein is heuristic rather than rigorous: a rigorous and a more general formulation is presented in Chapter 7. However, in view of expression (6.1.13) and (6.1.17), if $v(\mathbf{x})$ is a smooth function given by

$$(6.10.24) \quad v(\mathbf{x}) = \sqrt{\frac{\lambda(\mathbf{x}) + 2\mu(\mathbf{x})}{\rho(\mathbf{x})}},$$

equation (6.10.23) can be viewed as corresponding to P waves, and if $v(\mathbf{x})$ is a smooth function given by

$$(6.10.25) \quad v(\mathbf{x}) = \sqrt{\frac{\mu(\mathbf{x})}{\rho(\mathbf{x})}},$$

equation (6.10.23) can be viewed as corresponding to S waves. In general, for inhomogeneous continua, equations (6.1.4) cannot be split into two wave equations analogous to equations (6.1.12) and (6.1.16). In other words, the dilatational and rotational waves are coupled due to the inhomogeneity of the medium, as illustrated in Exercise 6.19. However, assuming sufficiently high frequency, there are two distinct wavefronts that propagate in an inhomogeneous continuum with speeds given by expressions (6.10.24) and (6.10.25).

²⁰Readers 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, while discussing physical applications, 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.

²¹Readers interested in an elegant formulation of the eikonal equation, which does not use weak inhomogeneity but stems from a geometrical argument, might refer to McOwen, R.C., (1996) *Partial differential equations: Methods and applications*: Prentice-Hall, Inc., pp. 34 – 35.

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 6.15 and 7.5. 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 (6.10.13) is only an approximation of the wave equation.

Transport equation. The second equation of system (6.10.20), namely,

$$(6.10.26) \quad 2\nabla A \cdot \nabla \psi + A \nabla^2 \psi = 0,$$

is the transport equation. For a given eikonal function, ψ , the transport equation describes the amplitude along the wavefront.

Concluding this chapter, let us mention that expression (6.10.18) is a zeroth-order term of the asymptotic series in $\exp(i\omega\psi) / (i\omega)^n$ given by²²

$$u(\mathbf{x}, \omega) \sim \exp[i\omega\psi(\mathbf{x})] \sum_{n=0}^N \frac{A_n(\mathbf{x})}{(i\omega)^n},$$

where \sim stands for “is asymptotically equivalent to”.²³ Hence, the results presented in Section 6.10.4 belong to the realm of asymptotic methods, which play an important role in seismology. In Section 7.2.4, we will discuss briefly ray theory in the context of asymptotic methods.

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.

²²Readers 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.

²³For 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.

The derivation of the wave equation shown in this chapter is rooted in the balance of linear momentum. This derivation formulates wave propagation 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.²⁴



²⁴Interested 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.

6.11. Exercises

EXERCISE 6.1. *Show the details of the derivation of Lemma 6.4.1.*

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 (6.4.4), $\partial y/\partial x = \partial z/\partial x = 1$, we obtain

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z}.$$

Consequently,

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} \right) \\ &= \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} \right) \frac{\partial y}{\partial x} + \frac{\partial}{\partial z} \left(\frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} \right) \frac{\partial z}{\partial x} \\ &= \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial y \partial z} + \frac{\partial^2 u}{\partial z \partial y} + \frac{\partial^2 u}{\partial z^2} \\ (6.11.1) \quad &= \frac{\partial^2 u}{\partial y^2} + 2 \frac{\partial^2 u}{\partial z \partial y} + \frac{\partial^2 u}{\partial z^2}, \end{aligned}$$

where, again, we used 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,

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= \frac{\partial}{\partial t} \left(v \frac{\partial u}{\partial y} - v \frac{\partial u}{\partial z} \right) \\ &= \frac{\partial}{\partial y} \left(v \frac{\partial u}{\partial y} - v \frac{\partial u}{\partial z} \right) \frac{\partial y}{\partial t} + \frac{\partial}{\partial z} \left(v \frac{\partial u}{\partial y} - v \frac{\partial u}{\partial z} \right) \frac{\partial z}{\partial t} \\ &= \left(v \frac{\partial^2 u}{\partial y^2} - v \frac{\partial^2 u}{\partial y \partial z} \right) v + \left(v \frac{\partial^2 u}{\partial z \partial y} - v \frac{\partial^2 u}{\partial z^2} \right) (-v) \\ &= v^2 \left(\frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial y \partial z} - \frac{\partial^2 u}{\partial z \partial y} + \frac{\partial^2 u}{\partial z^2} \right) \\ (6.11.2) \quad &= v^2 \left(\frac{\partial^2 u}{\partial y^2} - 2 \frac{\partial^2 u}{\partial z \partial y} + \frac{\partial^2 u}{\partial z^2} \right). \end{aligned}$$

where the equality of mixed partial derivatives was used. Inserting expressions (6.11.1) and (6.11.2) into equation (6.4.1), 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 was 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 6.4.2.*

SOLUTION 6.2. Considering the equality of mixed partial derivatives, we can write

$$\frac{\partial}{\partial y} \left[\frac{\partial u(y, z)}{\partial z} \right] = \frac{\partial}{\partial z} \left[\frac{\partial u(y, z)}{\partial y} \right] = 0.$$

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

$$\begin{aligned}
 u(y, z) &= \int F(y) \, dy \\
 &= f(y) + a(z),
 \end{aligned}$$

where $a(z)$ is the integration constant with respect to dy , and

$$\begin{aligned}
 u(y, z) &= \int G(z) \, dz \\
 &= g(z) + b(y),
 \end{aligned}$$

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. *Prove the following theorem.*

THEOREM 6.11.1. *Solutions of the wave equation given by*

$$(6.11.3) \quad \frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0,$$

with the boundary conditions given by

$$(6.11.4) \quad u(0, t) = 0$$

and

$$(6.11.5) \quad u(1, t) = 0$$

that correspond to a string of unit length whose ends are fixed, satisfy the conservation of the wave-function energy that is defined by

$$\mathcal{E}(t) := \int_0^1 \left[\left(\frac{\partial u}{\partial t} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] dx.$$

SOLUTION 6.3. PROOF. Let us consider energy \mathcal{E} at two arbitrary instances t_1 and t_2 . The conservation of energy implies

$$\int_0^1 \left[\left(\frac{\partial u}{\partial t} \Big|_{t_1} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t_1} \right)^2 \right] dx = \int_0^1 \left[\left(\frac{\partial u}{\partial t} \Big|_{t_2} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t_2} \right)^2 \right] dx.$$

To prove this equation is tantamount to showing that

$$(6.11.6) \quad \int_0^1 \left[\left(\frac{\partial u}{\partial t} \Big|_{t_1} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t_1} \right)^2 \right] dx - \int_0^1 \left[\left(\frac{\partial u}{\partial t} \Big|_{t_2} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t_2} \right)^2 \right] dx$$

vanishes for all t . Let us write expression (6.11.6) as

$$\int_0^1 \left\{ \left[\left(\frac{\partial u}{\partial t} \Big|_{t_1} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t_1} \right)^2 \right] - \left[\left(\frac{\partial u}{\partial t} \Big|_{t_2} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \Big|_{t_2} \right)^2 \right] \right\} dx.$$

Noticing that we can view the term in braces as a result of a definite integral with respect to t , we can write

$$(6.11.7) \quad \int_0^1 \int_{t_2}^{t_1} \frac{d}{dt} \left[\left(\frac{\partial u}{\partial t} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] dt dx.$$

Using identity (6.11.9) stated in Lemma 6.11.2, below, we can write expression (6.11.7) as

$$2v^2 \int_0^1 \int_{t_2}^{t_1} \frac{d}{dx} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right) dt dx.$$

Changing the order of integration, we get

$$2v^2 \int_{t_2}^{t_1} \int_0^1 \frac{d}{dx} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right) dx dt = 2v^2 \int_{t_2}^{t_1} \left[\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right]_{x=0}^{x=1} dt.$$

Applying the limits to the integrated term, we can write

$$(6.11.8) \quad 2v^2 \left[\int_{t_2}^{t_1} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right) \Big|_{x=1} - \int_{t_2}^{t_1} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right) \Big|_{x=0} \right] dt.$$

Boundary conditions (6.11.4) and (6.11.5) imply that $\partial u / \partial t|_{x=0} = 0$ and $\partial u / \partial t|_{x=1} = 0$, respectively. Physically, no displacement implies no velocity. Hence, expression (6.11.8) vanishes identically and we conclude that expression (6.11.6) vanishes, as required. Thus, the energy of wave function $u(x, t)$ is conserved. \square

LEMMA 6.11.2. *If $u(x, t)$ satisfies wave equation (6.11.3), then the following identity holds:*

$$(6.11.9) \quad \frac{d}{dt} \left[\left(\frac{\partial u}{\partial t} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] = 2v^2 \frac{d}{dx} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right).$$

PROOF. Let us rewrite identity (6.11.9) as

$$(6.11.10) \quad \frac{1}{2} \frac{d}{dt} \left[\left(\frac{\partial u}{\partial t} \right)^2 + v^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] - v^2 \frac{d}{dx} \left(\frac{\partial u}{\partial t} \frac{\partial u}{\partial x} \right) = 0.$$

To prove this identity, we need to show that the left-hand side vanishes. Differentiating the left-hand side using the chain rule, we get

$$\frac{1}{2} \left(2 \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} + 2v^2 \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} \right) - v^2 \left(\frac{\partial^2 u}{\partial x \partial t} \frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial x^2} \right).$$

Using the equality of mixed partial derivatives to cancel the terms containing $\partial^2 u / \partial t \partial x$ and $\partial^2 u / \partial x \partial t$, we get

$$\frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} - v^2 \frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} \left(\frac{\partial^2 u}{\partial t^2} - v^2 \frac{\partial^2 u}{\partial x^2} \right).$$

If $u(x, t)$ satisfies wave equation (6.11.3), then the term in parentheses in the above expression vanishes identically and we obtain

$$\frac{\partial u}{\partial t} \left(\frac{\partial^2 u}{\partial t^2} - v^2 \frac{\partial^2 u}{\partial x^2} \right) = 0.$$

Thus, the left-hand side of equation (6.11.10) vanishes and, hence, identity (6.11.9) is proved, as required. \square

EXERCISE 6.4. *Perform Fourier's transform of equation (6.5.4); namely,*

$$(6.11.11) \quad \sum_{i=1}^3 \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial x_i^2} - \frac{1}{v^2} \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial t^2} = 0,$$

with \mathbf{x} and \mathbf{k} being the variables of transformation.

SOLUTION 6.4. Since equation (6.11.11) is a differential equation, let us consider Fourier's transforms of derivatives. To find the transform of $\partial u(\mathbf{x}, t) / \partial x_j$, we write

$$\frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial u(\mathbf{x}, t)}{\partial x_j} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}.$$

Since we are differentiating with respect to a single component of \mathbf{x} , let us write explicitly

$$\frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\partial u(\mathbf{x}, t)}{\partial x_j} e^{-i\mathbf{k} \cdot \mathbf{x}} dx_j dx_k dx_l,$$

which allows us to use the integration by parts on the innermost integral. Therein, we let $j, k, l \in \{1, 2, 3\}$ to consider the partial derivative with respect to any of the components of \mathbf{x} . Integrating by parts, we get

$$\frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(u(\mathbf{x}, t) e^{-i\mathbf{k} \cdot \mathbf{x}} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} u(\mathbf{x}, t) \frac{\partial}{\partial x_j} e^{-i\mathbf{k} \cdot \mathbf{x}} dx_j \right) dx_k dx_l.$$

We assume that the integrated term vanishes at $\pm\infty$ to get

$$-\frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u(\mathbf{x}, t) \frac{\partial}{\partial x_j} e^{-i\mathbf{k} \cdot \mathbf{x}} dx_j dx_k dx_l.$$

Differentiating the exponential with respect to x_j , we get

$$\frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u(\mathbf{x}, t) ik_j e^{-i\mathbf{k} \cdot \mathbf{x}} dx_j dx_k dx_l,$$

which we rewrite as

$$ik_j \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} u(\mathbf{x}, t) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}.$$

We recognize that this expression contains the definition of Fourier's transform. Thus we write this expression as $ik_j \tilde{u}(\mathbf{k}, t)$, which is Fourier's transform of $\partial u(\mathbf{x}, t) / \partial x_j$. Following a procedure analogous to the one shown above, we obtain the transform of $\partial^2 u(\mathbf{x}, t) / \partial x_j^2$; namely,

$$(6.11.12) \quad \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial^2 u(\mathbf{x}, t)}{\partial x_j^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} = (ik_j)^2 \tilde{u}(\mathbf{k}, t) = -k_j^2 \tilde{u}(\mathbf{k}, t).$$

Having derived the formula for Fourier's transform of the second derivative, let us consider equation (6.11.11). We write

$$\frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \left(\left(\sum_{j=1}^3 \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial x_j^2} - \frac{1}{v^2} \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial t^2} \right) e^{-i\mathbf{k} \cdot \mathbf{x}} \right) d\mathbf{x} = \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} 0 e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}.$$

Using the linearity of the integral operator and the fact that Fourier's transform of zero is zero, we write

$$\frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \sum_{j=1}^3 \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial x_j^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} - \frac{1}{v^2} \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial^2 u(x_1, x_2, x_3, t)}{\partial t^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} = 0.$$

Again, using the linearity, we write

$$\begin{aligned} \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial^2 u}{\partial x_1^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} + \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial^2 u}{\partial x_2^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} + \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial^2 u}{\partial x_3^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} \\ - \frac{1}{v^2} \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{\partial^2 u}{\partial t^2} e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} = 0. \end{aligned}$$

Let us consider the first three integrals. In view of expression (6.11.12), we write them as

$$-k_1^2 \tilde{u}(\mathbf{k}, t) - k_2^2 \tilde{u}(\mathbf{k}, t) - k_3^2 \tilde{u}(\mathbf{k}, t) = -(k_1^2 + k_2^2 + k_3^2) \tilde{u}(\mathbf{k}, t).$$

We recognize that the the sum in parentheses is the scalar product of vector \mathbf{k} with itself. Thus it is the squared length of \mathbf{k} , which we denote by $|\mathbf{k}|^2$. Let us consider the fourth integral. Using the commutativity of the integral and differential operators, we write

$$-\frac{1}{v^2} \frac{\partial^2}{\partial t^2} \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} u(\mathbf{x}, t) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} = 0.$$

The term to be differentiated is the definition of Fourier's transform of u ; namely, \tilde{u} . Combining these results, we write Fourier's transform of equation (6.11.11) as

$$|\mathbf{k}|^2 \tilde{u}(\mathbf{k}, t) + \frac{1}{v^2} \frac{\partial^2 \tilde{u}(\mathbf{k}, t)}{\partial t^2} = 0,$$

which is equivalent to equation (6.5.5).

EXERCISE 6.5. Solve equation (6.5.5); namely,

$$\frac{\partial^2 \tilde{u}(\mathbf{k}, t)}{\partial t^2} + v^2 |\mathbf{k}|^2 \tilde{u}(\mathbf{k}, t) = 0,$$

with initial conditions given by equations (6.5.6) and (6.5.7); namely,

$$\tilde{u}(\mathbf{k}, t)|_{t=0} = \tilde{\gamma}(\mathbf{k})$$

and

$$\left. \frac{\partial \tilde{u}(\mathbf{k}, t)}{\partial t} \right|_{t=0} = \tilde{\eta}(\mathbf{k}),$$

respectively.

SOLUTION 6.5. Since the derivatives of $\tilde{u}(\mathbf{k}, t)$ are taken only with respect to a single variable, we can approach this partial differential equation as if it were an ordinary differential equation. Invoking the theory of second-order ordinary differential equations with constant coefficients, we write the general solution as

a linear combination of $\exp (i v|\mathbf{k}| t)$ and $\exp (-i v|\mathbf{k}| t)$; namely,

$$\tilde{u}(\mathbf{k}, t) = F(\mathbf{k}) \exp \{i v|\mathbf{k}| t\} + G(\mathbf{k}) \exp \{-i v|\mathbf{k}| t\},$$

which is expression (6.5.8). Since we are dealing with a partial differential equation with derivatives in terms of t alone, F and G are constants with respect to t but can depend on \mathbf{k} .

To obtain explicit expressions for F and G , we invoke the initial conditions. We get

$$\tilde{u}(\mathbf{k}, t)|_{t=0} = F(\mathbf{k}) + G(\mathbf{k}) = \tilde{\gamma}(\mathbf{k})$$

and

$$\begin{aligned} \left. \frac{\partial \tilde{u}(\mathbf{k}, t)}{\partial t} \right|_{t=0} &= i v|\mathbf{k}| F(\mathbf{k}) \exp \{i v|\mathbf{k}| t\}|_{t=0} - i v|\mathbf{k}| G(\mathbf{k}) \exp \{-i v|\mathbf{k}| t\}|_{t=0} \\ &= i v|\mathbf{k}| F(\mathbf{k}) - i v|\mathbf{k}| G(\mathbf{k}) = \tilde{\eta}(\mathbf{k}). \end{aligned}$$

Let us write this system of equations as

$$\begin{bmatrix} 1 & 1 \\ i v|\mathbf{k}| & -i v|\mathbf{k}| \end{bmatrix} \begin{bmatrix} F(\mathbf{k}) \\ G(\mathbf{k}) \end{bmatrix} = \begin{bmatrix} \tilde{\gamma}(\mathbf{k}) \\ \tilde{\eta}(\mathbf{k}) \end{bmatrix}.$$

Using Cramer's rule, we obtain

$$F(\mathbf{k}) = \frac{\det \begin{bmatrix} \tilde{\gamma}(\mathbf{k}) & 1 \\ \tilde{\eta}(\mathbf{k}) & -i v|\mathbf{k}| \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 \\ i v|\mathbf{k}| & -i v|\mathbf{k}| \end{bmatrix}} = \frac{-i v|\mathbf{k}| \tilde{\gamma}(\mathbf{k}) - \tilde{\eta}(\mathbf{k})}{-2i v|\mathbf{k}|} = \frac{1}{2} \left(\tilde{\gamma}(\mathbf{k}) + \frac{1}{i v|\mathbf{k}|} \tilde{\eta}(\mathbf{k}) \right)$$

and

$$G(\mathbf{k}) = \frac{\det \begin{bmatrix} 1 & \tilde{\gamma}(\mathbf{k}) \\ i v|\mathbf{k}| & \tilde{\eta}(\mathbf{k}) \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 \\ i v|\mathbf{k}| & -i v|\mathbf{k}| \end{bmatrix}} = \frac{\tilde{\eta}(\mathbf{k}) - i v|\mathbf{k}| \tilde{\gamma}(\mathbf{k})}{-2i v|\mathbf{k}|} = \frac{1}{2} \left(\tilde{\gamma}(\mathbf{k}) - \frac{1}{i v|\mathbf{k}|} \tilde{\eta}(\mathbf{k}) \right),$$

which are expressions (6.5.9) and (6.5.10), respectively.

EXERCISE 6.6. Find Fourier's transform of distribution (6.5.13), namely,

$$(6.11.13) \quad \alpha(\tau) = \frac{(2\pi)^3}{4\pi v^2 t} \iint_{S(\mathbf{0}, vt)} \tau d\zeta.$$

In other words, find $\tilde{\alpha}(\tau)$, where symbol $\tilde{}$ denotes the transformed entity.

SOLUTION 6.6. Let us consider the definition of Fourier's transform given by

$$\tilde{f}(\mathbf{k}) = \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} f(\mathbf{x}) \exp \{-i \mathbf{k} \cdot \mathbf{x}\} d\mathbf{x},$$

where f is a function. If f is considered as a distribution, the effect of this distribution on a test function, τ , is defined as

$$f(\tau) = \iiint_{\mathbb{R}^3} f(\mathbf{x}) \tau(\mathbf{x}) d\mathbf{x},$$

and the corresponding Fourier's transform is defined by letting $\tau = \exp\{-i\mathbf{k} \cdot \mathbf{x}\} / (2\pi)^3$. Distribution (6.11.13) has the effect of Dirac's delta with support on sphere S centred at the origin and with radius vt ; in other words, we can view this distribution as a three-dimensional Dirac's delta.²⁵ We write

$$(6.11.14) \quad \tilde{\alpha} = \frac{1}{4\pi v^2 t} \iint_{S(\mathbf{0}, vt)} \exp\{-i\mathbf{k} \cdot \mathbf{x}\} d\zeta(\mathbf{x}).$$

To integrate, we will use the spherical coordinates. Since the integrand does not depend on the orientation, we rotate the coordinate system to be such that \mathbf{k} points in the x_1 -axis direction. Hence, $\mathbf{k} = [|\mathbf{k}|, 0, 0]$, where $|\mathbf{k}|$ stands for the length of \mathbf{k} . Thus, we write expression (6.11.14) as

$$(6.11.15) \quad \tilde{\alpha} = \frac{1}{4\pi v^2 t} \iint_{S(\mathbf{0}, vt)} \exp\{-i|\mathbf{k}|x_1\} d\zeta(\mathbf{x}).$$

Recalling that the radius is vt , we write the relation between the coordinates as

$$x_1 = vt \cos \alpha,$$

$$x_2 = vt \sin \alpha \cos \beta$$

and

$$x_3 = vt \sin \alpha \sin \beta,$$

where α is the latitude and β is the azimuth. The element of surface for the sphere is

$$d\zeta(\mathbf{x}) = (vt)^2 \sin \alpha d\alpha d\beta.$$

Returning to integral (6.11.15), we write this integral in spherical coordinates as

$$\tilde{\alpha}(\tau) = \frac{1}{4\pi v^2 t} \int_0^{2\pi} \int_0^\pi \exp\{-i|\mathbf{k}|vt \cos \alpha\} (vt)^2 \sin \alpha d\alpha d\beta$$

Simplifying, we write

$$\tilde{\alpha}(\tau) = \frac{t}{4\pi} \int_0^{2\pi} \int_0^\pi \exp\{-i|\mathbf{k}|vt \cos \alpha\} \sin \alpha d\alpha d\beta.$$

²⁵Readers interested in Dirac's delta for two and three dimensions might refer to Barton, G., (1989) Elements of Green's functions and propagation: Potentials, diffusion and waves: Oxford Science Publications, pp. 30 – 32.

Let us consider the inner integral. Setting $m = \exp \{-i |\mathbf{k}| vt \cos \alpha\}$, we get

$$\frac{1}{i |\mathbf{k}| vt} \int dm = \frac{m}{i |\mathbf{k}| vt},$$

which means that

$$\begin{aligned} \int_0^\pi \exp \{-i |\mathbf{k}| vt \cos \alpha\} \sin \alpha d\alpha &= \left. \frac{\exp \{-i |\mathbf{k}| vt \cos \alpha\}}{i |\mathbf{k}| vt} \right|_0^\pi \\ &= \frac{\exp \{i |\mathbf{k}| vt\} - \exp \{-i |\mathbf{k}| vt\}}{i |\mathbf{k}| vt}. \end{aligned}$$

Thus,

$$\tilde{\alpha}(\tau) = \frac{t}{4\pi} \int_0^{2\pi} \frac{\exp \{i |\mathbf{k}| vt\} - \exp \{-i |\mathbf{k}| vt\}}{i |\mathbf{k}| vt} d\beta.$$

Since the integrand does not depend on the azimuth, we write

$$\begin{aligned} \tilde{\alpha}(\tau) &= \frac{t}{4\pi} \frac{\exp \{i |\mathbf{k}| vt\} - \exp \{-i |\mathbf{k}| vt\}}{i |\mathbf{k}| vt} \int_0^{2\pi} d\beta \\ &= \frac{1}{4\pi} \frac{\exp \{i |\mathbf{k}| vt\} - \exp \{-i |\mathbf{k}| vt\}}{i |\mathbf{k}| v} [\beta]_0^{2\pi} \\ &= \frac{1}{4\pi} \frac{\exp \{i |\mathbf{k}| vt\} - \exp \{-i |\mathbf{k}| vt\}}{i |\mathbf{k}| vt} 2\pi, \end{aligned}$$

which is Fourier's transform of distribution (6.11.13). Thus, we write

$$(6.11.16) \quad \tilde{\alpha}(\tau) = \frac{\exp \{i |\mathbf{k}| vt\} - \exp \{-i |\mathbf{k}| vt\}}{2i |\mathbf{k}| v},$$

which is the fraction in solution (6.5.12).

EXERCISE 6.7. Using the result of Exercise 6.6, show that the inverse Fourier's transform of expression (6.5.12); namely,

$$(6.11.17) \quad \tilde{u}(\mathbf{k}, t) = \tilde{\gamma}(\mathbf{k}) \frac{d}{dt} \frac{\exp \{iv |\mathbf{k}| t\} - \exp \{-iv |\mathbf{k}| t\}}{2iv |\mathbf{k}|} + \tilde{\eta}(\mathbf{k}) \frac{\exp \{iv |\mathbf{k}| t\} - \exp \{-iv |\mathbf{k}| t\}}{2iv |\mathbf{k}|}$$

is expression (6.5.14); namely,

$$u(\mathbf{x}, t) = \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{x}, vt)} \eta(\mathbf{y}) d\zeta(\mathbf{y}) + \frac{d}{dt} \frac{2\pi^2}{v^2 t} \iint_{S(\mathbf{x}, vt)} \gamma(\mathbf{y}) d\zeta(\mathbf{y}).$$

SOLUTION 6.7. Since $\tilde{\gamma}$ does not depend on t , we write expression (6.11.17) as

$$\tilde{u}(\mathbf{k}, t) = \frac{d}{dt} \left(\tilde{\gamma}(\mathbf{k}) \frac{\exp \{iv |\mathbf{k}| t\} - \exp \{-iv |\mathbf{k}| t\}}{2iv |\mathbf{k}|} \right) + \tilde{\eta}(\mathbf{k}) \frac{\exp \{iv |\mathbf{k}| t\} - \exp \{-iv |\mathbf{k}| t\}}{2iv |\mathbf{k}|}.$$

As derived in Exercise 6.6, expression (6.11.16) is the inverse Fourier's transform of distribution (6.11.13). Thus, we write

$$\tilde{u}(\mathbf{k}, t) = \frac{d}{dt} (\tilde{\gamma}(\mathbf{k}) \tilde{\alpha}(\mathbf{k})) + \tilde{\eta}(\mathbf{k}) \tilde{\alpha}(\mathbf{k}).$$

Taking the inverse Fourier's transform by using the fact that the product of two functions in the \mathbf{k} -domain is their convolution in the \mathbf{x} -domain, we write

$$u(\mathbf{x}, t) = \frac{d}{dt} (\gamma(\mathbf{x}) \star \alpha(\mathbf{x})) + \eta(\mathbf{x}) \star \alpha(\mathbf{x}).$$

Invoking the definition of convolution, we write

$$u(\mathbf{x}, t) = \frac{d}{dt} \iiint_{\mathbb{R}^3} \gamma(\mathbf{x} - \mathbf{y}) \alpha(\mathbf{y}) d\mathbf{y} + \iiint_{\mathbb{R}^3} \eta(\mathbf{x} - \mathbf{y}) \alpha(\mathbf{y}) d\mathbf{y}.$$

In view of the commutativity of convolution, we write

$$u(\mathbf{x}, t) = \frac{d}{dt} \iiint_{\mathbb{R}^3} \gamma(\mathbf{y}) \alpha(\mathbf{x} - \mathbf{y}) d\mathbf{y} + \iiint_{\mathbb{R}^3} \eta(\mathbf{y}) \alpha(\mathbf{x} - \mathbf{y}) d\mathbf{y},$$

Since α is Dirac's delta on a sphere whose centre is the origin, $\alpha(\mathbf{x} - \mathbf{y})$ is the corresponding Dirac's delta with support on the translated sphere whose centre is at \mathbf{x} . Recalling expression (6.11.13), we write the effect of this Dirac's delta on γ and η as

$$u(\mathbf{x}, t) = \frac{2\pi^2}{v^2 t} \left(\frac{d}{dt} \iint_{S(\mathbf{x}, vt)} \gamma(\mathbf{y}) d\zeta(\mathbf{y}) + \iint_{S(\mathbf{x}, vt)} \eta(\mathbf{y}) d\zeta(\mathbf{y}) \right),$$

which is equivalent to expression (6.5.14), as required.

EXERCISE 6.8. Using the fact that η is a function of s_1 and s_2 only, write integral

$$(6.11.18) \quad \iint_{S(\mathbf{x}, vt)} \eta(s_1, s_2) d\zeta(\mathbf{s}),$$

where S is a sphere with radius vt centred at \mathbf{x} , as

$$(6.11.19) \quad 2vt \iint_{D(\mathbf{x}, vt)} \frac{\eta(s_1, s_2)}{\sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}} ds_1 ds_2,$$

where D is the disc with radius vt centred at \mathbf{x} .

SOLUTION 6.8. Let us write the sphere that is centered at \mathbf{x} and whose radius is vt as

$$(6.11.20) \quad (s_1 - x_1)^2 + (s_2 - x_2)^2 + (s_3 - x_3)^2 = (vt)^2,$$

where \mathbf{s} and \mathbf{x} are coordinates of the three-dimensional space. Since η does not depend on s_3 , we consider the $s_1 s_2$ -plane. The element of surface for the sphere, $d\zeta(\mathbf{s})$, projected onto this plane is

$$d\zeta(\mathbf{s}) = \sqrt{1 + \left(\frac{\partial s_3}{\partial s_1}\right)^2 + \left(\frac{\partial s_3}{\partial s_2}\right)^2} ds_1 ds_2.$$

We solve equation (6.11.20) for s_3 to get

$$s_3 = \pm \sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2} + x_3,$$

where the signs denote s_3 that corresponds to the upper or lower hemisphere. Since η is independent of s_3 , let us use only the upper hemisphere and, later on, multiply our result by 2. Herein, we write

$$\frac{\partial s_3}{\partial s_1} = \frac{-(s_1 - x_1)}{\sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}}$$

and

$$\frac{\partial s_3}{\partial s_2} = \frac{-(s_2 - x_2)}{\sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}}.$$

Consequently,

$$d\zeta(\mathbf{s}) = \sqrt{1 + \frac{(s_1 - x_1)^2}{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2} + \frac{(s_2 - x_2)^2}{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}} ds_1 ds_2.$$

Expressing the radicand using the common denominator, we get

$$d\zeta(\mathbf{s}) = \frac{vt}{\sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}} ds_1 ds_2.$$

Hence we write integral (6.11.18) as

$$\iint_{S(\mathbf{x}, vt)} \eta(s_1, s_2) d\zeta(\mathbf{s}) = 2vt \iint_{(s_1 - x_1)^2 + (s_2 - x_2)^2 \leq (vt)^2} \frac{\eta(s_1, s_2)}{\sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}} ds_1 ds_2,$$

where the factor of 2 on the right-hand side accounts for both the upper and lower hemispheres. Denoting $(s_1 - x_1)^2 + (s_2 - x_2)^2 \leq (vt)^2$ by $D(\mathbf{x}, vt)$, which stands for the disc with radius vt that is centred at \mathbf{x} , we obtain

$$2vt \iint_{D(\mathbf{x}, vt)} \frac{\eta(s_1, s_2)}{\sqrt{(vt)^2 - (s_1 - x_1)^2 - (s_2 - x_2)^2}} ds_1 ds_2,$$

which is equation (6.11.19), as required.

EXERCISE 6.9. *Show that*

$$(6.11.21) \quad \int_{-\infty}^{\infty} \exp\{i(x - \zeta)k - k^2 t\} dk = \frac{\exp\left\{-\frac{(x - \zeta)^2}{4t}\right\}}{\sqrt{t}} \int_{-\infty}^{\infty} \exp\{-s^2\} ds.$$

SOLUTION 6.9. Let $k = a + ib$, where $a \in (-\infty, \infty)$ is the variable of integration and $b \in \mathbb{R}$ is fixed and given by $(x - \zeta)/2t$; hence, $dk = da$. Furthermore, let $a = s/\sqrt{t}$, and hence, $da = ds/\sqrt{t}$ and $dk = ds/\sqrt{t}$. Also, $k = s/\sqrt{t} + i(x - \zeta)/2t$, which we write on the common denominator as $k = (2s\sqrt{t} + i(x - \zeta))/2t$. Thus, we write the exponential of the above integral as

$$\begin{aligned} i(x - \zeta)k - k^2t &= i(x - \zeta) \frac{2s\sqrt{t} + i(x - \zeta)}{2t} - \left(\frac{2s\sqrt{t} + i(x - \zeta)}{2t} \right)^2 t \\ &= \frac{2i(x - \zeta)s\sqrt{t} - (x - \zeta)^2}{2t} - \frac{(2s\sqrt{t} + i(x - \zeta))^2}{4t}, \end{aligned}$$

which we rewrite as

$$\begin{aligned} \frac{1}{2t} \left(2i(x - \zeta)s\sqrt{t} - (x - \zeta)^2 - \frac{1}{2} (2s\sqrt{t} + i(x - \zeta))^2 \right) \\ = \frac{1}{2t} \left(2i(x - \zeta)s\sqrt{t} - (x - \zeta)^2 - \frac{1}{2} (4s^2t + 4is\sqrt{t}(x - \zeta) - (x - \zeta)^2) \right). \end{aligned}$$

Simplifying, we get

$$\frac{1}{2t} \left(2i(x - \zeta)s\sqrt{t} - (x - \zeta)^2 - 2s^2t - 2is\sqrt{t}(x - \zeta) + \frac{1}{2}(x - \zeta)^2 \right).$$

Canceling and gathering the common terms, we obtain

$$\frac{1}{2t} \left(-\frac{1}{2}(x - \zeta)^2 - 2s^2t \right) = -s^2 - \frac{(x - \zeta)^2}{4t}.$$

Returning to integral (6.11.21), we write it as

$$\int_{-\infty}^{\infty} \exp \left\{ -s^2 - \frac{(x - \zeta)^2}{4t} \right\} \frac{1}{\sqrt{t}} ds = \int_{-\infty}^{\infty} \exp \{-s^2\} \exp \left\{ -\frac{(x - \zeta)^2}{4t} \right\} \frac{1}{\sqrt{t}} ds.$$

Since x , ζ and t are independent of s , we get

$$\frac{\exp \left\{ -\frac{(x - \zeta)^2}{4t} \right\}}{\sqrt{t}} \int_{-\infty}^{\infty} \exp \{-s^2\} ds.$$

EXERCISE 6.10. Evaluate

$$(6.11.22) \quad \int_{-\infty}^{\infty} e^{-s^2} ds.$$

SOLUTION 6.10. This integrand is not a derivative of an elementary function, which is a function that can be formed from the algebraic, exponential, logarithmic or trigonometric functions by a finite number of operations consisting of addition, subtraction, multiplication and composition of functions. To find the

antiderivative, let us consider

$$(6.11.23) \quad I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy,$$

to be evaluated in polar coordinates, where

$$x(r, \alpha) = r \cos \alpha$$

and

$$y(r, \alpha) = r \sin \alpha.$$

Since the Jacobian of transformation is

$$\det \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \alpha} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \alpha} \end{bmatrix} = \det \begin{bmatrix} \cos \alpha & -r \sin \alpha \\ \sin \alpha & r \cos \alpha \end{bmatrix} = r (\cos^2 \alpha + \sin^2 \alpha) = r,$$

we write

$$I = \int_0^{2\pi} \int_0^{\infty} r e^{-r^2} dr d\alpha,$$

with the limits of integration covering the entire xy -plane. Let us consider the inner integral. Letting $r^2 = m$, which implies that $dr = dm/2r$, we write

$$\int_0^{2\pi} \int_0^{\infty} r e^{-r^2} dr = \frac{1}{2} \int_0^{2\pi} e^{-m} dm = \frac{1}{2} [-e^{-m}]_0^{\infty} = \frac{1}{2}.$$

Hence,

$$I = \frac{1}{2} \int_0^{2\pi} d\alpha = \pi.$$

Let us return to integral (6.11.23), and write it as

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} dx dy = \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy.$$

Since x and y are the dummy variables of integration, we denote them by s , to write

$$I = \left(\int_{-\infty}^{\infty} e^{-s^2} ds \right)^2,$$

where the integral in parentheses is integral (6.11.22). Thus,

$$\int_{-\infty}^{\infty} e^{-s^2} ds = \sqrt{I} = \sqrt{\pi},$$

which is the required result.

EXERCISE 6.11. Solve equations (6.7.7) and (6.7.9) for f_1 in terms of w .

SOLUTION 6.11. Solving equation (6.7.7), namely,

$$v_1 f_1' (v_1 t) - v_1 w' (-v_1 t) = -v_2 g_2' (-v_2 t),$$

for g_2' , we get

$$g_2' (-v_2 t) = \frac{v_1}{v_2} [w' (-v_1 t) - f_1' (v_1 t)].$$

Substituting into equation (6.7.9) and simplifying, we obtain

$$(\rho_1 v_1 + \rho_2 v_2) f_1' (v_1 t) = (\rho_2 v_2 - \rho_1 v_1) w' (-v_1 t).$$

Integrating both sides with respect to t and simplifying, we get

$$f_1 (v_1 t) = \frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2} w (-v_1 t),$$

as required.

EXERCISE 6.12. ²⁶Find the derivative of Heaviside's function, which is given by

$$h(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}.$$

SOLUTION 6.12. Using expression (6.8.1), we get

$$\int_{-\infty}^{\infty} h'(t) \tau(t) dt = - \int_0^{\infty} \tau'(t) dt = - \tau(t)|_0^{\infty}.$$

Since τ has compact support, $\tau(\infty) = 0$ and, hence, $-\tau(t)|_0^{\infty} = \tau(t)|_{t=0}$. Thus we can write the derivative of Heaviside's function as

$$(6.11.24) \quad \int_{-\infty}^{\infty} h'(t) \tau(t) dt = \tau(t)|_{t=0},$$

which is a well-defined quantity.

REMARK 6.11.3. To interpret this result, let us consider the key property defining Dirac's delta, δ , namely,

$$(6.11.25) \quad \int_{-\infty}^{\infty} \delta(t) \tau(t) dt = \tau(t)|_{t=0}.$$

Examining expressions (6.11.24) and (6.11.25), we can write

$$\int_{-\infty}^{\infty} h'(t) \tau(t) dt = \int_{-\infty}^{\infty} \delta(t) \tau(t) dt,$$

²⁶See also Section 6.8.2.

which means that $h'(t)$ behaves like $\delta(t)$. In the distributional sense, Dirac's delta can be regarded as the derivative of Heaviside's function.

The result of Exercise 6.12 is consistent with our intuition developed in standard differentiation. For $t \neq 0$, the derivative vanishes as it does in classical calculus, namely, $h'(t) = 0$. At the jump discontinuity, where $t = 0$, the derivative corresponds to Dirac's delta, which again is quite intuitive if we regard the value of δ at this point as being infinite.

EXERCISE 6.13. Consider wave equation (6.4.1). Using solution (6.9.2), obtain equation (6.9.4).

SOLUTION 6.13. In view of solution (6.9.2), namely, $u(x, t) = \dot{u}(x) \exp(-i\omega t)$, consider the position derivatives, namely,

$$(6.11.26) \quad \frac{\partial^2 u(x, t)}{\partial x^2} = \frac{\partial^2 \dot{u}(x)}{\partial x^2} \exp(-i\omega t),$$

and the time derivatives, namely,

$$(6.11.27) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = -\omega^2 \dot{u}(x) \exp(-i\omega t).$$

Substituting expressions (6.11.26) and (6.11.27) into equation (6.4.1), and dividing by the exponential factor, we obtain a function of a single variable,

$$\frac{d^2 \dot{u}(x)}{dx^2} + \left(\frac{\omega}{v}\right)^2 \dot{u}(x) = 0,$$

which is equation (6.9.4), as required.

EXERCISE 6.14. Consider equation (6.10.3). In view of transformation (6.10.4), let

$$(6.11.28) \quad \dot{u}(x, z, t) \equiv u(\xi, \varsigma, t),$$

where $\xi := x/v_x$ and $\varsigma := z/v_z$. Using the chain rule, show that equation (6.10.3) is equivalent to equation (6.10.5).

SOLUTION 6.14. Taking the derivative of both sides of equation (6.11.28) with respect to x , we obtain

$$\frac{\partial \dot{u}(x, z, t)}{\partial x} = \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} = \frac{1}{v_x} \frac{\partial u(\xi, \varsigma, t)}{\partial \xi},$$

and

$$(6.11.29) \quad \frac{\partial^2 \dot{u}(x, z, t)}{\partial x^2} = \frac{1}{v_x^2} \frac{\partial^2 u(\xi, \varsigma, t)}{\partial \xi^2}.$$

Similarly, taking the derivative of both sides of equation (6.11.28) with respect to z , we obtain

$$(6.11.30) \quad \frac{\partial^2 \dot{u}(x, z, t)}{\partial z^2} = \frac{1}{v_z^2} \frac{\partial^2 u(\xi, \varsigma, t)}{\partial \varsigma^2},$$

while, taking the derivative of both sides of equation (6.11.28) with respect to t , we get

$$(6.11.31) \quad \frac{\partial^2 \dot{u}(x, z, t)}{\partial t^2} = \frac{\partial^2 u(\xi, \varsigma, t)}{\partial t^2}.$$

We can always write equation (6.10.3) as

$$(6.11.32) \quad \frac{\partial^2 u(\xi, \varsigma, t)}{\partial \xi^2} + \frac{\partial^2 u(\xi, \varsigma, t)}{\partial \varsigma^2} = \frac{\partial^2 u(\xi, \varsigma, t)}{\partial t^2},$$

where ξ and ς are the variables of differentiation. Substituting expressions from equations (6.11.29), (6.11.30) and (6.11.31) into (6.11.32), we obtain equation (6.10.5), as required.

EXERCISE 6.15. ²⁷Consider 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 $\eta = n_1 x_1 + n_2 x_2 + n_3 x_3 - vt$, 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

$$(6.11.33) \quad (\nabla u)^2 \equiv \left(\frac{\partial u}{\partial x_1} \right)^2 + \left(\frac{\partial u}{\partial x_2} \right)^2 + \left(\frac{\partial u}{\partial x_3} \right)^2 = \frac{1}{v^2} \left(\frac{\partial u}{\partial t} \right)^2.$$

SOLUTION 6.15. 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 (6.11.33), 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

$$\left(\frac{\partial f}{\partial \eta} \right)^2 (n_1^2 + n_2^2 + n_3^2) = \left(\frac{\partial f}{\partial \eta} \right)^2.$$

This equality is justified since for the unit vector, \mathbf{n} , we have $n_1^2 + n_2^2 + n_3^2 = 1$.

EXERCISE 6.16. Show that the characteristic equation of wave equation (6.4.1) is, in general, satisfied by $u(x, t) = f(\eta)$, where $\eta = x \pm vt$.

²⁷See also Section 6.10.4 and Exercise 7.5.

SOLUTION 6.16. We could view the solution of this exercise as a special case of Exercise 6.15. However, if we wish to write it out, we get

$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial \eta} \frac{\partial \eta}{\partial x} = \frac{\partial f}{\partial \eta}$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial f}{\partial \eta} \frac{\partial \eta}{\partial t} = \pm v \frac{\partial f}{\partial \eta}.$$

In view of equation (6.11.33), the characteristic equation of equation (6.4.1) is

$$(6.11.34) \quad \left(\frac{\partial u}{\partial x} \right)^2 - \frac{1}{v^2} \left(\frac{\partial u}{\partial t} \right)^2 = 0.$$

Inserting the above expressions for $\partial u/\partial x$ and $\partial u/\partial t$ into the left-hand side of equation (6.11.34), we get

$$\left(\frac{\partial u}{\partial x} \right)^2 - \frac{1}{v^2} \left(\frac{\partial u}{\partial t} \right)^2 = \left(\frac{\partial f}{\partial \eta} \right)^2 - \frac{1}{v^2} \left(\pm v \frac{\partial f}{\partial \eta} \right)^2.$$

Performing algebraic manipulations, we obtain

$$\left(\frac{\partial f}{\partial \eta} \right)^2 - \left(\frac{\partial f}{\partial \eta} \right)^2 = 0,$$

as required. Thus, functions $f(x \pm vt)$ are solutions of equation (6.11.34), which is the characteristic equation of wave equation (6.4.1).

REMARK 6.11.4. We note that although both wave equation (6.4.1) and characteristic equation (6.11.34) have a general solution that possesses the same form, these equations are not equivalent to one another. Mathematically, equation

$$\frac{\partial u}{\partial x} = \pm \frac{1}{v} \frac{\partial u}{\partial t},$$

whose solutions are functions $u(x, t) = f(x \pm vt)$, does not suffice to describe such wave phenomena as reflection. Also, in the context of this chapter, the above equation is not the result of inserting stress-strain equations into Cauchy's equations of motion. Hence, the above equation is not rooted in physical concepts of the balance of linear momentum.

EXERCISE 6.17. ²⁸Show that expressions (6.4.4) are particular solutions of the characteristic equation of wave equation (6.4.1).

SOLUTION 6.17. We could view the solution of this exercise as a special case of Exercise 6.16. However, if we wish to write it out, inserting into the left-hand side of characteristic equation (6.11.34) expressions (6.4.4), which we can write as $u(x, t) = x \pm vt$, we get

$$\left[\frac{\partial(x \pm vt)}{\partial x} \right]^2 - \frac{1}{v^2} \left[\frac{\partial(x \pm vt)}{\partial t} \right]^2 = (1)^2 - \frac{1}{v^2} (\pm v)^2.$$

Performing algebraic manipulations, we obtain $1 - 1 = 0$, as required. Thus, expressions (6.4.4) are solutions of equation (6.11.34), which is the characteristic equation of wave equation (6.4.1).

²⁸See also Section 6.4.1.

EXERCISE 6.18. *In view of Section 6.9, considering the reduced form of equation (6.10.13) 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 system (6.10.20).*

SOLUTION 6.18. Considering a single spatial dimension, the reduced form of equation (6.10.13) is

$$\frac{d^2 u(x)}{dx^2} + \left[\frac{\omega}{v(x)} \right]^2 u(x) = 0.$$

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

$$\frac{d^2 A}{dx^2} + A\omega^2 \left(\frac{1}{v^2} - \frac{d\psi}{dx} \frac{d\psi}{dx} \right) + i\omega \left[2 \frac{dA}{dx} \frac{d\psi}{dx} + A \frac{d^2 \psi}{dx^2} \right] = 0,$$

which is analogous to equation (6.10.19) and, hence, leads to system (6.10.20), as required.

EXERCISE 6.19. ²⁹*Using stress-strain equations (6.1.1) and Cauchy's equations of motion (6.1.2), obtain equations of motion for an isotropic inhomogeneous continuum. Discuss these equations in the context of equations (6.1.4).*

SOLUTION 6.19. Considering an inhomogeneous continuum, where Lamé's parameters are functions of position, and in view of definition (1.4.6), we can write equations (6.1.1) as

$$(6.11.35) \quad \sigma_{ij} = \lambda(\mathbf{x}) \delta_{ij} \sum_{k=1}^3 \frac{\partial u_k}{\partial x_k} + \mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j \in \{1, 2, 3\},$$

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 (6.1.2) as

$$(6.11.36) \quad \rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^3 \frac{\partial \sigma_{ij}}{\partial x_j}, \quad i \in \{1, 2, 3\},$$

which are equations of motion for an isotropic inhomogeneous continuum. Using equations (6.11.35), we can write equations (6.11.36) 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

$$\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i} \left[\lambda(\mathbf{x}) \sum_{k=1}^3 \frac{\partial u_k}{\partial x_k} \right] + \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left[\mu(\mathbf{x}) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right],$$

²⁹See also Sections 6.1.1 and 6.10.4.

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

$$\begin{aligned} (6.11.37) \quad \rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} &= (\lambda + \mu) \frac{\partial}{\partial x_i} \sum_{j=1}^3 \frac{\partial u_j}{\partial x_j} + \mu \left(\sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2} \right) u_i \\ &\quad + \frac{\partial \lambda}{\partial x_i} \sum_{j=1}^3 \frac{\partial u_j}{\partial x_j} + \sum_{j=1}^3 \frac{\partial \mu}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \end{aligned}$$

where $i \in \{1, 2, 3\}$. These are equations of motion for an isotropic inhomogeneous continuum.³⁰

Examining equations (6.11.37), we notice that if ρ , λ and μ are constants — as is the case for homogeneous continua — equations (6.11.37) reduce to equations (6.1.4), as expected. Also we notice that invoking definitions (1.4.18) and (1.5.2) as well as identity (6.1.7) 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, $\mathbf{\Psi}$. 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 (6.11.36) using only φ and $\mathbf{\Psi}$. Consequently, we cannot split equations (6.11.37) 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.

³⁰Readers 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 2* and in *Part 3*.

7.1. Formulation of equations

In Chapter 6, the wave equation is derived by considering Cauchy's equations of motion (2.8.1) and expressing the stress-tensor components therein in terms of stress-strain equations (5.12.4), 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 (2.8.1) 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 (2.8.1), consider equations

$$(7.1.1) \quad \rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^3 \frac{\partial \sigma_{ij}}{\partial x_j}, \quad i \in \{1, 2, 3\}.$$

In equations (7.1.1), due to the inhomogeneity of the continuum, mass density, $\rho(\mathbf{x})$, is a function of position; unlike for a homogeneous continuum, where ρ is constant. The components of the displacement vector, \mathbf{u} , as well as the stress-tensor components, σ_{ij} , are also functions of position. Since \mathbf{u} and σ_{ij} are also functions of \mathbf{x} in homogeneous continua, to emphasize the distinction between a homogeneous and an inhomogeneous continuum, we explicitly state the \mathbf{x} -dependence only for the mass density, ρ , and for the elasticity tensor, c_{ijkl} , below. The stress-strain equations that account for the inhomogeneity of the continuum are expressed by rewriting equations (3.2.1) as

$$(7.1.2) \quad \sigma_{ij} = \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \varepsilon_{kl}, \quad i, j \in \{1, 2, 3\},$$

where the elasticity tensor is a functions of position. In view of the properties of the elasticity tensor that were discussed in Chapter 5, c_{ijkl} also describes the anisotropy of a given continuum. Thus, $\rho(\mathbf{x})$ and $c_{ijkl}(\mathbf{x})$ fully describe a given anisotropic inhomogeneous continuum. Using definition (1.4.6), we can rewrite stress-strain equations (7.1.2) as

$$(7.1.3) \quad \sigma_{ij} = \frac{1}{2} \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right), \quad i, j \in \{1, 2, 3\}.$$

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 (7.1.3) into equations of motion (7.1.1), we obtain

$$(7.1.4) \quad \rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left[\frac{1}{2} \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right) \right] \\ = \frac{1}{2} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \left(\frac{\partial^2 u_k}{\partial x_j \partial x_l} + \frac{\partial^2 u_l}{\partial x_j \partial x_k} \right) \right],$$

where $i \in \{1, 2, 3\}$. Equations (7.1.4) are equations of motion in anisotropic inhomogeneous continua. For isotropic continua, these equations reduce to equations (6.11.37), and for isotropic homogeneous continua, they further reduce to equations (6.1.4). Equations (7.1.4) are complicated differential equations and, in general, we are unable to find their solutions analytically.

7.2. Formulation of solutions

7.2.1. Introductory comments. Below, we formulate solutions to equations (7.1.4). To elucidate the meaning of our trial solution, we show three related approaches that lead to similar results.

7.2.2. Trial-solution formulation: General wave. To investigate equations (7.1.4), let us consider the trial solution that is a function of position, \mathbf{x} , and time, t , given by

$$(7.2.1) \quad \mathbf{u}(\mathbf{x}, t) = \mathbf{A}(\mathbf{x}) f(\eta),$$

where \mathbf{A} is a vector function of position, \mathbf{x} , and f is a scalar function whose argument is given by

$$(7.2.2) \quad \eta = \omega [\psi(\mathbf{x}) - t],$$

with ω being a constant with units of frequency. Function $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$ is called the eikonal function. Since both \mathbf{u} and \mathbf{A} have units of distance, we require f to be dimensionless. Since commonly $f(\eta)$ is an exponential function, as in expression (6.10.15), we made η dimensionless.

To see the physical meaning of this trial solution, consider function $f(\eta)$ in the context of trial solutions (6.10.6), (6.10.14) and (6.10.15). 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, equation

$$(7.2.3) \quad \psi(\mathbf{x}) - t = t_i,$$

where t_i is a constant, describes a moving wavefront. Function f gives the waveform as a function of time, and \mathbf{A} is the spatially variable amplitude of this waveform.

Inserting trial solution (7.2.1) into equations (7.1.4), using the symmetries of the elasticity tensor, c_{ijkl} , and the equality of mixed partial derivatives, we obtain

$$\begin{aligned} \omega^2 \rho(\mathbf{x}) A_i(\mathbf{x}) \frac{d^2 f}{d\eta^2} &= \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left\{ f \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x}) \frac{\partial^2 A_k}{\partial x_j \partial x_l} \right] \right. \\ &+ \omega \frac{df}{d\eta} \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} A_l \frac{\partial \psi}{\partial x_k} c_{ijkl}(\mathbf{x}) \left(\frac{\partial A_l}{\partial x_j} \frac{\partial \psi}{\partial x_k} + \frac{\partial A_k}{\partial x_l} \frac{\partial \psi}{\partial x_j} + A_l \frac{\partial^2 \psi}{\partial x_j \partial x_k} \right) \right] \\ &\left. + \omega^2 \frac{d^2 f}{d\eta^2} \left[c_{ijkl}(\mathbf{x}) A_k \frac{\partial \psi}{\partial x_j} \frac{\partial \psi}{\partial x_l} \right] \right\}, \end{aligned}$$

where $i \in \{1, 2, 3\}$. Taking all terms to the right-hand side of the equations and, therein, gathering the terms with the same derivatives of f , we get

$$\begin{aligned} (7.2.4) \quad 0 &= \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left\{ f \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x}) \frac{\partial^2 A_k}{\partial x_j \partial x_l} \right] \right. \\ &+ \omega \frac{df}{d\eta} \left[\frac{\partial}{\partial x_j} \left(c_{ijkl}(\mathbf{x}) A_l \frac{\partial \psi}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \frac{\partial A_k}{\partial x_l} \frac{\partial \psi}{\partial x_j} \right] \\ &\left. + \omega^2 \frac{d^2 f}{d\eta^2} \left[c_{ijkl}(\mathbf{x}) A_k \frac{\partial \psi}{\partial x_j} \frac{\partial \psi}{\partial x_l} \right] \right\} - \omega^2 \rho(\mathbf{x}) A_i(\mathbf{x}) \frac{d^2 f}{d\eta^2}, \end{aligned}$$

where $i \in \{1, 2, 3\}$.

Concisely, equation (7.2.4) can be stated as

$$a(\mathbf{x}) f'' + b(\mathbf{x}) f' + c(\mathbf{x}) f = 0,$$

where, in view of $f = f(\eta)$ being a single-variable function, we write $f'' := d^2 f / d\eta^2$ and $f' := df / d\eta$. For equation (7.2.4) to be satisfied by an arbitrary f with its first and second derivatives, each of the

coefficients — $a(\mathbf{x})$, $b(\mathbf{x})$ and $c(\mathbf{x})$ — must be zero.¹ As we can see from the physical meaning of trial solution (7.2.1), we require the arbitrariness of f in order to allow any function to describe the waveform. However, we need f , f' and f'' to be linearly independent, as illustrated in Exercise 7.1.

Setting each of the three coefficients to zero and assuming that $\omega \neq 0$, we obtain three systems of three equations, namely,

$$(7.2.5) \quad \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{\partial \psi}{\partial x_j} \frac{\partial \psi}{\partial x_l} A_k - \rho(\mathbf{x}) A_i = 0, \quad i \in \{1, 2, 3\},$$

$$(7.2.6) \quad \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[\frac{\partial}{\partial x_j} \left(c_{ijkl}(\mathbf{x}) A_l \frac{\partial \psi}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \frac{\partial A_k}{\partial x_l} \frac{\partial \psi}{\partial x_j} \right] = 0, \quad i \in \{1, 2, 3\}$$

and

$$(7.2.7) \quad \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left(\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x}) \frac{\partial^2 A_k}{\partial x_j \partial x_l} \right) = 0, \quad i \in \{1, 2, 3\},$$

which correspond to $a(\mathbf{x})$, $b(\mathbf{x})$ and $c(\mathbf{x})$, respectively. Equations (7.2.5), (7.2.6) and (7.2.7) constitute an overdetermined system of equations for $\psi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$ that results from inserting trial solution (7.2.1) into equations (7.1.4). We obtain a unique solution for functions $\psi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$ using equations (7.2.5) and (7.2.6). Equation (7.2.7) is a degenerate form of equations that result from the asymptotic-series nature of the trial solution, which will be introduced in Section 7.2.4.

7.2.3. Trial-solution formulation: Harmonic wave. Herein, we will follow the approach analogous to the one used in Section 6.10.4. It allows us to illustrate that equations (7.2.5), (7.2.6) and (7.2.7) result from the high-frequency approximation.

Recall equations (7.1.4); namely,

$$(7.2.8) \quad \rho(\mathbf{x}) \frac{\partial^2 u_i(\mathbf{x}, t)}{\partial t^2} = \frac{1}{2} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \left(\frac{\partial^2 u_k}{\partial x_j \partial x_l} + \frac{\partial^2 u_l}{\partial x_j \partial x_k} \right) \right],$$

where $i \in \{1, 2, 3\}$. We wish equations (7.2.8) to be expressed in terms of \mathbf{x} and ω in order to consider the limit of ω tending to infinity. For this purpose we will perform Fourier's transform of equations (7.2.8) with t and ω being the variables of transformation. Invoking Fourier's transform as illustrated in Section 6.9.2, we get

$$(7.2.9) \quad (i\omega)^2 \rho(\mathbf{x}) \tilde{u}_i(\mathbf{x}, \omega) = \frac{1}{2} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \left(\frac{\partial \tilde{u}_k}{\partial x_l} + \frac{\partial \tilde{u}_l}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \left(\frac{\partial^2 \tilde{u}_k}{\partial x_j \partial x_l} + \frac{\partial^2 \tilde{u}_l}{\partial x_j \partial x_k} \right) \right],$$

¹Readers interested in an analogous formulation of the three vanishing terms of equation (7.2.4) might refer to Červený, V., (2001) *Seismic ray theory*: Cambridge University Press, p. 55, pp. 57 – 58 and pp. 62 – 63.

where

$$\tilde{u}_i(\mathbf{x}, \omega) := \frac{1}{2\pi} \int_{-\infty}^{\infty} u_i(\mathbf{x}, t) \exp(-i\omega t) dt,$$

and $i \in \{1, 2, 3\}$. In a manner analogous to those used in Sections 6.10.4 and 7.2.2, let us consider the trial solution given by

$$(7.2.10) \quad \tilde{\mathbf{u}}(\mathbf{x}, \omega) = \mathbf{A}(\mathbf{x}) \exp\{i\omega[\psi(\mathbf{x})]\},$$

which is a generalization of expression (6.10.15) and a particular case of expression (7.2.1). However, with trial solution (7.2.10), we cannot use exactly the method presented in Section 7.2.2, as illustrated in Exercise 7.1. This is due to the fact that — in view of expressions (7.2.1) and (7.2.2) — we see that $\eta = i\omega[\psi(\mathbf{x})]$, and hence $d^2f(\eta)/d\eta^2$, $df(\eta)/d\eta$ and $f(\eta)$ are not linearly independent; they are equal to each other. In the method presented below we will take derivatives with respect to x_i , not with respect to the whole argument, η .

Inserting expression (7.2.10) into equations (7.2.9), differentiating and factoring out the common term, $\exp(i\omega\psi)$, we get

$$(7.2.11) \quad -\omega^2 \exp[i\omega\psi(\mathbf{x})] \rho A_i = \frac{1}{2} \exp[i\omega\psi(\mathbf{x})] \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[\frac{\partial c_{ijkl}}{\partial x_j} \left(\left(\frac{\partial A_k}{\partial x_l} + \frac{\partial A_l}{\partial x_k} + i\omega \left(\frac{\partial \psi}{\partial x_l} A_k + \frac{\partial \psi}{\partial x_k} A_l \right) \right) \right) + c_{ijkl} \left(\left(\frac{\partial^2 A_k}{\partial x_j \partial x_l} + \frac{\partial^2 A_l}{\partial x_j \partial x_k} \right) \right. \right. \\ \left. \left. + i\omega \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_k + \frac{\partial \psi}{\partial x_j} \frac{\partial A_l}{\partial x_k} + \frac{\partial^2 \psi}{\partial x_j \partial x_k} A_l + \frac{\partial \psi}{\partial x_l} \frac{\partial A_k}{\partial x_j} + \frac{\partial \psi}{\partial x_k} \frac{\partial A_l}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_l}{\partial x_k} \right) \right. \right. \\ \left. \left. - \omega^2 \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_k + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_l \right) \right) \right],$$

with $i \in \{1, 2, 3\}$. Canceling the exponential term and considering only the real part, we write

$$-\omega^2 \rho A_i = \frac{1}{2} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left\{ \frac{\partial c_{ijkl}}{\partial x_j} \left(\frac{\partial A_k}{\partial x_l} + \frac{\partial A_l}{\partial x_k} \right) \right. \\ \left. + c_{ijkl} \left[\left(\frac{\partial^2 A_k}{\partial x_j \partial x_l} + \frac{\partial^2 A_l}{\partial x_j \partial x_k} \right) - \omega^2 \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_k + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_l \right) \right] \right\},$$

where $i \in \{1, 2, 3\}$. Dividing both sides by ω and letting ω tend to infinity, we get

$$\frac{1}{2} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_k + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_l \right) - \rho A_i = 0, \quad i \in \{1, 2, 3\}.$$

Since j, k and l are summation indices and c_{ijkl} is symmetric in k and l , we can write these three equations as

$$\sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_k - \rho(\mathbf{x}) A_i = 0, \quad i \in \{1, 2, 3\},$$

which are equations (7.2.5). In view of the limit of ω , the presented method is the high-frequency approximation.

Considering the imaginary part, we obtain

$$\begin{aligned} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[\frac{\partial c_{ijkl}}{\partial x_j} \left(\frac{\partial \psi}{\partial x_l} A_k + \frac{\partial \psi}{\partial x_k} A_l \right) \right. \\ \left. + c_{ijkl} \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_k + \frac{\partial \psi}{\partial x_j} \frac{\partial A_l}{\partial x_k} + \frac{\partial^2 \psi}{\partial x_j \partial x_k} A_l + \frac{\partial \psi}{\partial x_l} \frac{\partial A_k}{\partial x_j} + \frac{\partial \psi}{\partial x_k} \frac{\partial A_l}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_l}{\partial x_k} \right) \right] = 0, \end{aligned}$$

where $i \in \{1, 2, 3\}$, which are equations (7.2.6) with partial derivatives $\partial/\partial x_j$ taken.

Also examining equations (7.2.11), we see that — in a manner analogous to the one used in Section 7.2.2 — we could also obtain equations (7.2.5) by setting the coefficients of ω^2 to zero.

7.2.4. Asymptotic-series formulation. To gain insight into the fact that trial solution (7.2.10) is an asymptotic series, and, hence, the resulting equations belong to the asymptotic ray theory, let us consider the following formulation. As discussed in Section 6.10.4, the trial solution is the zeroth-order term of an asymptotic series. Expression (7.2.10) is the zeroth-order term of

$$(7.2.12) \quad \tilde{\mathbf{u}}(\mathbf{x}, \omega) \sim \exp[i\omega\psi(\mathbf{x})] \left\{ \mathbf{A}_0(\mathbf{x}) + \frac{\mathbf{A}_1(\mathbf{x})}{i\omega} + \sum_{n=2}^{\infty} \frac{\mathbf{A}_n(\mathbf{x})}{(i\omega)^n} \right\},$$

where \sim stands for “is asymptotically equivalent to”. This is a series whose bases are $\exp(i\omega\psi)/(i\omega)^n$, where $n \in \{0, \dots, \infty\}$. Series (7.2.12) is an asymptotic series of $\tilde{\mathbf{u}}$; herein, the defining property of the asymptotic series is

$$(7.2.13) \quad \lim_{\omega \rightarrow \infty} \left\{ \frac{(i\omega)^N}{\exp[i\omega\psi(\mathbf{x})]} \left[\tilde{\mathbf{u}}(\mathbf{x}, \omega) - \exp[i\omega\psi(\mathbf{x})] \sum_{n=0}^N \frac{\mathbf{A}_n(\mathbf{x})}{(i\omega)^n} \right] \right\} = 0,$$

for all N . This statement implies that the difference between the exact solution and the partial sum of N terms of the series vanishes as ω tends to infinity — hence, the name “high-frequency approximation”; this difference tends to zero more rapidly than $(i\omega)^N / \exp(i\omega\psi)$ tends to infinity. For each function that has an asymptotic series, coefficients \mathbf{A}_n are unique, as illustrated in Exercise 7.2.

Prior to inserting the asymptotic series into the equations of motion, let us explain our motivation for using series (7.2.12). In particular, let us discuss the importance of the exponential term. What is the reason for

considering an asymptotic series given by the product of the exponential term and the summation,

$$(7.2.14) \quad \exp[i\omega\psi(\mathbf{x})] \sum_{n=0}^N \frac{\mathbf{A}_n(\mathbf{x})}{(i\omega)^n},$$

rather than the summation itself? An insight into the physical reason is given by the meanings of ψ and \mathbf{A} , which are associated with wavefronts and amplitudes, respectively. This is, however, an insight achieved by examining the solutions to the equations of motion. The mathematical reason to begin with such a series to obtain the solution can be explained as follows. The bases of our asymptotic expansion, which we will use in differential equations, are $\exp(i\omega\psi)/(i\omega)^n$. Consider the derivatives of expression (7.2.14). Each time we take a derivative of the exponential term with respect to x_i , we get $i\omega(\partial\psi/\partial x_i)\exp(i\omega\psi)$. The derivatives of expression (7.2.14) are linear combinations of $\exp(i\omega\psi)/(i\omega)^n$; we get such combinations even if we consider only the zeroth-order term of series (7.2.12), as we can see by examining expression (7.2.11).

Note that the units of the n th term in the denominator of summation (7.2.14) are s^{-n} , where s stands for seconds. Hence, the units of the numerator, \mathbf{A}_n , must change with n in such a way that all the terms in the series have the same units.

Inserting the first two terms of series (7.2.12) into the left-hand side of equations (7.2.9), we write

$$(7.2.15) \quad (i\omega)^2 \rho(\mathbf{x}) \exp[i\omega\psi(\mathbf{x})] \left[A_{0_i}(\mathbf{x}) + (i\omega)^{-1} A_{1_i}(\mathbf{x}) \right], \quad i \in \{1, 2, 3\}.$$

Inserting the first two terms of series (7.2.12) into a first-derivative term on the right-hand side of equations (7.2.9), we write

$$\frac{\partial \tilde{u}_k}{\partial x_l} = \frac{\partial}{\partial x_l} \left[\exp[i\omega\psi(\mathbf{x})] \left(A_{0_k} + \frac{A_{1_k}}{i\omega} \right) \right].$$

Differentiating, factoring out the common term, $\exp(i\omega\psi)$, and gathering the terms with common powers of $i\omega$, we get

$$(7.2.16) \quad \frac{\partial \tilde{u}_k}{\partial x_l} = \exp[i\omega\psi(\mathbf{x})] \left[(i\omega)^{-1} \left(\frac{\partial A_{1_k}}{\partial x_l} \right) + (i\omega)^0 \left(\frac{\partial \psi}{\partial x_l} A_{1_k} + \frac{\partial A_{0_k}}{\partial x_l} \right) + (i\omega)^1 \left(\frac{\partial \psi}{\partial x_l} A_{0_k} \right) \right].$$

Performing the same operations for $\partial \tilde{u}_l/\partial x_k$, adding together the two results, and gathering the terms with common powers of $i\omega$, we get

$$\begin{aligned} \frac{\partial \tilde{u}_k}{\partial x_l} + \frac{\partial \tilde{u}_l}{\partial x_k} &= \exp[i\omega\psi(\mathbf{x})] \\ &\left[(i\omega)^{-1} \left(\frac{\partial A_{1_k}}{\partial x_l} + \frac{\partial A_{1_l}}{\partial x_k} \right) + (i\omega)^0 \left(\frac{\partial \psi}{\partial x_l} A_{1_k} + \frac{\partial \psi}{\partial x_k} A_{1_l} + \frac{\partial A_{0_k}}{\partial x_l} + \frac{\partial A_{0_l}}{\partial x_k} \right) + (i\omega)^1 \left(\frac{\partial \psi}{\partial x_l} A_{0_k} + \frac{\partial \psi}{\partial x_k} A_{0_l} \right) \right]. \end{aligned}$$

To insert the first two terms of series (7.2.12) into a second-derivative term, we write

$$\frac{\partial^2 \tilde{u}_k}{\partial x_j \partial x_l} = \frac{\partial}{\partial x_j} \frac{\partial \tilde{u}_k}{\partial x_l},$$

where $\partial \tilde{u}_k / \partial x_l$ is given in expression (7.2.16). Differentiating, factoring out the common term, $\exp(i\omega\psi)$, and gathering the terms with common powers of $i\omega$, we get

$$\begin{aligned} \frac{\partial^2 \tilde{u}_k}{\partial x_j \partial x_l} = \exp\{i\omega[\psi(\mathbf{x})]\} & \left[(i\omega)^{-1} \left(\frac{\partial^2 A_{1_k}}{\partial x_j \partial x_l} \right) + (i\omega)^0 \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_{1_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial A_{1_k}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{1_k}}{\partial x_l} + \frac{\partial^2 A_{0_k}}{\partial x_j \partial x_l} \right) \right. \\ & \left. + (i\omega)^1 \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_{0_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{1_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial A_{0_k}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{0_k}}{\partial x_l} \right) + (i\omega)^2 \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{0_k} \right) \right]. \end{aligned}$$

Performing the same operations for $\partial^2 \tilde{u}_l / \partial x_j \partial x_k$, adding together the two results, and gathering the terms with common powers of $i\omega$, we get

$$\begin{aligned} \frac{\partial^2 \tilde{u}_k}{\partial x_j \partial x_l} + \frac{\partial^2 \tilde{u}_l}{\partial x_j \partial x_k} = \exp[i\omega\psi(\mathbf{x})] & \left[(i\omega)^{-1} \left(\frac{\partial^2 A_{1_k}}{\partial x_j \partial x_l} + \frac{\partial^2 A_{1_l}}{\partial x_j \partial x_k} \right) \right. \\ & + (i\omega)^0 \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_{1_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial A_{1_k}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{1_k}}{\partial x_l} + \frac{\partial^2 A_{0_k}}{\partial x_j \partial x_l} \right. \\ & \quad \left. + \frac{\partial^2 \psi}{\partial x_j \partial x_k} A_{1_l} + \frac{\partial \psi}{\partial x_k} \frac{\partial A_{1_l}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{1_l}}{\partial x_k} + \frac{\partial^2 A_{0_l}}{\partial x_j \partial x_k} \right) \\ & + (i\omega)^1 \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_{0_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{1_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial A_{0_k}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{0_l}}{\partial x_k} \right. \\ & \quad \left. + \frac{\partial^2 \psi}{\partial x_j \partial x_k} A_{0_l} + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_{1_l} + \frac{\partial \psi}{\partial x_k} \frac{\partial A_{0_l}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{0_k}}{\partial x_l} \right) \\ & \left. + (i\omega)^2 \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{0_k} + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_{0_l} \right) \right]. \end{aligned}$$

Returning to equations (7.2.9), we can write the right-hand side as

$$(7.2.17) \quad \frac{1}{2} \exp[i\omega\psi(\mathbf{x})] \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left[a(\mathbf{x}) (i\omega)^2 + b(\mathbf{x}) (i\omega)^1 + c(\mathbf{x}) (i\omega)^0 + d(\mathbf{x}) (i\omega)^{-1} \right],$$

where

$$a(\mathbf{x}) = c_{ijkl}(\mathbf{x}) \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{0_k} + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_{0_l} \right),$$

$$\begin{aligned} b(\mathbf{x}) = \frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} & \left(\frac{\partial \psi}{\partial x_l} A_{0_k} + \frac{\partial \psi}{\partial x_k} A_{0_l} \right) \\ & + c_{ijkl}(\mathbf{x}) \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_{0_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{1_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial A_{0_k}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{0_l}}{\partial x_k} \right. \\ & \quad \left. + \frac{\partial^2 \psi}{\partial x_j \partial x_k} A_{0_l} + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_{1_l} + \frac{\partial \psi}{\partial x_k} \frac{\partial A_{0_l}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{0_k}}{\partial x_l} \right), \end{aligned}$$

$$\begin{aligned}
c(\mathbf{x}) = \frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} & \left(\frac{\partial \psi}{\partial x_l} A_{1_k} + \frac{\partial \psi}{\partial x_k} A_{1_l} + \frac{\partial A_{0_k}}{\partial x_l} + \frac{\partial A_{0_l}}{\partial x_k} \right) \\
& + c_{ijkl}(\mathbf{x}) \left(\frac{\partial^2 \psi}{\partial x_j \partial x_l} A_{1_k} + \frac{\partial \psi}{\partial x_l} \frac{\partial A_{1_k}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{1_k}}{\partial x_l} + \frac{\partial^2 A_{0_k}}{\partial x_j \partial x_l} \right. \\
& \left. + \frac{\partial^2 \psi}{\partial x_j \partial x_k} A_{1_l} + \frac{\partial \psi}{\partial x_k} \frac{\partial A_{1_l}}{\partial x_j} + \frac{\partial \psi}{\partial x_j} \frac{\partial A_{1_l}}{\partial x_k} + \frac{\partial^2 A_{0_l}}{\partial x_j \partial x_k} \right)
\end{aligned}$$

and

$$d(\mathbf{x}) = \frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \left(\frac{\partial A_{1_k}}{\partial x_l} + \frac{\partial A_{1_l}}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \left(\frac{\partial^2 A_{1_k}}{\partial x_j \partial x_l} + \frac{\partial^2 A_{1_l}}{\partial x_j \partial x_k} \right),$$

with $i \in \{1, 2, 3\}$. Considering the left-hand side of equations (7.2.9) that includes the two terms of series (7.2.12), which is given in expression (7.2.15), we write it as

$$(7.2.18) \quad \exp[i\omega\psi(\mathbf{x})] \left[(i\omega)^2 \rho(\mathbf{x}) A_{0_i}(\mathbf{x}) + \omega \rho(\mathbf{x}) A_{1_i}(\mathbf{x}) \right], \quad i \in \{1, 2, 3\}.$$

Comparing the right-hand and left-hand sides, which are given by expressions (7.2.17) and (7.2.18), respectively, we cancel $\exp(i\omega\psi)$. Then, moving all the terms to one side, we write the resulting coefficient of $(i\omega)^2$ as

$$\frac{1}{2} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \left(\frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{0_k} + \frac{\partial \psi}{\partial x_k} \frac{\partial \psi}{\partial x_j} A_{0_l} \right) - \rho(\mathbf{x}) A_{0_i}(\mathbf{x}), \quad i \in \{1, 2, 3\}.$$

Since j, k and l are summation indices and c_{ijkl} is symmetric in k and l , we can write these three expressions as

$$\sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{0_k}(\mathbf{x}) - \rho(\mathbf{x}) A_{0_i}(\mathbf{x}), \quad i \in \{1, 2, 3\}.$$

For the resulting equations of motion to be satisfied for all $i\omega$, we require all the coefficients to be zero. Again considering the $(i\omega)^2$ coefficient, we write

$$(7.2.19) \quad \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_{0_k} - \rho(\mathbf{x}) A_{0_i} = 0, \quad i \in \{1, 2, 3\},$$

which are equations (7.2.5) in terms of \mathbf{A}_0 . These equations will lead us to the eikonal equation whose solution is $\psi(\mathbf{x})$; this equation is associated with the path of the propagating wave.

We could also write the coefficients of $(i\omega)^{-1}$, $(i\omega)^0$ and $(i\omega)^1$ and set them to zero. The solutions of corresponding equations are $\mathbf{A}_n(\mathbf{x})$. These equations are associated with the amplitude of the propagating wave and are called transport equations.

Let us return to series (7.2.12). Inserting more terms of this series into equations (7.2.9) would result in terms with higher negative powers of $i\omega$ in expression (7.2.17). Coefficients $a(\mathbf{x})$ and $b(\mathbf{x})$ would remain unchanged; in particular, equations (7.2.5) would remain the same — these equations are always expressed in the zeroth-order term of series (7.2.12). Keeping this property in mind, we can rewrite equations (7.2.19)

as

$$\sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{\partial \psi}{\partial x_l} \frac{\partial \psi}{\partial x_j} A_k - \rho(\mathbf{x}) A_i = 0, \quad i \in \{1, 2, 3\},$$

which are equations (7.2.5). Coefficient $c(\mathbf{x})$ would have terms with \mathbf{A}_0 , \mathbf{A}_1 and \mathbf{A}_2 ; coefficient $d(\mathbf{x})$ would have terms with \mathbf{A}_1 , \mathbf{A}_2 and \mathbf{A}_3 , but would no longer have any term with \mathbf{A}_0 . Each coefficient would have three \mathbf{A} s. Thus, a new coefficient, $e(\mathbf{x})$, would have terms with \mathbf{A}_2 , \mathbf{A}_3 and \mathbf{A}_4 . The pattern would continue.

7.3. Eikonal equation

In order to obtain $\psi(\mathbf{x})$, we turn our attention to equations (7.2.5), from which we can factor out the components of vector $\mathbf{A}(\mathbf{x})$. Hence, we rewrite equation (7.2.5) as

$$(7.3.1) \quad \sum_{k=1}^3 \left(\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{\partial \psi}{\partial x_j} \frac{\partial \psi}{\partial x_l} - \rho(\mathbf{x}) \delta_{ik} \right) A_k(\mathbf{x}) = 0, \quad i \in \{1, 2, 3\}.$$

In view of expression (6.10.16), let us denote

$$(7.3.2) \quad p_j := \frac{\partial \psi}{\partial x_j}, \quad j \in \{1, 2, 3\},$$

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 (7.2.2) and considering a three-dimensional continuum. Therein, ψ is a function relating position variables, x_1 , x_2 and x_3 , to the traveltimes, 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 (7.3.2), we can write equations (7.3.1) as

$$(7.3.3) \quad \sum_{k=1}^3 \left(\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) p_j p_l - \rho(\mathbf{x}) \delta_{ik} \right) A_k(\mathbf{x}) = 0, \quad i \in \{1, 2, 3\}.$$

Equations (7.3.3) are referred to as Christoffel's equations.

In Chapter 9, we discuss equations (7.3.3) 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 (7.3.3).

We know from linear algebra that equations (7.3.3) have nontrivial solutions if and only if

$$(7.3.4) \quad \det \left[\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) p_j p_l - \rho(\mathbf{x}) \delta_{ik} \right] = 0, \quad i, k \in \{1, 2, 3\}.$$

Assuming that $p^2 \neq 0$, we can write determinant (7.3.4) as

$$(7.3.5) \quad (p^2)^3 \det \left[\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} - \frac{\rho(\mathbf{x})}{p^2} \delta_{ik} \right] = 0, \quad i, k \in \{1, 2, 3\}.$$

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 (7.3.4), $p^2 = 0$ would result in $\det [\rho(\mathbf{x}) \delta_{ik}] = 0$, which would imply $\rho(\mathbf{x}) = 0$.

Expression (7.3.5) 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

$$(7.3.6) \quad \left[p^2 - \frac{1}{v_1^2(\mathbf{x}, \mathbf{p})} \right] \left[p^2 - \frac{1}{v_2^2(\mathbf{x}, \mathbf{p})} \right] \left[p^2 - \frac{1}{v_3^2(\mathbf{x}, \mathbf{p})} \right] = 0,$$

where $1/v_i^2$ are the roots of polynomial (7.3.5). The existence of three roots implies the existence of three types of waves, which can propagate in anisotropic continua.

The matrix with entries

$$(7.3.7) \quad \Gamma_{ik} = \left[\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} \right], \quad i, k \in \{1, 2, 3\},$$

which appears in equation (7.3.5) is called Christoffel's matrix. It is symmetric and positive-definite, as shown in Exercises 7.3 and 7.4, respectively. Since the matrix is symmetric, the roots are real; since it is positive-definite, they are positive and, hence, the values of v_i are real.

Now, let us consider a given root of equation (7.3.6). Each root is the eikonal equation for a given type of wave, namely,

$$(7.3.8) \quad p^2 = \frac{1}{v_i^2(\mathbf{x}, \mathbf{p})}, \quad i \in \{1, 2, 3\}.$$

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_i is the function describing phase velocity. This velocity is a function of position, \mathbf{x} , and the direction of \mathbf{p} . Hence, equation (7.3.8) applies to anisotropic inhomogeneous continua and can be viewed as an extension of equation (6.10.23), which

²Readers 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.

is valid for isotropic inhomogeneous continua. As shown above, due to the positive definiteness of Γ , function v_i is real-valued — a property of the body-wave velocities in an elastic continuum, which are discussed in Chapter 9.

Considering two adjacent wavefronts, we can view equation (7.3.8) 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 (7.3.8) 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 9. We can explicitly see the homogeneity of function v in Section 6.10.3, where we discussed waves in the context of elliptical velocity dependence. Therein, vector \mathbf{p} appears as a ratio of its components, namely, the directional dependence is given by expression (6.10.9).

Furthermore, as shown explicitly in Chapter 9, 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 (7.3.2), we can rewrite equation (7.3.8) as

$$(7.3.9) \quad [\nabla \psi(\mathbf{x})]^2 = \frac{1}{v^2(\mathbf{x}, \mathbf{p})}.$$

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 (7.2.6) is the transport equation. This transport equation possesses a vectorial form that is valid for anisotropic inhomogeneous continua. It is analogous to the scalar transport equation (6.10.26), 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

³Readers interested in a formulation relating the eikonal equation to Huygens' principle might refer to Arnold, V.I., (1989) *Mathematical methods of classical mechanics* (2nd edition): Springer-Verlag, pp. 248 – 252, and to Lanczos, C., (1949/1986) *The variational principles in mechanics*: Dover, pp. 269 – 270.

⁴Readers 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.

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.



7.4. Exercises

EXERCISE 7.1. Show that trial solution (7.2.1) with $f(\eta) = \exp \eta$, which corresponds to solution (7.2.10), does not allow us to obtain equations (7.2.5), (7.2.6) and (7.2.7).

SOLUTION 7.1. Letting $f(\eta) = \exp \eta$ in expression (7.2.1), inserting the resulting expression into equation (7.1.4), and factoring out the common term,

$$\exp \eta = f = \frac{df}{d\eta} = \frac{d^2 f}{d\eta^2},$$

we write

$$\begin{aligned} \exp \eta & \left(\sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \left\{ \left[\frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \frac{\partial A_k}{\partial x_l} + c_{ijkl}(\mathbf{x}) \frac{\partial^2 A_k}{\partial x_j \partial x_l} \right] \right. \right. \\ & \left. \left. + \left[\frac{\partial}{\partial x_j} \left(c_{ijkl}(\mathbf{x}) A_l \frac{\partial \psi}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \frac{\partial A_k}{\partial x_l} \frac{\partial \psi}{\partial x_j} \right] + \left[c_{ijkl}(\mathbf{x}) A_k \frac{\partial \psi}{\partial x_j} \frac{\partial \psi}{\partial x_l} \right] \right\} - \rho(\mathbf{x}) A_i \right) = 0, \end{aligned}$$

where $i \in \{1, 2, 3\}$. Since $\exp \eta \neq 0$, we obtain

$$\begin{aligned} \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 & \left\{ \frac{\partial c_{ijkl}(\mathbf{x})}{\partial x_j} \frac{\partial A_k}{\partial x_l} + \frac{\partial}{\partial x_j} \left(c_{ijkl}(\mathbf{x}) A_l \frac{\partial \psi}{\partial x_k} \right) + c_{ijkl}(\mathbf{x}) \left[\frac{\partial^2 A_k}{\partial x_j \partial x_l} + \frac{\partial A_k}{\partial x_l} \frac{\partial \psi}{\partial x_j} \right. \right. \\ & \left. \left. + A_k \frac{\partial \psi}{\partial x_j} \frac{\partial \psi}{\partial x_l} \right] \right\} - \rho(\mathbf{x}) A_i = 0, \quad i \in \{1, 2, 3\}. \end{aligned}$$

These are complicated differential equations for ψ and \mathbf{A} . Since f , f' and f'' are not linearly independent, we cannot split their corresponding terms into three distinct equations as we did in Section 7.2.2. Thus, we cannot proceed to solve them by methods that lead to the eikonal and transport equations.

EXERCISE 7.2. Expression (7.2.13), namely,

$$(7.4.1) \quad \lim_{\omega \rightarrow \infty} \left\{ \frac{(i\omega)^N}{\exp[i\omega\psi(\mathbf{x})]} \left[\tilde{\mathbf{u}}(\mathbf{x}, \omega) - \exp[i\omega\psi(\mathbf{x})] \sum_{n=0}^N \frac{\mathbf{A}_n(\mathbf{x})}{(i\omega)^n} \right] \right\} = 0,$$

allows us to determine uniquely all \mathbf{A}_n . Determine \mathbf{A}_0 , \mathbf{A}_1 and \mathbf{A}_2 .

SOLUTION 7.2. For $N = 0$ also $n = 0$, and we write expression (7.4.1) as

$$\lim_{\omega \rightarrow \infty} \left\{ \frac{1}{\exp[i\omega\psi(\mathbf{x})]} [\tilde{\mathbf{u}}(\mathbf{x}, \omega) - \exp[i\omega\psi(\mathbf{x})] \mathbf{A}_0(\mathbf{x})] \right\} = 0,$$

which means that

$$(7.4.2) \quad \mathbf{A}_0(\mathbf{x}) = \lim_{\omega \rightarrow \infty} \frac{\tilde{\mathbf{u}}(\mathbf{x}, \omega)}{\exp[i\omega\psi(\mathbf{x})]}.$$

For $N = 1$, $n = 0, 1$, and we write expression (7.4.1) as

$$\lim_{\omega \rightarrow \infty} \left\{ \frac{i\omega}{\exp[i\omega\psi(\mathbf{x})]} \left[\tilde{\mathbf{u}}(\mathbf{x}, \omega) - \exp[i\omega\psi(\mathbf{x})] \left[\mathbf{A}_0(\mathbf{x}) + \frac{\mathbf{A}_1(\mathbf{x})}{i\omega} \right] \right] \right\} = 0.$$

We can rewrite it as

$$\lim_{\omega \rightarrow \infty} \left\{ i\omega \left[\frac{\tilde{\mathbf{u}}(\mathbf{x}, \omega)}{\exp[i\omega\psi(\mathbf{x})]} - \left[\mathbf{A}_0(\mathbf{x}) + \frac{\mathbf{A}_1(\mathbf{x})}{i\omega} \right] \right] \right\} = 0$$

to get

$$\lim_{\omega \rightarrow \infty} \left\{ i\omega \left[\frac{\tilde{\mathbf{u}}(\mathbf{x}, \omega)}{\exp[i\omega\psi(\mathbf{x})]} - \mathbf{A}_0(\mathbf{x}) \right] - \mathbf{A}_1(\mathbf{x}) \right\} = 0.$$

This means that

$$(7.4.3) \quad \mathbf{A}_1(\mathbf{x}) = \lim_{\omega \rightarrow \infty} \left\{ i\omega \left[\frac{\tilde{\mathbf{u}}(\mathbf{x}, \omega)}{\exp[i\omega\psi(\mathbf{x})]} - \mathbf{A}_0(\mathbf{x}) \right] \right\},$$

where \mathbf{A}_0 is known from equation (7.4.2). For $N = 2$, $n = 1, 2, 3$, and we write expression (7.4.1) as

$$\lim_{\omega \rightarrow \infty} \left\{ \frac{(i\omega)^2}{\exp[i\omega\psi(\mathbf{x})]} \left[\tilde{\mathbf{u}}(\mathbf{x}, \omega) - \exp[i\omega\psi(\mathbf{x})] \left[\mathbf{A}_0(\mathbf{x}) + \frac{\mathbf{A}_1(\mathbf{x})}{i\omega} + \frac{\mathbf{A}_2(\mathbf{x})}{(i\omega)^2} \right] \right] \right\} = 0.$$

We can rewrite it as

$$\lim_{\omega \rightarrow \infty} \left\{ (i\omega)^2 \left[\frac{\tilde{\mathbf{u}}(\mathbf{x}, \omega)}{\exp[i\omega\psi(\mathbf{x})]} - \mathbf{A}_0(\mathbf{x}) \right] - i\omega \mathbf{A}_1(\mathbf{x}) - \mathbf{A}_2(\mathbf{x}) \right\},$$

which means that

$$(7.4.4) \quad \mathbf{A}_2(\mathbf{x}) = \lim_{\omega \rightarrow \infty} \left\{ (i\omega)^2 \left[\frac{\tilde{\mathbf{u}}(\mathbf{x}, \omega)}{\exp[i\omega\psi(\mathbf{x})]} - \mathbf{A}_0(\mathbf{x}) \right] - i\omega \mathbf{A}_1(\mathbf{x}) \right\},$$

where \mathbf{A}_0 and \mathbf{A}_1 are known from equations (7.4.2) and (7.4.3), respectively. Continuing this process we obtain uniquely all \mathbf{A}_n .

EXERCISE 7.3. Show that matrix (7.3.7) is symmetric.

SOLUTION 7.3. To show that matrix (7.3.7) is symmetric, we need to show that $\Gamma_{ik} = \Gamma_{ki}$, which is equivalent to showing that

$$(7.4.5) \quad \sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} = \sum_{j=1}^3 \sum_{l=1}^3 c_{kjil}(\mathbf{x}) \frac{p_j p_l}{p^2}, \quad i, k \in \{1, 2, 3\}.$$

In view of symmetries (4.2.2), we can write the left-hand side of equation (7.4.5) as

$$\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} = \sum_{j=1}^3 \sum_{l=1}^3 c_{klij}(\mathbf{x}) \frac{p_j p_l}{p^2}, \quad i, k \in \{1, 2, 3\}.$$

Examining the right-hand sides of the above equations, we see that — since both $j = 1, 2, 3$ and $l = 1, 2, 3$ — each term that appears in

$$\sum_{j=1}^3 \sum_{l=1}^3 c_{kjil}(\mathbf{x}) \frac{p_j p_l}{p^2}$$

also appears in

$$\sum_{j=1}^3 \sum_{l=1}^3 c_{klij}(\mathbf{x}) \frac{p_j p_l}{p^2}.$$

Thus, we can write

$$\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} = \sum_{j=1}^3 \sum_{l=1}^3 c_{kjil}(\mathbf{x}) \frac{p_j p_l}{p^2}, \quad i, k \in \{1, 2, 3\},$$

which is equation (7.4.5), as required. We conclude that Christoffel's matrix (7.3.7) is symmetric as a consequence of symmetries (4.2.2) of the elasticity matrix, which result from the existence of the strain-energy function.

EXERCISE 7.4. *Show that matrix (7.3.7) is positive-definite.*

SOLUTION 7.4. To show that matrix (7.3.7) is positive-definite, we need to show that

$$(7.4.6) \quad \sum_{i=1}^3 \sum_{k=1}^3 \left[\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} \right] w_i w_k > 0,$$

for an arbitrary nonzero vector, \mathbf{w} . To do so, let us recall equation (4.3.2), which we can rewrite as

$$\sum_{m=1}^6 \sum_{n=1}^6 C_{mn} \varepsilon_m \varepsilon_n > 0,$$

where ε is an arbitrary nonzero vector. Invoking formula (3.2.5), we write

$$(7.4.7) \quad \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \varepsilon_{ij} \varepsilon_{kl} > 0.$$

To proceed, we will use the fact that although, in view of definition (1.4.6), ε_{ij} is a symmetric tensor, inequality (7.4.7) remains valid for general second-rank tensors. To show this fact, let us return to Section 1.5 and recall that any second-rank tensor can be written as a sum of symmetric and antisymmetric tensors. As we will show below, the summation of the antisymmetric parts vanishes. Consider expression

$$\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \eta_{ij} a_{kl},$$

where η is an antisymmetric tensor, $\eta_{ij} = -\eta_{ji}$, and a_{kl} is an arbitrary one. Thus, we can write

$$\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \eta_{ij} a_{kl} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 (-c_{ijkl} \eta_{ji} a_{kl}).$$

Using the symmetry of the elasticity tensor given in expression (3.2.3), which is a result of the symmetry of stress tensor, we write

$$\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \eta_{ij} a_{kl} = \sum_{j=1}^3 \sum_{i=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 (-c_{jikl} \eta_{ji} a_{kl}).$$

Since i and j are summation indices we can rename them so that i becomes j and vice versa. Hence, we write

$$\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \eta_{ij} a_{kl} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 (-c_{ijkl} \eta_{ji} a_{kl}) = - \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \eta_{ij} a_{kl},$$

which implies

$$\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} \eta_{ij} a_{kl} = 0.$$

In other words, the summation of the antisymmetric parts vanishes and, hence, inequality (7.4.7) is valid whether or not the second-rank tensor is symmetric. In particular, we can write

$$(7.4.8) \quad \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} a_{ij} a_{kl} > 0.$$

Since a_{ij} is arbitrary, we consider

$$a_{ij} = w_i \frac{p_j}{|p|}$$

to write inequality (7.4.8) as

$$\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 c_{ijkl} w_i \frac{p_j}{|p|} w_k \frac{p_l}{|p|} > 0.$$

Using the algebraic properties of summation and multiplication, we write

$$\sum_{i=1}^3 \sum_{k=1}^3 \left[\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) \frac{p_j p_l}{p^2} \right] w_i w_k > 0,$$

which is inequality (7.4.6), as required. We conclude that Christoffel's matrix (7.3.7) is positive-definite as a consequence of the positive definiteness of the elasticity matrix, which results from the stability conditions, and the symmetry of the stress tensor, which results from the balance of angular momentum.

EXERCISE 7.5. *In view of Exercise 6.15, 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 (6.11.33) to be satisfied is the eikonal equation given by*

$$(\nabla \psi)^2 = \frac{1}{v^2}.$$

SOLUTION 7.5. Considering the argument of f given by $\eta = v_0 [\psi(\mathbf{x}) - t]$, we obtain

$$\frac{\partial u}{\partial x_i} = \frac{\partial f}{\partial \eta} \frac{\partial \eta}{\partial x_i} = v_0 \frac{\partial f}{\partial \eta} \frac{\partial \psi}{\partial x_i}, \quad i \in \{1, 2, 3\},$$

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 (6.11.33), 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 7.4.1. If v is constant, Exercise 7.5 is reduced to Exercise 6.15.

EXERCISE 7.6. Derive eikonal equation (6.10.22) in two spatial dimensions as the requirement describing a surface in the space of the independent variables, x, z and t , on which the initial conditions do not specify uniquely the second partial derivatives for equation (6.10.13) in these two spatial dimensions.

SOLUTION 7.6. In two spatial dimensions, we explicitly write equation (6.10.13) as

$$(7.4.9) \quad \frac{\partial^2 u(x, z, t)}{\partial x^2} + \frac{\partial^2 u(x, z, t)}{\partial z^2} = \frac{1}{v^2(x, z)} \frac{\partial^2 u(x, z, t)}{\partial t^2}.$$

Let the required surface be $\psi(x, z) = t$. On this surface, $u = u(x, z, \psi(x, z))$ and the initial conditions are the value of u given by

$$(7.4.10) \quad u(x, z, \psi(x, z)) = g(x, z)$$

and the value of the directional derivative of u given by

$$(7.4.11) \quad \frac{\partial u(x, z, \psi(x, z))}{\partial \mathbf{n}} = h(x, z),$$

where \mathbf{n} is a vector normal to the level-set given by $\psi(x, z) - t = 0$, namely,

$$\begin{aligned} \mathbf{n} &= \left[\frac{\partial [\psi(x, z) - t]}{\partial x}, \frac{\partial [\psi(x, z) - t]}{\partial z}, \frac{\partial [\psi(x, z) - t]}{\partial t} \right] \\ &= \left[\frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial z}, -1 \right]. \end{aligned}$$

First let us consider the first derivatives, namely, $\partial u/\partial x$, $\partial u/\partial z$ and $\partial u/\partial t$. Condition (7.4.10) provides two equations. To get them, we differentiate condition (7.4.10) with respect to x , to get

$$\frac{\partial}{\partial x} u(x, z, \psi(x, z)) = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} \frac{\partial \psi}{\partial x} = \frac{\partial g}{\partial x},$$

where we used the fact that ψ is tantamount to t . Differentiating with respect to z , we get

$$\frac{\partial}{\partial z} u(x, z, \psi(x, z)) = \frac{\partial u}{\partial z} + \frac{\partial u}{\partial t} \frac{\partial \psi}{\partial z} = \frac{\partial g}{\partial z}.$$

We can explicitly write condition (7.4.11) as

$$\begin{aligned} \left[\frac{\partial u}{\partial x}, \frac{\partial u}{\partial z}, \frac{\partial u}{\partial t} \right] \cdot \mathbf{n} &= \left[\frac{\partial u}{\partial x}, \frac{\partial u}{\partial z}, \frac{\partial u}{\partial t} \right] \cdot \left[\frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial z}, -1 \right] \\ &= \frac{\partial u}{\partial x} \frac{\partial \psi}{\partial x} + \frac{\partial u}{\partial z} \frac{\partial \psi}{\partial z} - \frac{\partial u}{\partial t} = h(x, z), \end{aligned}$$

which is the third equation. Now, we can write these three equations as a system of linear algebraic equations given by

$$\begin{bmatrix} 1 & 0 & \frac{\partial \psi}{\partial x} \\ 0 & 1 & \frac{\partial \psi}{\partial z} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial z} & -1 \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial z} \\ \frac{\partial u}{\partial t} \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial z} \\ h \end{bmatrix}.$$

Since, in general, the right-hand side is not zero, to uniquely solve this system for $\partial u/\partial x$, $\partial u/\partial z$ and $\partial u/\partial t$, we require that the determinant of the coefficient matrix be nonzero; we require that

$$(7.4.12) \quad \left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial z} \right)^2 + 1 \neq 0.$$

This condition is always satisfied. Hence, we can always uniquely solve for the first partial derivatives of u . Let us denote these solutions as $\partial u/\partial x := U_x(x, t)$, $\partial u/\partial z := U_z(x, t)$ and $\partial u/\partial t := U_t(x, t)$.

To find the second partial derivatives, we take partial derivatives of these three solutions with respect to x and with respect to z . We obtain

$$\frac{\partial U_x}{\partial x} = \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial x}(x, z, \psi(x, z)) \right] = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x \partial t} \frac{\partial \psi}{\partial x}$$

and

$$\frac{\partial U_x}{\partial z} = \frac{\partial}{\partial z} \left[\frac{\partial u}{\partial x}(x, z, \psi(x, z)) \right] = \frac{\partial^2 u}{\partial z \partial x} + \frac{\partial^2 u}{\partial x \partial t} \frac{\partial \psi}{\partial z}.$$

In view of the equality of mixed partial derivatives, $\partial U_x/\partial z = \partial U_z/\partial x$, so we proceed directly to

$$\frac{\partial U_z}{\partial z} = \frac{\partial}{\partial z} \left[\frac{\partial u}{\partial z}(x, z, \psi(x, z)) \right] = \frac{\partial^2 u}{\partial z^2} + \frac{\partial^2 u}{\partial z \partial t} \frac{\partial \psi}{\partial z}.$$

Recalling that the third argument, ψ , is tantamount to t , we get

$$\frac{\partial U_t}{\partial x} = \frac{\partial}{\partial x} \left[\frac{\partial u}{\partial t} (x, z, \psi(x, z)) \right] = \frac{\partial^2 u}{\partial x \partial t} + \frac{\partial^2 u}{\partial t^2} \frac{\partial \psi}{\partial x}$$

and

$$\frac{\partial U_t}{\partial z} = \frac{\partial^2 u}{\partial z \partial t} + \frac{\partial^2 u}{\partial t^2} \frac{\partial \psi}{\partial z}.$$

As the last equation for the second partial derivatives, we take equation (7.4.9). These six equations can be written as a system of linear algebraic equations given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \frac{\partial \psi}{\partial x} \\ 0 & 0 & 0 & 1 & 0 & \frac{\partial \psi}{\partial z} \\ 0 & 1 & 0 & 0 & \frac{\partial \psi}{\partial z} & 0 \\ 0 & 0 & \frac{\partial \psi}{\partial x} & 0 & 0 & 1 \\ 0 & 0 & \frac{\partial \psi}{\partial z} & 0 & 1 & 0 \\ 1 & 1 & -\frac{1}{v^2} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 u}{\partial x^2} \\ \frac{\partial^2 u}{\partial z^2} \\ \frac{\partial^2 u}{\partial t^2} \\ \frac{\partial^2 u}{\partial x \partial z} \\ \frac{\partial^2 u}{\partial z \partial t} \\ \frac{\partial^2 u}{\partial t \partial x} \end{bmatrix} = \begin{bmatrix} \frac{\partial U_x}{\partial x} \\ \frac{\partial U_x}{\partial z} \\ \frac{\partial U_z}{\partial z} \\ \frac{\partial U_t}{\partial x} \\ \frac{\partial U_t}{\partial z} \\ 0 \end{bmatrix},$$

where we used the equality of mixed partial derivatives.

For this system not to have a unique solution, we require that the determinant of the coefficient matrix be zero. Thus, we require that

$$(7.4.13) \quad \left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial z} \right)^2 = \frac{1}{v^2(x, z)},$$

which is eikonal equation (6.10.22) in two dimensions.

This derivation shows that eikonal equation (7.4.13) is the characteristic equation of equation (7.4.9). Since equation (7.4.9) is a second-order partial differential equation, we require that the initial condition uniquely specify the second partial derivatives. $\psi(x, z)$ is a surface on which we cannot specify the initial conditions so as to uniquely find the second partial derivatives; it is called the characteristic surface. Along this surface, the behaviour of u is predetermined by the original differential equation, which itself is rooted in the physical laws of equations of motion.

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 Aleksandrovich Kravtsov and Yuri Ilyich 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. This solution leads to a system of first-order linear ordinary differential equations that describe the curves that form the solution surface in the $\mathbf{x}\mathbf{p}$ -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, and which are the key equations of ray theory. Subsequently, we relate the orientations and magnitudes of vectors \mathbf{p} and $\dot{\mathbf{x}}$, which result in expressions relating phase and ray angles as well as phase and ray velocities. We conclude the chapter with two examples that illustrate the Hamiltonian approach.

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

¹In classical mechanics, the $\mathbf{x}\mathbf{p}$ -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.

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 (7.3.8), namely,

$$(8.1.1) \quad p^2 = \frac{1}{v^2(\mathbf{x}, \mathbf{p})},$$

where, $p^2 = \mathbf{p} \cdot \mathbf{p}$, and, in view of definition (7.3.2),

$$(8.1.2) \quad p_i := \frac{\partial \psi}{\partial x_i}, \quad i \in \{1, 2, 3\}.$$

We wish to solve this equation for $\mathbf{p}(\mathbf{x})$. In other words, at every point \mathbf{x} we are looking for vector $[\partial \psi / \partial x_1, \partial \psi / \partial x_2, \partial \psi / \partial x_3]$. Once, we get this vector, we can integrate it to obtain ψ — the solution of our partial differential equation.

The method that we are about to describe is specifically designed for the first-order partial differential equations that depend on the independent variables, \mathbf{x} , and on the first derivatives of function ψ , namely, $\nabla \psi(\mathbf{x})$. The eikonal equation is such a partial differential equation. This method does not apply to differential equations that also depend on function ψ itself.²

The solution of the eikonal equation is a surface in the $\mathbf{x}\mathbf{p}$ -space. We have a choice of several implicit descriptions of this surface as level sets of a function that we denote by $F(\mathbf{x}, \mathbf{p})$. The two obvious choices are

$$(8.1.3) \quad F(\mathbf{x}, \mathbf{p}) = p^2 - \frac{1}{v^2(\mathbf{x}, \mathbf{p})},$$

and

$$(8.1.4) \quad F(\mathbf{x}, \mathbf{p}) = p^2 v^2(\mathbf{x}, \mathbf{p}).$$

This way, in view of eikonal equation (8.1.1), the surfaces are the level sets of functions (8.1.3) or (8.1.4), given by

$$(8.1.5) \quad F(\mathbf{x}, \mathbf{p}) = 0,$$

and

$$(8.1.6) \quad F(\mathbf{x}, \mathbf{p}) = 1,$$

respectively. Since each formulation has different advantages, both are used in various sections of this book.

²Readers interested in the method of characteristics for solving general 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.

8.1.2. Characteristic equations. We seek to construct the solution given by $\mathbf{p} = \mathbf{p}(\mathbf{x})$, such that equation (8.1.5) or equation (8.1.6) is satisfied. In both cases, since $F(\mathbf{x}, \mathbf{p}(\mathbf{x}))$ is constant, it follows that $dF = 0$, where F is treated as a function of \mathbf{x} only.

Treating F as a function of \mathbf{x} only, which we denote by $F(\mathbf{x}, \mathbf{p}(\mathbf{x}))$, means that we constrain our consideration of F to $\mathbf{p}(\mathbf{x})$ — the solution we seek — as opposed to studying $F(\mathbf{x}, \mathbf{p})$, which refers to function F in the entire $\mathbf{x}\mathbf{p}$ -space.

Using the chain rule, we can explicitly state the differential of F as

$$dF[\mathbf{x}, \mathbf{p}(\mathbf{x})] = \sum_{i=1}^3 \frac{\partial F}{\partial x_i} dx_i + \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial F}{\partial p_j} \frac{\partial p_j}{\partial x_i} dx_i = 0.$$

Using definition (8.1.2), we can express it in terms of the eikonal function, ψ , as

$$dF[\mathbf{x}, \mathbf{p}(\mathbf{x})] = \sum_{i=1}^3 \frac{\partial F}{\partial x_i} dx_i + \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial F}{\partial p_j} \frac{\partial}{\partial x_i} \frac{\partial \psi}{\partial x_j} dx_i = 0.$$

Since $dx_i \neq 0$, using the equality of mixed partial derivatives and considering a given i , we can factor out dx_i and write

$$\frac{\partial F}{\partial x_i} + \sum_{j=1}^3 \frac{\partial F}{\partial p_j} \frac{\partial}{\partial x_j} \frac{\partial \psi}{\partial x_i} = 0, \quad i \in \{1, 2, 3\},$$

which are second-order partial differential equations. Again, using definition (8.1.2), we can rewrite these equations as

$$(8.1.7) \quad \frac{\partial F}{\partial x_i} + \sum_{j=1}^3 \frac{\partial F}{\partial p_j} \frac{\partial p_i}{\partial x_j} = 0, \quad i \in \{1, 2, 3\}.$$

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 will construct the solution surface as a union of these curves, which are commonly referred to as characteristics. To do so, we desire to obtain vectors tangent to the solution surface. To obtain these vectors, we will use geometrical properties of vectors in the context of the solution surface.

Let us consider a given $i \in \{1, 2, 3\}$. The corresponding equation among three equations (8.1.7) can be written as a scalar product of two vectors given by

$$(8.1.8) \quad \left[\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \frac{\partial F}{\partial p_3}, -\frac{\partial F}{\partial x_i} \right] \cdot \left[\frac{\partial p_i}{\partial x_1}, \frac{\partial p_i}{\partial x_2}, \frac{\partial p_i}{\partial x_3}, -1 \right] = 0.$$

Following the properties of the scalar product, we conclude that these two vectors are orthogonal to one another in the four-dimensional $x_1 x_2 x_3 p_i$ -space.

For a given $i \in \{1, 2, 3\}$, we can write the solution surface, $p_i = p_i(\mathbf{x})$, as a level set of the function given by

$$(8.1.9) \quad g_i(\mathbf{x}, p_i) = p_i(\mathbf{x}) - p_i,$$

where p_i is the value of function $p_i(\mathbf{x})$ at point \mathbf{x} . Herein, since the right-hand side of equation (8.1.9) vanishes identically, $g_i(\mathbf{x}, p_i)$ is a zero set. Using properties of the gradient operator, we can obtain vector normal to $g_i(\mathbf{x}, p_i)$, namely,

$$(8.1.10) \quad \mathbf{n}_i = \nabla g_i = \left[\frac{\partial g_i}{\partial x_1}, \frac{\partial g_i}{\partial x_2}, \frac{\partial g_i}{\partial x_3}, \frac{\partial g_i}{\partial p_i} \right].$$

In view of equation (8.1.9), we see that a vector normal to $g_i(\mathbf{x}, p_i)$ is also normal to the solution surface, $p_i = p_i(\mathbf{x})$. Thus, inserting $g_i(\mathbf{x}, p_i)$, given in expression (8.1.9), into expression (8.1.10), 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 vector

$$(8.1.11) \quad \mathbf{n}_i = \left[\frac{\partial p_i}{\partial x_1}, \frac{\partial p_i}{\partial x_2}, \frac{\partial p_i}{\partial x_3}, -1 \right],$$

which is normal to solution surface, $p_i = p_i(\mathbf{x})$ in the four-dimensional $x_1 x_2 x_3 p_i$ -space.

Examining equations (8.1.8) and (8.1.11), 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 . Hence, for a given $i \in \{1, 2, 3\}$, we have obtained vectors tangent to the solution surface, as desired.

Curves $[x_1(s), x_2(s), x_3(s), p_i(s)]$ that are in the solution surface and whose tangent vector is \mathbf{t}_i is the solution of a system of first-order ordinary differential equations, namely,

$$\left\{ \begin{array}{l} \frac{dx_1(s)}{ds} = \zeta \frac{\partial F}{\partial p_1} \\ \frac{dx_2(s)}{ds} = \zeta \frac{\partial F}{\partial p_2} \\ \frac{dx_3(s)}{ds} = \zeta \frac{\partial F}{\partial p_3} \\ \frac{dp_i(s)}{ds} = -\zeta \frac{\partial F}{\partial x_i} \end{array} \right.,$$

which we can concisely write as

$$(8.1.12) \quad \left\{ \begin{array}{l} \frac{dx_j(s)}{ds} = \zeta \frac{\partial F}{\partial p_j} \\ \frac{dp_i(s)}{ds} = -\zeta \frac{\partial F}{\partial x_i} \end{array} \right., \quad j \in \{1, 2, 3\},$$

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. System (8.1.12) describes curves that are associated

with solution surface $p_i = p_i(\mathbf{x})$ and — for a given $i \in \{1, 2, 3\}$ — exist in a four-dimensional space, $x_1 x_2 x_3 p_i$.

We note that the solutions of system (8.1.12) depend on the initial conditions, which we can write as $x_j(0) = x_j^0$ and $p_i(0) = p_i^0$, where $i, j \in \{1, 2, 3\}$. However, these initial conditions are not arbitrary; they must satisfy the differential equation given in expression (8.1.5) or (8.1.6).

Since the above derivation, which was shown for a given i , must hold for each $i \in \{1, 2, 3\}$, we can write equations (8.1.12) as

$$\begin{cases} \frac{dx_j}{ds} = \zeta \frac{\partial F}{\partial p_j} \\ \frac{dp_i}{ds} = -\zeta \frac{\partial F}{\partial x_i} \end{cases}, \quad i, j \in \{1, 2, 3\},$$

which, in view of i and j being the summation indices, we can restate as

$$(8.1.13) \quad \begin{cases} \frac{dx_i}{ds} = \zeta \frac{\partial F}{\partial p_i} \\ \frac{dp_i}{ds} = -\zeta \frac{\partial F}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

The solution of system (8.1.13) are curves that compose solution surface $\mathbf{p} = \mathbf{p}(\mathbf{x})$ in the six-dimensional $\mathbf{x}\mathbf{p}$ -space. Such curves are the characteristics of eikonal equation (8.1.1).

Hence, three second-order partial differential equations (8.1.7) become six first-order ordinary differential equations (8.1.13). These are the characteristic equations of eikonal equation (8.1.1). The solution of characteristic equations (8.1.13) consists of curves that compose the solution surface of eikonal equation (8.1.1).

8.1.3. Consistency of formulation. As stated in Section 8.1.1, there are two obvious forms of function F . Functions (8.1.3) and (8.1.4) 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 8.1.1. *Both formulations of the function given by expressions (8.1.3) and (8.1.4) result in the same characteristic curves.*

PROOF. Consider characteristic equations (8.1.13), namely,

$$\begin{cases} \frac{dx_i}{ds} = \zeta \frac{\partial F}{\partial p_i} \\ \frac{dp_i}{ds} = -\zeta \frac{\partial F}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Letting $F = p^2 v^2 (\mathbf{x}, \mathbf{p})$ and setting $\zeta = 1$, we note that equations (8.1.13) become

$$\begin{cases} \frac{dx_i}{ds} = 2 \left(p_i v^2 + p^2 v \frac{\partial v}{\partial p_i} \right) \\ \frac{dp_i}{ds} = -2 p^2 v \frac{\partial v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

We let $F = p^2 - 1/v^2 (\mathbf{x}, \mathbf{p})$ and equations (8.1.13) become

$$\begin{cases} \frac{dx_i}{ds} = 2\zeta \left(p_i + \frac{1}{v^3} \frac{\partial v}{\partial p_i} \right) \\ \frac{dp_i}{ds} = -2\zeta \frac{1}{v^3} \frac{\partial v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Equating the second equations of each set, we can write

$$p^2 v \frac{\partial v}{\partial x_i} = \zeta \frac{1}{v^3} \frac{\partial v}{\partial x_i}, \quad i \in \{1, 2, 3\}.$$

Solving for ζ , we obtain

$$\zeta = \frac{p^2 v \frac{\partial v}{\partial x_i}}{\frac{1}{v^3} \frac{\partial v}{\partial x_i}} = p^2 v^4, \quad i \in \{1, 2, 3\}.$$

Substituting $\zeta = p^2 v^4$ into the first equation of the second set, we obtain

$$\frac{dx_i}{ds} = 2p^2 v^4 \left(p_i + \frac{1}{v^3} \frac{\partial v}{\partial p_i} \right) = 2 \left(p_i p^2 v^4 + p^2 v \frac{\partial v}{\partial p_i} \right), \quad i \in \{1, 2, 3\},$$

which is equivalent to the first equation of the first set along $p^2 v^2 = 1$. □

Thus, following equations (8.1.13), both equations (8.1.5) and (8.1.6), 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 (8.1.13) in terms of time. Recall equation (7.2.3), which we can rewrite as

$$\psi(\mathbf{x}) = t + t_i,$$

where t denotes time and t_i is a constant. Differentiating with respect to s , we obtain

$$\frac{d\psi(\mathbf{x})}{ds} = \frac{dt}{ds}.$$

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 (8.1.13). If the parameter s is to be

equivalent to time, t , we require that

$$\frac{d\psi(\mathbf{x})}{ds} = 1.$$

We can restate the above condition as

$$(8.2.1) \quad \frac{d\psi(\mathbf{x})}{ds} = \sum_{i=1}^3 \frac{\partial \psi}{\partial x_i} \frac{dx_i}{ds} = 1.$$

Using definition (8.1.2), we rewrite condition (8.2.1) as

$$(8.2.2) \quad \frac{d\psi(\mathbf{x})}{ds} = \sum_{i=1}^3 p_i \frac{dx_i}{ds} = 1,$$

which is a condition for the time parametrization of characteristic equations (8.1.13).

8.2.2. Equations with variable scaling factor. In order to obtain the time parametrization of system (8.1.13) in the context of function (8.1.3), we can write $dx_i/ds = \zeta \partial F / \partial p_i$, where $F = p^2 - 1/v^2$, as

$$\frac{dx_i}{ds} = 2\zeta \left(p_i + \frac{1}{v^3} \frac{\partial v}{\partial p_i} \right), \quad i \in \{1, 2, 3\}.$$

In view of condition (8.2.2), we require that

$$\sum_{i=1}^3 p_i \frac{dx_i}{ds} = 2\zeta \sum_{i=1}^3 p_i \left(p_i + \frac{1}{v^3} \frac{\partial v}{\partial p_i} \right) = 2\zeta \left(p^2 + \frac{1}{v^3} \sum_{i=1}^3 p_i \frac{\partial v}{\partial p_i} \right) = 1.$$

Since v is homogeneous of degree 0 in the p_i , the summation on the right-hand side vanishes by Theorem A.2.1. Thus, we obtain

$$\sum_{i=1}^3 p_i \frac{dx_i}{ds} = 2\zeta p^2 = 1,$$

and solving for ζ , we immediately get $\zeta = 1/(2p^2)$.

Consequently, given function (8.1.3), the system of characteristic equations (8.1.13) that is parametrized in terms of time becomes

$$(8.2.3) \quad \begin{cases} \dot{x}_i = \frac{1}{2p^2} \frac{\partial F}{\partial p_i} \\ \dot{p}_i = -\frac{1}{2p^2} \frac{\partial F}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\},$$

where $\dot{x}_i := dx_i/dt$ and $\dot{p}_i := dp_i/dt$. Equations (8.2.3) are characteristic equations (8.1.13) whose scaling factor is a function of the p_i . In view of eikonal equation (8.1.1), we can also state this scaling factor as $v^2(\mathbf{x}, \mathbf{p})/2$.

An implication of this parametrization is shown in Exercise 8.8. An implication of another parametrization of characteristic equations (8.1.13) in the context of function (8.1.3) is shown in Exercise 8.7.

8.2.3. Equations with constant scaling factor. In order to obtain the time parametrization of system (8.1.13) in the context of function (8.1.4), we can write $dx_i/ds = \zeta \partial F / \partial p_i$, where $F = p^2 v^2$, as

$$\frac{dx_i}{ds} = 2\zeta \left(p_i v^2 + p^2 v \frac{\partial v}{\partial p_i} \right), \quad i \in \{1, 2, 3\}.$$

In view of condition (8.2.2), we require

$$\sum_{i=1}^3 p_i \frac{dx_i}{ds} = 2\zeta \sum_{i=1}^3 p_i \left(p_i v^2 + p^2 v \frac{\partial v}{\partial p_i} \right) = 2\zeta \left(p^2 v^2 + p^2 v \sum_{i=1}^3 p_i \frac{\partial v}{\partial p_i} \right) = 1.$$

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.2.1. Thus, we obtain

$$\sum_{i=1}^3 p_i \frac{dx_i}{ds} = 2\zeta = 1,$$

and solving for ζ , we immediately get $\zeta = 1/2$.

Consequently, given function (8.1.4), system (8.1.13) is parametrized in terms of time if

$$(8.2.4) \quad \begin{cases} \dot{x}_i = \frac{1}{2} \frac{\partial F}{\partial p_i} \\ \dot{p}_i = -\frac{1}{2} \frac{\partial F}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\},$$

where $\dot{x}_i := dx_i/dt$ and $\dot{p}_i := dp_i/dt$. Equations (8.2.4) are characteristic equations (8.1.13) whose scaling factor is the constant equal to $1/2$.

In view of functions (8.1.3) and (8.1.4), 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 8.1.1.

8.2.4. Formulation of Hamilton's ray equations. We now examine systems (8.2.3) and (8.2.4), 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 (8.2.4) as

$$(8.2.5) \quad \begin{cases} \dot{x}_i = \frac{\partial}{\partial p_i} \left(\frac{F}{2} \right) \\ \dot{p}_i = -\frac{\partial}{\partial x_i} \left(\frac{F}{2} \right) \end{cases}, \quad i \in \{1, 2, 3\}.$$

Let us denote

$$(8.2.6) \quad \mathcal{H} := \frac{F}{2},$$

where \mathcal{H} is referred to as the ray-theory Hamiltonian³. Now, we can write equations (8.2.5) as

$$(8.2.7) \quad \begin{cases} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Equations (8.2.7) 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. System (8.2.7) governs the signal trajectories in the $\mathbf{x}\mathbf{p}$ -space spanned by the position vectors, \mathbf{x} , and the phase-slowness vectors, \mathbf{p} .

The first set of equations of system (8.2.7) corresponds to the components of vectors tangent to curves $\mathbf{x}(t)$. These curves belong to the physical space. They are the trajectories along which signals propagate and, in the context of ray theory, they are rays.

The second set of equations of system (8.2.7) 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 (8.1.4), can be explicitly stated as

$$(8.2.8) \quad \mathcal{H} = \frac{1}{2} p^2 v^2(\mathbf{x}, \mathbf{p}).$$

It is a dimensionless quantity, unlike the classical-mechanics Hamiltonian, discussed in Chapter 13, which has units of energy. In view of eikonal equation (7.3.8), which states that $p^2 v^2 = 1$, and expression (8.2.8), we require that $\mathcal{H}(\mathbf{x}, \mathbf{p}) = 1/2$, along a ray.

8.3. Physical interpretation of Hamilton's ray equations and solutions

8.3.1. Equations.

Ray velocity. The first set of equations of system (8.2.7), namely, $\dot{x}_i = \partial \mathcal{H} / \partial p_i$, states the components of vector $\dot{\mathbf{x}}$, which is tangent to the ray, $\mathbf{x}(t)$. Since the right-hand sides of these equations are expressed in terms of the p_i , which are the components of the phase-slowness vector, this set of equations relates ray and phase velocities.

³In 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.

Ray orientation. The relations among the \dot{x}_i , which appear in the first set of equations of system (8.2.7), give us the orientation of the ray, $\mathbf{x}(t)$. Since the right-hand sides are expressed in terms of the p_i , this set of equations relates ray and wavefront orientations. This is illustrated in Section 8.5.5.

Wavefront orientation. The relations among the \dot{p}_i , which appear in the second set of equations of system (8.2.7), namely, $\dot{p}_i = -\partial\mathcal{H}/\partial x_i$, give us the orientation of the wavefront, $\psi(\mathbf{x})$. This is illustrated in Section 8.5.5.

8.3.2. Solutions.

Ray. The solution of the first set of equations of system (8.2.7) is $\mathbf{x}(t)$, which is the expression for the ray. This is illustrated in Section 8.5.6 on page 250.

Wavefront velocity. The solution of the second set of equations of system (8.2.7) is $\mathbf{p}(t)$, which is the expression for the vector normal to the wavefront, $\psi(\mathbf{x})$. The magnitude of this vector is the wavefront slowness. This is illustrated in Section 8.5.8.

Traveltime. Since both $\mathbf{x}(t)$ and $\mathbf{p}(t)$ are parametrized by time, t , we can obtain the expression for t by solving any $x_i(t)$ or $p_i(t)$ for t . This is illustrated in Section 8.5.7.

8.4. Relation between \mathbf{p} and $\dot{\mathbf{x}}$

8.4.1. General formulation. We wish to study the relation between the orientations and the magnitudes of vectors \mathbf{p} and $\dot{\mathbf{x}}$. Physically, \mathbf{p} is the vector normal to the wavefront and $\dot{\mathbf{x}}$ is the vector tangent to the ray. Mathematically, the components of these vectors are the variables of Legendre's transformation discussed in Appendix B and used in Chapter 11.

Consider a given point \mathbf{x} of the continuum and, therein, the directional dependence of \mathcal{H} . The first set of equations of system (8.2.7) is

$$(8.4.1) \quad \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad i \in \{1, 2, 3\}.$$

Inserting expression (8.2.8), namely,

$$\mathcal{H} = \frac{1}{2} p^2 v^2(\mathbf{x}, \mathbf{p}),$$

into equations (8.4.1) and using the equality resulting from the eikonal equation, namely, $p^2 v^2 = 1$, we obtain

$$(8.4.2) \quad \dot{x}_i = p_i v^2 + \frac{1}{v} \frac{\partial v}{\partial p_i}, \quad i \in \{1, 2, 3\},$$

where the phase-velocity function, v , is a function of the orientation of vector \mathbf{p} . Thus, we have obtained the relation between the components of vectors \mathbf{p} and $\dot{\mathbf{x}}$.

8.4.2. Phase and ray velocities. 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.⁴ We wish to find the magnitude of this vector, which can be written as

$$(8.4.3) \quad V := |\dot{\mathbf{x}}| = \sqrt{\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}} = \sqrt{\sum_{i=1}^3 \dot{x}_i^2}.$$

Using expression (8.4.2), we can write each term of the summation in radicand (8.4.3) as

$$\begin{aligned} (\dot{x}_i)^2 &= \left(p_i v^2 + \frac{1}{v} \frac{\partial v}{\partial p_i} \right)^2 \\ &= (p_i)^2 v^4 + 2p_i v \frac{\partial v}{\partial p_i} + \frac{1}{v^2} \left(\frac{\partial v}{\partial p_i} \right)^2, \quad i \in \{1, 2, 3\}. \end{aligned}$$

Performing the summation of the three terms, we obtain

$$\begin{aligned} \sum_{i=1}^3 \left[(p_i)^2 v^4 + 2p_i v \frac{\partial v}{\partial p_i} + \frac{1}{v^2} \left(\frac{\partial v}{\partial p_i} \right)^2 \right] &= v^4 \sum_{i=1}^3 (p_i)^2 + 2v \sum_{i=1}^3 p_i \frac{\partial v}{\partial p_i} \\ &\quad + \frac{1}{v^2} \sum_{i=1}^3 \left(\frac{\partial v}{\partial p_i} \right)^2 \\ &= v^4 \sum_{i=1}^3 (p_i)^2 + \frac{1}{v^2} \sum_{i=1}^3 \left(\frac{\partial v}{\partial p_i} \right)^2, \end{aligned}$$

where, since v is homogeneous of degree 0 in the p_i , the summation of $p_i (\partial v / \partial p_i)$ vanished due to Theorem A.2.1.

Thus, in view of equality $p^2 v^2 = 1$, we can write expression (8.4.3) as

$$V = \sqrt{v^2 + \frac{1}{v^2} (\nabla_{\mathbf{p}} v)^2},$$

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 rule and following the properties of logarithms, we obtain

$$(8.4.4) \quad V = \sqrt{v^2 + [\nabla_{\mathbf{p}} (\ln v)]^2}.$$

Expression (8.4.4) gives the magnitude of the signal velocity along the ray $\mathbf{x}(t)$. In expression (8.4.4), 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} .

⁴In 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 (7th edition)*: Cambridge University Press, pp. 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.

Two-dimensional case. To illustrate expression (8.4.4), 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 expression (8.5.3), namely,

$$(8.4.5) \quad \vartheta = \arctan \frac{p_1}{p_3}.$$

Hence, using expression (8.4.4), the magnitude of the ray velocity can be expressed in terms of the phase velocity and the phase angle.

Herein, using expression (8.4.4), we can write

$$(8.4.6) \quad V = \sqrt{v^2 + \left(\frac{\partial \ln v}{\partial p_1} \right)^2 + \left(\frac{\partial \ln v}{\partial p_3} \right)^2}.$$

We wish to express differential operators $\partial/\partial p_i$ in terms of the phase angle. Using the chain rule, we can write

$$(8.4.7) \quad \frac{\partial}{\partial p_1} = \frac{\partial \vartheta}{\partial p_1} \frac{\partial}{\partial \vartheta} = \frac{\partial \arctan \frac{p_1}{p_3}}{\partial p_1} \frac{\partial}{\partial \vartheta} = \frac{p_3}{p_1^2 + p_3^2} \frac{\partial}{\partial \vartheta} = p_3 v^2 \frac{\partial}{\partial \vartheta},$$

where $p_1^2 + p_3^2 = p^2 = 1/v^2$. Similarly, we obtain

$$(8.4.8) \quad \frac{\partial}{\partial p_3} = -p_1 v^2 \frac{\partial}{\partial \vartheta}.$$

Thus, expression (8.4.6) can be written as

$$\begin{aligned} V &= \sqrt{v^2 + \left(p_3 v^2 \frac{\partial \ln v}{\partial \vartheta} \right)^2 + \left(-p_1 v^2 \frac{\partial \ln v}{\partial \vartheta} \right)^2} \\ &= \sqrt{v^2 + (p_3^2 + p_1^2) v^4 \left(\frac{\partial \ln v}{\partial \vartheta} \right)^2} = \sqrt{v^2 + p^2 v^4 \left(\frac{\partial \ln v}{\partial \vartheta} \right)^2} \\ &= \sqrt{[v(\vartheta)]^2 + [v(\vartheta)]^2 \left(\frac{\partial \ln v(\vartheta)}{\partial \vartheta} \right)^2}. \end{aligned}$$

Following the chain rule, we obtain

$$(8.4.9) \quad V(\vartheta) = \sqrt{[v(\vartheta)]^2 + \left[\frac{\partial v(\vartheta)}{\partial \vartheta} \right]^2},$$

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 (8.4.9) gives the magnitude of the ray velocity in terms of these properties and as a function of the phase angle.

8.4.3. Phase and ray angles. To discuss 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 (8.4.5). 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

$$(8.4.10) \quad \theta = \arctan \frac{\dot{x}_1}{\dot{x}_3}.$$

In this two-dimensional case, expression (8.4.2) is

$$(8.4.11) \quad \dot{x}_i = p_i v^2 + \frac{1}{v} \frac{\partial v}{\partial p_i}, \quad i \in \{1, 3\}.$$

Hence, expression (8.4.10) becomes

$$\tan \theta = \frac{p_1 v^2 + \frac{1}{v} \frac{\partial v}{\partial p_1}}{p_3 v^2 + \frac{1}{v} \frac{\partial v}{\partial p_3}}.$$

We wish to express the differential operators $\partial/\partial p_i$ in terms of the phase angle. Recalling expression (8.4.7) and (8.4.8), we obtain

$$\tan \theta = \frac{p_1 v^2 + p_3 v \frac{\partial v}{\partial \vartheta}}{p_3 v^2 - p_1 v \frac{\partial v}{\partial \vartheta}} = \frac{p_1 + p_3 \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{p_3 - p_1 \frac{1}{v} \frac{\partial v}{\partial \vartheta}}.$$

Recalling expression (8.4.5), we divide both the numerator and the denominator by p_3 , to obtain

$$(8.4.12) \quad \tan \theta = \frac{\frac{p_1}{p_3} + \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{1 - \frac{p_1}{p_3} \frac{1}{v} \frac{\partial v}{\partial \vartheta}} = \frac{\tan \vartheta + \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{1 - \frac{\tan \vartheta}{v} \frac{\partial v}{\partial \vartheta}}.$$

Expression (8.4.12) relates the phase and the ray angles.

Note that, in view of standard formulations in polar coordinates, expression (8.4.12) gives the angle θ that corresponds to the vector normal to the curve $1/v(\vartheta)$.⁵ We refer to this curve as the phase-slowness curve. Furthermore, as shown in Exercise 11.3, ϑ corresponds to the vector normal to the curve given by $V(\theta)$, which is the ray-velocity curve. This angular relation between the phase-slowness and ray-velocity curves, together with expression (8.4.9), which relates the magnitudes of the phase and ray velocities, results in the property that we call polar reciprocity. In other words, the phase-slowness curve is the polar reciprocal of the ray-velocity curve, and vice-versa. In general, the phase-slowness and ray-velocity surfaces are the polar reciprocals of one another.⁶

The possibility of solving expression (8.4.12) explicitly for ϑ depends on function v . To understand this statement, consider the following description. The explicit solution of expression (8.4.12) for ϑ requires

⁵Interested readers might refer to Anton, H., (1984) Calculus with analytic geometry: John Wiley & Sons, pp. 730 – 731.

⁶Interested 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.

that we can solve expression (8.4.11) for the p_i in terms of the \dot{x}_i . Since expression (8.4.11) is derived from expression (8.4.1), we require the solvability of the latter expression for the p_i in terms of the \dot{x}_i .

Note that equation (8.4.1) is equivalent to equation (B.3.7) in Appendix B, which relates the variables used in Legendre's transformation. Thus, the possibility of expressing ϑ in terms of θ belongs to the study of this transformation.

For the velocity functions formulated in the context of the elasticity theory, we conjecture that expression (8.4.12) can be explicitly solved for ϑ if and only if v^2 is quadratic in the components of a vector that specifies the orientation of the wavefront. As we will see in Section 9.2.3, for SH waves in transversely isotropic continua, we get

$$v_{SH}^2 = \frac{1}{\rho} [C_{66} + (C_{44} - C_{66}) n_3^2],$$

which is a quadratic function in n_3 , with n_3 being a component of the unit vector normal to the wavefront; in other words, $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$. In the seismological context, this quadratic dependence is tantamount to an elliptical velocity dependence. Consequently, an explicit ray-velocity expression, $V = V(\theta)$, where θ is the ray angle, appears to be possible only for elliptical velocity dependence. This expression is illustrated in Exercise 8.5.

8.4.4. 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 8.4, the relationship between the magnitudes of the ray velocity, $V = |\dot{\mathbf{x}}|$, and phase velocity, $v = 1/|\mathbf{p}|$, is given by

$$(8.4.13) \quad V = \frac{v}{\mathbf{n} \cdot \mathbf{t}},$$

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 (8.4.13) shows that the phase-velocity vector is the projection of the ray-velocity vector onto the wavefront normal.⁷ 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 (8.4.13) as

$$(8.4.14) \quad V = \frac{v}{\cos(\theta - \vartheta)}.$$

⁷Readers 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.

Expression (8.4.14) conveniently involves all four entities discussed in this chapter, namely, ray velocity, V , phase velocity, v , ray angle, θ , and phase angle, ϑ .

8.5. Example: Elliptical anisotropy and linear inhomogeneity

8.5.1. Introductory comments. ⁸In this section, we study Hamilton's ray equations for a particular case of a wave that exhibits an elliptical velocity dependence with direction and a linear velocity dependence along one axis. This assumption, in the context of a two-dimensional continuum, allows us to conveniently illustrate the meaning of Hamilton's ray equations by considering analytic expressions for rays and traveltimes. Also, the same case will be treated in Section 14.3 in the context of Lagrange's ray equations. Thus, our examination of Sections 8.5 and 14.3 will allow us to investigate the same physical problem using the two different approaches that are available to study seismic ray theory.

We wish to emphasize that, as discussed in Section 6.10.3, the elliptical velocity dependence refers to the fact that infinitesimal wavefronts generated by a point source are elliptical. This behaviour of wavefronts is related to the properties of the continuum in which the given wave propagates. As shown in Section 7.3, three types of waves can propagate in anisotropic continua. In general, in a given continuum, each of the three waves exhibits a different infinitesimal wavefront. Thus, although in a particular continuum one of the three waves might exhibit an elliptical wavefront, the other two waves, in general, do not exhibit elliptical wavefronts. Often, for brevity, we refer to the elliptical velocity dependence as the elliptical anisotropy. However, it should be clear that elliptical anisotropy refers to the response of a given wave to the properties of the continuum, not to the material symmetry of the continuum itself.

8.5.2. Eikonal equation. As shown in expression (6.10.11), considering the xz -plane, we can write the phase velocity of a wave subjected to the elliptical anisotropy as

$$v(\vartheta) = \sqrt{v_x^2 \sin^2 \vartheta + v_z^2 \cos^2 \vartheta},$$

where v_x and v_z are the magnitudes of phase velocity along the x -axis and z -axis, respectively, and ϑ is the phase angle measured from the z -axis. For convenience of notation, we define parameter

$$(8.5.1) \quad \chi := \frac{v_x^2 - v_z^2}{2v_z^2},$$

which is a dimensionless quantity that vanishes in the isotropic case. Using this definition, we can solve for

$$v_x^2 = v_z^2 (1 + 2\chi),$$

and rewrite the expression for the phase velocity as

$$v(\vartheta) = v_z \sqrt{(1 + 2\chi) \sin^2 \vartheta + \cos^2 \vartheta}.$$

⁸This section is based on the work that was published by Rogister, Y., and Slawinski, M.A., (2005) Analytic solution of ray-tracing equations for a linearly inhomogeneous and elliptically anisotropic velocity model: *Geophysics*, **70**, D37 – D41.

If the wave is also subjected to the linear increase of velocity along the z -axis, we can write

$$(8.5.2) \quad v(\vartheta, z) = (a + bz) \sqrt{(1 + 2\chi) \sin^2 \vartheta + \cos^2 \vartheta},$$

where a and b are constants whose units are the units of velocity and the reciprocal of time, respectively. We refer to the velocity model described by this expression as the $ab\chi$ model.

Note that the meaning of the term “model” used herein, although consistent with the common use in seismology, is not the same as the meaning of models referred to in the footnote on page 7. Therein, the models constitute a complete physical picture in the context of a mathematical theory.

We wish to formulate the eikonal equation corresponding to equation (8.5.2). Following equation (8.1.1), we can write it as

$$p^2 := p_x^2 + p_z^2 = \frac{1}{(a + bz)^2 [(1 + 2\chi) \sin^2 \vartheta + \cos^2 \vartheta]}.$$

To express ϑ in terms of vector $\mathbf{p} = [p_x, p_z]$, we can write

$$(8.5.3) \quad \vartheta = \arctan \frac{p_x}{p_z},$$

which is equivalent to expression (6.10.9). Inserting the expression for ϑ into the above equation and using trigonometric identities, we get

$$p_x^2 + p_z^2 = \frac{p_x^2 + p_z^2}{(a + bz)^2 [(1 + 2\chi) p_x^2 + p_z^2]}.$$

Simplifying, we obtain

$$(8.5.4) \quad (a + bz)^2 [(1 + 2\chi) p_x^2 + p_z^2] = 1.$$

This is the eikonal equation that corresponds to elliptical anisotropy and linear inhomogeneity. Since the right-hand side is equal to unity, this expression is also the level set of function (8.1.4).

To avoid any confusion about the meaning of p_x , p_z , v_x and v_z , we refer the reader to Notation 6.10.1 on page 181.

To see that we are dealing with a differential equation, let us take a look at eikonal equation (8.5.4) and, in view of definition (7.3.2), rewrite it as

$$(8.5.5) \quad (a + bz)^2 \left[(1 + 2\chi) \left[\frac{\partial \psi(x, z)}{\partial x} \right]^2 + \left[\frac{\partial \psi(x, z)}{\partial z} \right]^2 \right] = 1.$$

We are looking for function $\psi(x, z)$. Rather than attempting to solve this nonlinear partial differential equation, we will study the system of ordinary differential equations that are the characteristic equations of equation (8.5.4). The solution of this system will provide us with information about the physical phenomenon that is governed by eikonal equation (8.5.4).

8.5.3. Hamilton's ray equations. Using expression (8.5.4), we write our ray-theory Hamiltonian as

$$\mathcal{H} = \frac{(a + bz)^2 [(1 + 2\chi) p_x^2 + p_z^2]}{2},$$

which, as we can see, is a dimensionless quantity. Following equations (8.2.7), we write our Hamilton's ray equations as

$$(8.5.6) \quad \begin{cases} \dot{x} = \frac{\partial \mathcal{H}(z, p_x, p_z)}{\partial p_x} = (a + bz)^2 (1 + 2\chi) p_x \\ \dot{z} = \frac{\partial \mathcal{H}(z, p_x, p_z)}{\partial p_z} = (a + bz)^2 p_z \\ \dot{p}_x = -\frac{\partial \mathcal{H}(z, p_x, p_z)}{\partial x} = 0 \\ \dot{p}_z = -\frac{\partial \mathcal{H}(z, p_x, p_z)}{\partial z} = -b(a + bz) [(1 + 2\chi) p_x^2 + p_z^2] \end{cases}.$$

These are the characteristic equations of equation (8.5.4).

Since we can write equation (8.5.4) as

$$(1 + 2\chi) p_x^2 + p_z^2 = \frac{1}{(a + bz)^2},$$

we can rewrite the last Hamilton's ray equation as

$$\dot{p}_z = -\frac{b}{a + bz}.$$

Since $\dot{p}_x = 0$, it immediately follows that $p_x(t) = \mathfrak{p}$, where \mathfrak{p} denotes a constant. Now, we can write the remaining three Hamilton's ray equations as a system of ordinary differential equations to be solved for x , z and p_z . These equations are

$$(8.5.7) \quad \frac{dx(t)}{dt} = [a + bz(t)]^2 (1 + 2\chi) \mathfrak{p},$$

$$(8.5.8) \quad \frac{dz(t)}{dt} = [a + bz(t)]^2 p_z(t)$$

and

$$(8.5.9) \quad \frac{dp_z(t)}{dt} = -\frac{b}{a + bz(t)}.$$

8.5.4. Initial conditions. To complete this system of differential equations, we need additional constraints. We choose to use the initial conditions, which correspond to the values of unknowns at the initial time. In other words, we need $x(t)$, $z(t)$, $p_x(t)$ and $p_z(t)$ at $t = 0$. While we already have $p_x(0) = \mathfrak{p}$, let us set $x(0) = 0$ and $z(0) = 0$ and $p_z(0) = p_{z0}$. The initial condition for $p_z(t)$ is not independent from the initial condition for $p_x(t)$. They are related by eikonal equation (8.5.4). Solving this equation for p_z

at $t = 0$, which corresponds to $z = 0$, we get

$$p_z(0) = \sqrt{\frac{1}{a^2} - (1 + 2\chi) [p_x(0)]^2}.$$

Since $p_x(0) = \mathfrak{p}$, the initial condition for p_z that obeys the eikonal equation is

$$(8.5.10) \quad p_z(0) = \sqrt{\frac{1}{a^2} - (1 + 2\chi) \mathfrak{p}^2}.$$

System (8.5.6) accompanied by the initial conditions has a clear meaning in the context of ray theory. We will discuss it in the next section.

8.5.5. Physical interpretation of equations and conditions.

Ray velocity. The first two equations of system (8.5.6) define the vector field in the xz -plane. Herein, solution $[x(t), z(t)]$ describes the path of a signal under elliptical velocity dependence with direction and a linear velocity dependence along the z -axis; this path is the ray. Below, we will discuss the physical information contained in the equations themselves.

Considering the first two equations of system (8.5.6), we can express the magnitude of the velocity of the signal along the ray as

$$V = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2},$$

where V is referred to as the ray velocity. We can explicitly write

$$V = \sqrt{\left([a + bz(t)]^2 (1 + 2\chi) \mathfrak{p}\right)^2 + \left([a + bz(t)]^2 p_z(t)\right)^2},$$

which we rewrite as

$$V = [a + bz(t)]^2 \sqrt{(1 + 2\chi)^2 \mathfrak{p}^2 + p_z(t)^2}.$$

This expression relates the ray velocity to the wavefront slowness.

Ray orientation. Considering the first two equations of system (8.5.6), we can express the direction of a ray as

$$\theta = \arctan \frac{dx}{dz},$$

where θ is measured from the z -axis and is referred to as the ray angle. Herein, by referring to system (8.5.6), we write

$$\theta = \arctan \frac{dx}{dz} = \arctan \frac{\dot{x}}{\dot{z}} = \arctan \left[(1 + 2\chi) \frac{p_x}{p_z} \right],$$

which we can rewrite as

$$\tan \theta = (1 + 2\chi) \frac{p_x}{p_z}.$$

Invoking expression (6.10.9), we can write

$$(8.5.11) \quad \tan \theta = (1 + 2\chi) \tan \vartheta,$$

which explicitly relates ray and phase angles.

Wavefront orientation. To gain more understanding of expression (6.10.9), which was already discussed between pages 181 and 183, we examine the last two equations of system (8.5.6).

We recall that contours of $\psi(x, z)$ correspond to wavefronts at given instants of time. Hence, in view of definition (7.3.2), $p_x := \partial\psi/\partial x$ and $p_z = \partial\psi/\partial z$ at a given point, $(x(t), z(t))$, are components of slowness with which a wavefront propagates at this point. Also, vector $\mathbf{p} = [p_x, p_z]$ at $(x(t), z(t))$ is normal to the wavefront at this point. In view of expression (8.5.3), we write

$$\vartheta = \arctan \frac{p_x}{p_z},$$

where ϑ is the phase angle, which gives the direction of wavefront propagation.

Initial conditions. Now, let us examine the physical meaning of the initial conditions. Setting

$$[x(0), z(0)] = [0, 0],$$

we fix the origin of the ray at the initial time. In other words, we locate the point source at the origin. In view of continuity of wavefronts and considering the inhomogeneity along the z -axis only, we know that

$$\mathbf{p} = \frac{\sin \vartheta}{v(\vartheta, z)}.$$

Considering $z = 0$ and using expression (8.5.2), we can write

$$(8.5.12) \quad p_x(0) = \mathbf{p} = \frac{\sin \vartheta_0}{a \sqrt{(1 + 2\chi) \sin^2 \vartheta_0 + \cos^2 \vartheta_0}},$$

where ϑ_0 denotes the take-off phase angle; in other words, the direction of the wavefront at the source.

8.5.6. Solution of Hamilton's ray equations.

General solution. We wish to solve system (8.5.6). Since we already know that $p_x = \mathbf{p}$, we must solve equations (8.5.7), (8.5.8) and (8.5.9) for $x(t)$, $z(t)$ and $p_z(t)$.

To do so, let us write the second equation of system (8.5.6) as

$$p_z = \frac{\frac{dz}{dt}}{[a + bz]^2}.$$

Differentiating with respect to t , we get

$$\frac{dp_z}{dt} = \frac{\frac{d^2z}{dt^2} (a + bz) - 2b \left(\frac{dz}{dt}\right)^2}{(a + bz)^3}.$$

Equating this result with the fourth equation of system (8.5.6) and rearranging, we get

$$(a + bz) \frac{d^2z}{dt^2} - 2b \left(\frac{dz}{dt}\right)^2 + b(a + bz)^2 = 0.$$

Letting $(a + bz)^2 = y$, we get

$$\frac{1}{2b} \frac{d^2 y}{dt^2} - \frac{3}{4yb} \left(\frac{dy}{dt} \right)^2 + by = 0.$$

Letting $y = a^2 \exp u$, we get

$$\frac{d^2 u}{dt^2} - \frac{1}{2} \left(\frac{du}{dt} \right)^2 + 2b^2 = 0.$$

Letting $du/dt = q$, we get

$$\frac{dq}{dt} - \frac{1}{2} q^2 + 2b^2 = 0.$$

We can rewrite this equation as

$$dt = \frac{dq}{\frac{1}{2} q^2 - 2b^2}.$$

Integrating both sides of this equation, we get

$$t + A_1 = -\frac{1}{b} \tanh^{-1} \frac{q}{2b},$$

where A_1 is an integration constant. Solving for q , we get

$$q = 2b \tanh(-b(t + A_1)).$$

To obtain u , we integrate and get

$$u = -2 \ln(\cosh(-b(t + A_1))) + A_2,$$

where A_2 is an integration constant. Hence,

$$(8.5.13) \quad y = a^2 \exp[A_2] \cosh^{-2}(-b(t + A_1)).$$

Since $z = (\sqrt{y} - a)/b$, we have the solution of the second equation of system (8.5.6), namely,

$$z(t) = \frac{1}{b} \frac{a \exp\left(\frac{A_2}{2}\right)}{\cosh(-b(t + A_1))} - \frac{a}{b}.$$

Also — in view of the second equation of system (8.5.6) — we have $p_z = \dot{z}/y$. Thus, we have the solution of the fourth equation of system (8.5.6). This solution is

$$p_z(t) = -\frac{\sinh(-b(t + A_1))}{a \exp\left(\frac{A_2}{2}\right)}.$$

We can write the remaining equation of system (8.5.6) as

$$\frac{dx}{dt} = y(1 + 2\chi)\mathfrak{p},$$

where y is given by expression (8.5.13). Integrating, we get

$$x(t) = \frac{\mathfrak{p}(1 + 2\chi)a^2 \exp A_2}{b} \tanh(b(t + A_1)) + A_3,$$

where A_3 is an integration constant. Thus, we can concisely write the general solution of system (8.5.6) as

$$(8.5.14) \quad \begin{cases} x(t) = \frac{\mathfrak{p}(1+2\chi)a^2 \exp A_2}{b} \tanh(b(t+A_1)) + A_3 \\ z(t) = \frac{1}{b} \frac{a \exp\left(\frac{A_2}{2}\right)}{\cosh(b(t+A_1))} - \frac{a}{b} \\ p_x(t) = \mathfrak{p} \\ p_z(t) = -\frac{\sinh(-b(t+A_1))}{a \exp\left(\frac{A_2}{2}\right)} \end{cases}.$$

Integration constants. Now, we can find the three integration constants. At $t = 0$, using the initial conditions discussed in Section 8.5.4, we can rewrite the solutions stated in the first two expressions of set (8.5.14) as

$$\begin{aligned} x(0) &= \frac{\mathfrak{p}(1+2\chi)a^2 \exp A_2}{b} \tanh(bA_1) + A_3 = 0, \\ z(0) &= \frac{1}{b} \frac{a \exp\left(\frac{A_2}{2}\right)}{\cosh(bA_1)} - \frac{a}{b} = 0. \end{aligned}$$

Also, considering $\dot{z} = 0$ and using the second equation of system (8.5.6) combined with solution $z(t)$ given in set (8.5.14) and evaluating at $t = 0$, we can write

$$p_z(0) = \frac{\dot{z}(0)}{a^2} = -\frac{\sinh(bA_1)}{a \exp\left(\frac{A_2}{2}\right)} = \sqrt{\frac{1}{a^2} - (1+2\chi)\mathfrak{p}^2},$$

where the right-hand side is given in expression (8.5.10).

Considering the last two equations, we have a system of two equations in two unknowns, A_1 and A_2 . Solving, we obtain

$$\begin{aligned} A_1 &= -\frac{1}{b} \tanh^{-1} \sqrt{1 - (1+2\chi)\mathfrak{p}^2 a^2}, \\ A_2 &= -\ln[(1+2\chi)\mathfrak{p}^2 a^2]. \end{aligned}$$

Inserting A_1 and A_2 into the equation for $x(0)$, we obtain

$$A_3 = \frac{\sqrt{1 - (1+2\chi)\mathfrak{p}^2 a^2}}{\mathfrak{p}b}.$$

Examining A_1 , A_2 and A_3 , we see that the units of A_1 are the units of time, A_2 is dimensionless and the units of A_3 are the units of distance. This is consistent with positions of the A_1 , A_2 and A_3 in system (8.5.14).

Unique solution. Having found A_1 , A_2 and A_3 , we can rewrite solutions (8.5.14) as

$$(8.5.15) \quad \left\{ \begin{array}{l} x(t) = \frac{1}{pb} \left[\tanh \left(bt - \tanh^{-1} \sqrt{1 - (1 + 2\chi) p^2 a^2} \right) + \sqrt{1 - (1 + 2\chi) p^2 a^2} \right] \\ z(t) = \frac{a}{b} \left[\frac{1}{pa \cosh \left(\tanh^{-1} \sqrt{1 - (1 + 2\chi) p^2 a^2} - bt \right) \sqrt{1 + 2\chi}} - 1 \right] \\ p_x(t) = p \\ p_z(t) = p \sqrt{1 + 2\chi} \sinh \left(\tanh^{-1} \sqrt{1 - (1 + 2\chi) p^2 a^2} - bt \right) \end{array} \right. ,$$

where, in view of expression (8.5.12), we have

$$p = \frac{\sin \vartheta_0}{a \sqrt{(1 + 2\chi) \sin^2 \vartheta_0 + \cos^2 \vartheta_0}}$$

with ϑ_0 being the take-off phase angle.

Thus — for a velocity model given by a , which is the velocity at $z = 0$, b , which describes the increase of velocity along the z -axis, and χ , which describes the elliptical velocity dependence with direction — we can choose the phase take-off angle, ϑ_0 , and, using the first two expressions of solutions (8.5.15), obtain the ray along which the signal generated at $(0, 0)$ propagates.

Geometrical interpretation. Examination of the first two expressions of solutions (8.5.15) allows us to learn about the shape of rays for the $ab\chi$ model. We can write each of these expressions as

$$pbx(t) - \sqrt{1 - p^2 a^2 (1 + 2\chi)} = \tanh \left(bt - \tanh^{-1} \sqrt{1 - (1 + 2\chi) p^2 a^2} \right)$$

and

$$\left[\frac{b}{a} z(t) + 1 \right] pa \sqrt{1 + 2\chi} = \frac{1}{\cosh \left(bt - \tanh^{-1} \sqrt{1 - (1 + 2\chi) p^2 a^2} \right)},$$

respectively. Squaring these two equations, adding them together, and using standard identities, we obtain

$$(8.5.16) \quad \frac{\left(x - \frac{\sqrt{1 - p^2 a^2 (1 + 2\chi)}}{pb} \right)^2}{\left(\frac{1}{pb} \right)^2} + \frac{\left(z + \frac{a}{b} \right)^2}{\left(\frac{1}{pb \sqrt{1 + 2\chi}} \right)^2} = 1.$$

This is the equation of an ellipse with a centre on the line given by $z = -a/b$. In other words, in the $ab\chi$ model, rays are elliptical arcs. In view of $v(z) = a + bz$, we conclude that the centre of the ellipse corresponds to the level where the velocity vanishes. This equation is identical to the one that we will obtain in Section 14.3.2 using Lagrange's, rather than Hamilton's, ray equations.

8.5.7. Solution of eikonal equation. Using solutions (8.5.15), we can obtain the graph of the solution of the original partial differential equation, namely, eikonal equation (8.5.4), as a parametric plot

$[x(t), z(t), t]$ for all \mathbf{p} that are consistent with the original equation. In the present case, we set $\vartheta_0 \in (-\pi/2, \pi/2)$ to get, following expression (8.5.12),

$$\mathbf{p} \in \left(-1/\left[a\sqrt{1+2\chi}\right], 1/\left[a\sqrt{1+2\chi}\right] \right).$$

In the case of the present example, we can also obtain an explicit analytic form of the solution of the original partial differential equation, namely, eikonal equation (8.5.4). In other words, we can use our results to obtain $\psi(x, z)$. To do so, we proceed in the following way. Since $t = \psi(x, z)$, let us solve the first equation of set (8.5.15) for t . We get

$$(8.5.17) \quad t(x; \mathbf{p}) = \frac{\tanh^{-1} \left[\mathbf{p}bx - \sqrt{1 - \mathbf{p}^2 a^2 (1 + 2\chi)} \right] + \tanh^{-1} \sqrt{1 - \mathbf{p}^2 a^2 (1 + 2\chi)}}{b}.$$

To express t in terms of x and z , and the parameters of a given $ab\chi$ model, we solve equation (8.5.16) for \mathbf{p} . We get

$$(8.5.18) \quad \mathbf{p}(x, z) = \frac{2x}{\sqrt{[x^2 + (1 + 2\chi)z^2][(2a + bz)^2(1 + 2\chi) + b^2x^2]}}.$$

Thus, expression (8.5.17) with \mathbf{p} given by expression (8.5.18) is the solution of equation (8.5.4). In the context of expressions (8.5.17) and (8.5.18), we can write this solution as $\psi = t(x, z)$.

We can also obtain the expression for t by solving the second equation of set (8.5.15), as shown in Exercise 8.10, where it appears as expression (8.7.17). Traveltime expression (8.5.17), stated above, is valid for the entire trajectories of all rays. Traveltime expression (8.7.17), stated below, is valid for only the downgoing segment of the rays. Both traveltime expressions are valid as long as the corresponding coordinates, x and z , respectively, are increasing; coordinate x increases for the entire ray, but coordinate z increases only for the downgoing segment of the ray. This distinction of the validity of the traveltime expressions will appear again, and will be discussed in more detail, in Section 14.3.3, in the context of integration along the x -axis and the z -axis.

8.5.8. Physical interpretation of solutions. Above, we discussed the rays and traveltimes, which are obtained from the solutions of Hamilton's ray equations (8.5.6). We also discussed wavefronts, which are obtained from the solution of eikonal equation (8.5.4). The physical meaning of these solutions is obvious. We did not yet explicitly use solution p_z , although we had to use all four equations to solve system (8.5.6). Now, we will explicitly use p_z to express the magnitude of the slowness of the wavefront propagation. Since the wavefront slowness is the reciprocal of the wavefront velocity, we write

$$v = \frac{1}{\sqrt{\mathbf{p} \cdot \mathbf{p}}} = \frac{1}{\sqrt{[p_x, p_z] \cdot [p_x, p_z]}} = \frac{1}{\sqrt{\mathbf{p}^2 + p_z^2}},$$

where v is referred to as the phase velocity.

8.6. Example: Isotropy and inhomogeneity

8.6.1. Parametric form. In most cases, one would solve Hamilton's ray equations given in expression (8.2.7) by numerical methods. Since Hamilton's ray equations are first-order linear ordinary differential equations, the solution can be obtained using standard computer tools.

As shown in Section (8.5), in particular cases system (8.2.7) allows us to analytically, rather than numerically, study ray theory in the context of anisotropic inhomogeneous continua. To gain further familiarity with this system, let us now consider a formulation for isotropic inhomogeneous continua, where eikonal equation (8.1.1) reduces to

$$(8.6.1) \quad p^2 = \frac{1}{v^2(\mathbf{x})},$$

which is eikonal equation (6.10.22).

To study ray equations in isotropic inhomogeneous continua, let us choose function (8.1.3), which becomes

$$(8.6.2) \quad F(\mathbf{x}) = p^2 - \frac{1}{v^2(\mathbf{x})}.$$

Using system (8.1.13), we can write the corresponding characteristic equations as

$$(8.6.3) \quad \begin{cases} \frac{dx_i}{ds} = 2\zeta p_i \\ \frac{dp_i}{ds} = -2\zeta \frac{1}{v^3} \frac{\partial v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Also, let us choose scaling factor ζ so that s is the arclength parameter. As shown in Exercise 8.6, we obtain the arclength parametrization of system (8.6.3) by letting $\zeta = v/2$. Furthermore, as shown in Exercise 8.7, system (8.6.3) can be restated as a single expression

$$(8.6.4) \quad \frac{d}{ds} \left[\frac{1}{v(\mathbf{x})} \frac{d\mathbf{x}}{ds} \right] = -\frac{\nabla v(\mathbf{x})}{v^2(\mathbf{x})},$$

where $\mathbf{x} = [x_1, x_2, x_3]$.

Equation (8.6.4) 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 .

8.6.2. Explicit form. Consider a three-dimensional isotropic inhomogeneous continuum where $\mathbf{x} = [x, y, z]$. Expression (8.6.4) can be explicitly written as three parametric equations for $x(s)$, $y(s)$ and

$z(s)$, namely,

$$(8.6.5) \quad \begin{aligned} \frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dx}{ds} \right) &= -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial x}, \\ \frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dy}{ds} \right) &= -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial y}, \\ \frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dz}{ds} \right) &= -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial z}, \end{aligned}$$

where s is the arclength parameter along the ray. Consequently, all three equations are related by $ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}$, where x , y and z are the orthonormal coordinates. Consequently, as shown in Exercise 8.11, instead of using the parametric form, under certain conditions related to the behaviour of the curve $\mathbf{x}(s)$, we can write equations (8.6.5) as two equations for $x(z)$ and $y(z)$, namely,

$$(8.6.6) \quad \frac{d}{dz} \left[\frac{1}{v(\mathbf{x})} \frac{\frac{dx}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}} \right] = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial x} \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}$$

and

$$(8.6.7) \quad \frac{d}{dz} \left[\frac{1}{v(\mathbf{x})} \frac{\frac{dy}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}} \right] = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial y} \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1},$$

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 (8.6.6) and (8.6.7) vanish and, for the resulting equations, we can obtain an analytic solution, as shown in Exercise 8.12. 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.⁹

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 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 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.

In Chapter 11, however, we will explore another formulation of rays using the 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

⁹Readers 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.

Part 3. 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.¹⁰ 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 13.2.2, 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.¹¹



¹⁰Readers interested in the WKBJ method might refer to Aki, K. and Richards, P.G., (2002) *Quantitative seismology* (2nd edition): University Science Books, pp. 434 – 437, (Box 9.6), and to Woodhouse, N.M.J., (1992) *Geometric quantization* (2nd edition): Oxford Science Publications, pp. 197 – 201 and pp. 236 – 249.

¹¹Readers 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, and to Batterman, R.W., (2002) *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*: Oxford University Press, Chapters 6 and 7.

8.7. Exercises

EXERCISE 8.1. Consider a three-dimensional isotropic inhomogeneous continuum. Using Hamilton's ray equations (8.2.7), show that, in isotropic continua, rays are orthogonal to wavefronts.

SOLUTION 8.1. Following expression (8.2.8), we can explicitly write Hamiltonian $\mathcal{H}(\mathbf{x}, \mathbf{p})$, in a three-dimensional isotropic inhomogeneous continuum, as

$$(8.7.1) \quad \mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} p^2 v^2(\mathbf{x}) = \frac{1}{2} [p_1, p_2, p_3] \cdot [p_1, p_2, p_3] v^2(x_1, x_2, x_3).$$

The corresponding Hamilton's ray equations (8.2.7) are

$$(8.7.2) \quad \begin{cases} \dot{x}_1 = p_1 v^2 \\ \dot{x}_2 = p_2 v^2 \\ \dot{x}_3 = p_3 v^2 \\ \dot{p}_1 = -p^2 v \frac{\partial v}{\partial x_1} \\ \dot{p}_2 = -p^2 v \frac{\partial v}{\partial x_2} \\ \dot{p}_3 = -p^2 v \frac{\partial v}{\partial x_3} \end{cases}.$$

Recalling definition (7.3.2), we can write the first three equations of system (8.7.2) as

$$(8.7.3) \quad [\dot{x}_1, \dot{x}_2, \dot{x}_3] = v^2 \left[\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, \frac{\partial \psi}{\partial x_3} \right].$$

The left-hand side of equation (8.7.3) 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 (8.7.3) as

$$\mathbf{t}|_{x_1, x_2, x_3} = v^2 (\nabla \psi)|_{x_1, x_2, x_3}.$$

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.

EXERCISE 8.2. Derive expression (8.4.12) using level-set function (8.1.4) and characteristic equations (8.1.13).

SOLUTION 8.2. In view of expression (8.4.10), the ray angle can be stated as

$$(8.7.4) \quad \tan \theta = \frac{\frac{dx_1}{ds}}{\frac{dx_3}{ds}},$$

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 (8.7.4) 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 (8.1.4), namely,

$$(8.7.5) \quad F(\mathbf{p}) = p^2 v^2(\mathbf{p}) = 1.$$

At a given point, expression (8.7.5) is not a function of \mathbf{x} , and, hence, $\partial F / \partial x_i = 0$. Thus, characteristic equations (8.1.13), are reduced to

$$\frac{dx_i}{ds} = \zeta \frac{\partial F}{\partial p_i} = 2\zeta \left(p_i v^2 + p^2 v \frac{\partial v}{\partial p_i} \right), \quad i \in \{1, 2\}.$$

Following expression (8.7.4), we can write the ray angle as

$$(8.7.6) \quad \tan \theta = \frac{2\zeta \left(p_1 v^2 + p^2 v \frac{\partial v}{\partial p_1} \right)}{2\zeta \left(p_3 v^2 + p^2 v \frac{\partial v}{\partial p_3} \right)} = \frac{p_1 v^2 + p^2 v \frac{\partial v}{\partial p_1}}{p_3 v^2 + p^2 v \frac{\partial v}{\partial p_3}}.$$

We wish to express the quantities on the right-hand side of expression (8.7.6) in terms of the phase angle, ϑ . Recalling expression (8.4.5), we can write the differential operator in the numerator as

$$\begin{aligned} \frac{\partial}{\partial p_1} &= \frac{\partial \vartheta}{\partial p_1} \frac{\partial}{\partial \vartheta} = \frac{\partial \arctan \frac{p_1}{p_3}}{\partial p_1} \frac{\partial}{\partial \vartheta} \\ &= \frac{\frac{1}{p_3}}{1 + \left(\frac{p_1}{p_3} \right)^2} \frac{\partial}{\partial \vartheta} = \frac{p_3}{p^2} \frac{\partial}{\partial \vartheta}. \end{aligned}$$

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 (8.7.6), we can rewrite it as

$$\tan \theta = \frac{p_1 v^2 + p_3 v \frac{\partial v}{\partial \vartheta}}{p_3 v^2 - p_1 v \frac{\partial v}{\partial \vartheta}} = \frac{p_1 + p_3 \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{p_3 - p_1 \frac{1}{v} \frac{\partial v}{\partial \vartheta}}.$$

Again, recalling expression (8.4.5), we divide both the numerator and the denominator by p_3 to obtain

$$\tan \theta = \frac{\frac{p_1}{p_3} + \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{1 - \frac{p_1}{p_3} \frac{1}{v} \frac{\partial v}{\partial \vartheta}} = \frac{\tan \vartheta + \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{1 - \frac{\tan \vartheta}{v} \frac{\partial v}{\partial \vartheta}},$$

which is expression (8.4.12), as required.

EXERCISE 8.3. *Derive expression (8.4.12) using level-set function (8.1.3) and characteristic equations (8.1.13).*

SOLUTION 8.3. Recall expression (8.7.4). Consider a given point in an anisotropic continuum and the level-set function that is given by expression (8.1.3), namely,

$$(8.7.7) \quad F(\mathbf{p}) = p^2 - \frac{1}{v^2(\mathbf{p})} = 0.$$

At a given point, expression (8.7.7) is not a function of \mathbf{x} , and, hence, $\partial F / \partial x_i = 0$. Thus, characteristic equations (8.1.13), are reduced to

$$\frac{dx_i}{ds} = \zeta \frac{\partial F}{\partial p_i} = 2\zeta \left(p_i + \frac{1}{v^3} \frac{\partial v}{\partial p_i} \right), \quad i \in \{1, 2\}.$$

Following expression (8.7.4), we can write the ray angle as

$$(8.7.8) \quad \tan \theta = \frac{p_1 + \frac{1}{v^3} \frac{\partial v}{\partial p_1}}{p_3 + \frac{1}{v^3} \frac{\partial v}{\partial p_3}}.$$

We wish to express the quantities on the right-hand side of expression (8.7.8) in terms of the phase angle, ϑ . In view of expression (8.4.5), we consider the differential operator in the numerator, namely,

$$\begin{aligned} \frac{\partial}{\partial p_1} &= \frac{\partial \vartheta}{\partial p_1} \frac{\partial}{\partial \vartheta} = \frac{\partial \arctan \frac{p_1}{p_3}}{\partial p_1} \frac{\partial}{\partial \vartheta} \\ &= \frac{\frac{1}{p_3}}{1 + \left(\frac{p_1}{p_3}\right)^2} \frac{\partial}{\partial \vartheta} = \frac{p_3}{p_1^2 + p_3^2} \frac{\partial}{\partial \vartheta}. \end{aligned}$$

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 (8.7.8), we can rewrite it as

$$\tan \theta = \frac{p_1 + \frac{1}{v^3} \frac{p_3}{p^2} \frac{\partial v}{\partial \vartheta}}{p_3 - \frac{1}{v^3} \frac{p_1}{p^2} \frac{\partial v}{\partial \vartheta}}.$$

Following eikonal equation (7.3.8), we can state $p^2 v^2 = 1$, and, hence, we can write

$$\tan \theta = \frac{p_1 + \frac{p_3}{v} \frac{\partial v}{\partial \vartheta}}{p_3 - \frac{p_1}{v} \frac{\partial v}{\partial \vartheta}}.$$

Again, recalling expression (8.4.5) and dividing both numerator and denominator by p_3 , we obtain

$$\tan \theta = \frac{\frac{p_1}{p_3} + \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{1 - \frac{p_1}{p_3} \frac{\partial v}{v \partial \vartheta}} = \frac{\tan \vartheta + \frac{1}{v(\vartheta)} \frac{\partial v(\vartheta)}{\partial \vartheta}}{1 - \frac{\tan \vartheta}{v(\vartheta)} \frac{\partial v(\vartheta)}{\partial \vartheta}},$$

which is expression (8.4.12), as required.

EXERCISE 8.4. Using expressions (8.4.9) and (8.4.12), derive expression (8.4.13).

SOLUTION 8.4. Using algebraic manipulation, we can write expression (8.4.12), namely,

$$\tan \theta = \frac{\tan \vartheta + \frac{1}{v} \frac{\partial v}{\partial \vartheta}}{1 - \frac{\tan \vartheta}{v} \frac{\partial v}{\partial \vartheta}},$$

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

$$(8.7.9) \quad \frac{\partial v}{\partial \vartheta} = v \tan (\theta - \vartheta).$$

Consider expression (8.4.9). In view of expression (8.7.9), 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

$$(8.7.10) \quad V = \frac{v}{\cos (\theta - \vartheta)},$$

which, notably, is expression (8.4.14). 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 (8.4.13), 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 (8.7.10) as

$$V = \frac{v}{\mathbf{n} \cdot \mathbf{t}},$$

which is expression (8.4.13), as required.

EXERCISE 8.5. *Derive a particular case of expression (8.4.12) that corresponds to the elliptical velocity dependence.*

SOLUTION 8.5. Inserting expression (6.10.11) into expression (8.4.9), we can write the magnitude of the ray-velocity vector as

$$(8.7.11) \quad V(\vartheta) = \sqrt{\frac{v_x^4 \tan^2 \vartheta + v_z^4}{v_x^2 \tan^2 \vartheta + v_z^2}}.$$

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 (6.10.11) into expression (8.4.12), we obtain

$$(8.7.12) \quad \tan \theta = \left(\frac{v_x}{v_z} \right)^2 \tan \vartheta,$$

which is the relation between the phase angle and the ray angle for elliptical velocity dependence. Expression (8.7.12) is analogous to expression (9.4.8), which corresponds to SH waves in transversely isotropic continua. Inserting expression (8.7.12) into expression (8.7.11), we can write the magnitude of the ray-velocity vector in terms of ray-related quantities, namely,

$$(8.7.13) \quad V(\theta) = V_z \sqrt{\frac{\tan^2 \theta + 1}{\left(\frac{V_z}{V_x}\right)^2 \tan^2 \theta + 1}},$$

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.

REMARK 8.7.1. Characteristic equations (8.1.13) can be parametrized by choosing various expressions for scaling factor ζ . Two typical examples are shown in Exercises 8.7 and 8.8, below. In both cases, we invoke function (8.1.3) and consider isotropic inhomogeneous continua. Hence, characteristic equations (8.1.13) become equations (8.6.3).

EXERCISE 8.6. ¹²Show that the arclength parametrization of system (8.6.3) requires $\zeta = v/2$.

SOLUTION 8.6. In general, if $\mathbf{x} = \mathbf{x}(s)$, using definition (8.1.2), we can write

$$\frac{d\psi(\mathbf{x})}{ds} = \sum_{i=1}^3 \frac{\partial \psi}{\partial x_i} \frac{dx_i}{ds} = \sum_{i=1}^3 p_i \frac{dx_i}{ds},$$

which, in view of characteristic equations (8.6.3), we can rewrite as

$$\frac{d\psi(\mathbf{x})}{ds} = \sum_{i=1}^3 2\zeta p_i p_i,$$

which we can immediately restate as

$$(8.7.14) \quad \frac{d\psi(\mathbf{x})}{ds} = 2\zeta p^2.$$

If s is the arclength parameter, then

$$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2},$$

and, hence,

$$(8.7.15) \quad \frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2},$$

where t stands for traveltime. Combining expressions (8.7.14) and (8.7.15), we obtain

$$(8.7.16) \quad \frac{d\psi(\mathbf{x})}{dt} = \frac{d\psi(\mathbf{x})}{ds} \frac{ds}{dt} = 2\zeta p^2 \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}.$$

¹²See also Section 8.6.1.

In view of condition (8.2.1) and since the square root gives the magnitude of velocity, we can rewrite equation (8.7.16) as

$$1 = 2\zeta p^2 v.$$

Solving for ζ , where — in view of equation (8.6.1) — 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 8.7. ¹³Letting $\zeta = v/2$, show that characteristic equations (8.6.3) can be reduced to equation (8.6.4), namely,

$$\frac{d}{ds} \left[\frac{1}{v(\mathbf{x})} \frac{d\mathbf{x}}{ds} \right] = -\frac{\nabla v(\mathbf{x})}{v^2(\mathbf{x})}.$$

SOLUTION 8.7. If $\zeta = v/2$, characteristic equations (8.6.3) become

$$\begin{cases} \frac{dx_i}{ds} = vp_i \\ \frac{dp_i}{ds} = -\frac{1}{v^2} \frac{\partial v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

The first equation of this system can be rewritten as

$$p_i = \frac{1}{v} \frac{dx_i}{ds}, \quad i \in \{1, 2, 3\}.$$

Hence, the second equation can be stated as

$$\frac{dp_i}{ds} = \frac{d}{ds} \left(\frac{1}{v} \frac{dx_i}{ds} \right) = -\frac{1}{v^2} \frac{\partial v}{\partial x_i}, \quad i \in \{1, 2, 3\}.$$

Thus, the system of characteristic equations can be written as a single expression

$$\frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{d\mathbf{x}}{ds} \right) = -\frac{\nabla v(\mathbf{x})}{v^2(\mathbf{x})},$$

where $\mathbf{x} = [x_1, x_2, x_3]$, which is equation (8.6.4), as required.

EXERCISE 8.8. ¹⁴Letting $\zeta = v^2/2$, show that characteristic equations (8.6.3) can be written as a system of equations given by

$$\begin{cases} \dot{x}_i = v^2 p_i \\ \dot{p}_i = -\frac{\partial}{\partial x_i} \ln v \end{cases}, \quad i \in \{1, 2, 3\}.$$

SOLUTION 8.8. As shown in Section 8.2, using function (8.1.3) and letting $\zeta = v^2/2$ results in the time parametrization of characteristic equations (8.1.13). Hence, characteristic equations (8.6.3) can be written

¹³See also Section 8.6.1.

¹⁴See also Section 8.2.2 and Exercise 13.3.

as

$$\begin{cases} \dot{x}_i := \frac{dx_i}{dt} = v^2 p_i \\ \dot{p}_i := \frac{dp_i}{dt} = -\frac{1}{v} \frac{\partial v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Following the chain rule, we can restate the second equation of this system to obtain

$$\begin{cases} \dot{x}_i = v^2 p_i \\ \dot{p}_i = -\frac{\partial}{\partial x_i} \ln v \end{cases}, \quad i \in \{1, 2, 3\},$$

as required.

REMARK 8.7.2. Lemma 8.1.1 shows that both functions (8.1.3) and (8.1.4) yield the same characteristics. Thus, in a seismological context, both functions result in the same rays. In view of Exercise 8.8, Exercise 8.9 illustrates this property for isotropic inhomogeneous continua.

EXERCISE 8.9. ¹⁵Using characteristic equations (8.2.4) and considering functions (8.1.4), show that, for isotropic inhomogeneous continua, we obtain the system of equations

$$\begin{cases} \dot{x}_i = v^2 p_i \\ \dot{p}_i = -\frac{\partial}{\partial x_i} \ln v \end{cases}, \quad i \in \{1, 2, 3\}.$$

SOLUTION 8.9. Considering functions (8.1.4) for isotropic inhomogeneous continua, characteristic equations (8.2.4), which are parametrized in terms of time, become

$$\begin{cases} \dot{x}_i = v^2 p_i \\ \dot{p}_i = -p^2 v \frac{\partial v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Since $p^2 v^2 = 1$, we can write

$$\begin{cases} \dot{x}_i = v^2 p_i \\ \dot{p}_i = -\frac{1}{v} \frac{\partial v}{\partial x_i} = -\frac{\partial}{\partial x_i} \ln v \end{cases}, \quad i \in \{1, 2, 3\},$$

which is also the solution of Exercise 8.8.

EXERCISE 8.10. Solve the second equation of set (8.5.15) for t .

SOLUTION 8.10. Consider the second equation of set (8.5.15), namely,

$$z(t) = \frac{a}{b} \left[\frac{1}{\mathfrak{p} a \cosh \left(\tanh^{-1} \sqrt{1 - (1 + 2\chi)} \mathfrak{p}^2 a^2 - bt \right) \sqrt{1 + 2\chi}} - 1 \right].$$

¹⁵See also Exercise 13.3.

We can rearrange this equation to get

$$\cosh \left(\tanh^{-1} \sqrt{1 - (1 + 2\chi) \mathbf{p}^2 a^2} - bt \right) = \frac{1}{\mathbf{p} (a + bz) \sqrt{1 + 2\chi}}.$$

Taking \cosh^{-1} of both sides, we write

$$\tanh^{-1} \sqrt{1 - (1 + 2\chi) \mathbf{p}^2 a^2} - bt = \cosh^{-1} \frac{1}{\mathbf{p} (a + bz) \sqrt{1 + 2\chi}}.$$

Solving for t , we obtain

$$(8.7.17) \quad t(z; \mathbf{p}) = \frac{1}{b} \left(\tanh^{-1} \sqrt{1 - (1 + 2\chi) \mathbf{p}^2 a^2} - \cosh^{-1} \frac{1}{\mathbf{p} (a + bz) \sqrt{1 + 2\chi}} \right),$$

as required.

REMARK 8.7.3. Expression (8.7.17) is valid for the downgoing segment of the ray, unlike expression (8.5.17), which is valid for the entire ray, as discussed on page 251.

EXERCISE 8.11. *Formally, show the steps leading from set (8.6.5) to equations (8.6.6) and (8.6.7).*

SOLUTION 8.11. The first two equations can be written as

$$\frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dx}{ds} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial x},$$

$$\frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dy}{ds} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial y},$$

which leads to

$$\frac{ds}{dz} \frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dx}{ds} \frac{ds}{dz} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial x} \frac{ds}{dz},$$

$$\frac{ds}{dz} \frac{d}{ds} \left(\frac{1}{v(\mathbf{x})} \frac{dy}{ds} \frac{ds}{dz} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial y} \frac{ds}{dz},$$

where we multiplied both sides of the equations by ds/dz , and we multiplied the factors inside the parentheses by unity in the form $(ds/dz) (dz/ds)$. The two equations can be immediately restated as

$$\frac{d}{dz} \left(\frac{1}{v(\mathbf{x})} \frac{dx}{dz} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial x} \frac{ds}{dz},$$

(8.7.18)

$$\frac{d}{dz} \left(\frac{1}{v(\mathbf{x})} \frac{dy}{dz} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial y} \frac{ds}{dz}.$$

We assume the invertibility 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

$$\begin{aligned}\frac{ds}{dz} &\equiv \frac{ds(x(z), y(z), z)}{dz} = \frac{\sqrt{[dx(z)]^2 + [dy(z)]^2 + [dz(z)]^2}}{dz} \\ &= \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + \left(\frac{dz}{dz}\right)^2}.\end{aligned}$$

Thus, since $ds/dz = 1/(dz/ds)$, equations (8.7.18) can be stated as

$$\frac{d}{dz} \left(\frac{1}{v(\mathbf{x})} \frac{\frac{dx}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial x} \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}$$

and

$$\frac{d}{dz} \left(\frac{1}{v(\mathbf{x})} \frac{\frac{dy}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}} \right) = -\frac{1}{v^2(\mathbf{x})} \frac{\partial v(\mathbf{x})}{\partial y} \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1},$$

which — as required — are equations (8.6.6) and (8.6.7), respectively.

EXERCISE 8.12. *Solve ray equations (8.6.6) and (8.6.7) for a vertically inhomogeneous continuum, where $v = v(z)$.*

SOLUTION 8.12. Since $v = v(z)$, the right-hand sides of equations (8.6.6) and (8.6.7) vanish. Consequently, we obtain

$$\begin{aligned}(8.7.19) \quad & \frac{d}{dz} \left[\frac{1}{v(z)} \frac{\frac{dx}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}} \right] = 0, \\ & \frac{d}{dz} \left[\frac{1}{v(z)} \frac{\frac{dy}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1}} \right] = 0.\end{aligned}$$

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 (8.7.19). In view of the vanishing of the total derivative, equation (8.7.19) can be restated as

$$(8.7.20) \quad \frac{1}{v(z)} \frac{\frac{dx}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + 1}} = p,$$

where \mathbf{p} is a constant. Equation (8.7.20), can be rewritten as

$$\left(\frac{dx}{dz}\right)^2 = \mathbf{p}^2 v^2 \left[\left(\frac{dx}{dz}\right)^2 + 1 \right].$$

Solving for dx/dz , we obtain

$$(8.7.21) \quad \frac{dx}{dz} = \frac{\mathbf{p}v}{\sqrt{1 - \mathbf{p}^2 v^2}},$$

and, hence we can state the solution as

$$x(z) = \int_{z_0}^z \frac{\mathbf{p}v(\xi)}{\sqrt{1 - \mathbf{p}^2 v^2(\xi)}} d\xi,$$

where ξ is the integration variable. This is a standard expression for a ray in vertically inhomogeneous continua, where, as shown in Exercise 8.13, $\mathbf{p} = \sin \theta / v(z)$.

EXERCISE 8.13. Consider equation (8.7.20). Show that $\mathbf{p} = \sin \theta / v(z)$.

SOLUTION 8.13. Since $dx/dz = \tan \theta$, following standard trigonometric identities, we can write equation (8.7.20) as

$$\mathbf{p} = \frac{1}{v(z)} \frac{\frac{dx}{dz}}{\sqrt{\left(\frac{dx}{dz}\right)^2 + 1}} = \frac{1}{v(z)} \frac{\tan \theta}{\sqrt{\tan^2 \theta + 1}} = \frac{\frac{\sin \theta}{\cos \theta}}{v(z) \sec \theta} = \frac{\sin \theta}{v(z)}.$$

EXERCISE 8.14. Consider a one-dimensional homogeneous continuum. Show that solution $x(t)$ of Hamilton's ray equations (8.2.7) corresponds to coordinates (6.4.4), which can be written as

$$x(t) = x_0 \pm vt.$$

SOLUTION 8.14. For a one-dimensional case, letting $x_1 \equiv x$ and $p_1 \equiv p$, we can write Hamilton's ray equations (8.2.7) as

$$\begin{cases} \dot{x} = pv^2 + p^2 v \frac{\partial v}{\partial p} \\ \dot{p} = -p^2 v \frac{\partial v}{\partial x} \end{cases},$$

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 (7.3.8) and since $v = \pm 1/p$, we can write

$$\dot{x} := \frac{dx}{dt} = \pm v.$$

Solving for dx , we obtain

$$dx = \pm v dt.$$

Integrating both sides, we obtain

$$x(t) = x_0 \pm vt,$$

as required and where x_0 is the integration constant.

EXERCISE 8.15. *Following ray equation (8.6.4), show that rays are straight lines in homogeneous continua.*¹⁶

SOLUTION 8.15. For homogeneous continua, v is constant and, hence, the right-hand side of ray equation (8.6.4) vanishes. Thus, we obtain

$$\frac{d}{ds} \left(\frac{1}{v} \frac{d\mathbf{x}}{ds} \right) = \mathbf{0}.$$

The vanishing of the total derivative implies that the term in parentheses can be written as

$$\frac{1}{v} \frac{d\mathbf{x}}{ds} = \mathbf{C},$$

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 $\mathbf{a} := \mathbf{C}v$.

¹⁶Readers interested in an insightful explanation of the straight-line appearance of the optical rays by the theory of quantum electrodynamics might refer to Feynman, R.P., (1985/2006) QED: The strange theory of light and matter: Princeton University Press, pp. 53 – 56, where we read that

the idea that light goes in a straight line is a convenient approximation to describe what happens in the world that is familiar to us.

CHAPTER 9

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.¹

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. In Chapters 7 and 8, 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 and Hamilton'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

¹Readers interested in philosophical aspects of this statement might refer to Steiner, M., (1998) The applicability of mathematics as a philosophical problem: Harvard University Press, pp. 1 – 11.

Also, we might remind ourselves that

La physique a bâti ses propres critères de validation: n'est valide à ses yeux que ce qui est confirmé par l'expérience; ni la cohérence logique ni l'élégance mathématique ne suffisent à elles seules à étayer une théorie physique, même si elles sont souvent de précieux indicateurs.*

Etienne Klein (1991) Conversations avec le Sphinx: Les paradoxes en physique

*Physics has established its own criteria of validation: In its eyes, only things confirmed by experiments are valid; neither logical consistency nor mathematical elegance suffice by themselves to support a physical theory, even though they are often valuable indicators.

isotropic continuum.² The chapter concludes with a discussion of the three corresponding phase-slowness surfaces and their intersections.

9.1. Explicit form of Christoffel's equations

We wish to study Christoffel's equations, shown in expression (7.3.3), namely,

$$(9.1.1) \quad \sum_{k=1}^3 \left(\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) p_j p_l - \rho(\mathbf{x}) \delta_{ik} \right) A_k(\mathbf{x}) = \mathbf{0}, \quad i \in \{1, 2, 3\},$$

in the context of a specific continuum. In other words, we wish to rewrite equations (9.1.1) 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,

$$(9.1.2) \quad p^2 = \frac{1}{v^2},$$

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 (9.1.1) as

$$(9.1.3) \quad p^2 \sum_{k=1}^3 \left(\sum_{j=1}^3 \sum_{l=1}^3 c_{ijkl}(\mathbf{x}) n_j n_l - \rho(\mathbf{x}) v^2 \delta_{ik} \right) A_k(\mathbf{x}) = \mathbf{0}, \quad i \in \{1, 2, 3\}.$$

We can state equations (9.1.3) in matrix notation as

$$(9.1.4) \quad p^2 [\Gamma(\mathbf{x}, \mathbf{n}) - \rho(\mathbf{x}) v^2 \mathbf{I}] \mathbf{A}(\mathbf{x}) = \mathbf{0},$$

where

$$(9.1.5) \quad \Gamma(\mathbf{x}, \mathbf{n}) = \begin{bmatrix} \sum_{j=1}^3 \sum_{l=1}^3 c_{1j1l}(\mathbf{x}) n_j n_l & \sum_{j=1}^3 \sum_{l=1}^3 c_{1j2l}(\mathbf{x}) n_j n_l & \sum_{j=1}^3 \sum_{l=1}^3 c_{1j3l}(\mathbf{x}) n_j n_l \\ \sum_{j=1}^3 \sum_{l=1}^3 c_{2j1l}(\mathbf{x}) n_j n_l & \sum_{j=1}^3 \sum_{l=1}^3 c_{2j2l}(\mathbf{x}) n_j n_l & \sum_{j=1}^3 \sum_{l=1}^3 c_{2j3l}(\mathbf{x}) n_j n_l \\ \sum_{j=1}^3 \sum_{l=1}^3 c_{3j1l}(\mathbf{x}) n_j n_l & \sum_{j=1}^3 \sum_{l=1}^3 c_{3j2l}(\mathbf{x}) n_j n_l & \sum_{j=1}^3 \sum_{l=1}^3 c_{3j3l}(\mathbf{x}) n_j n_l \end{bmatrix},$$

²Readers interested in an insightful formulation of Christoffel's equations and waves in a transversely isotropic continuum, which is exemplified by a hexagonal crystal, might also refer to Newnham, R.E., (2005) Properties of materials: Anisotropy, symmetry, structure: Oxford University Press, pp. 249 – 255.

which is equivalent to Christoffel's matrix (7.3.7). ³Using formula (3.2.5), we can state the entries of matrix $\Gamma(\mathbf{x}, \mathbf{n})$ in terms of the elasticity parameters $C_{mn}(\mathbf{x})$, to obtain

$$\begin{aligned}
 \Gamma_{11} &= C_{11}n_1^2 + C_{66}n_2^2 + C_{55}n_3^2 + 2(C_{16}n_1n_2 + C_{56}n_2n_3 + C_{15}n_1n_3), \\
 \Gamma_{22} &= C_{66}n_1^2 + C_{22}n_2^2 + C_{44}n_3^2 + 2(C_{26}n_1n_2 + C_{24}n_2n_3 + C_{46}n_1n_3), \\
 \Gamma_{33} &= C_{55}n_1^2 + C_{44}n_2^2 + C_{33}n_3^2 + 2(C_{45}n_1n_2 + C_{34}n_2n_3 + C_{35}n_1n_3), \\
 \Gamma_{12} &= \Gamma_{21} \\
 &= C_{16}n_1^2 + C_{26}n_2^2 + C_{45}n_3^2 \\
 &\quad + (C_{12} + C_{66})n_1n_2 + (C_{25} + C_{46})n_2n_3 + (C_{14} + C_{56})n_1n_3, \\
 \Gamma_{13} &= \Gamma_{31} \\
 &= C_{15}n_1^2 + C_{46}n_2^2 + C_{35}n_3^2 \\
 &\quad + (C_{14} + C_{56})n_1n_2 + (C_{36} + C_{45})n_2n_3 + (C_{13} + C_{55})n_1n_3, \\
 \Gamma_{23} &= \Gamma_{32} \\
 &= C_{56}n_1^2 + C_{24}n_2^2 + C_{34}n_3^2 \\
 &\quad + (C_{25} + C_{46})n_1n_2 + (C_{23} + C_{44})n_2n_3 + (C_{36} + C_{45})n_1n_3,
 \end{aligned}
 \tag{9.1.6}$$

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 (9.1.6) allow us to state Christoffel's equations for a given continuum. Hence, we can conveniently study behaviour of the continuum in terms of its properties, namely, its mass density and elasticity parameters.

System (9.1.4) is a homogeneous system of linear equations. In general, such a system 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.

³Note that it is also common to divide the elasticity parameters by mass density and, hence, to write Christoffel's equations (9.1.1) as

$$\sum_{k=1}^3 \left(\sum_{j=1}^3 \sum_{l=1}^3 \frac{c_{ijkl}(\mathbf{x})}{\rho(\mathbf{x})} p_j p_l - \delta_{ik} \right) A_k(\mathbf{x}) = 0, \quad i \in \{1, 2, 3\},$$

where, as we see in view of Exercise 2.6, the c_{ijkl}/ρ have units of velocity squared. The corresponding solvability condition can be written as

$$\det [\Gamma_{ik}(\mathbf{x}, \mathbf{p}) - \delta_{ik}] = 0, \quad i, k \in \{1, 2, 3\},$$

where the entries of matrix $\Gamma(\mathbf{x}, \mathbf{p})$ are

$$\Gamma_{ik}(\mathbf{x}, \mathbf{p}) := \sum_{j=1}^3 \sum_{l=1}^3 \frac{c_{ijkl}(\mathbf{x})}{\rho(\mathbf{x})} p_j p_l, \quad i, k \in \{1, 2, 3\}.$$

Each of the three eigenvalues of $\Gamma(\mathbf{x}, \mathbf{p})$, namely, $G_i(\mathbf{x}, \mathbf{p})$, where $i \in \{1, 2, 3\}$, results in an eikonal equation, which we can write as

$$G_i(\mathbf{x}, \mathbf{p}) = 1, \quad i \in \{1, 2, 3\},$$

and which is equivalent to equation (7.3.8).

We wish to examine the solvability of system (9.1.4). Since, for physically meaningful solutions, we require $p^2 \neq 0$, as discussed in Section 7.3, system (9.1.4) can be written as

$$(9.1.7) \quad [\Gamma(\mathbf{x}, \mathbf{n}) - \rho(\mathbf{x}) v^2 \mathbf{I}] \mathbf{A}(\mathbf{x}) = \mathbf{0}.$$

Hence, we can write the solvability condition of system (9.1.7) as

$$(9.1.8) \quad \det \begin{bmatrix} \Gamma_{11}(\mathbf{x}, \mathbf{n}) - \rho(\mathbf{x}) v^2 & \Gamma_{12}(\mathbf{x}, \mathbf{n}) & \Gamma_{13}(\mathbf{x}, \mathbf{n}) \\ \Gamma_{12}(\mathbf{x}, \mathbf{n}) & \Gamma_{22}(\mathbf{x}, \mathbf{n}) - \rho(\mathbf{x}) v^2 & \Gamma_{23}(\mathbf{x}, \mathbf{n}) \\ \Gamma_{13}(\mathbf{x}, \mathbf{n}) & \Gamma_{23}(\mathbf{x}, \mathbf{n}) & \Gamma_{33}(\mathbf{x}, \mathbf{n}) - \rho(\mathbf{x}) v^2 \end{bmatrix} = 0.$$

Determinantal equation (9.1.8) is an eigenvalue equation. We wish to learn about the associated eigenvalues and eigenvectors. Γ is a symmetric matrix as shown in Exercise 7.3. Consequently, we can invoke explicitly two theorems of linear algebra, namely,⁴

THEOREM 9.1.1. *Since Γ is symmetric, the corresponding eigenvalues are real.*

and

THEOREM 9.1.2. *Since Γ is symmetric, the corresponding eigenvectors are orthogonal to each other.*

In view of Theorem 9.1.1, the determinantal equation, stated in expression (9.1.8), has three real roots — the eigenvalues ρv_i^2 , where $i = 1, 2, 3$. Furthermore, in view of Theorem 9.1.2, the three corresponding eigenvectors are orthogonal to each other.

To recognize the physical meaning of the eigenvalues and eigenvectors of system (9.1.7), consider trial solution (7.2.1), which led to Christoffel's equations and which can be written as

$$(9.1.9) \quad \mathbf{u}(\mathbf{x}, t) = \mathbf{A}(\mathbf{x}) f\{v_0 [\psi(\mathbf{x}) - t]\}.$$

Examining expression (9.1.9) and in view of definition (7.3.2), namely, $p_j := \partial\psi/\partial x_j$, and expression (9.1.2), namely, $p^2 = 1/v^2$, 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 9.1.1, these velocities are real.

Also, as stated in trial solution (7.2.1), $\mathbf{A}(\mathbf{x})$ 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 9.1.2, each wave exhibits the displacement vector that is orthogonal to the displacement vectors of the other two waves.

The three displacement vectors are orthogonal to each other at a given point of the continuum only if all three corresponding wavefronts exhibit the same direction at that point. In seismological studies, if

⁴For proofs of Theorem 9.1.1 and Theorem 9.1.2, interested readers might refer to Anton, H., (1973) Elementary linear algebra: John Wiley & Sons, p. 289 and p. 399, and pp. 286 – 287, respectively.

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 (9.1.5), we can also conclude that, for a given wave in a continuum defined by stress-strain equations (7.1.2), 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 9.1.

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 $C_{mn}(\mathbf{x}) \equiv C_{mn}$.

9.2. Christoffel's equations and anisotropic continua

9.2.1. Introductory comments. We wish to study equation (9.1.8), which provides us with the phase velocities of the three waves within an anisotropic continuum. We also wish to examine the eigenvectors of the corresponding matrix $\mathbf{\Gamma}$, which are the displacement vectors.

Explicit expressions for these velocities in a generally anisotropic continuum can be obtained by inserting entries (9.1.6) into equation (9.1.8). 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 (9.1.7).

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 10, 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 of 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.

9.2.2. 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 (5.6.4).

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 (9.1.6) and in view of elasticity matrix (5.6.4), we note that system (9.1.7) becomes

$$(9.2.1) \quad \begin{bmatrix} C_{55} - \rho v^2 & C_{45} & 0 \\ C_{45} & C_{44} - \rho v^2 & 0 \\ 0 & 0 & C_{33} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

System (9.2.1) can be rewritten as

$$(9.2.2) \quad \begin{bmatrix} C_{55} - \rho v^2 & C_{45} \\ C_{45} & C_{44} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and

$$(9.2.3) \quad [C_{33} - \rho v^2] A_3 = 0.$$

The displacement vectors associated with equations (9.2.2) are contained in the x_1x_2 -plane. The displacement vector associated with equation (9.2.3) coincides with the x_3 -axis. Hence, the displacement directions associated with equations (9.2.2) are orthogonal to the direction of propagation, while the displacement direction associated with equation (9.2.3) 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.⁵

Note that this monoclinic example illustrates the fact that, along the symmetry axes, all waves propagate as pure-mode waves.

⁵Readers interested in an insightful description of the pure-mode waves might also refer to Newnham, R.E., (2005) Properties of materials: Anisotropy, symmetry, structure: Oxford University Press, pp. 256 – 258.

Phase velocities along symmetry axis. In order to obtain the phase velocity along the symmetry axis, consider equations (9.2.2). 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,

$$\rho^2 (v^2)^2 - [(C_{44} + C_{55}) \rho] (v^2) - (C_{45}^2 - C_{44}C_{55}) = 0,$$

and, hence, the velocities of the S waves are

$$(9.2.4) \quad v_{S_1} = \sqrt{\frac{(C_{44} + C_{55}) + \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2\rho}},$$

and

$$v_{S_2} = \sqrt{\frac{(C_{44} + C_{55}) - \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2\rho}}.$$

Also, consider equation (9.2.3). 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 (9.2.3), 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 (9.2.2), these vectors are contained in the x_1x_2 -plane. Inserting eigenvalue (9.2.4) into equations (9.2.2), we obtain

$$(9.2.5) \quad \begin{bmatrix} \frac{C_{55} - C_{44} - \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2} & C_{45} \\ C_{45} & \frac{C_{44} - C_{55} - \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In view of a three-dimensional continuum, we can write the nontrivial solution of system (9.2.5) as the displacement vector given by

$$(9.2.6) \quad \mathbf{A}_{S_1} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = b \begin{bmatrix} \frac{C_{55} - C_{44} + \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2C_{45}} \\ 1 \\ 0 \end{bmatrix},$$

where b is a nonzero constant. Hence, the angle that this vector makes with a coordinate axis in the x_1x_2 -plane is

$$(9.2.7) \quad \tan \Theta = \frac{A_1}{A_2} = \frac{C_{55} - C_{44} + \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2C_{45}}.$$

We can also find, in an analogous manner, the displacement vector that corresponds to the other S wave. It is given by

$$\mathbf{A}_{S_2} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = c \begin{bmatrix} -1 \\ \frac{C_{55} - C_{44} + \sqrt{(C_{44} - C_{55})^2 + 4C_{45}^2}}{2C_{45}} \\ 0 \end{bmatrix},$$

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 9.1.2, 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.⁶ Such directions are called the pure-mode directions. As illustrated herein, symmetry axes are pure-mode directions.

Natural coordinate systems. In Section 5.6.3, 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 9.2.1. *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 9.2.1, consider expression (9.2.7). Invoking the trigonometric identity given by

$$\tan(2\Theta) = \frac{2 \tan \Theta}{1 - \tan^2 \Theta},$$

we can restate expression (9.2.7) as

$$(9.2.8) \quad \tan(2\Theta) = \frac{2C_{45}}{C_{44} - C_{55}}.$$

⁶Readers interested in further description and additional references might refer to Helbig, K., (1994) Foundations of anisotropy for exploration seismics: Pergamon, p. 166.

Expression (9.2.8) is precisely expression (5.6.6), which allows us to express elasticity matrix (5.6.4) in a natural coordinate system to obtain matrix (5.6.7). To further illustrate this result, we notice that, using elasticity matrix (5.6.7), equations (9.2.1) become

$$(9.2.9) \quad \begin{bmatrix} \hat{C}_{55} - \rho v^2 & 0 & 0 \\ 0 & \hat{C}_{44} - \rho v^2 & 0 \\ 0 & 0 & \hat{C}_{33} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

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 (9.2.9), is the diagonal form of the square matrix shown in equation (9.2.2). In terms of a natural coordinate system, such a diagonalization is also obtained using equation (5.6.8), 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 (5.6.4), while \hat{C}_{44} and \hat{C}_{55} are entries of matrix (5.6.7). Also, examining systems (9.2.1) and (9.2.9), 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.

9.2.3. 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 (5.10.3). For notational convenience, letting

$$(9.2.10) \quad \frac{C_{11} - C_{12}}{2} = C_{66},$$

in matrix (5.10.3), we can write the entries of matrix Γ , given by expressions (9.1.6), as

$$\begin{aligned}
 \Gamma_{11} &= n_1^2 C_{11} + n_2^2 C_{66} + n_3^2 C_{44}, \\
 \Gamma_{22} &= n_1^2 C_{66} + n_2^2 C_{11} + n_3^2 C_{44}, \\
 \Gamma_{33} &= (n_1^2 + n_2^2) C_{44} + n_3^2 C_{33}, \\
 \Gamma_{12} &= \Gamma_{21} = n_1 n_2 (C_{11} - C_{66}), \\
 \Gamma_{13} &= \Gamma_{31} = n_1 n_3 (C_{13} + C_{44}), \\
 \Gamma_{23} &= \Gamma_{32} = n_2 n_3 (C_{13} + C_{44}).
 \end{aligned}
 \tag{9.2.11}$$

Note that, in view of expression (9.2.10), $C_{12} = C_{11} - 2C_{66}$. Thus, we could also write $\Gamma_{12} = \Gamma_{21} = n_1 n_2 (C_{12} + C_{66})$, which is consistent with the pattern of the last two lines of set (9.2.11). 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 (9.1.7) with entries (9.2.11). 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 (9.2.11) 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 (9.1.7) becomes

$$\begin{bmatrix} C_{11} - \rho v^2 & 0 & 0 \\ 0 & C_{66} - \rho v^2 & 0 \\ 0 & 0 & C_{44} - \rho v^2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.
 \tag{9.2.12}$$

By examining system (9.2.12), 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,

$$(9.2.13) \quad v_P = \pm \sqrt{\frac{C_{11}}{\rho}}.$$

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,

$$(9.2.14) \quad v_{SH} = \pm \sqrt{\frac{C_{66}}{\rho}}.$$

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,

$$(9.2.15) \quad v_{SV} = \pm \sqrt{\frac{C_{44}}{\rho}}.$$

Expressions (9.2.13), (9.2.14) and (9.2.15) 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 (9.2.11), we can write expression (9.1.8) as

$$\begin{aligned} \det [\mathbf{\Gamma} - \rho v^2 \mathbf{I}] &= [C_{66} (n_1^2 + n_2^2) + C_{44} n_3^2 - \rho v^2] \\ &\quad \{-C_{13}^2 (n_1^2 + n_2^2) n_3^2 - 2C_{13}C_{44} (n_1^2 + n_2^2) n_3^2 \\ &\quad + C_{33}C_{44}n_3^4 - C_{44} (n_1^2 + n_2^2) \rho v^2 - C_{33}n_3^2 \rho v^2 - C_{44}n_3^2 \rho v^2 \\ &\quad + C_{11} (n_1^2 + n_2^2) [C_{44} (n_1^2 + n_2^2) + C_{33}n_3^2 - \rho v^2] + \rho^2 v^4\}. \end{aligned}$$

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,

$$(9.2.16) \quad \begin{aligned} \det [\mathbf{\Gamma} - \rho v^2 \mathbf{I}] &= [C_{66} (1 - n_3^2) + C_{44} n_3^2 - \rho v^2] \\ &\quad \left\{ [C_{33}C_{44}n_3^4 - [2C_{13}C_{44} - C_{11}C_{33} + C_{13}^2] n_3^2 (1 - n_3^2) + C_{11}C_{44} (1 - n_3^2)^2] \right. \\ &\quad \left. + [(C_{11} - C_{33}) n_3^2 - (C_{11} + C_{44})] \rho v^2 + \rho^2 v^4 \right\}. \end{aligned}$$

Note that determinant (9.2.16) 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 (9.1.8) and, hence, setting expression (9.2.16) to zero, we immediately obtain the equation to be solved for the three velocities.

Solving the quadratic equation, shown in brackets in expression (9.2.16), and considering only the positive root, we obtain

$$(9.2.17) \quad v(\mathbf{n}) = \sqrt{\frac{C_{66}(1 - n_3^2) + C_{44}n_3^2}{\rho}}.$$

Setting $n_3 = 0$ and comparing to expressions (9.2.13), (9.2.14) and (9.2.15), we recognize expression (9.2.17) as corresponding to expression (9.2.14). Thus, we denote it as

$$(9.2.18) \quad v_{SH}(\mathbf{n}) = \sqrt{\frac{C_{66}(1 - n_3^2) + C_{44}n_3^2}{\rho}}.$$

Solving the biquadratic equation, shown in braces in expression (9.2.16), we obtain two solutions. Again, setting $n_3 = 0$, we recognize them as corresponding to expressions (9.2.13) and (9.2.15). 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

$$(9.2.19) \quad v_{qP}(\mathbf{n}) = \sqrt{\frac{(C_{33} - C_{11})n_3^2 + C_{11} + C_{44} + \sqrt{\Delta}}{2\rho}}$$

and

$$(9.2.20) \quad v_{qSV}(\mathbf{n}) = \sqrt{\frac{(C_{33} - C_{11})n_3^2 + C_{11} + C_{44} - \sqrt{\Delta}}{2\rho}},$$

where the discriminant, Δ , is

$$(9.2.21) \quad \Delta \equiv [(C_{11} - C_{33})n_3^2 - C_{11} - C_{44}]^2 - 4[C_{33}C_{44}n_3^4 - [2C_{13}C_{44} - C_{11}C_{33} + C_{13}^2]n_3^2(1 - n_3^2) + C_{11}C_{44}(1 - n_3^2)^2].$$

Note that since the n_3 component can be written as

$$(9.2.22) \quad n_3 = \cos \vartheta,$$

where ϑ is the phase angle, velocity expressions (9.2.18), (9.2.19) and (9.2.20) can be immediately stated in terms of the phase angle.

Let us examine expression (9.2.21). To do so, we rewrite it as

$$(9.2.23) \quad \Delta \equiv [(C_{11} - C_{44})(1 - n_3^2) - (C_{33} - C_{44})n_3^2]^2 + 4(C_{44} + C_{13})^2 n_3^2 (1 - n_3^2).$$

We see that Δ is a nonnegative quantity; hence, expressions (9.2.19) and (9.2.20) are real, as expected in view of the symmetry of matrix Γ . Expressions (9.2.19) and (9.2.20) are distinct from one another if $\Delta \neq 0$. Examining expression (9.2.23), we see that $\Delta = 0$ for the following cases: $n_3 = 0$ and $C_{11} = C_{44}$, $n_3 = \pm 1$ and $C_{33} = C_{44}$, as well as particular values of n_3 if $C_{44} = -C_{13}$. Since these equalities lead to peculiar results, which cannot be viewed as common in the idealized representation of continuum mechanics, we will assume that $C_{11} \neq C_{44}$, $C_{33} \neq C_{44}$, $C_{44} \neq -C_{13}$. From now on, we will consider Δ as a strictly positive quantity for all values of n_3 , except in a brief discussion at the end of Section 9.3.3.

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 (9.2.11), we can write the coefficient matrix of system (9.1.7) as

$$(9.2.24) \quad \begin{bmatrix} n_1^2 C_{11} + n_3^2 C_{44} - \rho v^2 & 0 & n_1 n_3 (C_{13} + C_{44}) \\ 0 & n_1^2 C_{66} + n_3^2 C_{44} - \rho v^2 & 0 \\ n_1 n_3 (C_{13} + C_{44}) & 0 & n_1^2 C_{44} + n_3^2 C_{33} - \rho v^2 \end{bmatrix}.$$

Considering equations (9.1.7) and in view of the coefficient matrix (9.2.24), we see that the second equation is not coupled with the remaining two. Hence, we can rewrite system (9.1.7) as

$$(9.2.25) \quad [n_1^2 C_{66} + n_3^2 C_{44} - \rho v^2(\mathbf{n})] A_2 = 0,$$

and

$$(9.2.26) \quad \begin{bmatrix} n_1^2 C_{11} + n_3^2 C_{44} - \rho v^2(\mathbf{n}) & n_1 n_3 (C_{13} + C_{44}) \\ n_1 n_3 (C_{13} + C_{44}) & n_1^2 C_{44} + n_3^2 C_{33} - \rho v^2(\mathbf{n}) \end{bmatrix} \begin{bmatrix} A_1 \\ A_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

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 (9.2.25) is parallel to the x_2 -axis, while the displacement vectors associated with equations (9.2.26) 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 (9.2.25). 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 (9.2.18), which can be written as

$$(9.2.27) \quad \rho v_{SH}^2(\mathbf{n}) = n_1^2 C_{66} + n_3^2 C_{44},$$

corresponds to the SH wave. Inserting expression (9.2.27) into equation (9.2.25), 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 (9.2.25). 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 (9.2.26) 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 (9.2.26) 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

$$(9.2.28) \quad \phi = \arctan \frac{A_1}{A_3}.$$

Using the second equation of system (9.2.26), we obtain

$$(9.2.29) \quad \frac{A_1}{A_3} = \frac{\rho v^2(\mathbf{n}) - n_1^2 C_{44} - n_3^2 C_{33}}{n_1 n_3 (C_{13} + C_{44})}.$$

Note that the same value of the displacement angle is obtained if we use the first equation of system (9.2.26), as illustrated in Exercise 9.5.

Since $\mathbf{n}^2 = n_1^2 + n_3^2 = 1$ and n_3 is given by expression (9.2.22), we can write expression (9.2.28) as

$$(9.2.30) \quad \phi = \arctan \frac{\rho v^2(\vartheta) - C_{44} \sin^2 \vartheta - C_{33} \cos^2 \vartheta}{(C_{13} + C_{44}) \sin \vartheta \cos \vartheta},$$

where $v(\vartheta)$ is given by expressions (9.2.19) or (9.2.20), together with expression (9.2.22), 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 (9.2.19) and (9.2.22) into expression (9.2.30). If we wish to find the displacement direction associated with the qSV wave, we insert expressions (9.2.20) and (9.2.22) into expression (9.2.30).

Examining expression (9.2.30), 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 9.4.1. 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 (9.2.29). In other words, we find the orientation but not the magnitude of the displacement vectors.

9.3. Phase-slowness surfaces

9.3.1. 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 10. 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.

9.3.2. Convexity of innermost sheet. For the elastic continua, the phase-slowness surface — which, for the transversely isotropic case, results from the bicubic equation given by expression (9.2.16) — 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.⁷

THEOREM 9.3.1. *In elastic continua, if the innermost sheet of the phase-slowness surface is detached, it is convex.*

There are particular cases in which the innermost phase-slowness sheet is not detached, such as the case of $C_{44} = -C_{13}$, mentioned above.⁸ Nevertheless, Theorem 9.3.1 is still valid, because the detachment is not necessary for the validity of the theorem. In other words, in elastic continua, the innermost sheet of the phase-slowness surface is convex.⁹

⁷Interested 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.

⁸For another example of a nondetached innermost sheet, readers might refer to Auld, B.A., (1973) *Acoustic fields and waves in solids*: John Wiley and Sons, Vol. I, p. 406.

⁹Interested readers might refer to Bucataru and Slawinski (2008) *in progress*

9.3.3. Intersection points. 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 sheets.

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 (9.2.18), (9.2.19) and (9.2.20), and using expression (9.2.22), the corresponding phase-slowness curves can be generated as a polar plot with the radius given by the reciprocal of the phase-velocity magnitude.

Note that the intersection points of the phase-slowness curves in the x_1x_3 -plane correspond to intersection lines of the phase-slowness sheets 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 (9.2.16). 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 (9.2.18) — into the biquadratic part of equation (9.2.16) — shown in braces — and simplifying, we obtain

$$(n_3^2 - 1) \left\{ (C_{66} - C_{11})(C_{44} - C_{66}) + \left[(C_{13} + C_{44})^2 - (C_{11} - C_{66})(C_{33} - 2C_{44} + C_{66}) \right] n_3^2 \right\} = 0,$$

which is an expression of the form

$$(9.3.1) \quad (n_3^2 - 1) (A + Bn_3^2) = 0,$$

where

$$A := (C_{66} - C_{11})(C_{44} - C_{66}),$$

and

$$B := (C_{13} + C_{44})^2 - (C_{11} - C_{66})(C_{33} - 2C_{44} + C_{66}).$$

Hence, immediate solutions of equation (9.3.1) are given by

$$n_3 = \pm 1,$$

which correspond to the propagation along the rotation-symmetry axis. Setting $n_3 = \pm 1$ in expressions (9.2.18) and (9.2.20), we can verify that these are the velocities of SH and qSV that are equal to one another.

The remaining solutions of equation (9.3.1) 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

$$(9.3.2) \quad n_3 = \pm \sqrt{\frac{(C_{11} - C_{66})(C_{44} - C_{66})}{(C_{13} + C_{44})^2 - (C_{11} - C_{66})(C_{33} - 2C_{44} + C_{66})}}.$$

- 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 (9.3.2) is of particular interest, because, in connection with expression (9.2.22), namely, $n_3 = \cos \vartheta$, it gives the value of the phase angle at which the intersection points occur, as shown in Exercise 9.2. 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 9.4.1, which follows Exercise 9.1.

In above derivations, we have excluded three particular cases, as stated on page 279. Let us comment on the most interesting one: $C_{44} = -C_{13}$. If $C_{44} = -C_{13}$, expressions (9.2.19) and (9.2.20) become

$$(9.3.3) \quad v_{qP}(\mathbf{n}) = \sqrt{\frac{(C_{33} - C_{11})n_3^2 + C_{11} + C_{44} + |(C_{11} - C_{44})(1 - n_3^2) - (C_{33} - C_{44})n_3^2|}{2\rho}}$$

and

$$(9.3.4) \quad v_{qSV}(\mathbf{n}) = \sqrt{\frac{(C_{33} - C_{11})n_3^2 + C_{11} + C_{44} - |(C_{11} - C_{44})(1 - n_3^2) - (C_{33} - C_{44})n_3^2|}{2\rho}}.$$

Expressions (9.3.3) and (9.3.4) are equal to one another for particular values of n_3 . To find these values, we set the right-hand side of expression (9.2.23) to zero, and get

$$(9.3.5) \quad n_3 = \pm \sqrt{\frac{C_{11} - C_{44}}{C_{11} + C_{33} - 2C_{44}}}.$$

At points corresponding to the values given by expression (9.3.5), the sheets of the phase-slowness surfaces that correspond to the qP and qSV waves are not smooth, due to the change of sign within the absolute values in expressions (9.3.3) and (9.3.4). Also, since n_3 is a component of a unit vector, we require that $|n_3| \in (0, 1)$, and thus expression (9.3.5) imposes constraints on elasticity parameters. To investigate these constraints, we rewrite this expression as

$$n_3 = \pm \frac{1}{\sqrt{1 + \frac{C_{33} - C_{44}}{C_{11} - C_{44}}}};$$

thus, we require the fraction in the denominator to be positive. Hence, we must assume that C_{33} and C_{11} are both greater, or both smaller, than C_{44} . A transversely isotropic continuum with these constraints

has the innermost sheet of the phase-slowness surface that is not detached from the other sheets. The displacement direction associated with this sheet means that the sheet corresponds to the qP wave. This correspondence is expected in view of expression (9.3.3) whose value is greater than the one of expression (9.3.4), for all n_3 except for the points given by expression (9.3.5) — the qP wave is the fastest one.¹⁰

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, as well as the ray angles and the displacement angles, discussed herein. As illustrated in Exercise 9.11, all three angles are related by analytic expressions. However, each angle plays a distinct role in theoretical formulations and the analysis of experimental measurements.



¹⁰Readers interested in the formulation of the displacement vectors for the waves in the transversely isotropic continuum discussed herein might refer to Bucataru and Slawinski (2008) *in progress*

9.4. Exercises

EXERCISE 9.1. ¹¹Formulate and solve equation (9.1.8) for isotropic continua.

SOLUTION 9.1. 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 (5.12.3) and following entries (9.1.6), we can write equation (9.1.8) as

$$(9.4.1) \quad \det \begin{bmatrix} \mu - \rho v^2 & 0 & 0 \\ 0 & \mu - \rho v^2 & 0 \\ 0 & 0 & \lambda + 2\mu - \rho v^2 \end{bmatrix} = 0,$$

to obtain

$$(9.4.2) \quad (\rho v^2 - \mu)^2 [\rho v^2 - (\lambda + 2\mu)] = 0.$$

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 (6.1.17) and (6.1.13), respectively.

REMARK 9.4.1. The first two solutions in Exercise 9.1 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 (9.4.2), 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 9.1 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 9.2. Given the values of the elasticity parameters of the Green-river shale¹², namely,

$$(9.4.3) \quad \begin{cases} C_{11} = 3.13 \times 10^{10} \text{ N/m}^2 \\ C_{13} = 0.34 \times 10^{10} \text{ N/m}^2 \\ C_{33} = 2.25 \times 10^{10} \text{ N/m}^2 \\ C_{44} = 0.65 \times 10^{10} \text{ N/m}^2 \\ C_{66} = 0.88 \times 10^{10} \text{ N/m}^2 \end{cases},$$

find the intersection points for the SH and qSV waves.

SOLUTION 9.2. Since $C_{44} \neq -C_{13}$, we can use expression (9.3.2), which refers to the SH and qSV waves only — herein, the qP wave is detached, as discussed on page 279. Combining expression (9.2.22)

¹¹See also Section 9.3.3

¹²These values are taken from Thomsen, L., (1986) Weak elastic anisotropy: Geophysics, **51**, 1954 – 1966.

and (9.3.2), we obtain

$$\vartheta = \arccos \sqrt{\frac{(C_{11} - C_{66})(C_{44} - C_{66})}{(C_{13} + C_{44})^2 - (C_{11} - C_{66})(C_{33} - 2C_{44} + C_{66})}} \approx 66^\circ.$$

The intersection points the SH and qSV waves occur at $\vartheta \approx 66^\circ$.

EXERCISE 9.3. *In view of Section 9.3.3, show that for isotropic continua, SH -wave velocity and SV -wave velocity coincide with one another for all directions.*

SOLUTION 9.3. As shown in elasticity matrix (5.12.3), for an isotropic continuum, we have

$$C_{11} = C_{22} = C_{33} = \lambda + 2\mu,$$

$$C_{13} = \lambda,$$

$$C_{44} = C_{66} = \mu,$$

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 9.3.3, if $A = B = 0$, the phase-slowness curves coincide for all directions.

EXERCISE 9.4. *Using expression (9.2.30), namely,*

$$(9.4.4) \quad \phi = \arctan \frac{\rho v^2(\vartheta) - C_{44} \sin^2 \vartheta - C_{33} \cos^2 \vartheta}{(C_{13} + C_{44}) \sin \vartheta \cos \vartheta},$$

show that, for P waves in isotropic continua, the phase angle, ϑ , and the displacement angle, ϕ , coincide.

SOLUTION 9.4. Considering the elasticity matrix for an isotropic continuum, namely, matrix (5.12.1), we see that

$$C_{13} = C_{11} - 2C_{44},$$

and

$$C_{11} = C_{33}.$$

Considering elasticity matrix (5.12.3) and expression (6.1.13), we can express the velocity of a P wave in an isotropic continuum as

$$v_P = \sqrt{\frac{C_{33}}{\rho}}.$$

Hence, expression (9.4.4) 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 9.5. *Expression (9.4.4) is obtained using the second equation of system (9.2.26). Verify that using the first equation of this system to obtain A_1/A_3 , we get the same result as shown in Exercise 9.4.*

SOLUTION 9.5. Using the first equation of system (9.2.26), we can write expression (9.2.28) as

$$(9.4.5) \quad \phi = \arctan \frac{A_1}{A_3} = \arctan \frac{(C_{13} + C_{44}) \sin \vartheta \cos \vartheta}{\rho v^2 (\vartheta) - C_{11} \sin^2 \vartheta - C_{44} \cos^2 \vartheta}.$$

In view of the isotropic-case expressions, stated in Exercise 9.4, we can rewrite expression (9.4.5) 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 9.6. *Using expression (9.4.4), 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 9.6. Considering the elasticity matrix for an isotropic continuum, namely, matrix (5.12.1), we see that

$$C_{13} = C_{11} - 2C_{44},$$

and

$$C_{11} = C_{33}.$$

Considering elasticity matrix (5.12.3) and expression (6.1.17), 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 9.4, expression (9.4.4) becomes

$$\phi = \arctan (-\cot \vartheta) = -\arctan (\cot \vartheta).$$

Using properties of the inverse trigonometric functions, we can rewrite this expression as

$$\phi = \arctan (\tan \vartheta) - \frac{\pi}{2} = \vartheta - \frac{\pi}{2},$$

as required and as expected from our discussion in Section 6.2.

EXERCISE 9.7. Using determinant (9.2.16) obtain expressions (9.2.13), (9.2.14) and (9.2.15).

SOLUTION 9.7. Consider the determinantal expression (9.2.16), namely,

$$\begin{aligned} \det [\mathbf{\Gamma} - \rho(\mathbf{x}) v^2 \mathbf{I}] &= [C_{66} (1 - n_3^2) + C_{44} n_3^2 - \rho v^2] \\ &\quad \left\{ [C_{33} C_{44} n_3^4 - [2C_{13} C_{44} - C_{11} C_{33} + C_{13}^2] n_3^2 (1 - n_3^2) + C_{11} C_{44} (1 - n_3^2)^2] \right. \\ &\quad \left. + [(C_{11} - C_{33}) n_3^2 - (C_{11} + C_{44})] \rho v^2 + \rho^2 v^4 \right\}. \end{aligned}$$

To consider propagation in the plane of transverse isotropy, we let $n_3 = 0$ to obtain

$$(9.4.6) \quad \det [\mathbf{\Gamma} - \rho(\mathbf{x}) v^2 \mathbf{I}] = (C_{66} - \rho v^2) [\rho^2 v^4 - (C_{11} + C_{44}) \rho v^2 + C_{11} C_{44}].$$

Setting expression (9.4.6) to zero, we obtain expressions (9.2.13), (9.2.14) and (9.2.15), as required.

EXERCISE 9.8. ¹³Show that *SH waves in transversely isotropic continua exhibit elliptical velocity dependence*.

SOLUTION 9.8. Consider expression (9.2.18). Recalling expression (9.2.22) and using trigonometric identities, we can write

$$v_{SH}(\vartheta) = \sqrt{\frac{C_{66}}{\rho} \sin^2 \vartheta + \frac{C_{44}}{\rho} \cos^2 \vartheta}.$$

Setting $\vartheta = 0$, we get $v_{SH}(0) = \sqrt{C_{44}/\rho}$, while setting $\vartheta = \pi/2$, we get $v_{SH}(\pi/2) = \sqrt{C_{66}/\rho}$, which can be denoted as v_z and v_x , respectively. Thus, we can write

$$v_{SH}(\vartheta) = \sqrt{v_x^2 \sin^2 \vartheta + v_z^2 \cos^2 \vartheta},$$

which is expression (6.10.11), giving the magnitude of phase velocity for the case of elliptical velocity dependence.

EXERCISE 9.9. Formulate Hamiltonian \mathcal{H} that corresponds to *SH waves in a transversely isotropic continuum*.

SOLUTION 9.9. In view of expression (8.2.8) and considering a given point of the continuum, we can write the corresponding ray-theory Hamiltonian as

$$\mathcal{H}(\mathbf{p}) = \frac{1}{2} p^2 v^2(\mathbf{p}).$$

Considering the *SH*-wave velocity given by expression (9.2.18), namely,

$$v_{SH}^2(\mathbf{n}) = \frac{C_{66} (1 - n_3^2) + C_{44} n_3^2}{\rho},$$

¹³See also Section 6.10.3

and since $n_i^2 = p_i^2/p^2$ and $n_1^2 = 1 - n_3^2$, we can write

$$v_{SH}^2(\mathbf{p}) = \frac{C_{66} \frac{p_1^2}{p^2} + C_{44} \frac{p_3^2}{p^2}}{\rho} = \frac{1}{p^2} \frac{C_{66} p_1^2 + C_{44} p_3^2}{\rho}.$$

Hence, we can write

$$(9.4.7) \quad \mathcal{H}_{SH}(\mathbf{p}) = \frac{1}{2} \frac{C_{66} p_1^2 + C_{44} p_3^2}{\rho}.$$

EXERCISE 9.10. *Using Legendre's transformation and expression (9.4.7), find the corresponding relation between the phase and the ray angles for SH waves in a transversely isotropic continuum.*

SOLUTION 9.10. As shown in expression (8.7.4), the ray angle is given by

$$\tan \theta = \frac{\frac{dx_1}{ds}}{\frac{dx_3}{ds}}.$$

Using time parametrization, we can immediately restate this expression as

$$\tan \theta = \frac{\frac{dx_1}{dt}}{\frac{dx_3}{dt}} \equiv \frac{\dot{x}_1}{\dot{x}_3},$$

where t denotes time. In view of transformation (B.3.7), we can write

$$\dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i}.$$

Thus, using expression (9.4.7), 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 (8.4.5), we can restate the above expression in terms of the phase angle, as

$$(9.4.8) \quad \tan \theta = \frac{C_{66}}{C_{44}} \tan \vartheta.$$

REMARK 9.4.2. Expression (9.4.8) 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 9.11. *Using expression (9.2.30), namely,*

$$(9.4.9) \quad \phi = \arctan \frac{\rho v^2(\vartheta) - C_{44} \sin^2 \vartheta - C_{33} \cos^2 \vartheta}{(C_{13} + C_{44}) \sin \vartheta \cos \vartheta},$$

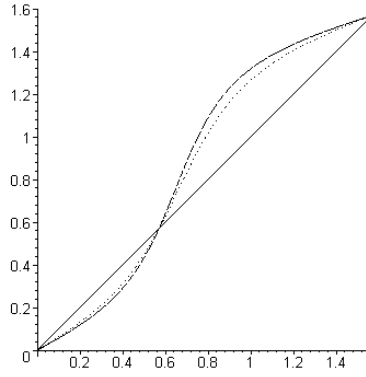


FIGURE 9.4.1. Ray, phase and displacement angles: The ray angle (dashed line) and the displacement angle (dotted line) are plotted as functions of the phase angle. The units of both axes are displayed in radians. This figure is the solution of Exercise 9.11.

and expression (8.4.12), which can be rewritten as

$$(9.4.10) \quad \theta = \arctan \frac{\tan \vartheta + \frac{1}{v(\vartheta)} \frac{\partial v(\vartheta)}{\partial \vartheta}}{1 - \frac{\tan \vartheta}{v(\vartheta)} \frac{\partial v(\vartheta)}{\partial \vartheta}},$$

as well as the elasticity parameters of the Green-river shale, shown in expressions (9.4.3), and its mass density, given by $\rho = 2310 \text{ kg/m}^3$, plot the displacement angles, ϕ , and the ray angle, θ , as a function of the phase angle, ϑ , for qP waves.

SOLUTION 9.11. Inserting phase-velocity expression (9.2.19) and expression (9.2.22), into expressions (9.4.9) and (9.4.10), we generate the plot of the displacement and the ray angles, respectively. This plot is shown in Figure 9.4.1.

REMARK 9.4.3. Figure 9.4.1 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 (9.4.9) and (9.4.10) — at the phase angle satisfying equation

$$\frac{\rho v_{qP}^2(\vartheta) - C_{44} \sin^2 \vartheta - C_{33} \cos^2 \vartheta}{(C_{13} + C_{44}) \sin \vartheta \cos \vartheta} = \frac{\tan \vartheta + \frac{1}{v(\vartheta)} \frac{\partial v(\vartheta)}{\partial \vartheta}}{1 - \frac{\tan \vartheta}{v(\vartheta)} \frac{\partial v(\vartheta)}{\partial \vartheta}}.$$

Examining Figure 9.4.1, 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.¹⁴

¹⁴Readers interested in relations among the phase angle, the ray angle and the displacement angle might also refer to Tsavkin, I., (2001) Seismic signatures and analysis of reflection data in anisotropic media: Pergamon, pp. 34 – 36.

CHAPTER 10

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.¹

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.

10.1. Angles at interface

10.1.1. 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.

¹Readers interested in philosophical implications of this statement might refer to Steiner, M., (1998) The applicability of mathematics as a philosophical problem: Harvard University Press, pp. 58 – 59 and pp. 105 – 114.

Recall Hamilton's ray equations (8.2.7), namely,

$$(10.1.1) \quad \begin{cases} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\},$$

where Hamiltonian \mathcal{H} is given by expression (8.2.8), namely,

$$(10.1.2) \quad \mathcal{H} = \frac{1}{2} p^2 v^2(\mathbf{x}, \mathbf{p}).$$

Examining equations (10.1.1) and expression (10.1.2), 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

$$\dot{p}_i \equiv \frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial x_i} = 0, \quad i \in \{1, 2\}.$$

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 (9.2.22) 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 \mathbf{p} . Now, since the magnitudes of phase slowness and phase velocity are the reciprocals of one another, we can write conserved quantity \mathbf{p} as

$$(10.1.3) \quad \mathbf{p} = \frac{\sin \vartheta}{v(\vartheta)},$$

where $v(\vartheta)$ gives the magnitude of phase velocity as a function of the phase angle.

Expression (10.1.3) is a general statement of Snell's law² in the context of phase angle and phase velocity. It is valid across interfaces between generally anisotropic continua. Since \mathbf{p} is a conserved quantity for a given solution curve $\mathbf{x}(t)$, which corresponds to a ray, we refer to \mathbf{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.³

10.1.2. Ray angles. We wish to use the continuity of the phase-slowness components to derive the relation between the ray angles across the interface.

Examining Hamilton's ray equations, an analytic relation between the phase angles and the ray angles was derived in Section 8.4.3 and given by expression (8.4.12). As discussed on page 241, this expression also states that at any point of the phase-slowness curve, the corresponding ray direction is always normal to this 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 (10.1.3) is generally true for $\vartheta \in (-\pi, \pi)$, obtaining analytic 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.⁴ This is a consequence of restrictions imposed by Legendre's transformation, which is discussed in Appendix B. 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 analytic expressions between the ray angles of incidence and transmission.

²Readers interested in discussion of Snell's law in the context of quantum electrodynamics might refer to Feynman, R.P., (1985/2006) QED: The strange theory of light and matter: Princeton University Press, pp. 38 – 45, for reflection, and pp. 50 – 52, for refraction. On page 56 of that book, we read that

the idea that light goes in a straight line is a convenient approximation to describe what happens in the world that is familiar to us; it's similar to the crude approximation that says when light reflects off a mirror, the angle of incidence is equal to the angle of reflection.

Readers interested in an insightful discussion of Snell's law, which includes both a geometrical and mechanical formulations, might refer to Polya, G., (1954) Mathematics and plausible reasoning, Vol. I: Induction and analogy in mathematics: Princeton University Press, pp. 142 – 155. Readers interested in Snell's law as an example of causation and determinism might refer to Weinert, F., (2005) The scientist as philosopher: Springer-Verlag, pp. 208 – 209.

³Readers 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.

⁴Readers 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.

10.1.3. 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

$$(10.1.4) \quad \begin{cases} f(p_x, p_z) = ({}_1v_x p_x)^2 + ({}_1v_z p_z)^2 = 1 \\ g(p_x, p_z) = ({}_2v_x p_x)^2 + ({}_2v_z p_z)^2 = 1 \end{cases} ,$$

for the medium of incidence and transmission, respectively, where v_x and v_z specify the horizontal and vertical phase velocities, respectively.

Note that either expression of set (10.1.4) is the equation of an ellipse in the $p_x p_z$ -plane, given by

$$(10.1.5) \quad \frac{p_x^2}{\left(\frac{1}{m v_x}\right)^2} + \frac{p_z^2}{\left(\frac{1}{m v_z}\right)^2} = 1, \quad m \in \{1, 2\},$$

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 (10.1.3), we can write

$$(10.1.6) \quad \mathbf{p} = p_x = \frac{\sin \vartheta}{v(\vartheta)} = \frac{\sin \vartheta_m}{\sqrt{{}_m v_x^2 \sin^2 \vartheta_m + {}_m v_z^2 \cos^2 \vartheta_m}}, \quad m \in \{1, 2\},$$

where, in view of elliptical velocity dependence, $v(\vartheta)$ is given by expression (6.10.11).

Conserved quantity in terms of ray angles and ray velocities. We wish to express conserved quantity (10.1.6) 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 (10.1.4) can be restated as

$$(10.1.7) \quad \begin{cases} f(p_x, p_z) = ({}_1V_x p_x)^2 + ({}_1V_z p_z)^2 = 1 \\ g(p_x, p_z) = ({}_2V_x p_x)^2 + ({}_2V_z p_z)^2 = 1 \end{cases} .$$

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 (10.1.7), we get the corresponding expression for a ray angle in elliptical velocity dependence, namely,

$$(10.1.8) \quad \cos \theta_m = \frac{mV_z^2 p_z}{\sqrt{(mV_x^2 p_x)^2 + (mV_z^2 p_z)^2}}, \quad m \in \{1, 2\}.$$

To invoke the conserved quantity, $\mathbf{p} = p_x$, we would like to explicitly solve equations (10.1.8) for p_x .

Using expressions of set (10.1.7), we can write

$$(10.1.9) \quad p_z = \frac{\sqrt{1 - (mV_x p_x)^2}}{mV_z}, \quad m \in \{1, 2\},$$

and, hence, inserting expressions (10.1.9) into equations (10.1.8), we get

$$(10.1.10) \quad \cos \theta_m = \frac{mV_z \sqrt{1 - (mV_x p_x)^2}}{\sqrt{mV_x^4 p_x^2 + mV_z^2 [1 - (mV_x p_x)^2]}}, \quad m \in \{1, 2\}.$$

Solving equations (10.1.10) for p_x , we obtain

$$p_x^2 = \frac{mV_z^2 \sin^2 \theta_m}{mV_x^2 (mV_z^2 \sin^2 \theta_m + mV_x^2 \cos^2 \theta_m)}, \quad m \in \{1, 2\}.$$

Simplifying, we can write

$$p_x^2 = \frac{1}{mV_x^2 \left[\left(\frac{mV_x}{mV_z} \right)^2 \cot^2 \theta_m + 1 \right]}, \quad m \in \{1, 2\}.$$

Consequently, the conserved quantity, $\mathbf{p} = p_x$, can be written as

$$(10.1.11) \quad \mathbf{p} = \frac{1}{mV_x \sqrt{\left(\frac{mV_x}{mV_z} \right)^2 \cot^2 \theta_m + 1}}, \quad m \in \{1, 2\},$$

which is conserved quantity (10.1.6) stated in terms of ray angles and ray velocities.

Note that we can write expression (10.1.6) as

$$(10.1.12) \quad \mathbf{p} = \frac{1}{m v_x \sqrt{\left(\frac{m v_z}{m v_x} \right)^2 \cot^2 \vartheta_m + 1}}, \quad m \in \{1, 2\},$$

which allows us to see the similarity of form between expressions (10.1.6) and (10.1.11). Notice, however, that in expression (10.1.11), we have V_x/V_z , while, in expression (10.1.12), we have v_z/v_x .

In general, as shown in Exercise 10.1, expressions (10.1.6) and (10.1.11) are equivalent to one another. For the isotropic case, as shown in Exercise 10.2, expressions (10.1.6) and (10.1.11) become identical.

Following expression (10.1.11) and in view of set (10.1.7), we can write

$$(10.1.13) \quad {}_1V_x^2 \left[\left(\frac{{}_1V_x}{{}_1V_z} \right)^2 \cot^2 \theta_1 + 1 \right] = {}_2V_x^2 \left[\left(\frac{{}_2V_x}{{}_2V_z} \right)^2 \cot^2 \theta_2 + 1 \right],$$

where the subscripts 1 and 2 correspond to the medium of incidence and transmission, respectively. Equation (10.1.13) can be viewed as a statement of Snell's law for elliptical velocity dependence, expressed in terms of ray angles and ray velocities.

10.2. Amplitudes at interface

10.2.1. Kinematic and dynamic boundary conditions.

Introductory comments. In Section 10.1, 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

$$(10.2.1) \quad \mathbf{u} = [0, u_2, 0].$$

Considering the oscillatory nature of waves and in view of expression (6.10.14), we can write the nonzero component of displacement as

$$(10.2.2) \quad u_2 = A \exp [i\omega (\mathbf{p} \cdot \mathbf{x} - t)],$$

where A denotes the amplitude of the displacement and $\exp [\cdot]$ is the phase factor.

Kinematic boundary conditions. The kinematic boundary conditions require the equality of the sum of displacements across 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 (10.2.2) and setting the amplitude of the incident signal to unity, we can write the kinematic boundary condition as

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \exp [i\omega (\mathbf{p}^i \cdot \mathbf{x} - t)] + \begin{bmatrix} 0 \\ A_r \\ 0 \end{bmatrix} \exp [i\omega (\mathbf{p}^r \cdot \mathbf{x} - t)] = \begin{bmatrix} 0 \\ A_t \\ 0 \end{bmatrix} \exp [i\omega (\mathbf{p}^t \cdot \mathbf{x} - t)],$$

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

$$(10.2.3) \quad \exp [i\omega (\mathbf{p}^i \cdot \mathbf{x} - t)] + A_r \exp [i\omega (\mathbf{p}^r \cdot \mathbf{x} - t)] = A_t \exp [i\omega (\mathbf{p}^t \cdot \mathbf{x} - t)].$$

Dynamic boundary conditions. The dynamic boundary conditions require the equality of the traction components across the interface, which is tantamount to the equality of the stress-tensor components. We will deduce this equality from Newton's third law of motion. To do so, we require the welded contact in order for the adjacent points to remain in constant contact across the interface, and we view the continuum on either side of the interface as two distinct bodies acting on one another.⁵

Since the stress-tensor components are the components of the traction acting on a plane with a given orientation, let us recall the sign convention described on page 41.

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.

⁵Readers interested in an insightful explanation of the requirement of the constant contact between two bodies acting on one another might refer to Schutz, B., (2003) Gravity from the ground up: Cambridge University Press, pp. 11 – 12.

For convenience, let us consider an interface coinciding with the x_1x_2 -plane, and let traction $\mathbf{T}^{(\mathbf{n})}$ act on this plane in such a way that $T_1^{(\mathbf{n})}$, $T_2^{(\mathbf{n})}$ and $T_3^{(\mathbf{n})}$ are positive; \mathbf{n} stands for the outward normal of the plane, it points in the positive direction of the x_3 -axis. Thus, $T_1^{(\mathbf{n})}$, $T_2^{(\mathbf{n})}$ and $T_3^{(\mathbf{n})}$ point in the positive directions of the x_1 -axis, the x_2 -axis and the x_3 -axis, respectively. Newton's third law of motion implies that $T_1^{(-\mathbf{n})}$, $T_2^{(-\mathbf{n})}$ and $T_3^{(-\mathbf{n})}$ are equal in magnitude to their counterparts on the other side of the interface, and they point in the negative directions of the corresponding axes. Since $\mathbf{T}^{(-\mathbf{n})}$ acts on a plane whose outward normal points in the negative direction of the x_3 -axis, the three components of $\mathbf{T}^{(-\mathbf{n})}$ are positive. Hence, we can write $T_i^{(\mathbf{n})} = T_i^{(-\mathbf{n})}$. In other words, Newton's third law implies the equality of the traction components across the interface.

To express this result in terms of the stress-tensor components, let us recall the index convention described on page 41 and illustrated in Figure 2.5.1. Following this convention, both vectors $\mathbf{T}^{(\mathbf{n})}$ and $\mathbf{T}^{(-\mathbf{n})}$ are described by σ_{31} , σ_{32} and σ_{33} shown, respectively, by the solid arrows on the upper horizontal plane and by the dashed arrows on the lower horizontal plane in Figure 2.5.1. Since we would have reached the same conclusion of equality if we had chosen a different orientation of the interface or different signs of the traction components, we have shown that Newton's third law implies the equality of the stress-tensor components, which are the dynamic boundary conditions. Let us use these conditions to discuss the displacement amplitudes at a given interface, say, the x_1x_2 -plane.

For an interface coinciding with the x_1x_2 -plane, $n_1 = n_2 = 0$. Thus, in view of expression (2.5.15) on page 45, we see that the corresponding traction components are identically zero. Consequently, their equality is satisfied trivially. As shown above, the equality of the nonzero components of traction implies the equality of the stress-tensor components given by

$$(10.2.4) \quad \begin{bmatrix} \sigma_{31}^I \\ \sigma_{32}^I \\ \sigma_{33}^I \end{bmatrix} = \begin{bmatrix} \sigma_{31}^{II} \\ \sigma_{32}^{II} \\ \sigma_{33}^{II} \end{bmatrix},$$

where the superscript I indicates the medium of incidence and reflection, and the superscript II indicates the medium of transmission. Furthermore, in view of the symmetry of the stress tensor stated in Theorem 2.7.1 on page 52, we can write

$$\sigma_{13}^I = \sigma_{31}^I = \sigma_{13}^{II} = \sigma_{31}^{II}$$

and

$$\sigma_{23}^I = \sigma_{32}^I = \sigma_{23}^{II} = \sigma_{32}^{II}.$$

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 (10.2.4) in terms of elasticity parameters and mass density. Recalling definition (1.4.6) and considering stress-strain equations (4.2.8) with the elasticity

matrix for a transversely isotropic continuum, given by matrix (5.10.3), we can write

$$\sigma_{13} = C_{44} \left(\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} \right),$$

$$\sigma_{23} = C_{44} \left(\frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2} \right)$$

and

$$\sigma_{33} = C_{13} \frac{\partial u_1}{\partial x_1} + C_{13} \frac{\partial u_2}{\partial x_2} + C_{33} \frac{\partial u_3}{\partial x_3},$$

where — using the stress-tensor symmetry — we chose to write σ_{13} and σ_{23} rather than σ_{31} and σ_{32} . Considering displacement vector (10.2.1), where $u_1 = u_3 = 0$, we can write explicitly the three conditions as

$$(10.2.5) \quad \begin{aligned} \sigma_{13} &= 0, \\ \sigma_{23} &= C_{44} \frac{\partial u_2}{\partial x_3} \end{aligned}$$

and

$$\sigma_{33} = C_{13} \frac{\partial u_2}{\partial x_2}.$$

Since $\sigma_{13} = 0$, the first condition is satisfied identically, and the remaining dynamic boundary conditions are

$$(10.2.6) \quad \sigma_{23}^I = \sigma_{23}^{II}$$

and

$$(10.2.7) \quad \sigma_{33}^I = \sigma_{33}^{II}.$$

Considering the displacement-vector components for the incident, reflected and transmitted waves in view of expression (10.2.5), we can write boundary condition (10.2.6) as

$$(10.2.8) \quad C_{44}^I \left(\frac{\partial u_2^i}{\partial x_3} + \frac{\partial u_2^r}{\partial x_3} \right) = C_{44}^{II} \frac{\partial u_2^t}{\partial x_3},$$

where the left-hand side contains the contributions of both the incident and reflected waves. Invoking expression (10.2.2), we can write equation (10.2.8) as

$$(10.2.9) \quad C_{44}^I (\omega p_3^i \exp [i\omega (\mathbf{p}^i \cdot \mathbf{x} - t)] + \omega p_3^r A_r \exp [i\omega (\mathbf{p}^r \cdot \mathbf{x} - t)]) = \omega p_3^t C_{44}^{II} A_t \exp [i\omega (\mathbf{p}^t \cdot \mathbf{x} - t)].$$

In an analogous manner, we consider boundary condition (10.2.7) to obtain

$$C_{13}^I (\omega p_2^i \exp [i\omega (\mathbf{p}^i \cdot \mathbf{x} - t)] + \omega p_2^r A_r \exp [i\omega (\mathbf{p}^r \cdot \mathbf{x} - t)]) = \omega p_2^t C_{13}^{II} A_t \exp [i\omega (\mathbf{p}^t \cdot \mathbf{x} - t)].$$

Since the phase-slowness vector, \mathbf{p} , is contained in the x_1x_3 -plane, $p_2 = 0$. Hence, this condition is satisfied identically.

Thus, for SH waves propagating across the interface separating two transversely isotropic continua in welded contact, equation (10.2.3) is the only kinematic boundary condition and equation (10.2.9) is the only dynamic boundary condition. These equations form the system to be solved for the reflection and transmission amplitudes in terms of properties of continua on either side of the interface; notably, the mass density is implicitly present in these conditions since it is contained in expressions for the phase-slowness vectors, \mathbf{p} .

10.2.2. 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 (10.2.3) and (10.2.9). 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 (10.2.3) and (10.2.9) simplify to

$$(10.2.10) \quad 1 + A_r = A_t,$$

and

$$(10.2.11) \quad C_{44}^I (p_3^i + A_r p_3^r) = C_{44}^{II} p_3^t A_t,$$

respectively.

We can further simplify condition (10.2.11). 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 (10.2.11) becomes

$$(10.2.12) \quad C_{44}^I p_3^i (1 - A_r) = C_{44}^{II} p_3^t A_t.$$

Now, it is convenient to explicitly include mass density in condition (10.2.12). Since p_3 is a vertical component of the phase-slowness vector, recalling expression (9.2.18), we can write

$$(10.2.13) \quad C_{44}^I \frac{\cos \vartheta_i}{\sqrt{\frac{C_{66}^I \sin^2 \vartheta_i + C_{44}^I \cos^2 \vartheta_i}{\rho_1}}} (1 - A_r) = C_{44}^{II} \frac{\cos \vartheta_t}{\sqrt{\frac{C_{66}^{II} \sin^2 \vartheta_t + C_{44}^{II} \cos^2 \vartheta_t}{\rho_2}}} A_t.$$

Equations (10.2.10) and (10.2.13) form a system of two equations to be solved for the two unknowns, namely, the reflection and transmission amplitudes. These solutions are

$$(10.2.14) \quad A_r(\vartheta) = \frac{\frac{\sqrt{\rho_1} C_{44}^I \cos \vartheta_i}{\sqrt{C_{66}^I \sin^2 \vartheta_i + C_{44}^I \cos^2 \vartheta_i}} - \frac{\sqrt{\rho_2} C_{44}^{II} \cos \vartheta_t}{\sqrt{C_{66}^{II} \sin^2 \vartheta_t + C_{44}^{II} \cos^2 \vartheta_t}}}{\frac{\sqrt{\rho_1} C_{44}^I \cos \vartheta_i}{\sqrt{C_{66}^I \sin^2 \vartheta_i + C_{44}^I \cos^2 \vartheta_i}} + \frac{\sqrt{\rho_2} C_{44}^{II} \cos \vartheta_t}{\sqrt{C_{66}^{II} \sin^2 \vartheta_t + C_{44}^{II} \cos^2 \vartheta_t}}},$$

and

$$(10.2.15) \quad A_t(\vartheta) = \frac{2 \frac{\sqrt{\rho_1} C_{44}^I \cos \vartheta_i}{\sqrt{C_{66}^I \sin^2 \vartheta_i + C_{44}^I \cos^2 \vartheta_i}}}{\frac{\sqrt{\rho_1} C_{44}^I \cos \vartheta_i}{\sqrt{C_{66}^I \sin^2 \vartheta_i + C_{44}^I \cos^2 \vartheta_i}} + \frac{\sqrt{\rho_2} C_{44}^{II} \cos \vartheta_t}{\sqrt{C_{66}^{II} \sin^2 \vartheta_t + C_{44}^{II} \cos^2 \vartheta_t}}}.$$

Expressions (10.2.14) and (10.2.15) 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.⁶

Interpretation of expressions. Examining expressions (10.2.14) and (10.2.15), 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 (10.2.14) 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 10.5.

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 (10.1.6), 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 (10.2.14) and (10.2.15), we conclude that, in such a case, A_r and A_t are complex numbers.

⁶Readers interested in an insightful explanation of the reflection and transmission amplitudes — as functions of the properties of the media and the incidence angle — in the context of the theory of quantum electrodynamics might refer to Feynman, R.P., (1985/2006) QED: The strange theory of light and matter: Princeton University Press, pp. 107 – 109.

Let us consider the transmitted wave. Returning to expression (10.2.2), we can write it as

$$\begin{aligned}
 (10.2.16) \quad u_2^t &= A_t \exp [i\omega (|\mathbf{p}^t| \cos \vartheta_t z + |\mathbf{p}^t| \sin \vartheta_t x - t)] \\
 &= A_t \exp [i\omega (|\mathbf{p}^t| i |\cos \vartheta_t| z + |\mathbf{p}^t| \sin \vartheta_t x - t)] \\
 &= A_t \exp (-\omega |\mathbf{p}^t| |\cos \vartheta_t| z) \exp [i (|\mathbf{p}^t| \sin \vartheta_t x - t)].
 \end{aligned}$$

Expression (10.2.16) 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 10.6.

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

$$(10.2.17) \quad A_r(\vartheta) = |A_r| \exp(i\kappa),$$

where $|A_r|$ is the magnitude and κ is the angle in the complex plane. In view of expressions (10.2.2) and (10.2.17), as well as using the fact that $|A_r| = 1$, we can write the nonzero component of displacement of the reflected wave as

$$(10.2.18) \quad u_2^r = \exp(i\kappa) \exp[i\omega(\mathbf{p} \cdot \mathbf{x} - t)] = \exp\{i[\kappa + \omega(\mathbf{p} \cdot \mathbf{x} - t)]\},$$

where $\exp\{\cdot\}$ is the phase factor. Consequently, examining expression (10.2.18) and following the sign convention used for the phase factor in expression (10.2.2), we see that if $\kappa > 0$, the reflected wave is phase-delayed relative to the incident wave. This is the consequence of the fact that positive κ results in the phase factor being evaluated at an earlier time. In other words, $\exp[i(\kappa - \omega t)]$ lags $\exp(-i\omega t)$ in time. Similarly, if $\kappa < 0$, the reflected wave is phase-advanced.⁷

Expressions in terms of incidence phase angle. As shown in Section 10.1, the incidence and transmission angles can be expressed in terms of one another. Consequently, we wish to state expressions (10.2.14) and (10.2.15) in terms of the phase angle of incidence only.

Recall conserved quantity (10.1.3). Let the phase velocity be given by expression (9.2.18), and the phase angle be stated by expression (9.2.22). Thus, we can write

$$(10.2.19) \quad \frac{\sin \vartheta_i}{\sqrt{\frac{C_{66}^I \sin^2 \vartheta_i + C_{44}^I \cos^2 \vartheta_i}{\rho_1}}} = \frac{\sin \vartheta_t}{\sqrt{\frac{C_{66}^{II} \sin^2 \vartheta_t + C_{44}^{II} \cos^2 \vartheta_t}{\rho_2}}}.$$

Solving equation (10.2.19) for the angle of transmission, yields

$$(10.2.20) \quad \vartheta_t = \arcsin \sqrt{\frac{\rho_1 C_{44}^{II} \sin^2 \vartheta_i}{[\rho_2 (C_{66}^I - C_{44}^I) - \rho_1 (C_{66}^{II} - C_{44}^{II})] \sin^2 \vartheta_i + \rho_2 C_{44}^I}}.$$

⁷Readers interested in phase shifts might refer to Aki, K., and Richards, P.G., (2002) Quantitative seismology (2nd edition): University Science Books, pp. 149 – 157.

Hence, by inserting expression (10.2.20) into expressions (10.2.14) and (10.2.15), 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 (10.2.14) and (10.2.15) in terms of the ray angle of incidence, rather than the phase angle of incidence. Following equation (9.4.8), we can express the phase angle in terms of the ray angle as

$$(10.2.21) \quad \vartheta_i = \arctan \left(\frac{C_{44}^I}{C_{66}^I} \tan \theta_i \right).$$

Consequently, by inserting expression (10.2.21) into expression (10.2.20), and inserting the resulting expression into expressions (10.2.14) and (10.2.15), we can state the latter expressions in terms of the ray angle of incidence only.

Turning points. Herein, we will consider reflection amplitudes for a ray in a continuum where the velocity increases monotonically with depth. For this purpose — with no loss of generality — we can use the expression derived in Exercise 10.3, which corresponds to an isotropic case, namely,

$$A_r = \frac{\rho_1 v_1 \cos \vartheta_i - \rho_2 v_2 \cos \vartheta_t}{\rho_1 v_1 \cos \vartheta_i + \rho_2 v_2 \cos \vartheta_t}.$$

We can write

$$\cos \vartheta_i = \sqrt{1 - \sin^2 \vartheta_i}$$

and

$$\cos \vartheta_t = \sqrt{1 - \sin^2 \vartheta_t}$$

Invoking

$$\mathbf{p} = \frac{\sin \vartheta_i}{v_1} = \frac{\sin \vartheta_t}{v_2},$$

we obtain

$$A_r = \frac{\rho_1 v_1 \sqrt{1 - \mathbf{p}^2 v_1^2} - \rho_2 v_2 \sqrt{1 - \mathbf{p}^2 v_2^2}}{\rho_1 v_1 \sqrt{1 - \mathbf{p}^2 v_1^2} + \rho_2 v_2 \sqrt{1 - \mathbf{p}^2 v_2^2}}.$$

Let $\rho_1 = \rho - \Delta\rho$, $\rho_2 = \rho + \Delta\rho$, $v_1 = v - \Delta v$ and $v_2 = v + \Delta v$. We get

$$A_r = \frac{(\rho - \Delta\rho)(v - \Delta v) \sqrt{1 - \mathbf{p}^2 (v - \Delta v)^2}}{(\rho - \Delta\rho)(v - \Delta v) \sqrt{1 - \mathbf{p}^2 (v - \Delta v)^2} + (\rho + \Delta\rho)(v + \Delta v) \sqrt{1 - \mathbf{p}^2 (v + \Delta v)^2}} - \frac{(\rho + \Delta\rho)(v + \Delta v) \sqrt{1 - \mathbf{p}^2 (v + \Delta v)^2}}{(\rho - \Delta\rho)(v - \Delta v) \sqrt{1 - \mathbf{p}^2 (v - \Delta v)^2} + (\rho + \Delta\rho)(v + \Delta v) \sqrt{1 - \mathbf{p}^2 (v + \Delta v)^2}}.$$

Consider the root given by

$$\sqrt{1 - \mathbf{p}^2 (v - \Delta v)^2} = \sqrt{1 - \mathbf{p}^2 [v^2 - 2v\Delta v + (\Delta v)^2]}.$$

Assuming Δv is small, we ignore the second-order term and write

$$\sqrt{1 - \mathbf{p}^2 (v - \Delta v)^2} \approx \sqrt{1 - \mathbf{p}^2 (v^2 - 2v\Delta v)}.$$

As ϑ_i approaches $\pi/2$, \mathbf{p} tends to $1/v$. Thus, we can write

$$\sqrt{1 - \mathbf{p}^2 (v - \Delta v)^2} \approx \sqrt{1 - \frac{1}{v^2} [v^2 - 2v\Delta v]} = \sqrt{-2\frac{\Delta v}{v}}.$$

Similarly, consider the other root, we get

$$\sqrt{1 - \mathbf{p}^2 (v + \Delta v)^2} \approx \sqrt{2\frac{\Delta v}{v}}.$$

Neglecting linear terms in $\Delta\rho$ and Δv , we obtain

$$A_r \approx \frac{\sqrt{2\frac{\Delta v}{v}} - \sqrt{-2\frac{\Delta v}{v}}}{\sqrt{2\frac{\Delta v}{v}} + \sqrt{-2\frac{\Delta v}{v}}}.$$

Rearranging, we write

$$A_r \approx \frac{\sqrt{2\frac{\Delta v}{v}} - i\sqrt{2\frac{\Delta v}{v}}}{\sqrt{2\frac{\Delta v}{v}} + i\sqrt{2\frac{\Delta v}{v}}}.$$

Simplifying, we get

$$A_r \approx \frac{1 - i}{1 + i}.$$

Multiplying both the numerator and the denominator by the complex conjugate, we obtain $A_r \approx -i$.

To understand the physical meaning of this result, we can also write $A_r = \exp(-i\pi/2)$, and consider the phase factor, namely, $\exp(-i\omega t)$. The effect of A_r is

$$A_r \exp(-i\omega t) = \exp\left(-i\frac{\pi}{2}\right) \exp(-i\omega t) = \exp\left[-i\left(\frac{\pi}{2} + \omega t\right)\right],$$

which is a phase advance at the turning point.

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 9.2.3. 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.⁸

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 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.⁹



⁸Readers 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.

⁹Readers 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 (2nd 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.

10.3. Exercises

EXERCISE 10.1. ¹⁰Show that expressions (10.1.6) and (10.1.11) are equivalent to one another.

SOLUTION 10.1. Consider expression (10.1.11). In view of the symmetry of an ellipse, we know that $V_x = v_x$ and $V_z = v_z$. Hence, we can write

$$\mathbf{p} = \frac{1}{v_x \sqrt{\left(\frac{v_x}{v_z}\right)^2 \cot^2 \theta + 1}}.$$

Recalling expression (8.7.12), we express the ray angle in terms of the phase angle to obtain

$$\mathbf{p} = \frac{1}{v_x \sqrt{\left(\frac{v_x}{v_z}\right)^2 \frac{1}{\tan^2 \left\{ \arctan \left[\left(\frac{v_x}{v_z}\right)^2 \tan \vartheta \right] \right\}} + 1}}.$$

Using trigonometric identities, we get

$$\mathbf{p} = \frac{1}{v_x \sqrt{\left(\frac{v_x}{v_z}\right)^2 \frac{1}{\left(\frac{v_x}{v_z}\right)^4 \tan^2 \vartheta} + 1}} = \frac{1}{\sqrt{v_z^2 \cot^2 \vartheta + v_x^2}}.$$

Multiplying both numerator and denominator by $\sin \vartheta$, we obtain

$$\mathbf{p} = \frac{\sin \vartheta}{\sqrt{v_z^2 \cos^2 \vartheta + v_x^2 \sin^2 \vartheta}}.$$

In view of expression (6.10.11), we can immediately write

$$\mathbf{p} = \frac{\sin \vartheta}{v(\vartheta)},$$

which is expression (10.1.6), as required.

EXERCISE 10.2. Show that in isotropic continua, expressions (10.1.6) and (10.1.11) are identical to one another.

SOLUTION 10.2. Consider expression (10.1.11). In isotropic continua, $V := V_x = V_z$. Hence, we can write

$$\mathbf{p} = \frac{1}{V \sqrt{\cot^2 \theta + 1}}.$$

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

$$\mathbf{p} = \frac{\sin \vartheta}{v},$$

which is the isotropic form of expression (10.1.6), as required.

¹⁰Also see Section 14.6

EXERCISE 10.3. *Following expression (10.2.14), state the expressions for the reflection and transmission amplitudes for isotropic continua in terms of mass density, ρ , and velocity, v .*

SOLUTION 10.3. In view of matrices (5.12.1) and (5.12.3), 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, $\vartheta_t = \arcsin[(v_2/v_1) \sin \vartheta_i]$, we can express both A_r and A_t in terms of the angle of incidence, ϑ_i .

EXERCISE 10.4. *Using expressions (10.2.14) and (10.2.15), state the expressions for the reflection and transmission amplitudes for normal incidence in terms of mass density, ρ , and velocity, $v_{SH}(0)$.*

SOLUTION 10.4. Consider expressions (10.2.14) and (10.2.15). 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 (9.2.18) and (9.2.22), we obtain $v := v_{SH}(0) = \sqrt{C_{44}/\rho}$ and, hence, we can restate the above expressions as

$$(10.3.1) \quad A_r(0) = \frac{\rho_1 \sqrt{\frac{C_{44}^I}{\rho_1}} - \rho_2 \sqrt{\frac{C_{44}^{II}}{\rho_2}}}{\rho_1 \sqrt{\frac{C_{44}^I}{\rho_1}} + \rho_2 \sqrt{\frac{C_{44}^{II}}{\rho_2}}} = \frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2},$$

and

$$(10.3.2) \quad A_t(0) = \frac{2\rho_1 \sqrt{\frac{C_{44}^I}{\rho_1}}}{\rho_1 \sqrt{\frac{C_{44}^I}{\rho_1}} + \rho_2 \sqrt{\frac{C_{44}^{II}}{\rho_2}}} = \frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2},$$

where $v_1 \equiv v_{SH}(0)$ in the medium of incidence and $v_2 \equiv v_{SH}(0)$ in the medium of transmission.

REMARK 10.3.1. Expressions (10.3.1) and (10.3.2) are identical to expressions (6.7.12) and (6.7.13), respectively, which were derived in a different way in Chapter 6.

EXERCISE 10.5. Consider expression

$$(10.3.3) \quad \langle \mathfrak{E} \rangle = \frac{1}{2} \rho v \omega^2 A^2,$$

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 10.4, show that the energy is conserved.

SOLUTION 10.5. 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 (10.3.3), 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

$$(10.3.4) \quad 1 = A_r^2 + \frac{\rho_2 v_2}{\rho_1 v_1} A_t^2.$$

Inserting expressions (10.3.1) and (10.3.2) into expression (10.3.4), we get

$$\begin{aligned} 1 &= \left(\frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2} \right)^2 + \frac{\rho_2 v_2}{\rho_1 v_1} \left(\frac{2\rho_1 v_1}{\rho_1 v_1 + \rho_2 v_2} \right)^2 \\ &= \left(\frac{\rho_1 v_1 - \rho_2 v_2}{\rho_1 v_1 + \rho_2 v_2} \right)^2 + \frac{4\rho_1 v_1 \rho_2 v_2}{(\rho_1 v_1 + \rho_2 v_2)^2} = 1, \end{aligned}$$

as required.

EXERCISE 10.6. ¹¹Show that if $\sin \vartheta_t > 1$, the magnitude of A_r is equal to unity.

SOLUTION 10.6. If $\sin \vartheta_t > 1$, then $\cos \vartheta_t$ is a pure imaginary number. In that case, expression (10.2.14) is of the form

$$A_r = \frac{a - bi}{a + bi}.$$

¹¹See also Section 10.2.2

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.

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 first-order ordinary differential equations, which are expressed in terms of Hamiltonian \mathcal{H} and exist in the $\mathbf{x}\mathbf{p}$ -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

In this chapter, we transform Hamilton’s ray equations into Lagrange’s ray equations. Lagrange’s ray equations allow us to study ray theory in the realm of the calculus of variations. Thus, this chapter can be viewed as a transition between *Part 2* and *Part 3* of the book.

Readers who are not familiar with Legendre’s transformation might find it useful to study this chapter together with Appendix B.

11.1. Legendre’s transformation of Hamiltonian

In view of Appendix B, where we discuss Legendre’s transformation, we follow expression (B.3.8) to consider a new function given by

$$(11.1.1) \quad \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \sum_{j=1}^3 p_j(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}_j - \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}})),$$

¹In 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.

where \mathcal{L} is referred to as the ray-theory Lagrangian² corresponding to a given \mathcal{H} , and $p_j(\mathbf{x}, \dot{\mathbf{x}})$ is a solution of

$$(11.1.2) \quad \dot{x}_j = \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial p_j}, \quad j \in \{1, 2, 3\},$$

which is equation (B.3.7). Hence, in view of Appendix B, \mathcal{L} is Legendre's transformation of \mathcal{H} . Now, we wish to rewrite Hamilton's ray equations (11.2.1) in terms of Lagrangian (11.1.1).

11.2. Formulation of Lagrange's ray equations

To obtain Lagrange's ray equations, recall Hamilton's ray equations (8.2.7), namely,

$$(11.2.1) \quad \begin{cases} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Using expression (11.1.1), 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

$$(11.2.2) \quad \begin{aligned} \frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i} &= \frac{\partial}{\partial x_i} \left[\sum_{j=1}^3 p_j(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}_j - \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}})) \right] \\ &= \sum_{j=1}^3 \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i} \dot{x}_j - \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}}))}{\partial x_i} - \sum_{j=1}^3 \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}}))}{\partial p_j} \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i} \end{aligned}$$

and

$$(11.2.3) \quad \begin{aligned} \frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i} &= \frac{\partial}{\partial \dot{x}_i} \left[\sum_{j=1}^3 p_j(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}_j - \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}})) \right] \\ &= \sum_{j=1}^3 \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i} \dot{x}_j + p_i(\mathbf{x}, \dot{\mathbf{x}}) - \sum_{j=1}^3 \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p}(\mathbf{x}, \dot{\mathbf{x}}))}{\partial p_j} \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i}, \end{aligned}$$

respectively.

Using Hamilton's ray equations (11.2.1), we can restate expressions (11.2.2) and (11.2.3) as

$$(11.2.4) \quad \frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i} = \sum_{j=1}^3 \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i} \dot{x}_j + \dot{p}_i(\mathbf{x}, \dot{\mathbf{x}}) - \sum_{j=1}^3 \dot{x}_j \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i}, \quad i \in \{1, 2, 3\},$$

²In this book we use four distinct Lagrangians, which are denoted by \mathcal{L} , \mathcal{F} , \mathbb{L} and \mathbb{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.

and

$$(11.2.5) \quad \frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i} = \sum_{j=1}^3 \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i} \dot{x}_j + p_i(\mathbf{x}, \dot{\mathbf{x}}) - \sum_{j=1}^3 \dot{x}_j \frac{\partial p_j(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i}, \quad i \in \{1, 2, 3\},$$

respectively.

Examining expressions (11.2.4) and (11.2.5), we see that the first and the third terms on the right-hand sides cancel each other. Hence, $\partial \mathcal{L} / \partial x_i = \dot{p}_i$ and $\partial \mathcal{L} / \partial \dot{x}_i = p_i$. Thus, we conclude that

$$\frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial x_i} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial \dot{x}_i} \right), \quad i \in \{1, 2, 3\},$$

which we can rewrite as

$$(11.2.6) \quad \frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = 0, \quad i \in \{1, 2, 3\},$$

where \mathcal{L} is given in expression (11.1.1). System (11.2.1) is composed of six first-order ordinary differential equations in t to be solved for $\mathbf{x}(t)$ and $\mathbf{p}(t)$. As a result of Legendre's transformation, we expressed this system as three second-order ordinary differential equations (11.2.6) in t . These equations constitute a system of three second-order ordinary differential equations to be solved for $\mathbf{x}(t)$, which is the curve corresponding to the ray. We refer to these equations as Lagrange's ray equations.

Note that when we introduced $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ in expression (11.1.1), $\dot{\mathbf{x}}$ denoted a new variable, which, *a priori*, had no association with \mathbf{x} . However, if we consider the solution of system (11.2.1), 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 (11.1.2) together with the first equation of system (11.2.1), we see that $d\mathbf{x}(t)/dt = \dot{\mathbf{x}}(t)$. Consequently, at the end, our initial abuse of notation did no harm and, rather, might be viewed as insightful. In other words, depending on the context, $\dot{\mathbf{x}}$ can be viewed as an independent variable or as a function of t .

In view of this derivation, system (11.2.6) is equivalent to system (11.2.1). 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 11.2 and in Appendix B. This leads to the following proposition.

PROPOSITION 11.2.1. *Rays, parametrized by time, can be obtained either by solving Hamilton's ray equations (11.2.1) or by solving Lagrange's ray equations (11.2.6).*

We note that, in view of Legendre's transformation, the derivation of Lagrange's ray equations from Hamilton's ray equations requires regularity of Hamiltonian \mathcal{H} , namely,

$$\det \left[\frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j} \right] \neq 0, \quad i, j \in \{1, 2, 3\},$$

which is a necessary condition for Legendre's transformation to be a local diffeomorphism. This limitation is discussed in Section 13.1.2.

11.3. Beltrami's identity

For our subsequent work, we notice that we can write all the equations of system (11.2.6) as a single equation, namely,

$$(11.3.1) \quad \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\sum_{i=1}^3 \dot{x}_i \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \mathcal{L} \right) = 0.$$

Equation (11.3.1) is also valid for an n -dimensional case, where $i \in \{1, \dots, n\}$. The verification of this expression, for the two-dimensional case, is shown in Exercise 11.1.

We refer to equation (11.3.1) 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 (11.3.1) 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.

Closing remarks

To describe rays in anisotropic inhomogeneous continua, we can use either Hamilton's ray equations or Lagrange's ray equations. Herein, Hamilton's ray equations are directly rooted in fundamental physical principles discussed in Chapter 2, while Lagrange's ray equations are based on these principles via Legendre's transformation, which links the two systems. Thus, Lagrange's ray equations are subject to the singularities of this transformation.

However, we can also treat Lagrange's ray equations in their own right without invoking Hamilton's ray equations. Lagrange's ray equations belong to the realm of variational methods. Hence, in *Part 3*, we will introduce the tools of the calculus of variations, which allow us to base these equations on Fermat's variational principle.



11.4. Exercises

EXERCISE 11.1. ³Considering a two-dimensional continuum, verify that, given Lagrangian \mathcal{L} that satisfies Lagrange's ray equations (11.2.6), Beltrami's identity (11.3.1) is also satisfied.

SOLUTION 11.1. 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 (11.3.1) can be written as

$$(11.4.1) \quad \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\dot{x} \frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{z} \frac{\partial \mathcal{L}}{\partial \dot{z}} - \mathcal{L} \right) = 0.$$

Differentiating the left-hand side of equation (11.4.1), we obtain

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\dot{x} \frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{z} \frac{\partial \mathcal{L}}{\partial \dot{z}} - \mathcal{L} \right) &= \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\dot{x} \frac{\partial \mathcal{L}}{\partial \dot{x}} \right) + \frac{d}{dt} \left(\dot{z} \frac{\partial \mathcal{L}}{\partial \dot{z}} \right) - \frac{d\mathcal{L}}{dt} \\ &= \frac{\partial \mathcal{L}}{\partial t} + \ddot{x} \frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{x} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) + \ddot{z} \frac{\partial \mathcal{L}}{\partial \dot{z}} + \dot{z} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{z}} \right) \\ &\quad - \left[\frac{\partial \mathcal{L}}{\partial x} \dot{x} + \frac{\partial \mathcal{L}}{\partial z} \dot{z} + \frac{\partial \mathcal{L}}{\partial \dot{x}} \ddot{x} + \frac{\partial \mathcal{L}}{\partial \dot{z}} \ddot{z} + \frac{\partial \mathcal{L}}{\partial t} \right] \\ &= \dot{x} \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) - \frac{\partial \mathcal{L}}{\partial x} \right] + \dot{z} \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{z}} \right) - \frac{\partial \mathcal{L}}{\partial z} \right] = 0, \end{aligned}$$

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 corresponds to a ray equation from system (11.2.6), namely,

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = 0 \\ \frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{z}} \right) = 0 \end{cases}.$$

EXERCISE 11.2. Assuming that Hamiltonian \mathcal{H} and Lagrangian \mathcal{L} do not explicitly depend on t , following expression (11.1.1) and using equations (11.2.6), derive Hamilton's ray equations (8.2.7).

SOLUTION 11.2. Consider $\mathcal{H}(x, p)$ and $\mathcal{L}(x, \dot{x})$. Following expression (11.1.1), the differential of \mathcal{H} becomes

$$d\mathcal{H} = \sum_{i=1}^3 dp_i \dot{x}_i + \sum_{i=1}^3 p_i d\dot{x}_i - \sum_{i=1}^3 \frac{\partial \mathcal{L}}{\partial x_i} dx_i - \sum_{i=1}^3 \frac{\partial \mathcal{L}}{\partial \dot{x}_i} d\dot{x}_i.$$

In view of expression (B.3.2), 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

$$(11.4.2) \quad d\mathcal{H} = \sum_{i=1}^3 \dot{x}_i dp_i - \sum_{i=1}^3 \frac{\partial \mathcal{L}}{\partial x_i} dx_i.$$

³See also Section 14.5

Also, the differential of \mathcal{H} , can be formally written as

$$(11.4.3) \quad d\mathcal{H} = \sum_{i=1}^3 \frac{\partial \mathcal{H}}{\partial p_i} dp_i + \sum_{i=1}^3 \frac{\partial \mathcal{H}}{\partial x_i} dx_i.$$

Equating the corresponding terms of expression (11.4.2) and its formal statement (11.4.3), we can write

$$\left\{ \begin{array}{l} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \frac{\partial \mathcal{L}}{\partial x_i} = -\frac{\partial \mathcal{H}}{\partial x_i} \end{array} \right., \quad i \in \{1, 2, 3\}.$$

Invoking Lagrange's ray equation (11.2.6) and recalling expression (B.3.2), we can write

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{dp_i}{dt} \equiv \frac{\partial \mathcal{L}}{\partial x_i} - \dot{p}_i = 0, \quad i \in \{1, 2, 3\}.$$

Hence, we obtain

$$\dot{p}_i = \frac{\partial \mathcal{L}}{\partial x_i}, \quad i \in \{1, 2, 3\}.$$

Thus, we can write

$$\left\{ \begin{array}{l} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \end{array} \right., \quad i \in \{1, 2, 3\},$$

which are Hamilton's ray equations (8.2.7), as required.

EXERCISE 11.3. ⁴In view of the polar reciprocity of the phase-slowness curve and the ray-velocity curve, derive the equation that, while analogous to expression (8.4.12), relates phase angle to both ray velocities and ray angles, namely,

$$\tan \vartheta = \frac{\tan \theta - \frac{1}{V(\theta)} \frac{\partial V}{\partial \theta}}{1 + \frac{\tan \theta}{V(\theta)} \frac{\partial V}{\partial \theta}}.$$

SOLUTION 11.3. 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

$$\mathcal{L} = \frac{\dot{x}_1^2 + \dot{x}_3^2}{[V(\theta)]^2},$$

⁴See also Section 10.1.2

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

$$\frac{\partial}{\partial\dot{x}_1} = \frac{\partial\theta}{\partial\dot{x}_1} \frac{\partial}{\partial\theta} = \frac{\partial \arctan \frac{\dot{x}_1}{\dot{x}_3}}{\partial\dot{x}_1} \frac{\partial}{\partial\theta} = \frac{\frac{1}{\dot{x}_3}}{1 + \left(\frac{\dot{x}_1}{\dot{x}_3}\right)^2} \frac{\partial}{\partial\theta} = \frac{\dot{x}_3}{V^2} \frac{\partial}{\partial\theta},$$

and

$$\frac{\partial}{\partial\dot{x}_3} = \frac{\partial\theta}{\partial\dot{x}_3} \frac{\partial}{\partial\theta} = \frac{\partial \arctan \frac{\dot{x}_1}{\dot{x}_3}}{\partial\dot{x}_3} \frac{\partial}{\partial\theta} = -\frac{\frac{\dot{x}_1}{\dot{x}_3^2}}{1 + \left(\frac{\dot{x}_1}{\dot{x}_3}\right)^2} \frac{\partial}{\partial\theta} = -\frac{\dot{x}_1}{V^2} \frac{\partial}{\partial\theta}.$$

Consider the expression for the phase-slowness components and for the ray-theory Lagrangian. Using the quotient rule, we can write

$$p_1 = \frac{\partial\mathcal{L}}{\partial\dot{x}_1} = \frac{2\dot{x}_1 V^2 - 2(\dot{x}_1^2 + \dot{x}_3^2) V \frac{\partial V}{\partial\dot{x}_1}}{V^4} = \frac{2\dot{x}_1 - 2V \frac{\partial V}{\partial\dot{x}_1}}{V^2},$$

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

$$p_1 = 2 \frac{\dot{x}_1 - \frac{\dot{x}_3}{V} \frac{\partial V}{\partial\theta}}{V^2}.$$

Similarly, we get

$$p_3 = 2 \frac{\dot{x}_3 + \frac{\dot{x}_1}{V} \frac{\partial V}{\partial\theta}}{V^2}.$$

Thus,

$$(11.4.4) \quad \tan \vartheta = \frac{p_1}{p_3} = \frac{\dot{x}_1 - \frac{\dot{x}_3}{V} \frac{\partial V}{\partial\theta}}{\dot{x}_3 + \frac{\dot{x}_1}{V} \frac{\partial V}{\partial\theta}} = \frac{\frac{\dot{x}_1}{\dot{x}_3} - \frac{1}{V} \frac{\partial V}{\partial\theta}}{1 + \frac{\frac{\dot{x}_1}{\dot{x}_3} \frac{\partial V}{\partial\theta}}{V}} = \frac{\tan \theta - \frac{1}{V} \frac{\partial V}{\partial\theta}}{1 + \frac{\tan \theta}{V} \frac{\partial V}{\partial\theta}},$$

as required, which shows that $1/v(\vartheta)$ and $V(\theta)$ are polar reciprocals of one another.

REMARK 11.4.1. Expression (11.4.4) 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 11.4. Using expressions (8.4.12) and (11.4.4) and following standard trigonometric identities, show that

$$\frac{\partial}{\partial\vartheta} \ln v = \frac{\partial}{\partial\theta} \ln V.$$

SOLUTION 11.4. Note that expression (8.4.12) can be written as

$$\tan \theta = \frac{\tan \vartheta + \frac{\partial}{\partial\vartheta} \ln v}{1 - \tan \vartheta \frac{\partial}{\partial\vartheta} \ln v} = \frac{\tan \vartheta + \tan [\arctan (\frac{\partial}{\partial\vartheta} \ln v)]}{1 - \tan \vartheta \tan [\arctan (\frac{\partial}{\partial\vartheta} \ln v)]},$$

which, following trigonometric identities, we can write as

$$(11.4.5) \quad \theta = \vartheta + \arctan \left(\frac{\partial}{\partial\vartheta} \ln v \right).$$

Similarly, expression (11.4.4) can be written as

$$(11.4.6) \quad \vartheta = \theta - \arctan \left(\frac{\partial}{\partial \theta} \ln V \right).$$

Solving expression (11.4.6) for θ and equating it to expression (11.4.5), we obtain

$$\frac{\partial}{\partial \vartheta} \ln v = \frac{\partial}{\partial \theta} \ln V,$$

as required.

Part 3

Variational formulation of rays

Introduction to Part 3

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.⁵

Richard Feynman (1967) The Character of Physical Law

The fundamental formulation of ray theory was presented in *Part 2*. This theory is based on Cauchy's equations of motion in anisotropic inhomogeneous continua and results in Hamilton's ray equations. Also, in *Part 2*, we used Legendre's transformation of the ray-theory Hamiltonian to obtain Lagrange's ray equations. Thus, within the limitations of this transformation, we have two equivalent forms of the ray equations.

In *Part 3*, 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.

In search of the stationarity condition for a definite integral that describes the traveltime of the signal between a source and a receiver, we will use the calculus of variations. 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.

⁵Readers 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.

The first scientific statement of a variational principle was formulated in optics by Pierre de Fermat in 1657.⁶ In its original formulation, this principle was referred to as the principle of least time.⁷ 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 1835.⁸

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.

⁶Interested readers might refer to Born, M., and Wolf, E., (1999) *Principles of optics* (7th edition): Cambridge University Press, p. xxvi.

⁷Readers interested in discussion of Fermat's principle in the context of quantum electrodynamics might refer to Feynman, R.P., (1985/2006) *QED: The strange theory of light and matter*: Princeton University Press, pp. 38 – 45 and pp. 50 – 52.

⁸Readers 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 quantum 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.¹

Leonhard Euler (1744) Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti²

Preliminary remarks

In Chapter 11, 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.

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.

¹Readers interested in the modern perspective on this statement might refer to Brown, J.R., (2001) Who rules in science: An opinionated guide to the wars: Harvard University Press, pp. 107 – 108:

During the heyday of fruitful interaction of science and religion, Kepler, Newton, Leibniz, and many others brought religious considerations to bear on scientific theorizing. [...] Their religious beliefs were not mere acts of faith, but the consequence of rational considerations. [...] It is no longer rational to set religious constraints on theorizing in biology or astronomy or anywhere else.

²Method of finding curved lines enjoying the maximum and minimum property; or the solution of the isoperimetric problem understood in the broadest sense

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.³

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

$$(12.1.1) \quad I = \int_a^b F(z(x), z'(x); x) dx.$$

Thus, we seek function $z(x)$ that makes integral (12.1.1) 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, $z(x)$, $z'(x) \equiv dz/dx$ 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

$$(12.1.2) \quad F(z(x), z'(x); x) \equiv F(\xi_1, \xi_2, \xi_3).$$

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 look 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 $\delta \int_a^b F dx = 0$.

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 $dz(x)$ refers to a differential change along a given curve for a change in x .⁴

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 11, respectively.⁵

³Readers interested in a definition of stationarity of a definite integral might refer to Arnold, V.I., (1989) *Mathematical methods of classical mechanics* (2nd edition): Springer-Verlag, p. 57.

⁴Readers 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.

⁵Readers interested in the relation between the explicit and parametric 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.

The condition of stationarity of integral (12.1.1) was derived by Euler in 1744.⁶ 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 (12.1.1). 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 (12.1.1) stationary. This stationarity condition is stated by the following theorem.

THEOREM 12.2.1. *Function $z(x)$ with the continuous first derivative on interval $[a, b]$ yields a stationary value of integral (12.1.1), namely,*

$$(12.2.1) \quad I = \int_a^b F(z(x), z'(x); x) dx,$$

in the class of functions with boundary conditions $z(a) = z_a$ and $z(b) = z_b$, if equation

$$(12.2.2) \quad \frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = 0$$

*is satisfied.*⁷

We refer to equation (12.2.2) as Euler's equation. Euler's equation (12.2.2) is a second-order ordinary differential equation.⁸

To see the connection between integral (12.2.1) and its stationarity condition, given by Euler's equation (12.2.2), consider the following heuristic argument.

Replace the integral by a finite sum of subdivisions given by $x_0, x_1, \dots, x_{n-1}, x_n$, 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 (12.2.1) can be written as

$$S_n = \sum_{i=0}^{n-1} F(z_{i+1}, z'_{i+1}; x_{i+1}) \Delta x,$$

where

$$(12.2.3) \quad z'_{i+1} := \frac{z_{i+1} - z_i}{\Delta x}.$$

⁶It 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 (2nd edition): Springer-Verlag, pp. 231 – 234.

⁷Readers interested in a rigorous proof of Theorem 12.2.1 might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2nd edition): Springer-Verlag, pp. 57 – 58.

⁸Readers 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.

Here S_n is viewed as a function of the $n - 1$ variables, z_1, \dots, 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

$$(12.2.4) \quad \frac{\partial S_n}{\partial z_i} = 0, \quad i \in \{1, \dots, n-1\}.$$

In view of expression (12.2.3), in the sum S_n , for any given $i \in (1, \dots, n-1)$, there are only two consecutive terms that explicitly contain a given z_i , namely,

$$(12.2.5) \quad F(z_i, z'_i; x_i) \Delta x + F(z_{i+1}, z'_{i+1}; x_{i+1}) \Delta x.$$

Applying stationarity condition (12.2.4), we take the derivative of expression (12.2.5) with respect to z_i and obtain

$$(12.2.6) \quad \left[\frac{\partial F}{\partial z}(z_i, z'_i; x_i) + \frac{\partial F}{\partial z'}(z_i, z'_i; x_i) \frac{\partial z'_i}{\partial z_i} \right] \Delta x + \left[\frac{\partial F}{\partial z'}(z_{i+1}, z'_{i+1}; x_{i+1}) \frac{\partial z'_{i+1}}{\partial z_i} \right] \Delta x = 0,$$

where $i \in \{1, \dots, n-1\}$.

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 (12.3.1).

In view of equation (12.2.6) and recalling expression (12.2.3), 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 (12.2.6) becomes

$$\left[\frac{\partial F}{\partial z}(z_i, z'_i; x_i) + \frac{\partial F}{\partial z'}(z_i, z'_i; x_i) \frac{1}{\Delta x} \right] \Delta x - \left[\frac{\partial F}{\partial z'}(z_{i+1}, z'_{i+1}; x_{i+1}) \frac{1}{\Delta x} \right] \Delta x = 0,$$

where $i \in \{1, \dots, n-1\}$. This equation can be rearranged to give

$$(12.2.7) \quad \frac{\partial F}{\partial z}(z_i, z'_i; x_i) - \frac{1}{\Delta x} \left[\frac{\partial F}{\partial z'}(z_{i+1}, z'_{i+1}; x_{i+1}) - \frac{\partial F}{\partial z'}(z_i, z'_i; x_i) \right] = 0,$$

where $i \in \{1, \dots, n-1\}$.

We now assume that as $\Delta x \rightarrow 0$ and $x_i \rightarrow x \in [a, b]$, z_i approaches $z(x)$ and $z'_i = (z_i - z_{i-1})/\Delta x$ approaches $z'(x)$. Then, equation (12.2.7) becomes

$$\frac{\partial F}{\partial z}(z(x), z'(x); x) - \frac{d}{dx} \left[\frac{\partial F}{\partial z'}(z(x), z'(x); x) \right] = 0,$$

which is Euler's equation (12.2.2), as required.

12.3. Beltrami's identity

A convenient form of Euler's equation (12.2.2) is Beltrami's identity, discussed in Section 11.3. In the two-dimensional case, where we look for the $z(x)$ that is a solution of Euler's equation (12.2.2), Beltrami's identity is equivalent to that equation. Hence, we can write

$$(12.3.1) \quad \frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = 0 = \frac{\partial F}{\partial x} + \frac{d}{dx} \left(z' \frac{\partial F}{\partial z'} - F \right).$$

In general, Beltrami's identity is not equivalent to the corresponding Euler's equations or Lagrange's ray equations. In Chapter 11, for instance, a single expression of Beltrami's identity (11.3.1) refers to three equations of system (11.2.6) and, hence, by itself, cannot give a unique solution of system (11.2.6). A verification of Beltrami's identity (12.3.1) 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 (12.3.1) vanishes. Important consequences of this simplification are discussed in Chapter 14 in the context of ray parameters.

12.4. Generalizations of Euler's equation

12.4.1. Introductory comments. Integral (12.1.1) 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.

12.4.2. 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,

$$I = \int_{a_y}^{b_y} \int_{a_x}^{b_x} F(z(x, y), z_x, z_y; x, y) dx dy,$$

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, $z(x, y)$, that renders I stationary. In this case, Euler's equation becomes

$$(12.4.1) \quad \frac{\partial F}{\partial z} - \left[\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial z_x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial z_y} \right) \right] = 0,$$

which is a second-order partial differential equation. The generalization for n variables follows the same pattern, thereby giving

$$(12.4.2) \quad \frac{\partial F}{\partial z} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial z_{x_i}} \right) = 0,$$

where $z = z(x_1, \dots, x_n)$ and $z_{x_i} := \partial z / \partial x_i$, with $i \in \{1, \dots, n\}$.

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 13.2.4.

12.4.3. 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,

$$I = \int_a^b F(y(x), y'(x), z(x), z'(x); x) dx.$$

Thus, we look for smooth curves $y(x)$ and $z(x)$ that render I stationary, subject to constraints

$$(12.4.3) \quad \begin{cases} y(a) = a_1 \\ z(a) = a_2 \\ y(b) = b_1 \\ z(b) = b_2 \end{cases},$$

where a_i and b_i are constants. In this case, Euler's equations become a system of second-order ordinary differential equations

$$\begin{cases} \frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0 \\ \frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = 0 \end{cases}.$$

The generalization for n functions follows the same pattern thereby giving a system of n equations,

$$(12.4.4) \quad \frac{\partial F}{\partial \zeta_i} - \frac{d}{dx} \left(\frac{\partial F}{\partial \zeta'_i} \right) = 0, \quad i \in \{1, \dots, n\},$$

where $\zeta_i = \zeta_i(x)$ and $\zeta'_i = d\zeta_i(x)/dx$.

12.4.4. 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,

$$I = \int_a^b F(z(x), z'(x), z''(x); x) dx.$$

Thus, we look for a smooth curve $z(x)$ that renders I stationary, subject to constraints

$$(12.4.5) \quad \begin{cases} z(a) = a_1 \\ z'(a) = a_2 \\ z(b) = b_1 \\ z'(b) = b_2 \end{cases},$$

where a_i and b_i are constants. In this case, Euler's equation becomes

$$\frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) + \frac{d^2}{dx^2} \left(\frac{\partial F}{\partial z''} \right) = 0.$$

This is a fourth-order ordinary differential equation. The generalization for n th-order derivatives follows the same pattern to yield

$$\frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) + \frac{d^2}{dx^2} \left(\frac{\partial F}{\partial z''} \right) + \dots + (-1)^n \frac{d^n}{dx^n} \left(\frac{\partial F}{\partial z^{(n)}} \right) = 0,$$

which is an ordinary differential equation of order $2n$.

12.5. Special cases of Euler's equation

12.5.1. Introductory comments. There are cases where, due to the explicit absence of certain arguments or to the particular form of integral (12.1.1), Euler's equation (12.2.2) 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 (12.1.2), 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.

12.5.2. 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 (12.2.2) is reduced to

$$\frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = 0.$$

The vanishing of the total derivative implies that

$$(12.5.1) \quad \frac{\partial F}{\partial z'} = C_1,$$

where C_1 denotes a constant. Thus, $z(x)$ is obtained as a solution of first-order ordinary differential equation (12.5.1).

12.5.3. 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 (12.5.1) as

$$(12.5.2) \quad \frac{dF(z')}{dz'} = C_1.$$

Denoting $dF(z')/dz'$ as $f(z')$, we can write equation (12.5.2) as

$$f(z') = C_1.$$

Assuming that $df/dz' \neq 0$, we can consider inverse function f^{-1} . Thus, we can write

$$z' = f^{-1}(C_1).$$

Recalling that $z' \equiv dz/dx$ and denoting $f^{-1}(C_1) = C_2$, we can write

$$(12.5.3) \quad \frac{dz}{dx} = C_2.$$

This is a first-order ordinary differential equation, whose solution,

$$z = C_2x + C_3,$$

is obtained directly by integration.

Thus, finding the curve which gives a stationary value of $\int_a^b F(z')dx$ 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.

12.5.4. Independence of x . Let us consider an integrand that is explicitly independent of x , namely, $F = F(z, z')$. Using Beltrami's identity (12.3.1), we obtain

$$\frac{d}{dx} \left(z' \frac{\partial F}{\partial z'} - F \right) = 0.$$

The vanishing of the total derivative implies that

$$(12.5.4) \quad z' \frac{\partial F}{\partial z'} - F = C,$$

where C denotes a constant. Thus, $z(x)$ is obtained as a solution of first-order ordinary differential equation (12.5.4).

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 (12.5.4) is a ray parameter, discussed in Chapter 14.

12.5.5. Total derivative. Let integrand $F(x, z, z')$ be a total derivative of function $f(x, z)$ with respect to x , namely,

$$(12.5.5) \quad F(z, z'; x) = \frac{df(x, z)}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} z'.$$

Consider the left-hand side of Euler's equation (12.2.2). Inserting function (12.5.5), we obtain

$$\frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = \frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2} z' - \frac{d}{dx} \left(\frac{\partial f}{\partial z} \right) = \frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2} z' - \left(\frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2} z' \right) = 0.$$

Thus, equation (12.2.2) is identically satisfied. Consequently, if F is a total derivative, Euler's equation (12.2.2) is satisfied by any $z(x)$. In other words, if a variational problem involves the integral of a total differential, namely,

$$\delta \int_a^b df(x, z) = 0,$$

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 12.3.⁹ 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.

12.5.6. Function of x and z .

Euler's equation. In physically motivated problems, we often encounter an integral given by $\int_a^b h(\mathbf{x}) ds$, which is an integral of function h , whose value depends on position \mathbf{x} along the arclength element ds . Such an integral represents a certain quantity measured along a trajectory that connects points a and b . Considering the two-dimensional case and assuming that the trajectory can be expressed as $z = z(x)$, we can write such an integral as

$$(12.5.6) \quad \int_{a_x}^{b_x} h(x, z) \sqrt{1 + \left(\frac{dz}{dx} \right)^2} dx.$$

Thus, Euler's equation (12.2.2) becomes

$$\frac{\partial}{\partial z} \left[h(x, z) \sqrt{1 + (z')^2} \right] - \frac{d}{dx} \left[\frac{\partial}{\partial z'} \left(h(x, z) \sqrt{1 + (z')^2} \right) \right] = 0,$$

⁹In 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.

where $z' \equiv dz/dx$. Performing partial derivatives, we obtain

$$\frac{\partial h(x, z)}{\partial z} \sqrt{1 + (z')^2} - \frac{d}{dx} \left[h(x, z) \frac{z'}{\sqrt{1 + (z')^2}} \right] = 0.$$

Then, by the product rule, we get

$$\frac{\partial h(x, z)}{\partial z} \sqrt{1 + (z')^2} - \left[\frac{dh(x, z)}{dx} \frac{z'}{\sqrt{1 + (z')^2}} + h(x, z) \frac{d}{dx} \frac{z'}{\sqrt{1 + (z')^2}} \right] = 0.$$

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

$$\frac{\partial h}{\partial z} \sqrt{1 + (z')^2} - \frac{\partial h}{\partial x} \frac{z'}{\sqrt{1 + (z')^2}} - \frac{\partial h}{\partial z} \frac{(z')^2}{\sqrt{1 + (z')^2}} - h \frac{z''}{[1 + (z')^2]^{\frac{3}{2}}} = 0.$$

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

$$(12.5.7) \quad \frac{\partial h}{\partial z} - \frac{\partial h}{\partial x} \frac{dz}{dx} - h \frac{\frac{d^2 z}{dx^2}}{1 + \left(\frac{dz}{dx}\right)^2} = 0.$$

To study equation (12.5.7) 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 (12.5.7), we obtain

$$(12.5.8) \quad v \frac{d^2 z}{dx^2} - \frac{\partial v}{\partial x} \left(\frac{dz}{dx} \right)^3 + \frac{\partial v}{\partial z} \left(\frac{dz}{dx} \right)^2 - \frac{\partial v}{\partial x} \frac{dz}{dx} + \frac{\partial v}{\partial z} = 0,$$

where we assume $v(x, z) \neq 0$. In such a case, integral (12.5.6), which can be rewritten as

$$(12.5.9) \quad \int_{a_x}^{b_x} \frac{\sqrt{1 + (z')^2}}{v(x, z)} dx,$$

represents the traveltime between two points. Thus, a solution of equation (12.5.8) is a ray, $z(x)$, satisfying Fermat's principle of stationary traveltime.

If $v(x, z)$ is given by a constant, integral (12.5.9) is explicitly independent of x and z , and it corresponds to the case discussed in Section 12.5.3. In such a case, equation (12.5.8) reduces to $d^2z/dx^2 = 0$, 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 (12.5.9).

Geometrical interpretation and physical meaning. Integral (12.5.6) 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 (12.5.6) 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 (12.5.7), 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 (12.5.7) 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 (12.5.8) 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 12.5.3, 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 12.6.1. If an n th-order differential equation

$$(12.6.1) \quad f(x, z, z', \dots, z^{(n)}) = 0,$$

can be transformed to the equivalent form

$$\frac{d}{dx}g(x, z, z', \dots, z^{(n-1)}) = 0,$$

we see that

$$(12.6.2) \quad g(x, z, z', \dots, z^{(n-1)}) = C,$$

where C is a constant. Expression (12.6.2) is a “first integral” of equation (12.6.1).

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 by Emmy Noether in Göttingen in 1918 in her paper entitled “Invariante Variationsprobleme”. We can state this theorem in the following way.¹⁰

THEOREM 12.6.2. *To every change of dependent or independent variables that leaves the integral of the Lagrangian invariant, there corresponds a conservation law.*

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 conserved quantity is a ray parameter, which we will discuss 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 (11.2.6) belong to the realm of Euler's equations. The parametric form of Euler's equation (12.2.2) corresponds to a system of two Euler's equations, namely,

$$(12.7.1) \quad \begin{cases} \frac{\partial G}{\partial x} - \frac{d}{dt} \left(\frac{\partial G}{\partial \dot{x}} \right) = 0 \\ \frac{\partial G}{\partial z} - \frac{d}{dt} \left(\frac{\partial G}{\partial \dot{z}} \right) = 0 \end{cases},$$

where $G = G(x, z, \dot{x}, \dot{z})$ with $\dot{x} := dx/dt$ and $\dot{z} := dz/dt$. A solution of system (12.7.1) is a curve in the xz -plane given by $[x(t), z(t)]$ that corresponds to variational problem

$$\delta \int G dt = 0.$$

To see the relation between G and F , which is stated in integral (12.1.1), we can write $dt = dx/\dot{x}$ and $z' := dz/dx$. Hence, $G(x, z, \dot{x}, \dot{z}) = F(z, \dot{z}/\dot{x}, x) \dot{x}$.

¹⁰Readers 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.

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,

$$(12.7.2) \quad \frac{\partial G}{\partial x_i} - \frac{d}{dt} \frac{\partial G}{\partial \dot{x}_i} = 0, \quad i \in \{1, \dots, n\},$$

whose solution is a curve in the \mathbf{x} -space given by $\mathbf{x}(t)$. Examining systems (11.2.6) and (12.7.2), we conclude that Lagrange's ray equations (11.2.6) possess the form of Euler's equations (12.7.2).

The fact that Euler's and Lagrange's equations have equivalent forms is the reason why equations of the form (11.2.6) and (12.7.2) 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.¹¹ 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. Also in Chapter 13, we will discuss Hamilton's variational principle of stationary action, which we will use to derive the wave equation. In Chapter 14, we will show that first integrals, which correspond to conserved quantities along the rays, can be used in raytracing techniques.



¹¹Readers interested in geodesic fields and its implication to minima and maxima might refer to Kreyszig, E., (1959/1991) Differential geometry: Dover, pp. 162 – 168.

12.8. Exercises

EXERCISE 12.1. *In view of Euler's equation (12.2.2), verify Beltrami's identity (12.3.1).*

SOLUTION 12.1. Consider $F = F(x, z, z')$ and Beltrami's identity (12.3.1). We can write

$$\begin{aligned} \frac{\partial F}{\partial x} + \frac{d}{dx} \left(z' \frac{\partial F}{\partial z'} - F \right) &= \frac{\partial F(x, z, z')}{\partial x} + \frac{d}{dx} \left(z' \frac{\partial F(x, z, z')}{\partial z'} \right) - \frac{dF(x, z, z')}{dx} \\ &= \frac{\partial F}{\partial x} + z'' \frac{\partial F}{\partial z'} + z' \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) - \left(\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} z' + \frac{\partial F}{\partial z'} z'' \right) \\ &= z' \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) - \frac{\partial F}{\partial z} z' = -z' \left[\frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) \right], \end{aligned}$$

where the terms in brackets is Euler's equation (12.2.2). Thus,

$$\frac{\partial F}{\partial x} + \frac{d}{dx} \left(z' \frac{\partial F}{\partial z'} - F \right) = 0,$$

as required.

EXERCISE 12.2. *Following the argument used to justify Theorem 12.2.1, derive the explicit form of Beltrami's identity (12.3.1).*

SOLUTION 12.2. To obtain Beltrami's identity, consider term (12.2.5), namely,

$$F(z_i, z'_i; x_i) \Delta x + F(z_{i+1}, z'_{i+1}; x_{i+1}) \Delta x,$$

where $i \in \{1, \dots, n-1\}$. Differentiating with respect to x , we obtain

$$\begin{aligned} &\left[\frac{\partial F}{\partial x}(z_i, z'_i; x_i) + \frac{\partial F}{\partial z'}(z_i, z'_i; x_i) \frac{\partial z'_i}{\partial x_i} \right] \Delta x + F(z_i, z'_i; x_i) \frac{\partial \Delta x}{\partial x_i} \\ (12.8.1) \quad &+ \frac{\partial F}{\partial z'}(z_{i+1}, z'_{i+1}; x_{i+1}) \frac{\partial z'_{i+1}}{\partial x_i} \Delta x + F(z_{i+1}, z'_{i+1}; x_{i+1}) \frac{\partial \Delta x}{\partial x_i}. \end{aligned}$$

Recalling expression (12.2.3) and the appropriate expression for Δx , we can write

$$\begin{aligned} \frac{\partial z'_i}{\partial x_i} &= \frac{\partial}{\partial x_i} \left(\frac{z_i - z_{i-1}}{x_i - x_{i-1}} \right) = -\frac{z_i - z_{i-1}}{(x_i - x_{i-1})^2} = -\frac{z_i - z_{i-1}}{(\Delta x)^2}, \\ \frac{\partial \Delta x}{\partial x_i} (z_i, z'_i; x_i) &= \frac{\partial}{\partial x_i} (x_i - x_{i-1}) = 1, \\ \frac{\partial z'_{i+1}}{\partial x_i} &= \frac{\partial}{\partial x_i} \left(\frac{z_{i+1} - z_i}{x_{i+1} - x_i} \right) = \frac{z_{i+1} - z_i}{(x_{i+1} - x_i)^2} = \frac{z_{i+1} - z_i}{(\Delta x)^2}, \\ \frac{\partial \Delta x}{\partial x_i} (z_{i+1}, z'_{i+1}; x_{i+1}) &= \frac{\partial}{\partial x_i} (x_{i+1} - x_i) = -1. \end{aligned}$$

Hence, expression (12.8.1) becomes

$$\begin{aligned} & \left[\frac{\partial F}{\partial x} (z_i, z'_i; x_i) - \frac{\partial F}{\partial z'} (z_i, z'_i; x_i) \frac{z_i - z_{i-1}}{(\Delta x)^2} \right] \Delta x + F (z_i, z'_i; x_i) \\ & + \frac{\partial F}{\partial z'} (z_{i+1}, z'_{i+1}; x_{i+1}) \frac{z_{i+1} - z_i}{(\Delta x)^2} \Delta x - F (z_{i+1}, z'_{i+1}; x_{i+1}). \end{aligned}$$

As in stationarity condition (12.2.4), we have a system of $n - 1$ equations, namely,

$$\begin{aligned} \frac{\partial S_n}{\partial x_i} &= 0 \\ &= F (z_i, z'_i; x_i) - F (z_{i+1}, z'_{i+1}; x_{i+1}) \\ &\quad - \frac{\partial F}{\partial z'} (z_i, z'_i; x_i) \frac{z_i - z_{i-1}}{\Delta x} + \frac{\partial F}{\partial z'} (z_{i+1}, z'_{i+1}; x_{i+1}) \frac{z_{i+1} - z_i}{\Delta x} \\ &\quad + \frac{\partial F}{\partial x} (z_i, z'_i; x_i) \Delta x, \end{aligned}$$

where $i \in \{1, \dots, n - 1\}$. Dividing both sides of each equation by Δx and using the appropriate definition of Δx , we can write

$$\begin{aligned} 0 &= \frac{F (z_i, z'_i; x_i) - F (z_{i+1}, z'_{i+1}; x_{i+1})}{\Delta x} \\ &\quad + \frac{\frac{\partial F}{\partial z'} (z_{i+1}, z'_{i+1}; x_{i+1}) \frac{z_{i+1} - z_i}{\Delta x} - \frac{\partial F}{\partial z'} (z_i, z'_i; x_i) \frac{z_i - z_{i-1}}{\Delta x}}{\Delta x} \\ &\quad + \frac{\partial F}{\partial x} (z_i, z'_i; x_i), \end{aligned}$$

where $i \in \{1, \dots, n - 1\}$. Letting $\Delta x \rightarrow 0$, 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

$$0 = -\frac{dF}{dx} + \frac{d}{dx} \left(\frac{\partial F}{\partial z'} z' \right) + \frac{\partial F}{\partial x}.$$

Rearranging and using the linearity of the differential operator, we get

$$\frac{\partial F}{\partial x} + \frac{d}{dx} \left[z' \frac{\partial F}{\partial z'} - F \right] = 0,$$

which is Beltrami's identity (12.3.1), as required.

EXERCISE 12.3. ¹²*Prove the following lemma.*

NOTATION 12.8.1. To state Lemma 12.8.2, below, we use the parametric form of the variational problems, rather than the explicit form used in this chapter.

LEMMA 12.8.2. *The two variational problems given by*

$$(12.8.2) \quad \delta \int F(\mathbf{x}, \dot{\mathbf{x}}) dt = 0$$

and

$$(12.8.3) \quad \delta \int \left[cF(\mathbf{x}, \dot{\mathbf{x}}) + \sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right] dt = 0$$

have the same function $\mathbf{x}(t)$ that renders the corresponding integrals stationary, if $f_j(\mathbf{x})$ are the components of a gradient of a function and c is a nonzero constant.

SOLUTION 12.3. We show two different proofs of Lemma 12.8.2. 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. [Proof A] To prove that $\mathbf{x}(t)$ is the same for variational problems (12.8.2) and (12.8.3), we reduce problem (12.8.3) to problem (12.8.2). Consider the integral of variational problem (12.8.3). In view of the linearity of the integral operator, we can write

$$\begin{aligned} \int \left[cF(\mathbf{x}, \dot{\mathbf{x}}) + \sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right] dt &= c \int F(\mathbf{x}, \dot{\mathbf{x}}) dt + \int \sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j dt \\ &= c \int F(\mathbf{x}, \dot{\mathbf{x}}) dt + \int \sum_{j=1}^n f_j(\mathbf{x}) dx_j. \end{aligned}$$

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

$$(12.8.4) \quad \int \sum_{j=1}^n f_j(\mathbf{x}) dx_j = \int \sum_{j=1}^n \frac{\partial g(\mathbf{x})}{\partial x_j} dx_j.$$

Since integral (12.8.4) is the integral of total differential

$$\frac{\partial g(\mathbf{x})}{\partial x_1} dx_1 + \dots + \frac{\partial g(\mathbf{x})}{\partial x_n} dx_n = dg(\mathbf{x}),$$

the value of integral (12.8.4) 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 (12.8.3) to variational problem (12.8.2) and, hence, the proof is complete.

¹²See also Section 12.5.5

[Proof B] Consider variational problem (12.8.2). The corresponding Euler's equations are

$$(12.8.5) \quad \frac{\partial F}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{x}_i} \right) = 0, \quad i \in \{1, \dots, n\}.$$

Now, consider variational problem (12.8.3). The corresponding Euler's equations are

$$\frac{\partial \left[cF + \sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \left[cF + \sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial \dot{x}_i} \right) = 0, \quad i \in \{1, \dots, n\},$$

which we can write as

$$(12.8.6) \quad c \left[\frac{\partial F}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{x}_i} \right) \right] + \frac{\partial \left[\sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \left[\sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial \dot{x}_i} \right) = 0,$$

where $i \in \{1, \dots, n\}$. To prove that solution $\mathbf{x}(t)$ is the same for equations (12.8.5) and (12.8.6), we prove that these two systems of equations are equivalent to one another. Recalling that $c \neq 0$, to prove that equations (12.8.5) and (12.8.6) are equivalent, we need to show that

$$(12.8.7) \quad \frac{\partial \left[\sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial x_i} = \frac{d}{dt} \left(\frac{\partial \left[\sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial \dot{x}_i} \right), \quad i \in \{1, \dots, n\}.$$

Consider the left-hand side of equation (12.8.7). Using the linearity of the differential operator, we can write

$$(12.8.8) \quad \frac{\partial \left[\sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial x_i} = \sum_{j=1}^n \frac{\partial f_j(\mathbf{x})}{\partial x_i} \dot{x}_j, \quad i \in \{1, \dots, n\}.$$

Consider the right-hand side of equation (12.8.7). Using the linearity of the differential operator and taking into account the fact that the only term of $\sum f_j(\mathbf{x}) \dot{x}_j$ that is dependent on \dot{x}_i is the term where $j = i$, we obtain

$$(12.8.9) \quad \frac{d}{dt} \left(\frac{\partial \left[\sum_{j=1}^n f_j(\mathbf{x}) \dot{x}_j \right]}{\partial \dot{x}_i} \right) = \frac{df_i(\mathbf{x})}{dt} = \sum_{j=1}^n \frac{\partial f_i(\mathbf{x})}{\partial x_j} \dot{x}_j, \quad i \in \{1, \dots, n\}.$$

Examining the coefficients of \dot{x}_j in expressions (12.8.8) and (12.8.9), we see that we need to show the equality given by

$$(12.8.10) \quad \frac{\partial f_j(\mathbf{x})}{\partial x_i} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}, \quad i, j \in \{1, \dots, n\}.$$

Recall that $[f_1(\mathbf{x}), \dots, f_n(\mathbf{x})]$ are the components of ∇g , for some function $g(\mathbf{x})$, namely, $[\partial g/\partial x_1, \dots, \partial g/\partial x_n]$.

Thus, we can write the left-hand side of equation (12.8.10) as

$$\frac{\partial f_j(\mathbf{x})}{\partial x_i} = \frac{\partial}{\partial x_i} \left[\frac{\partial g(\mathbf{x})}{\partial x_j} \right] = \frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j}, \quad i, j \in \{1, \dots, n\}.$$

Analogously, we can write the right-hand side of equation (12.8.10) as

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j} = \frac{\partial}{\partial x_j} \left[\frac{\partial g(\mathbf{x})}{\partial x_i} \right] = \frac{\partial^2 g(\mathbf{x})}{\partial x_j \partial x_i}, \quad i, j \in \{1, \dots, n\}.$$

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. \square

EXERCISE 12.4. Consider a variational problem given by integral (12.5.6). 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 12.5.6, provide a geometrical illustration.

SOLUTION 12.4. The variational problem in question is

$$(12.8.11) \quad \delta \int_0^1 \sqrt{1 + \left(\frac{dz}{dx} \right)^2} dx = 0.$$

Since integral (12.8.11) depends only on z' , in view of Section 12.5.3, the extremizing function is a straight line given by $z(x) = x$. Inserting $z = x$ into integral (12.8.11), we obtain the distance along the extremizing function, namely,

$$\int_0^1 \sqrt{2} dx = \sqrt{2},$$

as also expected from Pythagoras' theorem. Now, let us calculate the distance along another curve, for instance, $z(x) = x^2$. Integral (12.8.11) becomes

$$\int_0^1 \sqrt{1 + (2x)^2} dx = \frac{1}{4} \left[2x \sqrt{1 + (2x)^2} + \text{Arcsinh}(2x) \right]_0^1 \approx 1.48,$$

which is greater than $\sqrt{2}$, as expected.

In view of Section 12.5.6, integral (12.8.11) 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 (12.2.2) as the corresponding Hamilton's equations in dz/dx and dp/dx .

SOLUTION 12.5. Consider integrand (12.1.1), namely, $F(z, z'; x)$. In view of Legendre's transformation, discussed in Appendix B, let the variable of transformation be denoted by

$$(12.8.12) \quad p := \frac{\partial F}{\partial z'},$$

and the new function be

$$(12.8.13) \quad H(z, p; x) = pz' - F(z, z'; x),$$

which is the Hamiltonian corresponding to F . Hence, by the duality of Legendre's transformation, we can write

$$(12.8.14) \quad z' = \frac{\partial H}{\partial p}.$$

Invoking Euler's equation (12.2.2) and in view of expression (12.8.12), we can write

$$(12.8.15) \quad \frac{\partial F}{\partial z} = \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = \frac{dp}{dx} \equiv p'.$$

Hence, using expression (12.8.13) to express function F in terms of function H , we obtain

$$(12.8.16) \quad p' = \frac{\partial F}{\partial z} = \frac{\partial}{\partial z} [pz' - H(z, p; x)] = -\frac{\partial H}{\partial z}.$$

Thus, using equations (12.8.14) and (12.8.16), we can write a system of first-order ordinary differential equations in dz/dx and dp/dx , namely,

$$(12.8.17) \quad \begin{cases} z' = \frac{\partial H}{\partial p} \\ p' = -\frac{\partial H}{\partial z} \end{cases},$$

which are the required Hamilton's equations.

EXERCISE 12.6. In view of Exercise 12.5, prove the following theorem.

THEOREM 12.8.3. For an integral given by expression (12.1.1), namely,

$$\int_a^b F[z(x), z'(x); x] dx,$$

if F does not explicitly depend on x , the corresponding Hamiltonian, H , is the first integral of equation (12.2.2).

SOLUTION 12.6. PROOF. We can formally write

$$(12.8.18) \quad \frac{dH(p, z; x)}{dx} = \frac{\partial H}{\partial p} \frac{dp}{dx} + \frac{\partial H}{\partial z} \frac{dz}{dx} + \frac{\partial H}{\partial x}.$$

Invoking equations (12.8.17), the first two terms on the right-hand side of equation (12.8.18) vanish and, hence, this equation is reduced to

$$\frac{dH(p, z; x)}{dx} = \frac{\partial H}{\partial x}.$$

In view of expression (12.8.13), H does not depend on x explicitly if and only if F does not depend on x explicitly. In such a case, we obtain

$$\frac{dH}{dx} = 0,$$

and, hence, H is constant. Using expressions (12.8.12) and (12.8.13), 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 (12.2.2) that is given by expression (12.5.4). □

CHAPTER 13

Variational principles

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 11, 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 of stationary action.

13.1. Fermat's principle

NOTATION 13.1.1. 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. ¹In 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 13.1.2. *Rays are the solutions of the variational problem*

$$(13.1.1) \quad \delta \int_A^B \frac{ds}{V(\mathbf{x}, \mathbf{n})} = 0,$$

where ds is an arclength element and $V(\mathbf{x}, \mathbf{n})$ is the ray velocity in direction $\mathbf{n} = d\mathbf{x}/ds$ at point \mathbf{x} . A and B are the fixed endpoints of this variational problem.

Note that, in expression (13.1.1) 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 13.1.2, we must show that the solution of variational problem (13.1.1) is equivalent to the solution of Hamilton's ray equations (8.2.7); namely,

$$(13.1.2) \quad \begin{cases} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \end{cases}, \quad i \in \{1, \dots, n\}.$$

Let us investigate the properties of the Hamiltonian that is given by expression (8.2.8); 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

$$(13.1.3) \quad \mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} p^2 v^2 \left(\mathbf{x}, \frac{\mathbf{p}}{|\mathbf{p}|} \right),$$

where $|\mathbf{p}|$ is the magnitude of the phase-slowness vector.

By examining expression (13.1.3), 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 13.1.3. *Hamiltonian $\mathcal{H}(\mathbf{x}, \mathbf{p})$, given by expression (13.1.3), is conserved along the ray.*

¹This section is based on the work that was published by Bóna, A., and Slawinski, M.A., (2003) Fermat's principle for seismic rays in elastic media. *Journal of Applied Geophysics* **54**, 445 – 451.

PROOF. Differentiating $\mathcal{H}(\mathbf{x}, \mathbf{p})$ with respect to t , we get

$$\frac{d\mathcal{H}}{dt} = \sum_{i=1}^n \frac{\partial \mathcal{H}}{\partial x_i} \dot{x}_i + \sum_{i=1}^n \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i + \frac{\partial \mathcal{H}}{\partial t}.$$

Since \mathcal{H} does not explicitly depend on time, we write

$$\frac{d\mathcal{H}}{dt} = \sum_{i=1}^n \frac{\partial \mathcal{H}}{\partial x_i} \dot{x}_i + \sum_{i=1}^n \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i.$$

Using system (13.1.2), we obtain

$$\frac{d\mathcal{H}}{dt} = - \sum_{i=1}^n \dot{p}_i \dot{x}_i + \sum_{i=1}^n \dot{x}_i \dot{p}_i = 0,$$

as required. \square

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 (7.3.8), must be satisfied along the rays. Hence, in view of this equation, which states that $p^2 v^2 = 1$, and expression (8.2.8), we require that

$$(13.1.4) \quad \mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2}$$

along a ray.

13.1.3. Variational equivalent of Hamilton's ray equations. To show that rays obtained from Hamilton's ray equations (13.1.2) are solutions of variational problem (13.1.1), we express these equations in the context of the calculus of variations. As stated in Section 12.7, Lagrange's ray equations (11.2.6); namely,

$$(13.1.5) \quad \frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = 0, \quad i \in \{1, \dots, n\},$$

possess the form of Euler's equations. Consequently, in view of Chapter 12, we can state the following proposition.

PROPOSITION 13.1.4. *Rays are the solutions of the variational problem*

$$(13.1.6) \quad \delta \int \mathcal{L} dt = 0,$$

where the ray-theory Lagrangian \mathcal{L} is given by expression (11.1.1), namely

$$(13.1.7) \quad \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \sum_{j=1}^n p_j(\mathbf{x}, \dot{\mathbf{x}}) \dot{x}_j - \mathcal{H}(\mathbf{x}, \mathbf{p}).$$

13.1.4. Properties of Lagrangian \mathcal{L} . To examine variational formulations (13.1.1) and (13.1.6), we must study the properties of Lagrangian \mathcal{L} , given by expression (13.1.7), in terms of the corresponding Hamiltonian, \mathcal{H} . We begin by stating the following lemma.

LEMMA 13.1.5. *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.2.1, stated in Appendix A, we obtain

$$\mathcal{L}(\mathbf{x}(t), \dot{\mathbf{x}}(\mathbf{x}(t), \mathbf{p}(t))) = 2\mathcal{H} - \mathcal{H} = \mathcal{H},$$

which completes the proof. \square

In view of the conserved value of Hamiltonian \mathcal{H} , as shown in Lemma 13.1.3, and following expression (13.1.4), we obtain the following corollary of Lemma 13.1.5.

COROLLARY 13.1.6. *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 13.1.7. *If Hamiltonian $\mathcal{H}(\mathbf{x}, \mathbf{p})$ is homogeneous of degree 2 in the p_i , then Lagrangian $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ is homogeneous of degree 2 in the \dot{x}_i .*

PROOF. By Lemma 13.1.5, $\mathcal{H}(\mathbf{x}, \mathbf{p}) = \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}(\mathbf{x}, \mathbf{p}))$, 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

$$\dot{x}'_i = \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p}')}{\partial p'_i} = \frac{\partial \mathcal{H}(\mathbf{x}, a\mathbf{p})}{\partial (ap_i)}, \quad i \in \{1, \dots, n\}.$$

By the homogeneity of \mathcal{H} and the property of the differential operator, we can write

$$\dot{x}'_i = \frac{\partial \mathcal{H}(\mathbf{x}, a\mathbf{p})}{\partial (ap_i)} = \frac{a^2 \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial p_i}}{\frac{\partial (ap_i)}{\partial p_i}} = \frac{a^2 \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial p_i}}{a}, \quad i \in \{1, \dots, n\}.$$

Hence,

$$\dot{x}'_i = a \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial p_i}, \quad i \in \{1, \dots, n\},$$

which, in view of Hamilton's ray equations, given by system (13.1.2), can be stated as

$$\dot{x}'_i = a\dot{x}_i, \quad i \in \{1, \dots, n\}.$$

Consequently, we can write

$$\mathcal{L}(\mathbf{x}, a\dot{\mathbf{x}}) = \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}'),$$

which, by Lemma 13.1.5, yields

$$\mathcal{L}(\mathbf{x}, a\dot{\mathbf{x}}) = \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}') = \mathcal{H}(\mathbf{x}, \mathbf{p}') = \mathcal{H}(\mathbf{x}, a\mathbf{p}) = a^2 \mathcal{H}(\mathbf{x}, \mathbf{p}) = a^2 \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}),$$

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 . \square

Lemma 13.1.7 implies that variational problem (13.1.6) 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 (13.1.2) are parametrized by time; hence, the solutions of system (13.1.5) 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 13.1.6.

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 (13.1.1), is parametrization independent. This results from the fact that the integrand in expression (13.1.1) is homogeneous of degree 1 in the \dot{x}_i .

Let us consider a Lagrangian given by

$$(13.1.8) \quad \mathcal{F} = \sqrt{2\mathcal{L}},$$

where \mathcal{L} is given by expression (13.1.7). Note that, following Definition A.1.1, 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 (13.1.5) are also the solutions of the equations given by

$$(13.1.9) \quad \frac{\partial \mathcal{F}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{F}}{\partial \dot{x}_i} \right) = 0, \quad i \in \{1, \dots, n\}.$$

This is stated by the following lemma.

LEMMA 13.1.8. *A solution of equations (13.1.5) that satisfies the condition given in Corollary 13.1.6, where \mathcal{L} is given by expression (13.1.7), is also a solution of equations (13.1.9), where $\mathcal{F} = \sqrt{2\mathcal{L}}$.*

PROOF. Inserting $\mathcal{L} = \mathcal{F}^2/2$ into equations (13.1.5), we obtain

$$\begin{aligned} \frac{\partial}{\partial x_i} \left(\frac{\mathcal{F}^2}{2} \right) - \frac{d}{dt} \left[\frac{\partial}{\partial \dot{x}_i} \left(\frac{\mathcal{F}^2}{2} \right) \right] &= \mathcal{F} \frac{\partial \mathcal{F}}{\partial x_i} - \frac{d}{dt} \left[\mathcal{F} \frac{\partial \mathcal{F}}{\partial \dot{x}_i} \right] \\ &= \mathcal{F} \left[\frac{\partial \mathcal{F}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{F}}{\partial \dot{x}_i} \right) \right] - \frac{d\mathcal{F}}{dt} \frac{\partial \mathcal{F}}{\partial \dot{x}_i} \\ &= 0, \quad i \in \{1, \dots, n\}. \end{aligned}$$

Since $\mathcal{L} = 1/2$ along a ray, as shown in Corollary 13.1.6, then $\mathcal{F} = 1$ and, hence, $d\mathcal{F}/dt = 0$ along the solutions of equations (13.1.5). Thus, equations (13.1.5) become equations (13.1.9), as required. \square

Consequently, equations (13.1.9) can also be viewed as Lagrange's ray equations.

If we can show that

$$(13.1.10) \quad \mathcal{F} = \frac{|\dot{\mathbf{x}}|}{V(\mathbf{x}, \frac{\dot{\mathbf{x}}}{|\dot{\mathbf{x}}|})},$$

where $|\dot{\mathbf{x}}| = ds/dt$ and $\dot{\mathbf{x}}/|\dot{\mathbf{x}}| = \mathbf{n}$, then we prove Theorem 13.1.2, since the right-hand side of equation (13.1.10) is the integrand of equation (13.1.1).

13.1.6. Ray velocity. In order to show that the right-hand side of equation (13.1.10) is the integrand of equation (13.1.1), we must formulate ray velocity in a variational context. Since, as shown in Lemma 13.1.7, Lagrangian \mathcal{L} is homogeneous 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 13.1.6, 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

$$(13.1.11) \quad V(\mathbf{x}, \mathbf{n}) := |\dot{\mathbf{x}}| = \frac{1}{\sqrt{2\mathcal{L}(\mathbf{x}, \mathbf{n})}},$$

which is consistent with expression (8.4.3).

Now, we are ready to complete the proof of Theorem 13.1.2.

13.1.7. Proof of Fermat's principle.

PROOF. By Lemma 13.1.8, rays are the solutions of Euler's equations stated in system (13.1.9). Consequently, rays are the solutions of variational problem

$$(13.1.12) \quad \delta \int \mathcal{F} dt = 0.$$

In view of expression (13.1.8), we can restate this variational problem as

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = \delta \int \sqrt{2\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})} dt = 0.$$

Since, as stated in Lemma 13.1.7, \mathcal{L} is homogeneous of degree 2 in the \dot{x}_i , we can write

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = \delta \int \sqrt{2|\dot{\mathbf{x}}|^2 \mathcal{L}(\mathbf{x}, \mathbf{n})} dt = \delta \int |\dot{\mathbf{x}}| \sqrt{2\mathcal{L}(\mathbf{x}, \mathbf{n})} dt = 0.$$

In view of expressions (13.1.11) and since $|\dot{\mathbf{x}}| dt = ds$, we conclude that

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = \delta \int \frac{ds}{V(\mathbf{x}, \mathbf{n})} = 0.$$

Hence, the solutions of Hamilton's ray equations that correspond to rays are the solutions of variational problem (13.1.1). \square

Theorem 13.1.2 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,

$$\det \left[\frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j} \right] \neq 0, \quad i, j \in \{1, \dots, n\}.$$

In other words, we are unable to prove Theorem 13.1.2 at the inflection points of the phase-slowness surface. As stated in Theorem 9.3.1, for an elastic continuum defined by constitutive equations (3.2.1), 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 13.1.2 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. Hamilton's principle: Example

13.2.1. 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.² In this section, however, we will illustrate the analogy with classical mechanics by deriving the wave equation using Hamilton's variational principle.

13.2.2. 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*³. In the standard formulation of classical mechanics, *vis viva*, which underlies the concept of action, can be viewed as twice the kinetic energy.

The commonly accepted definition of action is

$$(13.2.1) \quad \mathbb{A} := \int_{t_1}^{t_2} \mathbb{L} \, dt,$$

where \mathbb{L} is the classical-mechanics Lagrangian that is defined by

$$(13.2.2) \quad \mathbb{L} := T - U,$$

with T and U denoting the kinetic energy and the potential energy, respectively.⁴

²Interested readers might refer to Basdevant, J-L., (2007) Variational principles in physics: Springer-Verlag, where the author discusses these principle in such branches of physics as thermodynamics and quantum mechanics.

³*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. Readers interested in the principle of *vis viva* as seen by early mechanicians might refer to Kuhn, T.S., (1996) The structure of scientific revolutions (3rd edition): The University of Chicago Press, pp. 189 – 191.

⁴Readers interested in developments of the definition of action might refer to Ekeland, I., (2000) Le meilleur des mondes possibles: Mathématiques et destinée: Seuil, pp. 57 – 98, or — for the English version — to Ekeland, I., (2006) The best of all possible worlds: Mathematics and destiny, pp. 44 – 78.

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*⁵, 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.⁶

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.⁷

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 (13.2.1), finding a stationary value of action is equivalent to variational problem

$$(13.2.3) \quad \delta \mathbb{A} = \delta \int_{t_1}^{t_2} \mathbb{L} dt = 0.$$

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 13.2.4. 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.

⁵Agreement of various laws of nature which until now appeared incompatible

⁶Action 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.

⁷Hamilton, W.R., (1833) On a general method of expressing the paths of light, and of the planets, by the coefficients of a characteristic function: Dublin University Review

13.2.3. 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 13.2.4.

Considering Hamilton's principle, stated in equation (13.2.3), 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

$$(13.2.4) \quad \frac{\partial \mathbb{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathbb{L}}{\partial \dot{x}_i} \right) = 0, \quad i \in \{1, 2, 3\},$$

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 (13.2.4) 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 (13.2.4) 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 (13.2.4) and recalling that T is assumed to be a function of velocity alone, we obtain

$$\frac{\partial \mathbb{L}}{\partial x_i} = \frac{\partial [T(\dot{\mathbf{x}}) - U(\mathbf{x})]}{\partial x_i} = -\frac{\partial U}{\partial x_i} =: F_i, \quad i \in \{1, 2, 3\},$$

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 (13.2.4) and recalling that U is assumed to be a function of position alone, we obtain

$$(13.2.5) \quad \frac{\partial \mathbb{L}}{\partial \dot{x}_i} = \frac{\partial [T(\dot{\mathbf{x}}) - U(\mathbf{x})]}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} =: p_i, \quad i \in \{1, 2, 3\},$$

which is the expression for a component of momentum.

Since the first term of equations (13.2.4) 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 (13.2.4) are equivalent to Newton's second law of motion; namely,

$$F_i - \frac{dp_i}{dt} = 0, \quad i \in \{1, 2, 3\}.$$

Also, as shown in Exercises 13.6 and 13.7, we can derive Hamilton's equations of motion from Newton's laws of motion.

To gain further insight into Lagrange's equations of motion, let us invoke Beltrami's identity (11.3.1) to write

$$\frac{\partial \mathbb{L}}{\partial t} + \frac{d}{dt} \left(\sum_{i=1}^n \dot{x}_i \frac{\partial \mathbb{L}}{\partial \dot{x}_i} - \mathbb{L} \right) = 0.$$

If \mathbb{L} does not depend explicitly on time — a concept used in Exercise 13.2, below, in the context of ray theory — this equation becomes

$$\frac{d}{dt} \left(\sum_{i=1}^n \dot{x}_i \frac{\partial \mathbb{L}}{\partial \dot{x}_i} - \mathbb{L} \right) = 0,$$

which implies that the term in parentheses is constant, C .⁸ If \mathbb{L} is homogeneous of degree 2 in the \dot{x}_i , in view of Theorem A.2.1 from Appendix A, we obtain $2\mathbb{L} - \mathbb{L} = C$. Thus, Lagrangian \mathbb{L} is conserved along trajectory $\mathbf{x}(t)$. This constant of motion can be viewed also as the conserved energy of the system. To see that — in view of expression (13.2.5), and expression (B.3.1) from Appendix B — we write the conserved energy as the classical-mechanics Hamiltonian, namely,

$$\mathbb{H} = \sum_{i=1}^n \dot{x}_i p_i - \mathbb{L}.$$

Herein, $\mathbb{H} = T + U$: the sum of kinetic and potential energies. We illustrate the relation between this expression and the definition of \mathbb{L} stated by expression (13.2.2) in Exercise 13.5.

Lagrange's equations of motion (13.2.4) 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.⁹

13.2.4. Wave equation.

Continuous systems and Lagrangian density. A seismological application of stationary-action principle (13.2.3) and, consequently, of Lagrange's equations of motion for elastic continua, is exemplified by the derivation of the wave equation.¹⁰ 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

$$(13.2.6) \quad \mathbb{L} = \iiint \mathfrak{L} dx_1 dx_2 dx_3,$$

where \mathfrak{L} is the Lagrangian density

$$\mathfrak{L} = \mathfrak{L} \left(u, \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial t}, x_i, t \right), \quad i = \{1, 2, 3\}.$$

⁸Readers interested in a different derivation of this conserved quantity might refer to Basdevant, J-L., (2007) Variational principles in physics: Springer-Verlag, pp. 54 – 55.

⁹Readers 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.

¹⁰Readers 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.

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, \dots, 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

$$(13.2.7) \quad F = \frac{k}{\Delta x} u,$$

where k denotes a constant.

Note that the term $k/\Delta x$ has the units of $[N/m]$ and expression (13.2.7) 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. Following equation (4.5.1), we write the potential energy of a segment of the string as

$$U_i = \int_0^{\Delta u} F du = \frac{k}{\Delta x} \int_0^{\Delta u} u du = \frac{1}{2} \frac{k}{\Delta x} (\Delta u)^2,$$

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

$$(13.2.8) \quad U = \sum_{i=1}^n U_i = \frac{1}{2} k \sum_{i=1}^n \frac{(u_i - u_{i-1})^2}{\Delta x}.$$

If $n \rightarrow \infty$ and $\Delta x \rightarrow 0$ in such a way that $(n+1)\Delta x = l$, the potential energy, U , can be written as

$$U = \frac{1}{2} k \sum_{i=1}^{\infty} \left[\frac{u(x_i, t) - u(x_{i-1}, t)}{x_i - x_{i-1}} \right]^2 (x_i - x_{i-1}),$$

where $x_i - x_{i-1} \equiv \Delta x$. 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

$$(13.2.9) \quad U = \frac{1}{2} k \int_0^l \left[\frac{\partial u(x, t)}{\partial x} \right]^2 dx.$$

The kinetic energy, T , for the entire string containing n discrete mass points, each of which has a mass m , is

$$(13.2.10) \quad T = \frac{1}{2} m \sum_{i=1}^n \left[\frac{\partial u(x_i, t)}{\partial t} \right]^2.$$

Rearranging and using the limit, we can write

$$(13.2.11) \quad T = \frac{1}{2} \frac{m}{\Delta x} \sum_{i=1}^{\infty} \left[\frac{\partial u(x_i, t)}{\partial t} \right]^2 \Delta x = \frac{1}{2} \rho \int_0^l \left[\frac{\partial u(x, t)}{\partial t} \right]^2 dx,$$

where $\rho := \lim_{\Delta x \rightarrow 0} m/\Delta x$, is the mass density of the one-dimensional continuum.

Since the kinetic energy, given in expression (13.2.11), is not a function of position, we can invoke the classical-mechanics Lagrangian, given by expression (13.2.2). Thus, using expressions (13.2.9) and (13.2.11), we can write

$$(13.2.12) \quad \begin{aligned} \mathbb{L}(x, t) &= T(\dot{x}) - U(x) \\ &= \frac{1}{2} \rho \int_0^l \left[\frac{\partial u(x, t)}{\partial t} \right]^2 dx - \frac{1}{2} k \int_0^l \left[\frac{\partial u(x, t)}{\partial x} \right]^2 dx \\ &= \int_0^l \left\{ \frac{\rho}{2} \left[\frac{\partial u(x, t)}{\partial t} \right]^2 - \frac{k}{2} \left[\frac{\partial u(x, t)}{\partial x} \right]^2 \right\} dx. \end{aligned}$$

Since we are presently dealing with a one-dimensional continuum, considering expression (13.2.6), we can write

$$(13.2.13) \quad \mathbb{L} = \int_0^l \mathfrak{L} dx,$$

where, in view of integral (13.2.12), \mathfrak{L} is the Lagrangian density given by

$$(13.2.14) \quad \mathfrak{L} \equiv \frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2 - \frac{k}{2} \left(\frac{\partial u}{\partial x} \right)^2.$$

To invoke a variational formulation, in view of expression (13.2.13) and following equation (13.2.3), we can write

$$\delta \int_0^t \mathbb{L} dt = \delta \int_0^t \int_0^l \mathfrak{L} dx dt = 0.$$

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 (12.4.1), we can write the stationarity condition as

$$(13.2.15) \quad \frac{\partial \mathfrak{L}}{\partial u} - \left[\frac{\partial}{\partial t} \left(\frac{\partial \mathfrak{L}}{\partial u_t} \right) + \frac{\partial}{\partial x} \left(\frac{\partial \mathfrak{L}}{\partial u_x} \right) \right] = 0,$$

where $u_t := \partial u / \partial t$ and $u_x := \partial u / \partial x$.

Equation (13.2.15) is Lagrange's equation of motion for a one-dimensional continuum.

Inserting the Lagrangian density, stated in expression (13.2.14), into equation (13.2.15) 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 $[k] = [kg\ m/s^2]$ 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 $\iint \mathcal{L} dx dt$ stationary. This illustrates the fact that a solution of Euler's equation involving two variables is a surface, as stated in Section 12.4.2.

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 (8.2.7) — 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 9.3.1, 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.¹¹



¹¹Interested 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.

13.3. Exercises

EXERCISE 13.1. Consider a traveltime integral in an anisotropic inhomogeneous continuum; namely,

$$(13.3.1) \quad \check{C} = \int_a^b \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt.$$

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,

$$ds = \frac{df}{dt} dt =: \dot{f} dt,$$

and

$$(13.3.2) \quad \dot{f} = \frac{ds}{dt}.$$

Consider $\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}})$, where

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{ds} \frac{ds}{dt},$$

which, in view of expression (13.3.2), can be written as

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{ds} \frac{df}{dt} =: \mathbf{x}' \dot{f}.$$

For the value of the integral (13.3.1) to be independent of parametrization, we require

$$(13.3.3) \quad \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = \mathcal{F}(\mathbf{x}, \mathbf{x}') ds.$$

Consider the left-hand side of equation (13.3.3). Since $\dot{\mathbf{x}} = \mathbf{x}' \dot{f}$ and $dt = ds/\dot{f}$, we can write it as

$$\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = \mathcal{F}(\mathbf{x}, \mathbf{x}' \dot{f}) \frac{ds}{\dot{f}}.$$

If \mathcal{F} is homogeneous of degree 1 in $\dot{\mathbf{x}}$, following Definition A.1.1, stated in Appendix A, we obtain

$$\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = \dot{f} \mathcal{F}(\mathbf{x}, \mathbf{x}') \frac{ds}{\dot{f}} = \mathcal{F}(\mathbf{x}, \mathbf{x}') ds,$$

which is equation (13.3.3), as required.

REMARK 13.3.1. Exercise 13.1 shows that the general statement of Fermat's principle, namely,

$$\delta \int \mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = 0,$$

is independent of parametrization. This is the case since $\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}}) dt = ds/V$ is homogeneous of degree 1 in $\dot{\mathbf{x}}$. Note that ds is homogeneous of degree 1 in $\dot{\mathbf{x}}$, while V is homogeneous of degree 0 in $\dot{\mathbf{x}}$.¹²

¹²See also Section 11.3

EXERCISE 13.2. In view of Lemma 13.1.7, use Beltrami's identity (11.3.1), namely,

$$(13.3.4) \quad \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\sum_{i=1}^n \dot{x}_i \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \mathcal{L} \right) = 0,$$

to show that Lagrangian \mathcal{L} is conserved along the ray.

SOLUTION 13.2. Since \mathcal{L} , given by expression (13.1.7), does not explicitly depend on t , equation (13.3.4) becomes $d(\sum_{i=1}^n \dot{x}_i \partial \mathcal{L} / \partial \dot{x}_i - \mathcal{L}) / dt = 0$, 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 13.1.7, \mathcal{L} is homogeneous of degree 2 in the \dot{x}_i , in view of Theorem A.2.1, 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 8.8 and 8.9; namely,

$$(13.3.5) \quad \begin{cases} \frac{dx_i}{dt} = v^2 p_i \\ \frac{dp_i}{dt} = -\frac{\partial \ln v}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Express system (13.3.5) as three second-order equations.

SOLUTION 13.3. Solving the first equation of system (13.3.5) for the components of the phase-slowness vector, we get

$$p_i = \frac{1}{v^2(\mathbf{x})} \frac{dx_i}{dt}, \quad i \in \{1, 2, 3\}.$$

Differentiating with respect to t , we can write

$$\frac{dp_i}{dt} = \frac{d}{dt} \left[\frac{1}{v^2(\mathbf{x})} \frac{dx_i}{dt} \right], \quad i \in \{1, 2, 3\},$$

which we can equate to the second equation of set (13.3.5) to obtain

$$(13.3.6) \quad \frac{\partial \ln v}{\partial x_i} + \frac{d}{dt} \left[\frac{1}{v^2(\mathbf{x})} \frac{dx_i}{dt} \right] = 0, \quad i \in \{1, 2, 3\},$$

as required.

EXERCISE 13.4. Consider the traveltime integral in an isotropic inhomogeneous continuum. Show that equations (13.3.6) 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 $\dot{\mathbf{x}} \equiv d\mathbf{x}/dt$. We invoke equations (13.1.9), namely,

$$(13.3.7) \quad \frac{\partial \mathcal{F}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{F}}{\partial \dot{x}_i} \right) = 0, \quad i \in \{1, 2, 3\},$$

which are a parametric form of Euler's equations. Considering integrand \mathcal{F} and the first term of equations (13.3.7), we obtain

$$\frac{\partial \mathcal{F}}{\partial x_i} = -\frac{1}{V^2(\mathbf{x})} \frac{\partial V(\mathbf{x})}{\partial x_i} \sqrt{\sum_{i=1}^3 \dot{x}_i \dot{x}_i}, \quad i \in \{1, 2, 3\}.$$

Using the fact that t denotes time, and, hence, as shown in expression (8.4.3),

$$(13.3.8) \quad \sqrt{\sum_{i=1}^3 \dot{x}_i \dot{x}_i} =: V(\mathbf{x}),$$

where V is the magnitude of ray velocity, we can write

$$\frac{\partial \mathcal{F}}{\partial x_i} = -\frac{1}{V(\mathbf{x})} \frac{\partial V(\mathbf{x})}{\partial x_i}, \quad i \in \{1, 2, 3\}.$$

Using the chain rule, we can rewrite this expression as

$$(13.3.9) \quad \frac{\partial \mathcal{F}}{\partial x_i} = -\frac{\partial}{\partial x_i} \ln V(\mathbf{x}), \quad i \in \{1, 2, 3\}.$$

Considering integrand \mathcal{F} and the second term of equations (13.3.7), we obtain

$$\frac{\partial \mathcal{F}}{\partial \dot{x}_i} = \frac{1}{V(\mathbf{x})} \frac{\dot{x}_i}{\sqrt{\sum_{i=1}^3 \dot{x}_i \dot{x}_i}}, \quad i \in \{1, 2, 3\},$$

which, using expression (13.3.8), we can write as

$$(13.3.10) \quad \frac{\partial \mathcal{F}}{\partial \dot{x}_i} = \frac{1}{V^2(\mathbf{x})} \frac{dx_i}{dt}, \quad i \in \{1, 2, 3\}.$$

Consequently, using expressions (13.3.9) and (13.3.10), we can write Euler's equations (13.3.7) as

$$(13.3.11) \quad \frac{\partial \ln V}{\partial x_i} + \frac{d}{dt} \left[\frac{1}{V^2(\mathbf{x})} \frac{dx_i}{dt} \right] = 0, \quad i \in \{1, 2, 3\}.$$

Since in isotropic continua, phase and ray velocities coincide, namely, $V \equiv v$, equations (13.3.11) are equivalent to equations (13.3.6), as required.

REMARK 13.3.2. 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.¹³

¹³Readers 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.

EXERCISE 13.5. Recall the classical-mechanics Lagrangian given in expression (13.2.2), namely,

$$(13.3.12) \quad \mathbb{L} := T - U,$$

where T and U are the kinetic and potential energies. Let the classical-mechanics Hamiltonian be

$$(13.3.13) \quad \mathbb{H} := T + U.$$

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

$$\mathbb{H} = \sum_{i=1}^n p_i \dot{x}_i - \mathbb{L},$$

where $p_i = m v_i$, with v_i being a component of velocity given by $v_i = dx_i/dt \equiv \dot{x}_i$. Hence, we can write

$$(13.3.14) \quad \mathbb{H} = m \sum_{i=1}^n \dot{x}_i^2 - \mathbb{L} = m v^2 - \mathbb{L},$$

where v stands for the magnitude of velocity. Recalling definitions (13.3.12) and (13.3.13), we can write expression (13.3.14) as

$$T + U = m v^2 - (T - U),$$

where T and U are the kinetic and potential energies, respectively. Simplifying, we obtain

$$T = \frac{1}{2} m v^2,$$

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,

$$(13.3.15) \quad m \frac{d^2 x_i}{dt^2} = - \frac{\partial U(\mathbf{x})}{\partial x_i}, \quad i \in \{1, 2, 3\},$$

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

$$(13.3.16) \quad p_i := m \frac{dx_i}{dt}, \quad i \in \{1, 2, 3\}.$$

Differentiating both sides of equations (13.3.16) with respect to t , we obtain

$$\frac{dp_i}{dt} = m \frac{d^2 x_i}{dt^2}, \quad i \in \{1, 2, 3\},$$

which are equations (13.3.15). Hence, Newton's second law of motion can be written as a set of two first-order differential equations,

$$(13.3.17) \quad \begin{cases} \frac{dx_i}{dt} = \frac{p_i}{m} \\ \frac{dp_i}{dt} = -\frac{\partial U}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

EXERCISE 13.7. ¹⁴Using expression (13.3.13), show that equations (13.3.17), obtained in Exercise 13.6, correspond to Hamilton's equations of motion that are given by

$$\begin{cases} \frac{dx_i}{dt} = \frac{\partial \mathbb{H}}{\partial p_i} \\ \frac{dp_i}{dt} = -\frac{\partial \mathbb{H}}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

SOLUTION 13.7. Consider expression (13.3.13). Using the standard expression for the kinetic energy, as well as the definition of linear momentum, we can write this expression as

$$(13.3.18) \quad \mathbb{H} = T + U = \frac{1}{2}m \left(\frac{dx_i}{dt} \right)^2 + U = \frac{1}{2m}p_i^2 + U, \quad i \in \{1, 2, 3\}.$$

Differentiating equations (13.3.18) with respect to both the p_i and the x_i , we obtain

$$\begin{cases} \frac{\partial \mathbb{H}}{\partial p_i} = \frac{p_i}{m} \\ \frac{\partial \mathbb{H}}{\partial x_i} = \frac{\partial U}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Using Newton's second law of motion, which is stated in expression (13.3.17), we obtain

$$\begin{cases} \frac{dx_i}{dt} = \frac{\partial \mathbb{H}}{\partial p_i} \\ \frac{dp_i}{dt} = -\frac{\partial \mathbb{H}}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\},$$

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 $v = dz/dt$, we can write the classical-mechanics Lagrangian as

$$\mathbb{L} = T - U = \frac{1}{2}m \left(\frac{dz}{dt} \right)^2 - mgz.$$

¹⁴See also Section 13.2.3

Consider the action integral

$$\mathbb{A} = \int_{t_0}^{t_1} \mathbb{L} \, dt.$$

Hamilton's principle implies

$$\delta \int_{t_0}^{t_1} (T - U) \, dt = \delta \int_{t_0}^{t_1} \left(\frac{1}{2} m (\dot{z})^2 - mgz \right) dt = 0,$$

where $\dot{z} := dz/dt$. Invoking Euler's equation, which corresponds to Lagrange's equations of motion (13.2.4), we obtain

$$\frac{\partial \mathbb{L}}{\partial z} - \frac{d}{dt} \left(\frac{\partial \mathbb{L}}{\partial \dot{z}} \right) = -mg - \frac{d}{dt} (m\dot{z}) = -mg - m\ddot{z} = 0,$$

which can be written as

$$\frac{d^2 z(t)}{dt^2} = -g,$$

where g is the free-fall acceleration, as required.

EXERCISE 13.9. *Following Section 13.2.4, derive a one-dimensional wave equation for transverse waves.*

SOLUTION 13.9. Let the transverse displacements of masses be u_0, \dots, 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

$$(13.3.19) \quad dU = \mu \left[\sqrt{(\Delta x)^2 + (u_i - u_{i-1})^2} - \Delta x \right],$$

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 (13.3.19) as

$$dU = \mu \Delta x \left[\sqrt{1 + \left(\frac{u_i - u_{i-1}}{\Delta x} \right)^2} - 1 \right].$$

Expanding the square root as a power series gives us

$$dU = \mu \Delta x \left[\frac{1}{2} \left(\frac{u_i - u_{i-1}}{\Delta x} \right)^2 - \frac{1}{8} \left(\frac{u_i - u_{i-1}}{\Delta x} \right)^4 + \dots \right].$$

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

$$dU \approx \frac{\mu}{2} \Delta x \left(\frac{u_i - u_{i-1}}{\Delta x} \right)^2.$$

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 $n \rightarrow \infty$ and $\Delta x \rightarrow 0$ 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

$$U = \frac{1}{2}\mu \sum_{i=1}^{\infty} \left[\frac{\partial u(x_i, t)}{\partial x} \right]^2 \Delta x.$$

Thus, in the limit, we obtain

$$U = \frac{1}{2}\mu \int_0^l \left[\frac{\partial u(x, t)}{\partial x} \right]^2 dx.$$

The kinetic energy is given by expression (13.2.11); namely,

$$T = \frac{1}{2}\rho \int_0^l \left[\frac{\partial u(x, t)}{\partial t} \right]^2 dx.$$

Thus, using \mathbb{L} , given in expression (13.2.2), we can write

$$\mathbb{L} = T - U = \int_0^l \left[\frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2 - \frac{\mu}{2} \left(\frac{\partial u}{\partial x} \right)^2 \right] dx.$$

Invoking Hamilton's principle, stated in expression (13.2.3), we obtain

$$\delta \int_0^t \mathbb{L} dt = \delta \int_0^t \int_0^l \left[\frac{\rho}{2} \left(\frac{\partial u}{\partial t} \right)^2 - \frac{\mu}{2} \left(\frac{\partial u}{\partial x} \right)^2 \right] dx dt = 0.$$

Using the corresponding Euler's equation (12.4.1), we get

$$(13.3.20) \quad \frac{\partial^2 u}{\partial x^2} = \frac{1}{\frac{\mu}{\rho}} \frac{\partial^2 u}{\partial t^2}.$$

Equation (13.3.20) 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 .

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 11, 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 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.

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.

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

$$(14.1.1) \quad \check{C} = \int_A^B \frac{\sqrt{1 + (z')^2}}{V(x, z, z')} dx =: \int_A^B F dx,$$

where $z' := dz/dx$. Since $dz/dx = \cot \theta$, 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 (14.1.1) 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 (14.1.1) is of the type given by integral (12.1.1), in general, we can obtain such a curve using Euler's equation (12.2.2). 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 (14.1.1) that are pertinent to seismic investigations.

14.2. Ray parameters as first integrals²

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' , traveltime integral (14.1.1) becomes

$$(14.2.1) \quad \check{C} = \int_A^B \frac{\sqrt{1 + (z')^2}}{V(z, z')} dx \equiv \int_A^B F(z, z') dx.$$

Since traveltime integral (14.2.1) does not exhibit an explicit dependence on x , to obtain the ray, we use Beltrami's identity (12.3.1), namely,

$$\frac{\partial F}{\partial x} + \frac{d}{dx} \left(z' \frac{\partial F}{\partial z'} - F \right) = 0,$$

which immediately leads to

$$(14.2.2) \quad z' \frac{\partial F}{\partial z'} - F = C,$$

where C is a constant. Expression (14.2.2) is first integral (12.5.4) and C is a conserved quantity along the ray.

²This 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.

We wish to study this conserved quantity. Inserting integrand F , given in integral (14.2.1), into expression (14.2.2), we obtain

$$(14.2.3) \quad z' \frac{\partial \left(\frac{\sqrt{1+(z')^2}}{V(z, z')} \right)}{\partial z'} - \frac{\sqrt{1+(z')^2}}{V(z, z')} = -\frac{V' z' \sqrt{1+(z')^2}}{V^2} - \frac{1}{V \sqrt{1+(z')^2}} = C,$$

where, for convenience, we denote $V := V(z, z')$ and $V' := \partial V / \partial z'$. The chain rule implies $-V' / V^2 = \partial [1/V(z, z')] / \partial z'$ and, hence, we get

$$(14.2.4) \quad \frac{\partial}{\partial z'} \left(\frac{1}{V} \right) z' \sqrt{1+(z')^2} - \frac{1}{V \sqrt{1+(z')^2}} = C.$$

Expression (14.2.4) is a first integral of Euler's equation (12.2.2) for traveltime integral (14.2.1).

In order to express the first integral in terms of the ray angle, we use $z' \equiv dz/dx = \cot \theta$, where θ is the ray angle measured from the z -axis. Also, the differential operator in expression (14.2.4) can be restated as $\partial / \partial z' = (\partial \theta / \partial z') \partial / \partial \theta$. Hence, using trigonometric identities, we obtain another form of expression (14.2.4) given by

$$(14.2.5) \quad p = \cos \theta \frac{\partial}{\partial \theta} \left[\frac{1}{V(z, \theta)} \right] + \frac{\sin \theta}{V(z, \theta)},$$

where $p = -C$ and where V and θ denote ray velocity and ray angle, respectively. Since p is conserved along a given ray, $z(x)$, we refer to this conserved quantity as ray parameter. Expression (14.2.5) is the ray parameter for anisotropic vertically inhomogeneous continua.

For expression (14.2.5) 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: Elliptical anisotropy and linear inhomogeneity

14.3.1. 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. Also, the same case was treated in Section 8.5 in the context of Hamilton's ray equations. Thus, our examination of Sections 8.5 and 14.3 will allow us to investigate the same physical problem using the two different approaches that are available to study seismic ray theory.

Since Euler's equation (12.2.2), or its Beltrami's identity (12.3.1), is a second-order ordinary differential equation, in view of Definition 12.6.1, first integral (14.2.4) and ray parameter (14.2.5) 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 (14.2.5) 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.

14.3.2. Rays.

Derivation. To obtain an analytic expression for a ray, we wish to use an exact ray-velocity expression to be inserted into expression (14.2.5). For this purpose, we consider expression (8.7.13), namely,

$$(14.3.1) \quad V(\theta) = V_z \sqrt{\frac{1 + \tan^2 \theta}{1 + \left(\frac{V_z}{V_x}\right)^2 \tan^2 \theta}},$$

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

$$(14.3.2) \quad \chi := \frac{V_x^2 - V_z^2}{2V_z^2},$$

where V_x and V_z stand for the magnitude of the horizontal and the vertical ray velocities, respectively. This definition of χ is consistent with definition (8.5.1) since — for elliptical velocity dependence — the ray and phase velocities are equal to one another along the axes of the ellipse.

Using χ , we can write expression (14.3.1) 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

$$(14.3.3) \quad V(\theta, z) = V_z(z) \sqrt{\frac{1 + 2\chi}{1 + 2\chi \cos^2 \theta}}.$$

³For 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.

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 (14.3.3) as

$$(14.3.4) \quad V(\theta, z) = (a + bz) \sqrt{\frac{1 + 2\chi}{1 + 2\chi \cos^2 \theta}},$$

where a and b are positive constants.

Inserting expression (14.3.4) into expression (14.2.5), we obtain

$$(14.3.5) \quad \begin{aligned} p &= \cos \theta \frac{\partial}{\partial \theta} \left[\frac{1}{(a + bz) \sqrt{\frac{1 + 2\chi}{1 + 2\chi \cos^2 \theta}}} \right] + \frac{\sin \theta}{(a + bz) \sqrt{\frac{1 + 2\chi}{1 + 2\chi \cos^2 \theta}}} \\ &= \frac{1}{(a + bz) \sqrt{1 + 2\chi}} \left(\cos \theta \frac{\partial}{\partial \theta} \sqrt{1 + 2\chi \cos^2 \theta} + \sin \theta \sqrt{1 + 2\chi \cos^2 \theta} \right) \\ &= \frac{\sin \theta}{(a + bz) \sqrt{1 + 2\chi} \sqrt{1 + 2\chi \cos^2 \theta}}. \end{aligned}$$

To obtain an expression for a ray, we wish to state ray parameter (14.3.5) in terms of position variables x and z . Dividing both the numerator and the denominator by $\sin \theta$, we rewrite expression (14.3.5) as

$$p = \frac{1}{(a + bz) \sqrt{1 + 2\chi} \sqrt{1 + \frac{2\chi + 1}{\tan^2 \theta}}}.$$

Squaring both sides and rearranging, we obtain

$$(14.3.6) \quad \frac{1}{\tan^2 \theta} = \frac{1 - p^2 (a + bz)^2 (1 + 2\chi)}{p^2 (a + bz)^2 (1 + 2\chi)^2}.$$

Since $1/\tan^2 \theta = (dz/dx)^2$, we have a first-order ordinary differential equation, which is a special case of first integral (12.5.4). We can write equation (14.3.6) as

$$(14.3.7) \quad \frac{dz}{dx} = \frac{\sqrt{1 - p^2 (a + bz)^2 (1 + 2\chi)}}{p (a + bz) (1 + 2\chi)},$$

which can be restated as

$$(14.3.8) \quad dx = \frac{p (a + bz) (1 + 2\chi)}{\sqrt{1 - p^2 (a + bz)^2 (1 + 2\chi)}} dz.$$

⁴Readers 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.

To integrate equation (14.3.8), 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,

$$\int_0^x d\xi = \int_0^z \frac{\mathbf{p}(a + b\zeta)(1 + 2\chi)}{\sqrt{1 - \mathbf{p}^2(a + b\zeta)^2(1 + 2\chi)}} d\zeta,$$

where ξ and ζ are the integration variables, we obtain

$$(14.3.9) \quad x = \frac{1}{\mathbf{p}b} \left[\sqrt{1 - \mathbf{p}^2 a^2 (1 + 2\chi)} - \sqrt{1 - \mathbf{p}^2 (a + bz)^2 (1 + 2\chi)} \right],$$

which describes the ray given by $x(z)$ for elliptical velocity dependence with direction and a linear velocity dependence with depth.

Interpretation. To obtain a geometrical interpretation of equation (14.3.9), we rearrange it and write

$$(14.3.10) \quad \frac{\left(x - \frac{\sqrt{1 - \mathbf{p}^2 a^2 (1 + 2\chi)}}{\mathbf{p}b} \right)^2}{\left(\frac{1}{\mathbf{p}b} \right)^2} + \frac{\left(z + \frac{a}{b} \right)^2}{\left(\frac{1}{\mathbf{p}b\sqrt{1 + 2\chi}} \right)^2} = 1.$$

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

$$(14.3.11) \quad \left[\frac{\sqrt{1 - \mathbf{p}^2 a^2 (1 + 2\chi)}}{\mathbf{p}b}, -\frac{a}{b} \right].$$

Equation (14.3.10) is identical to equation (8.5.16), which we obtained in Section 8.5.6 using Hamilton's ray equations.

In a seismological notation, with the z -axis being vertical and pointing downwards, the centre of the ellipse is located on the horizontal line positioned a/b units above the x -axis. In view of $v(z) = a + bz$, this line corresponds to the level where the velocity vanishes. Ellipse (14.3.10) passes through the origin, as can be verified by setting $z = 0$ in equation (14.3.9). The segment of the ellipse that is below the x -axis corresponds to the ray. This interpretation is consistent with our discussion in Section 8.5, as expected.

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. In such a case, the ray is a straight line, as also shown in Exercise 14.3. For a signal propagating along the z -axis, $\theta = 0$ and, following expression (14.3.5), $\mathbf{p} = 0$. Hence, in view of expression (14.3.11), the x -coordinate of the centre of the ellipse is located infinitely far from the source. In such a case, the ray is a vertical straight line. For the isotropic case, where $\chi = 0$, equation (14.3.10) reduces to the expression for a circle. In such a case, the rays are circular arcs.

14.3.3. Traveltimes. We can use ray parameter (14.2.5) to obtain the traveltime along the corresponding ray. For this purpose, we wish to rewrite integral (14.2.1) to include the ray parameter for a given source-receiver pair.

Integral (14.2.1) can be viewed as $\int ds/V$, where ds is the arclength element along the ray. In the xz -plane, the arclength element can be written as $ds = dx/\sin \theta$ or as $ds = dz/\cos \theta$, where θ is the ray angle. We choose the former case, since along the x -axis a ray is expressed as a single-valued function; this is not the case along the z -axis due to the downgoing and upgoing signal. Thus, traveltime integral (14.2.1) between the source at $(0, 0)$ and the receiver at (X, Z) can be written as

$$(14.3.12) \quad \check{C} = \int_0^X \frac{dx}{V(z, \theta) \sin \theta},$$

where $V(z, \theta)$ is given by expression (14.3.4).

To integrate, we must express $\sin \theta$ in terms of constants a, b, χ, \mathfrak{p} , and integration variable x . Combining expressions (14.3.4) and (14.3.5), we write

$$V(z, \theta) \sin \theta = (a + bz) \sqrt{\frac{1 + 2\chi}{1 + 2\chi \cos^2 \theta}} \mathfrak{p} (a + bz) \sqrt{1 + 2\chi} \sqrt{1 + 2\chi \cos^2 \theta} = \mathfrak{p} (a + bz)^2 (1 + 2\chi),$$

which is the denominator in integral (14.3.12). To proceed with integration, we must express z in terms of x . To do so, we write expression (14.3.9) as

$$\left(\mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right)^2 = 1 - \mathfrak{p}^2 (a + bz)^2 (1 + 2\chi).$$

Combining the last two equations, we get

$$\left(\mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right)^2 = 1 - \mathfrak{p} V(z, \theta) \sin \theta.$$

Solving for $V \sin \theta$, we obtain

$$(14.3.13) \quad V(z, \theta) \sin \theta = \frac{1 - \left(\mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right)^2}{\mathfrak{p}},$$

which is the denominator in integral (14.3.12) expressed in terms of variable x . Inserting expression (14.3.13) into integral (14.3.12) and using the fact that \mathfrak{p} is constant for a given source-receiver pair, we write

$$(14.3.14) \quad \check{C} = \mathfrak{p} \int_0^X \frac{dx}{1 - \left(\mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right)^2}.$$

Integrating between $z = 0$ and $z = Z$, while treating \mathfrak{p} as a constant, we obtain — as shown in Exercise 14.5 — the expression for the value of the traveltime; namely,

$$(14.3.15) \quad \check{C} = \frac{\tanh^{-1} \left[\mathfrak{p}bX - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right] + \tanh^{-1} \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}{b},$$

which is expression (8.5.17), as could be expected. We 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 (14.3.9) as

$$(14.3.16) \quad X = \frac{1}{\mathfrak{p}b} \left[\sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} - \sqrt{1 - \mathfrak{p}^2 (a + bZ)^2 (1 + 2\chi)} \right].$$

Solving for \mathfrak{p} , we obtain

$$(14.3.17) \quad \mathfrak{p} = \frac{2X}{\sqrt{[X^2 + (1 + 2\chi) Z^2] \left[(2a + bZ)^2 (1 + 2\chi) + b^2 X^2 \right]}}.$$

Expression (14.3.17) is identical to expression (8.5.18), which we obtained in Section 8.5.7 using Hamilton's ray equations. Herein, however, \mathfrak{p} is treated as a constant for a particular choice of X and Z .

Studying the properties of the continuum in terms of a , b and χ , we can use expression (14.3.15) with \mathfrak{p} given by expression (14.3.17) to obtain the traveltime between the source and the receiver. These expressions are convenient to use for inverse problems that are based on traveltime measurements.⁵ For such a study, we might wish to know if the receiver has been reached by a downgoing or an upgoing signal travelling along an elliptical arc. As shown in Exercise 14.7, the subsurface receiver at (X, Z) is reached by the downgoing signal from the source at $(0, 0)$ if

$$X < \sqrt{\frac{1 + 2\chi}{b}} (2a + bZ) Z.$$

For the value of X equal to the right-hand side, the signal is at its deepest point when it reaches the receiver. Also as shown in Exercise 14.7, the curve distinguishing between the downgoing and upgoing arrivals is a hyperbola whose asymptote is

$$x = \sqrt{1 + 2\chi} \left(z + \frac{a}{b} \right).$$

As shown in Section 6.10.3, by using a linear transformation of coordinates we can treat elliptical velocity dependence as an isotropic case. Consequently, we can also obtain the traveltime and ray-parameter expressions by such a transformation, as illustrated in Exercise 14.8.

⁵Interested readers might refer to Slawinski, M.A., Wheaton, C.J., and Powojowski, M. (2004) VSP traveltime inversion for linear inhomogeneity and elliptical anisotropy: *Geophysics*, **69**, 373 – 377.

14.4. Rays in isotropic continua

In Sections 14.2 and 14.3, we studied ray equations in two-dimensional anisotropic continua and obtained analytic expressions for rays and traveltimes. The availability of analytic expressions resulted from the assumption of homogeneity along the x -axis and, hence, from the existence of a first 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

$$(14.4.1) \quad \check{C} = \int_A^B \frac{\sqrt{1 + (z')^2}}{V(x, z)} dx.$$

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 12.5.6, the corresponding ray equation, which results from Euler's equation (12.2.2), is given by equation (12.5.8), namely,

$$(14.4.2) \quad V \frac{d^2 z}{dx^2} - \frac{\partial V}{\partial x} \left(\frac{dz}{dx} \right)^3 + \frac{\partial V}{\partial z} \left(\frac{dz}{dx} \right)^2 - \frac{\partial V}{\partial x} \frac{dz}{dx} + \frac{\partial V}{\partial z} = 0,$$

where, due to the isotropy of the continuum, phase and ray velocities coincide, namely, $v \equiv V$. Equation (14.4.2) 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 (14.1.1). An analogous parametric representation can be given in terms of $x(t)$, $z(t)$, $\dot{x} := dx/dt$ and $\dot{z} := dz/dt$. Then, the traveltime integral is

$$(14.5.1) \quad \check{C} = \int \frac{ds}{V} = \int \frac{\sqrt{\dot{x}^2 + \dot{z}^2}}{V(x, z, \dot{x}, \dot{z})} dt := \int \mathcal{F} dt,$$

where \mathcal{F} is a two-dimensional form of expression (13.1.10).

In view of the principle of stationary traveltime, we can use Lagrange's ray equations (13.1.9). In the two-dimensional case, discussed herein, these equations constitute the system

$$(14.5.2) \quad \begin{cases} \frac{\partial \mathcal{F}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{F}}{\partial \dot{x}} \right) = 0 \\ \frac{\partial \mathcal{F}}{\partial z} - \frac{d}{dt} \left(\frac{\partial \mathcal{F}}{\partial \dot{z}} \right) = 0 \end{cases},$$

where \mathcal{F} denotes the integrand of the traveltime integral.

Also, the equations of system (14.5.2) are related by Beltrami's identity; namely,

$$(14.5.3) \quad \frac{\partial \mathcal{F}}{\partial t} + \frac{d}{dt} \left(\dot{x} \frac{\partial \mathcal{F}}{\partial \dot{x}} + \dot{z} \frac{\partial \mathcal{F}}{\partial \dot{z}} - \mathcal{F} \right) = 0.$$

The justification for this form of Beltrami's identity is shown in Exercise 11.1.

In view of Theorem A.2.1, 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.2.1 that

$$(14.5.4) \quad \mathcal{F} = \dot{x} \frac{\partial \mathcal{F}}{\partial \dot{x}} + \dot{z} \frac{\partial \mathcal{F}}{\partial \dot{z}}.$$

Consequently, the term in parentheses of Beltrami's identity (14.5.3) vanishes and equation (14.5.3) 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 (14.5.2). If \mathcal{F} is not explicitly dependent on x , the first equation becomes $\partial \mathcal{F} / \partial \dot{x} = \mathbf{p}$, where \mathbf{p} is a conserved quantity. This conserved quantity is equivalent to ray-parameter expression (14.2.5), as shown in Exercise 14.4.

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 (8.2.7); namely,

$$\begin{cases} \dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \end{cases}, \quad i \in \{1, 2, 3\}.$$

Before discussing ray parameters, to gain more insight into Hamilton's ray equations let us consider \mathcal{H} itself. If \mathcal{H} does not explicitly depend on t , it is conserved along the ray, as stated by Lemma 13.1.3.

Also, in view of expression (13.3.13), we can infer that in classical mechanics the Hamiltonian that is independent of time implies the conservation of energy.

Let us return to ray parameters. 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 $dp_i/dt = 0$. To elucidate the consequences of this statement, recall expression (8.2.8); 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 (11.2.6), namely,

$$(14.6.1) \quad \frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) = 0, \quad i \in \{1, 2, 3\},$$

if \mathcal{L} does not explicitly depend on the x_i , the equation of system (14.6.1) that corresponds to the given subscript i is reduced to

$$(14.6.2) \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) = 0,$$

which implies that the term in parentheses of expression (14.6.2) is constant. In view of Legendre's transformation, following expressions (B.3.2), shown in Appendix B, we can write

$$(14.6.3) \quad p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i},$$

where $\mathbf{p} = p_i$ is the conserved quantity along the solution curve $\mathbf{x}(t)$.

Expressions given by $dp_i/dt = 0$ and $d(\partial \mathcal{L}/\partial \dot{x}_i)/dt = 0$, 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

$$\mathbf{p} = \frac{\sin \vartheta}{v(z, \vartheta)} = \cos \theta \frac{\partial}{\partial \theta} \left[\frac{1}{V(z, \theta)} \right] + \frac{\sin \theta}{V(z, \theta)},$$

where the relation between the magnitudes of phase and ray velocities is given by expressions (8.4.9), while the relation between the phase and ray angles is given by expression (8.4.12). An example illustrating this equivalence is shown in Exercise 10.1.

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.



14.7. Exercises

EXERCISE 14.1. *Using polar coordinates, formulate the conserved quantity for radially inhomogeneous continua, where the travelt ime integral is given by*

$$\int_a^b \frac{\sqrt{r^2 + (r')^2}}{V(r)} d\xi.$$

Explain the physical context of the conserved quantity.

SOLUTION 14.1. Consider Beltrami's identity given by expression

$$\frac{\partial F}{\partial \xi} + \frac{d}{d\xi} \left(r' \frac{\partial F}{\partial r'} - F \right) = 0,$$

where F denotes the integrand of the travelt ime integral and $r' := dr/d\xi$. Due to the explicit absence of the latitude angle, ξ , we obtain

$$r' \frac{\partial F}{\partial r'} - F = C,$$

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 travelt ime integral's invariance to the latitude angle. In other words, the velocity field consists of concentric circles.

REMARK 14.7.1. 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

$$(14.7.1) \quad C = - \frac{r \sin \theta}{V(r)},$$

which is a standard form of the ray parameter for radially inhomogeneous continua.⁶ Note that ray parameter (14.7.1) has different units than ray parameter (14.2.5).

EXERCISE 14.2. *Given expression (14.3.5), examine how the value of the anisotropy parameter χ affects the maximum depth of a ray.*

SOLUTION 14.2. Solving equation (14.3.5) for z , we obtain

$$z = \frac{1}{pb} \left(\frac{\sin \theta}{\sqrt{1 + 2\chi} \sqrt{1 + 2\chi \cos^2 \theta}} - pa \right).$$

⁶Readers interested in travelt ime expressions for rays whose ray parameter is given by expression (14.7.1) might refer to Kennett, B.L.N., (2001) The seismic wavefield, Vol. I: Introduction and theoretical development: Cambridge University Press, pp. 171 – 174.

The maximum depth is given by setting $\theta = \pi/2$. Thus, we get

$$(14.7.2) \quad z_{\max} = \frac{1}{pb} \left(\frac{1}{\sqrt{1+2\chi}} - pa \right).$$

To state expression (14.7.2) in terms of the take-off ray angle, we set $z = 0$ in expression (14.3.5), and denote the corresponding θ as θ_0 . Hence, expression (14.3.5) becomes

$$(14.7.3) \quad p = \frac{\sin \theta_0}{a\sqrt{1+2\chi}\sqrt{1+2\chi \cos^2 \theta_0}}.$$

Inserting expression (14.7.3) into expression (14.7.2), we obtain

$$z_{\max} = \frac{a}{b} \left(\frac{\sqrt{1+2\chi \cos^2 \theta_0}}{\sin \theta_0} - 1 \right),$$

which gives the maximum depth for a given take-off 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 14.7.2. 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. Using expression (14.3.9) show that in homogeneous continua rays are straight lines.

SOLUTION 14.3. Consider equation (14.3.9); namely,

$$x(z; a, b) = \frac{\sqrt{1 - p^2 a^2 (1 + 2\chi)} - \sqrt{1 - p^2 (a + bz)^2 (1 + 2\chi)}}{pb}.$$

For a homogeneous continuum, $b = 0$; hence, we can write

$$\lim_{b \rightarrow 0} x(z; a, b) = \lim_{b \rightarrow 0} \frac{\sqrt{1 - p^2 a^2 (1 + 2\chi)} - \sqrt{1 - p^2 (a + bz)^2 (1 + 2\chi)}}{pb}.$$

Since both the numerator and the denominator vanish, we invoke de l'Hôpital's rule to write

$$\lim_{b \rightarrow 0} x(z; a, b) = \lim_{b \rightarrow 0} \frac{\partial [\sqrt{1 - p^2 a^2 (1 + 2\chi)} - \sqrt{1 - p^2 (a + bz)^2 (1 + 2\chi)}]}{\frac{\partial b}{\partial (pb)}}.$$

Performing the differentiation, we get

$$\lim_{b \rightarrow 0} x(z; a, b) = \lim_{b \rightarrow 0} \frac{p(a + bz)(1 + 2\chi)z}{\sqrt{1 - p^2 (a + bz)^2 (1 + 2\chi)}}.$$

Taking the limit, we obtain

$$\lim_{b \rightarrow 0} x(z; a, b) = \frac{pa(1 + 2\chi)}{\sqrt{1 - p^2 a^2 (1 + 2\chi)}} z.$$

Since the term given by the fraction is constant and there is no additional term, this is the equation of a straight line going through the origin, $(x, z) = (0, 0)$.

To interpret this result, let us rewrite the above expression as

$$z(x) = \frac{\sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}{\mathfrak{p} a (1 + 2\chi)} x,$$

which describes the ray in an elliptically anisotropic homogeneous continuum. Let us also recall expression (14.3.5) and consider it for $b = 0$. Thus we get

$$\mathfrak{p} = \frac{\sin \theta}{a \sqrt{1 + 2\chi} \sqrt{1 + 2\chi \cos^2 \theta}}.$$

Combining these two expressions, we get

$$z(x) = \frac{\sqrt{1 - \left(\frac{\sin \theta}{\sqrt{1 + 2\chi} \sqrt{1 + 2\chi \cos^2 \theta}} \right)^2}}{\frac{\sin \theta \sqrt{1 + 2\chi}}{\sqrt{1 + 2\chi \cos^2 \theta}}} x.$$

Simplifying, we obtain

$$z(x) = \cot \theta x,$$

where $\cot \theta$ is the slope of the line. This is true for both isotropic homogeneous and anisotropic homogeneous cases; χ does not appear in the final result.

EXERCISE 14.4. *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 (14.2.5).*

SOLUTION 14.4. Using the first equation of system (14.5.2) and considering the case where x is not explicitly present in the integrand, \mathcal{F} , we obtain the conserved quantity given by

$$(14.7.4) \quad \mathfrak{p} = \frac{\partial \mathcal{F}}{\partial \dot{x}}.$$

In view of \mathcal{F} given in expression (14.5.1), we obtain

$$(14.7.5) \quad \mathfrak{p} = \frac{1}{V} \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{z}^2}} + \sqrt{\dot{x}^2 + \dot{z}^2} \frac{\partial}{\partial \dot{x}} \left(\frac{1}{V} \right).$$

In order to state expression (14.7.5) 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 (14.7.5), 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 (14.2.5), as required.

EXERCISE 14.5. *Integrate traveltime expression (14.3.14), namely,*

$$(14.7.6) \quad \mathfrak{p} \int_0^X \frac{dx}{1 - \left(\mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right)^2},$$

using the fact that \mathfrak{p} is constant.

SOLUTION 14.5. Let us make the following substitution.

$$\zeta := \mathfrak{p}bx - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}.$$

Thus, we rewrite integral (14.7.6) as

$$\frac{1}{b} \int_{-\sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}^{\mathfrak{p}bX - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}} \frac{d\zeta}{1 - \zeta^2}.$$

Using partial fractions, we write

$$\frac{1}{2b} \int_{-\sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}^{\mathfrak{p}bX - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}} \left(\frac{1}{1 - \zeta} + \frac{1}{1 + \zeta} \right) d\zeta.$$

Integrating, we get

$$\frac{1}{2b} \ln \frac{1 + \zeta}{1 - \zeta} \Big|_{-\sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}^{\mathfrak{p}bX - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}.$$

Evaluating, we obtain

$$\frac{1}{2b} \left(\ln \frac{1 - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} + \mathfrak{p}bX}{1 + \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} - \mathfrak{p}bX} - \ln \frac{1 - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}{1 + \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}} \right).$$

Using an identity for hyperbolic functions, namely,

$$\tanh^{-1} \zeta = \frac{1}{2} \ln \frac{1 + \zeta}{1 - \zeta},$$

we can also write

$$\frac{\tanh^{-1} \left[\mathfrak{p}bX - \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)} \right] + \tanh^{-1} \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}{b}.$$

EXERCISE 14.6. *In a manner analogous to the one described in Section 14.3.3, derive the traveltime expression by integrating traveltime integral (14.2.1) along the z -axis. Discuss the validity of the resulting expression.*

SOLUTION 14.6. In the xz -plane, the arclength element can be written as $ds = dz / \cos \theta$, where θ is the ray angle. Hence, traveltime integral (14.2.1) between the source at $(0, 0)$ and the receiver at (X, Z) is

$$(14.7.7) \quad \check{C} = \int_0^Z \frac{dz}{V(z, \theta) \cos \theta},$$

where $V(z, \theta)$ is given by expression (14.3.4). Hence, we can explicitly write

$$(14.7.8) \quad \check{C} = \int_0^Z \frac{dz}{(a + bz) \sqrt{\frac{1+2\chi}{1+2\chi \cos^2 \theta}} \cos \theta}.$$

To integrate, we must express $\cos \theta$ in terms of constants a, b, χ, \mathfrak{p} , and integration variable z . Using expression (14.3.5) and trigonometric identities, we obtain

$$(14.7.9) \quad \cos \theta = \sqrt{\frac{1 - \mathfrak{p}^2 (a + bz)^2 (1 + 2\chi)}{1 + 2\chi \mathfrak{p}^2 (a + bz)^2 (1 + 2\chi)}}.$$

Inserting expression (14.7.9) into integral (14.7.8), after algebraic manipulation, we obtain

$$(14.7.10) \quad \check{C} = \int_0^Z \frac{dz}{(a + bz) \sqrt{1 - \mathfrak{p}^2 (a + bz)^2 (1 + 2\chi)}}.$$

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,

$$(14.7.11) \quad \check{C} = \frac{1}{b} \ln \left[\frac{a + bZ}{a} \frac{1 + \sqrt{1 - \mathfrak{p}^2 a^2 (1 + 2\chi)}}{1 + \sqrt{1 - \mathfrak{p}^2 (a + bZ)^2 (1 + 2\chi)}} \right],$$

with \mathfrak{p} given by expression (14.3.17).

REMARK 14.7.3. Expression (14.7.11) is valid for the downgoing rays only, which could be illustrated by plotting this expression versus x for a set value of Z ; such a plot would exhibit a cusp at x that corresponds to the case of the signal reaching the receiver, (X, Z) , at its deepest point — just prior to its beginning an upward path. Between the source and the cusp, the receiver is reached by a downgoing signal travelling along an elliptical trajectory. Beyond the cusp, the receiver is reached by an upgoing signal travelling along

this trajectory. The cusp in such a plot is the result of the traveltime beyond the cusp being calculated with the opposite sign than the traveltime before the cusp, due to the change of the direction of integration that is set by dz .

EXERCISE 14.7. *In view of Remark 14.7.3 and considering a source at $(0, 0)$, derive an expression for the curve in the xz -plane that separates the receiver locations, (X, Z) , reached by the downgoing signal from those reached by the upgoing one.*

SOLUTION 14.7. At a given depth (set value of Z), the receivers at different horizontal locations (varying values of X) are reached either by a downgoing or upgoing signal. The receiver at $(0, Z)$, which is directly below the source, is reached by the downgoing signal whose takeoff ray angle is $\theta = 0$. As X grows, so does θ , until the takeoff ray angle corresponds to the ray that horizontally grazes the receiver at (X, Z) ; the takeoff angle is at its maximum. From that point on, the receiver is reached by the upgoing signal travelling along an elliptical arc. Consequently, the ray takeoff angle begins to decrease, and the receiver at an infinite horizontal distance would be reached by a ray whose takeoff angle is nearly zero — the ray is nearly vertical at takeoff. In view of expression (8.5.11), we see that the increase or decrease in the ray angle corresponds to the increase or decrease of the phase angle, since $\chi > -1/2$, as we can deduce from definition (8.5.1). Also, in view of expression (8.5.12), we see that the increase or decrease of the phase takeoff angle corresponds to the increase or decrease of ray parameter. Hence, to find point x at depth Z at which the ray angle reaches its largest value, let us write expression (14.3.17) as

$$p(x; Z) = \frac{2x}{\sqrt{[x^2 + (1 + 2\chi) Z^2] \left[(2a + bZ)^2 (1 + 2\chi) + b^2 x^2 \right]}},$$

and consider its derivative, namely,

$$\frac{dp}{dx} = \frac{8a^2 Z^2 (1 + 2\chi) + 8abZ^3 (1 + 2\chi) + 2b^2 \left(Z^4 (1 + 2\chi)^2 - x^4 \right)}{\left(\left(x^2 + (1 + 2\chi)^2 Z^2 \right) \left(b^2 x^2 + (1 + 2\chi) (2a + bZ)^2 \right) \right)^{\frac{3}{2}}}.$$

To find x that corresponds to the maximum of the takeoff angle, we set the numerator to zero and proceed to rearrange the resulting equation using the fact that a, b, χ, Z are real and a, b, Z are positive. Thus we obtain the required solution, namely,

$$x = \sqrt{\frac{1 + 2\chi}{b}} (2a + bZ) Z.$$

We can rewrite this solution as

$$\frac{x^2}{\left(\frac{a}{b}\right)^2 (1 + 2\chi)} - \frac{\left(z + \frac{a}{b}\right)^2}{\left(\frac{a}{b}\right)^2} = 1,$$

which is an expression for hyperbola. To find its asymptote, we set the right-hand side to zero, and solve for x to obtain

$$x = \sqrt{1 + 2\chi} \left(z + \frac{a}{b} \right).$$

EXERCISE 14.8. Consider the traveltimes expression for the downgoing signal between the source at $(0, 0)$ and the receiver at (X, Ξ) in an isotropic and linearly inhomogeneous continuum in the $x\zeta$ -plane, namely,⁷

$$(14.7.12) \quad \check{C} = \frac{1}{b} \ln \left[\frac{\alpha + b\Xi}{\alpha} \frac{1 + \sqrt{1 - \mathbf{p}^2 \alpha^2}}{1 + \sqrt{1 - \mathbf{p}^2 (\alpha + b\Xi)^2}} \right],$$

where

$$(14.7.13) \quad \mathbf{p} = \frac{2X}{\sqrt{(X^2 + \Xi^2) [(2\alpha + b\Xi)^2 + b^2 X^2]}}.$$

Using a transformation of coordinates, derive equations (14.7.11) and (14.3.17).

SOLUTION 14.8. Since equations (14.7.11) and (14.3.17) deal with an elliptical velocity dependence, consider the magnitudes of the horizontal and vertical velocities given by

$$(14.7.14) \quad v_x = a\sqrt{1 + 2\chi},$$

with χ given by expression (14.3.2), and by

$$(14.7.15) \quad v_z = a,$$

respectively.

Note that, since v_x and v_z are the magnitudes of velocities along the symmetry axes, expressions (14.7.14) and (14.7.15) are the same for both phase and ray velocities.

Regardless of the inhomogeneity of the model, infinitesimal wavefronts resulting from any point source within the medium are ellipses with axes $(dt) v_x$ and $(dt) v_z$, where dt is the traveltimes increment and v_x and v_z are the magnitudes of velocity at a given point. We can write such a wavefront as

$$\frac{(dx)^2}{(dt)^2 v_x^2} + \frac{(dz)^2}{(dt)^2 v_z^2} = 1,$$

which, using expressions (14.7.14) and (14.7.15), we can rewrite as

$$(14.7.16) \quad \frac{(dx)^2}{1 + 2\chi} + (dz)^2 = (dt)^2 a^2.$$

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

$$(14.7.17) \quad \zeta = z\sqrt{1 + 2\chi}.$$

⁷Readers interested in the derivation of expressions (14.7.12) and (14.7.13), might refer to Slotnick, M.M., (1959) Lessons in seismic computing: Society of Exploration Geophysicists, Lesson 37.

Thus, in view of expression (14.7.17), we let $z = \zeta/\sqrt{1+2\chi}$ in expression (14.7.16) to write it as

$$(14.7.18) \quad (dx)^2 + (d\zeta)^2 = (dt)^2 \alpha^2,$$

where

$$(14.7.19) \quad \alpha = a\sqrt{1+2\chi}$$

is the velocity in the $x\zeta$ -plane. Equation (14.7.18) describes infinitesimal circular wavefronts in the $x\zeta$ -plane.

Since equations (14.7.11) and (14.3.17) deal with the magnitude of the velocity that increases linearly, let us set $v(\zeta) = \alpha + b\zeta$. Note that the units of b are $[1/s]$. Consequently, its value does not depend on the scaling of position coordinates. Substituting expression (14.7.19) into expressions (14.7.12) and (14.7.13), as well as — in view of expression (14.7.17) — letting $\Xi = Z\sqrt{1+2\chi}$, we obtain expressions (14.7.11) and (14.3.17), as required.

EXERCISE 14.9. *In view of Lemma 12.8.2, 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 14.9. 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,

$$(14.7.20) \quad V(z, z') = A(z) B(z') = \frac{A(z)}{1 + C \frac{z'}{\sqrt{1+(z')^2}}}.$$

Consider traveltime integral

$$\check{C} = \int_{x_1}^{x_2} \frac{\sqrt{1 + (z')^2}}{V(z, z')} dx = \int_{x_1}^{x_2} \frac{\sqrt{1 + (z')^2}}{\frac{A(z)}{1 + C \frac{z'}{\sqrt{1+(z')^2}}}} dx.$$

Upon algebraic manipulations, we obtain

$$\check{C} = \int_{x_1}^{x_2} \frac{\sqrt{1 + (z')^2} + C z'}{A(z)} dx \equiv \int_{x_1}^{x_2} F dx.$$

To find the ray, we invoke Euler's equation (12.2.2) to obtain

$$(14.7.21) \quad \frac{\partial F}{\partial z} - \frac{d}{dx} \left(\frac{\partial F}{\partial z'} \right) = - \left(\sqrt{1 + (z')^2} + C z' \right) \frac{\frac{\partial A}{\partial z}}{A^2} - \frac{d}{dx} \left[\frac{1}{A} \left(\frac{z'}{\sqrt{1 + (z')^2}} + C \right) \right] = 0.$$

Considering only factors which contain C , gives us

$$-C z' \frac{\frac{\partial A}{\partial z}}{A^2} + C z' \frac{\frac{\partial A}{\partial z}}{A^2} = 0.$$

Thus, Euler's equation is independent of C . In view of expression (14.7.20), the ray resulting from equation (14.7.21) is the same for both $V(z, z')$ and $A(z)$.

Part 4

Appendices

Introduction to Part 4

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 Grant 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 2* with those discussed in *Part 3*.

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 (8.2.8); namely,

$$(A.1.1) \quad \mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} p^2 v^2(\mathbf{x}, \mathbf{p}),$$

where $p^2 \equiv \mathbf{p} \cdot \mathbf{p}$, is homogeneous of degree 2 in \mathbf{p} . To see this property, consider Definition A.1.1.

DEFINITION A.1.1. A real function $f(x_1, \dots, x_n)$ is homogeneous of degree r in the variables x_1, \dots, x_n if

$$f(cx_1, \dots, cx_n) = c^r f(x_1, \dots, x_n),$$

for every real number c . If $f(cx_i) = |c|^r f(x_i)$, where $i \in \{1, \dots, n\}$, we say that f is absolute-value homogeneous of degree r in the x_i .

REMARK A.1.2. 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.1.1) as

$$(A.1.2) \quad \mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} p^2 v^2 \left(\mathbf{x}, \frac{\mathbf{p}}{|\mathbf{p}|} \right),$$

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.1.1, 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.1.2) is homogeneous of degree 2 in \mathbf{p} , since

$$(A.1.3) \quad \begin{aligned} \mathcal{H}(\mathbf{x}, c\mathbf{p}) &= \frac{1}{2} [(c\mathbf{p}) \cdot (c\mathbf{p})] v^2 \left(\mathbf{x}, \frac{c\mathbf{p}}{|c\mathbf{p}|} \right) = \frac{1}{2} c^2 p^2 v^2 \left(\mathbf{x}, \frac{c\mathbf{p}}{|c\mathbf{p}|} \right) \\ &= \frac{c^2}{2} p^2 v^2 \left(\mathbf{x}, \frac{\mathbf{p}}{|\mathbf{p}|} \right) = c^2 \mathcal{H}(\mathbf{x}, \mathbf{p}). \end{aligned}$$

We can also illustrate Definition A.1.1 by the following straightforward example.

EXAMPLE A.1.3. Consider function

$$(A.1.4) \quad f(x_1, x_2, x_3) = x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2.$$

Let

$$\begin{aligned} f(cx_1, cx_2, cx_3) &= cx_1 cx_2 cx_3 + cx_1 (cx_2)^2 + cx_2 (cx_3)^2 \\ &= c^3 (x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2) \\ &= c^3 f(x_1, x_2, x_3). \end{aligned}$$

Thus, in view of Definition A.1.1, 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.2.1. 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

¹Euler's homogeneous-function theorem can be stated in the following way.

¹Interested readers might refer to Olmsted, J.M.H., (1961) Advanced calculus: Prentice-Hall, Inc., p. 272.

THEOREM A.2.1. *If function $f(x_1, \dots, x_n)$ is homogeneous of degree r in x_1, \dots, x_n , then*

$$(A.2.1) \quad \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1, \dots, x_n) x_i = r f(x_1, \dots, x_n).$$

PROOF. In view of Definition A.1.1, we can write

$$(A.2.2) \quad f(cx_1, \dots, cx_n) = c^r f(x_1, \dots, x_n).$$

Differentiating both sides of equation (A.2.2) with respect to c , we obtain

$$(A.2.3) \quad \sum_{i=1}^n f_i(cx_1, \dots, cx_n) \frac{\partial (cx_i)}{\partial c} = r c^{r-1} f(x_1, \dots, x_n),$$

where f_i denotes the derivative of function f with respect to its i th argument. To obtain the expression stated in Theorem A.2.1, we consider a particular case where $c = 1$. Letting $c = 1$, we can rewrite equation (A.2.3) as

$$(A.2.4) \quad \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1, \dots, x_n) x_i = r f(x_1, \dots, x_n),$$

which is equation (A.2.1), as required. □

To illustrate Theorem A.2.1, we can study function (A.1.4), as shown in the following example.

EXAMPLE A.2.2. Using function (A.1.4), 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.2.1) as

$$\begin{aligned} \sum_{i=1}^3 \frac{\partial f}{\partial x_i} x_i &= (x_2 x_3 + x_2^2) x_1 + (x_1 x_3 + 2x_1 x_2 + x_3^2) x_2 + (x_1 x_2 + 2x_2 x_3) x_3 \\ &= 3(x_1 x_2 x_3 + x_1 x_2^2 + x_2 x_3^2) \\ (A.2.5) \quad &= 3f(x_1, x_2, x_3). \end{aligned}$$

Expression (A.2.5) is the right-hand side of equation (A.2.1) for a function that is homogeneous of degree 3 in the x_i , as expected from Theorem A.2.1.

Equation (A.2.1) is often invoked in this book. For instance, in the proof of Lemma 13.1.5 — knowing that \mathcal{H} is homogeneous of degree 2 in \mathbf{p} , as shown in expression (A.1.3) — 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.2.1) in the context of physics.

EXAMPLE A.2.3. Following the standard classical-mechanics formulation, let the kinetic energy be

$$T(v) = \frac{1}{2}mv^2,$$

where m and v denote mass and velocity, respectively. In view of Definition A.1.1, T is homogeneous of degree 2 in v since

$$T(cv) = \frac{1}{2}m(cv)^2 = \frac{c^2}{2}mv^2 = c^2T(v),$$

where c denotes a constant. Thus, following Theorem A.2.1, 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 (8.2.8), is homogeneous of degree 2 in \mathbf{p} . $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$, given in expression (11.1.1), is homogeneous of degree 2 in $\dot{\mathbf{x}}$. $\mathcal{F}(\mathbf{x}, \dot{\mathbf{x}})$, given in expression (13.1.8), 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 13.1.2, which is the statement of Fermat's principle.

Certain functions used in our studies exhibit no homogeneity. For instance, traveltime integrand $F(z, z')$, given in expression (14.2.1), is not homogeneous in either variable.

Euler's homogeneous-function theorem is explicitly used in Chapters 8, 11 and 13.



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 planes¹. 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

$$(B.1.1) \quad y = f(x_1, \dots, x_n).$$

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 $y = u_1 x_1 + \dots + u_n x_n - v$, where u_1, \dots, u_n and v are real numbers that define the plane uniquely and, hence, can be viewed as coordinates of the

¹Readers 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.

plane. The transformation from the equation of a surface, given by equation (B.1.1), to the equation that describes all of its tangent planes, given by

$$v = g(u_1, \dots, u_n),$$

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, \dots, u_n) , we have different values of v . In other words, v is not a single-valued function of (u_1, \dots, 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,

$$(B.1.2) \quad v = ux - f(x),$$

and

$$(B.1.3) \quad u = \frac{df}{dx},$$

respectively.

To derive 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.1.3) 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

$$(B.1.4) \quad g(u) = ux(u) - f(x(u)).$$

Function $g(u)$ is Legendre's transformation of $f(x)$, where u and x are the transformation variables related by equation (B.1.3). This construction is illustrated by the following example.

EXAMPLE B.1.1. Let $f(x) = x^2$. Then, following equation (B.1.3), we can write

$$(B.1.5) \quad u = \frac{df}{dx} = 2x.$$

Hence, solving for x , we obtain

$$x = \frac{u}{2}.$$

Consequently, following equation (B.1.4), we write

$$g(u) = ux(u) - f(x(u)) = \frac{u^2}{2} - \left(\frac{u}{2}\right)^2 = \frac{u^2}{4}.$$

Therefore, $g(u) = u^2/4$ is Legendre's transformation of $f(x) = x^2$, where u and x are the transformation variables related by equation (B.1.5).

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

$$\frac{dh}{dx} = u - \frac{df}{dx} = 0,$$

which gives

$$u = \frac{df}{dx}.$$

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.1.1.

EXAMPLE B.2.1. Let $g(u) = u^2/4$. In view of equation (B.1.4), consider a new function given by

$$\begin{aligned} f(x) &= ux - g(u) \\ &= ux - \frac{u^2}{4}, \end{aligned} \tag{B.2.1}$$

where the new independent variable is

$$x = \frac{dg}{du} = \frac{u}{2}. \tag{B.2.2}$$

Herein, we can uniquely express u in terms of x ; namely, $u = 2x$. Hence, we can write function (B.2.1) as

$$f(x) = 2x^2 - \frac{(2x)^2}{4} = x^2,$$

as expected from Example B.1.1. Therefore, $f(x) = x^2$ is Legendre's transformation of $g(u) = u^2/4$, where x and u are the transformation variables related by equation (B.2.2).

²For a proof of this duality, readers might refer to Arnold, V.I., (1989) Mathematical methods of classical mechanics (2nd edition): Springer-Verlag, p. 63.

B.3. Transformation between Lagrangian \mathcal{L} and Hamiltonian \mathcal{H}

NOTATION B.3.1. 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.

REMARK B.3.2. *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.*

NOTATION B.3.3. In this appendix, to show the generality of the formulation, all expressions are derived for an n -dimensional space.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 point \mathbf{x} , variables $\dot{\mathbf{x}}$ and \mathbf{p} are the active variables of transformation.

Let $\mathcal{L} = \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$. In view of expression (B.1.4), consider a new function given by

$$(B.3.1) \quad \mathcal{H}(\mathbf{x}, \mathbf{p}) = \sum_{i=1}^n \dot{x}^i p_i - \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}),$$

where, in view of expression (B.1.3), the new variables are

$$(B.3.2) \quad p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}^i}, \quad i \in \{1, \dots, n\}.$$

Following expression (B.3.1), we can write the differential of $\mathcal{H}(\mathbf{x}, \mathbf{p})$ as

$$(B.3.3) \quad \begin{aligned} d\mathcal{H} &= \sum_{i=1}^n (p_i dx^i + \dot{x}^i dp_i) - \sum_{i=1}^n \left(\frac{\partial \mathcal{L}}{\partial x^i} dx^i + \frac{\partial \mathcal{L}}{\partial \dot{x}^i} d\dot{x}^i \right) \\ &= \sum_{i=1}^n \left(p_i dx^i + \dot{x}^i dp_i - \frac{\partial \mathcal{L}}{\partial x^i} dx^i - \frac{\partial \mathcal{L}}{\partial \dot{x}^i} d\dot{x}^i \right). \end{aligned}$$

In view of expression (B.3.2), the first and the last term in expression (B.3.3) cancel one another. Thus, we obtain

$$(B.3.4) \quad d\mathcal{H} = \sum_{i=1}^n \left(\dot{x}^i dp_i - \frac{\partial \mathcal{L}}{\partial x^i} dx^i \right).$$

³Readers 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.

Also, we can formally write the differential of $\mathcal{H}(\mathbf{x}, \mathbf{p})$ as

$$(B.3.5) \quad d\mathcal{H} = \sum_{i=1}^n \left(\frac{\partial \mathcal{H}}{\partial x^i} dx^i + \frac{\partial \mathcal{H}}{\partial p_i} dp_i \right),$$

which is a statement of the chain rule.

Equating the right-hand sides of equations (B.3.4) and (B.3.5), we can write

$$(B.3.6) \quad \sum_{i=1}^n \left(\dot{x}^i dp_i - \frac{\partial \mathcal{L}}{\partial x^i} dx^i \right) = \sum_{i=1}^n \left(\frac{\partial \mathcal{H}}{\partial x^i} dx^i + \frac{\partial \mathcal{H}}{\partial p_i} dp_i \right).$$

By examining equation (B.3.6), we conclude that

$$(B.3.7) \quad \dot{x}^i = \frac{\partial \mathcal{H}}{\partial p_i}, \quad i \in \{1, \dots, n\}.$$

Examining expressions (B.3.2) and (B.3.7), we recognize the duality of these expressions. Thus, we can write the counterpart of expression (B.3.1), namely,

$$(B.3.8) \quad \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \sum_{i=1}^n \dot{x}^i p_i - \mathcal{H}(\mathbf{x}, \mathbf{p}),$$

where the active variables are given by expression (B.3.7).

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 11.

Recall Lagrange's ray equations (11.2.6), which, in general, can be written as

$$(B.4.1) \quad \frac{\partial \mathcal{L}}{\partial x^i} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^i}, \quad i \in \{1, \dots, n\}.$$

By examining equation (B.3.6), we conclude that

$$(B.4.2) \quad -\frac{\partial \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})}{\partial x^i} = \frac{\partial \mathcal{H}(\mathbf{x}, \mathbf{p})}{\partial x^i}, \quad i \in \{1, \dots, n\},$$

where relations between $\dot{\mathbf{x}}$ and \mathbf{p} are given by expressions (B.3.2) and (B.3.7). Hence, using expressions (B.3.2) and (B.4.2), we can write equation (B.4.1) as

$$-\frac{\partial \mathcal{H}}{\partial x^i} = \frac{dp_i}{dt}, \quad i \in \{1, \dots, n\},$$

which can be immediately restated as

$$(B.4.3) \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x^i}, \quad i \in \{1, \dots, n\}.$$

Thus, we conclude that using the new function, given in expression (B.3.1), and the new variables, given in expression (B.3.2), we obtain expression (B.3.7), while, invoking Lagrange's ray equations (11.2.6), we obtain expression (B.4.3). We notice that the system composed of equations (B.3.7) and (B.4.3) are Hamilton's ray equations (8.2.7); namely,

$$(B.4.4) \quad \begin{cases} \dot{x}^i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x^i} \end{cases}, \quad i \in \{1, \dots, n\}.$$

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.4.4) 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.3.7) 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 Hamiltonian formulation discussed in *Part 2* with the Lagrangian one discussed in *Part 3*. The transformation is explicitly used in Chapters 11 and 13.



⁴In 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* (2nd edition): Springer-Verlag, pp. 1 – 6.

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.0.1. 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
D	material time derivative
δ	variation
δ_{ij}	Kronecker's delta
ϵ_{ijk}	permutation

\in	“belongs to a set”
\rightarrow	“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
$f(\cdot) _a$	function $f(\cdot)$ evaluated at $\cdot = a$

C.2. Physical quantities

C.2.1. Greek letters.

ε_{kl}	strain tensor see expression (1.4.6)
ϑ	phase angle see expressions (9.2.22) and (8.4.5)
θ	ray angle see expressions (8.7.4) and (8.4.12)
κ	compressibility see expressions (5.12.10)
λ	Lamé's parameter see expressions (5.12.2)
μ	Lamé's parameter, also known as rigidity modulus see expressions (5.12.2)
ν	Poisson's ratio see expressions (5.14.31) and (5.14.34)
ξ_{ij}	rotation tensor see expression (1.5.1)
Ψ	rotation vector see expression (1.5.2)
ρ	mass density see expression (2.1.1)
σ_{ij}	stress tensor see expression (2.5.2) and Figure 2.5.1
φ	dilatation see expression (1.4.18)
ϕ	displacement angle see expression (9.2.28)
ω	angular frequency see expression (6.9.2)

C.2.2. Roman letters.

c_{ijkl}	elasticity tensor see expression (3.2.1)
C_{mn}	elasticity-matrix entries, also known as elasticity parameters see expression (4.2.8)
E	Young's modulus see Remark 5.14.7
\mathcal{F}	ray-theory Lagrangian, absolute-value homogeneous of degree 1 in the x_i see expression (13.1.8)
F	ray-theory Lagrangian, inhomogeneous see expression (14.1.1)
\mathcal{H}	ray-theory Hamiltonian, homogeneous of degree 2 in the p_i see expression (8.2.8)
\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 (11.1.1)
\mathbb{L}	classical-mechanics Lagrangian see expression (13.2.2)
W	strain energy see expressions (4.1.3) and (4.1.1)

Index

- absolute-value homogeneous function, *see also* Euler's
 homo-geneous-function theorem *and* homogeneous
 function
- arclength element, 374
- definition, 391
- ray-theory Lagrangian \mathcal{F} , 347, 374, 394
- traveltime integral, 374
- acoustic impedance, 172, 173
- action, *see also* Hamilton's principle
 - classical mechanics, 254, 322, 350, 351, 362
 - definition, **350**
 - least, 351
 - quantum mechanics, 254
 - stationary, 188, 343, 351, 353
- action-at-a-distance, **39**, 40, 64
- adiabatic process, 88
- amplitude
 - displacement, 184, 276, 277
 - interface, 291, 296–298, 300, 301, 304, 305, 307, 308
 - measurement, 271
 - signal, 141
 - transport equation, 187
 - wavefront, 211
- angle
 - deformation, 21, 24–26
 - displacement, 280, 284, 290, 403
 - isotropy, 286, 287
 - Euler's, 105, 106
 - group, *see* ray
 - incidence, 291, 293
 - elliptical velocity dependence, 301–303
 - latitude, 377
 - phase, **182**, 229, 240, 241, 243, 256–258, 278, 280, 284,
 290–293, 316, 375, 403
 - elliptical velocity dependence, 182, 243, 259, 289,
 294, 303, 306
 - interface, 292, 293, 301, 302, 307
 - intersection point, 283
 - isotropy, 286, 287, 306
 - ray parameter, 375
- phase-advance, 302
- phase-delay, 302
- polarization, *see* displacement
- ray, 229, 241–243, 255–258, 284, 289, 290, 293, 316,
 317, 375, 378, 400, 403
 - elliptical velocity dependence, 259, 289, 294–296,
 303, 306, 368, 371, 381
 - interface, 293, 303
 - isotropy, 306
 - ray parameter, 367, 375, 377, 379
 - traveltime, 366
- reflection, 291, 293, 304
- rotation, 103, 109, 112, 114, 130, 402
 - natural coordinate system, 106, 274, 275
- transmission, 291, 293, 304
 - elliptical velocity dependence, 302
- angular frequency, 177, 211, 308, 403
- angular momentum
 - rate of change, 52
- ansatz, *see* trial solution
- antisymmetry
 - rotation tensor, 25, 32
 - tensor, 26, 224
- arclength
 - element, 331, 344, 371, 381

- parameter, 252, 253, 259
- asymptotic
 - method, 149, 187
 - relation, 401
 - series, 142, 187, 212, 214
 - units, 215
 - solution, 209
- atomic structure, 8, 68
- Babich, Vasilii M. (1930–), 142
- balance principle, 33, 36, 55
 - angular momentum, 33–36, 50–52, 54, 55, 67, 85, 225
 - spatial description, 35
 - conservation principle, 34
 - electric charge, 33
 - energy, 33, 55, 79, 82, 301, 308
 - irreversibility
 - entropy, 33
 - linear momentum, 4, 33–36, 39, 40, 43, 45, 46, 50–52, 55, 86, 87, 188, 205
 - spatial description, 35
 - magnetic flux, 33
 - mass, 4, 33–36, 50, 52, 55, 86, 87, *see also* conservation of mass
- basis
 - symmetric matrix, 95
 - trigonometric polynomial, 115, 129
- Beltrami's identity, **314, 327**
 - classical-mechanics Lagrangian, 352
 - Euler's equation, 327, 330, 336
 - Lagrange's equations of motion, 352
 - ray parameter, 366, 368, 377
 - ray-theory Lagrangian \mathcal{F} , 374
 - ray-theory Lagrangian \mathcal{L} , 314, 315, 347, 358
- Beltrami, Eugenio (1835–1900), 314
- Benndorf, Hans (1870–1953), xvii
- Bernoulli, Daniel (1700–1782), 154
- Bernoulli, Johannes (1667–1748), 322
- body force, **39**, 47, 48, 50, 54, 55
 - Cauchy's equations of motion, 47, 74
 - conservation of linear momentum, 39
 - conservative system, 47
 - stress tensor, 43
 - Cauchy's tetrahedron, 43
 - surface force, 39, 48
- body wave, *see* wave
- boundary condition
 - wave equation, 156
- boundary conditions, 325, 326, 354, 362
 - dynamic, 296–298, 300, 304, 305
 - kinematic, 296, 297, 300, 304, 305
- calculus of variations, xxii, 141, 311, 314, 322, 389
 - conserved quantity, 374
 - Lagrange's ray equations, 345
 - physical application, 343
 - ray theory, 323, 334, 335, 356
 - stationarity condition, 321, 323
 - definite integral, 323
 - minimum/maximum, 335
- Cauchy's equations of motion, *see also* Hamilton's
 - equations of motion, Lagrange's equations of motion *and* Newton's law of motion
 - balance of angular momentum, 53
 - balance of linear momentum, 33, 46, 87, 141
 - body force, 47, 50
 - conservation of linear momentum, 40, 46
 - constitutive equations, 143
 - elastic continuum, 143
 - Hamilton's ray equations, 321, 400
 - one-dimensional continuum, 74
 - stress-strain equations, 56, 141
 - anisotropic inhomogeneous, 179, 187, 209
 - isotropic homogeneous, 144, 187, 209
 - isotropic inhomogeneous, 206, 207
 - surface force, 47
 - three-dimensional continuum, 55
 - units, 47
 - unknowns, 33, 47
 - wave equation, 56, 73, 150, 205
- Cauchy's first law of motion, 47
- Cauchy's second law of motion, 54
- Cauchy's stress principle, 11, 38, 41
- Cauchy's stress tensor, 45, *see also* stress tensor
- Cauchy's tetrahedron, 42, 43, *see also* tetrahedron
- Cauchy, Augustin-Louis (1789–1857), 4, 8, 38
- causality, 292, 321
- central force, 50–52

- Červený, Vlastislav (1932–), 142
- characteristic equation
 - eikonal equation, 225, 226, 228
 - plane wave, 204
 - wave equation, 155, 204, 205
- characteristics
 - curve, 155, 230, 231, 233, 234, 261
 - equations, 229, 231, 233–235, 252, 255–257, 259–261, 358, 359
 - method, 229, 230
- Christoffel's equations, 218, 267–272, 275, 276
- Christoffel's matrix, 219, 269
 - positive definiteness, 219, 220, 225
 - symmetry, 219, 224
 - orthogonal eigenvectors, 270
 - real eigenvalues, 270
 - transversely isotropic continuum, 276
- classical continuum mechanics, 54, 63
- classical mechanics, 350
 - action, 254, 343
 - geometrical optics, 350
 - Hamilton's equations
 - conservation of energy, 375
 - Lagrange's equations, 335
 - least action, 351
 - ray theory, 254
 - stationary action, 351
 - wave equation, 350
- compact support, 160, 168, 174, 176, 202
- complex
 - conjugate
 - reflection, 309
 - function
 - eikonal and transport equations, 185
 - number
 - evanescent wave, 302
 - reflection and transmission, 301
 - transverse isotropy, 130
 - plane
 - phase shift, 302
- compressibility, 120, 121, 403
- compression, 24, 41, 42, *see also* tension
- condensed-matter physics, 88
- conservation of mass, *see also* balance principle
 - balance of mass, 35
 - balance principles, 33, 55, 86
 - equation of continuity, 4, 33, 35, 87
 - moving-volume integral, 36, 52
 - spatial description, 35
 - symmetry of stress tensor, 52
- conservation principle
 - balance principle, 34
 - space-time, 36
- conservative system, 47, 80, **80**, 89, 351
- conserved quantity, 237, 292–295, 302, 304, 366, 367, 372, 374, 375, 377, 379
- constant phase
 - phase velocity, 219
 - wavefront, 181, 292
- constitutive equations, *see also* Hooke's law *and* stress-strain equations
 - balance of energy, 55
 - elastic continuum, 55, 56, 349
 - elastic material, 56
 - empirical relation, 63
 - Hooke's law, 65
 - isotropic homogeneous continuum, 143
 - matrix form, 68
 - phenomenology, 86
 - principle of determinism, 64
 - principle of local action, 64
 - principle of material frame indifference, 64
 - principle of objectivity, 64
 - stress-strain equations, 70, 73
- constructive interference, 149, 349
- continuity, *see also* equality
 - function, 57
 - phase, 291, 292, 296, 304
 - wavefront, 292
 - phase slowness, 293
- continuity equation, *see* equation of continuity
- continuum mechanics, 3
 - anisotropy/inhomogeneity, 68
 - atomic structure, 8
 - axiomatic format, 9
 - primitive concepts, 9
 - balance principle, 33
 - Cauchy's stress principle, 11, 38

- deformation, 26, 33
- displacement, 11
- emergence, 8
- exact solution, 4
- granular structure, 8
- history, 4
- Hooke's law, 65
- nonrelativistic, 10, 64
- particle mechanics, 38
- physical reality, 7
- primitive concepts, 9
 - consistency, 10
 - manifold of physical experience, 10, 11, 64
 - material body, 10
 - system of internal forces, 11
 - weak deductive completeness, 10
- seismology, xxi, 3, 63
- stress tensor, 45
- convective derivative, 14
- coordinate plane
 - orthotropic continuum, 107
 - traction, 41, 42, 48
- coordinate system, *see also* material coordinates and spatial coordinates
- characteristics, 154
- curvilinear, 398
 - Laplacian, 146
 - strain tensor, 21
- elasticity parameter, 87
- natural, 106
 - definition, **105**
 - generally anisotropic continuum, 105
 - isotropic continuum, 118
 - monoclinic continuum, 105, 106, 110, 272, 274, 275
 - orthotropic continuum, 107, 108
 - pure-mode direction, 107
 - tetragonal continuum, 111
- orientation
 - Euler's angle, 106
 - strain energy, 113
- orthonormal, 29, 41, 94, 105, 106, 398
 - calculus of variations, 333
 - Jacobian, 12
 - Laplacian, 145, 146
 - raytracing, 370
 - point symmetry, 102
 - polar, 241, 377
 - reference
 - material symmetry, 99
 - tetrahedron, 44
 - transformation, 24, 29, 31, 59, 60, 93, 103, 372
 - isotropic continuum, 120
 - isotropic tensor, 119, 120
 - Jacobian, 94, 124
 - material symmetry, 93, 94, 100, 101
 - orthogonal, 94, 124
 - strain energy, 83, 84
 - transverse isotropy, 129
- coupling, *see also* decoupling
 - wave, 305
- covector, 398
- crystal lattice, 68, 123
- crystallography, 123
- cubic continuum, xxv, 106, 116, 117, 122
- Curie's symmetry group, 122
- curl, 145, 151–153, 401
 - S* wave, 147, 150
 - displacement potential, 151, 153
 - rotation, 25, 26, 32
 - shape change, 145
- d'Alembert, Jean Le Rond (1717–1783), 154, 173
- dashpot, 71
- Debye, Pieter Debye (1884–1966), 142
- decoupling, *see also* coupling
 - SH* wave
 - transversely isotropic continuum, 305
 - wave, 305
- dielectric tensor, 66
- diffeomorphism
 - Legendre's transformation, 314
- differential equation, xxii, 74, 75, 389
 - balance of angular momentum, 53
 - balance principles, 63, 87
 - continuum mechanics, 3
 - mathematical physics, 143
 - method of characteristics, 229
 - ordinary

- Beltrami's identity, 368
- calculus of variations, 321
- characteristic equations, 230, 233
- Euler's equation, 325, 327–330, 373
- first integral, 333, 334
- Hamilton's equations, 341, 360, 361
- Hamilton's ray equations, 237, 311, 313
- Lagrange's ray equations, 311, 313
- Newton's second law, 360
- ray parameter, 368, 369
- partial, 231, 233
 - anisotropic inhomogeneous continuum, 210
 - Cauchy's equations of motion, 47
 - eikonal equation, 187, 220, 229, 230
 - equation of continuity, 35
 - Euler's equation, 327, 328
 - system, 87
 - wave equation, 148, 155, 176, 179, 183
- WKB method, 254
- dilatation, 24, 145, 403
 - P wave, 146
 - definition, **24**
 - displacement potential, 152
 - equations of motion, 145, 207
 - Lamé's parameters, 120
 - scalar, 31
 - trace, 31
- dilatational wave, *see* wave
- Dirac's delta, **202**
 - derivative, 203
 - distribution, 174
- displacement, xxii, 270–272
 - P wave, 183
 - SH wave, 183
 - SV wave, 183
 - amplitude, 17
 - direction
 - measurement, 271
 - gradient, 17, 20, 26, 30, 31
 - infinitesimal, 15–18, 28, 271
 - Lagrange's density, 353
 - longitudinal, 354
 - material point, 11, 12, 17, 271
 - potential, 143, 149, 151, *see also* scalar potential *and* vector potential
 - reduced wave equation, 177
 - strain tensor, 21, 23
 - transverse, 362, 363
 - vector, 18, 19, 25, 47, 66, 207, 210, 270–274, 279, 280
 - SH wave, 279
 - S wave, 280
 - qP wave, 280
 - conservation of linear momentum, 39
 - conservation of mass, 33, 36
 - coordinate transformation, 29
 - curl, 25, 32
 - direction, 107, 218, 267, 271–276, 279–281, 283–285, 287, 305
 - divergence, 24
 - equations of motion, 49, 50, 144
 - interface, 291, 296, 297, 299, 301, 302, 304
 - plane wave, 148, 149
 - strain energy, 81, 89
 - strain tensor, 70
 - velocity, 17, 156, 157
- distribution, **174**, *see also* distribution theory *and* generalized function
 - wave equation
 - solution, 38, 188
- distribution theory, 174, 188, *see also* distribution *and* generalized function
- divergence, 35, 53, 145, 147, 152, 401
 - P wave, 146, 150
 - dilatation, 24, 25, 152
 - displacement potential, 151, 152
 - space-time, 36
 - stress tensor, 47
 - volume change, 145
- divergence theorem
 - balance of angular momentum, 53
 - balance of linear momentum, 46
 - conservation of mass, 34
 - time derivative of volume integral, 37
- eigenspace
 - anisotropic continuum, 285
 - isotropic continuum, 285

- eigenvalue
 - phase velocity, 218, 270, 273, 283
 - solvability condition, 269
 - stability conditions, 85, 137
- eigenvector
 - displacement, 218, 270, 271, 274, 283
- eikonal equation, **186, 219, 269**
 - anisotropic inhomogeneous continuum, 186, 209, 217, 218, 220, 221, 225, 226, 229, 230, 236–238, 257, 267, 345
 - Cauchy's equations of motion, 56, 73
 - characteristics, 233
 - inhomogeneous continuum, 226, 228
 - isotropic homogeneous continuum, 264
 - isotropic inhomogeneous continuum, 184, 187, 252, 359
 - scaling factor, 235
 - solution surface, 233
 - wave equation, 185, 187
- eikonal function, 184, 187, 211, 220, 231, 255
- elastic continuum, xxii, 33, 79, *see also* elastic material
 - constitutive equations, 56, 63
 - crystal lattice, 123
 - elasticity parameters, 88
 - equations of motion, 141, 143
 - Hooke's law, 83
 - Legendre's transformation, 400
 - linear, 66, 83, 85, **87**
 - phase-slowness surface, 281, 349
 - convexity, 281
 - point symmetry, 102
 - seismology, 4
 - strain energy, 79–83, 87, 89
 - stress-strain equations, 73, 141
 - symmetry classes
 - partial ordering, 122
 - variational principle, 350, 353, 356
 - wave phenomena, 143
- elastic material, 56, 63, 65, 66, 73, 79, *see also* elastic continuum
- elastic medium, xxii, 3, 63, 141, 188
- elastic wave, *see* wave
- elasticity constant, *see* elasticity parameter
- elasticity matrix, **69, 83**, *see also* elasticity tensor
 - arbitrary coordinate system, 105
 - Christoffel's equations
 - isotropic continuum, 285
 - monoclinic continuum, 272, 275
 - transversely isotropic continuum, 275
- displacement direction
 - isotropic continuum, 286, 287
 - monoclinic continuum, 274, 275
 - transversely isotropic continuum, 279
- elasticity tensor, 68
- formulation, 68, 69
- interface, 299
- invariants, 119, 126
- material symmetry, 93, 94, 99–102, 122
 - generally anisotropic continuum, 103
 - isotropic continuum, 118
 - monoclinic continuum, 104, 106, 127
 - orthotropic continuum, 107, 108
 - tetragonal continuum, 111
 - transversely isotropic continuum, 112, 114
- natural coordinate system, 105
- phase-slowness curve, 286
- positive definite
 - Christoffel's matrix, 225
 - isotropic continuum, 120, 136, 137
 - stability conditions, 84, 85, 91
- strain energy
 - Christoffel's matrix, 224
- symmetry, 122
 - isotropic continuum, 130
 - strain energy, 83, 99
- elasticity parameter, **83**, 404
 - anisotropic continuum
 - Christoffel's equations, 268, 269
 - invariants, 119
 - phase velocity, 267, 271
 - ray velocity, 240
- anisotropic inhomogeneous continuum, 179
 - eikonal equation, 220
- condensed-matter physics, 88
- elastic continuum, 86, 87
- elastic material, 87
- Green-river shale, 285, 290
- interface, 291, 298, 301
- material symmetry, 88, 104, 122, 123

- isotropic continuum, 118, 119, 132
- orthotropic continuum, 108
- tetragonal continuum, 111
- transversely isotropic continuum, 112
- natural coordinate system, 105, 106, 274
- normalized, 269
- phase-slowness curve, 282
- phase-slowness surface, 281
- spring constant, 89
- stability conditions, 84
- strain energy, 83, 84
- temperature dependence, 88
- traveltime, 284
- elasticity tensor, 66, 404, *see also* elasticity matrix
 - anisotropic inhomogeneous continuum, 210
 - components, 66
 - independence, 68, 82, 83
 - stress-strain equations, 70
 - units, 66
 - vanishing, 104, 109
- elasticity matrix, 68
- equations of motion
 - anisotropic inhomogeneous continuum, 211
- isotropic continuum, 120
- property tensor, 66
- rank, 60, 66, 119
- symmetry, 67, 93, 122
 - strain energy, 79, 81
- elasticity theory
 - elastic continuum, 5, 86
 - elasticity tensor
 - isotropic continuum, 120
 - Hamilton's ray equations, 264
 - linearity, 26, 65, 70
 - ray angle, 242
 - ray velocity, 367
 - strain tensor, 20
 - stress tensor, 45
 - stress-strain equations, 63
 - stress/strain, 8, 18, 38
- electric charge, 66
- electric displacement, 66
- electric field, 66
- electromagnetic theory, 154, 331
- elliptical anisotropy, 243, **243**, 244, 367
- elliptical velocity dependence, **183**, 243, **243**
 - interface
 - amplitude, 296
 - angle, 291, 293–296
 - isotropy, 372, 383
 - phase velocity, 368
 - phase-slowness surface, 400
 - ray angle, 258, 259, 289, 317
 - ray parameter, 368
 - ray velocity, 242, 258, 367, 368
 - rays, 243, 246, 367, 370, 379
 - transversely isotropic continuum, 182, 288
 - phase slowness, 181
 - phase velocity, 288
 - traveltime, 243, 246, 367
 - wave equation, 179, 180, 182
 - phase velocity, 182
- ellipticity, 368, *see also* elliptical velocity dependence
- emergence, 8
- energy, 4, 55, 79, 82, 84, 88, 89
 - evanescent wave, 302
 - incident wave, 173, 296, 301, 308
 - kinetic, 159, 350–355, 360, 361, 363, 394
 - potential, 4, 79, 80, 159, 350, 351, 353, 354, 362, 363
 - strain, 79–82, 84–90, 92, 99, 112, 118, 130, 135, 136, 159, 354, 404
 - Christoffel's matrix, 224
 - coordinate transformation, 83, 84, 113, 116
 - units, 159, 237
 - wave function, 159, 160, 191, 192
- entropy, 33
- envelope
 - plane-wave solution
 - elliptical velocity dependence, 181
 - tangent lines, 396
 - tangent planes, 395
- equality, *see also* continuity
 - displacement, 291, 296, 297, 304
 - mixed partial derivatives
 - displacement potential, 151–153
 - eikonal equation, 211, 231
 - equations of motion, 207
 - Euler's equation, 340

- strain energy, 81
 - theorem, 81
 - wave equation, 144, 146, 147, 149, 189, 190, 192
- phase slowness, 292, 300
- stress tensor, 297, 298
- traction, 291, 296–298, 304
- equation of continuity
 - balance principles, 4, 33, 87
 - constitutive equations, 56, 70, 74
 - derivation, 35, 36
 - elastic continuum, 141
 - material time derivative, 37, 57, 58
 - unknowns, 33, 47, 55, 70
- equivoluminal wave, *see* wave
- Euler's angle, 105, 106
- Euler's equation, 323, **325**, **334**, 338, 339, 359, 366, 373, 385
 - Beltrami's identity, 327, 336
 - first integral, 333, 365, 373
 - formulation, 325, 327
 - generalizations, 327–329
 - Hamilton's equations, 341
 - Lagrange's equations of motion, 352, 362
 - Lagrange's ray equations, 334, 335, 345, 348
 - parametric form, 334, 335, 358, 359
 - ray parameter, 367, 368
 - special cases, 329, 331, 332, 366
 - wave equation, 355, 356, 363
- Euler's homogeneous-function theorem, 391, 392, 394, *see also* absolute-value homogeneous function *and* homogeneous function
 - characteristic equations, 235, 236
 - classical-mechanics Lagrangian, 353
 - example, 393
 - Fermat's principle, 346, 347
 - Hamilton's ray equations, 229
 - kinetic energy, 394
 - Legendre's transformation, 389
 - proof, 392
 - ray-theory Lagrangian \mathcal{F} , 374
 - ray-theory Lagrangian \mathcal{L} , 358
 - strain energy, 90
- Euler, Leonhard (1707–1783), 4, 154, 173, 325
- Euler-Lagrange equation, 325, 335, *see also* Euler's equation, Lagrange's equations of motion *and* Lagrange's ray equations
- Eulerian description, 12, *see also* spatial description
- evolution equation
 - solution
 - domain of dependence, 168
 - range of influence, 168
- Fermat's principle, 350, *see also* variational principles
 - calculus of variations, 333, 356, 359
 - interface, 304
 - proof, 343, 348
 - limitations, 349, 356
 - parameter independence, 347, 357, 394
 - statement, **344**
 - quantum electrodynamics, 322
 - ray parameter, 366
 - ray theory, 335
 - variational principles, 314, 322, 350
- Fermat, Pierre de (1601–1665), 322, 344
- first integral, 333–335, 341, 342, 365–369, 373, 376
- fluid, 47, 120, 147
- forces vives, 365, *see also* vis viva
- Fourier's transform, 177–179
- gauge invariance, 331
- gauge transformation, 150
- Gauss's divergence theorem, *see* divergence theorem
- generalized function, 174, **174**, 188, *see also* distribution *and* distribution theory
- generally anisotropic continuum, 103, 105, 122, 123, 179, 182
 - elasticity matrix, 103
 - natural coordinate system, 105
 - phase velocity, 271
 - point symmetry, 103
 - ray angle, 293
 - Snell's law, 293
- geometrical optics, 349, 350
- gradient, 145, 151, 153, 184, 255, 294, 295, 401
 - calculus of variations, 338
 - characteristic equations, 232
 - displacement, 16, 17, 30, 31

- displacement potential, 150
- displacement vector, 26
- eikonal function, 184, 185, 218, 255
- material time derivative, 14
- phase-velocity function, 239
- scalar field, 21
- vector field, 21
- velocity, 17, 263
- grains, 3, 8, 68
- gravitation, 48
- gravitational force, 39
- gravity, 48, 361
- Green, George (1793–1841), 4
- Green-river shale
 - elasticity parameters, 285
 - pure-mode direction, 290
 - ray and displacement angles, 290
- group angle, *see* angle
- group velocity, *see* ray velocity
- halfspace
 - anisotropic homogeneous continuum, 149
 - elliptical velocity dependence, 294
- Hamilton's equations
 - Euler's equation, 341
- Hamilton's equations of motion, 352, 361
- Hamilton's principle, 188, 335, 343, 349–351, **351**, 352, 361–363, *see also* action and variational principles
- Hamilton's ray equations, *see also* Hamilton's equations, Hamilton's equations of motion and Lagrange's ray equations
- analytical solution, xxvi, 243, 245, 247, 251
- anisotropic inhomogeneous continuum, 229, 267, 311, 314, 324
 - conserved quantity, 292, 293, 365, 374, 375
 - Fermat's principle, 343–347, 349, 356
 - ray, 313, 314, 376
 - ray angle, 293
- elliptical velocity dependence, 183, 245, 250, 370, 372
- high frequency, 253
- isotropic homogeneous continuum, 264
- isotropic inhomogeneous continuum, 255
- Legendre's transformation, 312, 313, 315, 316, 389, 399, 400
- method of characteristics, 236, 237
- ray theory, xxvii, 243, 253, 321, 365, 367
- Hamilton, William Rowan (1805–1865), 142, 322, 351
- Hamilton-Jacobi equation, 220
- Hamiltonian, 341
 - classical-mechanics, 237, 353, 360, 361, 375, 404
 - definition, **360**
 - ray-theory, 237, 255, 288, 292, 311, 313, 315, 321, 343–346, 349, 356, 391, 395, 398, 404
 - definition, **237, 245**
- heat equation, 163
- Heaviside's function, 71, 72, 75, 76, 202, 203
- Heckmann diagram, 66
- Helmholtz von, Hermann (1821 – 1894), 177
- Helmholtz's decomposition, 149
- Helmholtz's equation, 177
- Herman's theorem, 116
- hexagonal continuum, *see* transversely isotropic continuum
- high frequency
 - approximation, 180, 253, 254
 - anisotropic inhomogeneous continuum, xxii
 - eikonal equation, 186, 212, 214
 - equations of motion, 141
 - inhomogeneous continuum
 - wave equation, 186
 - signal, 141
- Hilbert, David (1862–1943), 322
- homogeneous continuum
 - anisotropic continuum, 182
 - differential equations, 74
 - displacement potential, 151
 - elliptical velocity dependence, 180
 - equations of motion, 143–146, 148, 151, 207, 210
 - Hamilton's ray equations, 264
 - layered, 291, 292, 331, 365, 372, 375
 - plane wave, 148, 149
 - rays, 265, 330, 378, 379
 - stress-strain equations, 209
 - wave equation, 141, 179, 183, 187, 209, 350, 351
 - extension, 179
- homogeneous equation
 - Christoffel's equations
 - nontrivial solution, 269

- homogeneous function, *see also* absolute-value
 - homogeneous function *and* Euler's
 - homogeneous-function theorem
- arclength element, 357, 374
- classical-mechanics Lagrangian, 353
- definition, 391
- example, 392
- kinetic energy, 394
- multivariable function, 394
- parameter independence, 347
- phase velocity, 220, 235, 236, 239, 344, 392
- ray velocity, 357, 374
- ray-theory Hamiltonian, 344, 346, 391–394, 404
- ray-theory Lagrangian \mathcal{F} , 347, 357, 374, 394
- ray-theory Lagrangian \mathcal{L} , 314, 346–349, 358, 394, 404
- strain energy, 81, 86, 89, 90
- traveltime integral, 374
- homogeneous-function theorem, *see* Euler's
 - homogeneous-function theorem
- Hooke's law, 65, 75, 83, 354, *see also* constitutive
 - equations *and* stress-strain equations
 - constitutive equations, 65
 - history, 65
 - linearity, 66
- Hooke, Robert (1635–1703), 65
- Hookean solid, 65, 70, 72
- Huygens' principle, 166, 177, 220, 292
- hydrostatic pressure, 46, 121
- incident wave, *see* wave
- initial condition
 - characteristic equations, 233
 - ray, 370
 - wave equation, 154–156, 163
- integral equation
 - balance of angular momentum, 51
 - balance of linear momentum, 46
 - conservation of angular momentum, 52
 - conservation of linear momentum, 39
 - conservation of mass, 34, 35
 - vanishing of integrand, 57
 - wave propagation, 188
- interface, 149, 291, 293
 - amplitude, 296
 - boundary conditions, 296–298, 300, 301, 304, 305
 - conserved quantity, 291, 292, 294
 - energy, 296
 - energy transmission, 301, 302
 - halfspace, 294
 - Snell's law, 292, 293
 - wave, 187
 - wavefront, 292
 - welded contact, 296
- interface wave, *see* wave
- inverse
 - derivative, 30
 - function, 330
 - Legendre's transformation, 399
 - mapping, 12
 - matrix, 101
 - problem, 284, 372
 - transformation, 27, 60
 - transpose, 29
 - trigonometric function, 287
- isothermal process, 88
- isotropic continuum, 106, 118, 122
 - characteristic equations, 259, 261, 264, 358
 - Christoffel's equation, 285
 - displacement, 271, 272, 285–287
 - displacement potential, 151
 - eikonal equation, 186, 220
 - elasticity matrix, 130, 131
 - elasticity parameters, 118
 - Lamé's parameters, 119
 - elasticity tensor, 119, 120
 - equations of motion, 141, 143–145, 187, 207, 209, 210
 - Euler's equation, 330, 332, 333, 358, 359
 - Fermat's principle, 359
 - interface, 296, 307
 - intersection point, 283, 286
 - Lamé's parameters, 118, 132
 - natural coordinate system, 118
 - orthogonal transformation, 118
 - Poisson's ratio, 132, 133
 - ray, 370, 373, 378, 384
 - ray equations, 252, 253, 255
 - ray parameter, 306
 - rigidity modulus, 133

- strain energy, 135, 136
- stress-strain equations, 119, 120, 131, 132, 209
- transport equation, 220
- wave equation, 143, 146, 148, 149
 - extension, 179
- Young's modulus, 134
- isotropic tensor
 - definition, **119**
 - isotropic continuum, 119, 120, 131
- Jacobian, **27**, 94, 124, **124**, 402
- Keller, Joseph B. (1923–), 142
- Kelvin-Voigt model, 71, 72
- kinetic energy, *see* energy
- Kirchhoff, Gustav (1824–1887), 177
- Kronecker's delta, 54, 91, 96, 97, 131, 132, 134, 136, 144, 206, 401
- Lagrange's equations of motion, 51, 350, 352, 353, 362
- Lagrange's ray equations
 - anisotropic inhomogeneous continuum, 314, 324
 - Fermat's principle, 343, 345, 347, 348, 365
 - ray, 313, 376
 - singularity, 314
 - two-dimensional, 373
- Beltrami's identity, 314, 315, 327
- conserved quantity, 293, 375
- Euler's equation, 323, 334, 335
- Euler-Lagrange equation, 335
- Legendre's transformation, 312, 313, 316, 321, 389, 399, 400
- ray theory, xxvi, xxvii, 243, 311, 321
- Lagrange, Joseph-Louis (1736–1813), 154, 324, 328
- Lagrangian
 - classical-mechanics, 351, 353, 360, 361, 404
 - continuous system, 353, 355, 363
 - definition, **350**
 - Noether's theorem, 334
 - ray-theory, 356
 - Fermat's principle, 343
 - ray-theory \mathcal{F} , 348, 357–359, 404
 - definition, **347**
 - Fermat's principle, 348, 349
 - ray-theory \mathcal{L} , 311, 312, 314–317, 345, 347, 348, 358, 395, 398, 400, 404
 - definition, **312**
 - properties, 345–347
 - ray-theory F , 404
 - definition, **366**
- Lagrangian density, 353, 355, 356
- Lagrangian description, 12, *see also* material description
- Lamé's parameters, 118–121, 131, 132, 135, 136, 146, 147, 206, 286, 403
- Lamé's theorem, 152
- Laplace's equation, 163
- layered medium, 183, 275, 331, 366
- Legendre's transformation, xxii, 400
 - classical mechanics, 360
 - conserved quantity, 375
 - duality, 313, 341
 - elliptical velocity dependence, 242, 293
 - Euler's equation, 341
 - Euler's homogeneous-function theorem, 389
 - Fermat's principle, 346, 356
 - regularity, 349
 - formulation
 - definition, 395
 - duality, 397
 - geometrical context, 395, 396
 - limitations, 396
 - ray theory, 399, 400
 - ray-theory Lagrangian \mathcal{L} , 398
 - single variable, 396, 397
 - singularity, 400
- Hamilton's ray equations, 237
- Lagrange's ray equations, 311, 313, 314, 321, 389
- ray angle, 289, 316
- ray velocity
 - elliptical velocity dependence, 367
- ray-theory Hamiltonian
 - regularity, 313, 314
- ray-theory Lagrangian \mathcal{L} , 311, 312
- transformation variable, 238
- level curve
 - phase slowness, 294
- level set
 - characteristic equations, 229–231, 255, 256

- eikonal function, 184, 211, 255
- wavefront, 211, 218
- linear momentum
 - balance, *see* balance principle
 - Hamilton's equations of motion, 360, 361
 - Legendre's transformation, 360
 - rate of change, 39, 40, 43, 51
- linear stress-strain relation, 63, *see also* constitutive
 - equations, Hooke's law *and* stress-strain equations
- linear velocity dependence, 243, 246, 367–370, 384
- linear-momentum density, 51
- linearity
 - differential operator
 - balance of angular momentum, 53
 - Beltrami's identity, 337
 - displacement potential, 151–153
 - equations of motion, 207
 - Euler's equation, 339
 - wave equation, 144, 146, 147
 - integral operator
 - balance of linear momentum, 46
 - calculus of variations, 338
 - superposition principle, 29
- linearized theory, 15–17
 - balance of angular momentum, 52, 53
 - Cauchy's equations of motion, 47, 49, 50, 70
 - elasticity, 26
- longitudinal direction, *see* pure-mode direction
- longitudinal wave, *see* wave
- mass, *see also* balance principle
 - density, 47, 403
 - Cauchy's equations of motion, 49
 - Cauchy's tetrahedron, 43
 - Christoffel's equations, 267, 269, 271
 - conservation of mass, 33, 34
 - elastic continuum, 83, 87
 - equations of motion, 55, 56
 - Green-river shale, 290
 - inhomogeneous continuum, 206, 210
 - interface, 298, 300, 301, 307
 - material time derivative, 36
 - normalization, 269
 - one-dimensional continuum, 355
 - phase velocity, 220, 267
 - primitive concepts, 10
 - ray theory, 291
 - ray velocity, 240
 - units, 61
 - wave equation, 146, 147
 - gravitational, 64
 - inertial, 64
- material coordinates, 12, 14, 22, 24, 27, 70
- material description, 11–13, 15, 27, 35
 - definition, **11, 12**
- material point, 10–14, 17, 27, 271
 - definition, **10**
- material symmetry, 93, 94, 99, 100, 102, 112, 113, 123, *see also* point symmetry, generally aniso-tropic
 - continuum, monoclinic continuum, orthotropic continuum, tetragonal continuum, transversely isotropic continuum *and* isotropic continuum
 - definition, **93**
 - elasticity matrix, **99**
 - strain energy, **113**
- material time derivative, 13–16, 401
 - operator, 14, **14**, 37, 40, 47, 52, 57, 58
- Mauertuis, Pierre-Louis Moreau de (1601–1665), 322, 351
- Maxwell model, 72
- Maxwell's equations, 154
- methodological holism, 8
- methodological individualism, 8
- momentum phase space, *see* **xp**-space
- monoclinic continuum, 103, 108, 117, 122, 267, 271, 272
 - displacement, 272
 - elasticity matrix, 104
 - natural coordinate system, 105, 106, 274
 - phase velocity, 272
 - pure-mode direction, 107
- natural coordinate system, *see* coordinate system
- Newton's law of motion, 36, 352
 - second, 39, 40, 49, 50, 154, 350, 352, 360, 361
 - third, 42, 43, 50, 51, 297, 298
 - strong, 51, **51**, 52, 54
 - weak, **40**, 51

- Noether's theorem, 334, 375, *see also* conserved quantity,
first integral *and* ray parameter
- Noether, Emmy (1882–1935), 334
- Noll, Walter (1925–), 4
- orthogonal matrix, 29, 94, 101
- orthorhombic continuum, *see* ortho-tropic continuum
- orthotropic continuum, 106–108, 111, 117, 122
elasticity matrix, 108
- oscillatory motion, 177
- Ostrogradsky theorem, *see* divergence theorem
- P* wave, *see* wave
- particle, 8, 10–12, 321, 349, 352, 353
- particle mechanics, 12, 36, 38, 39, 352
- permutation symbol, 53, 54, 401
- phase advance, 302
- phase delay, 302
- phase factor
interface
evanescent wave, 302
phase shift, 302
plane wave, 297
sign convention, 302
- trial solution
anisotropic inhomogeneous continuum, 211
elliptical velocity dependence, 177, 304
isotropic inhomogeneous continuum, 184
plane wave, 184
- phase shift, 302
- phase slowness, *see also* phase-slow-ness
covector, 398
magnitude
Christoffel's equations, 268
eikonal equation, 229
elliptical velocity dependence, 182
Snell's law, 292
rate of change
Hamilton's ray equations, 237
- phase space
momentum, *see* $\mathbf{x}\mathbf{p}$ -space
velocity, *see* $\mathbf{x}\dot{\mathbf{x}}$ -space
- phase velocity, 17, **181**
eikonal equation, 219
- elliptical velocity dependence, 258, 259
eikonal equation, 243
wave equation, 181, 182
- expression
monoclinic continuum, 273
transversely isotropic continuum, 277, 278
- Hamilton's ray equations
ray parameter, 375
- magnitude, 239, 240, 242, 243, 268, 271, 281
elliptical velocity dependence, 288
interface, 292–294, 302
- phase-slowness, *see also* phase slowness
curve, 241
interface, 293–295, 300
intersections, 282, 284, 286
polar reciprocity, 241, 316
- surface, 268, 281
bicubic equation, 281
Fermat's principle, 349, 356
interface, 293, 295
intersections, 282
Legendre's transformation, 400
polar reciprocity, 241
properties, 281
sheets, 281, 283
- vector, 184, 257
direction, 219, 271
eikonal function, 184
elliptical velocity dependence, 181, 182
gradient, 239
interface, 292, 293, 296, 299–301
isotropic inhomogeneous continuum, 358
Legendre's transformation, 316, 317
magnitude, 264, 292
surface, 281
wave equation, 181
wavefront, 218
- Planck's constant, 254
- plane wave, *see* wave
- point symmetry, 102, 103, 107, 108, 121
- Poisson's ratio, 132–134, 136, 137, 403
- polar reciprocal, 182, 241, 317
- potential energy, *see* energy
- pressure, *see also* hydrostatic pressure

- pressure wave, *see* wave
- primary wave, *see* wave
- pure-mode direction, 107, 274, **274**, 290
- qP wave, *see* wave
- qS wave, *see* wave
- qSV wave, *see* wave
- quantum electrodynamics, 265, 293, 301, 322
- quantum mechanics, 254, 322, 350
- ray angle, *see* angle
- ray equation, *see* Hamilton's ray equations *and* Lagrange's ray equations
- ray parameter, 365, *see also* conserved quantity *and* first integral
 - Beltrami's identity, 327
 - first integral, 334, 365, 376
 - elliptical velocity dependence, 368
 - generally inhomogeneous continuum, 376
 - Hamilton's ray equations, 237, 293
 - integration, 368
 - Lagrange's ray equations, 374, 379
 - lateral symmetry, 331
 - anisotropic continuum, 366, 367
 - elliptical velocity dependence, 368
 - ray, 369
 - ray equations, 374, 375
 - spherical symmetry, 377
 - traveltime, 371
- ray theory, xxi, xxii, xxvi, 141, 142, 179
 - asymptotic methods, 187
 - calculus of variations, 321, 323, 332, 334, 335
 - eikonal equation, 209
 - Fermat's principle, 343, 356
 - first integral, 334
 - Hamilton's ray equations, xxvii, 229, 252, 253, 267, 321, 365
 - interface, 291
 - Lagrange's ray equations, xxvii, 254, 321, 335, 365
 - natural coordinate system
 - pure-mode direction, 107
 - ray, **237**
 - ray parameter, 365, 374
 - variational principles, 343, 350
 - wave theory, 254
 - WKB method, 254
- ray velocity, 17, 183, 229, 237, 239, **239**, 243
 - conserved quantity, 294–296
 - curve
 - elliptical velocity dependence, 317
 - polar reciprocity, 241, 316, 317
 - two-dimensional continuum, 241
 - expression, 293
 - elliptical velocity dependence, 242, 258, 259, 317, 368, 400
 - linear velocity dependence, 368, 369
 - function
 - homogeneity, 374
 - time invariance, 374
 - magnitude, 239, 242
 - elliptical velocity dependence, 246
 - Fermat's principle, 344, 348
 - isotropic continuum, 306, 373
 - isotropic inhomogeneous continuum, 359
 - two-dimensional continuum, 240
 - phase velocity, 375
 - ray parameter, 375
 - elliptical velocity dependence, 367, 368
 - two-dimensional continuum, 367
 - surface, 241
 - traveltime integral
 - two-dimensional continuum, 366
 - vector, 242, 258, 398
 - magnitude, 259
 - wavefront slowness, 246
- rayfield, 375
- raytracing, 314, 335, 356, 373
- receiver, *see also* source
 - constructive interference, 349
 - displacement vector, 271, 296
 - ray parameter, 371, 372
 - traveltime, 321, 371, 372, 381, 383
- reductionism, 8
- reflected wave, *see* wave
- reflection angle, *see* angle
- rigidity, 120, **120**, 121, 147, 403
- Rivlin, Ronald (1915–2005), 4
- rotation

- tensor, 25, 403
 - antisymmetry, 25, 26
- vector, 25, 26, 403
 - S wave, 147
 - curl, 32
 - displacement potential, 152, 153
 - equations of motion, 145, 146, 207
- rotational wave, *see* wave
- Rudzki, Maurycy Pius (1862–1916), xvii, xxii, xxiii
- Runge, Carl (1856–1927), 142
- S wave, *see* wave
- scalar potential, *see also* vector potential
 - classical mechanics, 360
 - displacement, 150–152
- scaling factor
 - characteristic equations, **232**, 234–236, 252, 259
- scattering, 168, 172, 173
- Schwartz, Laurent (1915–2002), 174
- secondary wave, *see* wave
- SH wave, *see* wave
- shear wave, *see* wave
- singularity, *see also* caustic
 - Legendre’s transformation, 314, 356, 400
- Snell’s law, *see also* reflection angle *and* transmission angle
 - elliptical velocity dependence, 296
 - generally anisotropic continuum, 293
 - isotropic continuum, 307
 - philosophical insight, 293
 - physical insight, 293
 - quantum electrodynamics, 293
- Snell, Willebrord Van Roijen (1591–1626), 142
- Sobolev, Sergei Lvovich (1908–1989), 173, 174
- solid, 121, 137, 147
- solution surface, 229–233
- solvability condition, 269, 270, 273
- Sommerfeld, Arnold (1868–1951), 142
- source, *see also* receiver
 - coordinate system, 370
 - displacement vector, 271
 - elliptical velocity dependence, 370
 - plane wave, 149, 305
 - ray parameter, 371, 372
 - traveltime, 321, 371, 372, 381, 383
- spatial coordinates, 12, 14, 15, 17, 27, 28, 70
- spatial description, 11–13, 15–17, 27, 35
 - definition, **11**, **12**
- spring, 65, 71
- spring constant, 89
- stability conditions, 84, 85
 - Christoffel’s matrix, 225
 - elasticity matrix, 84, 85, 121
 - isotropic continuum, 136, 137
 - transversely isotropic continuum, 127–129
- stationarity, *see* calculus of variations *and* variational principles
- steady-state equation, 163, 177
- stiffness matrix, *see* elasticity matrix
- stiffness tensor, *see* elasticity tensor
- Stokes, George Gabriel (1819–1903), 70
- Stokesian fluid, 70–72
- strain energy, *see* energy
- strain tensor, **21**, 403
 - components
 - column matrix, 69
 - displacement vector, 70
 - elasticity tensor, 67
 - independence, 67, 68
 - isotropic continuum, 134, 135
 - square matrix, 98
 - strain energy, 81, 89, 91, 115
 - stress-strain equations, 66, 70, 73
 - transformation, 98, 99, 113
 - definition
 - infinitesimal displacement, 21
- deformation, 7, 18
- derivation, 18–21
- equations of motion
 - anisotropic inhomogeneous continuum, 210
 - isotropic homogeneous continuum, 144
- Hooke’s law, 65
- physical meaning, 21
 - components, 21
 - deformation, 21
 - length change, 22
 - shape change, 24
 - volume change, 24

- property tensor, 66
- rank, 21, 60, 66
- rotation tensor, 26
- symmetry, 21
 - constitutive equations, 68
 - elasticity tensor, 67
 - matrix form, 74
 - stress-strain equations, 67, 120, 131, 132
- trace, 24, 31
- units, 21, 66
- stress tensor, **45**, 403
 - Cauchy's equations of motion, 46
 - components, 47, 62
 - anisotropic inhomogeneous continuum, 210
 - column matrix, 69
 - compressional, 46
 - Hooke's law, 66
 - independence, 47, 54, 67
 - interface, 297, 298
 - isotropic continuum, 134, 136
 - shear, 46
 - square matrix, 95
 - stress-strain equations, 67, 68, 70, 73, 82
 - tensile, 46
 - transformation, 95, 97, 101, 127
 - conservation of linear momentum, 40
 - couple, 51
 - derivation, 41–45
 - equations of equilibrium, 47
 - equations of motion
 - anisotropic inhomogeneous continuum, 209
 - isotropic homogeneous continuum, 144
 - force, 18, 33
 - Hooke's law, 65, 66
 - nonsymmetric
 - balance of angular momentum, 54
 - property tensor, 66
 - rank, 59, 66
 - strain energy, 80
 - symmetry, 54, 55, 61, 298, 299
 - balance of angular momentum, 52–54, 225
 - constitutive equations, 68
 - elasticity tensor, 67
 - stress-strain equations, 67
- traction
 - direction, 45, 46
 - orientation, 45
 - units, 61, 66
- stress vector, *see* traction
- stress-strain equations, *see also* constitutive equations *and*
 - Hooke's law
- Cauchy's equations of motion, 74, 87, 141, 179, 209
- Christoffel's equations, 271
- constitutive equations, 73
- dilatation, 24
- displacement vector, 70
- elastic continuum, 85, 87, 94, 141
- elastic material, 73
- equations of motion
 - anisotropic inhomogeneous continuum, 209, 210
 - isotropic homogeneous continuum, 143, 144
 - isotropic inhomogeneous continuum, 206
- formulation, 65–69
 - Cauchy's approach, 66, 80
 - Green's approach, 66, 80
 - matrix form, 67, 69, 74, 102
 - tensor form, 68
 - vector form, 98
- generally anisotropic continuum, 103
- Hooke's law, 65
- infinitesimal displacement, 70
- interface, 298
- isotropic continuum, 119, 120, 131, 132
 - Lamé's parameters, 120, 121, 134
 - Poisson's ratio, 133
 - strain energy, 135
 - Young's modulus, 134
- linearity, 4, 73
- monoclinic continuum, 127
- orthotropic continuum, 107
- quotient rule, 60
- strain energy, 82–84, 91
 - spring constant, 89
- transformation, 93–100
 - invariance, 101
 - point symmetry, 102
- wave equation, 205
- superposition principle, 29

- surface force, **38**, 39
 - body force, 39, 48
 - Cauchy's equations of motion, 47
 - stress, 38
 - stress tensor, 41, 43
 - Cauchy's tetrahedron, 43
 - traction, 39
- surface wave, *see* wave
- symmetrization, 114
- symmetry, 12, 88, 93, 94, 122, 123, 127, 131, 268, 293, 305, *see also* antisymmetry, material symmetry *and* symmetrization
- angular velocity dependence, 367
- axis
 - five-fold, 112, 114
 - four-fold, 111
 - monoclinic continuum, 267, 271–274
 - tetragonal continuum, 111
 - transversely isotropic continuum, 275, 276, 278, 279, 282, 283, 296, 300, 301, 305
- Christoffel's equations, 270
- class
 - Curie's symmetry group, 122
 - elasticity parameters, 122
 - identification, 119
 - partial ordering, 122
 - relations among, 117, 121
 - symmetry group, 121
- conserved quantity, 375, 376
- coordinate transformation
 - strain energy, 84
- elasticity matrix, 118, 130
 - stability condition, 85, 137
 - strain energy, 83, 99
- elasticity tensor, 85, 93, 211
 - strain energy, 79, 81
 - stress-strain equations, 67, 68
- elliptical velocity dependence, 293, 294, 306
- group, 94, 102, 103, 107, 109, 111–114, 116, 118, 127
 - cubic continuum, 117
- matrix, 95
- plane, 105
 - monoclinic continuum, 105, 272
 - orthotropic continuum, 106, 107
 - tetragonal continuum, 111
- reflection, 104, 127
- rotation
 - transversely isotropic continuum, 282
- strain energy, 112, 116
- strain tensor, 21, 66–69, 89, 98, 120, 131, 132
 - stress-strain equations, 67, 69
- stress tensor, 47, 52, 54, 55, 61, 66–69, 95, 298, 299
 - Christoffel's matrix, 225
 - stress-strain equations, 67, 69
- tensor, 26, 224
- tangent space, *see* $\mathbf{x}\mathbf{x}$ -space
- Taylor's series
 - displacement, 16, 19, 22
- tension, 41, 42, 65, 362, *see also* compression
- test function, 164, 174–176
- tetragonal continuum, 106, 111, 117, 122
 - elasticity matrix, 111
- tetrahedron, 42–44, 48
- torque, 51
- total differential
 - calculus of variations, 331, 338
 - strain energy, 82
- traction, 11, 38, 41–46, 59, 60, 62, 291, 296–298, 304
 - Cauchy's stress principle, 38
 - definition, **38**
 - surface force, 38, 39
- transformation, *see also* Legendre's transformation
 - coordinate, 24, 31, 83, 84
 - material symmetry, 93
 - material/spatial, 27
- matrix, 27, 29–31, 59
 - generally anisotropic continuum, 121
 - isotropic continuum, 118, 121
 - Jacobian, 27, 124
 - orthotropic continuum, 107
 - point symmetry, 102
 - strain energy, 113, 114
 - tetragonal continuum, 111
 - transversely isotropic continuum, 112, 130
- orthogonal, 93, 94, 103, 109, 125
 - eigenvalues, 119
 - elasticity matrix, 99–101

- orthonormal coordinate system, 93, 94
 - strain tensor, 98, 99
 - stress tensor, 95, 97
 - stress tensor, 60
- transmission angle, *see* angle
- transmitted wave, *see* wave
- transport equation
 - anisotropic inhomogeneous continuum, 217, 220
 - isotropic inhomogeneous continuum, 187, 220
- transverse wave, *see* wave
- transversely isotropic continuum, 106, 112, 122, 123, 268, 272, 275, 291
- SH* wave
 - ray-theory Hamiltonian, 288
- S* wave, 288
- Christoffel's equations, 275, 276
 - displacement, 276, 279
 - phase velocity, 278
- elasticity matrix, 112
- elasticity parameters, 276
- elliptical velocity dependence
 - phase velocity, 182
 - ray velocity, 242, 259
- interface, 296, 304, 305
 - amplitude, 300, 301
 - boundary conditions, 296, 299, 300
- phase-slowness curve
 - intersection point, 282
- phase-slowness surface, 281, 282
- rotation invariance, 116
- stability conditions, 127, 129
- traveltime
 - eikonal function, 218
 - elliptical velocity dependence, 251, 261, 262, 371, 380–382
 - downgoing signal, xxvii, 251, 262, 371, 372, 381–383
 - upgoing signal, xxvii, 251, 371, 372, 381, 382
- equation
 - isotropic inhomogeneous continuum, 373
- expression
 - elliptical velocity dependence, 372, 381, 383
 - ray parameter, 376
- Hamilton's ray equations, 243
- integral, 321, 333, 357
 - conserved quantity, 365
 - elliptical velocity dependence, 371, 381
 - Fermat's principle, 343
 - first integral, 367
 - invariance, 384
 - isotropic inhomogeneous continuum, 358
 - lateral homogeneity, 331, 366
 - parametric form, 374
 - polar coordinates, 377
 - ray parameter, 365
 - total derivative, 331
 - two-dimensional continuum, 366, 373, 374
 - variational principle, 343
- inverse problem, 284
- ray parameter, 371
- ray theory, 141
 - variational principle, 321
- traveltime increment
 - wavefront
 - elliptical velocity dependence, 383
- trial solution
 - anisotropic inhomogeneous continuum, 209–212
 - equations of motion, 220
 - Christoffel's equations, 270
 - displacement vector, 270
 - eikonal function, 184, 211
 - elliptical velocity dependence, 181
 - harmonic wave, 176
 - high frequency, 253
 - plane wave, 184
 - reduced wave equation, 177, 206
 - wave equation, 176, 184
 - weak inhomogeneity, 184, 185
- trigonal continuum, xxv, 106, 109, 122
- trigonometric polynomial, 113–115, 129, 130
- Truesdell, Clifford (1919–2000), 4
- variational principles, *see also* action, Fermat's principle *and* Hamilton's principle
 - least traveltime, 322
 - stationary action, 335, 350
 - stationary traveltime, xxii, 304, 333, 335, 343, 344, 349, 356, 365, 373
- vector potential, *see also* scalar potential

- displacement, 150–153
- vis viva*, 350, *see also* forces vives
- volume-preserving vector field, 147
- wave
 - P , 146
 - displacement potential, 152
 - S , 147
 - displacement potential, 152
 - P , 24, 146, 147, 154, 183, 186, 187
 - displacement direction, 272, 273, 285, 286
 - displacement potential, 149, 152
 - eikonal equation, 186
 - plane, 149
 - speed, 147
 - transversely isotropic continuum, 276, 290
 - velocity, 273, 276, 277, 286
- SH
 - amplitude, 296, 301, 304
 - elliptical velocity dependence, 288, 296
 - interface, 296
 - transversely isotropic continuum, 182, 277, 279, 280, 282, 285, 286, 288, 289, 291, 305
 - velocity, 282, 283, 286, 288
 - vertical inhomogeneity, 183
- SV , 183
 - transversely isotropic continuum, 277, 282, 285, 286
 - velocity, 277, 286
- S , 26, 147, 154, 187
 - displacement direction, 271–274, 285, 287
 - displacement potential, 150, 153
 - eikonal equation, 186
 - plane, 149
 - rigidity, 147
 - speed, 147
 - transversely isotropic continuum, 276, 277, 282
 - velocity, 273, 287
- S_1 , 273, 282
- S_2 , 273, 282
- qP
 - displacement direction, 290
 - Fermat's principle, 356
 - transversely isotropic continuum, 279, 280, 305
- qSV
 - transversely isotropic continuum, 279, 280, 282, 305
 - velocity, 282, 283
- qS
 - Fermat's principle, 356
- body, xxii, 187, 220
- dilatational, 147, 186, 207
- elastic, xvii
- equivoluminal, 147
- evanescent, 302
- harmonic, 176, 177, 212
- incident, 296, 297
 - amplitude, 301
 - boundary condition, 299, 300
 - displacement direction, 301
 - energy, 296, 308
 - phase, 302
- interface, 187
- longitudinal, 154, 356
- monochromatic
 - interface, 300, 308
 - trial solution, 181
- plane, 143, 148, 149, 176, 184, 204
- pressure, 147
- primary, 147
- pure-mode, 272, 276, 280, 290
- reflected, 297
 - amplitude, 298
 - boundary condition, 299, 300
 - displacement direction, 301
 - energy, 308
 - evanescent wave, 302
 - phase, 302
- rotational, 147, 186, 207
- secondary, 147
- seismic, xvii, xxi, xxii, 3, 7, 8
 - body, 141
 - continuum mechanics, 33
 - infinitesimal displacement, 15
 - isotropic continuum, 143
- shear, 147, 283
- standing, 177
- surface, 187
- transmitted, 296, 297
 - amplitude, 298, 301

- boundary condition, 299, 300
- displacement, 302
- energy, 308
- transverse, 154, 356, 362, 363
- wave equation
 - P wave, 24, 146, 186
 - S wave, 26, 147
 - approximation
 - generally anisotropic continuum, 182
 - balance of linear momentum, 188
 - Cauchy's equations of motion, 56, 73, 75
 - isotropic homogeneous continuum, 141, 187, 209
 - characteristic equation, 204, 205
 - displacement potential, 149
 - P wave, 152
 - S wave, 153
 - elliptical velocity dependence, 179
 - extension, 179
 - elliptical velocity dependence, 180–182
 - generally anisotropic continuum, 182
 - weak inhomogeneity, 183, 187
 - Fourier's transform, 177–179
 - Hamilton's principle, 188, 335, 350, 351, 353, 363
 - P wave, 356
 - S wave, 356, 362
 - reduced, 176, 177, 179, 184, 203
 - solution
 - d'Alembert, 154, 157, 189
 - distribution, 173–176, *see also* weak solution
 - domain of dependence, 161
 - plane wave, 204
 - propagation speed, 146, 147, 157, 160, 161
 - range of influence, 161, 162
 - scattering, 168
 - spatial dimension, 163
 - theory of distributions, 188
 - stress-strain equations, 119, 143
- wave function
 - P wave, 146
 - S wave, 147
 - dilatation, 146, 152
 - generic, 154
 - rotation vector, 147, 152
 - scalar potential, 152
 - vector potential, 152
- wave mechanics, 350
- wave theory, xxi
 - ray theory, 254
- wavefield, 3, 229
 - plane wave, 149
- waveform
 - trial solution, 211, 212
- wavefront
 - anisotropic continuum
 - infinitesimal, 243
 - anisotropic inhomogeneous continuum, 204
 - amplitude, 184, 187
 - constant phase, 219
 - displacement vector, 270, 271
 - eigenvalue, 218
 - eikonal equation, 220, 229
 - eikonal function, 211, 218
 - orientation, 220, 271, 272, 274, 278, 392
 - phase slowness, 218, 219
 - phase velocity, 239, 270
 - phase-slowness vector, 218, 219, 237–240, 242, 258, 268, 392
 - shape, 184
 - elliptical velocity dependence, 180–182
 - infinitesimal, 243, 383, 384
 - orientation, 181, 182, 242
 - phase slowness, 181
 - phase velocity, 181
 - interface, 292, 305
 - isotropic homogeneous continuum
 - plane, 184
 - isotropic inhomogeneous continuum, 255
 - eikonal function, 184
 - planar, 184
 - shape, 184
 - ray theory, 141, 321
 - wave equation
 - three spatial dimensions, 165
 - two spatial dimensions, 166
- wavelength
 - displacement, 17, 271
 - ray theory, 141
 - short, 185

- weak inhomogeneity, 183
- weak derivatives, 174
- weak inhomogeneity, 179, 183
 - eikonal equation, **186**
 - frequency, **185**
 - wavelength, **183**
- weak solution, 175
 - wave equation, 173, *see also* distribution
- welded contact, 297, **297**, 300, 305
- well-posed problem, 158, 159
- WKB method, 254

- $\mathbf{x}\mathbf{p}$ -space, 229, 230, 233, 237, 311
- $\mathbf{x}\dot{\mathbf{x}}$ -space, 311

- Young's modulus, 132, 134, 136, 404

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.

Abbott, M.B., (1966) An introduction to the method of characteristics: Elsevier

Achenbach, J.D., (1973) Wave propagation in elastic solids: North Holland

Achenbach, J.D., Gautesen, 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 géométrique (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 (*2nd* 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

Appel, W., (2005) Mathématiques pour la physique et les physiciens (*3rd* edition): H & K Editions

Appell, P., (1905) Cours de mécanique à l'usage des élèves de la classe de mathématique spéciale: Gauthier-Villars, Paris

Arianrhod, R., (2005) Einstein's heroes: Imagining the world through the language of mathematics: Oxford University Press

Armstrong, M.A., (1988) Groups and symmetry: Springer-Verlag

- 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
- Arnold, V.I., (2004) *Lectures on partial differential equations*: 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
- Barton, G., (1989) *Elements of Green's functions and propagation: Potentials, diffusion and waves*: Oxford Science Publications
- Båth, M., and Berkhout, A.J., (1984) *Mathematical aspects of seismology*: Geophysical Press
- Basdevant, J.-L., (2007) *Variational principles in physics*: Springer-Verlag
- Basdevant, J.-L., (2005) *Principes variationnelles et dynamique*: Vuibert
- Batterman, R.W., (2002) *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*: Oxford University Press
- Bellman, R., (1966/2003) *Perturbation techniques in mathematics, engineering and physics*: Dover
- Bender, C.M., and Orszag, S.A., (1978/1999) *Advanced mathematical methods for scientists and engineers: Asymptotic methods and perturbation theory*: Springer-Verlag
- 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 (27th 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
- Boccara, N., (1997) Distributions: Ellipses
- Bóna, A., Bucataru, I., Slawinski, M.A., (2007) Coordinate-free classification of elasticity tensor: Journal of Elasticity **87**(2-3), 109–132.
- Bóna, A., Bucataru, I., and Slawinski, M.A., (2005) Characterization of elasticity-tensor symmetries using SU(2). Journal of Elasticity **75**(3), 267 – 289.
- Bucataru, I., and Slawinski, M.A., (2005) Generalized orthogonality of rays and wavefronts in anisotropic inhomogeneous media. Nonlinear analysis: Series B Real-world applications **6**, 111 – 121.
- Bóna, A., Bucataru, I., and Slawinski, M.A., (2004) Material symmetries of elasticity tensor. The Quarterly Journal of Mechanics and Applied Mathematics **57**(4), 583 – 598.
- Bóna, A., and Slawinski, M.A., (2003) Fermat's principle for seismic rays in elastic media. Journal of Applied Geophysics **54**, 445 – 451.
- 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 (7th edition): Cambridge University Press
- Bos, L.P., Gibson, P.C., Kotchetov, M., and Slawinski, M.A., (2004) Classes of anisotropic media: a tutorial: Studia geophysica and geodætica, **48**, pp. 265 – 287.
- Brekhovskikh, L.M., and Goncharov, V., (1982/1994) Mechanics of continua and wave dynamics: Springer-Verlag
- Bridgman, P.W., (1927) The logic of modern physics: The MacMillan Co.
- Bronson, R., (1973) Modern introductory differential equations: Schaum's Outlines, McGraw-Hill, Inc.
- Brown, J.R., (2001) Who rules in science: An opinionated guide to the wars: Harvard University Press
- Brown, J.R., (1991) The laboratory of the mind: Thought experiments in the natural sciences: Routledge
- 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

- Bunge, M., (2006) Chasing reality: Strife over realism: University of Toronto Press
- Bunge, M., (2003) Emergence and convergence: Qualitative novelty and the unity of knowledge: University of Toronto Press
- Bunge, M. A., (1998) Philosophy of Science, Vol. I: From problem to theory (Revised edition): Transaction Publishers
- Bunge, M. A., (1998) Philosophy of Science, Vol. II: From explanation to justification (Revised edition): Transaction Publishers
- Bunge, M., (1973) Philosophy of physics: D. Reidel Publishing Co.
- Bunge, M., (1973) Method, model and matter: D. Reidel Publishing Co.
- Bunge, M., (1967) Foundations of physics: Springer-Verlag
- Bunge, M., (editor), (1964) Critical approaches to science and philosophy: Transaction Publishers
- 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
- Chapman, C., (2004) Fundamentals of seismic wave propagation: Cambridge University Press
- Cheney, W., (2001) Analysis for applied mathematics: Springer-Verlag
- 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 (2nd edition): Springer-Verlag
- Coirier, J., (1997) Mécanique des milieux continues: Concepts de base: Cours et exercices corrigés: Dunod
- Coirier, J., (2001) Mécanique des milieux continues: Aide-mémoire: Dunod
- Coleman, B.D., and Noll, W., (1961) Foundations of linear viscoelasticity: *Revue of modern physics*, **33**, 239 – 249.
- 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
- Dahlen, F.A., and Tromp, J., (1998) Theoretical global seismology: Princeton University Press
- de Bruijn, N.G., (1958/1981) Asymptotic methods in analysis: Dover
- de Hoop, A. T., (1995) Textbook of radiation and scattering of waves: Academic Press, Inc.
- 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., (1971/1994) Differential forms and applications: Springer-Verlag
- do Carmo, M.P., (1976) Differential geometry of curves and surfaces: Prentice-Hall, Inc.
- Doran, C., and Lasenby, A., (2003) Geometric algebra for physicists, Cambridge University Press
- Dugas, R., (1950/1996) Histoire de la mécanique: Éditions Jacques Gabay
- Duistermaat, J.J., (1996) Fourier Integral operators: Birkhäuser
- Einstein, A., and Infeld, L., (1938) Evolution of physics from early concepts to relativity and quanta: Simon & Schuster
- Eisenhart, L. P., (1925/1997) Riemannian Geometry: Princeton University Press
- Ekeland, I., (2000) Le meilleur des mondes possibles: Mathématiques et destinée: Seuil
- Elmore, W.C., and Heald, M.A., (1969/1985) Physics of waves: Dover
- Elsogolc, 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

- Eringen, A.C., (1999) Microcontinuum field theories I: Foundations and solids: Springer-Verlag
- Evans, L.C., (1998) Partial differential equations: AMS
- 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., (1985/2006) QED: The strange theory of light and matter: Princeton University Press
- Feynman, R.P., (1967) The character of physical law: The MIT Press
- Feynman, R.P., Leighton, R.B., and Sands, M., (1963/1989) Feynman's lectures on physics: Addison-Wesley Publishing Co.
- Flanders, H., (1963/1989) Differential forms with applications to the physical sciences: Dover
- Folland, G.B., (1995) Introduction to partial differential equations (2nd edition): Princeton University Press
- Forest, S., (2006) Milieux continus généralisés et matériaux hétérogènes: Mines Paris
- 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.
- Garrity, T.A., (2001) All the mathematics you missed [but need to know for graduate school]: Cambridge University Press
- 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
- Hadsell, F., (1995) Tensors of geophysics for mavericks and mongrels: SEG
- Hadsell, F., and Hansen, R., (1999) Tensors of geophysics: Generalized functions and curvilinear coordinates: SEG
- Hankins, T.L., (1980) Sir William Rowan Hamilton: The Johns Hopkins University Press
- 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.
- Hesse, M.B., (1962/2005) Forces and fields: The concept of action at a distance in the history of physics: Dover
- 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
- Isakov, V., (1998) Inverse problems for partial differential equations: Springer-Verlag
- Jaynes, E.T. (2003) Probability theory: The logic of science: Cambridge University Press
- Jeans, J., (1943/1981) Physics and philosophy: Dover
- Jeffreys, H., and Jeffreys, B., (1946/1999) Methods of Mathematical Physics: Cambridge University Press
- John, F., (1982) Partial differential equations (4th edition): Springer-Verlag
- John, F., (1955/2004) Plane waves and spherical means applied to partial differential equations: Dover

- 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
- Kennett, B.L.N., (2002) The seismic wavefield, Vol. II: Interpretation of seismograms on regional and global scales: Cambridge University Press
- Kilmister, C.W., and Reeve, J.E., (1966) Rational mechanics: Longmans
- Kittel, C. (2005) Introduction to solid state physics (8th edition): John Wiley and Sons
- Klein, E., (1991) Conversations avec le Sphinx: Les paradoxes en physique: Editions Albin Michel
- 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., (2005) Geometrical optics in engineering physics: Alpha Science International Ltd.
- 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., (2004) 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
- Kuhn, T.S., (1996) The structure of scientific revolutions (3rd edition): The University of Chicago Press
- 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
- Landau, L.D., and Lifshitz, E.M., (1976) Course of theoretical physics, Vol. 1: Mechanics (3rd edition): Elsevier
- Landau, L.D., and Lifshitz, E.M., (1975) Course of theoretical physics, Vol. 2: The classical theory of fields (4th edition): Elsevier
- Landau, L.D., and Lifshitz, E.M., (1986) Course of theoretical physics, Vol. 7: Theory of elasticity (3rd edition): Elsevier
- 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
- Mach, E., (1893/1989) The science of mechanics: A critical and historical account of its development: Open Court Publishing Co.
- 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 (2nd 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
- Nahin, P.J., (2006) Dr. Euler's fabulous formula: Cures many mathematical ills: Princeton University Press
- Neusel, M.D., and Smith, L., (2002) Invariant theory of finite groups: American Mathematical Society
- Newnham, R.E., (2005) Properties of materials: Anisotropy, symmetry, structure: Oxford University Press
- 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
- Okasha, S., (2002) Philosophy of science: A very short introduction: Oxford University Press
- Olmsted, J.M.H., (1961) Advanced calculus: Prentice-Hall, Inc.
- Olver, P.J., (1993) Applications of Lie groups to differential equations (*2nd* edition): Springer-Verlag
- Pesic, P., (2003) Abel's proof: An essay on the sources and meaning of mathematical unsolvability: The MIT Press
- 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

- Polya, G., (1954) Mathematics and plausible reasoning, Vol. I: Induction and analogy in mathematics: Princeton University Press
- Polya, G., (1968) Mathematics and plausible reasoning, Vol. II: Patterns of plausible inference: Princeton University Press
- 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
- Pujol, J., (2003) Elastic wave propagation and generation in seismology: Cambridge University Press
- 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
- Renardy, M., and Rogers, R.C., (1993) An introduction to partial differential equations: Springer-Verlag
- 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) Lecture notes on continuum mechanics: Memorial University of Newfoundland
- Rochester, M.G., (1997) Lecture notes on mathematical physics: Memorial University of Newfoundland
- Rogister, Y., (2000) Etude théorique des déformations globales de la terre (Ph.D. thesis): Université de Liège
- Rogister, Y., and Slawinski, M.A., (2005) Analytic solution of ray-tracing equations for a linearly inhomogeneous and elliptically anisotropic velocity model: Geophysics, **70**, D37 – D41.
- Ross, S.L., (1989) Introduction to ordinary differential equations (4th 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
- Rychlewski, J., (1985) On Hooke's law: Prikl. Matem. Mekhan., **48** (3), 420 – 435.
- Rymarz, C., (1993) Mechanika ośrodkow ciągłych: PWN — Polish Scientific Publishers
- Sagan, H., (1969/1992) Introduction to the calculus of variations: Dover

- Scheerer, P.M., (1999) Introduction to seismology: Cambridge University Press
- Schey, H.M., (2005) div, grad, curl, and all that: An informal text on vector calculus (4th edition): W.W. Norton & Co.
- 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., (2003) Gravity from the ground up: 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.
- Seaborn, J.B., (2002) Mathematics for the physical sciences: Springer-Verlag
- Sedov, L.I., (1971) A course in continuum mechanics: Wolters-Noordhoff Publishing
- 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., Wheaton, C.J., and Powojowski, M. (2004) VSP travelttime inversion for linear inhomogeneity and elliptical anisotropy: Geophysics, **69**, 373 – 377.
- 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

- Snieder, R., (2004) A guided tour of mathematical methods for the physical sciences: Cambridge University Press
- 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.
- Steiner, M., (1998) The applicability of mathematics as a philosophical problem: Harvard University Press
- 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 Griffith, B.A., (1949) Principles of mechanics: McGraw-Hill, Inc.
- Synge, J.L., and Schild, A., (1949/1978) Tensor calculus: Dover
- Szekeres, P., (2004) A course in modern mathematical physics: Groups, Hilbert space and differential geometry: Cambridge University Press
- 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
- Truesdell, C.A., (1966) Six lectures on modern natural philosophy: Springer-Verlag
- 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
- van Brunt, B. (2006) *The calculus of variations*: Springer-Verlag
- Wang, C.C., and Truesdell, C., (1973) *Introduction to rational elasticity*: Noordhoff International Publishing
- Weinert, F., (2005) *The scientist as philosopher*: Springer-Verlag
- 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., (1939/1946) *The classical groups: Their invariants and representations*: Princeton University Press
- 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 (2nd 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 (4th edition)*: Brooks/Cole Publishing Co.
- Zwillinger D., (1992) *Book of differential equations (2nd edition)*: Academic Press



Michael A. Slawinski is a professor in the Department of Earth Sciences at Memorial University in St. John's, Newfoundland, and an adjunct professor in the Department of Mathematics and Statistics at the University of Calgary.

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.

Since 1998, Slawinski has been the director of The Geomechanics Project, a theoretical-research group whose research focuses on seismic ray theory in the context of differential geometry and continuum mechanics.