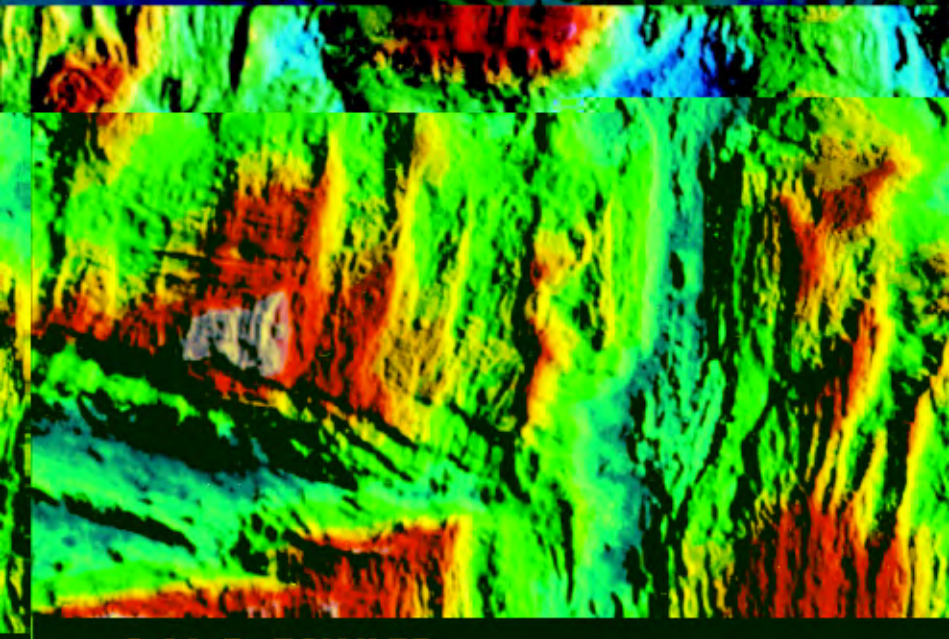
Second Edition

The Solid Earth

An Introduction to Global Geophysics



C. M. R. FOWLER

CAMBRIDGE

The Solid Earth

From the citation for the Prestwich Medal of the Geological Society, 1996 (awarded for the contribution made by *The Solid Earth* to geophysics teaching and research) by the then President Professor R. S. J. Sparks F.R.S.

The Prestwich Medal is given for major contributions to earth science, and provides an opportunity for the Society to recognise achievements in areas that can lie outside the terms of reference of its other awards. This year, the Prestwich Medal has been given to Mary Fowler for the contribution of her book, *The Solid Earth*, which has had an enormous impact. The book has been acclaimed by today's leading geophysicists. There is consensus that, although there are many books covering various aspects of geophysics, there are only a small number that can be seen as landmarks in the subject. Mary's book has been compared to Jeffreys's *The Earth* and Holmes's *Physical Geology*.

The Solid Earth is recognised by her peers as a monumental contribution. In this book she displays a wide knowledge of a very broad range of geological and geophysical topics at a very high level. The book provides a balanced and thoroughly researched account which is accessible to undergraduates as well as to active researchers.

The book has been described as one of the outstanding texts in modern earth sciences.

(*Geoscientist*, Geological Society, 1996, Vol. **6**, No. 5, p. 24.)

The Solid Earth

An Introduction to Global Geophysics

Second Edition

C. M. R. Fowler
Royal Holloway
University of London

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge, CB2 2RU, UK
40 West 20th Street, New York, NY 10011–4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

© Cambridge University Press 2005

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2005

Printed in the United Kingdom at the University Press, Cambridge

Typefaces Times NR 10/13 pt. and Universe *System* LaTeX 2ε [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data
Fowler, C. M. R.
The solid earth: an introduction to global geophysics / C. M. R. Fowler. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0 521 58409 4 (hardback) – ISBN 0 521 89307 0 (paperback)
1. Geophysics. 2. Earth. I. Title.
QC806.F625 2004
550 – dc22 2003065424

ISBN 0 521 58409 4 hardback
ISBN 0 521 89307 0 paperback

TO MY FAMILY

Magna opera domini exoquisita in omnes voluntates ejus.

The works of the Lord are great, sought out of all them that have pleasure therein.

Psalm 111.2: at the entrance to the old Cavendish Laboratories, Cambridge.

Contents

Preface to the first edition *page* xi
Preface to the second edition xv
Acknowledgements to the first edition xvi
Acknowledgements to the second edition xviii

1 Introduction **1**
References and bibliography 3

2 Tectonics on a sphere: the geometry of plate tectonics **5**
2.1 Plate tectonics 5
2.2 A flat Earth 11
2.3 Rotation vectors and rotation poles 14
2.4 Present-day plate motions 15
2.5 Plate boundaries can change with time 24
2.6 Triple junctions 26
2.7 Absolute plate motions 32
Problems 37
References and bibliography 40

3 Past plate motions **43**
3.1 The role of the Earth's magnetic field 43
3.2 Dating the oceanic plates 54
3.3 Reconstruction of past plate motions 67
Problems 93
References and bibliography 94

4 Seismology Measuring the interior **100**
4.1 Waves through the Earth 100
4.2 Earthquake seismology 111
4.3 Refraction seismology 140
4.4 Reflection seismology 157
Problems 178
References and bibliography 186

5 Gravity **193**
5.1 Introduction 193
5.2 Gravitational potential and acceleration 193
5.3 Gravity of the Earth 196
5.4 The shape of the Earth 198
5.5 Gravity anomalies 202
5.6 Observed gravity and geoid anomalies 213
5.7 Flexure of the lithosphere and the viscosity of the mantle 218
Problems 228
References and bibliography 230

6 Geochronology **233**
6.1 Introduction 233
6.2 General theory 234
6.3 Rubidium–strontium 244
6.4 Uranium–lead 247
6.5 Thorium–lead 249
6.6 Potassium–argon 251
6.7 Argon–argon 252
6.8 Samarium–neodymium 254
6.9 Fission-track dating 258
6.10 The age of the Earth 262
Problems 265
References and bibliography 267

7 Heat **269**
7.1 Introduction 269
7.2 Conductive heat flow 270
7.3 Calculation of simple geotherms 275
7.4 Worldwide heat flow: total heat loss from the Earth 285
7.5 Oceanic heat flow 288
7.6 Continental heat flow 298
7.7 The adiabat and melting in the mantle 303
7.8 Metamorphism: geotherms in the continental crust 308
Problems 321
References and bibliography 323

8 The deep interior of the Earth **326**
8.1 The internal structure of the Earth 326
8.2 Convection in the mantle 353
8.3 The core 371
References and bibliography 381

9 The oceanic lithosphere: ridges, transforms, trenches and oceanic islands **391**
9.1 Introduction 391
9.2 The oceanic lithosphere 397
9.3 The deep structure of mid-ocean ridges 409
9.4 The shallow structure of mid-ocean ridges 417
9.5 Transform faults 440
9.6 Subduction zones 458
9.7 Oceanic islands 487
Problems 492
References and bibliography 493

10 The continental lithosphere **509**
10.1 Introduction 509
10.2 The growth of continents 517
10.3 Sedimentary basins and continental margins 557
10.4 Continental rift zones 584
10.5 The Archaean 595
Problems 601
References and bibliography 602

Appendix 1 Scalars, vectors and differential operators 615
Appendix 2 Theory of elasticity and elastic waves 620
Appendix 3 Geometry of ray paths and inversion of earthquake body-wave time–distance curves 630
Appendix 4 The least-squares method 636
Appendix 5 The error function 638
Appendix 6 Units and symbols 640
Appendix 7 Numerical data 648
Appendix 8 The IASP91 Earth model 650
Appendix 9 The Preliminary Reference Earth Model, isotropic version – PREM 651
Appendix 10 The Modified Mercalli Intensity Scale (abridged version) 654
Glossary 655
Index 666

The colour plates are situated between pages 398 and 399.

Preface to the first edition

Geophysics is a diverse science. At its best it has the rigour of physics and the vigour of geology. Its subject is the Earth. How does the Earth work? What is its composition? How has it changed? Thirty years ago many of the answers to these questions were uncertain. We knew the gross structure of our planet and that earthquakes occurred, volcanoes erupted and high mountains existed, but we did not understand why. Today we have a general knowledge of the workings of the planet, although there is still much to be discovered.

My aim in writing this book was to convey in a fairly elementary way what we know of the structure and dynamics of the solid Earth. The fabric of geophysics has changed dramatically in the decades since the discovery of plate tectonics. The book places a strong emphasis on geophysical research since the initial formulation of plate theory, and the discussion centres on the crust and upper mantle. It also outlines the recent increases in our knowledge of the planet's deeper interior.

To whom is this book addressed? It is designed to serve as an introduction to geophysics for senior undergraduates in geology or physics and for graduate students in either subject who need to learn the elements of geophysics. My hope is that the book will give them a fairly comprehensive basis on which to build an understanding of the solid Earth.

Part of the challenge in writing a geophysics text is to make the book accessible to both types of student. For instance, some students enter the study of geophysics from a background in the Earth sciences, others from physics or mathematics: only a few enroll directly in geophysics programmes. Geology students tend to know about rocks and volcanoes, but possess only the basics of calculus. In contrast, students of physics have good mathematical skills, but do not know the difference between a basalt and a granite. I have attempted throughout the book to explain for the geologists the mathematical methods and derivations and to include worked examples as well as questions. I hope that this will make the book useful to students who have only introductory calculus. For the non-geologists, I have tried to limit or explain the abundant geological terminology. There is a glossary of terms, to rescue physics students lost in the undergrowth of nomenclature.

For more advanced students of either geological or physical training I have in places included more mathematical detail than is necessary for a basic

introductory course. This detail can easily be by-passed without either interrupting the continuity of the text or weakening the understanding of less mathematical students. Throughout the book I have attempted to give every step of logic so that students can understand why every equation and each conclusion is valid.

In general, I have tried to avoid the conventional order of textbooks in which geophysical theory comes first, developed historically, followed in later sections by interesting and concrete examples. For instance, because the book focuses to a large extent on plate-tectonic theory, which is basic to the study of the crust and mantle, this theory is introduced in its proper geophysical sense, with a discussion of rotation, motions on plate boundaries and absolute plate motions. Most geological texts avoid discussing this, relying instead on two-dimensional cartoons of ridges and subduction zones. I have met many graduate students who have no idea what a rotation pole is. Their instructors thought the knowledge irrelevant. Yet understanding tectonics on a sphere is crucial to geophysics because one cannot fully comprehend plate motion without it.

The next chapters of the book are concerned with past plate motions, magnetics, seismology and gravity. These are the tools with which plate tectonics was discovered. The exposition is not historical, although historical details are given. The present generation of geophysicists learned by error and discovery, but the next generation will begin with a complete structure on which to build their own inventions.

These chapters are followed by discussions of radioactivity and heat. The Earth is a heat engine, and the discovery of radioactivity radically changed our appreciation of the physical aspects of the planet's history, thermal evolution and dynamics. The study of isotopes in the Earth is now, perhaps unfairly, regarded as an area of geochemistry rather than of geophysics; nevertheless, the basic tools of dating, at least, should be part of any geophysics course. Understanding heat, on the other hand, is central to geophysics and fundamental to our appreciation of the living planet. All geology and geophysics, indeed the existence of life itself, depend on the Earth's thermal behaviour. Heat is accordingly discussed in some detail.

The final chapters use the knowledge built up in the earlier ones to create an integrated picture of the complex operation of the oceanic and continental lithosphere, its growth and deformation. The workshops of geology – ridges and subduction zones – are described from both geophysical and petrological viewpoints. Sedimentary basins and continental margins employ most of the world's geophysicists. It is important that those who explore the wealth or perils of these regions know the broader background of their habitat.

SI units have been used except in cases where other units are clearly more appropriate. Relative plate motions are quoted in centimetres or millimetres per year, not in metres per second. Geological time and ages are quoted in millions or billions of years (Ma or Ga) instead of seconds. Temperatures are quoted in

degrees Centigrade (°C), not Kelvin (K). Seismic velocities are in kilometres per second, not in metres per second.

Most geophysicists look for oil. Some worry about earthquakes or landslips, or advise governments. Some are research workers or teach at universities. Uniting this diversity is a deep interest in the Earth. Geophysics is a rigorous scientific discipline, but it is also interesting and fun. The student reader to whom this book is addressed will need rigour and discipline and often hard work, but the reward is an understanding of our planet. It is worth it.

Chapter 1
Introduction

Geophysics, the physics of the Earth, is a huge subject that includes the physics of space and the atmosphere, of the oceans and of the interior of the planet. The heart of geophysics, though, is the theory of the solid Earth. We now understand in broad terms how the Earth's surface operates, and we have some notion of the workings of the deep interior. These processes and the means by which they have been understood form the theme of this book. To the layperson, geophysics means many practical things. For Californians, it is earthquakes and volcanoes; for Texans and Albertans, it is oil exploration; for Africans, it is groundwater hydrology. The methods and practices of applied geophysics are not dealt with at length here because they are covered in many specialized textbooks. This book is about the Earth, its *structure* and *function* from surface to centre.

Our search for an understanding of the planet goes back millennia to the ancient Hebrew writer of the Book of Job and to the Egyptians, Babylonians and Chinese. The Greeks first measured the Earth, Galileo and Newton put it in its place, but the Victorians began the modern discipline of geophysics. They and their successors were concerned chiefly with understanding the structure of the Earth, and they were remarkably successful. The results are summarized in the magnificent book *The Earth* by Sir Harold Jeffreys, which was first published in 1924. Since the Second World War the function of the Earth's surface has been the focus of attention, especially since 1967 when geophysics was revolutionized by the discovery of *plate tectonics*, the theory that explains the function of the uppermost layers of the planet.

The rocks exposed at the surface of the Earth are part of the *crust* (Fig. 1.1). This crustal layer, which is rich in silica, was identified by John Milne (1906), Lord Rayleigh and Lord Rutherford (1907). It is on average 38 km thick beneath continents and 7–8 km thick beneath oceans. Beneath this thin crust lies the *mantle*, which extends down some 2900 km to the Earth's central *core*. The mantle (originally termed *Mantel* or 'coat' in German by Emil Wiechert in 1897, perhaps by analogy with Psalm 104) is both physically and chemically distinct from the crust, being rich in magnesium silicates. The crust has been derived from the mantle over the aeons by a series of melting and reworking processes. The boundary between the crust and mantle, which was delineated by Andrya

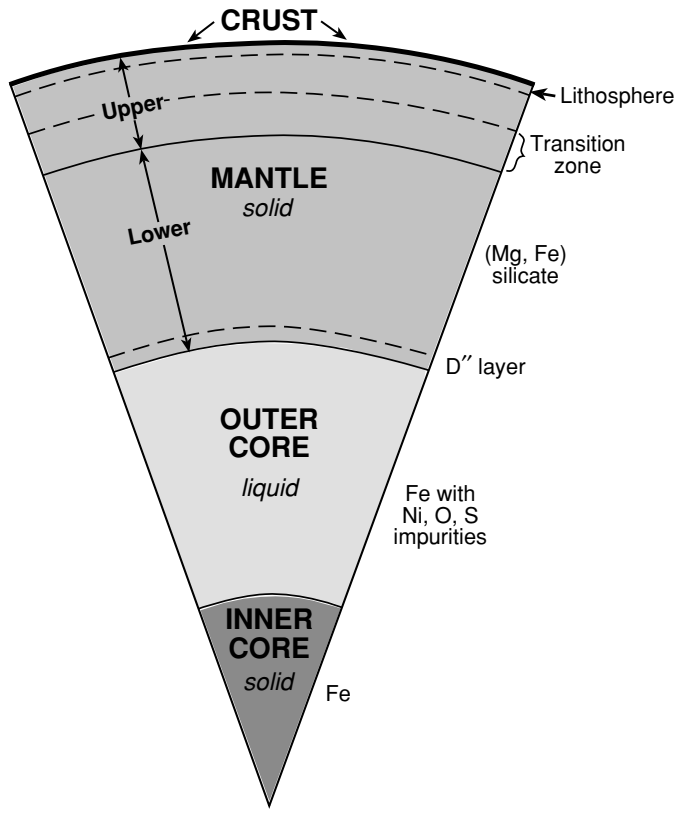


Figure 1.1. The major internal divisions of the Earth.

Mohorovičić in 1909, is termed the Mohorovičić discontinuity, or *Moho* for short. The core of the Earth was discovered by R. D. Oldham in 1906 and correctly delineated by Beno Gutenberg in 1912 from studies of earthquake data (Gutenberg 1913, 1914). The core is totally different, both physically and chemically, from the crust and mantle. It is predominantly iron with lesser amounts of other elements. The core was established as being fluid in 1926 as the result of work on tides by Sir Harold Jeffreys. In 1929 a large earthquake occurred near Buller in the South Island of New Zealand. This, being conveniently on the other side of the Earth from Europe, enabled Inge Lehmann, a Danish seismologist, to study the energy that had passed through the core. In 1936, on the basis of data from this earthquake, she was able to show that the Earth has an *inner core* within the liquid outer core. The inner core is solid.

The presence of ancient beaches and fossils of sea creatures in mountains thousands of feet above sea level was a puzzle and a stimulation to geologists from Pliny's time to the days of Leonardo and Hutton. On 20 February 1835, the young Charles Darwin was on shore resting in a wood near Valdivia, Chile, when suddenly the ground shook. In his journal *The Voyage of the Beagle* Darwin (1845) wrote that 'The earth, the very emblem of solidity, has moved beneath our feet

like a thin crust over a fluid.' This was the great Concepción earthquake. Several days later, near Concepción, Darwin reported that 'Captain Fitz Roy found beds of putrid mussel shells still adhering to the rocks, ten feet above high water level: the inhabitants had formerly dived at low-water spring-tides for these shells.' The volcanoes erupted. The solid Earth was active.

By the early twentieth century scientific opinion was that the Earth had cooled from its presumed original molten state and the contraction which resulted from this cooling caused surface topography: the mountain ranges and the ocean basins. The well-established fact that many fossils, animals and plants found on separated continents must have had a common source was explained by either the sinking of huge continental areas to form the oceans (which is, and was then recognized to be, impossible) or the sinking beneath the oceans of land bridges that would have enabled the animals and plants to move from continent to continent.

In 1915 the German meteorologist Alfred Wegener published a proposal that the continents had slowly moved about. This theory of *continental drift*, which accounted for the complementarity of the shapes of coastlines on opposite sides of oceans and for the palaeontological, zoological and botanical evidence, was accepted by some geologists, particularly those from the southern hemisphere such as Alex Du Toit (1937), but was generally not well received. Geophysicists quite correctly pointed out that it was physically impossible to move the continents through the solid rock which comprised the ocean floor. By the 1950s, however, work on the magnetism of continental rocks indicated that in the past the continents must have moved relative to each other; the *mid-ocean ridges*, the Earth's longest system of mountains, had been discovered, and continental drift was again under discussion. In 1962 the American geologist Harry H. Hess published an important paper on the workings of the Earth. He proposed that continental drift had occurred by the process of *seafloor spreading*. The mid-ocean ridges marked the limbs of rising convection cells in the mantle. Thus, as the continents moved apart, new seafloor material rose from the mantle along the mid-ocean ridges to fill the vacant space. In the following decade the theory of plate tectonics, which was able to account successfully for the physical, geological and biological observations, was developed. This theory has become the unifying factor in the study of geology and geophysics. The main difference between plate tectonics and the early proposals of continental drift is that the continents are no longer thought of as ploughing through the oceanic rocks; instead, the oceanic rocks and the continents are together moving over the interior of the Earth.

References and bibliography

Brush, S. J. 1980. Discovery of the earth's core. *Am. J. Phys*., **48**, 705–24.

Darwin, C. R. 1845. *Journal of Researches into the Natural History and Geology of the Countries Visited during the Voyage of H.M.S. Beagle round the World, under the Command of Capt. Fitz Roy R.N*., 2nd edn. London: John Murray.

Du Toit, A. 1937. *Our Wandering Continents*. Edinburgh: Oliver and Boyd.
Gutenberg, B. 1913. Über die Konstitution der Erdinnern, erschlossen aus Erdbebenbeobachtungen. *Phys. Zeit.*, **14**, 1217.
1914. Über Erdbebenwellen, VIIA. Beobachtungen an Registrierungen von Fernbeben in Göttingen und Folgerungen über die Konstitution des Erdkörpers. *Nachr. Ges. Wiss. Göttingen. Math. Phys.*, *Kl.* **1**, 1–52.
Hess, H. H. 1962. History of ocean basins. In A. E. J. Engel, H. L. James and B. F. Leonard, eds., *Petrologic Studies: A Volume in Honor of A. F. Buddington*. Boulder, Colorado: Geological Society of America, pp. 599–620.
Jeffreys, H. 1926. The rigidity of the Earth's central core. *Mon. Not. Roy. Astron. Soc. Geophys. Suppl.*, **1**, 371–83. (Reprinted in Jeffreys, H. 1971. *Collected Papers, Vol. 1*. New York: Gordon and Breach.)
1976. *The Earth*, 6th edn. Cambridge: Cambridge University Press.
Lehmann, I. 1936. P′. *Trav. Sci., Sect. Seis. U.G.G.I. (Toulouse)*, **14**, 3–31.
Milne, J. 1906. Bakerian Lecture – recent advances in seismology. *Proc. Roy. Soc.* A, **77**, 365–76.
Mohorovičić, A. 1909. Das Beben vom 8. X. 1909. *Jahrbuch met. Obs. Zagreb*, **9**, 1–63.
Oldham, R. D. 1906. The constitution of the earth as revealed by earthquakes. *Quart. J. Geol. Soc.*, **62**, 456–75.
Rutherford, E. 1907. Some cosmical aspects of radioactivity. *J. Roy. Astr. Soc. Canada*, May–June, 145–65.
Wegener, A. 1915. *Die Entstehung der Kontinente und Ozeane*.
1924. *The Origin of Continents and Oceans*. New York: Dutton.
Wiechert, E. 1897. Über die Massenvertheilung im Innern der Erde. *Nachr. Ges. Wiss. Göttingen*, 221–43.

General books

Anderson, R. N. 1986. *Marine Geology: A Planet Earth Perspective*. New York: Wiley.
Brown, G. C. and Mussett, A. E. 1993. *The Inaccessible Earth*, 2nd edn. London: Chapman and Hall.
Cattermole, P. and Moore, P. 1985. *The Story of the Earth.* Cambridge: Cambridge University Press.
Clark, S. P. J. 1971. *Structure of the Earth*. Englewood Cliffs, New Jersey: Prentice-Hall.
Cloud, P. 1988. *Oasis in Space: Earth History from the Beginning*. New York: Norton.
Cole, G. H. A. 1986. *Inside a Planet*. Hull: Hull University Press.
Holmes, A. 1965. *Principles of Physical Geology*. New York: Ronald Press.
Lowrie, W. 1997. *Fundamentals of Geophysics*, Cambridge: Cambridge University Press.
van Andel, T. H. 1994. *New Views on an Old Planet, Continental Drift and the History of Earth*, 2nd edn. Cambridge: Cambridge University Press.
Wyllie, P. J. 1976. *The Way the Earth Works*. New York: Wiley.

Chapter 2

Tectonics on a sphere: the geometry of plate tectonics

2.1 Plate tectonics

The Earth has a cool and therefore mechanically strong outermost shell called the *lithosphere* (Greek *lithos*, 'rock'). The lithosphere is of the order of 100 km thick and comprises the crust and uppermost mantle. It is thinnest in the oceanic regions and thicker in continental regions, where its base is poorly understood. The *asthenosphere* (Greek *asthenia*, 'weak' or 'sick') is that part of the mantle immediately beneath the lithosphere. The high temperature and pressure which exist at the depth of the asthenosphere cause its viscosity to be low enough to allow viscous flow to take place on a geological timescale (millions of years, not seconds!). If the Earth is viewed in purely mechanical terms, the mechanically strong lithosphere floats on the mechanically weak asthenosphere. Alternatively, if the Earth is viewed as a heat engine, the lithosphere is an outer skin, through which heat is lost by conduction, and the asthenosphere is an interior shell through which heat is transferred by convection (Section 7.1).

The basic concept of *plate tectonics* is that the lithosphere is divided into a small number of nearly rigid *plates* (like curved caps on a sphere), which are moving over the asthenosphere. Most of the deformation which results from the motion of the plates – such as stretching, folding or shearing – takes place along the edge, or boundary, of a plate. Deformation away from the boundary is not significant.

A map of the *seismicity* (earthquake activity) of the Earth (Fig. 2.1) outlines the plates very clearly because nearly all earthquakes, as well as most of the Earth's volcanism, occur along the plate boundaries. These *seismic belts* are the zones in which differential movements between the nearly rigid plates occur. There are seven main plates, of which the largest is the Pacific plate, and numerous smaller plates such as Nazca, Cocos and Scotia plates (Fig. 2.2).

The theory of plate tectonics, which describes the interactions of the lithospheric plates and the consequences of these interactions, is based on several important assumptions.

1. The generation of new plate material occurs by *seafloor spreading*; that is, new oceanic lithosphere is generated along the active mid-ocean ridges (see Chapters 3 and 9).
2. The new oceanic lithosphere, once created, forms part of a rigid plate; this plate may but need not include continental material.
3. The Earth's surface area remains constant; therefore the generation of new plate by seafloor spreading must be balanced by destruction of plate elsewhere.
4. The plates are capable of transmitting stresses over great horizontal distances without buckling, in other words, the relative motion between plates is taken up only along plate boundaries.

Plate boundaries are of three types.

1. Along *divergent* boundaries, which are also called accreting or constructive, plates are moving away from each other. At such boundaries new plate material, derived from the mantle, is added to the lithosphere. The divergent plate boundary is represented by the *mid-ocean-ridge system*, along the axis of which new plate material is produced (Fig. 2.3(a)).
2. Along *convergent* boundaries, which are also called consuming or destructive, plates approach each other. Most such boundaries are represented by the *oceanic-trench*, *island-arc* systems of *subduction zones* where one of the colliding plates descends into the mantle and is destroyed (Fig. 2.3(c)). The downgoing plate often penetrates the mantle to depths of about 700 km. Some convergent boundaries occur on land. Japan, the Aleutians and the Himalayas are the surface expression of convergent plate boundaries.
3. Along *conservative* boundaries, lithosphere is neither created nor destroyed. The plates move laterally relative to each other (Fig. 2.3(e)). These plate boundaries are represented by *transform faults*, of which the San Andreas Fault in California, U.S.A. is a famous example. Transform faults can be grouped into six basic classes (Fig. 2.4). By far the most common type of transform fault is the ridge–ridge fault (Fig. 2.4(a)), which can range from a few kilometres to hundreds of kilometres in length. Some very long ridge–ridge faults occur in the Pacific, equatorial Atlantic and southern oceans (see Fig. 2.2, which shows the present plate boundaries, and Table 8.3). Adjacent plates move relative to each other at rates up to about 15 cm yr^{-1}.

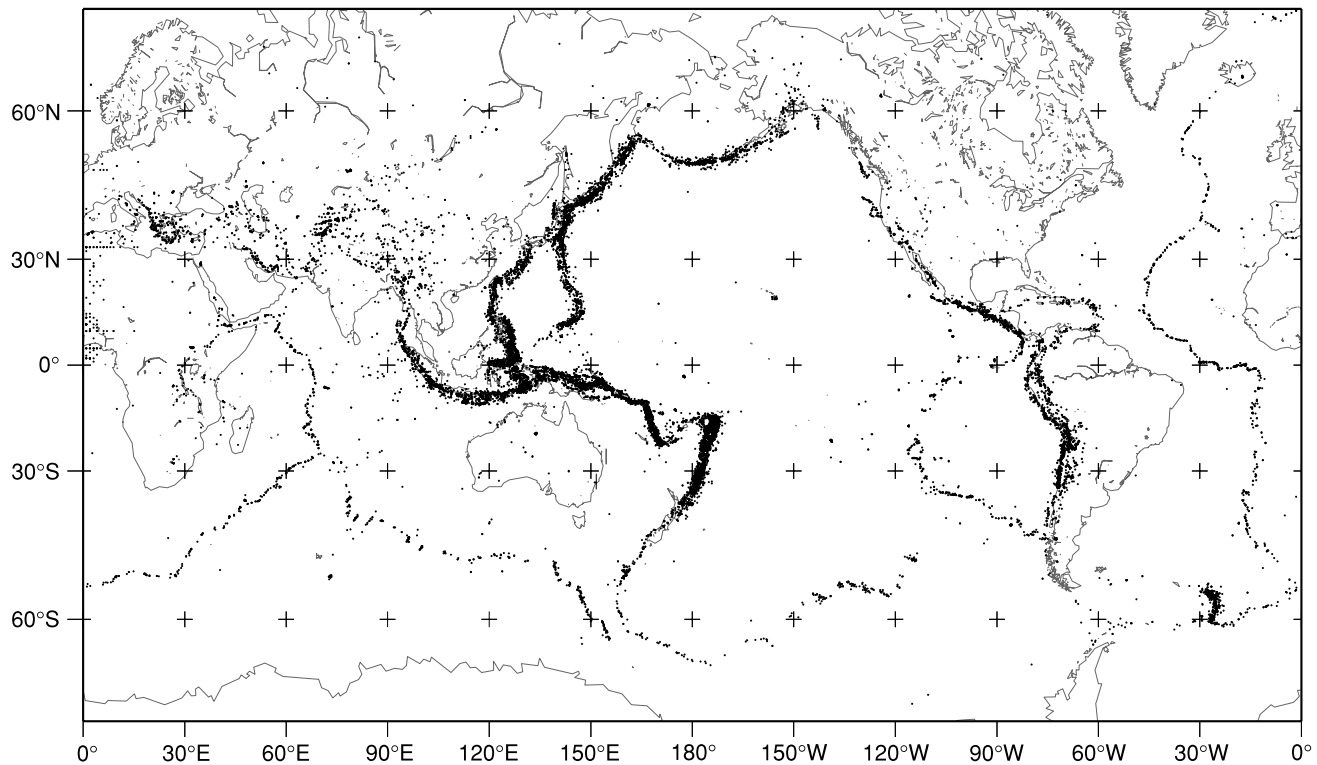


Figure 2.1. Twenty-three thousand earthquakes with magnitudes greater than 5.2 occurred between 1978 and 1989 at depths from 0 to 700 km. These earthquakes clearly delineate the boundaries of the plates. (From ISC catalogue.)

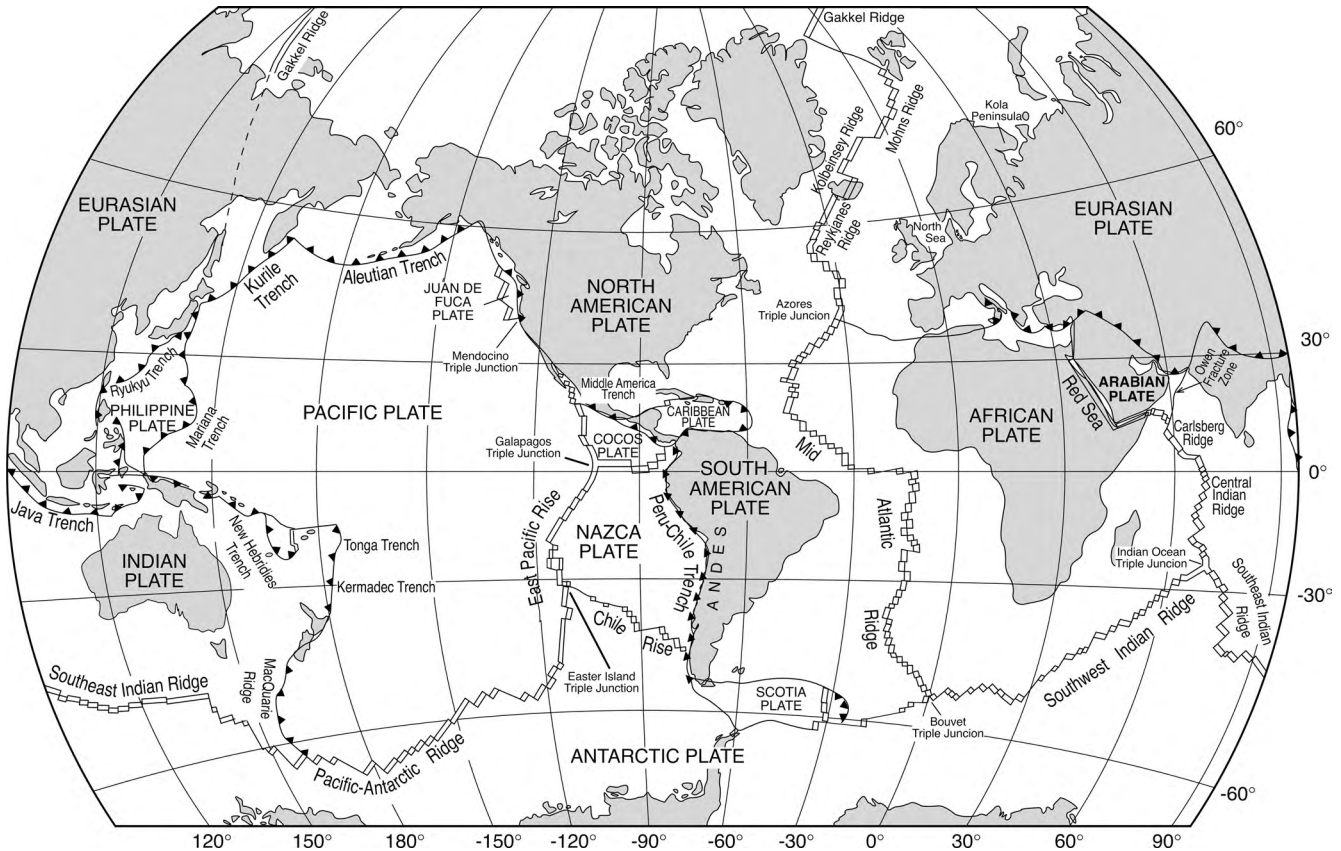


Figure 2.2. The major tectonic plates, mid-ocean ridges, trenches and transform faults.

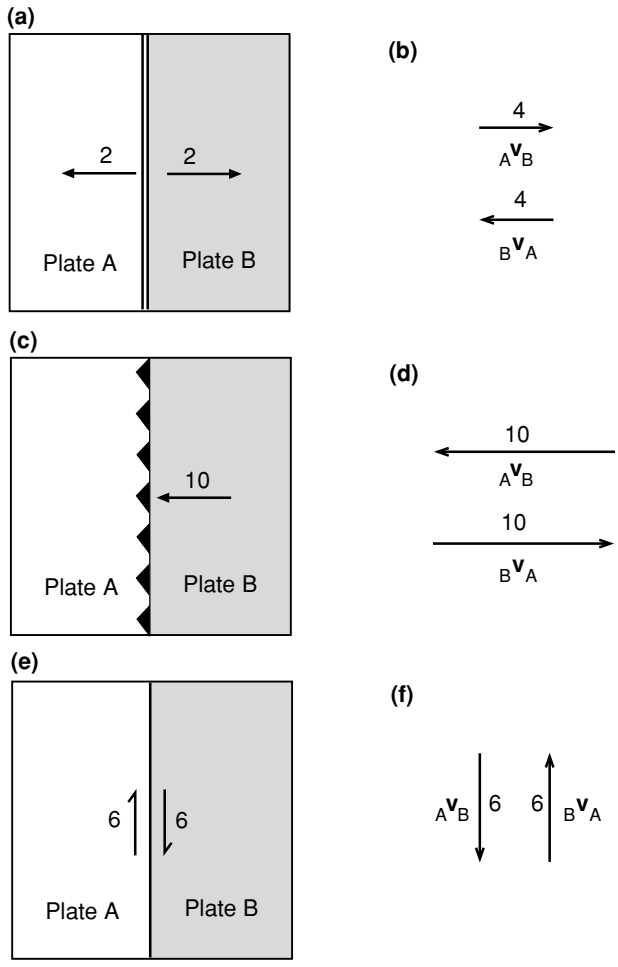


Figure 2.3. Three possible boundaries between plates A and B.

(a) A *constructive* boundary (mid-ocean ridge). The double line is the symbol for the ridge axis, and the arrows and numbers indicate the direction of spreading and relative movement of the plates away from the ridge. In this example the half-spreading rate of the ridge (half-rate) is 2 cm yr^{-1}; that is, plates A and B are moving apart at 4 cm yr^{-1}, and each plate is growing at 2 cm yr^{-1}. (b) The relative velocities $_A\mathbf{v}_B$ and $_B\mathbf{v}_A$ for the ridge shown in (a).

(c) A *destructive* boundary (subduction zone). The barbed line is the symbol for a subduction zone; the barbs are on the side of the overriding plate, pointing away from the subducting or downgoing plate. The arrow and number indicate the direction and rate of relative motion between the two plates. In this example, plate B is being subducted at 10 cm yr^{-1}. (d) The relative velocities $_A\mathbf{v}_B$ and $_B\mathbf{v}_A$ for the subduction zone shown in (c).

(e) A *conservative* boundary (transform fault). The single line is a symbol for a transform fault. The half-arrows and number indicate the direction and rate of relative motion between the plates: in this example, 6 cm yr^{-1}. (f) The relative velocities $_A\mathbf{v}_B$ and $_B\mathbf{v}_A$ for the transform fault shown in (e).

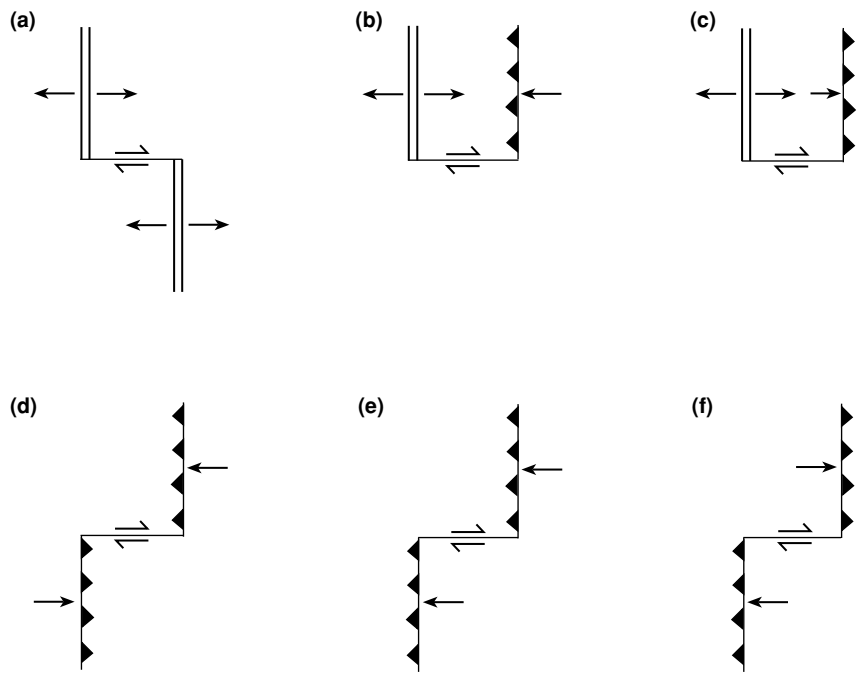


Figure 2.4. The six types of dextral (right-handed) transform faults. There are also six sinistral (left-handed) transform faults, mirror images of those shown here. (a) Ridge–ridge fault, (b) and (c) ridge–subduction-zone fault, (d), (e) and (f) subduction-zone–subduction-zone fault. (After Wilson (1965).)

The present-day rates of movement for all the main plates are discussed in Section 2.4.

Although the plates are made up of both oceanic and continental material, usually only the oceanic part of any plate is created or destroyed. Obviously, seafloor spreading at a mid-ocean ridge produces only oceanic lithosphere, but it is hard to understand why continental material usually is not destroyed at convergent plate boundaries. At subduction zones, where continental and oceanic materials meet, it is the oceanic plate which is subducted (and thereby destroyed). It is probable that, if the thick, relatively low-density continental material (the continental crustal density is approximately 2.8×10^3 kg m^{-3}) reaches a subduction zone, it may descend a short way, but, because the mantle density is so much greater (approximately 3.3×10^3 kg m^{-3}), the downwards motion does not continue. Instead, the subduction zone ceases to operate at that place and moves to a more favourable location. Mountains are built (*orogeny*) above subduction zones as a result of continental collisions. In other words, the continents are rafts of lighter material, which remain on the surface while the denser oceanic lithosphere is subducted beneath either oceanic or continental lithosphere. The discovery that plates can include both continental and oceanic parts, but that only the oceanic parts are created or destroyed, removed the main objection to the theory of *continental drift*, which was the unlikely concept that somehow continents were ploughing through oceanic rocks.

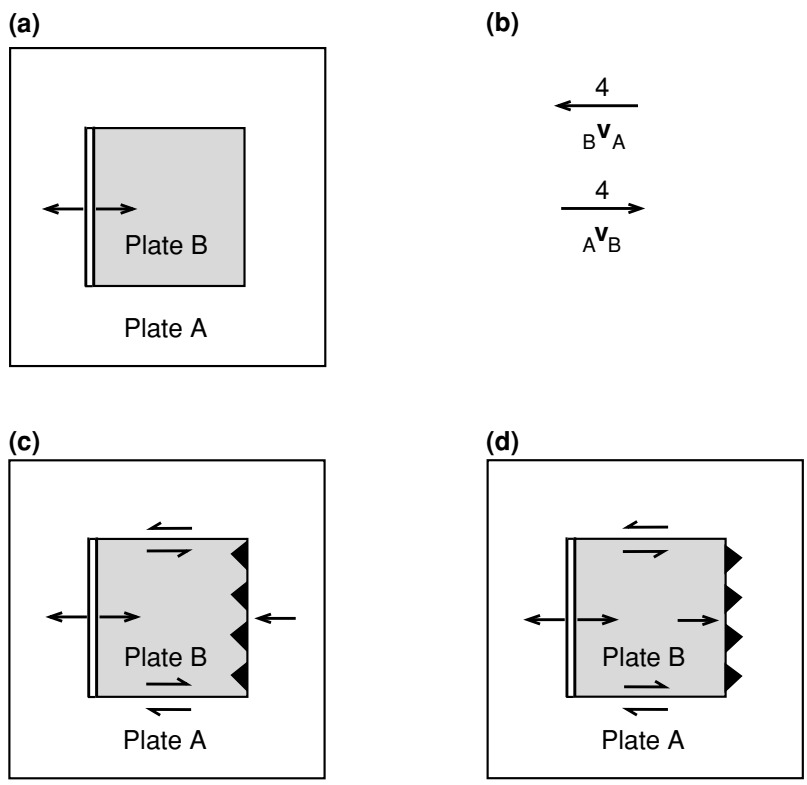


Figure 2.5. (a) A two-plate model on a flat planet. Plate B is shaded. The western boundary of plate B is a ridge from which seafloor spreads at a half-rate of 2 cm yr^{-1}. (b) Relative velocity vectors ${}_A\mathbf{v}_B$ and ${}_B\mathbf{v}_A$ for the plates in (a). (c) One solution to the model shown in (a): the northern and southern boundaries of plate B are transform faults, and the eastern boundary is a subduction zone with plate B overriding plate A. (d) An alternative solution for the model in (a): the northern and southern boundaries of plate B are transform faults, and the eastern boundary is a subduction zone with plate A overriding plate B.

2.2 A flat Earth

Before looking in detail at the motions of plates on the surface of the Earth (which of necessity involves some spherical geometry), it is instructive to return briefly to the Middle Ages so that we can consider a flat planet.

Figure 2.3 shows the three types of plate boundary and the ways they are usually depicted on maps. To describe the relative motion between the two plates A and B, we must use a vector that expresses their relative rate of movement (relative velocity). The velocity of plate A with respect to plate B is written ${}_B\mathbf{v}_A$ (i.e., if you are an observer on plate B, then ${}_B\mathbf{v}_A$ is the velocity at which you see plate A moving). Conversely, the velocity of plate B with respect to plate A is ${}_A\mathbf{v}_B$, and

$$ {}_A\mathbf{v}_B = -{}_B\mathbf{v}_A \tag{2.1} $$

Figure 2.3 illustrates these vectors for the three types of plate boundary.

To make our models more realistic, let us set up a two-plate system (Fig. 2.5(a)) and try to determine the more complex motions. The western boundary of plate B is a ridge that is spreading with a half-rate of 2 cm yr^{-1}. This information enables us to draw ${}_A\mathbf{v}_B$ and ${}_B\mathbf{v}_A$ (Fig. 2.5(b)). Since we know the

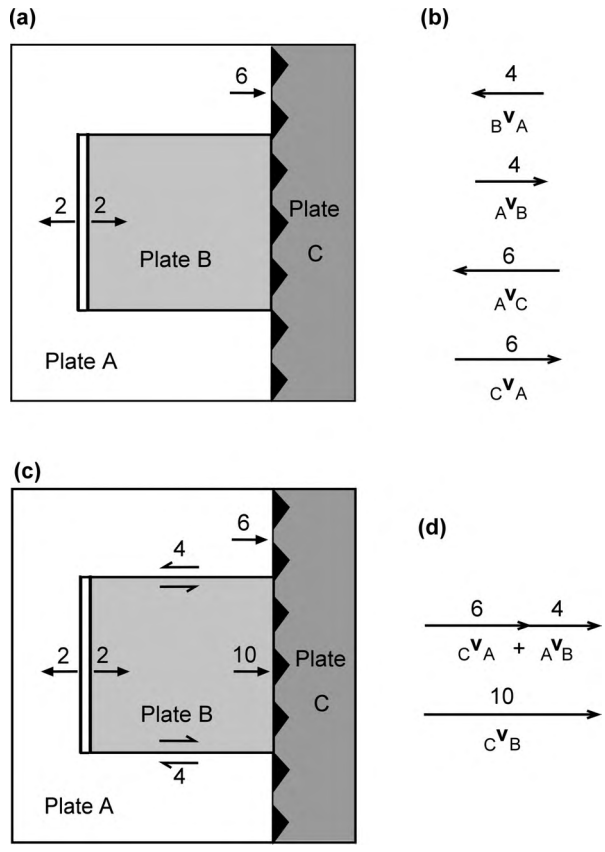


Figure 2.6. (a) A three-plate model on a flat planet. Plate A is unshaded. The western boundary of plate B is a ridge spreading at a half-rate of 2 cm yr^{-1}. The boundary between plates A and C is a subduction zone with plate C overriding plate A at 6 cm yr^{-1}. (b) Relative velocity vectors for the plates shown in (a). (c) The solution to the model in (a): the northern and southern boundaries of plate B are transform faults, and the eastern boundary is a subduction zone with plate C overriding plate B at 10 cm yr^{-1}. (d) Vector addition to determine the velocity of plate B with respect to plate C, $_C\mathbf{v}_B$.

shape of plate B, we can see that its northern and southern boundaries must be transform faults. The northern boundary is *sinistral*, or left-handed; rocks are offset to the left as you cross the fault. The southern boundary is *dextral*, or right-handed; rocks are offset to the right as you cross the fault. The eastern boundary is ambiguous: $_A\mathbf{v}_B$ indicates that plate B is approaching plate A at 4 cm y^{-1} along this boundary, which means that a subduction zone is operating there; but there is no indication as to which plate is being subducted. The two possible solutions for this model are shown in Figs. 2.5(c) and (d). Figure 2.5(c) shows plate A being subducted beneath plate B at 4 cm yr^{-1}. This means that plate B is increasing in width by 2 cm yr^{-1}, this being the rate at which new plate is formed at the ridge axis. Figure 2.5(d) shows plate B being subducted beneath plate A at 4 cm yr^{-1}, faster than new plate is being created at its western boundary (2 cm yr^{-1}); so eventually plate B will cease to exist on the surface of the planet.

If we introduce a third plate into the model, the motions become more complex still (Fig. 2.6(a)). In this example, plates A and B are spreading away from the ridge at a half-rate of 2 cm yr^{-1}, just as in Fig. 2.5(a). The eastern boundary of plates A and B is a subduction zone, with plate A being subducted beneath plate

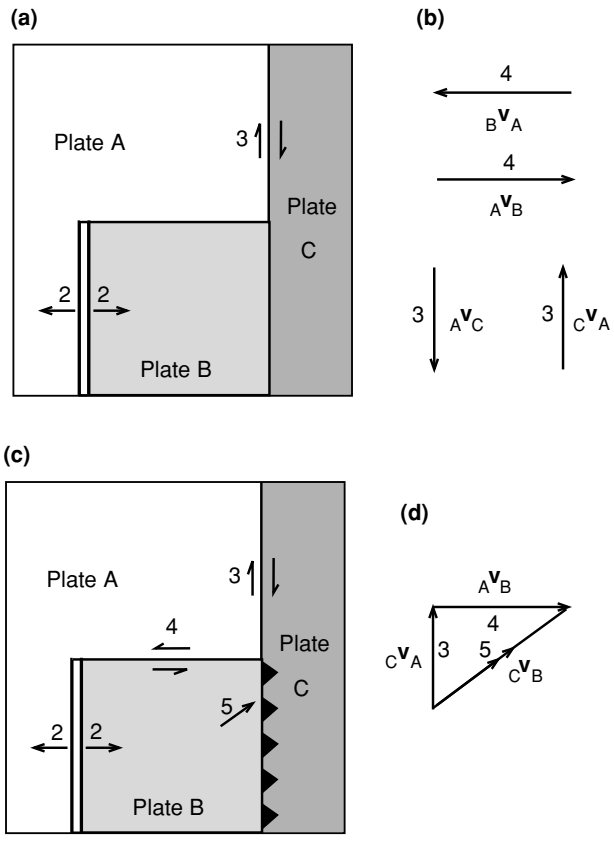


Figure 2.7. (a) A three-plate model on a flat planet. Plate A is unshaded. The western boundary of plate B is a ridge from which seafloor spreads at a half-rate of 2 cm yr^{-1}. The boundary between plates A and C is a transform fault with relative motion of 3 cm yr^{-1}. (b) Relative velocity vectors for the plates shown in (a). (c) The stable solution to the model in (a): the northern boundary of plate B is a transform fault with a 4 cm yr^{-1} slip rate, and the boundary between plates B and C is a subduction zone with an oblique subduction rate of 5 cm yr^{-1}. (d) Vector addition to determine the velocity of plate B with respect to plate C, $_C\mathbf{v}_B$.

C at 6 cm yr^{-1}. The presence of plate C does not alter the relative motions across the northern and southern boundaries of plate B; these boundaries are transform faults just as in Fig. 2.5. To determine the relative rate of plate motion at the boundary between plates B and C, we must use vector addition:

$$_C\mathbf{v}_B = {}_C\mathbf{v}_A + {}_A\mathbf{v}_B \tag{2.2}$$

This is demonstrated in Fig. 2.6(d): plate B is being subducted beneath plate C at 10 cm yr^{-1}. This means that the net rate of destruction of plate B is $10 - 2 = 8$ cm yr^{-1}; eventually, plate B will be totally subducted, and a simple two-plate subduction model will be in operation. However, if plate B were overriding plate C, it would be increasing in width by 2 cm yr^{-1}.

So far the examples have been straightforward in that all relative motions have been in an east–west direction. (Vector addition was not really necessary; common sense works equally well.) Now let us include motion in the north–south direction also. Figure 2.7(a) shows the model of three plates A, B and C: the western boundary of plate B is a ridge that is spreading at a half-rate of 2 cm yr^{-1}, the northern boundary of plate B is a transform fault (just as in the other examples)

and the boundary between plates A and C is a transform fault with relative motion of 3 cm yr^{-1}. The motion at the boundary between plates B and C is unknown and must be determined by using Eq. (2.2). For this example it is necessary to draw a vector triangle to determine ${}_{C}\mathbf{v}_{B}$ (Fig. 2.7(d)). A solution to the problem is shown in Fig. 2.7(c): plate B undergoes oblique subduction beneath plate C at 5 cm yr^{-1}. The other possible solution is for plate C to be subducted beneath plate B at 5 cm yr^{-1}. In that case, the boundary between plates C and B would not remain collinear with the boundary between plates B and C but would move steadily to the east. (This is an example of the instability of a *triple junction*; see Section 2.6.)

These examples should give some idea of what can happen when plates move relative to each other and of the types of plate boundaries that occur in various situations. Some of the problems at the end of this chapter refer to a flat Earth, such as we have assumed for these examples. The real Earth, however, is spherical, so we need to use some spherical geometry.

2.3 Rotation vectors and rotation poles

To describe motions on the surface of a sphere we use Euler's '*fixed-point*' theorem, which states that 'The most general displacement of a rigid body with a fixed point is equivalent to a rotation about an axis through that fixed point.'

Taking a plate as a rigid body and the centre of the Earth as a fixed point, we can restate this theorem: 'Every displacement from one position to another on the surface of the Earth can be regarded as a rotation about a suitably chosen axis passing through the centre of the Earth.'

This restated theorem was first applied by Bullard *et al.* (1965) in their paper on continental drift, in which they describe the fitting of the coastlines of South America and Africa. The 'suitably chosen axis' which passes through the centre of the Earth is called the *rotation axis*, and it cuts the surface of the Earth at two points called the *poles of rotation* (Fig. 2.8(a)). These are purely mathematical points and have no physical reality, but their positions describe the directions of motion of all points along the plate boundary. The magnitude of the angular velocity about the axis then defines the magnitude of the relative motion between the two plates. Because angular velocities behave as vectors, the relative motion between two plates can be written as $\boldsymbol{\omega}$, a vector directed along the rotation axis. The magnitude of $\boldsymbol{\omega}$ is ω, the angular velocity. The sign convention used is that a rotation that is clockwise (or right-handed) when viewed from the centre of the Earth along the rotation axis is positive. Viewed from outside the Earth, a positive rotation is anticlockwise. Thus, one rotation pole is positive and the other is negative (Fig. 2.8(b)).

Consider a point X on the surface of the Earth (Fig. 2.8(c)). At X the value of the relative velocity v between the two plates is

$$v = \omega R \sin\theta \tag{2.3}$$

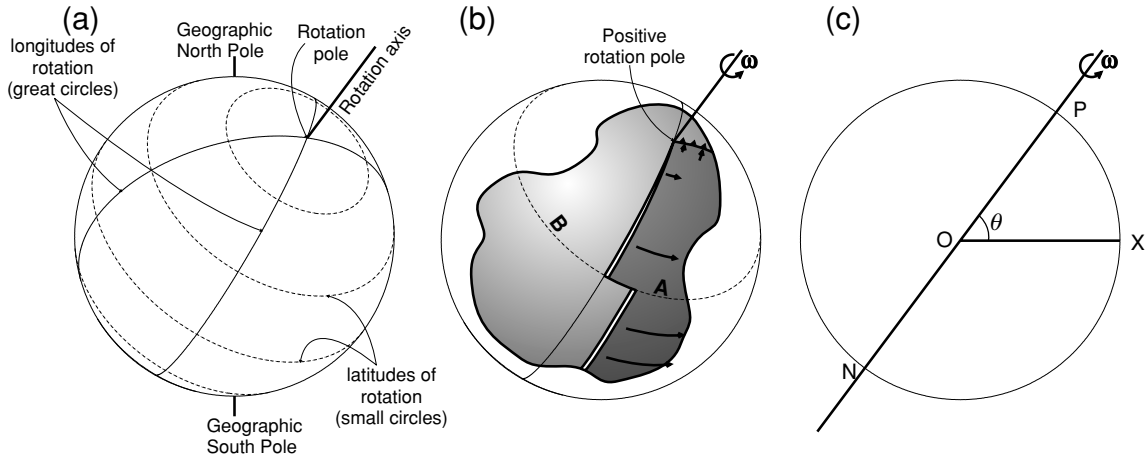


Figure 2.8. The movement of plates on the surface of the Earth. (a) The lines of latitude of rotation around the rotation poles are small circles (shown dashed) whereas the lines of longitude of rotation are great circles (i.e., circles with the same diameter as the Earth). Note that these lines of latitude and longitude are *not* the geographic lines of latitude and longitude because the poles for the geographic coordinate system are the North and South Poles, not the rotation poles. (b) Constructive, destructive and conservative boundaries between plates A and B. Plate B is assumed to be fixed so that the motion of plate A is relative to plate B. The visible rotation pole is positive (motion is anticlockwise when viewed from outside the Earth). Note that the spreading and subduction rates increase with distance from the rotation pole. The transform fault is an arc of a small circle (shown dashed) and thus is perpendicular to the ridge axis. As the plate boundary passes the rotation pole, the boundary changes from constructive to destructive, i.e. from ridge to subduction zone. (c) A cross section through the centre of the Earth O. P and N are the positive and negative rotation poles, and X is a point on the plate boundary.

where θ is the angular distance between the rotation pole P and the point X, and R is the radius of the Earth. This factor of $\sin\theta$ means that the relative motion between two adjacent plates changes with position along the plate boundary, in contrast to the earlier examples for a flat Earth. Thus, the relative velocity is zero at the rotation poles, where $\theta = 0^\circ$ and 180°, and has a maximum value of ωR at 90° from the rotation poles. If by chance the plate boundary passes through the rotation pole, the nature of the boundary changes from divergent to convergent, or vice versa (as in Fig. 2.8(b)). Lines of constant velocity (defined by $\theta =$ constant) are small circles about the rotation poles.

2.4 Present-day plate motions

2.4.1 Determination of rotation poles and rotation vectors

Several methods can be used to find the present-day *instantaneous poles of rotation* and *relative angular velocities* between pairs of plates. *Instantaneous* refers to a geological instant; it means a value averaged over a period of time ranging

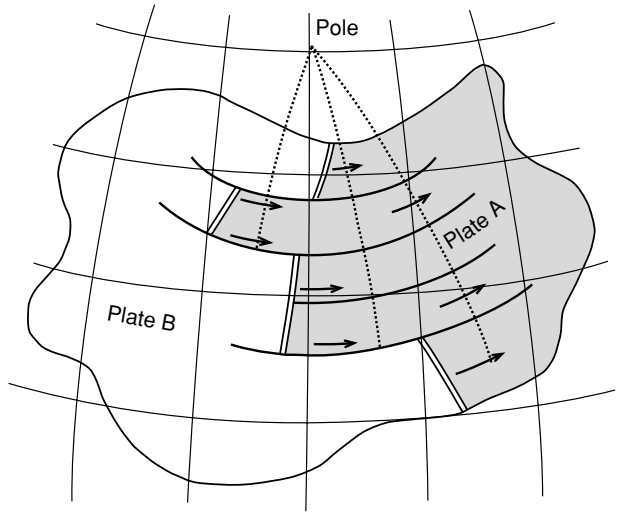


Figure 2.9. On a spherical Earth the motion of plate A relative to plate B must be a rotation about some pole. All the transform faults on the boundary between plates A and B must be small circles about that pole. Transform faults can be used to locate the pole: it lies at the intersection of the great circles which are perpendicular to the transform faults. Although ridges are generally perpendicular to the direction of spreading, this is not a geometric requirement, so it is not possible to determine the relative motion or locate the pole from the ridge itself. (After Morgan (1968).)

from a few years to a few million years, depending on the method used. These methods include the following.

1. A local determination of the direction of relative motion between two plates can be made from the strike of active transform faults. Methods of recognizing transform faults are discussed fully in Section 8.5. Since transform faults on ridges are much easier to recognize and more common than transform faults along destructive boundaries, this method is used primarily to find rotation poles for plates on either side of a mid-ocean ridge. The relative motion at transform faults is parallel to the fault and is of constant value along the fault. This means that the faults are arcs of small circles about the rotation pole. The rotation pole must therefore lie somewhere on the great circle which is perpendicular to that small circle. So, if two or more transform faults can be used, the intersection of the great circles is the position of the rotation pole (Fig. 2.9).
2. The spreading rate along a constructive plate boundary changes as the sine of the angular distance θ from the rotation pole (Eq. (2.3)). So, if the spreading rate at various locations along the ridge can be determined (from spacing of oceanic magnetic anomalies as discussed in Chapter 3), the rotation pole and angular velocity can then be estimated.
3. The analysis of data from an earthquake can give the direction of motion and the plane of the fault on which the earthquake occurred. This is known as a *fault-plane*

solution or a *focal mechanism* (discussed fully in Section 4.2.8). Fault-plane solutions for earthquakes along a plate boundary can give the direction of relative motion between the two plates. For example, earthquakes occurring on the transform fault between plates A and B in Fig. 2.8(b) would indicate that there is right lateral motion across the fault. The location of the pole and the direction, though not the magnitude, of the motion can thus be estimated.

4. Where plate boundaries cross land, surveys of displacements can be used (over large distances and long periods of time) to determine the local relative motion. For example, stream channels and even roads, field boundaries and buildings may be displaced.
5. Satellites have made it possible to measure instantaneous plate motions with some accuracy. One method uses a satellite laser-ranging system (SLR) to determine differences in distance between two sites on the Earth's surface over a period of years. Another method, very-long-baseline interferometry (VLBI), uses quasars for the signal source and terrestrial radio telescopes as the receivers. Again, the difference in distance between two telescope sites is measured over a period of years. Worldwide, the rates of plate motion determined by VLBI and SLR agree with geologically determined rates to within 2%.

   A third method of measuring plate motions utilizes the Global Positioning System (GPS) which was developed to provide real-time navigation and positioning using satellites. A worldwide network of GPS receivers with a precision suitable for geodynamics has been established (1 cm in positioning and $<10^{-3}$ arcsec in pole-position estimates). It is called the International GPS Service for Geodynamics (IGS), and is a permanent global network of receivers. Analysis of data from 1991–1996 shows that the agreement of GPS velocities with the geologically determined velocities for all but a few locations is to better than 95% confidence. This is another impressive corroboration of relative plate motions – *the plates are continually in motion*.

An estimate of the present-day plate motions, NUVEL-1A, made by using 277 measurements of ridge spreading rate, 121 oceanic transform-fault azimuths and 724 earthquake slip vectors is given in Table 2.1. Figure 2.10 shows velocities in southern California relative to North America as determined from geodetic measurements (including GPS and VLBI) between 1972 and 1995. The measured velocity of motion across this boundary was 50 mm yr^{-1}, compared with the 49 mm yr^{-1} predicted (Table 2.1). Thus again geological estimates based on measurements with a timescale of a million years agree with measurements made over a few years.

Although clear boundaries between rigid plates describe the relative motions and structures well, there are a few boundaries for which the term '*diffuse plate boundary*' is appropriate. The main examples are the North American and South American plate boundary from the Mid-Atlantic Ridge to the Caribbean and the boundary which subdivides the Indian plate.

It is important to realize that a rotation with a large angular velocity ω does not necessarily mean that the relative motion along the plate boundary is also large.

Table 2.1 *Rotation vectors for the present-day relative motion between some pairs of plates: NUVEL-1A*

	Positive-pole position		Angular velocity
Plates	Latitude	Longitude	(10^{-7} deg yr^{-1})
Africa–Antarctica	5.6°N	39.2°W	1.3
Africa–Eurasia	21.0°N	20.6°W	1.2
Africa–North America	78.8°N	38.3°E	2.4
Africa–South America	62.5°N	39.4°W	3.1
Australia–Antarctica	13.2°N	38.2°E	6.5
Pacific–Antarctica	64.3°S	96.0°E	8.7
South America–Antarctica	86.4°S	139.3°E	2.6
Arabia–Eurasia	24.6°N	13.7°E	5.0
India–Eurasia	24.4°N	17.7°E	5.1
Eurasia–North America	62.4°N	135.8°E	2.1
Eurasia–Pacific	61.1°N	85.8°W	8.6
Pacific–Australia	60.1°S	178.3°W	10.7
North America–Pacific	48.7°N	78.2°W	7.5
Cocos–North America	27.9°N	120.7°W	13.6
Nazca–Pacific	55.6°N	90.1°W	13.6
Nazca–South America	56.0°N	94.0°W	7.2

Note: The first plate moves anticlockwise with respect to the second plate as shown.
Source: After DeMets *et al.* (1990; 1994).

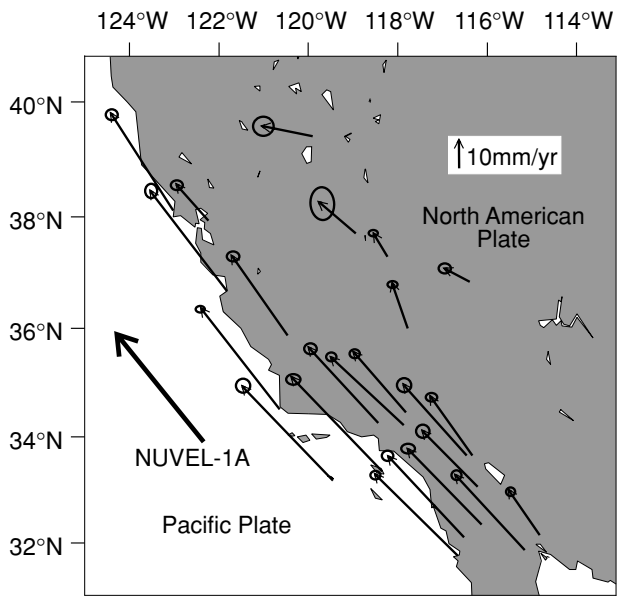


Figure 2.10. Motion of southern California with respect to North America. Error ellipses represent 95% confidence. (After Shen *et al.* (1997). Crustal deformation measured in Southern California, *EOS Trans. Am. Geophys. Un.*, **78** (43), 477 and 482, 1997. Copyright 1997 American Geophysical Union. Reprinted by permission of American Geophysical Union)

Table 2.2 *Symbols used in calculations involving rotation poles*

Symbol	Meaning	Sign convention
λ_p	Latitude of rotation pole P	°N positive
λ_x	Latitude of point X on plate boundary	°S negative
ϕ_p	Longitude of rotation pole P	°W negative
ϕ_x	Longitude of point X on plate boundary	°E positive
v	Velocity of point X on plate boundary	
v	Amplitude of velocity v	
β	Azimuth of the velocity with respect to north N	Clockwise positive
R	Radius of the Earth	
ω	Angular velocity about rotation pole P	

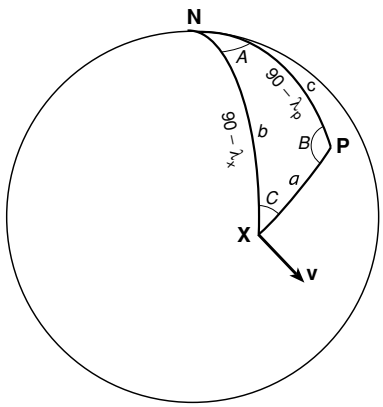


Figure 2.11. A diagram showing the relative positions of the positive rotation pole P and point X on the plate boundary. N is the North Pole. The sides of the spherical triangle NPX are all great circles, the sides NX and NP are lines of geographic longitude. The vector **v** is the relative velocity at point X on the plate boundary (note that **v** is perpendicular to PX). It is usual to quote the lengths of the sides of spherical triangles as angles (e.g., latitude and longitude when used as geographic coordinates).

The distance between rotation pole and plate boundary is important (remember the $\sin\theta$ factor multiplying the angular velocity ω in Eq. (2.3)). However, it is conventional to use the relative velocity at $\theta = 90^\circ$ when quoting a relative velocity for two plates, even though neither plate may extend 90° from the pole.

2.4.2 Calculation of the relative motion at a plate boundary

Once the instantaneous rotation pole and angular velocity for a pair of adjacent plates have been determined, they can be used to calculate the direction and magnitude of the relative motion at any point along the plate boundary.

The notation and sign conventions used in the following pages are given in Table 2.2. Figure 2.11 shows the relative positions of the North Pole N, positive rotation pole P and point X on the plate boundary (compare with Fig. 2.8(b)). In the spherical triangle NPX, let the angles $\widehat{\mathrm{XNP}} = A$, $\widehat{\mathrm{NPX}} = B$ and $\widehat{\mathrm{PXN}} = C$,

and let the angular lengths of the sides of the triangle be PX $= a$, XN $= b$ and NP $= c$. Thus, the angular lengths b and c are known, but a is not:

$$b = 90 - \lambda_x \tag{2.4}$$

$$c = 90 - \lambda_p \tag{2.5}$$

Angle A is known, but B and C are not:

$$A = \phi_p - \phi_x \tag{2.6}$$

Equation (2.3) is used to obtain the magnitude of the velocity at point X:

$$v = \omega R \sin a \tag{2.7}$$

The azimuth of the velocity, β is given by

$$\beta = 90 + C \tag{2.8}$$

To find the angles a and C needed for Eqs. (2.7) and (2.8), we use spherical geometry. Just as there are cosine and sine rules relating the angles and sides of plane triangles, there are cosine and sine rules for spherical triangles:

$$\cos a = \cos b \, \cos c + \sin b \, \sin c \, \cos A \tag{2.9}$$

and

$$\frac{\sin a}{\sin A} = \frac{\sin c}{\sin C} \tag{2.10}$$

Substituting Eqs. (2.4)–(2.6) into Eq. (2.9) gives

$$\begin{aligned} \cos a = {} & \cos(90 - \lambda_x) \, \cos(90 - \lambda_p) \\ & + \sin(90 - \lambda_x) \, \sin(90 - \lambda_p) \cos(\phi_p - \phi_x) \end{aligned} \tag{2.11}$$

This can then be simplified to yield the angle a, which is needed to calculate the velocity from Eq. (2.7):

$$a = \cos^{-1}[\sin \lambda_x \sin \lambda_p + \cos \lambda_x \cos \lambda_p \cos(\phi_p - \phi_x)] \tag{2.12}$$

Substituting Eqs. (2.5) and (2.6) into Eq. (2.10) gives

$$\frac{\sin a}{\sin(\phi_p - \phi_x)} = \frac{\sin(90 - \lambda_p)}{\sin C} \tag{2.13}$$

Upon rearrangement this becomes

$$C = \sin^{-1}\left(\frac{\cos \lambda_p \, \sin(\phi_p - \phi_x)}{\sin a}\right) \tag{2.14}$$

Therefore, if the angle a is calculated from Eq. (2.12), angle C can then be calculated from Eq. (2.14), and, finally, the relative velocity and its azimuth can be calculated from Eqs. (2.7) and (2.8). Note that the inverse sine function of Eq. (2.14) is double-valued.[1] Always check that you have the correct value for C.

[1] An alternative way to calculate motion along a plate boundary and to avoid the sign ambiguities is to use vector algebra (see Altman (1986) or Cox and Hart (1986), p. 154).

Example: calculation of relative motion at a plate boundary

Calculate the present-day relative motion at 28°S, 71°W on the Peru–Chile Trench using the Nazca–South America rotation pole given in Table 2.1. Assume the radius of the Earth to be 6371 km:

$$\lambda_x = -28^\circ, \quad \phi_x = -71^\circ$$
$$\lambda_p = 56^\circ, \quad \phi_p = -94^\circ$$
$$\omega = 7.2 \times 10^{-7} \text{ deg yr}^{-1} = \frac{\pi}{180} \times 7.2 \times 10^{-7} \text{ rad yr}^{-1}$$

These values are substituted into Eqs. (2.12), (2.14), (2.7) and (2.8) in that order, giving

$$a = \cos^{-1}[\sin(-28)\sin(56) + \cos(-28)\cos(56)\cos(-94+71)]$$
$$= 86.26^\circ \tag{2.15}$$

$$C = \sin^{-1}\left(\frac{\cos(56)\ \sin(-94+71)}{\sin(86.26)}\right)$$
$$= -12.65^\circ \tag{2.16}$$

$$v = \frac{\pi}{180} \times 7.2 \times 10^{-7} \times 6371 \times 10^5 \times \sin(86.26)$$
$$= 7.97 \text{ cm yr}^{-1} \tag{2.17}$$

$$\beta = 90 - 12.65$$
$$= 77.35^\circ \tag{2.18}$$

Thus, the Nazca plate is moving relative to the South American plate at 8 cm yr^{-1} with azimuth 77°; the South American plate is moving relative to the Nazca plate at 8 cm yr^{-1}, azimuth 257° (Fig. 2.2).

2.4.3 Combination of rotation vectors

Suppose that there are three rigid plates A, B and C and that the angular velocity of A relative to B, ${}_B\boldsymbol{\omega}_A$, and that of B relative to C, ${}_C\boldsymbol{\omega}_B$, are known. The motion of plate A relative to plate C, ${}_C\boldsymbol{\omega}_A$, can be determined by vector addition just as for the flat Earth:

$${}_C\boldsymbol{\omega}_A = {}_C\boldsymbol{\omega}_B + {}_B\boldsymbol{\omega}_A \tag{2.19}$$

(Remember that in this notation the first subscript refers to the 'fixed' plate.) Alternatively, since ${}_B\boldsymbol{\omega}_A = -{}_A\boldsymbol{\omega}_B$, Eq. (2.19) can be written as

$${}_A\boldsymbol{\omega}_B + {}_B\boldsymbol{\omega}_C + {}_C\boldsymbol{\omega}_A = 0 \tag{2.20}$$

The resultant vector ${}_C\boldsymbol{\omega}_A$ of Eq. (2.19) must lie in the same plane as the two original vectors ${}_B\boldsymbol{\omega}_A$ and ${}_C\boldsymbol{\omega}_B$. Imagine the great circle on which these two poles

Table 2.3 *Notation used in addition of rotation vectors*

Rotation vector	Magnitude	Latitude of pole	Longitude of pole
${}_{B}\boldsymbol{\omega}_{A}$	${}_{B}\omega_{A}$	λ_{BA}	ϕ_{BA}
${}_{C}\boldsymbol{\omega}_{B}$	${}_{C}\omega_{B}$	λ_{CB}	ϕ_{CB}
${}_{C}\boldsymbol{\omega}_{A}$	${}_{C}\omega_{A}$	λ_{CA}	ϕ_{CA}

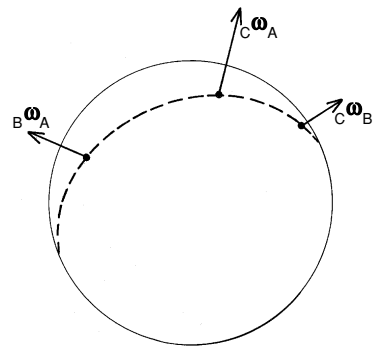


Figure 2.12. Relative-rotation vectors ${}_{B}\boldsymbol{\omega}_{A}$ and ${}_{C}\boldsymbol{\omega}_{A}$ for the plates A, B and C. The dashed line is the great circle on which the two poles lie. The resultant rotation vector is ${}_{C}\boldsymbol{\omega}_{A}$ (Eq. (2.19)). The resultant pole must also lie on the same great circle because the resultant rotation vector has to lie in the plane of the two original rotation vectors.

lie; the resultant pole must also lie on that same great circle (Fig. 2.12). Note that this relationship (Eqs. (2.19) and (2.20)) should be used only for infinitesimal movements or angular velocities, not for finite rotations. The theory of finite rotations is complex. (For a treatment of the whole theory of instantaneous and finite rotations, the reader is referred to Le Pichon *et al.* (1973).)

Let the three vectors ${}_{B}\boldsymbol{\omega}_{A}$, ${}_{C}\boldsymbol{\omega}_{B}$ and ${}_{C}\boldsymbol{\omega}_{A}$ be written as shown in Table 2.3. It is simplest to use a rectangular coordinate system through the centre of the Earth, with the x–y plane being equatorial, the x axis passing through the Greenwich meridian and the z axis passing through the North Pole, as shown in Fig. 2.13. The sign convention of Table 2.2 continues to apply. Then Eq. (2.19) can be written

$$x_{CA} = x_{CB} + x_{BA} \tag{2.21}$$

$$y_{CA} = y_{CB} + y_{BA} \tag{2.22}$$

$$z_{CA} = z_{CB} + z_{BA} \tag{2.23}$$

where x_{BA}, y_{BA} and z_{BA} are the x, y and z coordinates of the vector ${}_{B}\boldsymbol{\omega}_{A}$, and so on. Equations (2.21)–(2.23) become

$$x_{CA} = {}_{C}\omega_{B} \cos\lambda_{CB} \cos\phi_{CB} + {}_{B}\omega_{A} \cos\lambda_{BA} \cos\phi_{BA} \tag{2.24}$$

$$y_{CA} = {}_{C}\omega_{B} \cos\lambda_{CB} \sin\phi_{CB} + {}_{B}\omega_{A} \cos\lambda_{BA} \sin\phi_{BA} \tag{2.25}$$

$$z_{CA} = {}_{C}\omega_{B} \sin\lambda_{CB} + {}_{B}\omega_{A} \sin\lambda_{BA} \tag{2.26}$$

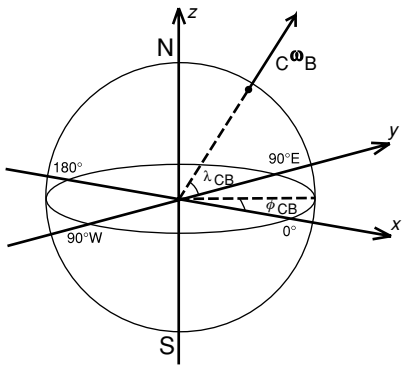


Figure 2.13. The rectangular coordinate system used in the addition of rotation vectors. The x–y plane is equatorial with the x axis passing through 0° Greenwich and the z axis through the North Pole. Notation and sign conventions are given in Table 2.3.

when the three rotation vectors are expressed in their x, y and z components. The magnitude of the resultant rotation vector, ${}_{\mathrm{C}}\omega_{\mathrm{A}}$, is

$$ {}_{\mathrm{C}}\omega_{\mathrm{A}} = \sqrt{x_{\mathrm{CA}}^2 + y_{\mathrm{CA}}^2 + z_{\mathrm{CA}}^2} \tag{2.27} $$

and the pole position is given by

$$ \lambda_{\mathrm{CA}} = \sin^{-1}\left(\frac{z_{\mathrm{CA}}}{{}_{\mathrm{C}}\omega_{\mathrm{A}}}\right) \tag{2.28} $$

and

$$ \phi_{\mathrm{CA}} = \tan^{-1}\left(\frac{y_{\mathrm{CA}}}{x_{\mathrm{CA}}}\right) \tag{2.29} $$

Note that this expression for ϕ_{CA} has an ambiguity of 180° (e.g., $\tan 30^\circ = \tan 210^\circ = 0.5774$, $\tan 110^\circ = \tan 290^\circ = -2.747$). This is resolved by adding or subtracting 180° so that

$$ x_{\mathrm{CA}} > 0 \quad \text{when } -90 < \phi_{\mathrm{CA}} < +90^\circ \tag{2.30} $$

$$ x_{\mathrm{CA}} < 0 \quad \text{when } |\phi_{\mathrm{CA}}| > 90^\circ \tag{2.31} $$

The problems at the end of this chapter enable the reader to use these methods to determine motions along real and imagined plate boundaries.

Example: addition of relative rotation vectors

Given the instantaneous rotation vectors in Table 2.1 for the Nazca plate relative to the Pacific plate and the Pacific plate relative to the Antarctic plate, calculate the instantaneous rotation vector for the Nazca plate relative to the Antarctic plate.

Plate	Rotation vector	Latitude of pole	Longitude of pole	Angular velocity (10^{-7} deg yr^{-1})
Nazca–Pacific	${}_{\mathrm{P}}\omega_{\mathrm{N}}$	55.6°N	90.1°W	13.6
Pacific–Antarctica	${}_{\mathrm{A}}\omega_{\mathrm{P}}$	64.3°S	96.0°E	8.7

To calculate the rotation vector for the Nazca plate relative to the Antarctic plate we apply Eq. (2 19):

$$ {}_{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{N}} = {}_{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{P}} + {}_{\mathrm{P}}\boldsymbol{\omega}_{\mathrm{N}} \tag{2.32} $$

Substituting the tabulated values into the equations for the x, y and z components of ${}_{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{N}}$ (Eqs. (2.24)–(2.26)) yields

$$ \begin{aligned} x_{\mathrm{AN}} &= 8.7\cos(-64.3)\cos(96.0) + 13.6\cos(55.6)\cos(-90.1) \\ &= -0.408 \end{aligned} \tag{2.33} $$

$$ \begin{aligned} y_{\mathrm{AN}} &= 8.7\cos(-64.3)\sin(96.0) + 13.6\cos(55.6)\sin(-90.1) \\ &= -3.931 \end{aligned} \tag{2.34} $$

$$ \begin{aligned} z_{\mathrm{AN}} &= 8.7\sin(-64.3) + 13.6\sin(55.6) \\ &= 3.382 \end{aligned} \tag{2.35} $$

The magnitude of the rotation vector ${}_{\mathrm{A}}\boldsymbol{\omega}_{\mathrm{N}}$ can now be calculated from Eq. (2.27) and the pole position from Eqs. (2.28) and (2.29):

$$ {}_{\mathrm{A}}\omega_{\mathrm{N}} = \sqrt{0.408^2 + 3.931^2 + 3.382^2} = 5.202 \tag{2.36} $$

$$ \lambda_{\mathrm{AN}} = \sin^{-1}\left(\frac{3.382}{5.202}\right) = 40.6 \tag{2.37} $$

$$ \phi_{\mathrm{AN}} = \tan^{-1}\left(\frac{-3.931}{-0.408}\right) = 180 + 84.1 \tag{2.38} $$

Therefore, the rotation for the Nazca plate relative to the Antarctic plate has a magnitude of 5.2×10^{-7} deg yr^{-1}, and the rotation pole is located at latitude 40.6°N, longitude 95.9°W.

2.5 Plate boundaries can change with time

The examples of plates moving upon a flat Earth (Section 2.2) illustrated that plates and plate boundaries do not stay the same for all time. This observation remains true when we advance from plates moving on a flat model Earth to plates moving on a spherical Earth. The formation of new plates and destruction of existing plates are the most obvious global reasons why plate boundaries and relative motions change. For example, a plate may be lost down a subduction zone, such as happened when most of the Farallon and Kula plates were subducted under the North American plate in the early Tertiary (see Section 3.3.3). Alternatively, two continental plates may coalesce into one (with resultant mountain building). If the position of a rotation pole changes, all the relative motions also change. A drastic change in pole position of say 90° would, of course, completely alter the status quo: transform faults would become ridges and subduction zones, and vice versa! Changes in the trends of transform faults and magnetic anomalies on the Pacific plate imply that the direction of seafloor spreading has changed there,

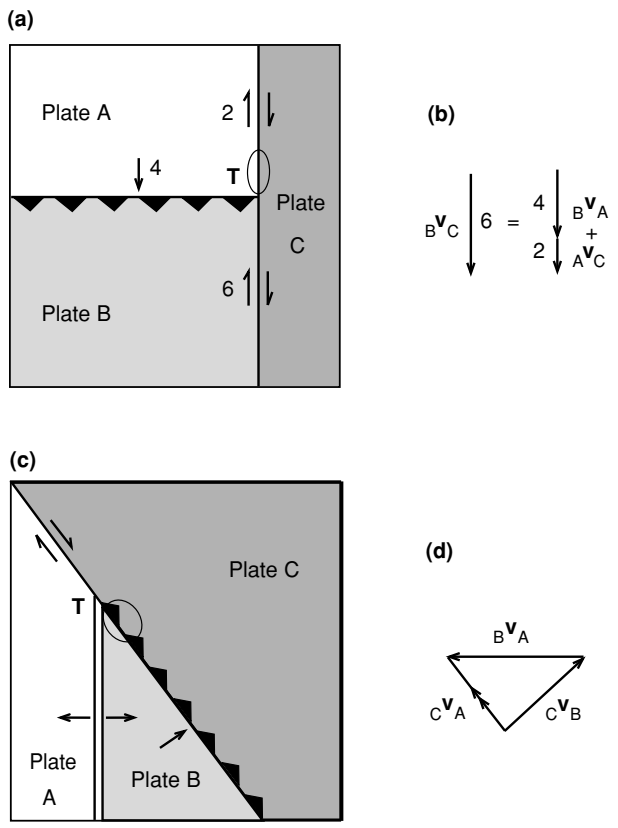


Figure 2.14. Two examples of a plate boundary that locally changes with time. (a) A three-plate model. Point T, where plates A, B and C meet, is the triple junction. The western boundary of plate C consists of transform faults. Plate B is overriding plate A at 4 cm yr^{-1}. The circled part of the boundary of plate C changes with time. (b) Relative-velocity vectors for the plates in (a). (c) A three-plate model. Point T, where plates A, B and C meet, is the triple junction. The boundary between plates A and B is a ridge, that between plates A and C is a transform fault and that between plates B and C is a subduction zone. The circled part of the boundary changes with time. (d) Relative-velocity vectors for the plates in (c). In these examples, velocity vectors have been used rather than angular-velocity vectors. This is justified, even for a spherical Earth, because these examples are concerned only with small regions in the immediate vicinity of the triple junctions, over which the relative velocities are constant.

and indicate that the Pacific–Farallon pole position changed slightly a number of times during the Tertiary.

Parts of plate boundaries can change locally, however, without any major 'plate' or 'pole' event occurring. Consider three plates A, B and C. Let there be a convergent boundary between plates A and B, and let there be strike–slip faults between plates A and C and plates B and C, as illustrated in Figs. 2.14(a) and (b). From the point of view of an observer on plate C, part of the boundary of C (circled) will change with time because the plate to which it is adjacent will change from plate A to plate B. The boundary will remain a dextral (right-handed) fault, but the slip rate will change from 2 cm yr^{-1} to 6 cm yr^{-1}. Relative to plate C, the subduction zone is moving northwards at 6 cm yr^{-1}. Another example of this type of plate-boundary change is illustrated in Figs. 2.14(c) and (d). In this case, the relative velocities are such that the boundary between plates A and C is a strike–slip fault, that between plates A and B is a ridge and that between plates B and C is a subduction zone. The motions are such that the ridge migrates slowly to the south relative to plate C, so the circled portion of plate boundary will change with time from subduction zone to transform fault.

These local changes in the plate boundary are a geometric, consequence of the motions of the three rigid plates rather than being caused by any disturbing outside event. A complete study of all possible interactions of three plates is made in the next section. Such a study is very important because it enables us to apply the theory of rigid geometric plates to the Earth and deduce past plate motions from evidence in the local geological record. We can also predict details of future plate interactions.

2.6 Triple junctions

2.6.1 Stable and unstable triple junctions

A *triple junction* is the name given to a point at which three plates meet, such as the points T in Fig. 2.14. A triple junction is said to be 'stable' when the relative motions of the three plates and the azimuth of their boundaries are such that the configuration of the junction does not change with time. The two examples shown in Fig. 2.14 are thus stable. In both cases the triple junction moves along the boundary of plate C, locally changing this boundary. The relative motions of the plates and triple junction and the azimuths and types of plate boundaries of the whole system do not change with time. An 'unstable' triple junction exists only momentarily before evolving to a different geometry. If four or more plates meet at one point, the configuration is always unstable, and the system will evolve into two or more triple junctions.

As a further example, consider a triple junction where three subduction zones meet (Fig. 2.15): plate A is overriding plates B and C, and plate C is overriding plate B. The relative-velocity triangle for the three plates at the triple junction is shown in Fig. 2.15(b). Now consider how this triple junction evolves with time. Assume that plate A is fixed; then the positions of the plates at some later time are as shown in Fig. 2.15(c). The dashed boundaries show the extent of the subducted parts of plates B and C. The subduction zone between plates B and C has moved north along the north–south edge of plate A. Thus, the original triple junction (Fig. 2.15(a)) was unstable; however, the new triple junction (Fig. 2.15(c)) is stable (meaning that its geometry and the relative velocities of the plates are unchanging), though the triple junction itself continues to move northwards along the north–south edge of plate A. The point X is originally on the boundary of plates A and B. As the triple junction passes X, an observer there will see a sudden change in subduction rate and direction. Finally, X is a point on the boundary of plates A and C.

In a real situation, the history of the northward passage of the triple junction along the boundary of plates A and C could be determined by estimating the time at which the relative motion between the plates changed at a number of locations along the boundary. If such time estimates increase regularly with position along the plate boundary, it is probable that a triple junction migrated along

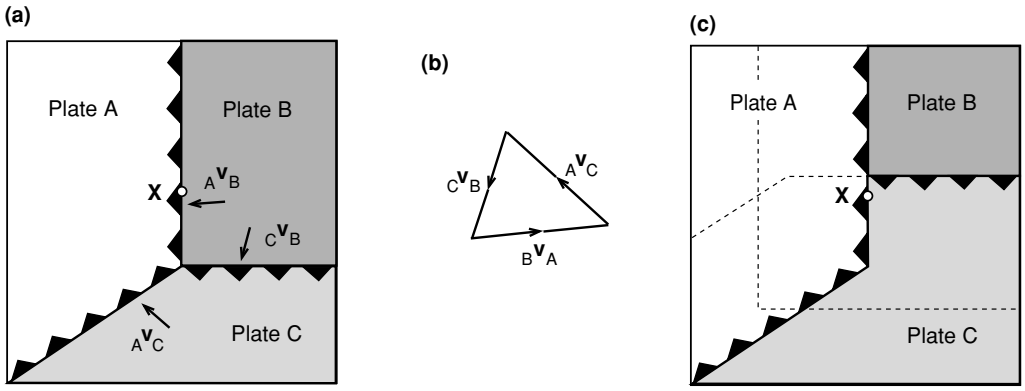
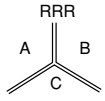
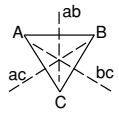
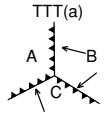
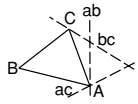
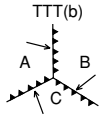
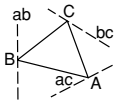
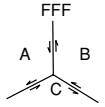
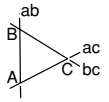
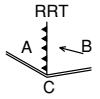
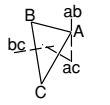
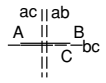
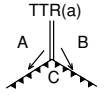
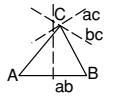
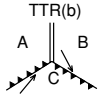
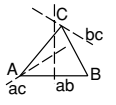


Figure 2.15. (a) A triple junction where three subduction zones intersect. Plate A overrides plates B and C, while plate C overrides plate B. $_A\mathbf{v}_B$, $_C\mathbf{v}_B$ and $_A\mathbf{v}_C$ are the relative velocities of the three plates in the immediate vicinity of the triple junction. (b) The relative-velocity triangle for (a). (c) The geometry of the three subduction zones at some time later than in (a). The dashed lines show where plates B and C would have been had they not been subducted. The point X in (a) was originally on the boundary between plates A and B; now it is on the boundary between plates A and C. The original triple junction has changed its form. (After McKenzie and Morgan (1969).)

the boundary. The alternative, a change in relative motion between the plates, would occur at one time. It can be seen that, although the original triple junction shown in Fig. 2.15(a) is not stable, it would be stable if $_A\mathbf{v}_C$ were parallel to the boundary between plates B and C. Then the boundary between B and C would not move in a north–south direction relative to A, so the geometry of the triple junction would be unchanging with time. The other configuration in which the triple junction would be stable occurs when the edge of the plate either side of the triple junction is straight. This is, of course, the final configuration illustrated in Fig. 2.15(c).

Altogether there are sixteen possible types of triple junction, all shown in Fig. 2.16. Of these sixteen triple junctions, one is always stable (the ridge–ridge–ridge junction) if oblique spreading is not allowed, and two are always unstable (the fault–fault–fault and fault–ridge–ridge junctions). The other thirteen junctions are stable under certain conditions. In the notation used to classify the types of triple junction, a ridge is written as R, a transform fault as F and a subduction zone (or trench) as T. Thus, a ridge–ridge–ridge junction is RRR, a fault–fault–ridge junction is FFR, and so on.

To examine the stability of any particular triple junction, it is easiest to draw the azimuths of the plate boundaries onto the relative velocity triangle. In Fig. 2.16 the lengths of the lines AB, BC and AC are proportional and parallel to the relative velocities $_A\mathbf{v}_B$, $_B\mathbf{v}_C$ and $_A\mathbf{v}_C$. Thus, the triangles are merely velocity triangles such as that shown in Fig. 2.15(b). The triple junction of Fig. 2.15, type

Geometry	Velocity triangle	Stability	Possible Examples
RRR 		All orientations stable	East Pacific Rise and Galapagos Rift Zone, Indian Ocean Triple Junction
TTT(a) 		Stable if ab, ac form a straight line, or if bc is parallel to the slip vector CA	Central Japan
TTT(b) 		Stable if the complicated general condition for ab, bc and ac to meeting at a point is satisfied	
FFF 		Unstable	
RRT 		ab must go through centroid of ABC	
RRF 		Unstable, evolves to FFR; but stable if ab and ac are perpendicular	
TTR(a) 		Stable if ab goes through C, or if ac, bc form a straight line	
TTR(b) 		Stable if complicated general conditions are satisfied	

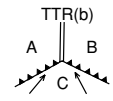
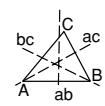
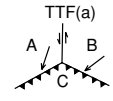
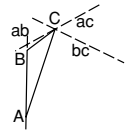
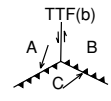
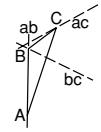
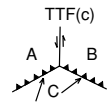
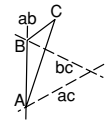
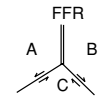
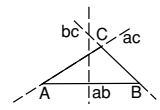
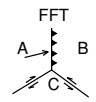
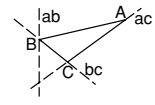
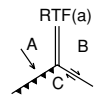
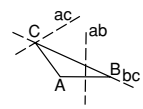
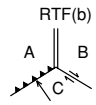
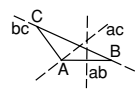
Geometry	Velocity triangle	Stability	Possible Examples
TTR(b) 		Stable if the angles between ab and ac, bc, respectively, are equal, or if ac, bc form a straight line	
TTF(a) 		Stable if ac, bc form a straight line, or if C lies on ab	Intersection of the Peru–Chile trench and the Chile Rise
TTF(b) 		Stable if bc, ab form a straight line, or if ac goes through B	
TTF(c) 		Stable if ab, ac form a straight line, or if ab, bc do so	
FFR 		Stable if C lies on ab, or if ac, bc form a straight line	Owen fracture zone and the Carlsberg Ridge, Chile Rise and the East Pacific Rise
FFT 		Stable if ab, bc form a straight line, or if ac, bc do so	San Andreas Fault and Mendocino fracture zone (Mendocino triple junction)
RTF(a) 		Stable if ab goes through C or if ac, bc form a straight line	Mouth of the Gulf of California (Rivera triple junction)
RTF(b) 		Stable if ac, ab cross on bc	

Figure 2.16. The geometry and stability of all possible triple junctions. In the categories represented by RRR, RTT, RTF and so on, R denotes ridge, T trench and F transform fault. The dashed lines ab, bc and ac in the velocity triangles represent velocities that leave the geometry of the boundary between plates A and B, B and C and A and C, respectively, unchanged. A triple junction is stable if ab, bc and ac meet at a point. Only an RRR triple junction (with ridges spreading symmetrically and perpendicular to their strikes) is always stable. (After McKenzie and Morgan (1969).)

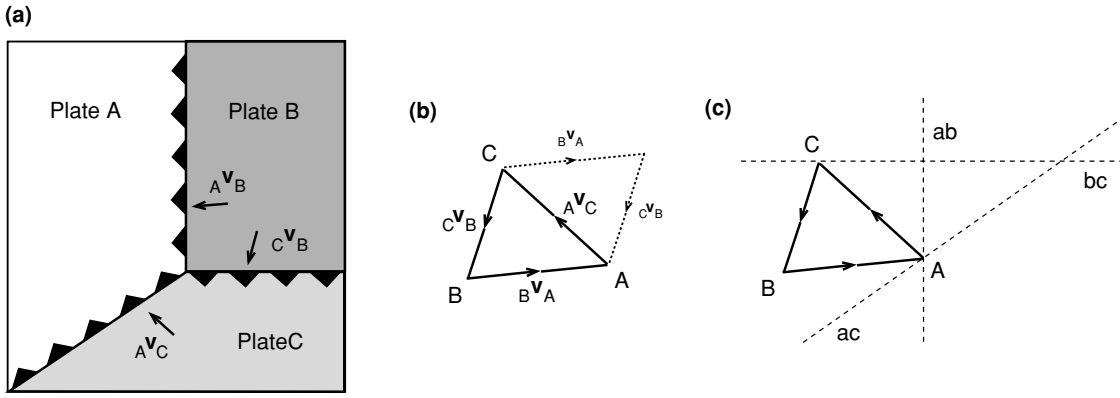


Figure 2.17. Determination of stabilty for a triple junction involving three subduction zones, TTT(a) of Fig. 2.16. (a) The geometry of the triple junction and relative velocities; this is the same example as Fig. 2.15(a). (b) Relative-velocity triangles (Fig. 2.15(b)). Sides BA, CB and AC represent $_B\mathbf{v}_A$, $_C\mathbf{v}_B$ and $_A\mathbf{v}_C$, respectively. The corner C represents the velocity of plate C. Thus for example, relative to plate A, the velocity of plate C, $_A\mathbf{v}_C$, is represented by the line from point A to point C, and the line from point B to point A represents $_B\mathbf{v}_A$, the velocity of plate A relative to plate B. (c) The dashed lines ab, ac, and bc drawn onto the velocity triangle ABC represent possible velocities of the boundary between plates A and B, plates A and C and plates B and C, respectively, which leave the geometry of those boundaries unchanged. The triple junction is stable if these three dashed lines intersect at a point. In this example, that would occur if ab were parallel to ac or if the velocity $_A\mathbf{v}_C$ were parallel to bc. If the geometry of the plate boundaries and the relative velocities at the triple junction do not satisfy either of these conditions, then the triple junction is unstable. If it is unstable, the geometry of the plate boundaries will change; this particular geometry and triple junction can exist only momentarily in geological time.

TTT(a) in Fig. 2.16, is shown in Fig. 2.17. The subduction zone between plates A and B does not move relative to plate A because plate A is overriding plate B. However, because all parts of the subduction zone look alike, any motion of the subduction zone parallel to itself would also satisfy this condition. Therefore, we can draw onto the velocity triangle a dashed line ab, which passes through point A and has the strike of the boundary between plates A and B (Fig. 2.17(c)). This line represents the possible velocities of the boundary between plates A and B which leave the geometry of these two plates unchanged. Similarly, we can draw a line bc that has the strike of the boundary between plates B and C and passes through point C (since the subduction zone is fixed on plate C) and a line ac that passes through point A and has the strike of the boundary between plates A and C. The point at which the three dashed lines ab, bc and ac meet represents the velocity of a stable triple junction. Clearly, in Fig. 2.17(c), these three lines do not meet at a point; therefore, this particular plate-boundary configuration is unstable. However, the three dashed lines would meet at a point (and the triple

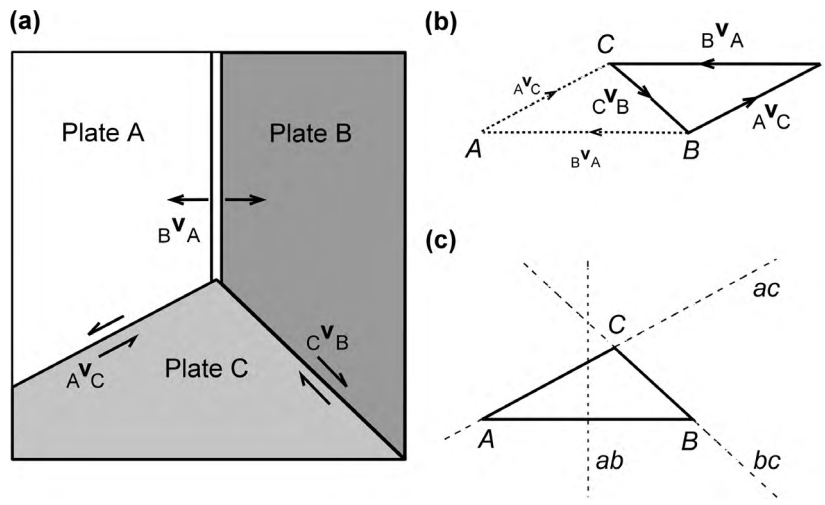


Figure 2.18. Determination of stability for a triple junction involving a ridge and two transform faults, FFR, (a) The geometry of the triple junction and relative velocities. (b) Relative- velocity triangles (notation as for Fig. 2.17(b)). (c) Dashed lines drawn onto the relative-velocity triangle as for Fig. 2.17(c). This triple junction is stable if point C lies on the line ab, or if ac and bc are collinear. The first condition is satisfied if the triangle ABC is isosceles (i.e., the two transform faults are mirror images of each other). The second condition is satisfied if the boundary of plate C with plates A and B is straight.

junction would be stable) if bc were parallel to AC or if ab and ac were parallel. This would mean that either the relative velocity between plates A and C, $_A\mathbf{v}_C$, was parallel to the boundary between plates B and C or the entire boundary of plate A was straight. These are the only two possible situations in which such a TTT triple junction is stable. By plotting the lines ab, bc and ac onto the relative-velocity triangle, we can obtain the stability conditions more easily than we did in Fig. 2.15.

Figure 2.18 illustrates the procedure for a triple junction involving a ridge and two transform faults (type FFR). In this case, ab, the line representing any motion of the triple junction along the ridge, must be the perpendicular bisector of AB. In addition bc and ac, representing motion along the faults, are collinear with BC and AC. This type of triple junction is stable only if line ab goes through point C (both transform faults have the same slip rate) or if ac and bc are collinear (the boundary of plate C is straight). Choosing ab to be the perpendicular bisector of AB assumes that the ridge is spreading symmetrically and at right angles to its strike. This is usually the case. However, if the ridge does not spread symmetrically and/or at right angles to its strike, then ab must be drawn accordingly, and the stability conditions are different.

Figure 2.16 gives the conditions for stability of the various types of triple junction and also gives examples of some of the triple junctions occurring around the Earth at present. Many research papers discuss the stability or instability of the Mendocino and Queen Charlotte triple junctions that lie off western North America. These are the junctions (at the south and north ends, respectively) of the Juan de Fuca plate with the Pacific and North American plates and so are subjects of particular interest to North Americans because they involve the San Andreas Fault in California and the Queen Charlotte Fault in British Columbia. Another triple junction in that part of the Pacific is the Galapagos triple junction, where the Pacific, Cocos and Nazca plates meet; it is an RRR junction and thus is stable.

2.6.2 The significance of triple junctions

Work on the Mendocino triple junction, at which the Juan de Fuca, Pacific and North American plates meet at the northern end of the San Andreas Fault, shows why the stability of triple junctions is important for continental geology. The Mendocino triple junction is an FFT junction involving the San Andreas Fault, the Mendocino transform fault and the Cascade subduction zone. It is stable, as seen in Fig. 2.16, provided that the San Andreas Fault and the Cascade subduction zone are collinear. It has, however, been suggested that the Cascade subduction zone is after all not exactly collinear with the San Andreas Fault and, thus, that the Mendocino triple junction is unstable. This instability would result in the northwards migration of the triple junction and the internal deformation of the continental crust of the western U.S.A. along pre-existing zones of weakness. It would also explain many features such as the clockwise rotation of major blocks, such as the Sierra Nevada, and the regional extension and eastward stepping of the San Andreas transform. The details of the geometry of this triple junction are obviously of great importance to the regional evolution of the entire western U.S.A. Much of the geological history of the area over approximately the past thirty million years may be related to the migration of the triple junction, so a detailed knowledge of the plate motions is essential background for any explanation of the origin of Tertiary structures in this region. This subject is discussed further in Section 3.3.3. The motions of offshore plates can produce major structural changes even in the continents.

The Dead Sea Fault is similar to the San Andreas Fault system in that it is an intra-continental plate boundary. It is the boundary between the Arabian and African plates and extends northwards from the Red Sea to the East Anatolian Fault (Fig. 10.18). It is a left-lateral strike–slip fault with a slip rate of $\sim$5 mm yr^{-1}. That such a major strike–slip boundary is located close to the continental edge, but still within the continent, is because that is where the plate is weakest – a thinned continental margin is weaker than both oceanic and continental lithosphere. The

NUVEL-1A rotation pole for Africa–Arabia is at 24°N, but, in order for the motion along the boundary to be strike–slip, either the pole should be some 6° further north, or Sinai must be moving separately relative to Africa.

2.7 Absolute plate motions

Although most of the volcanism on the Earth's surface is associated with the boundaries of plates, along the mid-ocean ridges and subduction zones, some isolated volcanic island chains occur in the oceans (Fig. 2.19(a)). These chains of oceanic islands are unusual in several respects: they occur well away from the plate boundaries (i.e., they are *intraplate volcanoes*); the chemistry of the erupted lavas is significantly different from that of both mid-ocean-ridge and subduction-zone lavas; the active volcano may be at one end of the island chain, with the islands ageing with distance from that active volcano; and the island chains appear to be arcs of small circles. These features, taken together, are consistent with the volcanic islands having formed as the plate moved over what is colloquially called a *hotspot*, a place where melt rises from deep in the mantle. Figure 2.19(b) shows four volcanic island and seamount chains in the Pacific Ocean. There is an active volcano at the southeastern end of each of the island chains. The Emperor–Hawaiian seamount chain is the best defined and most studied. The ages of the seamounts increase steadily from Loihi (the youngest and at present active) a submarine volcano off the southeast coast of the main island of Hawaii, northwestwards through the Hawaiian chain. There is a pronounced change in strike of the chain where the volcanic rocks are about forty-three million years (43 Ma) old. The northern end of the Emperor seamount chain near the Kamchatka peninsula of Russia is 78 Ma old. The change in strike at 43 Ma can most simply be explained by a change in the direction of movement of the Pacific plate over the hotspot at that time. The chemistry of oceanic-island lavas is discussed in Section 7.8.3 and the structure of the islands themselves in Section 9.7.

All the plate motions described so far in this chapter have been relative motions, that is, motions of the Pacific plate relative to the North American plate, the African plate relative to the Eurasian plate, and so on. There is no fixed point on the Earth's surface. Absolute plate motions are motions of the plates relative to some imaginary fixed point. One way of determining absolute motions is to suppose that the Earth's mantle moves much more slowly than the plates so that it can be regarded as nearly fixed. Such absolute motions can be calculated from the traces of the oceanic island chains or the traces of continental volcanism, which are assumed to have formed as the plate passed over a hotspot with its source fixed in the mantle. The absolute motion of a plate, the Pacific, for example, can be calculated from the traces of the oceanic island and seamount chains on it. The absolute motions of all the other plates can then be calculated from their motion relative to the Pacific plate. Repeating the procedure, using hotspot traces from other plates, gives some idea of the validity of the assumption that hotspots are fixed.

(a)

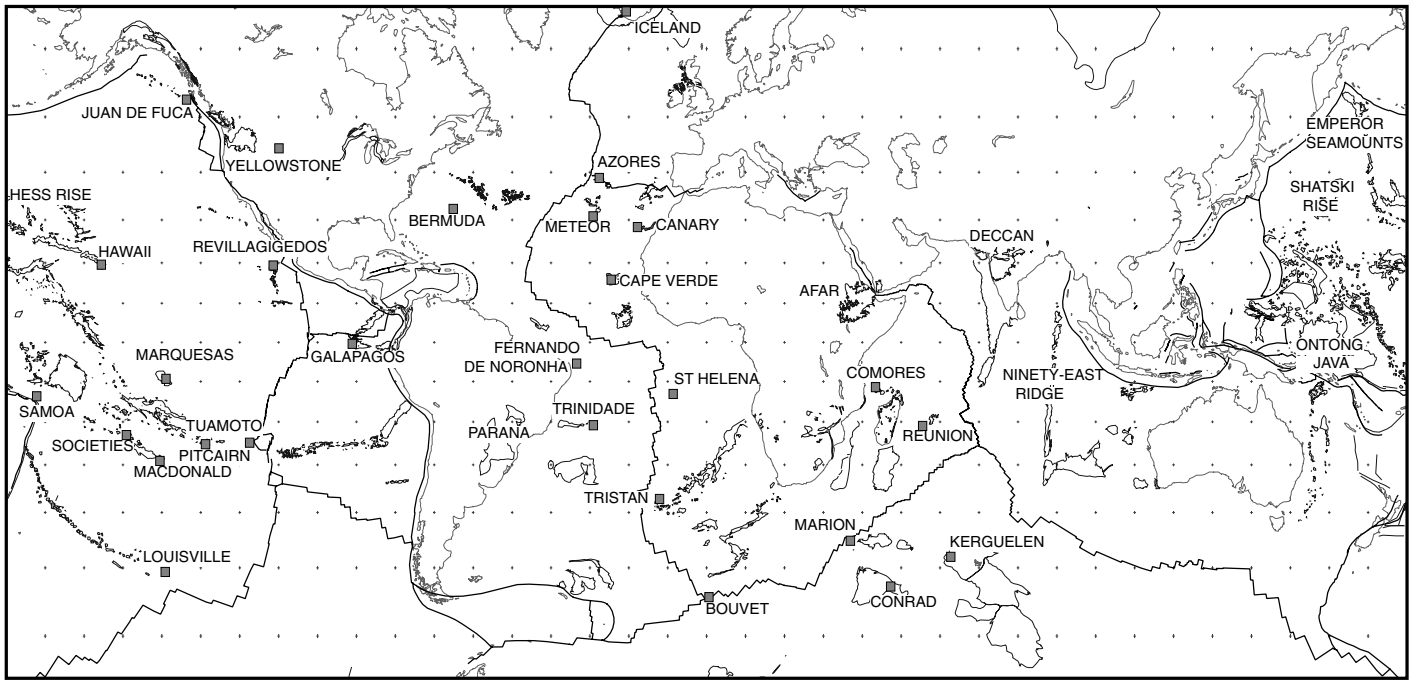


Figure 2.19. (a) The global distribution of hotspots (grey squares) and associated volcanic tracks. (After Norton, I. O. Global hotspot reference frames and plate motion. Geophysical Monograph 121, 339–57, 2000. Copyright 2000 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

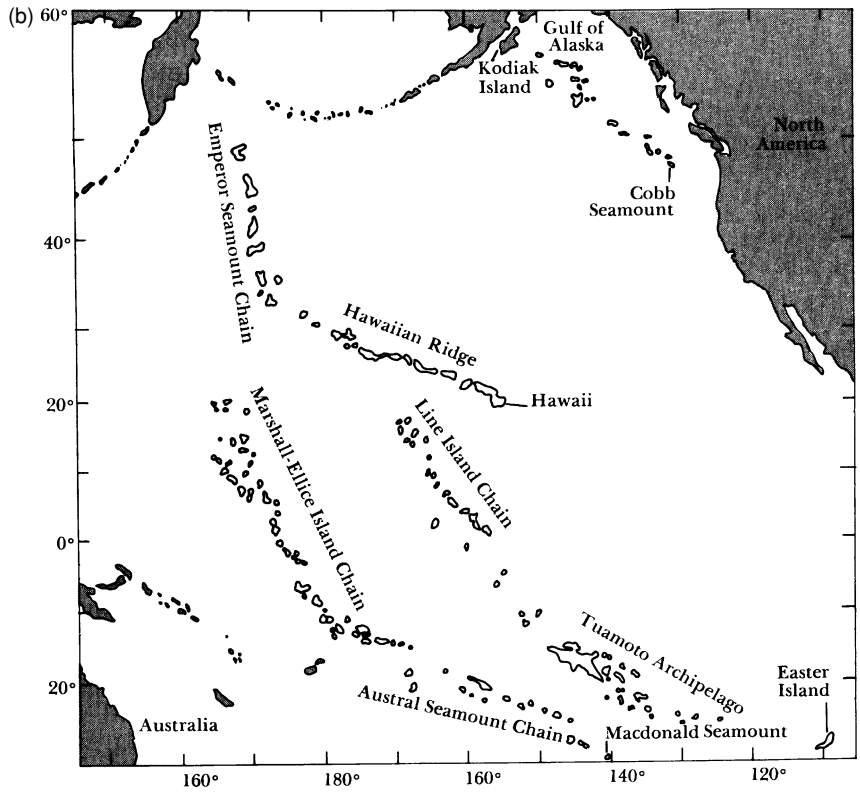


Figure 2.19. (b) Four volcanic island chains in the Pacific Ocean. The youngest active volcano is at the southeast end of each chain. (From Dalrymple *et al.* (1973).)

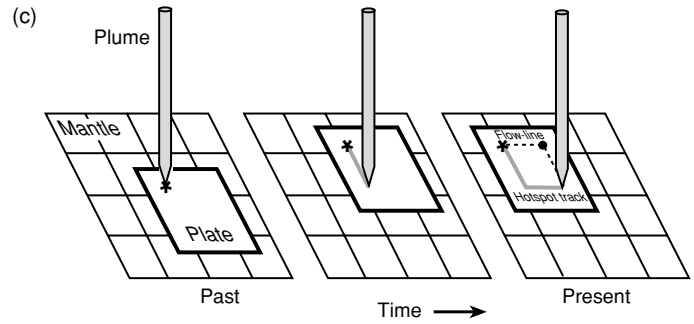


Figure 2.19. (c) A demonstration of the relative motions between the hotspot (fixed in the mantle) and the seamount chain on the overriding plate. The pencil represents the hotspot, which is fixed in the mantle (rectangular grid). The plate moves over the mantle and the pencil marks the line of seamounts (the 'hotspot track'). The star is the position of a seamount. After formation of the seamount, the plate moves north for one unit and then west for one unit so leaving a solid (pencil) line of seamounts, the hotspot track. The 'flow-line', the relative motion of the plate with respect to the hotspot, is the dashed line. (Reprinted with permission from *Nature* (Stein *Nature* **387**, 345–6) Copyright 1997 Macmillan Magzines Ltd.)

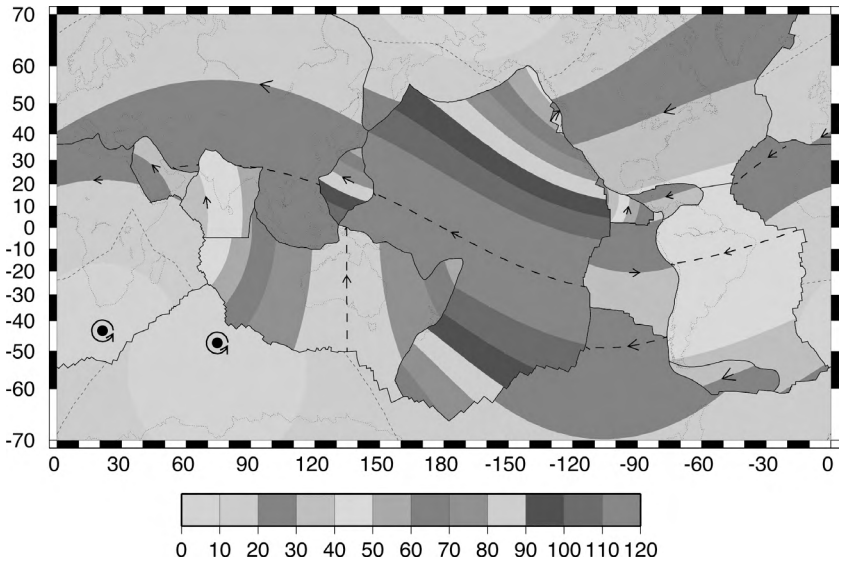


Figure 2.20. Absolute motions of the plates as determined from hotspot traces. HS3-NUVEL-1A is a set of angular velocities of fifteen plates relative to the hotspots. The hotspot dataset HS3 averages plate motion over the last 5.8 Ma. No hotspots are in significant relative motion. The 95% confidence limit is ±20–40 km/Ma^{-1} but can be 145 km/Ma^{-1}. (See Plate 1 for colour version). (From Gripp and Gordon (2002).)

The hotspot reference frame therefore is the motions of the plates relative to the hotspots, which are assumed to be fixed in the mantle. The 'hotspot track', the linear chain of volcanic islands and seamounts, is the path of the hotspot with respect to the overlying plate. The slow steady motion of the oceanic plate with respect to the hotspot (the flowline) is not marked by any feature, however. Figure 2.19(c) illustrates the difference between the hotspot track and the flowline. In order to use probable hotspot tracks to run the plate motions backwards, it is necessary to know the ages of the islands and seamounts, as we do for the Hawaiian chain. However, even if the ages of seamounts are unknown, the flow-lines can be used. By assigning a range of possible ages to any seamount, a series of flow-lines can be plotted (these will all follow the actual flow-line), producing a line that will go through the present position of the hotspot. (For example, when the seamount in Fig. 2.19(c) is erroneously assumed to be one unit old rather than two, the black dot shows the backtracked hotspot location – this lies on the flow-line but not on the hotspot track.) The process is then repeated for another seamount and the two sets of flow-lines compared. If the two seamounts formed at the same hotspot then the flow-lines will intersect at the location of that hotspot. If the seamounts were not formed by the same hotspot, the flow-lines should not intersect. Application of this method is improving our knowledge of hotspot locations and absolute plate motions.

Figure 2.20 shows a determination of the present absolute plate motions. Plate motions relative to hotspots cannot be estimated as accurately as can relative

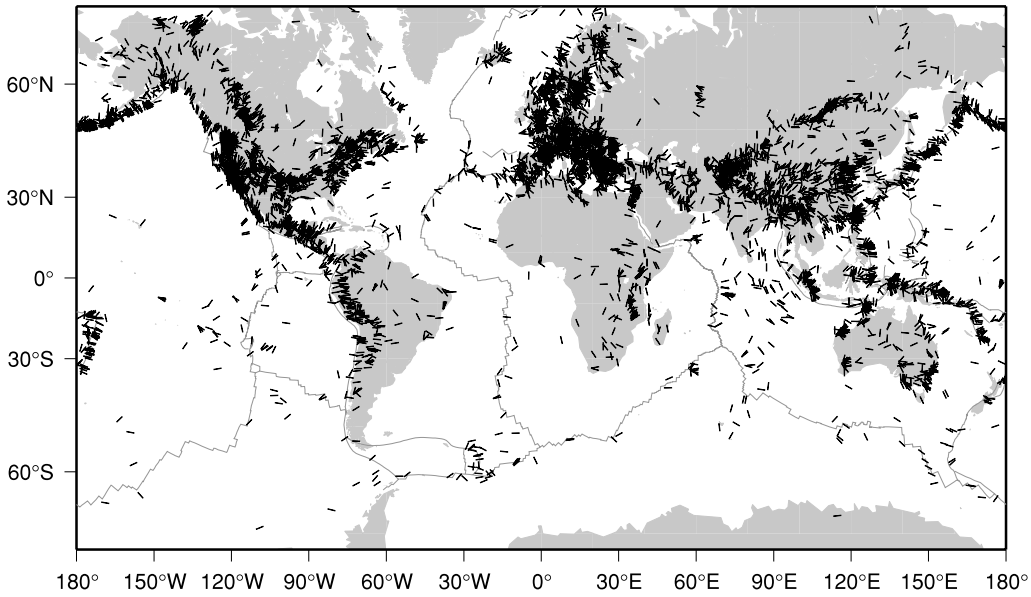


Figure 2.21. A generalised world stress map. Lines show the orientation of the maximum horizontal stress. A colour version of the map showing the tectonic regimes – normal faulting, strike–slip faulting and thrust faulting – is available on-line. (From Reinecker *et al.* 2003, available on-line at http://www.world-stress.org.)

plate motions. This is because hotspot tracks have average widths in excess of 100 km, which is orders of magnitude greater than the width of active transform faults (<2 km) and the width of the transition between magnetic anomalies (1–5 km). Continental plates tend to move more slowly than oceanic plates. Similarly, plates with more than a quarter of their perimeter subducting tend to move faster than plates that are hardly subducting anywhere. Values for the absolute motions obtained by different authors vary somewhat depending upon assumptions made and the particular relative plate motions used. Sources of uncertainty in determining absolute motions are the determination of the motions of the hotspot traces and assumptions about mantle dynamics and forces driving the plates. Nevertheless, a framework of motion can be developed even if some uncertainties remain. The Pacific and Indian plates are moving fast and the North American plate more slowly, while the Eurasian and African plates are hardly moving at all.

Figure 2.21 is a plot of the orientation of horizontal stress measured in the interior of the plates. The pattern of these intraplate stresses in the crust (stress is force per unit area) can be used to assess the forces acting on the plates. The direction of the stresses is correlated with the direction of the absolute plate motions (Fig. 2.20). This implies that the forces moving the plates around, which act along the edges of the plates (Section 7.10) are, to first order, responsible for the stresses in the lithosphere.

Problems

1. All plates A–D shown in Fig. 2.22 move rigidly without rotation. All ridges add at equal rates to the plates on either side of them; the rates given on the diagram are half the plate-separation rates. The trench forming the boundary of plate A does not consume A.

   Use the plate velocities and directions to determine by graphical means or otherwise (a) the relative motion between plates B and D and (b) the relative motion between the triple junction J and plate A. Where and when will J reach the trench? Draw a sketch of the geometry after this collision showing the relative velocity vectors and discuss the subsequent evolution. (From Cambridge University Natural Sciences Tripos 1B, 1982.)

2. Long ago and far away in a distant galaxy lived two tribes on a *flat* planet known as Emit-on (Fig. 2.23). The tribe of the dark forces lived on the islands in the cold parts of the planet near A, and a happy light-hearted people lived on the sunny beaches near B. Both required a constant supply of fresh andesite to survive and collected the shrimp-like animals that prospered near young magnetic-anomaly patterns, for their food. Using the map and your knowledge of any other galaxy, predict the future of these tribes. (From Cambridge University Natural Sciences Tripos 1B, 1980.)

3. Use the map in Fig. 2.2 to do the following.
   (a) Calculate the relative present-day motions at the tabulated locations, using or calculating the appropriate poles from those given in Table 2.1.
   (b) Plot the azimuth and magnitude of these relative motions on the map.
   (c) Plot the pole positions. Note how the relative motions change along plate boundaries as the angular distance from the pole changes.
   (d) Discuss the nature of the plate boundary between the Indian/Australian and Pacific plates.
   (e) Discuss the nature of the plate boundary between the Eurasian and African plates between the Azores and Gibraltar. How does this boundary differ in the Mediterranean?

   (From G. C. P. King, personal communication.)

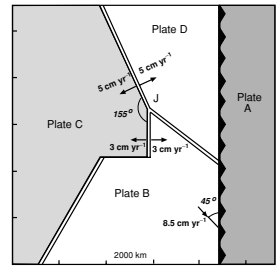


Figure 2.22. A map of part of a flat planet.

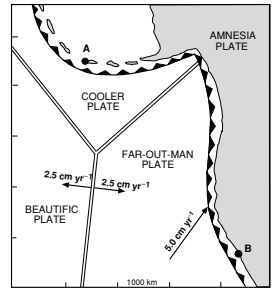


Figure 2.23. A map of part of Emit-on.

Latitude	Longitude	Location
54°N	169°E	W. Aleutian Trench
52°N	169°W	E. Aleutian Trench
38°N	122°W	San Francisco–San Andreas Fault
26°N	110°W	Gulf of California
13°S	112°W	East Pacific Rise
36°S	110°W	East Pacific Rise
59°S	150°W	Antarctic–Pacific Ridge
45°S	169°E	S. New Zealand
55°S	159°E	Macquarrie Island
52°S	140°E	Southeast Indian Ridge

(*cont.*)

Latitude	Longitude	Location
28°S	74°E	Southeast Indian Ridge
7°N	60°E	Carlsberg Ridge
22°N	38°E	Red Sea
55°S	5°E	Southwest Indian Ridge
52°S	5°E	Mid-Atlantic Ridge
9°N	40°W	Mid-Atlantic Ridge
35°N	35°W	Mid-Atlantic Ridge
66°N	18°W	Iceland
36°N	8°W	Gorringe Bank
35°N	25°E	E. Mediterranean
12°S	120°E	Java Trench
35°N	72°E	Himalayas
35°S	74°W	S. Chile Trench
4°S	82°W	N. Peru Trench
20°N	106°W	Middle America Trench

4. Define the pole of relative angular velocity for two plates and explain why it is a useful concept. Give two practical methods for finding poles of rotation from geophysical data. When these data are not available, how would you try to locate the poles? (Cambridge University Natural Sciences Tripos IB, 1977.)
5. Ares is a planet with just two plates, A and B (Fig. 2.24). Plate B comprises the lower hemisphere and plate A the upper, as shown. Points a, b and c lie on the equator, and point d is diametrically opposite b. The zero meridian passes through point a. The pole of rotation of plate A relative to plate B is at 45°N, 0°E. The amplitude of the angular velocity vector is 10^{-10} radians per terrestrial year. The radius of Ares is 3400 km.
   (a) What is the nature of the plate boundary between plates A and B?
   (b) State where magnetic lineations might be found and sketch the pattern that would be observed.
   (c) Calculate the relative velocity between plates A and B at locations a, b, c and d.
   (d) Sketch possible fault-plane solutions for earthquakes occurring at locations a, b, c and d.
   (e) Discuss the possible existence of such a two-plate planet.
   (f) Discuss briefly how the stability or instability of a two-plate tectonic system depends upon the pole position and/or relative size of the two plates.
6. The lithosphere of Tritekton, a recently discovered spherical satellite of Jupiter, consists of three plates, as shown in Fig. 2.25. Plate A is a hemisphere, and plates B and C are half a hemisphere each. Points a, b and c lie on its equator, and point d (not visible) lies diametrically opposite b. The zero meridian lies midway between b and d and passes through a. The pole of rotation B to A lies in plate B at latitude 30°N longitude 0°. The amplitude of the angular velocity vector through this pole is 3×10^{-9} radians per terrestrial year. The rotation pole C to B is at the north pole N, and the amplitude of the angular velocity vector is 6×10^{-9} rad yr^{-1}. The radius of Tritekton is 6000 km.

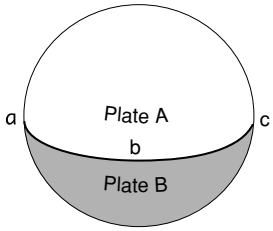


Figure 2.24. Planet Ares has just two plates.

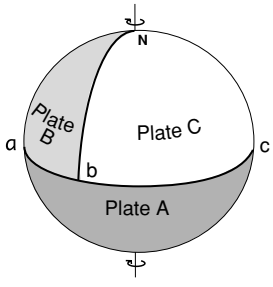


Figure 2.25. Tritekton has three plates.

(a) Find the coordinates of the angular velocity vector of C to A and its amplitude.
(b) Draw velocity triangles for the triple junctions at b and d.
(c) Tritekton has a magnetic field that reverses every ten million years. Draw the magnetic pattern created at the triple junction where extension is taking place. State whether it is at b or d.
(d) The features at each plate boundary are trenches, transform faults or spreading ridges. Describe their distribution along each plate boundary.

(Cambridge University Natural Sciences Tripos IB, 1979.)

7. In Fig. 2.26, the trench between B and C is consuming B only. The ridge between A and B is spreading symmetrically at right angles to its axis. The pole of rotation between A and C is fixed to C. The angular velocity of A with respect to C is in the direction shown, and is 2×10^{-8} rad yr^{-1} in magnitude.
   (a) Mark the poles of rotation and angular velocities of (i) plate B and (ii) the ridge axis with respect to plate C. (Remember that the pole of motion between two plates is that point which is stationary with respect to both of them.)
   (b) Show the direction and rate of consumption at X and Y.
   (c) How long will it take for the ridge to reach (i) X and (ii) Y?
   Sketch the history of the triple junction between A, B and C with respect to C.

   (Cambridge University Natural Sciences Tripos 1B, 1981.)
8. The three plates A, B and C meet at a ridge–ridge–ridge triple junction as shown in Fig. 2.27. The ridge between plates A and B has a half-spreading rate of 2 cm yr^{-1}. Calculate
   (a) the half-spreading rates of the other two ridges and
   (b) the motion of the triple junction relative to plate C.
9. All the plates of Problem 1 move rigidly without rotation. Discuss the difference between those motions and the motions that would take place were the plates on the surface of a spherical Earth. (For example, by how much would the plate-separation rates vary along the length of the ridge between plates B and C?)
10. Four flat plates are moving rigidly, without rotation, on a flat Earth as illustrated in Fig. 2.28.
    (a) Determine the relative motion vector of the Beautific-Joker's–Nasty (BJN) triple junction to the Albatross plate.
    (b) Determine the relative motion vector of the Beautific-Nasty–Erratic (BNE) triple junction to the Albatross Plate.
    (c) Using these vectors, draw the loci of the future positions of these junctions relative to the Albatross plate. Mark points at 10-Ma intervals.
    (d) At 10-Ma intervals redraw the positions of the plate boundaries until both the Joker's plate and the Nasty plate have been consumed at the subduction zone boundary by the Albatross plate.
    (e) Draw velocity triangles for the triple junctions at the times when they pass through points A and B.
    (f) Draw the magnetic-anomaly pattern formed by the BJN junction for 10 Ma, assuming that a recognizable anomaly appears at 2-Ma intervals.

    (From G. C. P. King, personal communication.)

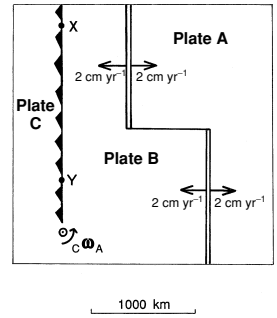


Figure 2.26. The map for Problem 7.

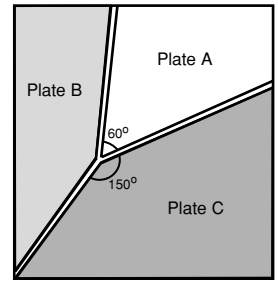


Figure 2.27. A ridge–ridge–ridge (RRR) triple junction.

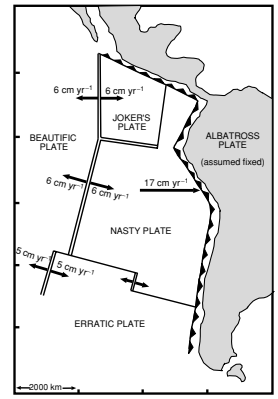


Figure 2.28. Plates on a flat Earth.

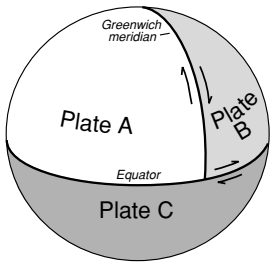


Figure 2.29. A mythical planet has three plates.

11. Figure 2.29 depicts a mythical planet whose tectonic features comprise three plates A, B and C with boundaries as shown. The plate boundary A–B is a fault with sense of motion as shown and relative velocity of 4 cm yr^{-1}, and that between B and C is also a fault, with velocity 3 cm yr^{-1}.
    (a) Using angular velocity vectors, find the pole of relative motion between plates A and C, giving its geographic coordinates. Calculate the relative angular velocity in degrees per year (the radius of the mythical planet is 6000 km).
    (b) Describe the tectonic nature of the boundary between plates A and C. Give details of places where magnetic lineations might be found (the mythical planet has a reversing magnetic field like the Earth's), what the maximum spreading rate might be expected to be, and any other geological and geophysical information you can deduce.
    (c) Consider the triple junction where plates A, B and C all meet on the visible side of the mythical planet. What will happen to the triple junction if this instantaneous velocity description continues for several million years? In particular, think about the way ridges spread symmetrically and trenches consume on only one side. Is the junction likely to persist as drawn? What about the junction on the other side of the mythical planet? Is it similar?

References and bibliography

Altman, S. L. 1986. *Rotations, Quaternions and Double Groups*. New York: Oxford University Press.

Argus, D. and Gordon, R. 1991. No-net rotation model of current plate velocities incorporating plate motion model NUVEL.1. *Geophys. Res. Lett.*, **18**, 2039–42.

Barazangi, M. and Dorman, J. 1969. World seismicity maps compiled from ESSA Coast and Geodetic Survey, epicenter data, 1961–1967. *Bull. Seism. Soc. Am.*, **59**, 369–80.

Båth, M. 1979. *Introduction to Seismology*, 2nd edn. Boston: Birkhauser.

Bullard, E. C., Everett, J. E. and Smith, A. G. 1965. Fit of continents around the Atlantic. In P. M. S. Blackett, E. C. Bullard and S. K. Runcorn, eds., *A Symposium on Continental Drift*, *Roy. Soc. London, Phil. Trans. Ser. A*, **258**, 41–75.

Carter, W. E. and Robertson, D. S. 1986. Studying the Earth by very-long-baseline-interferometry. *Sci. Am.*, **255**, 46–54.

Chase, C. G. 1978. Plate kinematics: the Americas, East Africa and the rest of the world. *Earth Planet. Sci. Lett.*, **37**, 355–68.

Chu, D. H. and Gordon, R. G. 1999. Evidence for motion between Nubia and Somalia along the southwest Indian ridge. *Nature*, **398**, 64–7.

Coode, A. M. 1965. A note on oceanic transcurrent faults. *Can. J. Earth Sci.*, **2**, 400–1.

Cox, A., ed. 1973. *Plate Tectonics and Geomagnetic Reversals*. San Fransisco, California: Freeman.

Cox, A. and Hart, R. B. 1986. *Plate Tectonics: How it Works*. Oxford: Blackwell Scientific.

Dalrymple, G. B., Silver, E. A. and Jackson, E. D. 1973. Origin of the Hawaiian islands. *Am. Scientist*, **61**, 294–308.

DeMets, C., Gordon, R. G., Argus, D. F. and Stein, S. 1990. Current plate motions. *Geophys. J. Int.*, **101**, 425–78.

1994. Effects of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. *Geophys. Res. Lett.*, **21**, 2191–4.

Dickinson, W. R. and Snyder, W. S. 1979. Geometry of triple junctions related to San Andreas transform. *J. Geophys. Res.*, **84**, 561–72.

Gordon, R. G. and Jurdy, D. M. 1986. Cenozoic global plate motions. *J. Geophys. Res.*, **91**, 12 389–406.

Gripp, A. E. and Gordon, R. G. 1990. Current plate velocities relative to the hotspots incorporating the NUVEL-1 global plate motion. *Geophys. Res. Lett.*, **17**, 1109–12.

2002. Young tracks of hotspots and current plate velocities. *Geophys. J. Int.* **150**, 321–61.

Gubbins, D. 1990. *Seismology and Plate Tectonics*. Cambridge: Cambridge University Press.

Harper, J. F. 1986. Mantle flow and plate motions. *Geophys. J. Roy. Astr. Soc.*, **87**, 155–71.

Irving, E. 1977. Drift of the major continental blocks since the Devonian. *Nature*, **270**, 304–9.

Kroger, P. M., Lyzenga, G. A., Wallace, K. S. and Davidson, J. M. 1987. Tectonic motion in western United States inferred from very long baseline interferometry measurements, 1980–1986. *J. Geophys. Res.*, **92**, 14 151–63.

Larson, K. M., Freymueller, J. T. and Philipsen, S. 1997. Global plate velocities from the Global Positioning System. *J. Geophys. Res.*, **102**, 9961–81.

Le Pichon, X., Francheteau, J. and Bonnin, J. 1973. *Plate Tectonics*. Vol. 6 of Developments in Geotectonics. Amsterdam: Elsevier.

Lowman, P. D. 1992. Geophysics from orbit: the unexpected surprise, *Endeavour*, **16**, (2), 50–8.

McKenzie, D. P. and Morgan, W. J. 1969. Evolution of triple junctions. *Nature*, **224**, 125–33.

McKenzie, D. P. and Parker, R. L. 1967. The north Pacific: an example of tectonics on a sphere. *Nature*, **216**, 1276–80.

Minster, J. B. and Jordan, T. H. 1978. Present-day plate motions. *J. Geophys. Res.*, **83**, 5331–54.

Molnar, P. and Stock, J. 1987. Relative motions of hotspots in the Pacific, Atlantic and Indian Oceans since Late Cretaceous time. *Nature*, **327**, 587–91.

Morgan, W. J. 1968. Rises, trenches, great faults and crustal blocks. *J. Geophys. Res.*, **73**, 1959–82.

1971. Convection plumes in the lower mantle. *Nature*, **230**, 42–3.

Norton, I. O. 2000. Global hotspot reference frames and plate motion. In M. A. Richards, R. G. Gordon and R. D. van der Hilst, eds., *The History and Dynamics of Global Plate Motions*. Geophysical Monograph 121. Washington: American Geophysical Union, pp. 339–57.

Reinecker, J., Heidbach, O. and Müller, B. 2003. The 2003 release of the World Stress Map (http://www.world-stress.org).

Royer, J.-Y. and Gordon, R. G. 1997. The motion and boundary between the Capricorn and Antarctic plates. *Science*, **277**, 1268–74.

Shen *et al.* 1997. Crustal deformation measured in southern California. *EOS Trans. Am. Geophys. Un.*, **78** (43), 477 and 482.

Stein, R. S. 1987. Contemporary plate motion and crustal deformation. *Rev. Geophys.*, **25**, 855–63.

1988. Plate tectonic prediction fulfilled. *Phys. Today* (January), S42–4.

Stein, S. A. 1997. Hot-spotting in the Pacific. *Nature*, **387**, 345–6.

Steinberger, B., Sutherland, R. and O'Connell, R. J. 2004. Prediction of Emperor-Hawaii seamount locations from a revised model of global plate motion and mantle flow. *Nature*, **430**, 167–73.

Stella, G. F., Dixon, T. H. and Mao, A. 2002. REVEL: a model for recent plate velocities from space geodesy. *J. Geophys. Res.* **107** (B4), doi: 10.1029/200JB000033.

Sykes, L. R. 1967. Mechanism of earthquakes and nature of faulting on the mid-ocean ridges. *J. Geophys. Res.*, **72**, 2131–53.

Wessel, P. and Kronke, L. 1997. A geometric technique for relocating hotspots and redefining absolute plate motions. *Nature*, **387**, 365–9.

Wilson, J. T. 1965. A new class of faults and their bearing on continental drift. *Nature*, **207**, 343–7.

Zatman, S., Gordon, R. G. and Richards, M. A. 2001. Analytic models for the dynamics of diffuse oceanic plate boundaries. *Geophys. J. Int.*, **145**, 145–56.

Ziegler, A. M., Scotese, C. R., McKerrow, W. S., Johnson, M. E. and Bambach, R. K. 1979. Paleozoic paleogeography. *Ann. Rev. Earth Planet. Sci.*, **7**, 473–502.

Zoback, M. L. 1992. First- and second-order patterns of stress in the lithosphere: the World Stress Map Project. *J. Geophys. Res.*, **97**, 11 703–28.

Zoback, M. L. *et al.* 1989. Global patterns of tectonic stress: a status report on the World Stress Map Project of the International Lithosphere Program. *Nature*, **341**, 291–8.

Zoback, M. D. and Zoback, L. M. 1991. Tectonic stress field of North America and relative plate motions. In D. B. Slemmons, E. R. Engdahl, M. D. Zoback and D. D. Blackwell, eds., *Neotectonics of North America*. Geological Society of America Decade Map Vol. 1. Boulder, Colorado: Geological Society of America, pp. 339–66.

Chapter 3
Past plate motions

3.1 The role of the Earth's magnetic field

3.1.1 Introduction

It should be clear from the preceding chapter that it is possible, without too much difficulty, to calculate the relative motions of pairs of plates at any location along their common boundary and to see what may occur in the future. This chapter deals with the past motions of the plates and shows how to reconstruct their previous interactions from evidence they have left.

Two important facts together make it possible to determine past plate motions. The first is that the Earth's magnetic field has not always had its present (normal) polarity with the 'north' magnetic pole close to the north geographic pole[1] and the 'south' magnetic pole close to the south geographic pole. Over geological history the magnetic field has intermittently reversed. Thus, there have been times in the past when the north magnetic pole has been located close to the present-day south geographic pole and the south magnetic pole has been located close to the present-day north geographic pole; then the field is said to be reversed. The second fact is that, under certain circumstances (discussed in Section 3.2) rocks can record the Earth's past (*palaeo*) magnetic field. Together, these facts enable us to estimate dates and past positions of the plates from magnetic measurements.

3.1.2 The Earth's magnetic field

To specify the geomagnetic field at any point on the Earth's surface both a magnitude and a direction are required: the geomagnetic field is a vector quantity. It is far from being constant either in magnitude or in direction and varies spatially over the surface of the Earth as well as in time. Systematic mapping of the magnetic field began some five hundred years ago with the magnetic compass measurements carried out by the early mariners. The internationally agreed values of the geomagnetic field are updated and published every few years as the *International Geomagnetic Reference Field* (IGRF). Figure 3.1(a) shows the

[1] The geomagnetic pole at present situated in the northern hemisphere is in fact a south pole since it attracts the north poles of magnets (compass needles)! See Fig. 3.2.

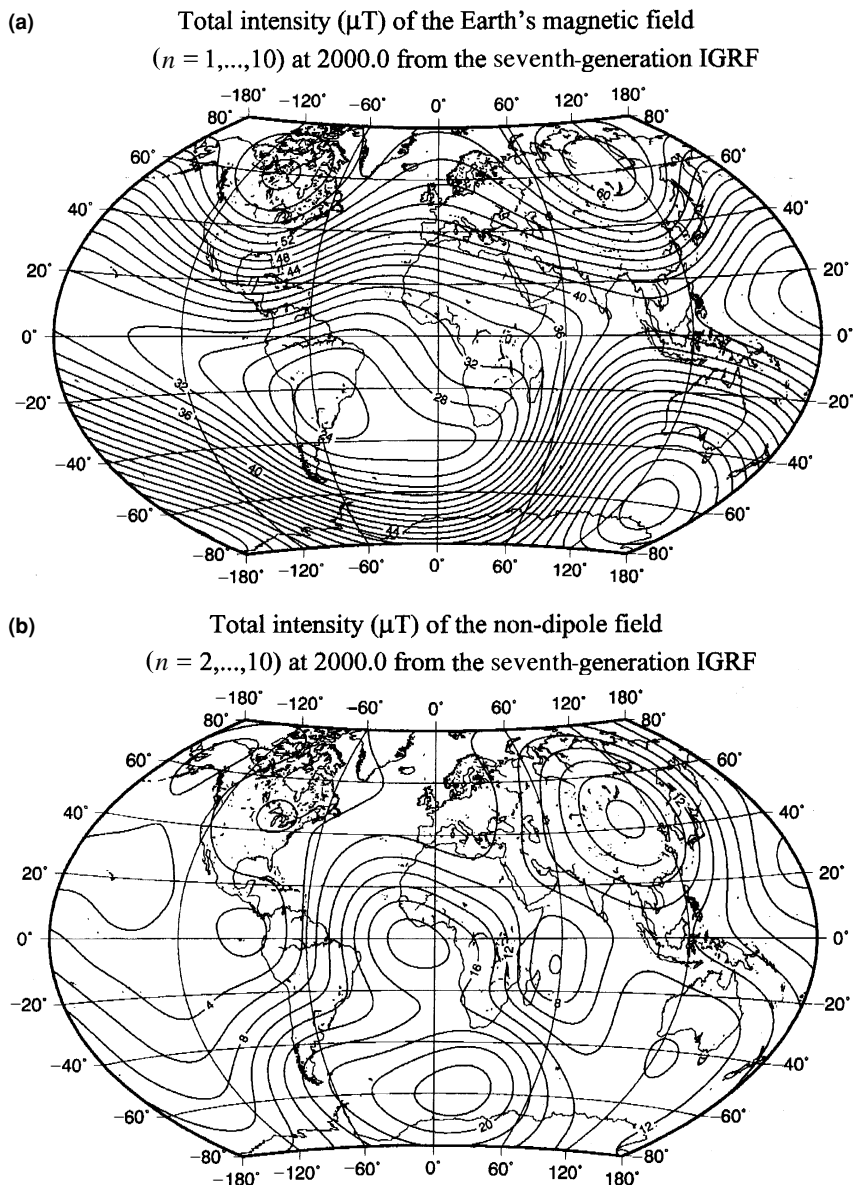


Figure 3.1. (a) The total intensity of the Earth's magnetic field in microteslas (μT). (b) The total intensity of the non-dipole component of the Earth's magnetic field (magnetic field – best-fitting dipole field) in microteslas (μT). The field is the International Geomagnetic Reference Field (IGRF seventh generation). The nanotesla (1 nT = 10^{-3} μT), an SI unit, is the same as gamma (γ), the magnetic field unit of the old c.g.s. system.

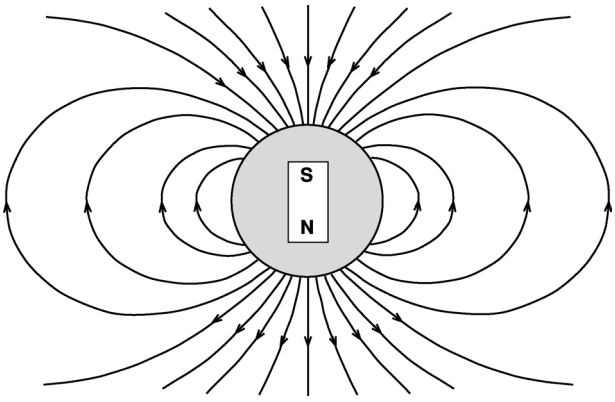


Figure 3.2. A magnetic dipole (e.g., a bar magnet) at the centre of the Earth has a magnetic field that is a good first approximation to the Earth's magnetic field. Note that the geomagnetic North Pole is so called because the north end of the compass needle points towards it (it is therefore actually a south magnetic pole). Such a dipole field would also be produced by a uniformly magnetized Earth, a uniformly magnetized core or particular current systems within the core (Fig. 8.23). The origin of the field and models of the core and possible reasons for the reversals of the Earth's magnetic field are discussed further in Section 8.3.2.

magnitude of the IGRF for 1995. To a first approximation the Earth's magnetic field is a *dipole field*.[2] This means that the Earth's magnetic field can be represented by a magnetic dipole situated at the centre of the Earth (imagine a bar magnet at the centre of the Earth, Fig. 3.2). This fact was first pointed out in 1600 by Sir William Gilbert.[3] The difference between the Earth's magnetic field and the best dipole field is termed the *non-dipole field*. Figure 3.1(b) shows the magnitude of the non-dipole field. At the Earth's surface the non-dipole field is small compared with the dipole field, though this is not the case for the field at the core–mantle boundary.

Today, the best-fitting dipole is aligned at about 11.5° to the Earth's geographic north–south axis (spin axis). The geomagnetic poles are the two points where the axis of this best-fitting dipole intersects the Earth's surface. Now at 79°N, 71°W and 79°S, 109°E, they are called the *geomagnetic north* and g*eomagnetic south poles*, respectively. The *geomagnetic equator* is the equator of the best-fitting dipole axis. The two points on the Earth's surface at which the magnetic field is vertical and has no horizontal component are called the *magnetic poles*, or *dip poles*. The present north magnetic pole is at 76°N, 101°W, and the south magnetic pole is at 66°S, 141°E. The *magnetic equator* is the line along which the magnetic field is horizontal and has no vertical component. If the field were

[2] The magnetic field produced by a bar magnet is a dipole field.

[3] William Gilbert was a distinguished physician. He was President of the College of Physicians and physician to Queen Elizabeth I and King James I. His book *De Magnete* is in print in translation.

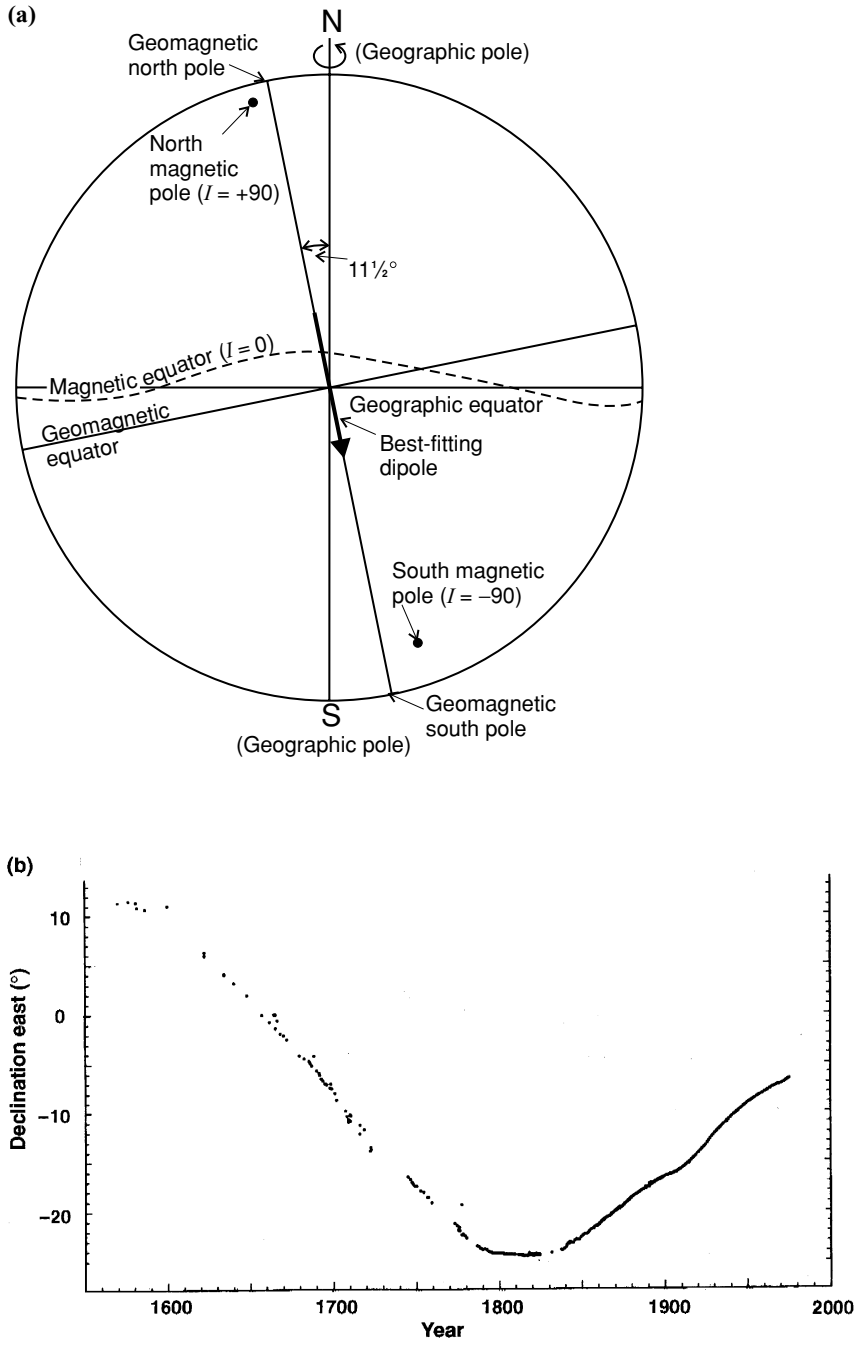


Figure 3.3. (a) A diagram illustrating the difference between the Earth's geographic, geomagnetic and magnetic poles and equator. (From McElhinny (1973).) (b) The change in magnetic declination as observed from London over four centuries. (From Malin and Bullard (1981).)

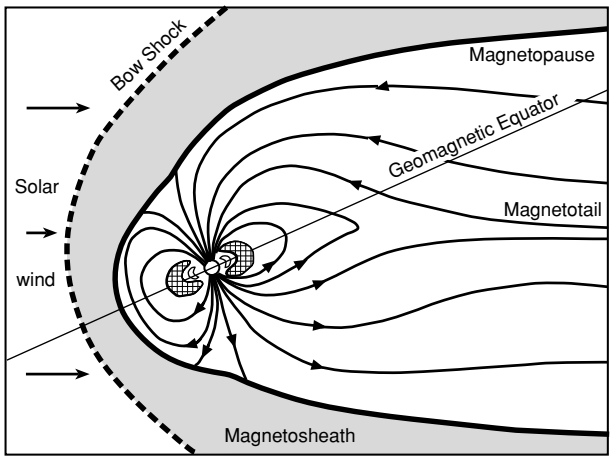


Figure 3.4. The Earth's magnetic field far from the Earth is controlled largely by the solar wind. The Van Allen radiation belts are the cross-hatched regions either side of the Earth. (After Hutton (1976).)

exactly a dipole field, these magnetic poles and this equator would be coincident with the geomagnetic north and south poles and equator. The various poles and equators are illustrated in Fig. 3.3.

Figure 3.4 shows the magnetic field lines around the Earth. The Sun plays the major part in the shape of the field far from the Earth and in the short-term variations of the field. The *solar wind*, a constant stream of ionized particles emitted by the Sun, confines the Earth's magnetic field to a region known as the *magnetosphere* and deforms the field lines so that the magnetosphere has a long 'tail', the *magnetotail*, which extends several million kilometres away from the Sun. A shock wave, the *bow shock*, is produced where the solar wind is slowed by interaction with the Earth's magnetic field. A turbulent zone within the bow shock is known as the *magnetosheath*, the inner boundary of which is called the *magnetopause*. The Earth's magnetic field shields the Earth from most of the incident radiation, and the atmosphere absorbs much of the remainder. Major sunspot activity causes changes in the solar wind, which in turn result in short-term fluctuations in the magnetic field known as *magnetic storms*. The *Van Allen radiation belts* are zones of charged particles trapped by the Earth's magnetic field. Any changes in the solar wind, and hence in the magnetic field, can allow these charged particles to enter the upper atmosphere, where they cause the spectacular auroral displays known as the *Northern* and *Southern Lights*. The active processes and interactions which take place between the magnetosphere and the solar wind are collectively referred to as 'space weather'. The Sun is also responsible for the *diurnal* (daily) variation in the Earth's magnetic field. This variation, which has an amplitude of less than 0.5% of the total field, is the main short-period variation in the Earth's magnetic field.

The long-term reversals of the magnetic field, mentioned in Section 3.1.1, are used to date the oceanic lithosphere. However, the geomagnetic field changes

on an intermediate timescale too: the pole wanders by a few degrees a century. This wandering has been measured throughout historical time and is termed *secular variation*. However, it appears that, on average over geological time, the geomagnetic field axis has been aligned along the Earth's spin axis (i.e., on average the geomagnetic poles have been coincident with the geographic poles). This means that, to a first approximation, the geomagnetic field can be modelled as the field of a dipole aligned along the geographic north–south axis. This assumption is critical to all palaeomagnetic work: if the dipole axis had wandered randomly in the past and had not, on average, been aligned along the geographic axis, all palaeomagnetic estimates of past positions of rock samples would be meaningless because they would be relative only to the position of the geomagnetic pole at the time each sample acquired its permanent magnetization and would have nothing at all to do with the geographic pole.

The *magnetic potential* from which the Earth's magnetic field is derived can be expressed as an infinite series of spherical harmonic functions. The first term in this series is the potential due to a dipole situated at the centre of the Earth. At any position **r**, from a dipole, the magnetic potential $V(\mathbf{r})$ is given by

$$V(\mathbf{r}) = \frac{1}{4\pi r^3}\mathbf{m}\cdot\mathbf{r} \tag{3.1}$$

where **m** is the dipole moment, a vector aligned along the dipole axis. For the Earth **m** is 7.94×10^{22} A m^2 in magnitude. The magnetic field $\mathbf{B}(\mathbf{r})$ at any position **r** can then be determined by differentiating the magnetic potential:

$$\mathbf{B}(\mathbf{r}) = -\mu_0 \nabla V(\mathbf{r}) \tag{3.2}$$

where $\mu_0 = 4\pi \times 10^{-7}$ kg m A^{-2} s^{-2} is the magnetic permeability of free space (A is the abbreviation for amp).

To apply Eq. (3.2) to the Earth, we find it most convenient to work in spherical polar coordinates (r, θ, ϕ), (r is the radius, θ the colatitude and ϕ the longitude or azimuth on the sphere, as shown in Fig. A1.4). The magnetic field $\mathbf{B}(\mathbf{r})$ is then written as $\mathbf{B}(\mathbf{r}) = (B_r, B_\theta, B_\phi)$ in this coordinate system. B_r is the radial component of the field, B_θ is the southerly component and B_ϕ is the easterly component. (See Appendix 1 for details of this and other coordinate systems.)

In spherical polar coordinates, if we assume that **m** is aligned along the *negative* z axis (see caption for Fig. 3.2), Eq. (3.1) is

$$\begin{aligned} V(\mathbf{r}) &= \frac{1}{4\pi r^3}\mathbf{m}\cdot\mathbf{r} \\ &= -\frac{mr\cos\theta}{4\pi r^3} \\ &= -\frac{m\cos\theta}{4\pi r^2} \end{aligned} \tag{3.3}$$

Substitution of Eq. (3.3) into Eq. (3.2) gives the three components (B_r, B_θ, B_ϕ) of the magnetic field due to a dipole at the centre of the Earth. The radial component

of the field is B_r:

$$\begin{aligned} B_r(r,\theta,\phi) &= -\mu_0 \frac{\partial V}{\partial r} \\ &= \frac{\mu_0 m \cos\theta}{4\pi} \frac{\partial}{\partial r}\left(\frac{1}{r^2}\right) \\ &= -\frac{2\mu_0 m \cos\theta}{4\pi r^3} \end{aligned} \tag{3.4}$$

The component of the field in the θ direction is B_θ:

$$\begin{aligned} B_\theta(r,\theta,\phi) &= -\mu_0 \frac{1}{r}\frac{\partial V}{\partial \theta} \\ &= \frac{\mu_0 m}{4\pi r^3}\frac{\partial}{\partial \theta}(\cos\theta) \\ &= -\frac{\mu_0 m \sin\theta}{4\pi r^3} \end{aligned} \tag{3.5}$$

The third component is B_ϕ:

$$\begin{aligned} B_\phi(r,\theta,\phi) &= -\mu_0 \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi} \\ &= 0 \end{aligned} \tag{3.6}$$

Note that, by symmetry, there can obviously be no field in the ϕ (east) direction. The total field strength at any point is

$$\begin{aligned} B(r,\theta,\phi) &= \sqrt{B_r^2 + B_\theta^2 + B_\phi^2} \\ &= \frac{\mu_0 m}{4\pi r^3}\sqrt{1+3\cos^2\theta} \end{aligned} \tag{3.7}$$

Along the north-polar axis ($\theta = 0$) the field is

$$\begin{aligned} B_r(r,0,\phi) &= -\frac{\mu_0 m}{2\pi r^3} \\ B_\theta(r,0,\phi) &= 0 \end{aligned} \tag{3.8}$$

On the equator ($\theta = 90^\circ$) the field is

$$\begin{aligned} B_r(r,90,\phi) &= 0 \\ B_\theta(r,90,\phi) &= -\frac{\mu_0 m}{4\pi r^3} \end{aligned} \tag{3.9}$$

and along the south-polar axis ($\theta = 180^\circ$) the field is

$$\begin{aligned} B_r(r,180,\phi) &= \frac{\mu_0 m}{2\pi r^3} \\ B_\theta(r,180,\phi) &= 0 \end{aligned} \tag{3.10}$$

If we define a constant B_0 as

$$B_0 = \frac{\mu_0 m}{4\pi R^3} \tag{3.11}$$

where R is the radius of the Earth, then Eqs. (3.4) and (3.5) give the components of the magnetic field at the Earth's surface:

$$B_r(R,\theta,\phi) = -2B_0\cos\theta \tag{3.12}$$

$$B_\theta(R,\theta,\phi) = -B_0\sin\theta \tag{3.13}$$

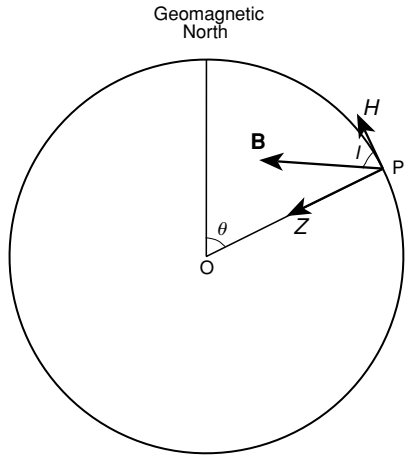


Figure 3.5. A cross section through the Earth, illustrating the components of the dipole field at a location P on the surface. O is the centre of the Earth, θ the colatitude of P, Z and H are the vertical and horizontal components of $\boldsymbol{B}$ and I is the angle of inclination.

It follows immediately that, at any point on the Earth's surface, the expression $B_r^2 + 4B_\theta^2$ has the constant value $4B_0^2$. B_0 is most simply visualized in practice as being the equatorial field of the best-fitting dipole field (see Eq. (3.9)). The strength of the field at the poles is about 6×10^{-5} teslas (T) or 6×10^4 nanoteslas (nT) (1 weber $\mathrm{m}^{-2} = 1$ T), and at the equator it is about 3×10^{-5} T (Fig. 3.1(a)).

In geomagnetic work, the inward radial component of the Earth's field is usually called Z (it is the downward vertical at the Earth's surface) and is positive. The horizontal magnitude (always positive) is called H.

Thus, for a dipole,

$$Z(R, \theta, \phi) = -B_r(R, \theta, \phi) \tag{3.14}$$

$$H(R, \theta, \phi) = |B_\theta(R, \theta, \phi)| \tag{3.15}$$

At the surface of the Earth the angle between the magnetic field and the horizontal is called the *inclination* I (Fig. 3.5):

$$\tan I = \frac{Z}{H} \tag{3.16}$$

Substituting for B_r and B_θ from Eqs. (3.12) and (3.13) gives

$$\begin{aligned} \tan I &= \frac{2\cos\theta}{\sin\theta} \\ &= 2\cot\theta \\ &= 2\tan\lambda \end{aligned} \tag{3.17}$$

where λ is the magnetic latitude ($\lambda = 90 - \theta$). Equation (3.17) makes it a simple matter to calculate the magnetic latitude, given the angle of inclination, and vice versa. Mariners use the angle of inclination for navigational purposes.

The angle of *declination* is the azimuth of the horizontal component of the magnetic field. It is measured in degrees east or west of north. Mariners call it the variation or magnetic variation of the compass.[4] Figure 3.3(b) illustrates

[4] The magnetic compass was a Chinese invention. Declination was described by Shen Kua in 1088.

how the declination has varied over the last four centuries. In the case of a dipole field aligned along the Earth's rotation axis, the declination would always be zero. Obtaining a very detailed image of the historical magnetic field will not be possible since the measurements were recorded in ship's logs and are therefore necessarily confined to the main shipping and exploration routes.

3.1.3 Magnetization of rocks

Rocks can become permanently magnetized by the Earth's magnetic field. This fact has enabled geophysicists to track past movements of the plates.

As any volcanic rock cools, it passes through a series of critical temperatures at which the various grains of iron minerals acquire spontaneous magnetization. These critical temperatures, called the *Curie points* or Curie temperatures, are different for each mineral (e.g., approximately 580 °C for magnetite (Fe_3O_4) and 680 °C for haematite (Fe_2O_3)). Once the temperature of the rock is lower than the *blocking temperature*, (which for most minerals is tens of degrees less than the Curie point) the magnetized grains cannot be reoriented. This means that the grains have their magnetic moments aligned with the direction the Earth's magnetic field had at the time of cooling. Both of these temperatures are much lower than the temperatures at which lavas crystallize (typically 800–1100 °C), which means that magnetization becomes permanent some time *after* lavas solidify. How long afterwards depends on the physical size and other properties of the intrusion or flow and the rate of cooling, which depends in turn on its environment (see Section 7.8). This type of permanent residual magnetization is called *thermoremanent magnetization* (TRM) and is considerably larger in magnitude than the magnetism induced in the basalt by the Earth's present field.

Sedimentary rocks can also acquire remanent magnetization even though they have never been as hot as 500 °C, but this remanent magnetization of sediments is generally very much less than that of igneous rocks. Sedimentary rocks can acquire magnetization in two ways: *depositional* or *detrital remanent magnetization* (DRM) and *chemical remanent magnetization* (CRM). Detrital remanent magnetization can be acquired, as indicated by its name, during the deposition of sedimentary rocks. If the sediments are deposited in still water, any previously magnetized small grains will align themselves with their magnetic moments parallel to the Earth's magnetic field as they fall. Large grains are unaffected. Chemical remanent magnetization is acquired *in situ after* deposition during the chemical growth of iron oxide grains, as in a sandstone. When the grains reach some critical size, they become magnetized in the direction of the Earth's field at that time. Chemical remanent magnetization is thus a secondary remanent magnetization whereas TRM and DRM are both primary remanent magnetizations dating from the time of formation of the rocks.

The degree to which a rock body can be magnetized by an external magnetic field is determined by the *magnetic susceptibility* of the rock. *Induced*

magnetization is the magnetization of the rock **M** which is induced when the rock is put into the Earth's magnetic field **B**; it is given by

$$\mu_0 \mathbf{M}(\mathbf{r}) = \chi \mathbf{B}(\mathbf{r}) \tag{3.18}$$

where χ is the magnetic susceptibility (a dimensionless physical property of the rock). Values of χ for basalts vary from about 10^{-4} to 10^{-1}, so the induced magnetization gives rise to a field that is very much weaker than the Earth's field. Thermoremanent magnetization is generally many times stronger than this induced magnetization. For any rock sample, the ratio of its remanent magnetization to the magnetization induced by the Earth's present field is called the *Königsberger ratio Q*. Measured values of *Q* for oceanic basalts are in the range 1–160. Thus, an effective susceptibility for TRM of basalt of about 10^{-3}–10^{-1} appears to be reasonable. This permanent TRM therefore produces a local field of perhaps 1% of the Earth's magnetic field. Effective susceptibilities for sedimentary rocks (DRM and CRM) are about two orders of magnitude less than those for basalt (TRM).

The relationship between the angle of inclination and the magnetic latitude (Eq. (3.17)) means that a measurement of the angle of inclination of the remanent magnetization of a suitable lava or sediment laid down on a continent immediately gives the magnetic palaeolatitude for the particular piece of continent. If the continent has not moved with respect to the pole since the rock cooled, then the magnetic latitude determined from the magnetization of the rock is the same as its present latitude. However, if the continent has moved or if the rock has been tilted, the magnetic latitude determined from the magnetization of the rock can be different from its present latitude. Thus, the angle of inclination provides a powerful method of determining the past latitudes (*palaeolatitudes*) of the continents. Unfortunately, it is not possible to use palaeomagnetic data to make a determination of palaeolongitude.

If the angles of declination and of inclination of our rock sample are measured, the position of the palaeomagnetic pole can be calculated. To do this, it is necessary to use spherical geometry, as in the calculations of Chapter 2. Figure 2.11 shows the appropriate spherical triangle if we assume N to be the present north pole, P the palaeomagnetic north pole and X the location of the rock sample. The cosine formula for a spherical triangle (e.g., Eq. (2.9)) gives the geographic latitude of the palaeomagnetic pole P, λ_p, as

$$\cos(90 - \lambda_p) = \cos(90 - \lambda_x)\cos(90 - \lambda) + \sin(90 - \lambda_x)\sin(90 - \lambda)\cos D \tag{3.19}$$

where λ_x is the geographic latitude of the sample location, D the measured remanent declination and λ the palaeolatitude (given by Eq. (3.17)). Simplifying Eq. (3.19) gives

$$\sin \lambda_p = \sin \lambda_x \sin \lambda + \cos \lambda_x \cos \lambda \cos D \tag{3.20}$$

Example: calculation of palaeomagnetic latitude

Magnetic measurements have been made on a basalt flow at present at 47°N, 20°E. The angle of inclination of the remanent magnetization of this basalt is 30°. Calculate the magnetic latitude of this site at the time the basalt was magnetized.

The magnetic latitude λ is calculated by using Eq. (3.17):

$$\tan I = 2 \tan \lambda$$

I is given as 30°, so

$$\begin{aligned} \lambda &= \tan^{-1}\left(\frac{\tan 30^\circ}{2}\right) \\ &= \tan^{-1}(0.2887) \\ &= 16.1^\circ \end{aligned}$$

Therefore, at the time the sample was magnetized, it was at a magnetic latitude of 16°N, which indicates that between then and now the site has moved 31° northwards to its present position at 47°N.

After λ_p has been calculated, the sine formula for a spherical triangle (e.g., Eq. (2.10)) can be used to give the difference between the longitudes of the palaeomagnetic pole and the sample location, $\phi_p - \phi_x$:

$$\begin{aligned} \sin(\phi_p - \phi_x) &= \frac{\sin(90 - \lambda)\sin \mathrm{D}}{\sin(90 - \lambda_p)} \\ &= \frac{\cos\lambda \sin \mathrm{D}}{\cos\lambda_p} \qquad \sin\lambda \geq \sin\lambda_p \sin\lambda_x \\ \sin(180 + \phi_p - \phi_x) &= \frac{\cos\lambda \sin \mathrm{D}}{\cos\lambda_p} \qquad \sin\lambda < \sin\lambda_p \sin\lambda_x \end{aligned} \tag{3.21}$$

Therefore, by using Eqs. (3.17) and (3.19)–(3.21), we can calculate past magnetic-pole positions.

If palaeomagnetic pole positions can be obtained from rocks of different ages on the same continent, these poles can be plotted on a map. Such a plot is called a *polar-wander path* (a name that is a 'fossil' of the older notion that it was the poles, not the continents, that drifted) and shows how the magnetic pole moved relative to that continent. If such polar-wander paths from two continents coincide, then the two continents cannot have moved relative to each other during the times shown. However, if the paths differ, there must have been relative motion of the continents. Figure 3.6 shows polar-wander paths for Europe and North America for the last 550 Ma. Although these two paths have almost the same shape, they are certainly not coincident. When the opening of the Atlantic Ocean is taken into account, however, the two paths can be rotated on top of each other, and they then do approximately coincide.

Example: calculation of latitude and longitude of the palaeomagnetic pole

If the angle of declination for the basalt flow of the previous example is 80°, calculate the latitude and longitude of the palaeomagnetic pole.

The latitude of the palaeomagnetic pole is calculated by using Eq. (3.20) with $\lambda = 16^\circ$ and $D = 80^\circ$:

$$\sin\lambda_p = \sin 47 \sin 16 + \cos 47 \cos 16 \cos 80$$

Thus, $\lambda_p = 18^\circ$.

The longitude of the palaeomagnetic pole can now be calculated from Eq. (3.21) after checking whether $\sin\lambda \geq \sin\lambda_p \sin\lambda_x$ or not:

$$\sin(\phi_p - \phi_x) = \frac{\cos 16 \sin 80}{\cos 18}$$

Thus, $\phi_p - \phi_x = 84^\circ$, and so $\phi_p = 104^\circ$. The position of the palaeomagnetic pole is therefore 18°N, 104°E.

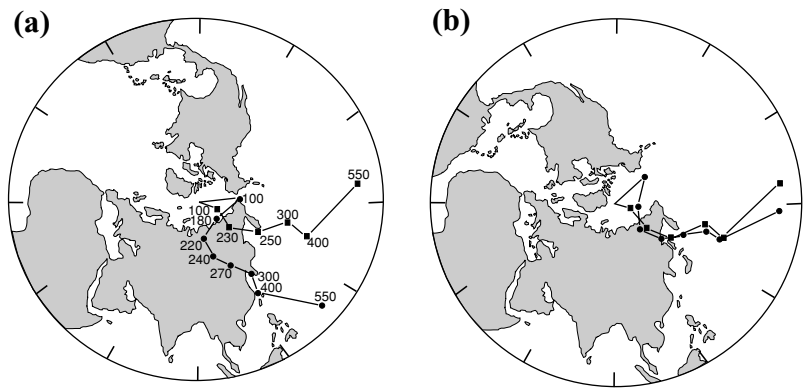


Figure 3.6. (a) Polar-wander curves for North America (circles) and Europe (squares). (b) Polar-wander curves for North America and Europe when allowance has been made for the opening of the Atlantic Ocean. The two curves are now almost coincident. (After McElhinny (1973).)

3.2 Dating the oceanic plates

3.2.1 Magnetic stripes

To use measurements of the magnetic field to gain information about the magnetization of the crust, it is first necessary to subtract the regional value of the geomagnetic field (e.g., IGRF 1995). What remains is the magnetic anomaly. Over the oceans, magnetic-field measurements are made by towing a magnetometer behind a ship. A magnetometer can be routinely towed while a research ship is on passage or doing other survey work that does not involve slow, tight manoeuvring. These marine instruments measure the magnitude B, but not the

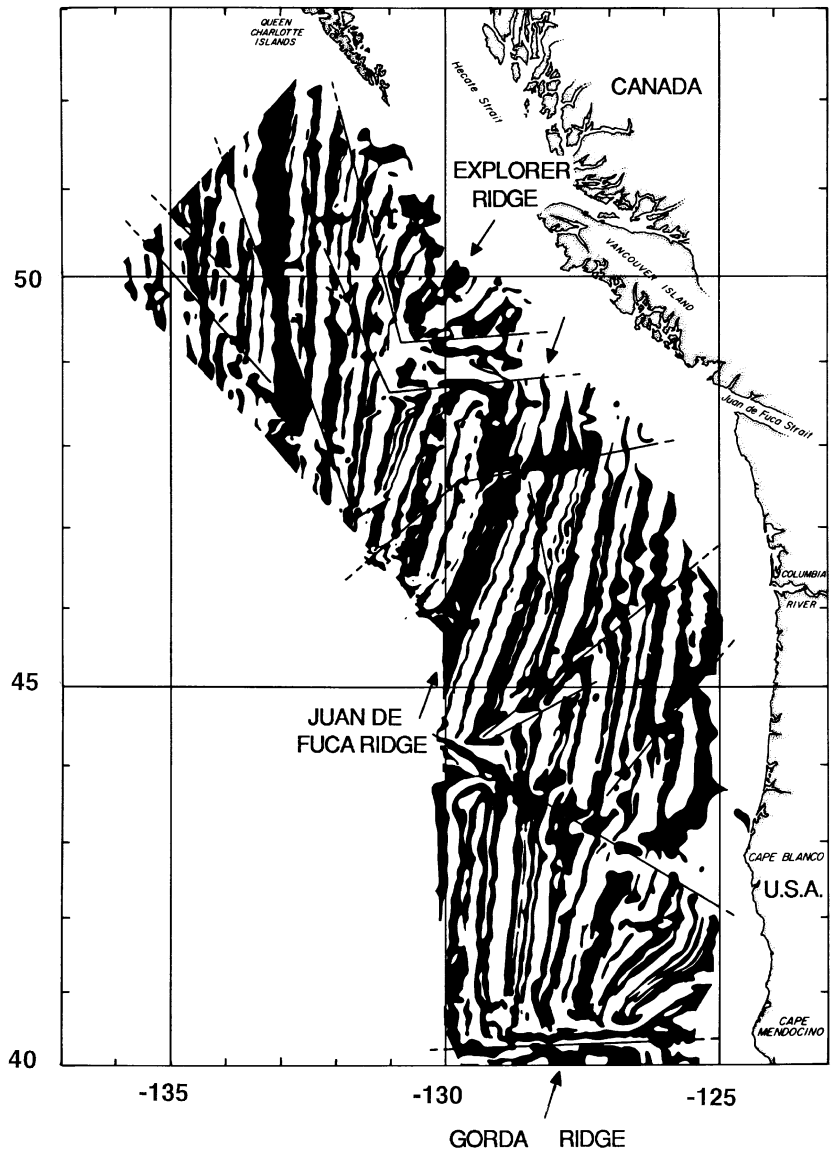


Figure 3.7. Anomalies of the total magnetic field southwest of Vancouver Island in the northeast Pacific Ocean. Positive anomalies are shown in black and negative anomalies in white. Arrows indicate the axes of the three mid-ocean ridges; straight lines indicate the main faults which offset the anomaly pattern. This was the first large-scale map to show the details of the magnetic anomalies over an active mid-ocean ridge. As such it was a vital piece of evidence in the development of the theories of seafloor spreading and then plate tectonics. Compare this with a more recent magnetic-anomaly map for the same area shown in Fig. 3.22. (After Raff and Mason (1961).)

direction, of the total field **B**. (Magnetometers that can measure both magnitude and direction are widely used on land as prospecting tools.) Marine magnetic anomalies are therefore anomalies in the magnitude (or total intensity) of the magnetic field.

The first detailed map of magnetic anomalies off the west coast of North America, published in 1961, showed what was then a surprising feature: alternate stripes of anomalously high and low values of the magnetic field stretching over the entire region (Fig. 3.7). All subsequent magnetic-anomaly maps show that these stripes are typical of oceanic regions. The stripes run parallel to and are

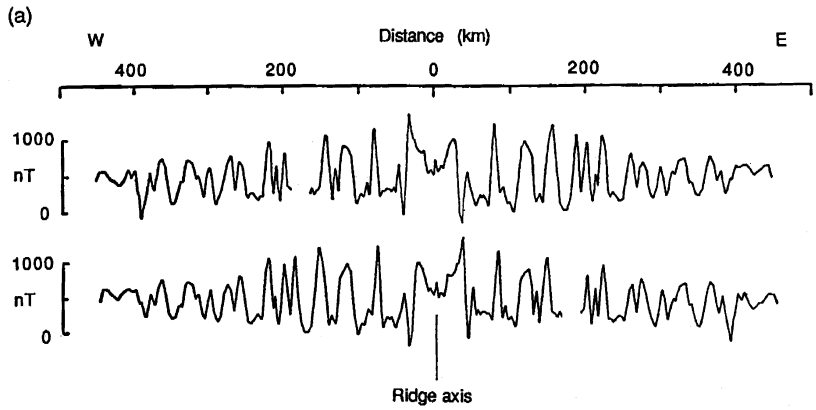


Figure 3.8. (a) A magnetic-anomaly profile across the Pacific–Antarctic Ridge plotted above its end-to-end reverse (mirror image) demonstrates the commonly observed symmetry of magnetic anomalies about the ridge axis. The half-spreading rate as determined from this profile is 4.5 cm yr^{-1}. Magnetic anomalies are generally about ±500 nT in amplitude, about 1% of the Earth's magnetic field. (After Pitman and Heirtzler (1966).)

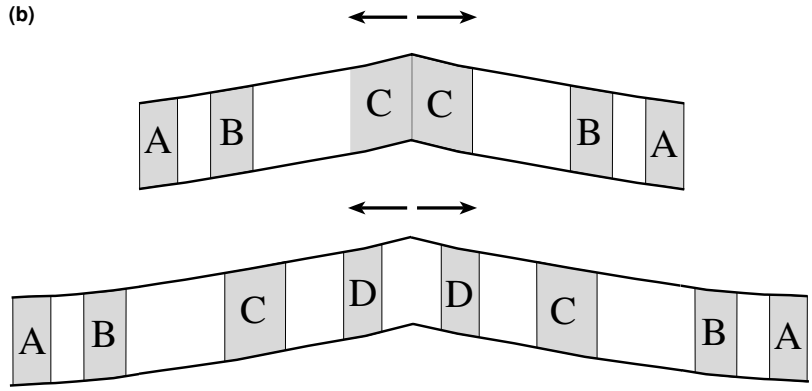


Figure 3.8. (b) A cross section through an idealized ridge illustrates the block model of normally (black) and reversely (white) magnetized material. In the upper drawing, the Earth's magnetic field is positive (normal) and has been for some time; thus, the positively magnetized blocks C have formed. Positively magnetized blocks B and A were formed during earlier times when the Earth's magnetic field also had normal polarity. The lower drawing shows the same ridge at a later time; another positive block D has formed and the magnetic field is in a period of reverse polarity.

generally symmetrical about the axes of the mid-ocean ridges. They are offset by fracture zones, and are a few tens of kilometres in width and typically ±500 nT in magnitude (Fig. 3.8(a)). This means that magnetic-anomaly maps are very useful in delineation of ridge axes and fracture zones.

The origin of these magnetic stripes was first correctly understood in 1963 by F. J. Vine and D. H. Matthews and independently by L. W. Morley. They realized

that the idea of seafloor spreading, which had then newly been proposed by H. H. Hess, coupled with the then recently discovered evidence that intermittent reversals of the Earth's magnetic field have taken place, provide the answer. The oceanic crust, which is formed along the axes of the mid-ocean ridges as mafic material wells up, acts as a double-headed magnetic tape recorder that preserves the past reversals of the magnetic field on each plate (Fig. 3.8(b)). The width of magnetic stripe is determined by the speed at which 'the tape' is moving (the half-spreading rate) and the length of time between the magnetic reversals. Thus, while the Earth's magnetic field is in its normal polarity, a block of oceanic crust is formed with a strong component of permanent magnetization aligned with the normal field. When the Earth's field is reversed, new oceanic crust will have a strong component of permanent magnetization aligned with the reversed field. In this way a magnetically normal- and reversed-striped oceanic crust is formed, with the stripes parallel to the ridge axis. To 'decode' a magnetic-anomaly pattern, however, it is necessary to know either when the Earth's field reversed, or the half-spreading rate of the ridge.

Example: variability of marine magnetic anomalies

Decoding magnetic anomalies is not as simple in practice as it sounds, partly because the magnetization of the oceanic crust does not conform to a perfect block model. The lava flows that make up the magnetized layer are not produced continuously along the ridge axis. Rather, they are extruded randomly both in time and in space within an *emplacement zone* that is centred on the ridge axis. The effect of this is shown in Fig. 3.9. The simple block model (Fig. 3.8) gives way to a much more complex structure as the width of the emplacement zone is increased. The variability in magnetic anomalies increases accordingly. An emplacement zone 10 km across, which seems to be appropriate for the Mid-Atlantic Ridge, explains the variability of Atlantic magnetic anomalies. Pacific magnetic anomalies are much less variable, partly because the emplacement zone seems to be narrower but mainly because the much faster spreading rate means that the polarity reversals are much more widely spaced.

Nevertheless, the block model for the magnetization of oceanic crust is widely used. All the anomalies shown in the rest of this chapter have been calculated for block models.

By 1966, researchers had established a *reversal timescale* extending back some 4 Ma by using potassium–argon isotopic dating (see Section 6.7) to fit a timescale to the magnetic-reversal sequence that had been measured in continental lava piles and on oceanic islands. This timescale of reversal was then used to interpret oceanic anomalies by calculating theoretical magnetic anomalies for assumed spreading rates and latitudes. Figure 3.10 shows the theoretical

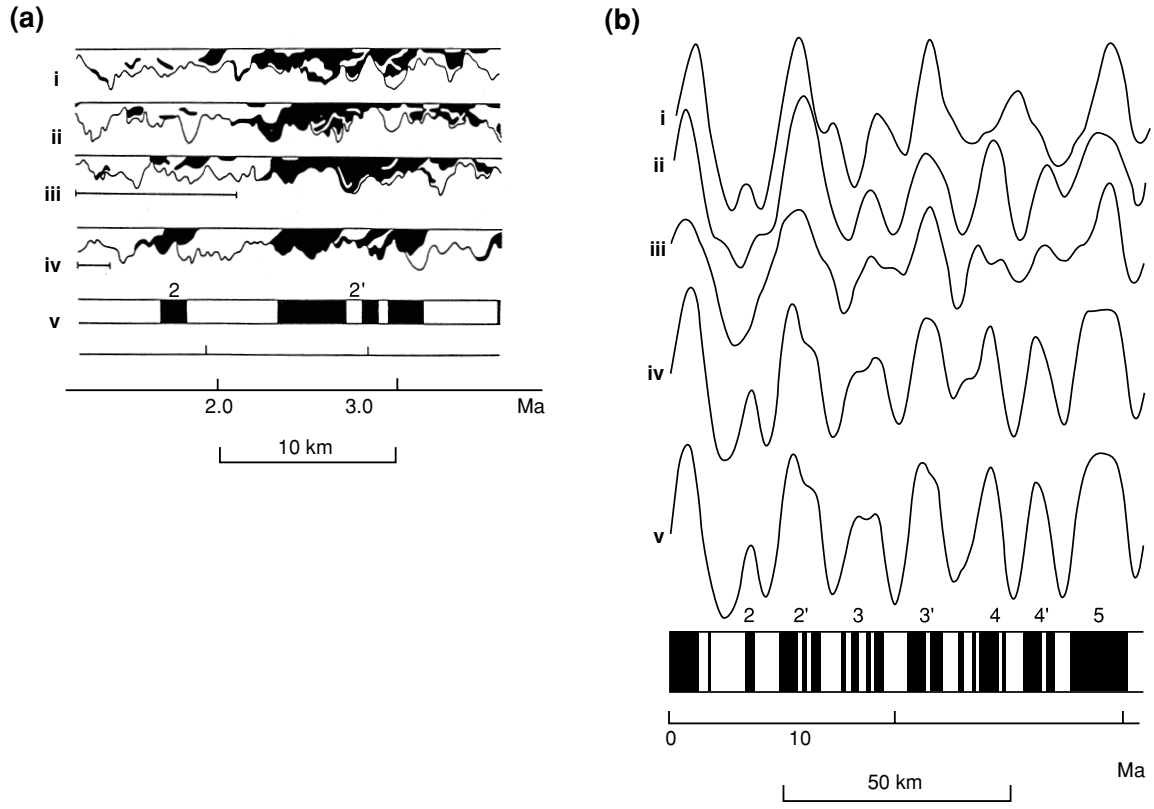


Figure 3.9. (a) A cross section through the upper part of the oceanic crust, showing the magnetized layer generated by extrusion of lava randomly in both time and space within an emplacement zone centred on the ridge axis. The half-spreading rate is 1 cm yr^{-1}. Black, normally magnetized rock; white, reversely magnetized rock. (b) Theoretical magnetic anomalies generated by this random-extrusion model. In both (a) and (b), (i), (ii) and (iii) are for an emplacement zone 10 km across, (iv) is for an emplacement zone 2 km across and (v) is for the block model in which the emplacement zone has zero width. (From Schouten and Denham (1979).)

anomalies calculated for the Juan de Fuca Ridge and the East Pacific Rise using this 1966 reversal timescale. These theoretical anomalies match the actual anomalies very well and were used to confirm the *Vine–Matthews hypothesis*. Figure 3.11 shows a more recent (and therefore more detailed) determination of the reversal sequence for the last 4 Ma.

A geomagnetic timescale extending back 80 Ma was first established in 1968 by assuming that the spreading rate in the South Atlantic had remained constant from 80 Ma until now. The observed anomaly pattern was matched with theoretical profiles computed for a sequence of normal and reversed magnetized blocks symmetrical about the Mid-Atlantic Ridge axis (Fig. 3.12). The timescale derived

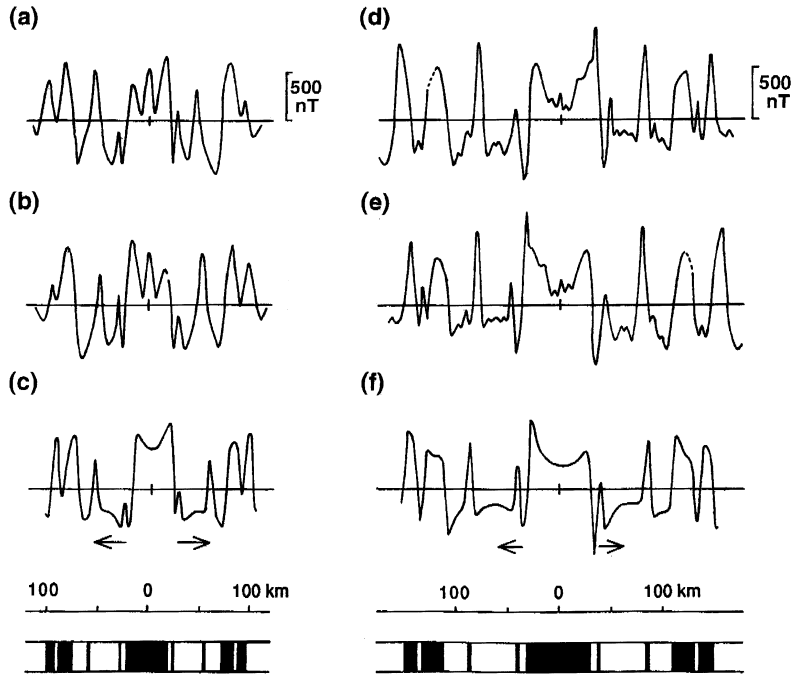


Figure 3.10. (a) The magnetic-anomaly profile over the Juan de Fuca Ridge at 46°N, southwest of Vancouver Island, Canada. (b) The profile in (a) is reversed; note the symmetry. (c) A model magnetic-anomaly profile calculated for this ridge assuming a half-spreading rate of 2.9 cm yr^{-1} and the magnetic-reversal sequence shown below. On the magnetic-reversal timescale, black blocks represent periods of normal polarity and white blocks periods of reverse polarity. (d) The magnetic-anomaly profile over the East Pacific Rise at 51°S. (e) The profile in (d) is reversed; note the symmetry. (f) A model magnetic-anomaly profile calculated for this ridge assuming a half-spreading rate of 4.4 cm yr^{-1} and the magnetic-reversal sequence shown below. The magnetic-reversal timescale is the same as that for the Juan de Fuca Ridge (see (c)), but with a different horizontal scale because of the faster spreading rate. (After Vine (1966).)

from this model and the whole seafloor-spreading hypothesis were spectacularly confirmed by drilling the ocean bottom as part of the Deep Sea Drilling Project (DSDP), an international enterprise that began in 1968 (see Section 9.2.1). During Leg 3 of the DSDP a series of holes was drilled into the basalt at the top of the oceanic crust right across the Atlantic at 30°S (Fig. 3.13). It was not possible to use radiometric methods to date the lavas sampled from the top of the crust because they were too altered; instead, the basal sediments were dated using fossils. The ages are therefore slightly younger than the lava ages would have been. Figure 3.13(b) shows these sediment ages plotted against the distances of their sites from the ridge axis. The straight line confirmed that spreading in the

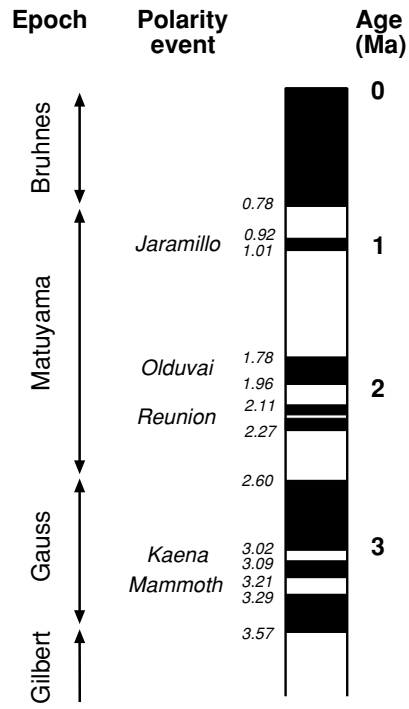


Figure 3.11. Details of the recent reversals of the Earth's magnetic field as determined from detailed radiometric dating of continental and oceanic-island lavas and palaeomagnetism of marine sediments. The *epochs*, time intervals during which the Earth's magnetic field was either predominantly normal or predominantly reversed in polarity, have been named after prominent scientists in the study of the Earth's magnetic field. (William Gilbert was a sixteenth-century English physician, Carl Friedrich Gauss was a nineteenth-century German mathematician, Bernard Brunhes (1906) was the first person to propose that the Earth's magnetic field was reversed at the time lavas were formed and Motonori Matuyama (1929) was the first person to attempt to date these reversals.) The *polarity events (subchrons)*, short fluctuations in the magnetic polarity, are named after the geographic location where they were first recognized (e.g., Olduvai Gorge, Tanzania, the site of the early hominid discoveries of Leakey; Mammoth, California, U.S.A.; and Jaramillo Creek, New Mexico, U.S.A.). (Based on McDougall *et al.* (1992) and Spell and McDougall (1992).)

South Atlantic had been continuous and had been occurring at a fairly steady rate for the last 80 Ma, and that the geomagnetic timescale (Fig. 3.12) was reasonably accurate. Since then, considerable effort has been put into ensuring that geological and magnetic timescales are as precise as possible. Figure 3.14 shows a timescale from the middle Jurassic to the present.

To use a geomagnetic timescale to date the oceanic plates, it is necessary to recognize specific anomalies. Fortunately, the reversal sequence is sufficiently irregular (Fig. 3.14) for this to be possible for the trained eye. The prominent anomalies up to age 83 Ma have been numbered from one to thirty-three. For ages 125–162 Ma they are labelled with the prefix M (M standing for Mesozoic). Particularly prominent is the long Magnetic Quiet Zone in the Cretaceous (83–124 Ma numbered C34), during which no reversals occurred.

Figure 3.12 showed profiles from the Pacific and their corresponding theoretical profiles, as well as the South Atlantic profiles. It is clear that, in contrast to the history of the Atlantic, spreading rates in the Pacific have changed markedly with time. This is also the case in the Indian Ocean. Figure 3.15 is an isochron map of the ocean floor. This shows that the ridges have undergone a number of changes both in rate and in direction of spreading in the past. The oldest parts of the ocean floor are Jurassic. This is an interesting fact in itself and the subject of conjecture about the density and stability of older oceanic lithosphere and causes of initiation of a subduction zone in any particular location.

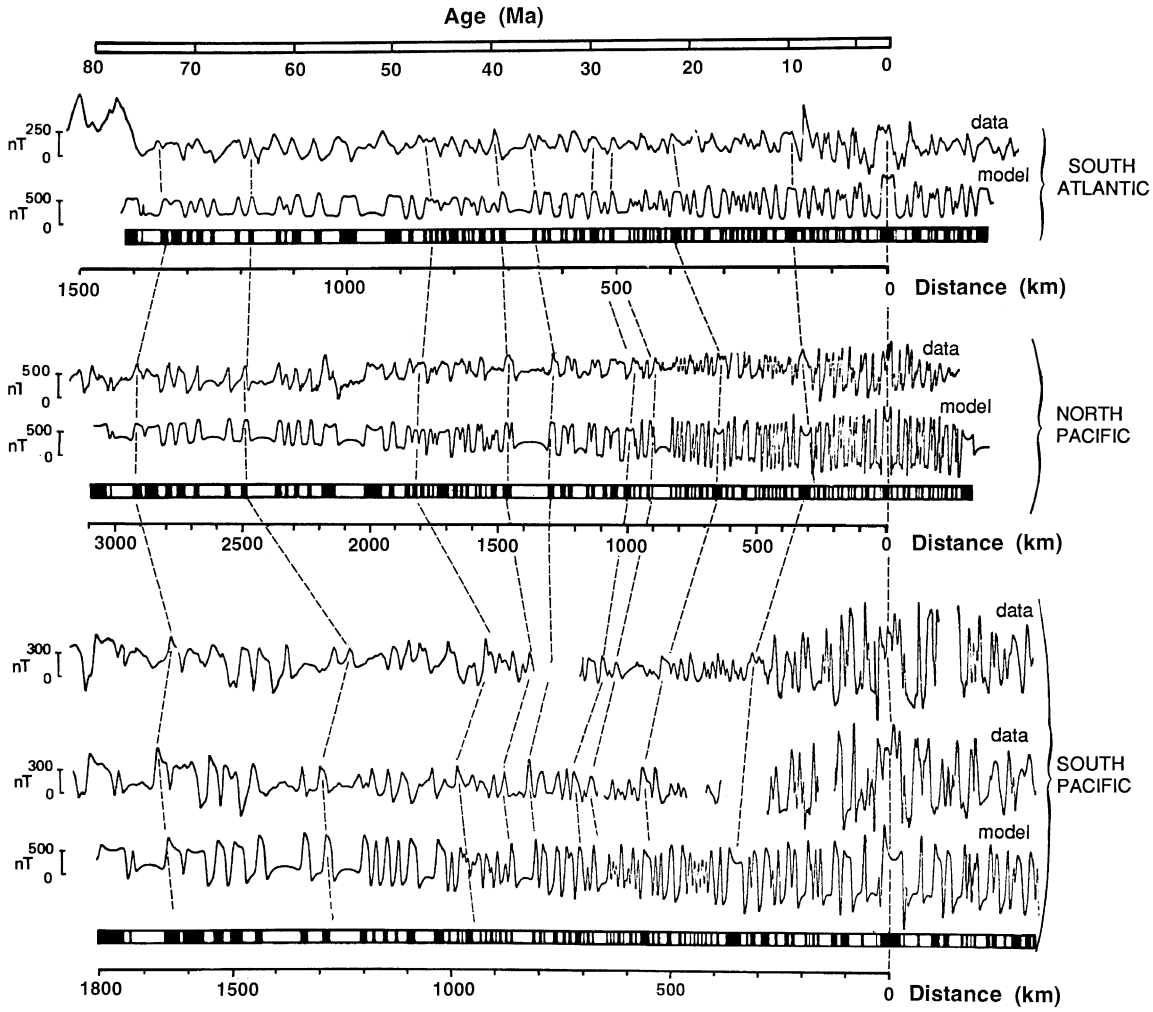


Figure 3.12. The geomagnetic-reversal timescale for the last 80 Ma as proposed in 1968. The magnetic-anomaly profiles with their model profiles and model reversal sequences are shown for the South Atlantic, North Pacific and South Pacific. The South Atlantic timescale was made by assuming that the spreading rate there has been constant for the last 80 Ma. In comparison, spreading in the Pacific has clearly been both faster and more irregular (note the different distance scales for the Pacific data). Dashed lines connect specific magnetic anomalies numbered as in Fig. 3.14. (After Heirtzler *et al.* (1968).)

3.2.2 Calculation of marine geomagnetic anomalies

The marine magnetic-anomaly patterns (e.g., Figs. 3.10 and 3.12) can give immediate values for the relative motion between two plates if specific anomalies can be identified and if the reversal timescale is known. The patterns can also be used to estimate the relative motion between the plate and the Earth's magnetic pole. This

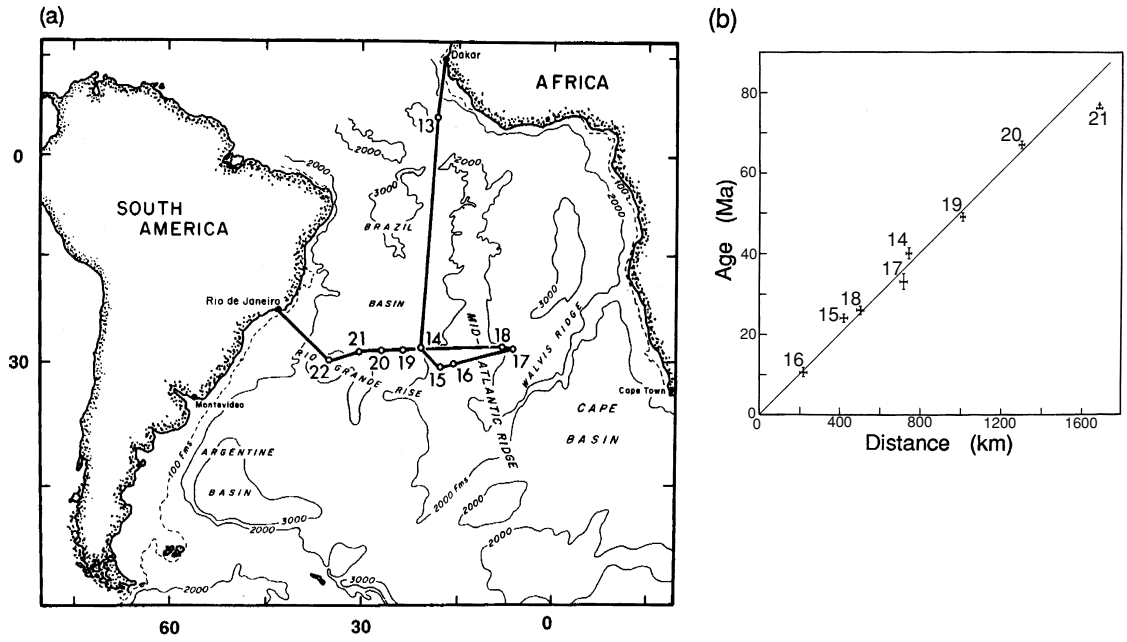


Figure 3.13. (a) Geographic locations of the drilling sites on Leg 3 of the Deep Sea Drilling Project. Site numbers refer to the order in which the holes were drilled. (b) The age of the sediment immediately above the basalt basement versus the distance of the drill site, as shown in (a), from the axis of the Mid-Atlantic Ridge. (From Maxwell *et al.* (1970).)

follows in the same way as that in which the remanent magnetization of continental lavas gives their magnetic latitude (Eq. (3.17)). The magnitude and direction of the remanent magnetization of oceanic crust depend on the latitude at which the crust was formed and are unaffected by later movements and position. This means that, although the magnetic anomaly resulting from this magnetization of the crust is dependent on the present location of that piece of crust, it can be used to determine the original latitude and orientation of the mid-ocean ridge, though not the longitude.

Imagine a mid-ocean ridge spreading symmetrically, producing infinitely long blocks of new crust (Fig. 3.16). The magnetic field measured above any block will include a contribution from the permanent magnetization of the block. Suppose first that, at the time the block was formed, the ridge was at the equator, where the magnetic field is north–south and has no vertical component (Eq. (3.9)), and was spreading east–west. In this case, the permanent magnetization of the block would be along the block, $\mathbf{M} = (0, M_y, 0)$, and so the magnetic field lines cannot leave the block (the block being infinitely long). It is therefore impossible for the magnetic field of this block to affect the magnetic field outside the block. Such a block would not produce an anomaly!

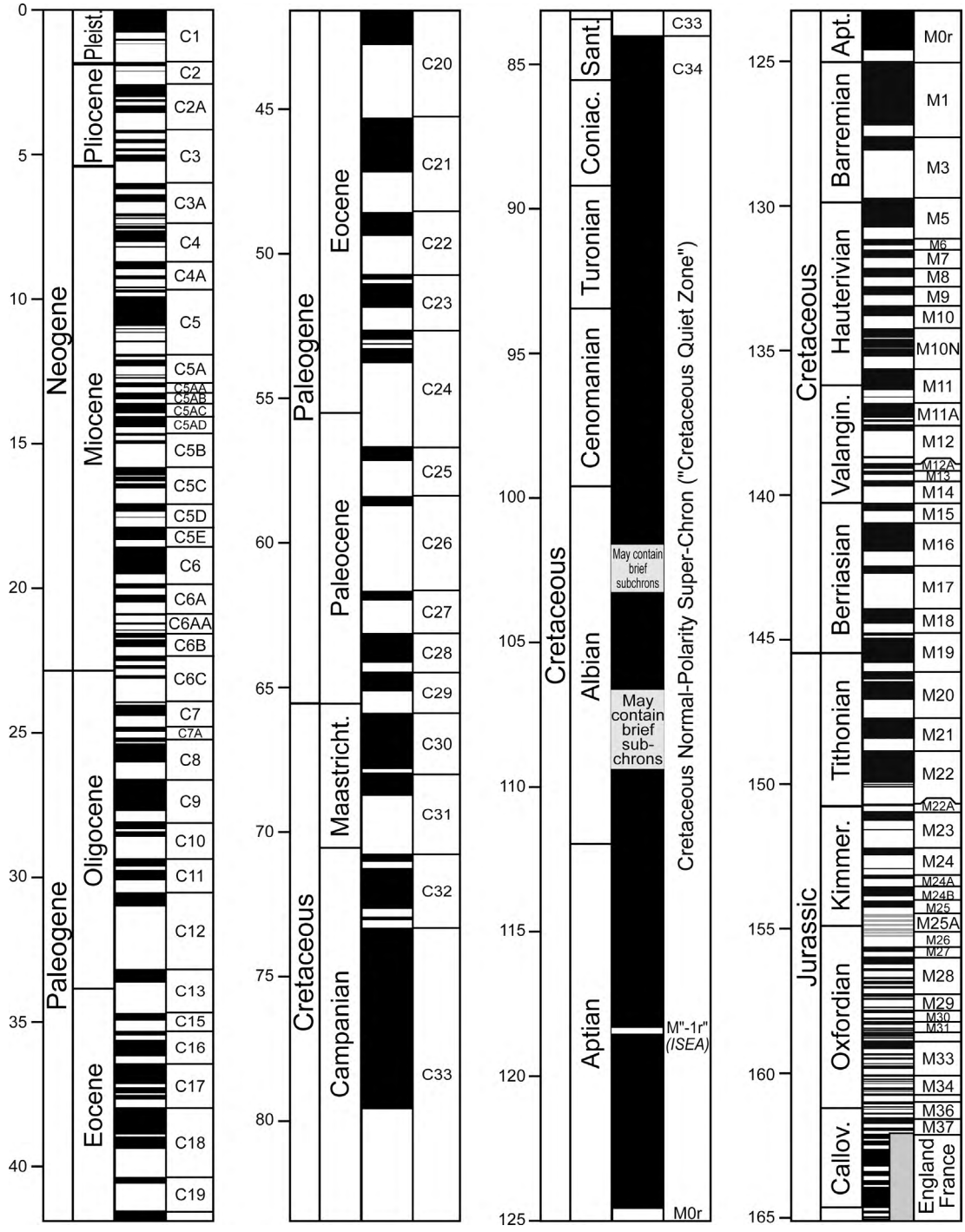


Figure 3.14. The geomagnetic-reversal time scale. Black indicates periods of normal polarity for the Earth's magnetic field; white, reversed-polarity periods. The grey pattern indicates uncertain polarity. The anomaly or chron numbers, C1–C34 and M0–M38, are on the right-hand side of the columns; the age (in Ma) is along the left-hand side. Note that the International scale uses not 'Tertiary' and 'Quaternary' but 'Paleogene' and 'Neogene'. (By Gabi Ogg, after Gradstein *et al.* 2004, *Geologic Time Scale 2004.*)

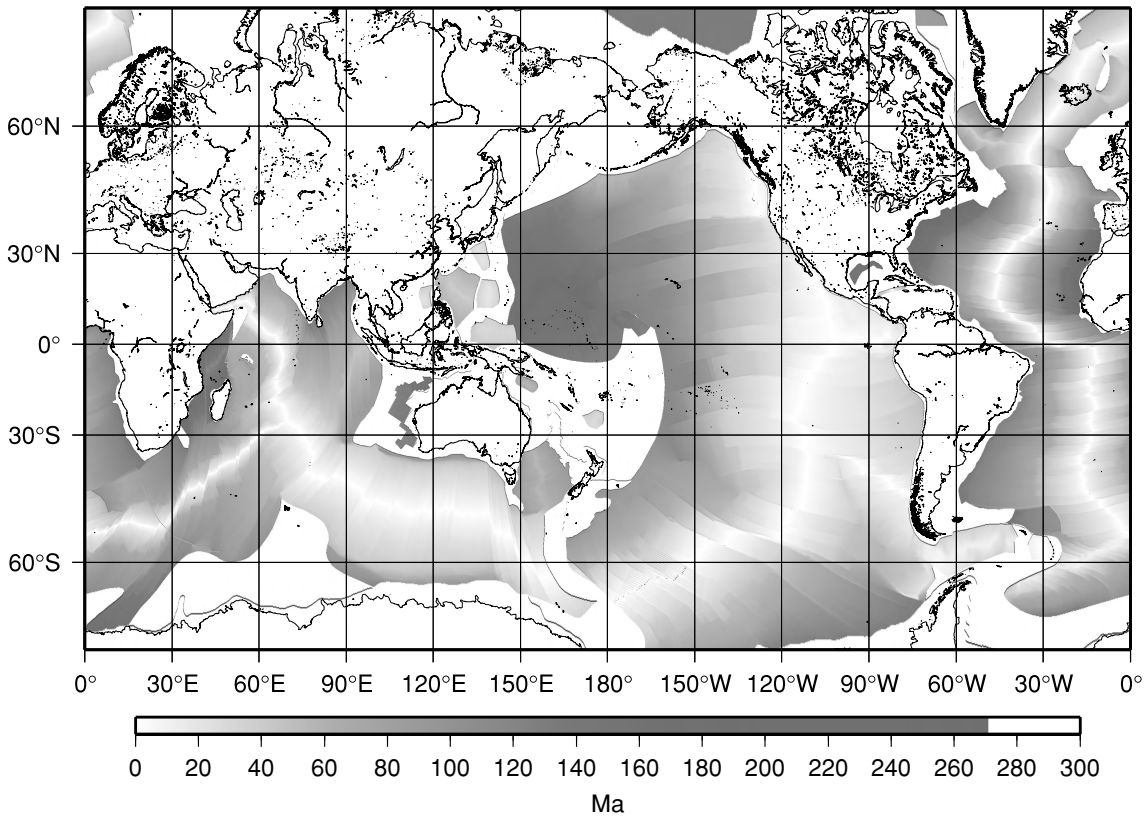


Figure 3.15. The age of the ocean floor determined using magnetic-anomaly data, basement ages from deep-sea drilling, anomaly timescales and rotation poles and angles. For a colour version see Plate 2. (R. D. Müller, personal communication 2004, after Muller *et al.* (2002).)

Now imagine another ridge, still spreading east–west, but not at the equator. Because, in this example (as in the previous one), the ridge is striking north–south, the horizontal remanent magnetization is in the y direction and has no component in the x direction (because $B_\phi = 0$). As before, magnetization in the y direction cannot affect the magnetic field outside the block (the block being infinitely long) and so it is only the vertical component of the remanent magnetization that affects the magnetic field outside the block. The vertical component of the magnetic field as measured today by a ship above the block will be either reduced or increased, depending on whether the magnetic field at the time the block was formed was reversed or normal. The magnetic anomalies produced by a symmetrical pattern of such blocks from an east–west spreading ridge are symmetrical.

For all spreading directions of the ridge that are not east–west, there is a component of the remanent magnetization in the x direction. The effect of this

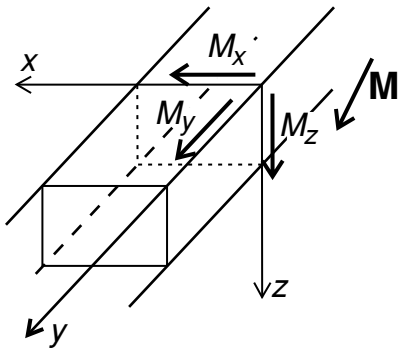


Figure 3.16. An infinitely long block of magnetized oceanic crust. The magnetization of the block **M** is resolved into three components (M_x, M_y and M_z). M_y is the horizontal component parallel to the block and so cannot contribute to the magnetic field outside the block. M_z (the vertical component) and M_x (the horizontal component perpendicular to the block) both contribute to the magnetic field outside the block and hence to the resultant magnetic anomaly produced by the block.

component M_x is complicated, but it usually produces an asymmetry in the magnetic anomalies. In the vicinity of the magnetic poles, where the Earth's magnetic field is almost vertical, the effects of the x component of magnetization are almost negligible; however, close to the magnetic equator the effect becomes important.

Calculation of synthetic magnetic anomalies for any general block model (e.g., Fig. 3.8(b)) is performed in the following way.

1. Calculate remanent magnetization in the x and z directions for each block by assuming the orientation and latitude at which blocks were formed.
2. Calculate the field produced by these blocks along a line perpendicular to the blocks and at a constant distance above them (i.e. sea-surface level).
3. Add this field to the Earth's magnetic field (e.g., IGRF 1995) at the block's present-day latitude. This is the field that would be 'measured'.
4. Calculate the difference in magnitude between this 'measured' field and the Earth's field. (For more details, see McKenzie and Sclater (1971).)

To visualize the effect of present-day latitude on the magnetic anomalies produced by a given block structure, consider a model oceanic crust formed by east–west spreading of a mid-ocean ridge at 40°S. Figure 3.17 shows the magnetic anomalies that this magnetized oceanic crust would produce if it later moved north from 40°S. The present-day latitude is clearly an important factor in the shapes of the anomalies; note in particular that, if the plate moves across the equator, positively magnetized blocks give rise to negative anomalies, and vice versa. (This is a result of B_r at today's latitude and M_z having opposite signs.)

Magnetic anomalies are dependent on the orientation of the ridge at which the crust was formed as well as the latitude. The orientation determines the relative values of M_x and M_y. Figure 3.18 shows anomalies that would be observed at 15°N. The magnetized blocks were produced by ridges at 40°S, striking in the three directions shown. The anomalies produced by the ridge striking N45°E and the ridge striking N45°W are identical. This ambiguity in the strike of the ridge

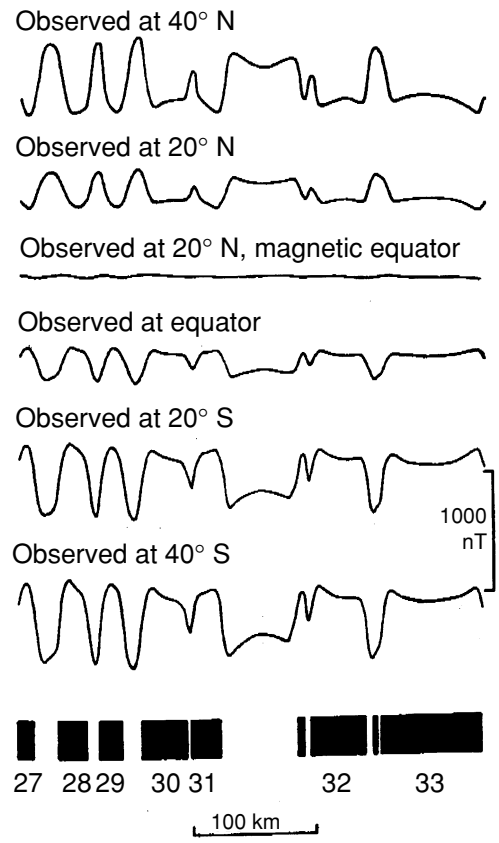


Figure 3.17. Synthetic magnetic-anomaly profiles generated by a ridge at 40°S, spreading east–west with a half-rate of 3 cm yr^{-1}. The magnetized crust subsequently moved northwards to the latitudes shown, leaving the blocks striking north–south. Note that the shapes of the anomalies depend upon the present latitude and that, if the magnetized crust moves across the equator, positively magnetized blocks give rise to negative anomalies. Black denotes periods of normal magnetic field; white denotes periods of reversed field. Numbers are anomaly numbers (see Fig. 3.14). (From McKenzie and Sclater (1971).)

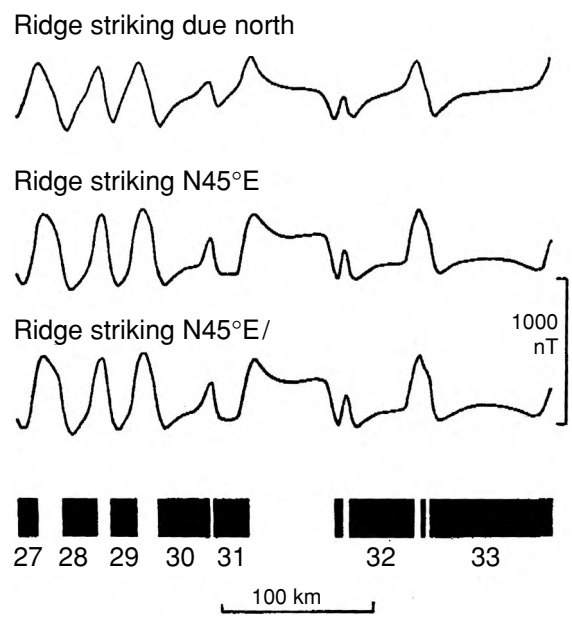


Figure 3.18. Synthetic magnetic-anomaly profiles produced by the magnetized blocks now striking east–west at 15°N. The magnetized oceanic crust was formed at 40°S by mid-ocean ridges spreading at 3 cm yr^{-1} half-rate and striking due north, N45°E and N45°W. Note that the anomalies for the N45°E ridge and the N45°W ridge are exactly the same. Numbers are anomaly numbers (see Fig. 3.14). (From McKenzie and Sclater (1971).)

is always there: Anomalies from a ridge striking $-\alpha$ could also have come from a ridge striking $+\alpha$.

The thickness of the magnetized blocks that give rise to the oceanic magnetic anomalies is not well determined. In general, the anomalies depend primarily on the product of the magnetic susceptibility and layer thickness (i.e., 400 m of material with a susceptibility of 0.05 gives rise to almost the same magnetic anomaly as 2 km of material with a susceptibility of 0.01). At one time it was thought that the magnetized layer was only the very top of the oceanic crust (then termed layer 2A, discussed in Section 9.2.1) but this is not now thought necessarily to be the case. The axial magnetization is generally greater than magnetization of old oceanic crust. This high-amplitude axial anomaly/magnetization is probably caused by highly magnetized extrusive basalts. The high magnetization of young basalts of 20–30 A m^{-1} decays rapidly within the first few million years. This initial decay is then followed by a slow long-term decay in amplitude with age until magnetization of 3–6 A m^{-1} is reached. This later decay may be due to low temperature and hydrothermal oxidation of the magnetitic mineral magnetite. Over old oceanic crust the main source of magnetic anomalies may be deeper gabbros, rather than the uppermost crust.

3.3 Reconstruction of past plate motions

3.3.1 Introduction

A magnetic-anomaly profile can be used to construct a magnetized block model of the ocean crust and to estimate the latitude and orientation of the mid-ocean ridge which produced it. When several profiles are available and the magnetic anomalies are plotted on a map, as in Fig. 3.7, such a block model is easily visualized. To determine the past movements of plates, a substantial amount of palaeomagnetic data is required. This data collection began with the development of magnetometers, by Bell Telephone Labs and Gulf Oil, to detect submarines. The instruments were required to measure magnetic fields to about 1 nT (about one part in 50 000), which was ideal for measuring marine magnetic anomalies.

Since the oldest oceanic lithosphere is Jurassic ($\sim$160 Ma old), magnetic-anomaly data can only be used to trace the past motions of the plates back to that time. Continental magnetic data and other geological data provide evidence for motions of the plates prior to the Jurassic, but the data are necessarily sparser and more difficult to interpret. The remainder of this chapter provides sections on the geological histories of the Pacific, Indian and Atlantic Oceans as established by deciphering magnetic anomalies. The full reconstruction of the geological history of an ocean is possible only if the ocean contains only ridges. For oceans such as the Indian and Pacific, in which subduction has taken place, there has been a loss of information, which prevents a full reconstruction. Deciphering the geological histories of the oceans, especially the early work by McKenzie and Sclater on the Indian Ocean, was one of the triumphs of plate tectonics.

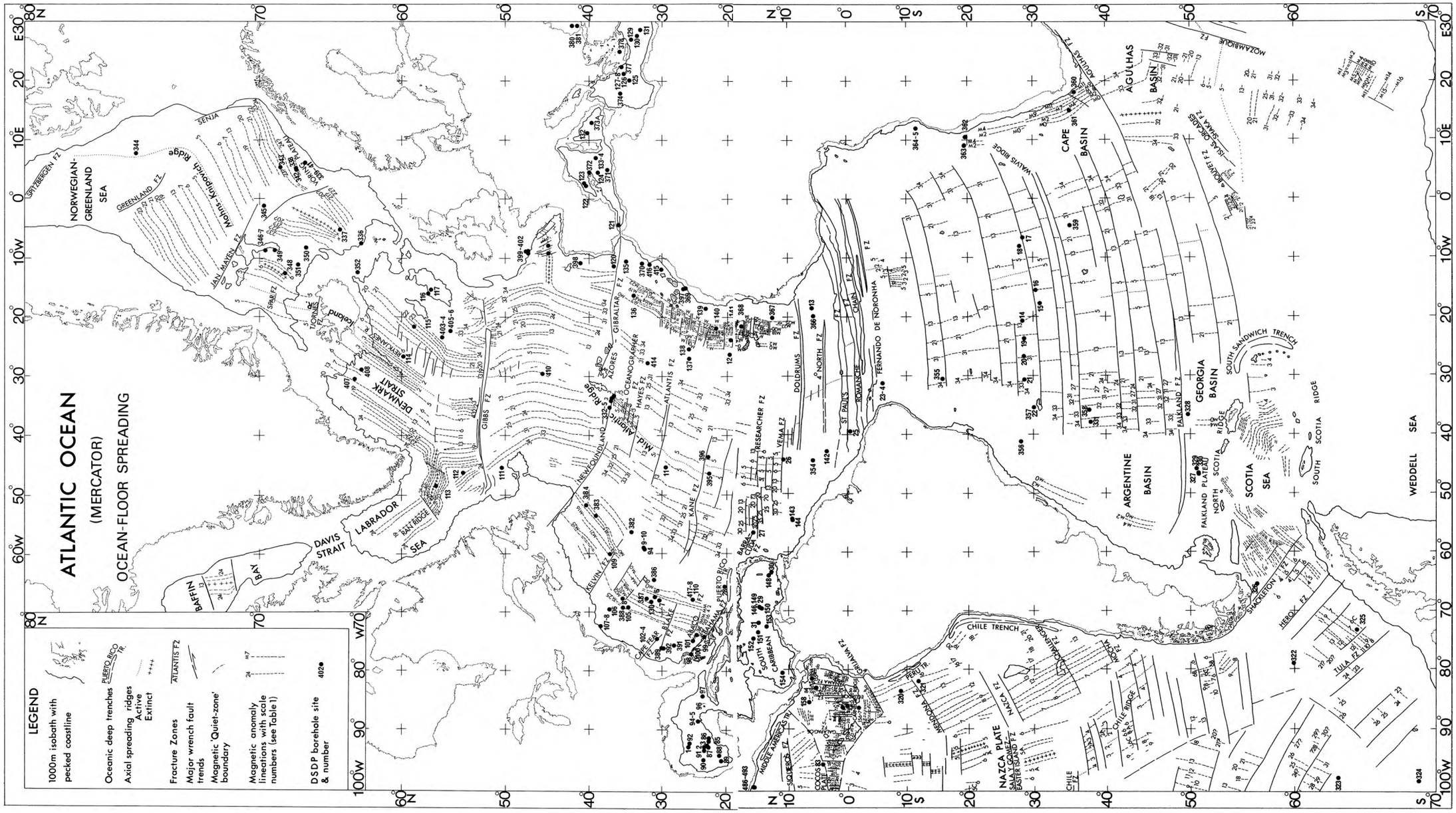


Figure 3.19. Magnetic anomalies, fracture zones and DSDP borehole sites in the Atlantic Ocean. (From Owen (1983).)

3.3.2 The Atlantic Ocean

The magnetic-anomaly map of the Atlantic Ocean is by far the simplest for the three major oceans (Fig. 3.19). The history of continental splitting and seafloor spreading is almost completely preserved because, apart from the short lengths of the Puerto Rico Trench and the South Sandwich Trench, there are no subduction zones. The Mid-Atlantic Ridge is the plate boundary between the Eurasian, African and North and South American plates. Despite the changes in pairs of plates, the poles and instantaneous rotation rates are such that the spreading rate of the Mid-Atlantic Ridge does not vary greatly along its length. The boundary between the North American and the South American plates is best described as a 'diffuse plate boundary' between the Mid-Atlantic Ridge and the Caribbean.

The oldest identified anomaly in the South Atlantic is M11, which occurs just off the west coast of South Africa. Thus, Africa and South America must have started separating shortly before this time (135 Ma). The oldest anomaly in the central Atlantic is M25, identified off the east coast of North America and the northwest coast of Africa. Africa and North America therefore started to separate during the mid Jurassic, probably soon after 180 Ma. This motion resulted in considerable faulting and folding in the Mediterranean region because the early stage of rifting between Eurasia and North America did not begin until much later ($\sim$120–140 Ma). Rifting then proceeded northwards in stages. The Reykjanes Ridge between Greenland and Eurasia started spreading at about 55 Ma. The northwest–southeast anomalies in the Labrador Sea between Canada and Greenland (anomalies 27–19) and extending northwards into Baffin Bay (anomalies 24–13) indicate that there was also an active ridge there from about 60 Ma until 43 or 35 Ma. Since this time, Greenland has not moved independently and has been part of the North American plate: the spreading that started with anomaly 24 has continued only along the Reykjanes Ridge. North of Iceland the Arctic ridge system includes the Kolbeinsey, Mohns, Knipovitch and Gakkel Ridges (Fig. 2.2).

3.3.3 The Indian Ocean

The magnetic-anomaly map of the Indian Ocean (Fig. 3.20) is considerably more complex than that of the Atlantic Ocean. The three present-day mid-ocean ridges – the Central Indian Ridge, Southwest Indian Ridge and Southeast Indian Ridge – intersect at the Indian Ocean (or Rodriguez) Triple Junction (Fig. 2.2), an RRR triple junction. The Southeast Indian Ridge is spreading fairly fast (3 cm yr^{-1} half-rate) and has smooth topography, whereas the Southwest Indian Ridge is spreading very slowly (0.6–0.8 cm yr^{-1} half-rate) and has rough topography and many long fracture zones. The Carlsberg Ridge starts in the Gulf of Aden and trends southeast. At the equator it is intersected and offset by many transform faults, so the net strike of the plate boundary between Africa and

India becomes almost north–south. This part of the plate boundary is called the Central Indian Ridge. The Southwest Indian Ridge, which extends from the Bouvet Triple Junction in the South Atlantic to the Indian Ocean Triple Junction, is the boundary between the African and Antarctic plates. The position of the present-day rotation pole for Africa relative to Antarctica at about 6°N, 39°W (Table 2.1) means that this ridge is offset by a series of very long transform faults. One of the main fault complexes, the Andrew Bain Fracture Zone, offsets the ridge axis by some 500 km. The half-spreading rate in this region is about 0.8 cm yr^{-1}. The African is subdivided into the Nubian and Somalian plates (Sect. 10.4.2). The triple junction between the Nubian, Somalian and Antarctic plates is located within the Andrew Bain Fracture zone at ∼30° E. The other major bathymetric features of the Indian Ocean are the Ninety-East Ridge and the Chagos–Maldive–Laccadive Ridge system. Both of these linear north–south submarine mountain chains are hotspot tracks – of the Kerguelen and Réunion hotspots, respectively.

The oldest magnetic anomalies in the Indian Ocean occur in the Wharton basin (M9–M25), which lies between Australia and the Java Trench, west of Western Australia (M0–M22), between Madagascar and east Africa (M2–M22), in the Somali basin (M13–M21) and north of Antarctica at 20°E (M1–M16). These anomalies indicate that the separation of Africa and Antarctica began by the time of anomaly M21 (150 Ma). At this time, east Antarctica and Madagascar (then joined) moved southwards away from Africa (Fig. 3.21(a)). In so doing, they generated the symmetrical anomalies in the Somali basin, as well as those off the east Antarctic coast and west of Madagascar, which appear to be the northern and southern halves of a symmetrical pattern. At about the time of anomaly M0, seafloor spreading in the Somali basin north of Madagascar stopped and a new ridge was initiated between Madagascar and Antarctica: the proto-Southwest Indian Ridge formed, separating Antarctica from Africa, Madagascar and India. From the time of anomaly 34 (∼90 Ma) until the present, spreading along this ridge appears to have been fairly constant, but magnetic-anomaly data from this ridge are very scarce.

The anomalies north of Australia, to the north and east of the Exmouth Plateau, M9–M25, are only the southern half of a symmetrical pattern; the northern part has been subducted by the Java Trench, which is currently active there. The oldest anomaly close to the continental margin north of Australia, M25, gives an estimate of the date of opening of this ocean as 155 Ma. A northeast–southwest pattern of anomalies (M11–M0) is present in the Bay of Bengal and west of Western Australia there is a symmetrical pattern of anomalies M0–M10. This indicates that a ridge was active here by M10 time (134 Ma). This ridge was the plate boundary between the Indian plate and the Australia–Antarctic plate. Australia and Antarctica were still joined at this stage, though continental extension had been taking place since the separation of India (Fig. 3.21(b)).

At about 96 Ma a major change in the plate motions occurred; and the major part of the Indian Ocean has been created since then. At this time (or shortly

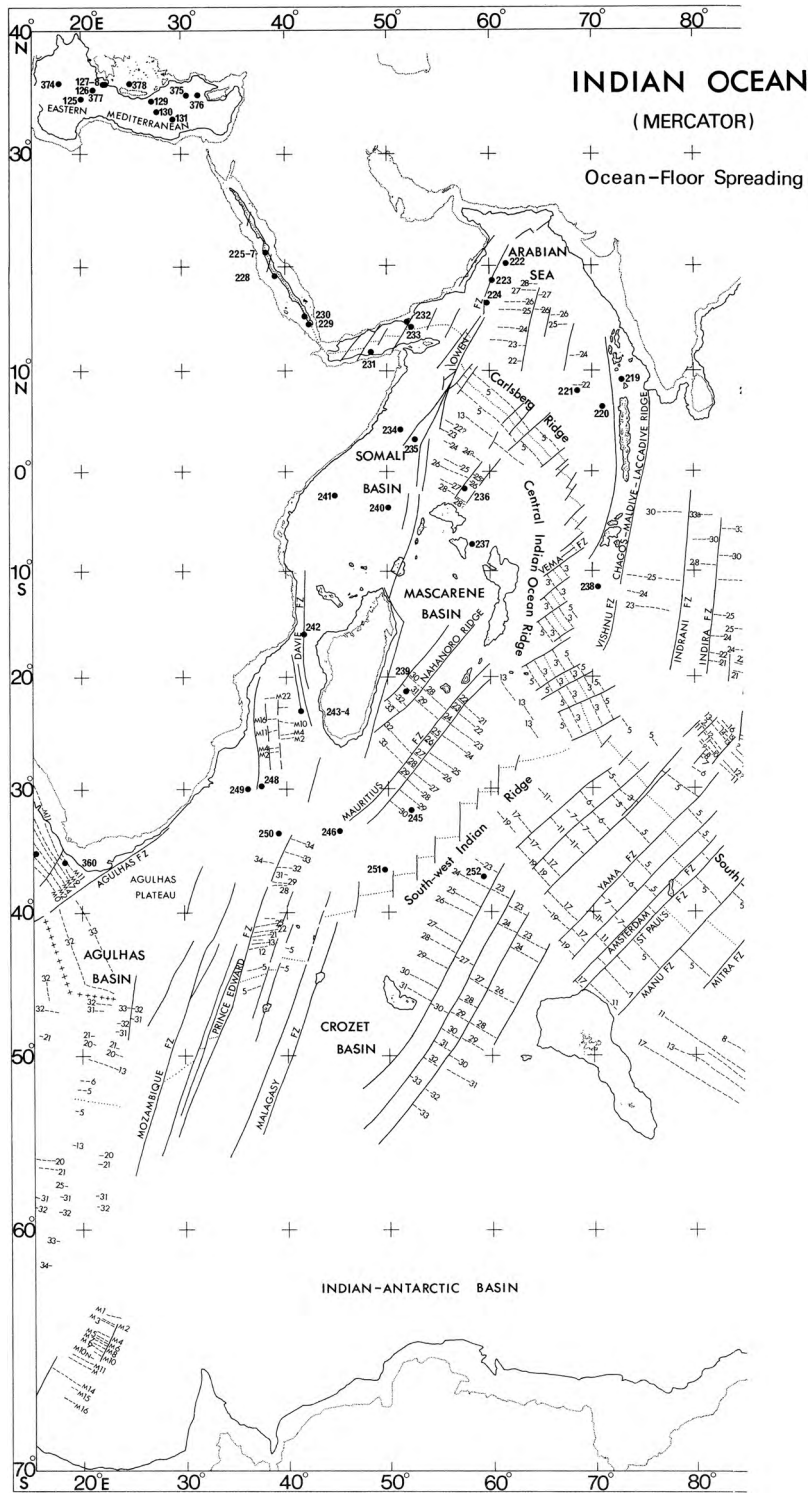
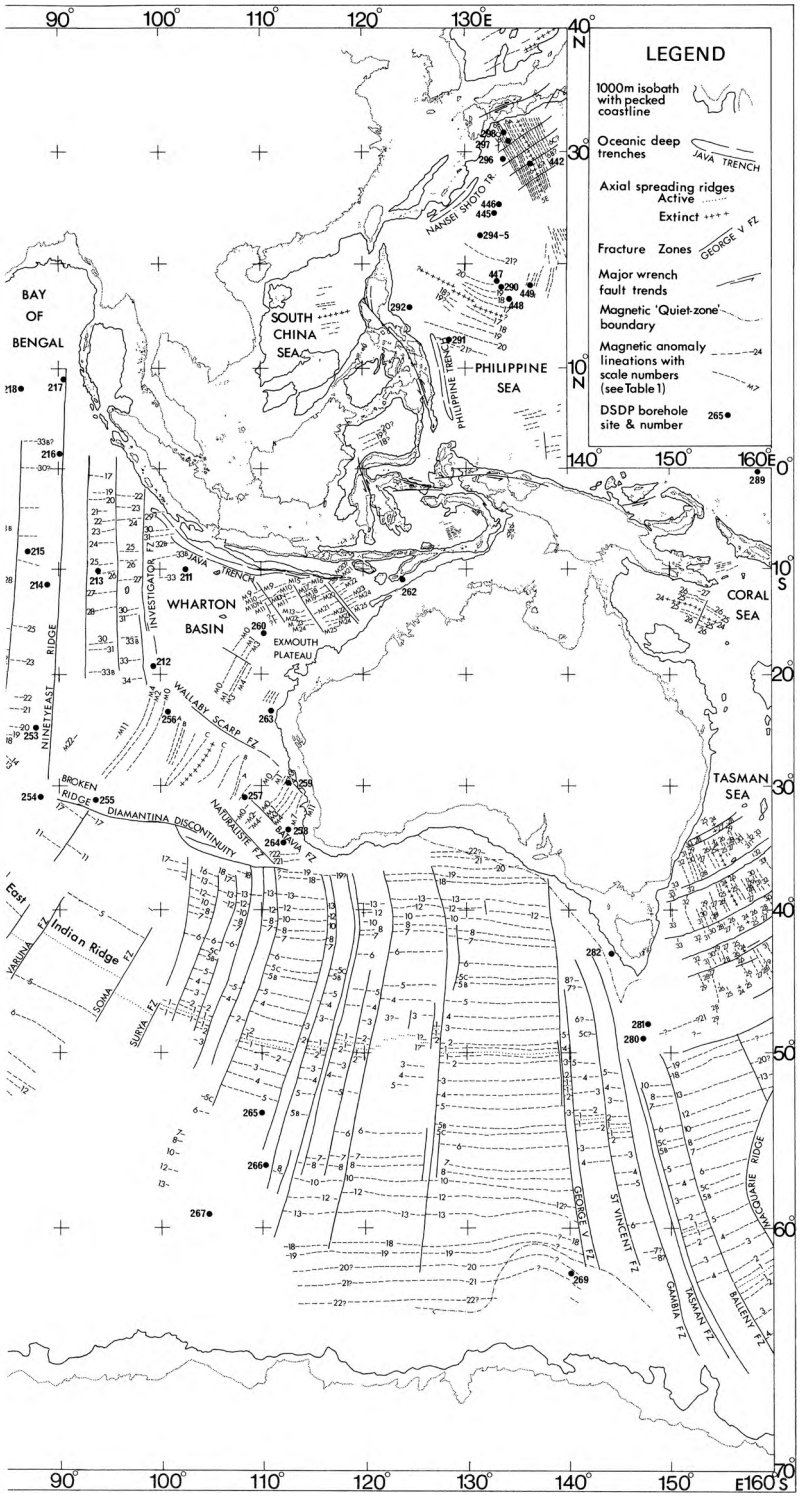


Figure 3.20. Magnetic anomalies, fracture zones and DSDP borehole sites in the Indian Ocean. (From Owen (1983).)


LEGEND
1000m isobath with pecked coastline
Oceanic deep trenches
JAVA TRENCH
Axial spreading ridges
Active
Extinct
Fracture Zones
GEORGE V FZ
Major wrench fault trends
Magnetic 'Quiet-zone' boundary
Magnetic anomaly lineations with scale numbers (see Table 1)
DSDP borehole site & number
BAY OF BENGAL
SOUTH CHINA SEA
PHILIPPINE SEA
PHILIPPINE TRENCH
NANSEI SHOTO TR.
JAVA TRENCH
INVESTIGATOR FZ
WHARTON BASIN
EXMOUTH PLATEAU
CORAL SEA
TASMAN SEA
NINETYEAST RIDGE
WALLABY SCARP FZ
BROKEN RIDGE
DIAMANTINA DISCONTINUITY
NATURALISTE FZ
East Indian Ridge
YARRINA FZ
SOMA FZ
SURYA FZ
ST VINCENT FZ
TASMAN FZ
GARRIA FZ
BALLENY FZ
MACQUARIE RIDGE
90° 100° 110° 120° 130° 140° 150° E160°
40° N 30° 20° 10° N 0° 10° S 20° 30° 40° 50° 60° 70° S

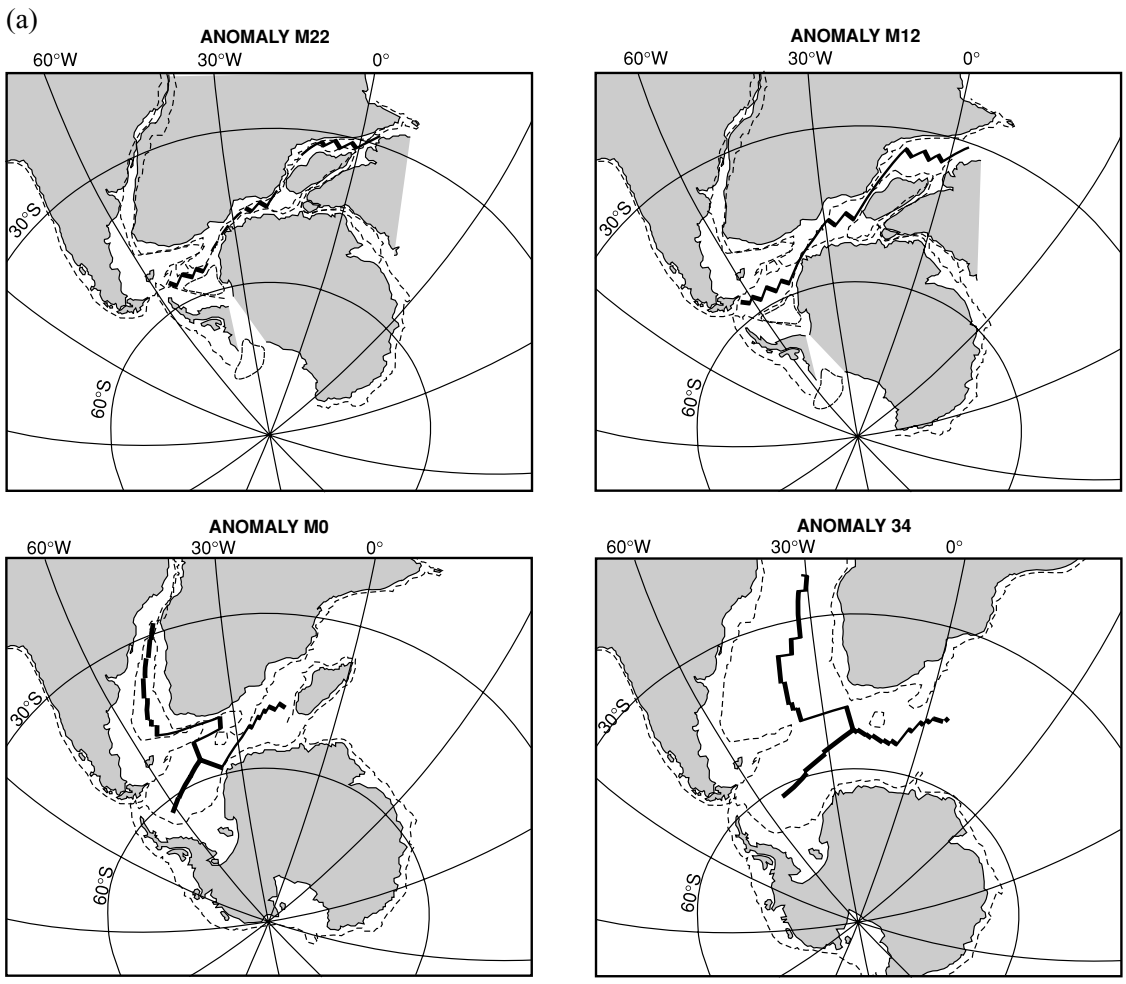


Figure 3.21. (a) Reconstructions of the positions of South America, Africa and Antarctica at the time of anomalies M22, M12, M0 and 34. (From Lawver *et al.* (1985).)

before, as anomaly information is not available for the Magnetic Quiet Zone), the Southeast Indian Ridge started spreading. This began the splitting of Australia from Antarctica. At the same time, India began to move rapidly northwards from Antarctica. The Carlsberg–Central Indian Ridge also started spreading, separating Africa from India. The half-spreading rate of the portion of the Southeast Indian Ridge between Australia and Antarctica was slow (perhaps only $0.5\,\mathrm{cm\,yr^{-1}}$), whereas the spacing of the same anomalies to the south and east of India indicates half-spreading rates of $10\,\mathrm{cm\,yr^{-1}}$ or more for the ridge between India and Antarctica (Fig. 3.21(b)). Figure 3.21(c) illustrates the relationship between the Réunion hotspot and the Deccan Traps, a very major flood-basalt

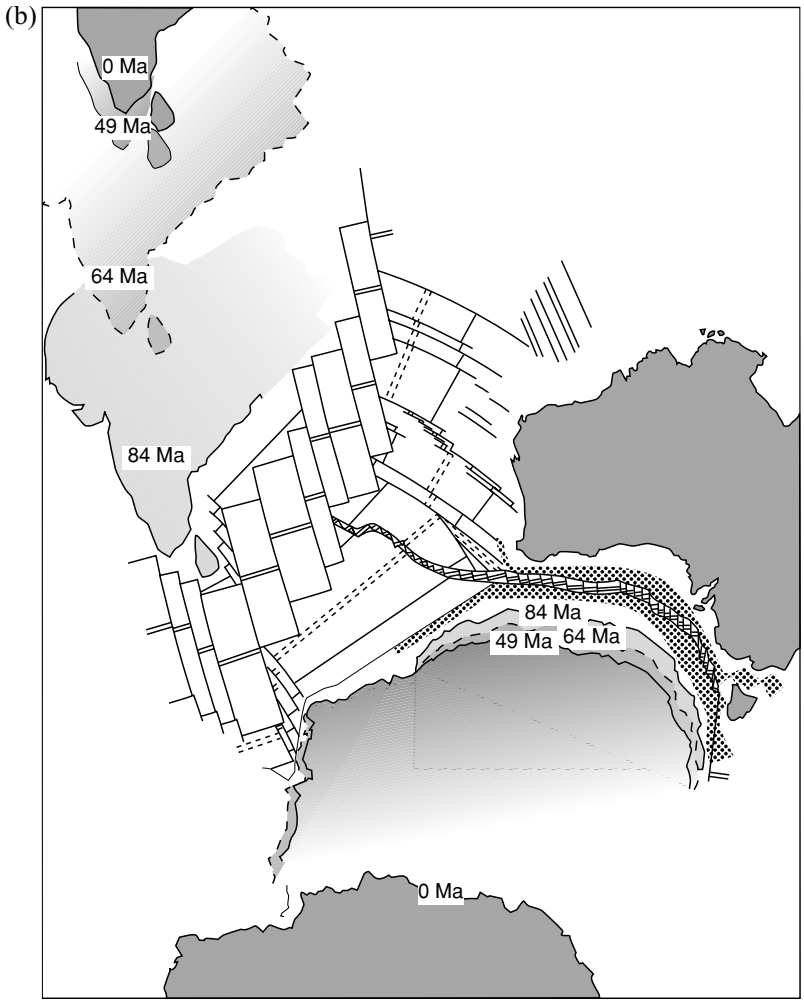


Figure 3.21. (b) Reconstructions of the positions of India and Antarctica relative to Australia 84, 64 and 49 Ma ago. The 84-Ma mid-ocean ridge system is shown with double solid lines. The abandoned 96-Ma mid-ocean ridge system is shown with double dashed lines. The continental extension between Australia and Antarctica is stippled. (After Powell *et al.* (1988).)

province in India. First, at about 65 Ma, the Deccan Traps[5] were erupted in a major and sudden outpouring of lava with a total volume of well over one million cubic kilometres. Then, as India moved northwards, the hotspot formed the Chagos–Maldive–Laccadive Ridge and the Mascarene Plateau. The hotspot track was subsequently split by spreading on the Central Indian Ridge. The volcanically active Réunion Island marks the present position of the hotspot (Fig. 2.19a).

At about the time of anomaly 19 (43 Ma), there was a major reorganization in the plate motions when the Indian and Asian continents collided (refer to

[5] The Deccan Traps, like the Parana in Brazil (at the western end of the Rio Grande Ridge, the Tristan hotspot track) are continental flood basalts. These huge volumes of tholeiitic basalt were erupted very rapidly and were a consequence of decompression melting (sometimes called pressure-release melting) of a rising hot mantle plume head when it reached the rifting continental lithosphere.

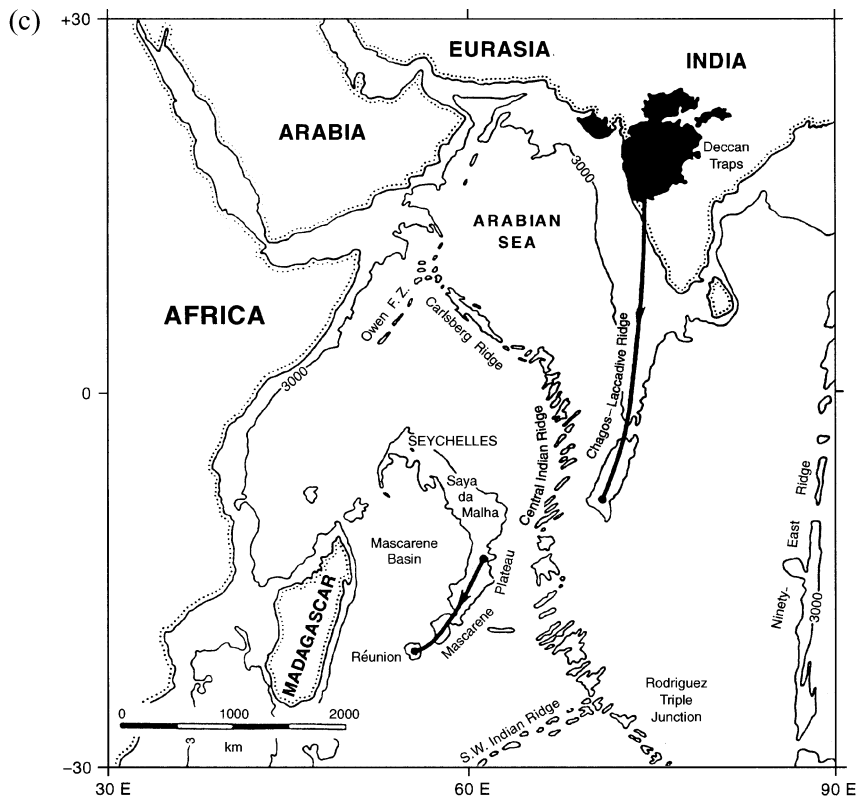


Figure 3.21. (c) The Mascarene Plateau and the Chagos–Maldive–Laccadive Ridge mark the track of the Réunion hotspot from the initial massive outpouring of basalt in the Deccan Traps to its present location beneath Réunion Island. (After White, R. and McKenzie, D. 'Magmatism at rift zones: the generation of volcanic continental margins and flood basalts, *J. Geophys. Res.*, **94**, 7685–729. Copyright 1989 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

Section 10.2.3 for details of continent–continent collisions). It is apparent in Fig. 3.20 that a second major change of spreading direction must have occurred on the Southeast Indian Ridge at this time. Older anomalies strike approximately east–west, whereas the younger anomalies strike northwest–southeast. Therefore, prior to the time of anomaly 19, the Southeast Indian Ridge was striking east–west and the east–west magnetic anomalies in the Arabian Sea, south of India and in the Bay of Bengal were all formed by the same ridge. Almost all the lineations formed on the north side of that section of the ridge which lay to the east of the Ninety-East Ridge (a hotspot trace) as well as the extinct ridge itself have been subducted by the Java Trench, the only subduction zone in the Indian Ocean. The ridge between Australia and Antarctica continued to spread very slowly. The fate of the ridge between Australia and India is not clear. It is possible that India and Australia then lay on the same plate. There has been only slight motion between India and Australia since about the time of anomaly 13. Very broad 'diffuse' plate boundaries subdivide the Indian plate into the Indian, Australian and Capricorn plates, with the Capricorn plate lying to the east of the Central Indian Ridge between 10°S and 30°S and extending east to ~80°E.

The fine details of the motions of the plates in this region and the exact positioning of the continents as they were prior to anomaly M25 when they formed the mega-continent Gondwanaland remain to be firmly established, but on a broad scale the motions are fairly well understood.

3.3.4 The Pacific Ocean

The magnetic-anomaly pattern off the west coast of North America was the first magnetic-anomaly pattern to be studied in detail (Fig. 3.7). An interpretation of these anomalies based on the plate motions and earthquakes in the area is shown in Fig. 3.22. The details of the plate boundaries and their relative motions and ages are much clearer. Although the ridges in this region are spreading relatively slowly, the plates are small, so the lithosphere was only about 10 Ma old when it was subducted beneath the North American plate. Another feature of the magnetic anomalies in this region is the difference of some 20° between the present-day trend of the Juan de Fuca Ridge and the strike of the older anomalies on the Pacific plate (the southwest part of Fig. 3.22(a)). This difference is caused by changes in the rotation pole and subsequent reorientation of the ridge. The diagonal *pseudofaults* which offset the magnetic anomalies in this region are also due to the adjustment of the ridge system to changes in the rotation poles (Fig. 3.23). Thus, when an area is studied in detail, the original questions may well be answered and theories validated, but usually new questions are also raised (in this instance, the new problem is the exact method by which ridges adjust to changes in rotation poles, or vice versa). The past plate motions in other parts of the Pacific are more difficult to interpret than those in the region flanking North America, where the presence of an active ridge system means that both sides of the anomaly pattern are preserved.

Further to the south and west on the Pacific plate, the oceanic anomaly pattern is, on a broad scale, fairly simple (Figs. 3.24 and 3.28). Anomalies strike almost north–south and are offset by fracture zones (see Section 9.5.1). The central part of the ocean was formed during a period which included the Magnetic Quiet Zone. Thus, there are not many anomalies to be observed. However, much farther north towards the Aleutian islands the pattern changes. The anomalies change direction so that they are striking approximately east–west. This feature is called the Great Magnetic Bight. The other main feature of the northern Pacific is that the north–south anomalies represent only the western half of the pattern and, except for the short ridge segments such as the Juan de Fuca Ridge, the mid-ocean ridge that created the oceanic plate no longer exists. This vanished ridge has been subducted under the North American plate. With it went much of the Farallon plate, the name given to the plate which once was to the east of the ridge and had the matching half of the symmetrical anomaly pattern (Fig. 3.24). For the Farallon plate and the Pacific–Farallon Ridge to have been subducted in this manner, the rate of subduction must have been greater than the rate at which the ridge

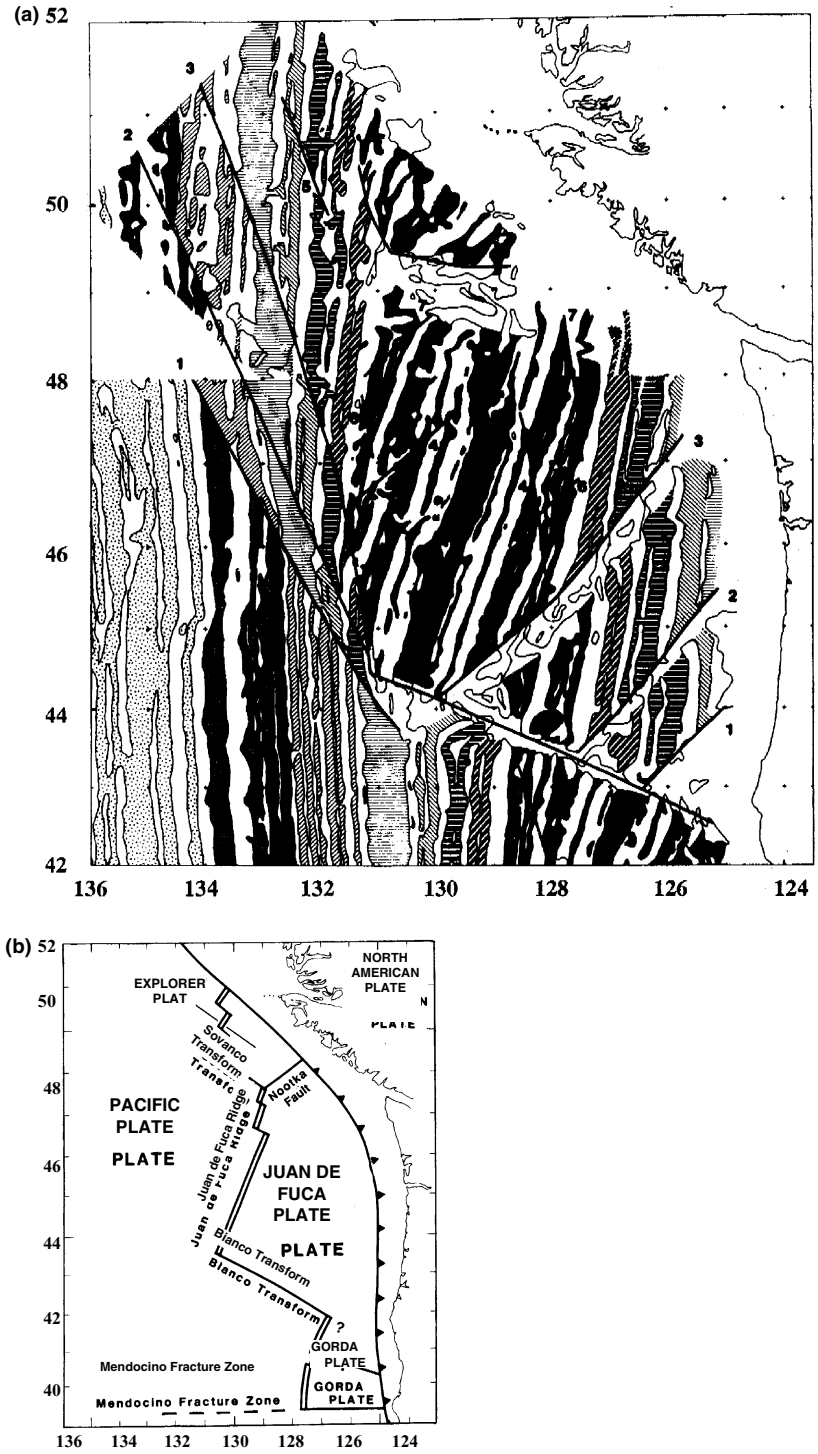


Figure 3.22. (a) Magnetic-anomaly data for the northeastern Pacific Ocean. Solid black lines are major offsets in the anomaly pattern (termed *pseudofaults*). This figure, which includes the data shown in Fig. 3.7, has been shaded so that the ages of the various anomalies stand out. (b) The location map for (a), showing the plates and plate boundaries. Juan de Fuca and Explorer plates are undergoing oblique subduction beneath the North American plate. There is left-lateral strike–slip motion along the Nootka fault and right-lateral strike–slip motion along the Queen Charlotte Fault (see Fig. 3.27). (After Wilson *et al.* (1984).)

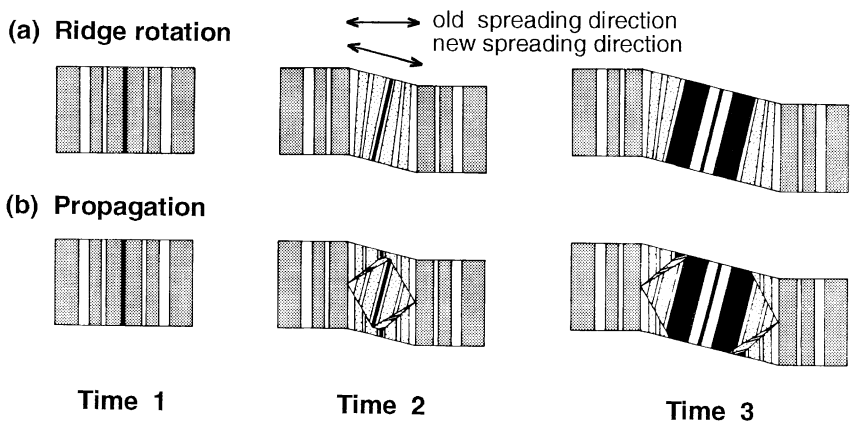


Figure 3.23. Magnetic-anomaly patterns for two possible mechanisms by which a mid-ocean ridge could adjust to changes in its spreading direction caused by a change in rotation pole. (a) Ridge rotation. In this model the spreading at the ridge is asymmetrical during the adjustment time. Half-spreading rates are not equal on both sides of the ridge and vary along the length of the ridge. (b) Propagating rifting. In this model a section of the ridge is assumed to jump to its new orientation. This 'propagating' ridge segment then lengthens and pseudofaults develop at oblique angles to the ridge as observed in Fig. 3.22(a). (From Wilson *et al.* (1984).)

was spreading. All that remains are short segments of the ridge and fragments of the Farallon plate, now the Juan de Fuca, Explorer, Gorda, Cocos and Nazca plates.

The change in direction of the magnetic anomalies in the Great Magnetic Bight region indicates that a third plate was involved (Fig. 2.16 shows that three ridges meeting at a triple junction would produce such an anomaly pattern). This third plate has been named the Kula plate. The Pacific, Farallon and Kula plates are thus assumed to have met at an RRR triple junction (the Kula Triple Junction). The spreading rates and directions of these ridges have to be determined from the anomalies and fracture zones in the vicinity of the Great Magnetic Bight. The Kula plate has been subducted beneath the North American plate and so no longer exists, athough a very small piece may be trapped on the Pacific plate in the western Aleutian arc. Both the other ridges have also been subducted. Putting all of this information together to determine the motions of the plates in the northern Pacific region for the last 80 Ma involves much spherical geometry and computing. An idealized flat-plate model illustrating the main features of the evolution of the northeastern Pacific was shown in Chapter 2, Problem 2. Figure 3.25 shows a reconstruction of the evolution of this region from 50 Ma until the present time. From 80 Ma until about 55 Ma there were four plates in this northern Pacific region (North America, Kula, Farallon and Pacific). Two of them were being subducted beneath the North American plate: the Kula plate in the north and the Farallon plate farther south. About 55 Ma ago the northern

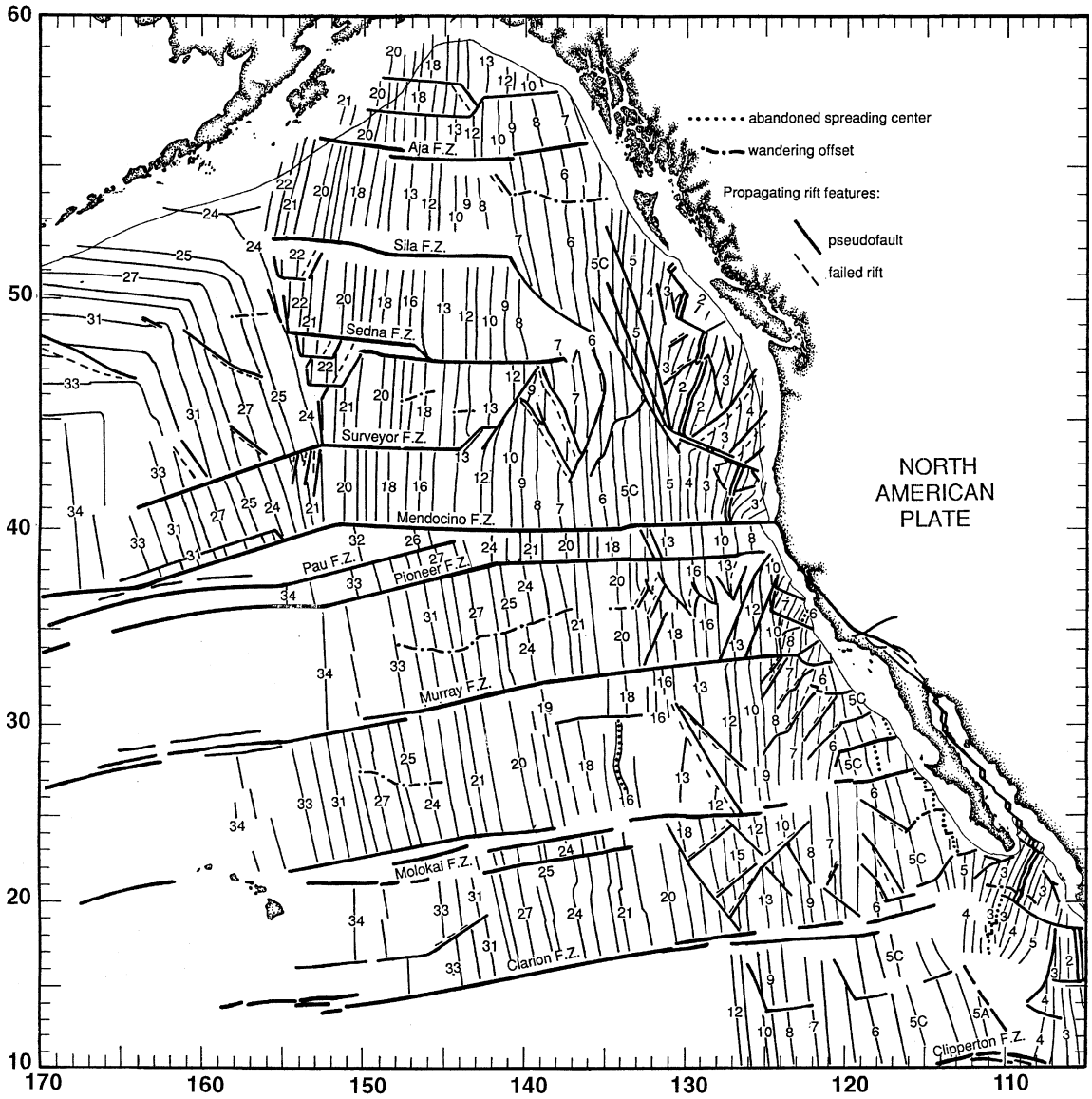


Figure 3.24. Magnetic anomalies in the northeastern Pacific. Numbers are anomaly numbers (Fig. 3.14), not ages. (From Atwater (1989).)

part of the Farallon plate broke off to form the Vancouver plate. The present Juan de Fuca plate is the remnant of this plate. The location of the boundary between the Vancouver and Farallon plates was rather variable but the relative motion was slow oblique compression and subduction of the Vancouver plate beneath North America continued to take place. At about 30 Ma, the situation changed when the ridge between the Farallon and Pacific plates first reached the

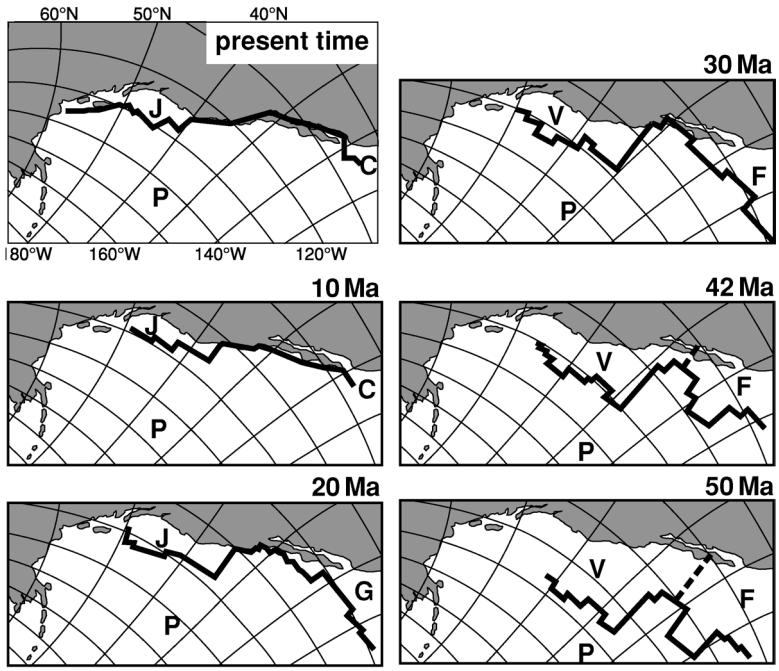


Figure 3.25. Reconstructions of the positions of the Pacific plate and the Farallon plate with respect to the North American plate (shaded). P, Pacific plate; F, Farallon plate; V, Vancouver plate; J, Juan de Fuca plate; G, Guadalupe plate; and C, Cocos plate. (From Atwater (1989), after Stock and Molnar (1988).)

subduction zone. This meant that subduction ceased on that part of the North American plate boundary where the North American plate was adjacent to the Pacific plate as the relative motion between the Pacific and North American plates was parallel to the boundary. Thus, the San Andreas Fault and the Mendocino Triple Junction were born. By 10 Ma, the Farallon–Guadalupe–Cocos plate was very small, and the San Andreas Fault system had lengthened. At about 9 Ma and 5 Ma, as discussed earlier, the strike of the Juan de Fuca Ridge changed by some 20° in total, resulting in the present configuration of the plates. The geological evolution of western North America was controlled by the motions of these oceanic plates. If the relative motion between the Pacific and North American plates is assumed to have been parallel to the subduction zone between the Farallon and North American plates (the present-day Cascadia Subduction Zone), the Mendocino Triple Junction, where the Farallon, North American and Pacific plates meet, must have been stable (Figs. 2.16, 3.26 and 3.27). However, since the Cascadia Subduction Zone is not, at present, collinear with the San Andreas Fault, the triple junction is unstable and may well always have been unstable (Fig. 3.27). The evolution of this unstable triple junction is shown in Fig. 3.26(e). If the three plates are assumed to be rigid, a hole must develop. Such a hole would presumably fill with rising mantle material from below and sediments from above. It would become in effect a microplate. Alternatively, if the continental North American plate is allowed to deform, then the triple junction evolves as shown in Fig. 3.26(f), with internal deformation involving both extension and rotation over

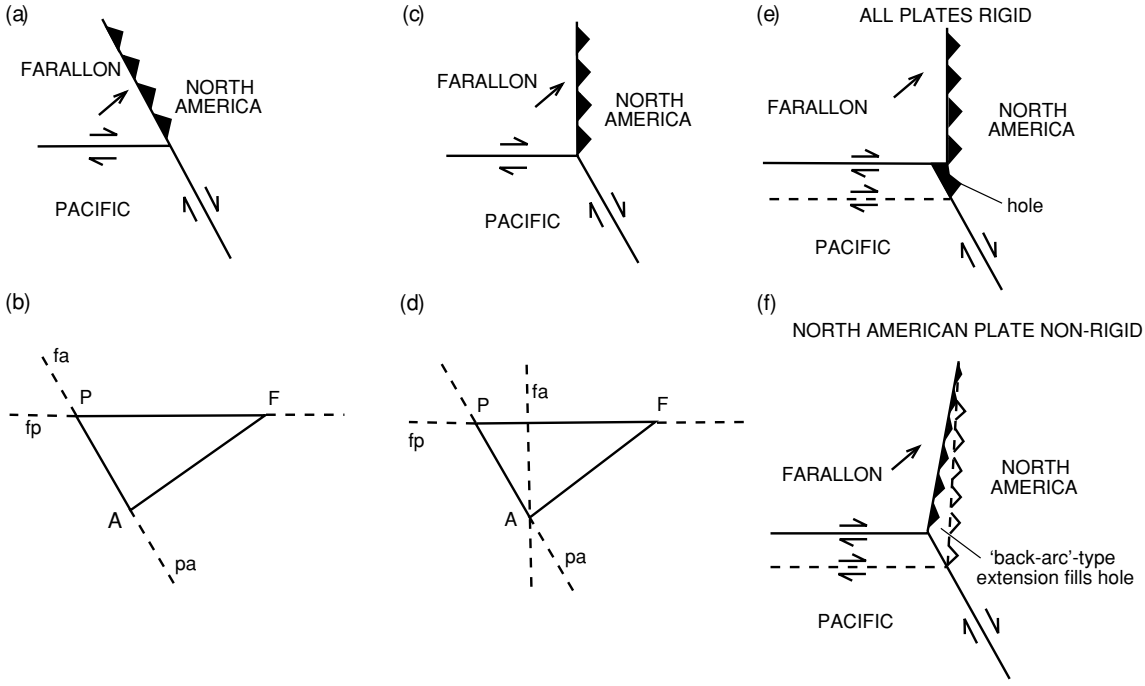


Figure 3.26. Evolution of the Mendocino Triple Junction (prior to change in Juan de Fuca rotation pole). (a) A stable Mendocino Triple Junction with the subduction zone and the fault being collinear, as in Fig. 3.25. (b) Relative-velocity vectors for the geometry in (a). (c) An unstable Mendocino Triple Junction; the subduction zone and the fault are not collinear. (d) Relative-velocity vectors for the geometry shown in (c). (e) Evolution of the triple junction shown in (c) when the three plates are assumed to be rigid. (f) Alternative evolution of the triple junction shown in (c) if the two oceanic plates are assumed to be rigid but the continental (North American) plate is allowed to deform. (From Ingersoll (1982).)

a wide zone to the east of both the subduction zone and the fault. The deformations which would be produced by this process are identical to those indicated by continental palaeomagnetic data and would account for the regional extension in the western U.S.A. (which began about 30 Ma ago when the Mendocino Triple Junction was formed), as well as for the eastward stepping of the San Andreas Fault with time, which has transferred parts of coastal California from the North American plate to the Pacific plate. This eastward stepping of the plate boundary has effectively been accomplished by the capture of sections of the partially subducted Farallon plate by the Pacific plate. This capture of subducted parts of oceanic plate subjected the overriding North American continental margin to distributed shear and extensional forces. Utimately parts of the overriding margin joined the Pacific plate and the plate boundary stepped eastwards, with regions, such as the Western Transverse Ranges, being caused to rotate and zones of extension developing as a result. Baja California has rotated clockwise and

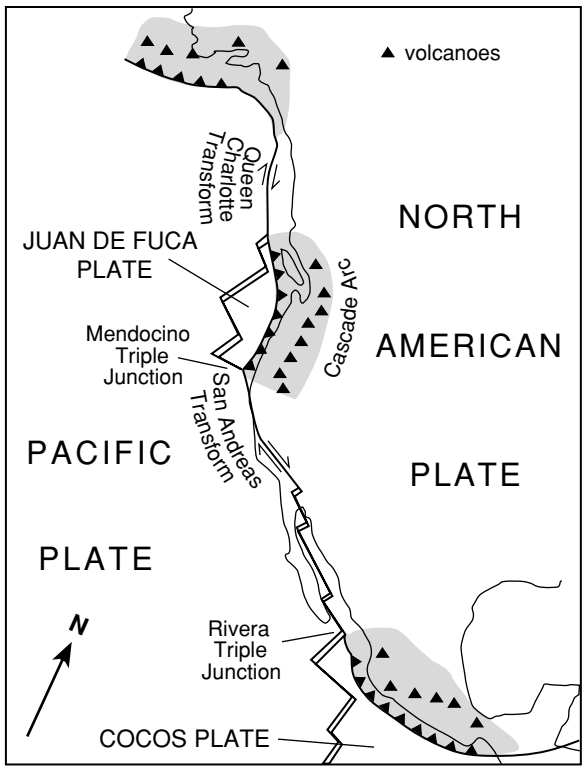


Figure 3.27. A present-day tectonic map of the western part of the North American plate. Shaded regions mark the extent of expected arc magmatism. (After Dickinson (1979).)

been translated northwards since its initial rift from North America at about 12–14 Ma. Details of the past plate motions and their effect on the interior of the North American plate will no doubt continue to be revised for many years, but the main features of the model as presented here are probably nearly correct.

Farther south, to the west of the Middle Americas Trench, magnetic lineations striking east–west are sandwiched between north–south anomalies (Fig. 3.28). This region has undergone a series of plate reorganizations since 30 Ma when the Pacific–Farallon Ridge first intersected the North American Subduction Zone. In general terms these reorganizations can best be described as the breaking up of large plates into smaller plates as the ridge was progressively subducted. At 30 Ma one plate (Farallon) lay to the east of the ridge. At 25 Ma the east–west anomalies indicate that the Cocos–Nazca Ridge (otherwise known as the Galapagos spreading centre) started spreading; so two plates, the Nazca plate and the Cocos or Guadalupe plate, lay to the east of the ridge. At about 12 Ma, the Cocos plate subdivided, spawning the tiny Rivera plate; and now three plates lie to the east of the ridge we call the East Pacific Rise (Fig. 2.2). A schematic flat model of the plates in this region was presented in Chapter 2, Problem 10.

The magnetic lineations in the northwestern Pacific region are much older and more complex than those of the northeastern Pacific (Fig. 3.28). As plates

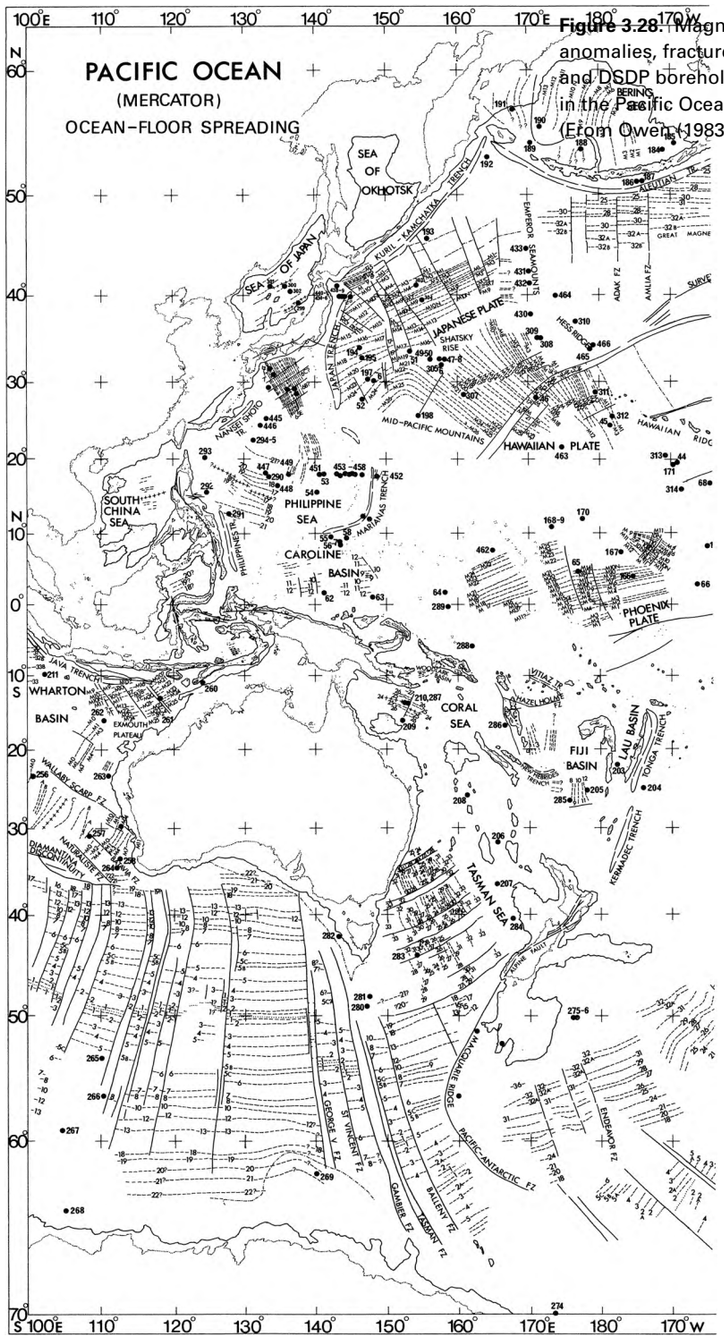
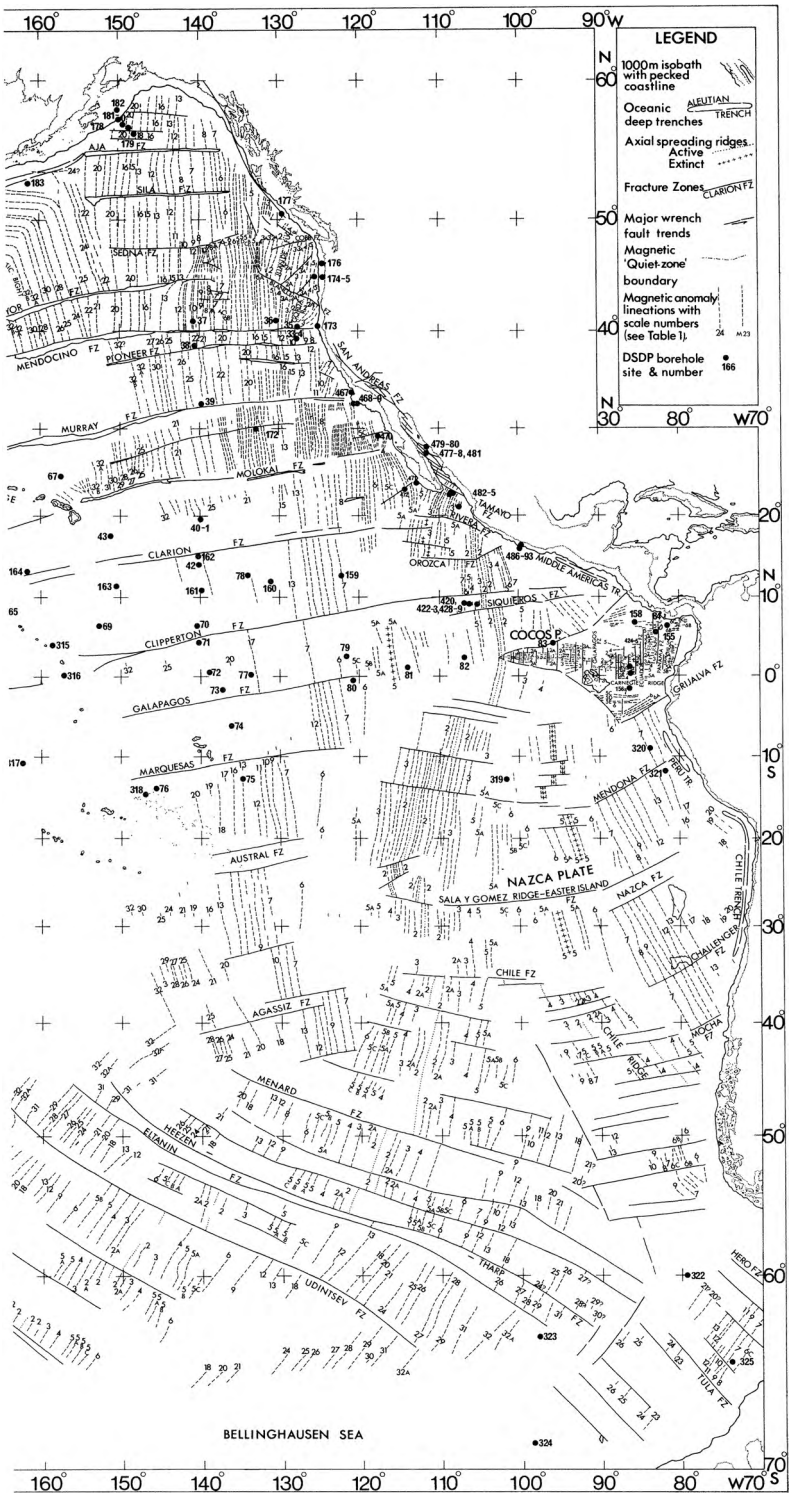


Figure 3.28. Magnetic anomalies, fracture zones and DSDP borehole sites in the Pacific Ocean. (From Owen (1983).)


LEGEND
1000m isobath with pecked coastline
Oceanic deep trenches
ALEUTIAN TRENCH
Axial spreading ridges.. Active
Extinct
Fracture Zones
CLARION FZ
Major wrench fault trends
Magnetic 'Quiet-zone' boundary
Magnetic anomaly lineations with scale numbers (see Table 1).
24
M23
DSDP borehole site & number
166
NAZCA PLATE
COCOS P.
BELLINGHAUSEN SEA
MENDOCINO FZ
MURRAY FZ
MOLOKAI FZ
CLARION FZ
CLIPPERTON FZ
GALAPAGOS FZ
MARQUESAS FZ
AUSTRAL FZ
SALA Y GOMEZ RIDGE-EASTER ISLAND FZ
NAZCA FZ
CHILE FZ
AGASSIZ FZ
MENARD FZ
HEEZEN FZ
ELIANIN FZ
UDINTSEV FZ
THAER FZ
MENDANA FZ
GRIJALVA FZ
SIQUEIROS FZ
OROZCO FZ
RIVERA FZ
TAMAYO FZ
MIDDLE AMERICA TR
PERU TR
CHILE TRENCH
CHALLENGER FZ
MOCHA FZ
HERO FZ
TUCA FZ
SAN ANDREAS FZ
PIONEER FZ
SEDNA FZ
SIBA FZ
AJA FZ
160° 150° 140° 130° 120° 110° 100° 90° W
70° W 80° 30° N
60° N 50° 40° 20° 10° N 0° 10° S 20° 30° 40° 50° 60° 70° S

become older, tectonic reconstructions generally become increasingly difficult and subject to error and, frequently, to speculation. The southwest–northeast lineations extending from the Japan trench towards the Aleutian trench are called the Japanese lineations. The northwest–southeast lineations to the south and west of the Emperor Seamounts and the Hawaiian Ridge are called the Hawaiian lineations. Both the Japanese and the Hawaiian lineations are identified as M1 to M29 inclusive. The east–west lineations that straddle the equator and extend from 160°E to 170°W are the Phoenix lineations (so named because of their proximity to the Phoenix Islands).

The Japanese and Hawaiian lineations form a well-defined magnetic bight, older than, but otherwise very similar to, the Great Magnetic Bight in the north-eastern Pacific. This older magnetic bight is thought to have been formed by an RRR triple junction, where the Farallon and Pacific plates met a third plate, the Izanagi plate (Fig. 3.29).

The Phoenix lineations are from anomalies Ml to M25 inclusive and were produced, between about 127 and 155 Ma, by an east–west striking Phoenix–Pacific Ridge possibly some 40° south of their present latitude. It has been suggested that this ridge extended far to the west and joined the ridge system of the Indian Ocean that gave rise to the magnetic anomalies north of the Exmouth Plateau (Section 3.3.3). Any younger lineations and the symmetrical southern half of these Phoenix lineations are no longer present here.

There are two magnetic bights in the southern Pacific, both less well defined than their northern counterparts. The first is at the eastern end of the Phoenix lineations where the lineations (M12–M4, 137–131 Ma) bend from approximately east–west to northwest–southeast. These northwest–southeast anomalies were probably produced by the Pacific–Farallon Ridge. The second magnetic bight is at about 40°S, 145°W, where anomalies 32–20 (45–70 Ma) bend from approximately southwest–northeast to northwest–southeast. Unfortunately, much of the evidence required to determine the details of these western Pacific plates and their relative motions has been swallowed by the hungry western Pacific subduction zones.

Figure 3.29 shows a series of reconstructions of the plates in the Pacific from 110 Ma to the present. The Magnetic Quiet Zone lasted from 124 to 84 Ma, so not much can be deduced about the motions of the plates during that period. The Kula plate came into existence then. It is not possible for the Kula plate to be the older Izanagi plate; rather, it is believed to have been a piece that broke off either the Farallon plate or the Pacific plate. The Farallon plate was very large indeed: between about 85 and 55 Ma ago the Pacific–Farallon Ridge extended for some 10 000 km.

The evolution of the western and southern Pacific was clearly very complex, with a number of changes in spreading centres. The process probably included the creation and subduction of whole plates. No doubt further detailed mapping of

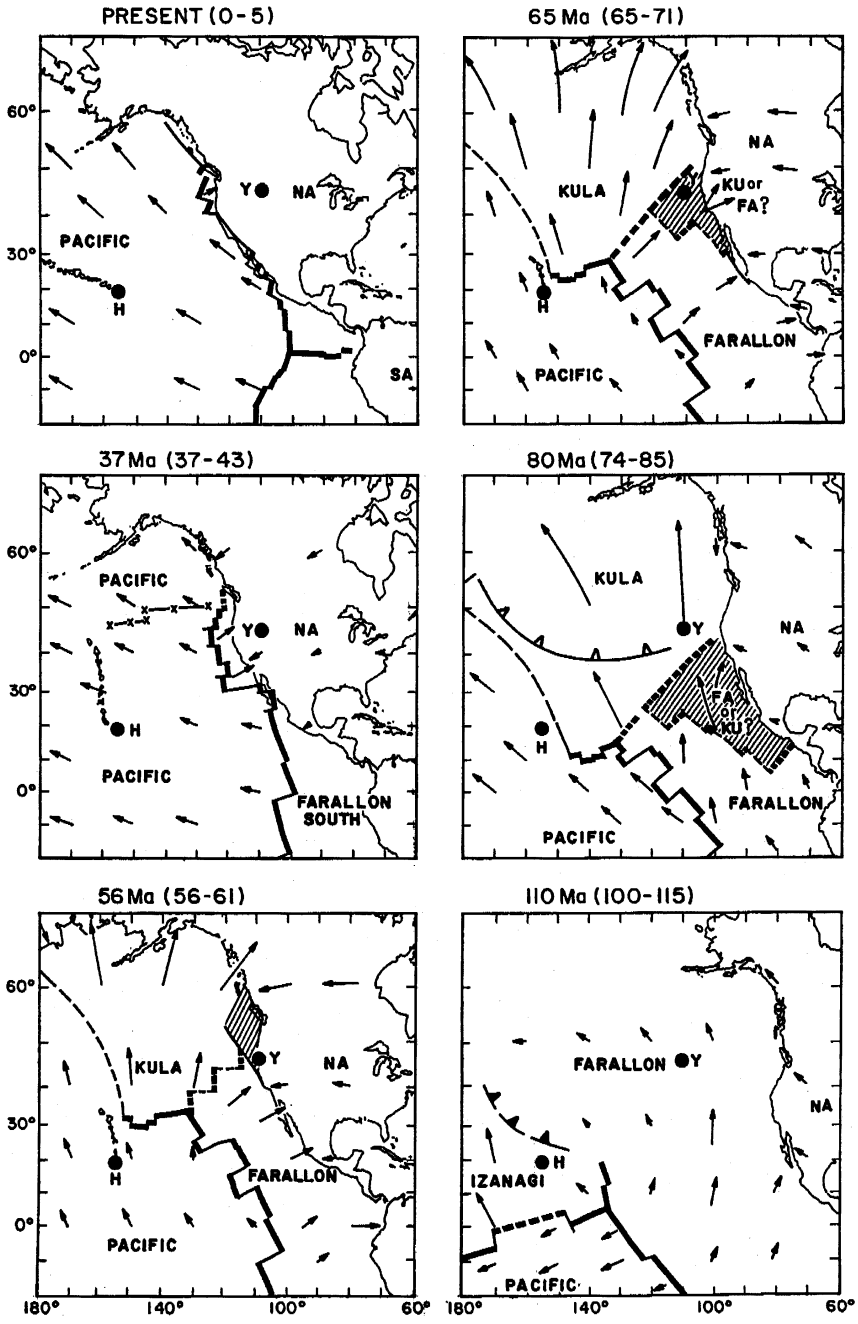


Figure 3.29. Reconstruction of the motions of the Pacific, Izanagi, Farallon, Kula and North American (NA) plates. Arrows indicate motion of the plates with respect to the hotspots. The hatched region at 56–80 Ma indicates the range of possible locations for the ridge between the Kula and Farallon plates. Y, Yellowstone hotspot; and H, Hawaiian hotspot. (From Atwater (1989).)

the seafloor, the lineations, palaeomagnetic measurements and dating of drilled and dredged basement samples will slowly refine and improve the picture of the history of the Pacific back to Jurassic time.

3.3.5 The continents

Figure 3.30, a series of snapshots of the continents, shows how they have moved relative to each other through the Phanerozoic. By the late Carboniferous to earliest Permian, the continents were all joined together and formed one supercontinent, which we call *Pangaea* (Greek, 'all lands'). The northern part of Pangaea, comprising today's North America, Greenland and Eurasia, has been named *Laurasia* or *Laurentia*, and the southern part of the continent, comprising South America, Africa, India, Antarctica and Australia, is called *Gondwanaland*. Laurasia and Gondwanaland became distinct during the Jurassic when Pangaea rifted in two. The wedge-shaped ocean between Laurasia and Gondwanaland is the *Tethys* Sea. (Tethys was the wife of Oceanus in Greek mythology.) It is this sea which has presumably been subducted beneath Laurasia as India and Africa have moved northwards. The Mediterranean, Caspian and Black Seas are the last vestiges of this ancient ocean, the completion of whose subduction resulted in the building of the Alpine, Carpathian and Himalayan mountain chains. All these regions contain scattered outcrops of ophiolites (a suite of rocks with chemical and lithological similarities to the oceanic crust, which may be examples of crust from ancient back-arc basins, Section 9.2.1). There is considerable current debate about the reconstruction of continental fragments prior to the Pangaea supercontinent (uncertainty increases with age) and there are many differing views on the arrangements of the current continental pieces into pre-Pangaean supercontinents. Rodonia was a Mid-to-Late Proterozoic (750–∼1000 Ma) supercontinent. It is proposed that, in the Late Proterozoic, it broke up into Pannotia, Siberia and North China. Pannotia then further split into Laurentia (the Precambrian core of North America), Gondwanaland and Baltica. It is proposed that, at about the same time, there was a major global glaciation ('snowball' Earth). The oldest reconstructions shown in Fig. 3.30, which are largely based on Dalziel (1997), are presented only as one example of what Precambrian geography may have been like, rather than as absolute fact.

Figure 3.30. Palaeocontinental maps showing the present-day continents in their previous positions. ANT, Antarctica; GR, Greenland; IND, India; M, Madagascar; NOAM, North America; and SOAM, South America. Solid lines, fracture zones and magnetic lineations. Dark grey shading, large igneous provinces (LIPs) volcanics produced at hot spots. (Reconstructions provided by Kylara Martin of the PLATES Project, Institute for Geophysics, University of Texas at Austin, Lawver *et al.* 2003.)

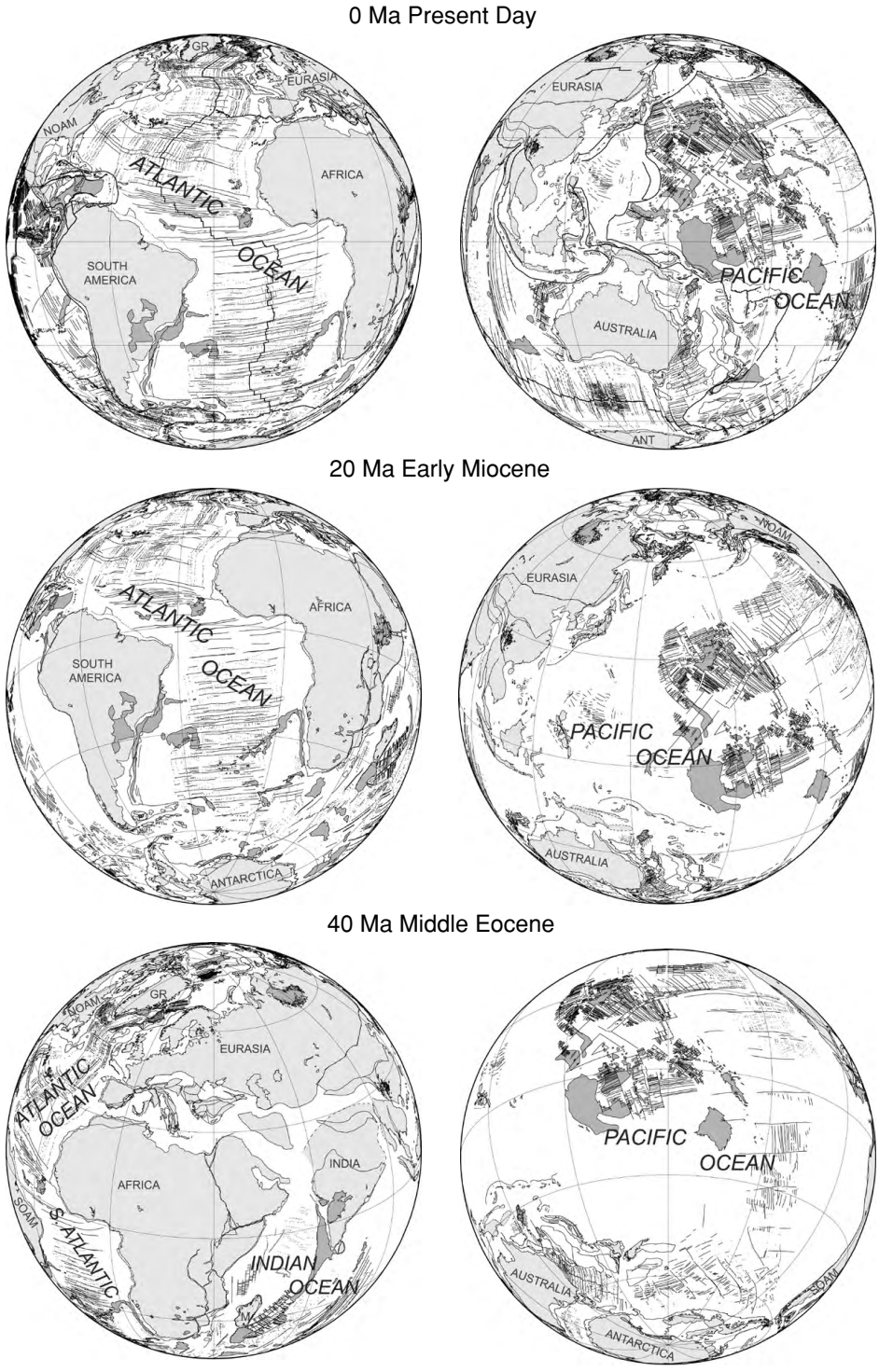

0 Ma Present Day
GR
EURASIA
NOAM
ATLANTIC
OCEAN
AFRICA
SOUTH
AMERICA
EURASIA
PACIFIC
OCEAN
AUSTRALIA
ANT
20 Ma Early Miocene
ATLANTIC
OCEAN
AFRICA
SOUTH
AMERICA
ANTARCTICA
NOAM
EURASIA
PACIFIC
OCEAN
AUSTRALIA
40 Ma Middle Eocene
NOAM
GR
EURASIA
ATLANTIC
OCEAN
INDIA
AFRICA
SOAM
S. ATLANTIC
INDIAN
OCEAN
PACIFIC
OCEAN
AUSTRALIA
SOAM
ANTARCTICA

Figure 3.30. (*cont.*)

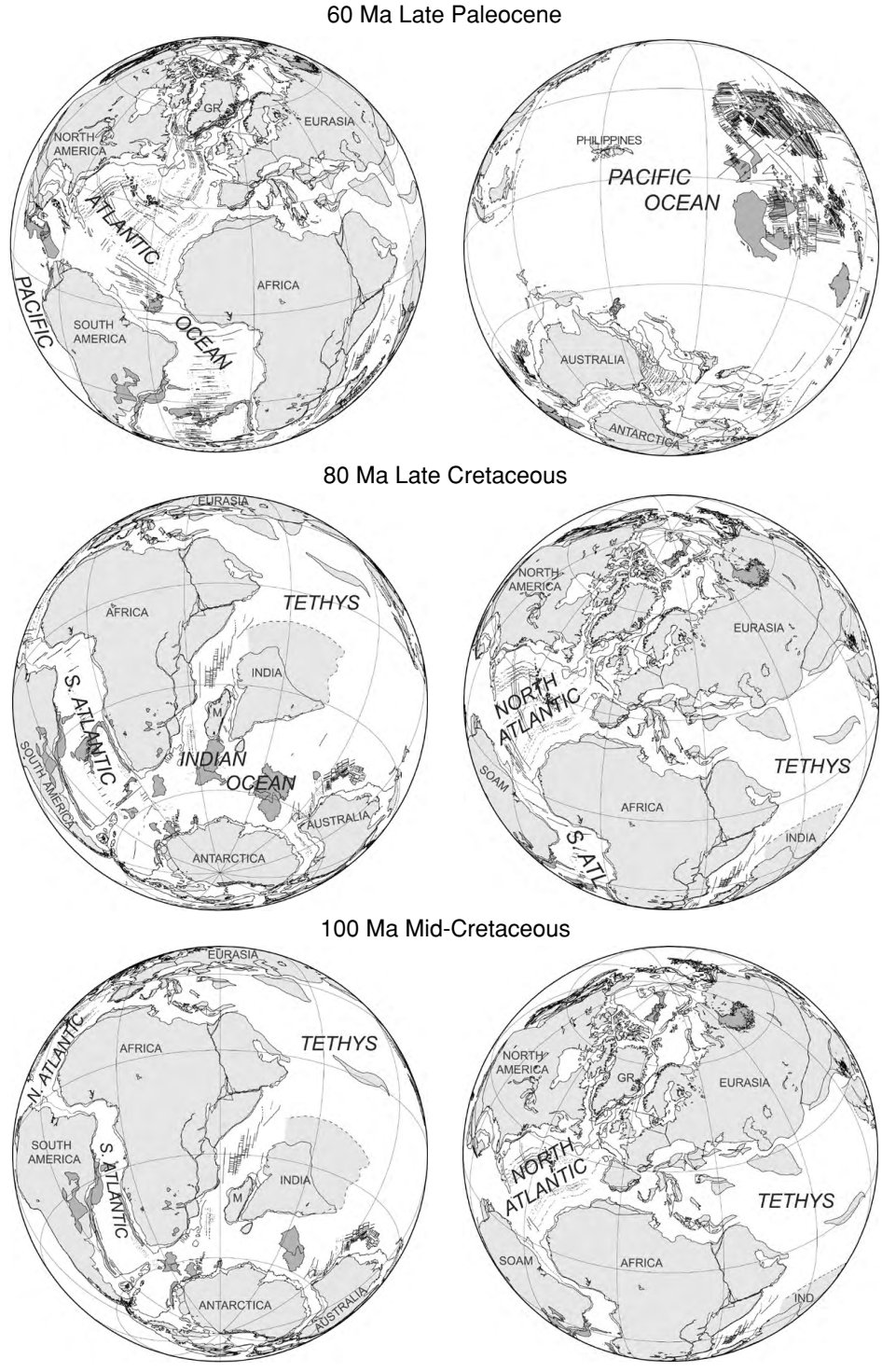


Figure 3.30. (*cont.*)

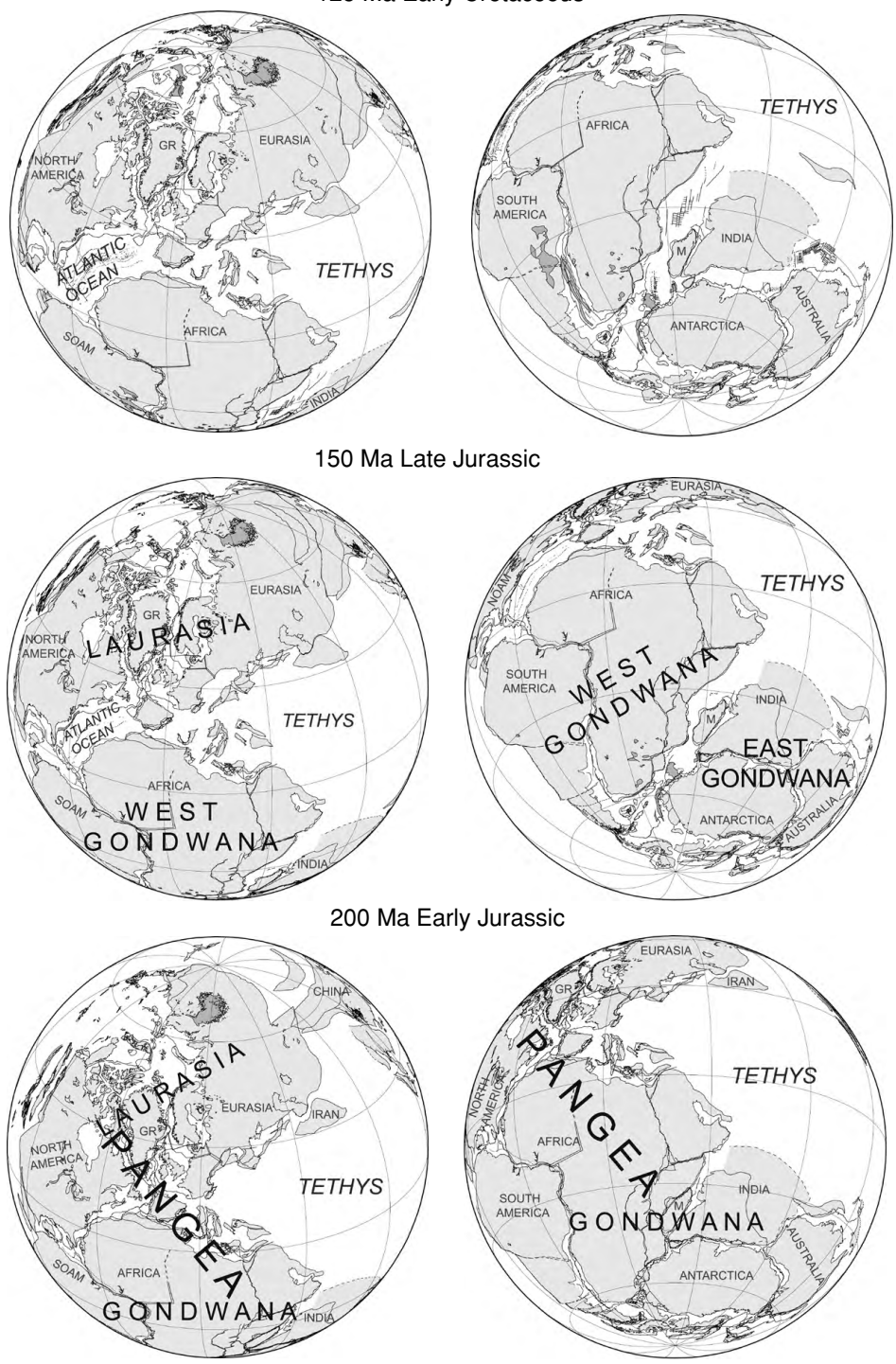
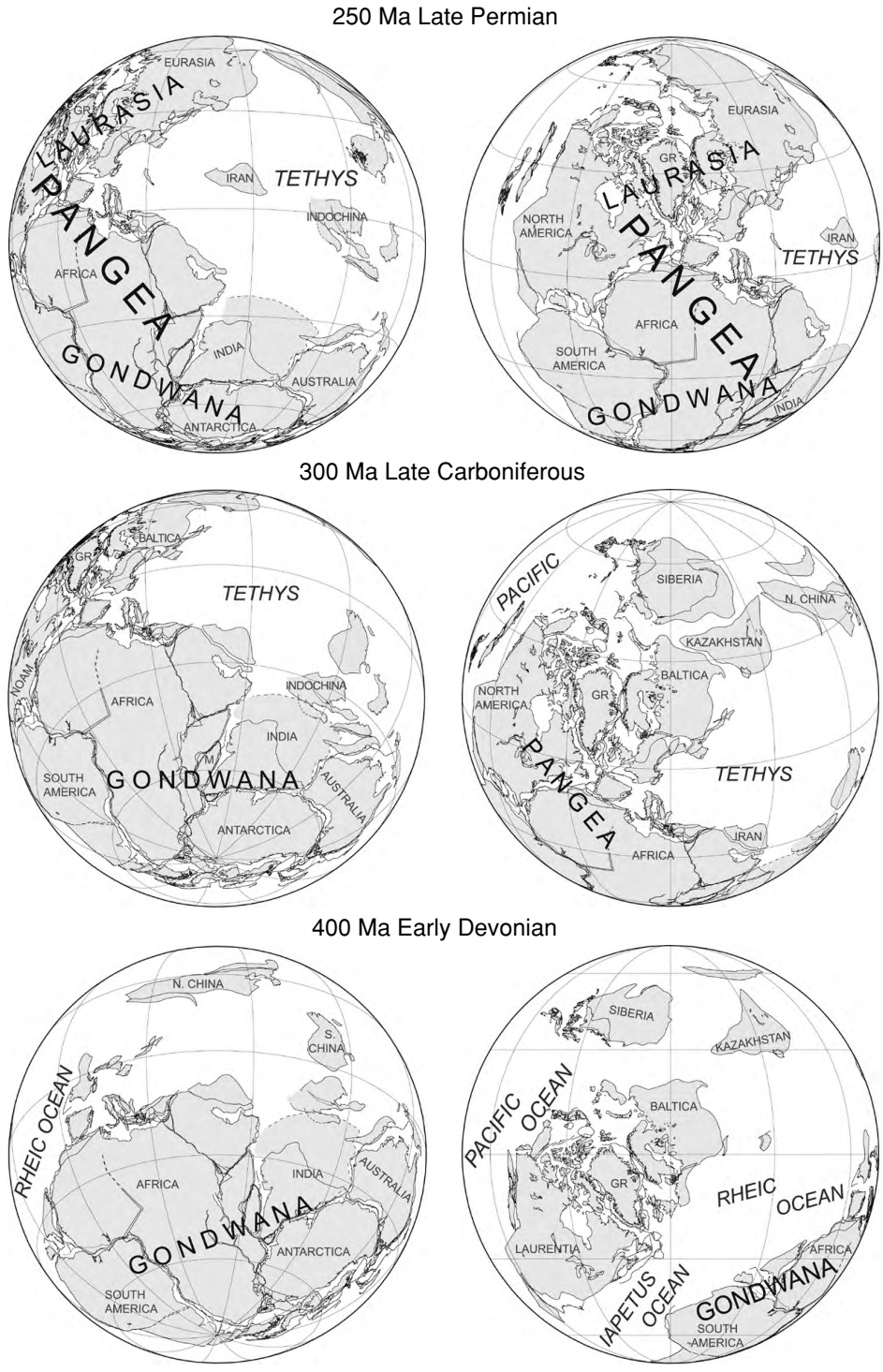


Figure 3.30. (*cont.*)



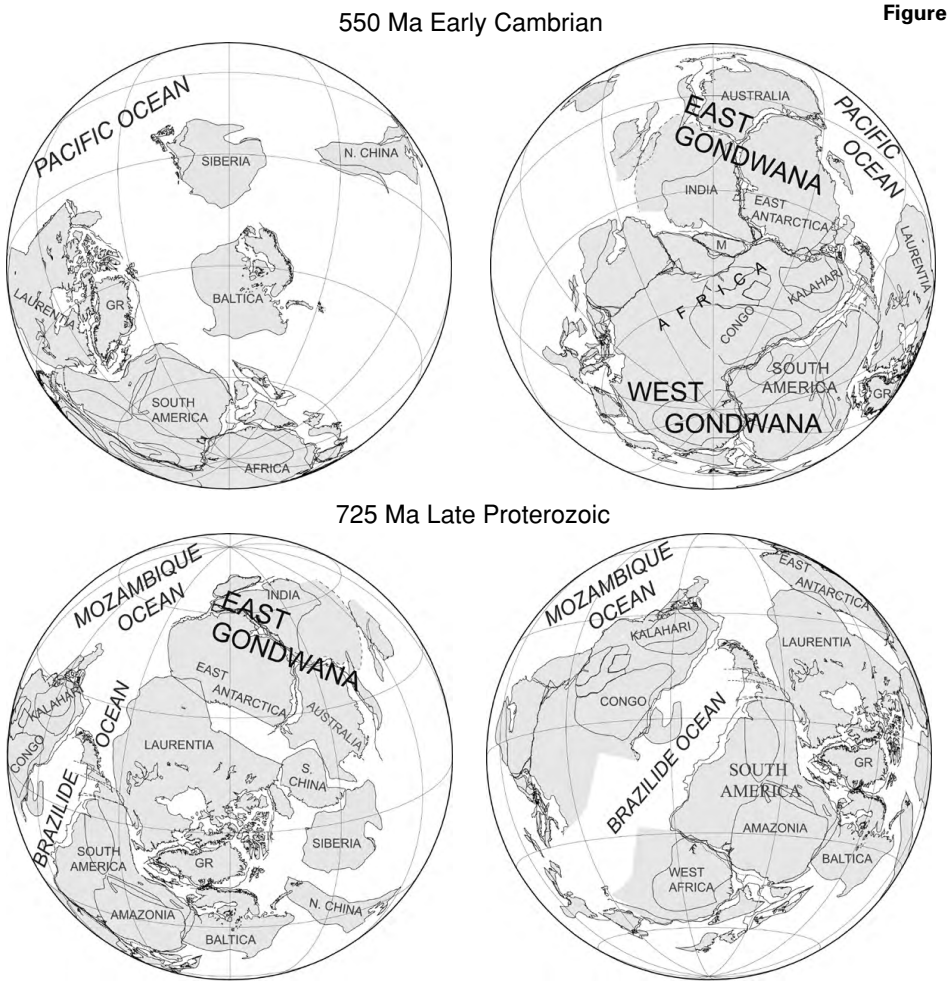


Figure 3.30. (*cont.*)

Problems

1. Assume that the Earth's magnetic field is a dipole aligned along the geographic north–south axis.
   (a) What is the angle of inclination at London, U.K. (51°N, 0°E)?
   (b) What is the angle of inclination at Canberra, Australia (35°S, 149°E)?
   (c) What is the angle of inclination on Spitzbergen (78°N, 16°E)?
   (d) What is the angle of inclination at Rio de Janeiro, Brazil (23°S, 43°W)?
   (e) If the angle of inclination is 76°, where are you?
   (f) If the angle of inclination is −36°, where are you?
2. Magnetic measurements have been made on some lavas found at 60°N, 90°W. The angle of inclination is measured to be 37°. At what magnetic latitude were these lavas erupted?

3. If the direction of magnetization of the lavas of Problem 2 is due west, calculate the position of the pole of the Earth's magnetic field at the time the lavas were erupted. What does this indicate about the continent on which these lavas occur?
4. If the measurement of the angle of inclination of the lavas of Problem 2 is in error by 5°, what is the subsequent error in the calculated palaeolatitude?
5. Down to what depth can (a) oceanic lithosphere and (b) continental lithosphere be permanently magnetized? (Hint: see Chapter 7.)
6. (a) Use Fig. 3.12 to calculate a spreading rate for the South Atlantic. Is this a half-spreading rate or a plate-separation rate?
   (b) Using Fig. 3.12, estimate minimum and maximum spreading rates for the Pacific during the last 80 Ma.
7. Calculate the minimum length of time that a reversal of the Earth's magnetic field lasted if it was detected from sea-surface magnetic data collected in (a) the Atlantic Ocean and (b) the equatorial Pacific Ocean.
8. Where is the oldest ocean floor, and what age is it? Discuss what may have happened to the rest of the ocean floor of this age and why.
9. At any given latitude, what orientation of ridge gives rise to the largest-amplitude magnetic anomalies?
10. At what latitude do magnetic anomalies due to magnetized oceanic crust have minimum amplitude?

References and bibliography

Atwater, T. 1970. Implications of plate tectonics for the Cenozoic tectonic evolution of western North America. *Bull. Geol. Soc. Am*., **81**, 3513–36.

1989. Plate tectonic history of the northeast Pacific and western North America. In E. L. Winterer, D. M. Hussong and R. W. Decker, eds., *The Geology of North America, Vol. N, The Eastern Pacific Ocean and Hawaii*. Boulder, Colorado: Geological Society of America, pp. 21–7.

Atwater, T. and Severinghaus, J. 1989. Tectonic maps of the northeast Pacific. In E. L. Winterer, D. M. Hussong and R. W. Decker, eds., *The Geology of North America, Vol. N, The Eastern Pacific Ocean and Hawaii*. Boulder, Coloredo: Geological Society of America, pp. 15–20.

Bambach, R. K., Scotese, C. R. and Ziegler, A. M. 1980. Before Pangea: the geographies of the Paleozoic world. *Am. Sci*., **68**, 26–38.

Barton, C. E. 1996. Revision of International Geomagnetic Reference Field released. *EOS Trans. Am. Geophys. Un*., **77** (16), 153.

1997. International Geomagnetic Reference Field: the seventh generation. *J. Geomag. Geoelectr*., **49**, 123–48.

Briden, J. C., Hurley, A. M. and Smith, A. G. 1981. Paleomagnetism and Mesozoic–Cenozoic paleocontinental world maps. *J. Geophys. Res*., **86**, 11 631–56.

Brunhes, B. 1906. Récherches sur la direction d'aimentation des roches volcaniques (1). *J. Physique, 4ième Sér*., **5**, 705–24.

Bullard, E. C. and Mason, R. G. 1963. The magnetic field over the oceans. In M. N. Hill, ed., *The Sea*, Vol. 3. New York: Wiley, pp. 175–217.

Candé, S. C. and Kent, D. V. 1992. A new geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *J. Geophys. Res*., **97**, 13 917–51.

1995. Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *J. Geophys. Res*., **100**, 6093–5.

Candé, S. C. and Mutter, J. C. 1982. A revised identification of the oldest sea-floor spreading anomalies between Australia and Antarctica. *Earth Planet. Sci. Lett*., **58**, 151–60.

Dalziel, I. W. D. 1997. Neoproterozoic–Paleozoic geography and tectonics: review, hypothesis, environmental speculation. *Geol. Soc. Am. Bull*., **109** (1), 16–42.

DeMets, C. and Royer, J.-Y. 2003. A new high-resolution model for India–Capricorn motion since 20 Ma: implications for chronology and magnitude of distributed crustal deformation in the Central Indian Basin. *Current Sci.*, **85**, 339–45.

Dickinson, W. R. 1979a. Cenozoic plate tectonic setting of the Cordilleran region in the United States. In J. M. Armentrout, M. R. Cole and H. TerBest Jr, eds., *Cenozoic Paleography of the Western United States*. Pacific Coast Paleogeography Symposium 3. Tulsa, Oklahoma: Society of Economic Palaeontologists and Mineralogists, pp. 1–13.

1979b. Plate tectonic evolution of North Pacific Rim. In S. Uyeda, R. Murphy and K. Kobayashi, eds., *Geodynamics of Western Pacific, Advances in Earth and Planetary Sciences*, 6. Tokyo: Centre for Academic Publications, pp. 1–19.

Ellis, R. M. *et al*. 1983. The Vancouver Island Seismic Project: a COCRUST onshore–offshore study of a convergent margin. *Can. J. Earth Sci*., **20**, 719–41.

Engebretson, D. C., Cox, A. and Gordon, R. G. 1985. *Relative Motion between Oceanic and Continental Plates in the Pacific Basin*. Geological Society of America Special Paper 206. Boulder, Colorado: Geological Society of America.

Falvey, D. A. and Mutter, J. C. 1981. Regional plate tectonics and evolution of Australia's passive continental margins. *Bur. Miner. Resour. J. Aust. Geol. Geophys*., **6**, 1–29.

Farrar, E. and Dixon, J. M. 1981. Early Tertiary rupture of the Pacific plate: 1700 km of dextral offset along the Emperor Trough–Line Islands lineament. *Earth Planet. Sci. Lett*., **53**, 307–22.

Fisher, R. L. and Goodwillie, A. M. 1997. The physiography of the Southwest Indian Ridge, *Marine Geophys. Res*., **19**, 451–5.

Fisher, R. L. and Sclater, J. G. 1983. Tectonic evolution of the southwest Indian Ocean ridge system since the mid-Cretaceous: plate mobility and stability of the pole of Antarctica/Africa for at least 80 My. *Geophys. J. Roy. Astr. Soc*., **73**, 553–76.

Gahagan, L. M. *et al*. 1988. Tectonic map of the ocean basins from satellite altimetry data. *Tectonophysics*, **155**, 1–26.

Garland, D. G. 1979. *Introduction to Geophysics: Mantle, Core and Crust*, 2nd edn. Philadelphia: Saunders.

Gilbert, W. 1600. *De magnete* (English translation by P. F. Mottelay, 1893, New York: Dover Books Inc.)

Gradstein, F. M., Ogg, J. G. and Smith, A. G. 2004. *A Geologic Time Scale*. Cambridge: Cambridge University Press.

Grantz, A., Johnson, L. and Sweeney, J. F., eds., 1990. *The Arctic Ocean Region, The Geology of North America*, Vol. L. Boulder, Colorado: Geological Society of America.

Handschumacher, D. W., 1976. Post-Eocene plate tectonics of the Eastern Pacific. In G. H. Sutton *et al*., eds., *The Geophysics of the Pacific Ocean Basin and its Margin*. Vol. 19 of

Geophysical Monographs of the American Geophysical Union. Washington: American Geophysical Union, pp. 177–202.

Harland, W. B., Cox A. V., Llewellyn, P. G., Pickton, C. A. G., Smith, A. G. and Walters, R. 1990. *A Geologic Time Scale 1989*. Cambridge: Cambridge University Press.

Harrison, C. G. A. 1987. Marine magnetic anomalies – the origin of the stripes. *Ann. Rev. Earth Planet. Sci*., **15**, 505–43.

Heirtzler, J. R., Dickson, G. O., Herron, E. M., Pitman III, W. C. and Le Pichon, X. 1968. Marine magnetic anomalies, geomagnetic field reversals and motions of the ocean floor and continents. *J. Geophys. Res*., **73**, 2119–36.

Hilde, T. W. C., Isezaki, N. and Wageman, J. M. 1976. Mesozoic sea-floor spreading in the North Pacific. In G. H. Sutton *et al*., eds., *The Geophysics of the Pacific Ocean Basin and its Margin*. Vol. 19 of Geophysical Monographs of the American Geophysical Union. Washington: American Geophysical Union, pp. 205–26.

Hutton, V. R. S. 1976. The electrical conductivity of the Earth and planets. *Rep. Prog. Phys*., **39**, 487–572.

Ingersoll, R. V. 1982. Triple-junction instability as cause for late Cenozoic extension and fragmentation of the western United States. *Geology*, **10**, 621–4.

Irving, E. 1983. Fragmentation and assembly of the continents, mid-Carboniferous to present. *Geophys. Surveys*, **5**, 299–333.

1988. The paleomagnetic confirmation of continental drift. *EOS Trans. Am. Geophys. Un*., **69** (44), 994–1014.

Irving, E. and Sweeney, J. F. 1982. Origin of the Arctic Basin. *Trans. Roy. Soc. Canada, Ser. IV*, **20**, 409–16.

Klitgord, K. D. and Mammerickx, J. 1982. Northern East Pacific Rise: magnetic anomaly and bathymetric framework. *J. Geophys. Res*., **87**, 6725–50.

Klitgord, K. D. and Schouten, H. 1986. Plate kinematics of the Central Atlantic. In P. R. Vogt and B. E. Tucholke, eds., *The Geology of North America, Vol. M, The Western North Atlantic Region*. Boulder, Colorado: Geological Society of America, pp. 3351–78.

Larson, R. L. and Chase, C. G. 1972. Late Mesozoic evolution of the western Pacific Ocean. *Bull. Geol. Soc. Am*., **83**, 3627–44.

Lawver, L. A., Sclater, J. G. and Meinke, L. 1985. Mesozoic and Cenozoic reconstructions of the South Atlantic. *Tectonophysics*, **114**, 233–54.

Lawver, L. A., Dalziel, I. W. D., Gahagan, L. M., Martin, K. M. and Campbell, D. A. 2003. The PLATES 2003 Atlas of Plate Reconstruction (750 Ma to Present Day). PLATES Progress Report No. 280-0703. University of Texas Institute for Geophysics Technical Report No. 190. Houston, Texas: University of Texas Press.

Lemaux II, J., Gordon, R. G. and Royer, J.-Y. 2002. Location of the Nubia–Somalia boundary along the Southwest Indian Ridge. *Geology*, **30**, 339–42.

Malin, S. R. C. and Bullard, E. 1981. The direction of the Earth's magnetic field at London, 1570–1975. *Phil. Trans. Roy. Soc*. A, **299**, 357–423.

Mammerickx, J. and Klitgord, K. D. 1982. Northern East Pacific Rise: evolution from 25 mybp to the present. *J. Geophys. Res*., **87**, 6751–9.

Matuyama Motonori 1929. On the direction of magnetisation of basalt in Japan, Tyôsen and Manchuria. *Japan Academy Proc*., **5**, 203–5.

Maxwell, A. E., Von Herzen, R. P., Hsü, K. J., Andrews, J. E., Saito, T., Percival Jr, S. F., Milow, E. D. and Boyce, R. E. 1970. Deep sea drilling in the South Atlantic. *Science*, **168**, 1047–59.

McDougall, I., Brown, F. H., Cerling, T. E., and Hillhouse, J. W. 1992. A reappraisal of the geomagnetic time scale to 4 Ma using data from the Turkana Basin, East Africa. *Geophys. Res. Lett*., **19**, 2349–52.

McElhinny, M. W. 1973. *Palaeomagnetism and Plate Tectonics*. Cambridge: Cambridge University Press.

McKenzie, D. P. 1973. The evolution of the Indian Ocean. *Sci. Am*., **228** (5), 62–72.

McKenzie, D. P. and Sclater, J. G. 1971. The evolution of the Indian Ocean since the late Cretaceous. *Geophys J. Roy. Astr. Soc*., **25**, 437–528.

Molnar, P., Pardo-Casas, F. and Stock, J. 1988. The Cenozoic and late Cretaceous evolution of the Indian Ocean Basin: uncertainties in the reconstructed positions of the Indian, African and Antarctic plates. *Basin Res.*, **1**, 23–40.

Morel, P. and Irving, E. 1981. Paleomagnetism and the evolution of Pangaea. *J. Geophys. Res*., **86**, 1858–72.

Morley, L. W. and Larochelle, A. 1964, Paleomagnetism as a means of dating geological events. In F. F. Osborne, ed., *Geochronology in Canada*. Royal Society of Canada Special Publication No. 8. Toronto: Toronto University Press, pp. 39–51.

Müller, R. D., Gaina, C., Roest, W., Clark, S. P. and Sdrolias, M. 2002. The evolution of global oceanic crust from Jurassic to present day: a global data integration. In *American Geophysical Union Fall Meeting Abstracts*. Washington: American Geophysical Union.

Müller, R. D. *et al*. 1997. Digital isochrons of the world's ocean floor. *J. Geophys. Res*. **102**, 3211–14.

Mutter, J. C., Hegarty, K. A., Candé, S. C. and Weissel, J. K. 1985. Breakup between Australia and Antarctica: a brief review in the light of new data. *Tectonophysics*., **114**, 255–79.

Ness, G., Levi, S. and Couch, R. 1980. Marine magnetic anomaly timescales for the Cenozoic and late Cretaceous: a précis, critique and synthesis. *Rev. Geophys. Space Phys*., **18**, 753–70.

Nicholson, C., Sorlien, C. C., Atwater, T., Crowell, J. C. and Luyendyk, B. P. 1994. Microplate capture, rotation of the Western Transverse Ranges, and initiation of the San Andreas transform as a low angle fault system. *Geology*, **22**, 491–5.

Norton, I. O. and Sclater, J. G. 1979. A model for the evolution of the Indian Ocean and the breakup of Gondwanaland. *J. Geophys. Res*., **84**, 6803–30.

Ogg, J. G. 1995. Magnetic polarity time scale of the Phanerozoic. In T. J. Ahrens, ed., Global *Earth Physics: A Handbook of Physical Constants*. Washington: American Geophysical Union, pp. 240–70.

Owen, H. G. 1983. *Atlas of Continental Displacement: 200 Million Years to the Present*. Cambridge: Cambridge University Press.

Parkinson, W. D. 1983. *Introduction to Geomagnetism*. Edinburgh: Scottish Academic Press.

Peddie, N. W. 1982. IGRF 1980: a report by IAGA Division Working Group 1. *Geophys. J. Roy. Astr. Soc*., **68**, 265–8.

Pitman III, W. C. and Heirtzler, J. R. 1966. Magnetic anomalies over the Pacific–Antarctic ridge. *Science*, **154**, 1164–71.

Powell, C. McA., Roots, S. R. and Veevers, J. J. 1988. Pre-breakup continental extension in East Gondwanaland and the early opening of the eastern Indian Ocean. *Tectonophysics*, **155**, 261–83.

Raff, A. D. and Mason, R. G. 1961. Magnetic survey off the west coast of North America, 40°N latitude to 50°N latitude. *Geol. Soc. Am. Bull*., **72**, 1267–70.

Richards, M. A., Gordon, R. G. and van der Hilst, R. D. eds. 2000. *The History and Dynamics of Global Plate Motions.* Geophysical Monograph 121. Washington: American Geophysical Union.

Royer, J.-Y. and Gordon, R. G. 1997. The motion and boundary between the Capricorn and Antarctic plates. *Science*, **277**, 1268–74.

Royer, J.-Y., Sclater, J. G. and Sandwell, D. T. 1989. A preliminary tectonic chart of the Indian Ocean. *Proc. Indian Acad. Sci., Earth, Planet Sci*., **98**, 7–24.

Schouten, H. and Denham, C. R. 1979. Modelling the oceanic source layer. In M. Talwani, G. C. A. Harrison and D. E. Hayes, eds., *Implications of Deep Drilling Results in the Atlantic Ocean: Ocean Crust*. Vol. 2 of Maurice Ewing Series. Washington: American Geophysical Union, pp. 151–9.

Schouten, H., Denham, C. and Smith, W. 1982. On the quality of marine magnetic anomaly sources and sea floor topography. *Geophys. J. Roy. Astr. Soc*., **70**, 245–59.

Schouten, J. A. 1971. A fundamental analysis of magnetic anomalies over oceanic ridges. *Mar. Geophys. Res*., **1**, 111–44.

Sclater, J. G., Fischer, R. L., Patriat, P., Tapscott, C. and Parsons, B. 1981. Eocene to recent development of the southwest Indian ridge, a consequence of the evolution of the Indian Ocean triple junction. *Geophys. J. Roy. Astr. Soc*., **64**, 587–604.

Sclater, J. G., Jaupart, C. and Galson, D. 1980. The heat flow through oceanic and continental crust and the heat loss of the Earth. *Rev. Geophys. Space Phys*., **18**, 269–311.

Sclater, J. G., Parsons, B. and Jaupart, C. 1981. Oceans and continents: similarities and differences in the mechanisms of heat loss. *J. Geophys. Res*., **86**, 11 535–52.

Scotese, C. R. 1984. An introduction to this volume: Paleozoic paleomagnetism and the assembly of Pangea. In R. Van der Voo, C. R. Scotese and N. Bonhommet, eds., *Plate Reconstruction from Paleozoic Paleomagnetism*. Geodynamics Series, Vol. 12. Washington: American Geophysical Union, pp. 1–10.

Scotese, C. R. and Sager, W. W., eds., 1988. Mesozoic and Cenozoic plate reconstructions. *Tectonophysics*, **155**.

Smith, A. G., Hurley, A. M. and Briden, J. C. 1981. *Phanerozoic Paleocontinental World Maps*. Cambridge: Cambridge University Press.

Spell, T. L. and McDougall, I. 1992. Revisions to the age of the Brunhes–Matuyama boundary and the Pleistocene geomagnetic polarity timescale. *Geophys. Res. Lett*., **19**, 1181–4.

Stein, S. and Freymuller, J. T. eds. 2002. *Plate Boundary Zones.* Geodynamics Series 30. Washington: American Geophysical Union.

Stock, J. and Molnar, P. 1988. Uncertainties and implications of the Late Cretaceous and Tertiary position of the North American relative to Farallon, Kula and Pacific plates. *Tectonics*, **8**, 1359–70.

Stock, J. M. and Hodges, K. V. 1989. Transfer of Baja California to the Pacific Plate. *Tectonics*, **8**, 99–115.

Tivey, M. A. and Tucholke, B. E. 1998. Magnetization of 0–29 Ma ocean crust on the Mid-Atlantic Ridge. *J. Geophys. Res*., **103**, 17 807–26.

Van der Voo, R., Peinado, J. and Scotese, C. R. 1984. A paleomagnetic reevaluation of Pangea reconstructions. In R. Van der Voo, C. R. Scotese and N. Bonhommet, eds., *Plate Reconstruction from Paleozoic Paleomagnetism*. Geodynamics Series, Vol. 12. Washington: American Geophysical Union, pp. 11–26.

Veevers, J. J. 1984. *Phanerozoic Earth History of Australia*. Oxford: Clarendon.
1986. Breakup of Australia and Antarctica estimated as mid-Cretaceous (95 ± 5 Ma) from magnetic and seismic data at the continental margin. *Earth Planet. Sci. Lett*., **77**, 91–9.
Vine, F. J. 1966. Spreading of the ocean floor: new evidence. *Science*, **154**, 1405–15.
Vine, F. J. and Matthews, D. H. 1963. Magnetic anomalies over oceanic ridges. *Nature*, **199**, 947–9.
Vine, F. J. and Wilson, J. T. 1965. Magnetic anomalies over a young oceanic ridge off Vancouver Island. *Science*, **150**, 485–9.
White, R. and McKenzie, D. 1989. Magmatism at rift zones: the generation of volcanic continental margins and flood basalts. *J. Geophys. Res.*, **94**, 7685–729.
Wilson, D. S., Hey, R. N., and Nishimura, C. 1984. Propagation as a mechanism of reorientation of the Juan de Fuca Ridge. *J. Geophys. Res*., **89**, 9215–25.
Zatman, S., Gordon, R. G. and Richards, M. A. 2001. Analytic models for the dynamics of diffuse oceanic plate boundaries. *Geophys. J. Int*., **145**, 145–56.

Chapter 4
Seismology Measuring the interior

4.1 Waves through the Earth

4.1.1 Introduction

Seismology is the study of the passage of elastic waves through the Earth. It is arguably the most powerful method available for studying the structure of the interior of the Earth, especially the crust and mantle. There are various other geophysical techniques, including the study of gravity, magnetism and the electrical properties of the Earth, which can be applied on scales ranging from the planet as a whole to large regions or small areas or even individual rock samples (Telford *et al*. 1990; Dobrin and Savit 1988); but seismology is probably the most widely used and the most informative. This chapter discusses the methods by which we obtain information about the interior of the planet from the study of elastic waves passing through the Earth.

Earthquake seismology is perhaps the best technique for investigating the Earth's interior. The study of earthquakes was of major significance in giving us our understanding of plate tectonics: earthquake foci have delineated the boundaries of the tectonic plates very accurately. It has also helped us to map the internal structure of our planet. The distribution of earthquakes shows us where the Earth is active (mostly near the surface), and the passage of seismic waves through the Earth allows us, as it were, to CAT-scan its interior.

When an earthquake or an explosion occurs within the Earth, part of the energy released takes the form of elastic waves that are transmitted through the Earth. These waves can be detected by an instrument called a *seismograph*,[1] consisting of a *seismometer*, which measures and amplifies the motion of the ground on which is rests, and a recorder, which transfers the data onto paper, magnetic tape or disc. The speed with which these elastic waves travel depends on the density and elastic moduli of the rocks through which the waves pass. There are two types of elastic waves: *body waves* and *surface waves*.

[1] The first seismograph capable of detecting distant events was developed by John Milne (1850–1913) in the 1890s.

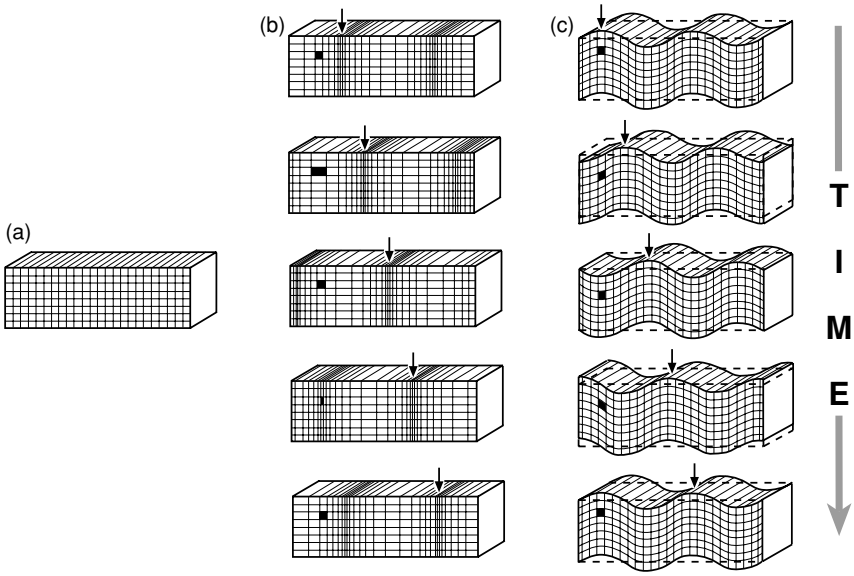


Figure 4.1. Successive stages in the deformation of (a) a block of material by (b) P-waves and (c) SV-waves. The sequences progress in time from top to bottom and the seismic wave is travelling through the block from left to right. An arrow marks the crest of the wave at each stage. (b) For P-waves, both the volume and the shape of the marked region change as the wave passes. (c) For S-waves, the volume remains unchanged and the region undergoes rotation only. The particle motion and deformation for SH-waves (horizontally polarized S-waves) are the same as for SV-waves (vertically polarized S-waves) but occur entirely in the horizontal plane instead of the vertical plane; imagine (c) rotated through 90°. (From Phillips (1968).)

4.1.2 Body waves

Body waves are seismic waves that travel through the body of the Earth. The propagation of body waves is similar to that of light: body waves are reflected and transmitted at interfaces where the seismic velocity and/or density change, and they obey Snell's law (Section 4.3.2). There are two types of body waves.

1. P-waves (P stands for primary, pressure or push–pull). These waves involve compression and rarefaction of the material as the wave passes through it but not rotation. P-waves are most correctly called dilatational or irrotational waves. They are the analogue in a solid of sound waves in air.
2. S-waves (S stands for secondary, shear or shake). These waves involve shearing and rotation of the material as the wave passes through it but no change in volume. S-waves are most correctly called rotational or equivoluminal waves.

Figure 4.1 shows the deformation undergone by a block of material when P- and S-waves pass through it. The P-wave particle motion is longitudinal, meaning that the particles making up the medium through which the P-wave is passing

vibrate about an equilibrium position, in the same direction as the direction in which the P-wave is travelling. In contrast, the particle motion of S-waves is transverse, that is, perpendicular to the direction of motion of the S-wave. The S-wave motion can be split into a horizontally polarized motion termed SH and a vertically polarized motion termed SV.

The *wave equation* is derived and discussed in Appendix 2. The three-dimensional *compressional wave equation* (Eq. (A2.48)) is

$$\frac{\partial^2 \phi}{\partial t^2} = \alpha^2 \nabla^2 \phi \tag{4.1}$$

where ϕ is the *scalar displacement potential* and α the speed at which dilatational waves travel. The three-dimensional *rotational wave equation* (Eq. (A2.49)) is

$$\frac{\partial^2 \boldsymbol{\psi}}{\partial t^2} = \beta^2 \nabla^2 \boldsymbol{\psi} \tag{4.2}$$

where $\boldsymbol{\psi}$ is the *vector displacement potential* and β the speed at which rotational waves travel. The *displacement* $\mathbf{u}$ of the medium by any wave is given by Eq. (A2.45):

$$\mathbf{u} = \nabla \phi + \nabla \wedge \boldsymbol{\psi} \tag{4.3}$$

The speeds at which dilatational and rotational waves travel, α and β, are termed the *P-wave* and *S-wave seismic velocities*. Often the symbols v_p and v_s are used instead of α and β.

The P- and S-wave velocities depend on the physical properties of the material through which the wave travels (Eqs. (A2.31), (A2.39) and (A2.44)):

$$\alpha = \sqrt{\frac{K + \frac{4}{3}\mu}{\rho}} \tag{4.4}$$

$$\beta = \sqrt{\frac{\mu}{\rho}} \tag{4.5}$$

where K is the bulk modulus or incompressibility, μ the shear modulus or rigidity and ρ the density.

The bulk modulus K, which is defined as the ratio of the increase in pressure to the resulting fractional change in volume, is a measure of the force per unit area required to compress material. The shear modulus μ is a measure of the force per unit area needed to change the shape of a material. Since P-waves involve change of volume and shape, α is a function of K and μ, whereas β is only a function of μ because S-waves involve no change in volume. Since the bulk modulus K must be positive, Eqs. (4.4) and (4.5) show that α is always greater than β, or, in other words, P-waves always travel faster than S-waves. The rigidity μ for a liquid is zero, a liquid has no rigidity and cannot sustain shear; Eq. (4.5) therefore indicates that S-waves cannot be propagated through liquids. Thus, S-waves cannot be transmitted through the Earth's liquid outer core.

k = -v(dp/dv)

k = rho(dp/drho)

mu = FL/AdelX

The dependence of α and β on density is not immediately obvious, but, in general, denser rocks have higher seismic velocities, contrary to what one would

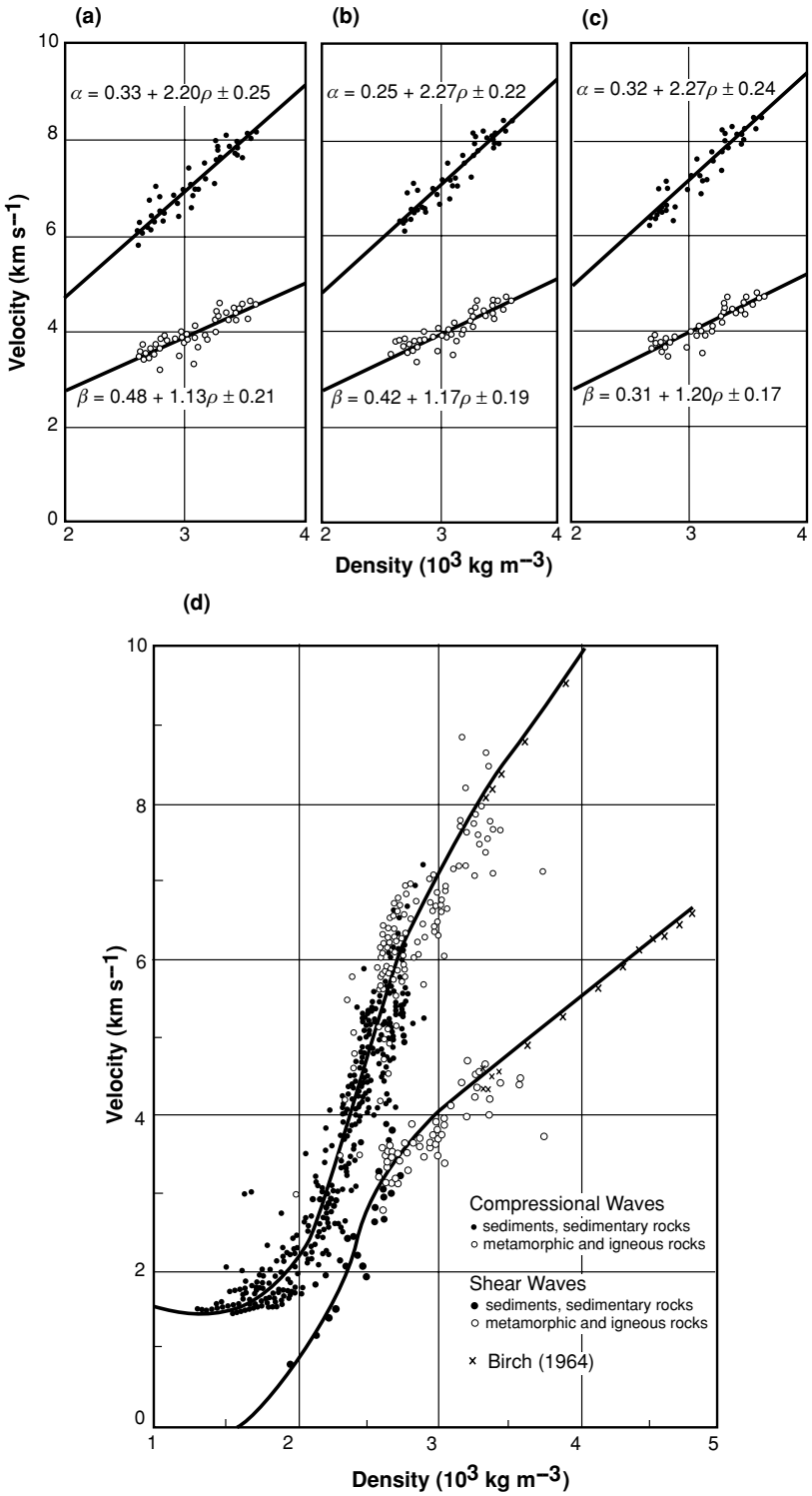


Figure 4.2. (a)–(c) Examples of the essentially linear relationship between seismic velocities and density known as Birch's law. Laboratory measurements made on crustal rocks at confining pressures of (a) 0.2 GPa, (b) 0.6 GPa and (c) 1.0 GPa, which correspond to approximate depths of 6, 18 and 30 km, are shown. (1 GPa = 10^9 Pa. The pascal, a unit of pressure, is equal to one newton per square metre: 1 Pa = 1 N m^{-2}. A unit of pressure still often used today is the bar: 1 bar = 10^6 dyne cm^{-2}. Thus, 1 GPa = 10^4 bar = 10 kbar). (After Gebrande (1982).) (d) The relationship between seismic velocities and density referred to as the Nafe–Drake curve after its originators. Open circles, values for igneous and metamorphic rocks; solid circles, values for sedimentary rocks; crosses show a linear model from Birch 1964. (After Ludwig *et al.* (1970).)

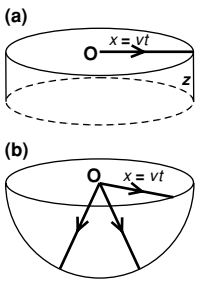


Figure 4.3. Amplitudes of surface waves and body waves. (a) A surface wave is travelling from a source O at speed v. After a time t the area of the cylindrical wavefront is $2\pi vtz = 2\pi xz$. The initial energy in the wave is now spread out over the area of this wavefront. Energy is proportional to the square of amplitude (Eq. (4.26)). Assuming that energy is conserved, the amplitude at a distance x is thus proportional to $x^{-1/2}$. (b) A body wave is travelling from a source O at speed v. After a time t the area of the spherical wavefront is $4\pi v^2t^2 = 4\pi x^2$. Assuming conservation of energy, the amplitude at distance x is proportional to x^{-1}.

expect from a first glance at Eqs. (4.4) and (4.5). This occurs because the elastic moduli K and μ are also dependent on ρ and increase more rapidly than $\rho^{1.0}$. A linear relationship, termed *Birch's law*, between density and seismic velocity is of the form

$$v = a\rho + b \tag{4.6}$$

where a and b are constants; it fits measurements from many crustal and mantle rocks fairly reasonably. Figure 4.2 shows examples of such linear relationships. The Nafe–Drake curve in Fig. 4.2(d) is another example of an empirical relationship between density and seismic velocity. This shows clearly that igneous and metamorphic crustal rocks generally have higher seismic velocities than those in sedimentary rocks. The ability to estimate density at a particular depth from the velocity (albeit with considerable uncertainty) can be useful when investigating the isostatic implications of structures determined seismologically, or in attempting to make gravity models for regions where something is already known about the seismic–velocity structure.

4.1.3 Surface waves

Surface waves are seismic waves that are guided along the surface of the Earth and the layers near the surface. They do not penetrate into the deep interior. Surface waves are generated best by shallow earthquakes. Nuclear explosions, although shallow, do not generate comparable surface waves. This fortunate fact is the basis for one criterion of discrimination between earthquakes and nuclear explosions (Section 4.2.4). Surface waves are larger in amplitude and longer in duration than body waves, and, because their velocities are lower than those of body waves, they arrive at seismographs after the main P- and S-waves. The reason for the relatively larger amplitude of surface waves is easy to understand (Fig. 4.3). The area of the cylindrical wavefront of a surface wave is proportional to x, the distance from its source, which means that the amplitude of the surface wave is inversely proportional to $\sqrt{x}$. In contrast, the wavefront of P- and S-waves at any time is spherical; therefore, the area of the wavefront is proportional to x^2, the square of the distance from the source. The amplitude of the body wave is therefore inversely proportional to x.

There are two types of surface waves,[2] both named after famous physicists: *Rayleigh waves*, sometimes descriptively called 'ground roll' in exploration seismology, are named after Lord Rayleigh, who predicted their existence in 1887; *Love waves* are named after A. E. H. Love, who predicted their existence in 1911.

[2] A third type of surface wave, the Stoneley wave, propagates along an interface between two media and is more correctly called an interface wave. Stoneley waves are not dispersive; they decrease in amplitude with distance from the interface and have a velocity between the lesser S-wave velocity and the greater Rayleigh-wave velocity.

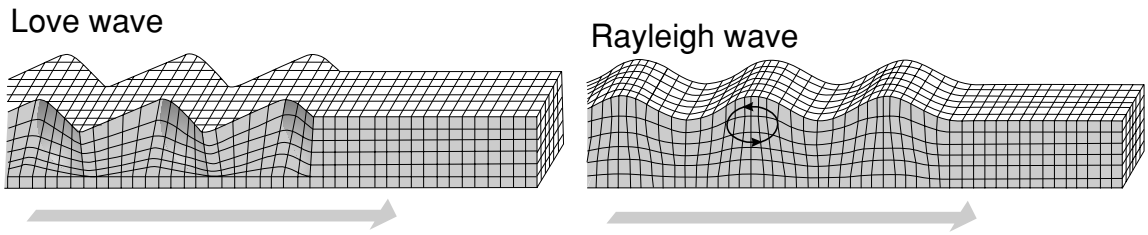


Figure 4.4. The particle motion for surface waves. (After Bolt (1976).)

Rayleigh waves are denoted by LR or R, and Love waves are denoted by LQ or Q (L for long; R for Rayleigh; Q for *Querwellen*, German, 'transverse waves'; Section 4.2.7). Rayleigh waves occur close to the surface of a semi-infinite medium. The particle motion for these waves is confined to a vertical plane containing the direction of propagation (Fig. 4.4). Near the surface of a uniform half-space it is a retrograde (anticlockwise for a wave travelling to the right) vertical ellipse. Rayleigh waves can therefore be recorded both by the vertical component and by the horizontal component of a seismometer. In contrast, Love waves occur when there is a general increase of S-wave velocity with depth. They propagate by multiple internal reflections of horizontally polarized S-waves (SH-waves) in this near-surface medium and thus propagate in a waveguide. The Love-wave particle motion is transverse and horizontal, so they can be recorded only by horizontal seismometers (Fig. 4.4).

The amplitude of Rayleigh waves decays exponentially with depth beneath the surface. The amplitude of any particular frequency component is dependent upon the following ratio:

$$\frac{\text{depth beneath the surface}}{\text{wavelength}}$$

Thus, for each frequency, the amplitude decreases by the same factor when the depth increases by a wavelength. This means that the longer-wavelength (longer-period or lower-frequency) surface waves contain more information about the deep velocity structure, and the shorter-wavelength (shorter-period or higher-frequency) surface waves yield information about the shallow structure. A good rule of thumb is that surface waves sample to a depth of their wavelength divided by three. Surface waves with periods less than about 50 s are used to determine gross crustal shear-wave velocity structures. However, even the highest-frequency (shortest-wavelength) surface waves cannot resolve the fine structure of the crust because their wavelength (velocity/frequency) is too great. Surface waves with longer periods are used to determine the mantle shear-wave velocity structure. For details of the mathematics of surface waves refer to a text such as officer (1974), Lay and Wallace (1995), Aki Richards (2001), Udías (1999), or Seth and Wysession (2003).

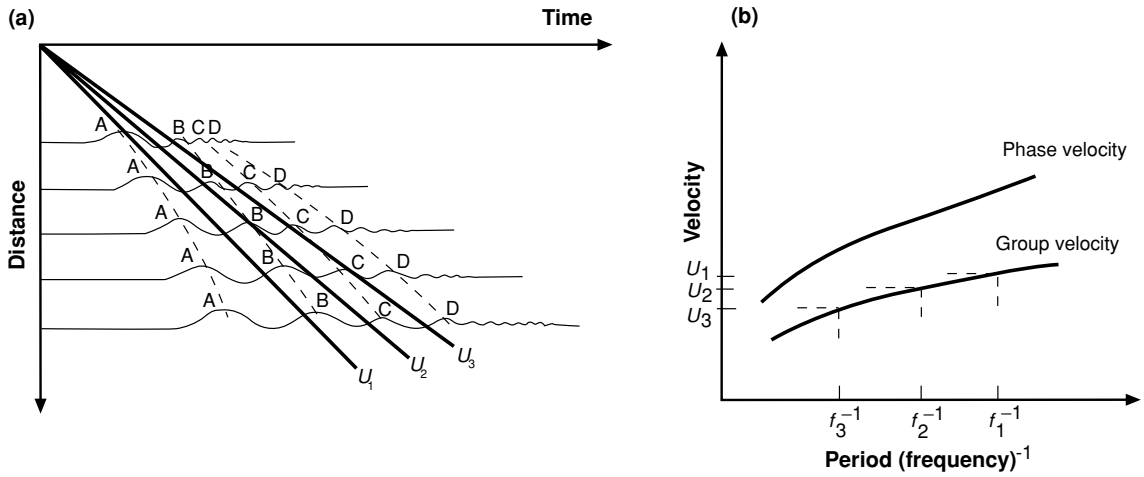


Figure 4.5. (a) Seismic records of dispersive surface waves obtained at increasing distances from the source (at the origin). The phases A, B, C and D are each associated with a different frequency from one record to the next – frequency decreases with increasing distance in this example. The inverse slope of the dashed line linking the phases is the phase velocity, which is a function of frequency. The straight solid lines have inverse slopes U_1, U_2 and U_3, the group velocities for frequency f_1, f_2 and f_3. The group velocity is a constant for each frequency and, in this example, decreases with increasing frequency. (From Officer (1974).) (b) Group- and phase-velocity dispersion curves for the example shown in (a).

Both Rayleigh waves and Love waves are, in practice, *dispersive*, which means that their velocities depend on frequency (i.e., different frequencies travel at different velocities). Dispersion means that a wave train changes shape as it travels (see Fig. 4.18(b) for an example). The first surface-wave energy to arrive at any seismometer is of those frequencies that have the greatest velocities. The other frequencies will arrive later, according to their velocities. Therefore, seismometers at increasingly greater distances from an earthquake (measured along the great-circle arc linking seismometer and earthquake) record wave trains that are increasingly long. The velocity with which surface-wave energy associated with a particular frequency travels is called the *group velocity*. The other velocity used in dealing with the surface waves is the *phase velocity*, which is defined as the velocity with which any particular phase (i.e., peak or trough) travels. The group velocity and the phase velocity are both functions of frequency.

Consider Fig. 4.5, which illustrates a record section[3] of surface waves. Notice that, on each seismogram, the lower-frequency waves arrive before the higher-frequency waves. The first phase to arrive on each record is peak A. Notice also that the frequency of peak A is not a constant from one record to the next: the frequency decreases (period increases) as the distance increases. The dashed

[3] A record section is a travel time versus distance display of seismograms.

curve linking peak A on each record defines the phase velocity for peak A. The phase velocity for the frequency of peak A at each distance is the inverse slope of the dashed curve at that distance. The dashed lines linking the subsequent peaks B, C and D also determine the phase velocity as a function of frequency. The slopes of all these dashed lines indicate that, in this example, the phase velocity decreases as the frequency of the surface waves increases (i.e., the phase velocity increases with the period).

The group velocity of surface waves is the velocity at which surface-wave energy of a given frequency travels. It is a constant for a given frequency. Therefore on Fig. 4.5(a) straight lines passing through the origin mark the surface-wave signal of any particular frequency on each successive record. Such lines for the three frequencies f_1, f_2 and f_3, where $f_1 < f_2 < f_3$, are shown in Fig. 4.5(a). The group velocities for these frequencies are U_1, U_2 and U_3, where $U_1 > U_2 > U_3$. In this example, the group velocity decreases as the frequency of these surface waves increases (i.e., the group velocity increases with the period).

A plot of velocity against period, called a *dispersion curve*, is the usual way of presenting this velocity–frequency information. Figure 4.5(b) illustrates dispersion curves appropriate for the records of Fig. 4.5(a). Notice that, in this example, the group velocity is less than the phase velocity. To repeat, energy travels with the group velocity, not the phase velocity.

Theoretically, the group velocity $U(f)$ and phase velocity $V(f)$ are linked by

$$U = V + f\,\frac{\mathrm{d}V}{\mathrm{d}f} \tag{4.7a}$$

where f is frequency. When expressed in terms of wavenumber, the group velocity $U(k)$ and phase velocity $V(k)$ are linked by

$$U = V + k\,\frac{\mathrm{d}V}{\mathrm{d}k} \tag{4.7b}$$

In the example shown in Fig. 4.5, $\mathrm{d}V/\mathrm{d}f$ is negative (phase velocity decreases as frequency increases), so $U < V$.

First, consider Rayleigh waves. In the ideal theoretical situation, in which the elastic properties of the Earth are constant with depth, Rayleigh waves are not dispersive and travel at a velocity of approximately 0.92β (i.e., slower than S-waves). However, the real Earth is layered, and, when the equations of motion for Rayleigh waves in a layered Earth are solved (which is far beyond the scope of this text), then Rayleigh waves are found to be dispersive (see Fig. 4.18(b)).

Next, consider Love waves. Love waves can exist only when the shear-wave velocity increases with depth or, in the layered case, when the shear-wave velocity of the overlying layer β_1 is less than that of the substratum β_2 (i.e., $\beta_1 < \beta_2$). Love waves are *always* dispersive. Their phase velocity is always between β_1 and β_2 in magnitude. The low-frequency (long-period) limit of the phase velocity is β_2, and the high-frequency (short-period) limit is β_1. Generally, Love-wave group

velocities are greater than Rayleigh-wave group velocities (Fig. 4.5(a)), which means that Love waves usually arrive before Rayleigh waves on seismograms (see Fig. 4.17(b)).

Dispersion curves contain much information about the velocity structure of the crust and upper mantle, but it is no simple matter to extract it. A *linearized inversion technique* (an iterative scheme that obtains a best fit between actual dispersion curves and the model curves which would be generated by particular velocity–depth structures), calculated by computer, is often used to obtain a velocity–depth structure appropriate for particular dispersion curves. Figure 4.6(a) shows standard continental and oceanic dispersion curves. Notice that, over much of the frequency range, oceanic paths are faster than continental paths. In Fig. 4.6(b) notice the low S-wave velocities starting at approximately 50–100 km depth. This low-velocity region is generally identified with the asthenosphere, the zone beneath the lithospheric plates, where the temperatures approach the solidus of mantle material and thus the material ceases to behave rigidly, becoming more ductile and able to creep.

4.1.4 Free oscillations of the Earth

Any mechanical system has a natural oscillation that can be excited, and the Earth is no exception. The Earth can oscillate in an indefinite number of *normal modes* of oscillation, rather like a giant bell. These free oscillations are standing waves: vibrations of the whole Earth with periods between about 100 s and 1 h. Although such oscillations had been predicted theoretically at the beginning of the twentieth century, the first definite measurement did not take place until 22 May 1960, when a large earthquake in Chile excited the oscillations sufficiently for them to be detected. Modern instrumentation is such that now free oscillations excited by large earthquakes can be recorded about twenty times each year. Free oscillations from the 1994 $M_w = 8.3$ earthquake in Bolivia were recorded for months afterwards.

There are two independent types of free oscillations: *spheroidal oscillations* (S) and *toroidal* or *torsional oscillations* (T). The general displacement for spheroidal oscillations has both radial and tangential components. The displacement for toroidal oscillations is always perpendicular to the radius vector and so is confined to the surfaces of concentric spheres within the Earth. Such oscillations, which involve only the crust and mantle (the outer core is liquid and so cannot sustain shear), do not change the shape or volume of the Earth. Both spheroidal and toroidal oscillations have an infinite number of modes (or, as in music, overtones). The notation used to describe free oscillations is ${}_nS_l$ and ${}_nT_l$. The first subscript, n, indicates the overtone: $n = 0$ is the fundamental mode while higher-frequency modes with $n > 0$ are overtones. The second subscript, l, the harmonic degree, indicates the number of nodes in latitude (places with zero

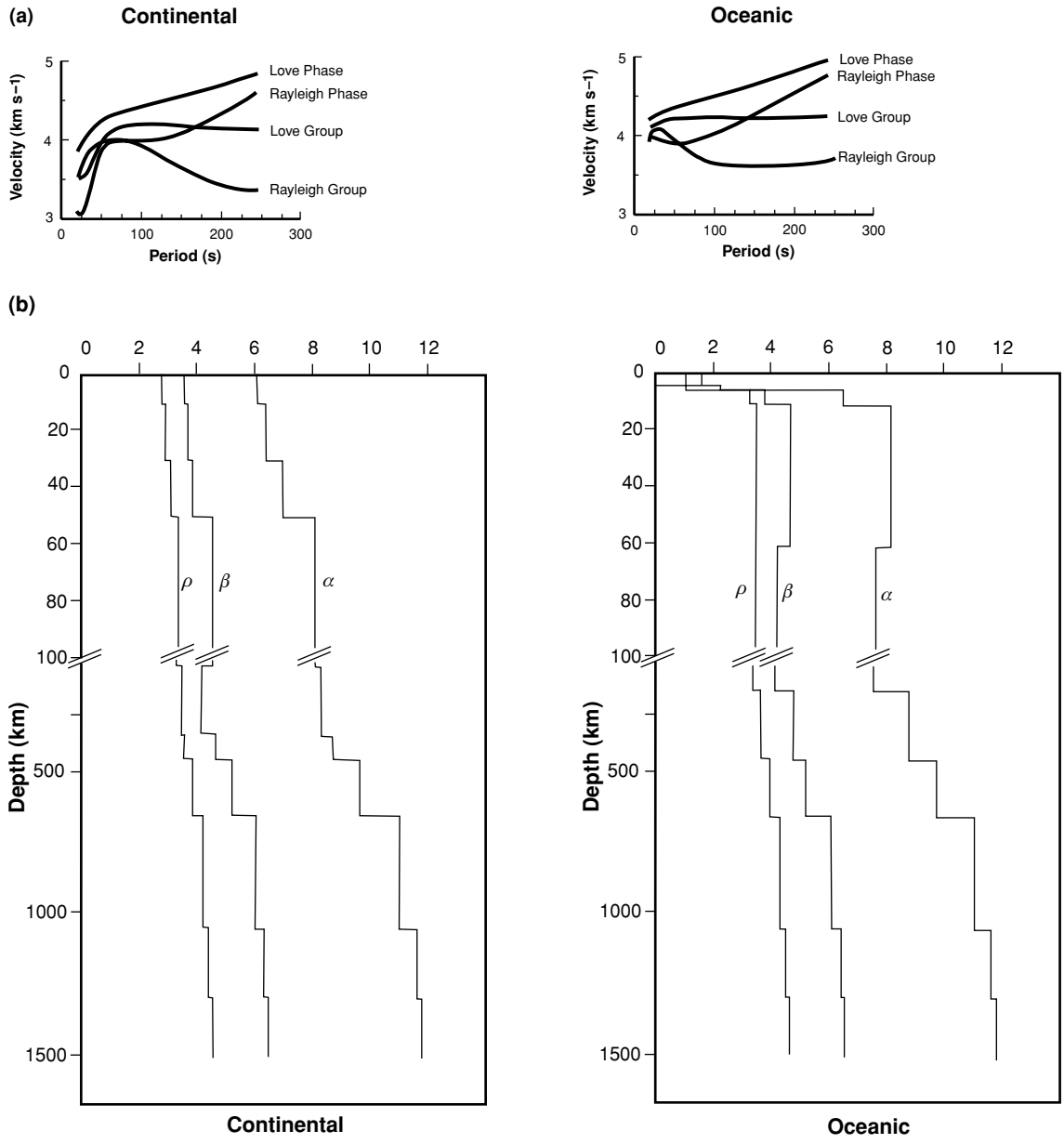


Figure 4.6. (a) Standard surface-wave dispersion curves for continental and oceanic paths. (b) Seismic velocity and density models for the dispersion curves in (a). Note the change in depth scale at 100 km. (After Knopoff and Chang (1977).)

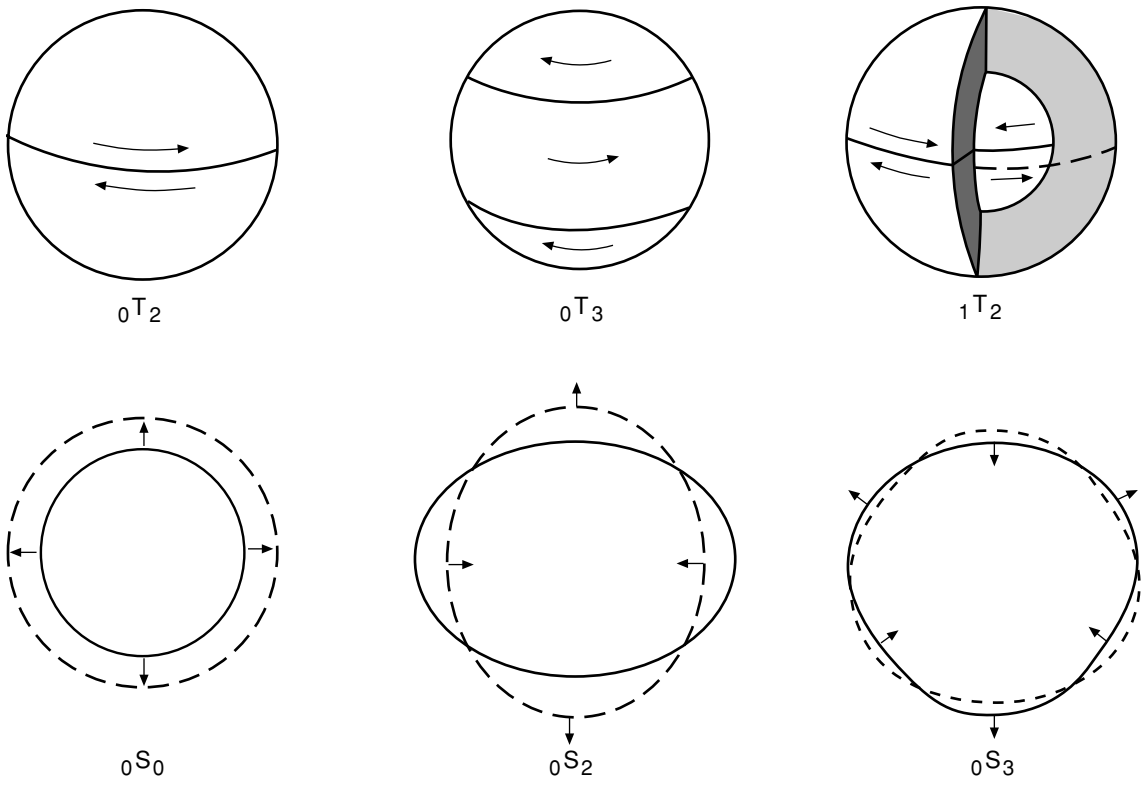


Figure 4.7. Free oscillation modes and periods: $_0T_2$ (44 min), $_0T_3$ (28 min), $_1T_2$ (13 min), $_0S_0$ (20 min), $_0S_2$ (54 min) and $_0S_3$ (36 min). In mode $_0T_2$, the opposite hemispheres twist in opposite directions. In mode $_1T_2$ the outside part of the sphere twists as in mode $_0T_2$, but the inner sphere twists in the opposite direction. (After Bullen and Bolt (1985).) To view some animations for harmonic degree 3, see http://earth.leeds.ac.uk/~greg/Sphar/index.html (from G. Houseman).

displacement). The simplest spheroidal oscillation is a purely radial motion, $_0S_0$, the so-called 'breathing' or 'balloon' mode of the Earth, which has a period of about 20 min (Fig. 4.7). Mode $_0S_2$, termed the 'football' mode of the Earth, is the lowest-frequency oscillation, with a period of almost an hour. Free oscillations can be detected by strainmeters (both S and T modes) and gravimeters used for measuring Earth tides (only S modes, since T modes do not involve radial motion, compression or dilatation), as well as by modern digital seismometers and accelerometers capable of recording very-long-period data.

The fundamental spheroidal modes are equivalent to the standing wave resulting from interference amongst all the Rayleigh waves, whereas the fundamental toroidal modes are equivalent to the standing wave resulting from interference amongst all the Love waves. This means that the periods of free oscillations can be used to extend the dispersion curves for surface waves out to the maximum

periods of ~2600 s (T mode) and ~3200 s (S mode). Other modes jointly contain (but not in a simple way) the body-wave phases. Thus the interference pattern obtained by summing together all the theoretical normal modes and including a model of a particular earthquake source yields a synthetic seismogram for that earthquake: these generally compare well with recordings.

4.2 Earthquake seismology

4.2.1 Location of earthquakes

The *earthquake focus* or hypocentre is the point in the Earth where the earthquake nucleated. It is specified by latitude, longitude and depth beneath the surface. The *earthquake epicentre* is the point on the Earth's surface vertically above the focus. For an earthquake beneath California, the focus might be at 37°N, 122°W and 10 km depth. The epicentre for this earthquake would be 37°N, 122°W. In fact, earthquakes do not occur exactly at points; rather the stress releases occur within small volumes or along fault planes.

Assume that an earthquake occurs at the surface at time t_0. There are three unknowns: time of origin, latitude and longitude. To determine these three unknowns, we need to know the arrival times of seismic waves at three nearby seismometers. If the shallow P-wave velocity is α and the shallow S-wave velocity is β, then the time taken for P-waves to travel from the focus to seismometer 1 at distance r_1 is r_1/α. Similarly, the travel time for the S-waves is r_1/β. The arrival time of P-waves at seismometer 1 is then $t_0 + r_1/\alpha$, and the arrival time of S-waves at seismometer 1 is $t_0 + r_1/\beta$. The difference between the P- and S-wave arrival times at seismometer 1, $t_{1,\mathrm{S-P}}$, is given by

$$t_{1,\mathrm{S-P}} = \frac{r_1}{\beta} - \frac{r_1}{\alpha} \tag{4.8}$$

and similarly for seismometers 2 and 3. If we assume values for α and β, we now have three linear equations with three unknowns, r_1, r_2 and r_3, which can easily be solved:

$$\begin{aligned} t_{1,\mathrm{S-P}} &= \frac{r_1}{\beta} - \frac{r_1}{\alpha} \\ t_{2,\mathrm{S-P}} &= \frac{r_2}{\beta} - \frac{r_2}{\alpha} \\ t_{3,\mathrm{S-P}} &= \frac{r_3}{\beta} - \frac{r_3}{\alpha} \end{aligned} \tag{4.9}$$

It would be simplest to solve this simple example graphically by drawing a map of the area, marking the seismometers and then drawing an arc of a circle of radius r_1 about seismometer 1, another radius r_2 about seismometer 2 and another radius r_3 about seismometer 3. The focus of the earthquake is then at the intersection of the three arcs (Fig. 4.8).

In reality, things are not quite so simple because the Earth is neither flat nor homogeneous with globally constant P- and S-wave velocities. In addition, there

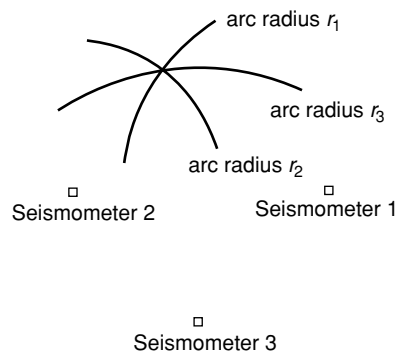


Figure 4.8. A map showing the location of seismometers 1, 2 and 3. The earthquake focus is at distances r_1 from seismometer 1, r_2 from seismometer 2 and r_3 from seismometer 3 (r_1, r_2 and r_3 are given by Eqs. (4.9)). The focus can be located by drawing an arc of a circle of radius r_1 about seismometer 1 and then repeating this for seismometers 2 and 3 (radii r_2 and r_3, respectively). The point at which the three arcs intersect is the focus.

is a fourth unknown, the focal depth: earthquakes do not conveniently occur at the surface. Errors in arrival times have to be allowed for. In practice, therefore, several seismometers are required in order to locate an earthquake. Locations are routinely calculated using as many travel times from as many stations as possible. The distance between the epicentre and the seismometer is called the *epicentral distance*. The shortest distance on the surface of a sphere between any two points on the sphere is along the great circle which intersects the two points. The epicentral distance is therefore the length of this great-circle arc and is usually quoted in degrees (except for local earthquake studies, in which case it would be quoted in kilometres).

With a measurement of t_{s-p} at any seismic station it is possible to make an initial estimate of the epicentral distance by using a compilation of the global travel times (Fig. 4.16). For example, when t_{s-p} is 3 min the epicentral distance is about 16° (1800 km), whereas when t_{s-p} is 5 min the epicentral distance is 30° (3300 km).

Focal depths can also be determined from measurement of the difference in travel time between the P phase and the pP phase (pP denotes a P-wave reflected at the Earth's surface in the vicinity of the earthquake, as in Fig. 4.9). The P phase travels along path FS and has arrival time t_P, whereas the pP phase travels along path FRS and arrives at time t_{pP}. At teleseismic distances FR is small compared with FS; so, to a first approximation, the length FR is given by

$$\mathrm{FR} = \frac{\alpha}{2}(t_{pP} - t_P) \tag{4.10}$$

For shear waves (using phases sS and S), the corresponding distance would be

$$\mathrm{FR} = \frac{\beta}{2}(t_{sS} - t_S) \tag{4.11}$$

The focal depth h is then given by

$$h = \mathrm{EF} = \mathrm{FR}\sin\theta \tag{4.12}$$

In practice, travel times of all the earthquake wave phases are tabulated: The *Jeffreys–Bullen* (1988) tables (often called J–B tables and shown graphically in Fig. 4.16) are probably the most complete (they are a revision of the first

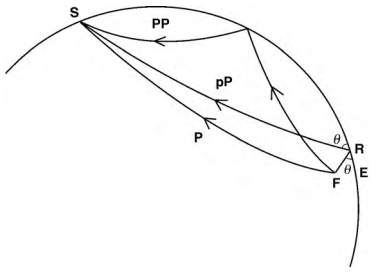


Figure 4.9. Two points on the Earth's surface satisfy the condition for a reflection of the P-wave (or S-wave) travelling through the mantle. (At a reflection point the angle of incidence equals the angle of reflection.) The first reflection point is close to the focus, and the second is midway between the focus and the seismometer. The first reflected ray is coded pP (or sS), and the second is coded PP (or SS). The difference between the arrival times of pP and P (the direct P-wave without reflection) (or sS and S) can be used to determine the focal depth (EF) of the earthquake. E is the epicentre, F the focus, S the seismometer and R the reflection point close to the focus. The diagram is purely schematic and not drawn to scale.

tabulation of travel times to be published). Even with a record from only one station, provided that the phases P, S and pP have been identified and the S – P time (Eq. (4.9)) and the pP – P time have been calculated, it is possible to use the tables to estimate both the epicentral distance and the focal depth.

Figure 4.10(a) shows how earthquake occurrence varies with focal depth and Fig. 4.10(b) shows how the maximum earthquake magnitude varies with focal depth. Most earthquakes are shallow, and none has been detected from depths greater than 700 km (see Section 4.2.3 on earthquake magnitudes and Section 9.6.3 on deep-focus earthquakes).

4.2.2 Aftershocks, foreshocks and swarms

An earthquake occurs as the result of a slow build-up of strain (deformation) in rock, usually caused by the relative motion of adjacent plates. When a fault or volume of rock can no longer resist movement, the stored strain energy is suddenly released, causing an earthquake.

Frequently, a strong earthquake is followed by a sequence of *aftershocks*, which can continue for months, although the number of aftershocks decays approximately as the reciprocal of time elapsed since the main earthquake. An earthquake is defined as an aftershock if its distance from the main earthquake is less than the length of that ruptured fault plane. Aftershocks occur during a period of readjustment following the main shock, in which small localized strains on the fault are released. Deep-focus earthquakes do not usually have aftershocks. The aftershocks are normally much smaller than the main earthquake and gradually decrease in magnitude. The largest aftershock, which often occurs within

aftershock

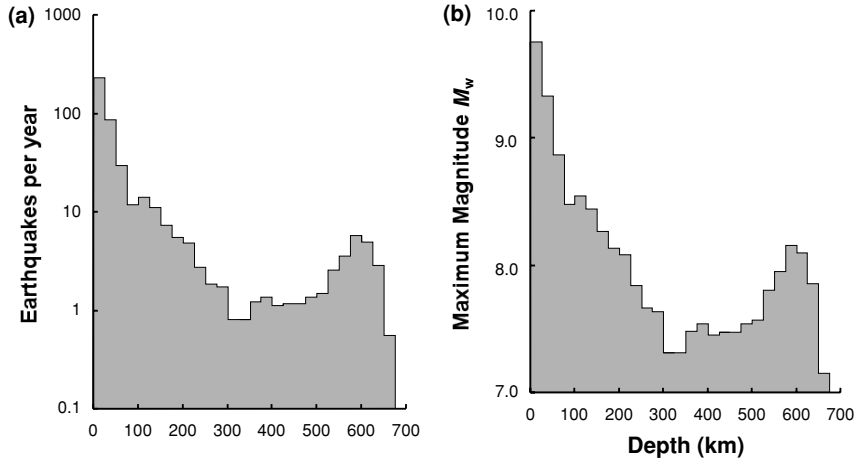


Figure 4.10. The variation of (a) earthquake occurrence and (b) maximum magnitude, with focal depth for earthquakes with $M_w > 5.45$ recorded during the period 1977–1997. Most of the earthquakes were shallow and resulted from brittle fracture and frictional sliding along fault planes. The number of earthquakes decreases exponentially with depth down to about 300 km. Seismic activity levels then stay low at just over one per year down to 500 km depth. There the activity increases, reaching almost six per year before ceasing completely at ~670 km. The largest earthquakes are shallow: the maximum magnitude decreases with depth down to ~300 km; it then remains at ~7.5 until 550 km, where it increases, reaching ~8.2 by 600 km. (After Frohlich (1998).)

hours of the main shock, usually reaches a magnitude about 1.2 units lower than that of the main shock (Båth 1979). (Earthquake magnitude is discussed in Section 4.2.3.)

Sometimes the main earthquake is preceded by a small *foreshock*. Unfortunately, it is not usually possible to identify it as a foreshock until after the main event has occurred. Earthquake *swarms* are yet another sequence of activity. In a swarm there is no main earthquake but instead a large number of small shocks, often with many occurring every day. The numbers of shocks build up slowly to a maximum and then die down again. Most earthquake swarms occur in volcanic areas, especially along the mid-ocean ridges, and these events usually have very shallow foci.

foreshock
swarms

Earthquakes do not account for all the motion that the plate-velocity models predict (Section 2.4). *Afterslip*, a term used to describe motion that takes place over some time after an earthquake, is so gradual that no seismic event is detected: it is *aseismic*. The recent establishment of the GPS satellite network (Section 2.4.1) means that continuously operating GPS receivers can now be deployed to investigate the occurrence of aseimic slip and afterslip along particular regions of interest. A 1994 $M_w = 7.6$ earthquake beneath Honshu,

afterslip
aseismic

Table 4.1 *Earthquake magnitudes and corresponding descriptors*

Earthquake magnitude	Earthquake descriptor	Approximate number per year
>8.0	Great	1–3
7.0–7.9	Major	<20
6.0–6.9	Strong	180
5.0–5.9	Moderate	1 800
4.0–4.9	Light	10 000
3.0–3.9	Minor	90 000
2.0–2.9	Very minor	1 000 000

Source: IRIS.

Japan, on the plate boundary with the subducting Pacific plate, was followed by a year during which the amount of displacement resulting from aseismic slip was comparable to the displacement that occurred during the earthquake. The next decades will provide much more data on deformation, slip rates and the release of strain in earthquakes around the world.

4.2.3 Earthquake magnitude and moment

The concept of earthquake magnitude was introduced by C. F. Richter in 1935. He was studying local earthquakes in southern California and proposed a particular logarithmic magnitude scale for them that was based on the maximum amplitude of the first arriving P-wave as measured by a 12-Hz Wood–Anderson seismometer. His was a local magnitude scale for southern California. Since then, measurements of magnitude have been extended globally to earthquakes at all depths and distances and to both body and surface waves. These are quantitative scales that give an estimate of the size of the earthquake. Richter's name is known around the world because press releases on earthquakes invariably quote magnitude as being 'on the Richter scale'. The descriptors generally associated with earthquakes of a particular magnitude are listed in Table 4.1.

Magnitude scales for earthquakes are all logarithmic and are based on measurements of the amplitude of the seismic waves. In addition there are *intensity scales* such as the Mercalli scale (Appendix 10), which are subjective and are based on the shaking and damage to buildings, breaking of glass, ground cracking, people running outside, and so on. Intensity scales must take into account different building codes and styles around the world and so are necessarily continent/country specific. Maps showing the extent of earthquake damage would normally utilize an intensity scale.

All the magnitude scales are of the form

$$M = \log_{10}\left(\frac{A}{T}\right) + q(\Delta, h) + a \tag{4.13}$$

where M is the magnitude, A the maximum amplitude of the wave (in 10^{-6} m), T the period of the wave (in seconds), q a function correcting for the decrease of amplitude of the wave with distance from the epicentre and focal depth, Δ the angular distance from seismometer to epicentre, h the focal depth of the earthquake and a an empirical constant related to the seismograph, its site location and subsurface characteristics.

It is important in all magnitude work to use the period ranges for which the particular equation has been determined; waves of different frequencies have different amplitude behaviours. Many determinations of the constant a and function q of Eq. (4.13) have been made, for various frequency bands, focal depths, geographic locations and so on. Båth (1981) provides a detailed review of earthquake magnitudes.

For shallow-focus (<50 km) teleseismic earthquakes (those with $20^\circ < \Delta < 160^\circ$) a *surface-wave magnitude* M_S can be defined as

$$M_S = \log_{10}\left(\frac{A}{T}\right)_{max} + 1.66\log_{10}\Delta + 3.3 \tag{4.14}$$

The amplitude A used is the maximum of the horizontal component of the Rayleigh wave in the period range 18–22 s. (If the vertical component is used instead, the empirical constant in the equation depends on the seismic layering beneath the recording station; and, if there is no layering, a constant term of 3.1 is appropriate.) Because earthquakes occurring at greater depths are not so efficient at generating surface waves, Eq. (4.14) gives too small a value for their magnitude. To allow for this, a correction must be applied to Eq. (4.14) for earthquakes with focal depths greater than 50 km:

$$(M_S)_{corrected} = M_S + \delta M_S(h) \tag{4.15}$$

where M_S is calculated from Eq. (4.14) and $\delta M_S(h)$ is the correction for focal depth h. The maximum value of δM_S(h), 0.4, is used for any earthquake with a focal depth greater than 90 km. At ranges shorter than 20° (approximately 2200 km), a correction, $\delta M_S(\Delta)$, must also be made to Eq. (4.14) to account for differences in absorption, scattering, geometrical spreading and dispersion:

$$(M_S)_{corrected} = M_S + \delta M_S(\Delta) \tag{4.16}$$

Estimates of this correction vary between 0.6 and 0.1. These corrections illustrate the fact that magnitude, although easy to measure, is not an exact description of an earthquake. Many variables are associated with the recording site and passage of the seismic signal through the Earth, as well as with the earthquake itself, all of which can have a large effect on magnitude. Differences of 0.2–0.3 units

in calculated magnitudes are not uncommon even under the most favourable conditions.

Because deep-focus earthquakes are not effective at generating surface waves, a better magnitude scale for them is based on body waves (P- and S-waves). A first suggestion by Gutenberg (1945) for *body-wave magnitude* m_b of shallow-focus earthquakes for P, PP and S waves of period 12 s was

$$m_b = \log_{10}\left(\frac{A}{12}\right) + 0.01\Delta + 5.9 \tag{4.17}$$

Many more recent determinations of the calibration function $q(\Delta, h)$ in Eq. (4.13) have been made since then, but the one generally used today is still the Gutenberg–Richter (Gutenberg and Richter 1956) calibration function:

$$m_b = \log_{10}\left(\frac{A}{T}\right)_{\max} + q(\Delta, h) \tag{4.18}$$

In this tabulation, $q(\Delta, h)$ for P-waves is not strongly dependent on focal depth and increases from ~6.0 at $\Delta = 10^\circ$ through ~6.5 at $\Delta = 80^\circ$ to ~8.0 at $\Delta = 110^\circ$. The period T of the first arriving P-wave that has travelled through the crust and mantle will generally be 0.1–5 s.

Values of m_b and M_S calculated for an earthquake are not usually the same: the two magnitude scales will not agree on the magnitude of an earthquake. This should not be a complete surprise; after all the two scales are based on different signals: m_b on body waves with period 1–5 s and M_S on surface waves with period 18–22 s. Generally, for a shallow-focus earthquake, the surface-wave magnitude will be more reliable than the body-wave magnitude. However, that the two scales are related is clear from Fig. 4.11(a). The data are scattered, but m_b and M_S appear to be linearly related. The m_b–M_S data plots vary appreciably from one earthquake region to the next, but a worldwide average of m_b–M_S relations is

$$m_b = 2.94 + 0.55M_S \tag{4.19}$$

The two magnitude scales coincide at magnitude 6.5. For small magnitudes, m_b is greater than M_S, and for large magnitudes, m_b is less than M_S.

The *seismic moment* M_0 of an earthquake is defined by

$$M_0 = \mu A u \tag{4.20}$$

where μ is the shear modulus, A the area of the fault and u the average displacement on the fault. The seismic moment is a physical quantity with a unique value for any earthquake. It can be determined by observations and estimates of the fault-plane area and displacement but can also be expressed in terms of the low-frequency (<0.005 Hz) amplitude spectra of surface waves. This is important since it means that M_0, as calculated from seismograms, can be used as the basis for a consistent magnitude scale. M_0 is replacing M_S as the standard measure of the long-period source spectra. Plots of log M_0 against M_S and of log A against log M_0 and the fault length against log M_0 are shown in Fig. 4.12. The linear trend

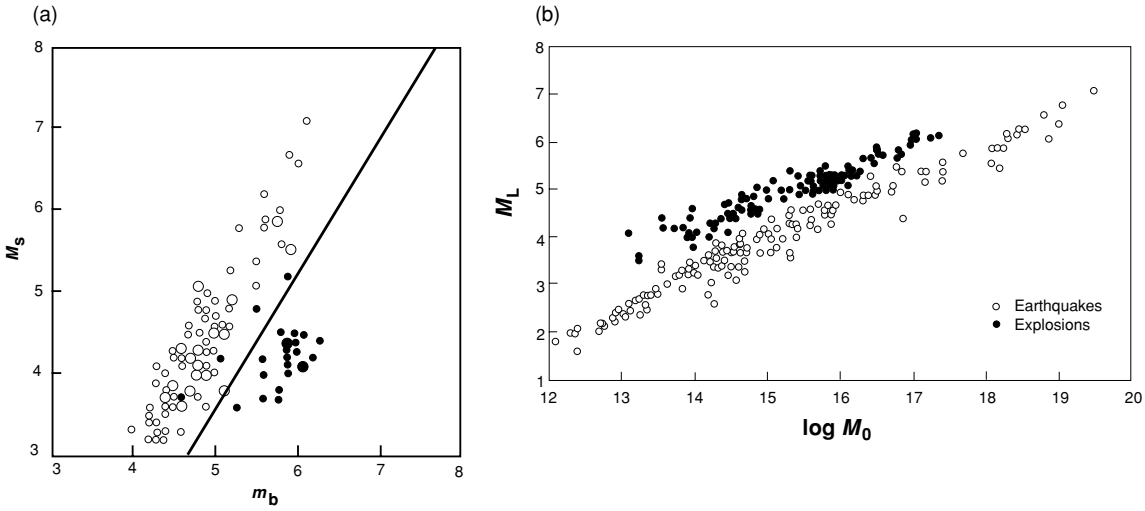


Figure 4.11. (a) A plot of surface-wave magnitude M_S against body-wave magnitude m_b for twenty-six nuclear explosions (solid circles) and ninety-nine earthquakes (open circles), all in Eurasia. The solid line is the discriminant line $m_b = 2.87 + 0.60M_S$. Nuclear explosions are less efficient at generating surface waves than are earthquakes with the same body-wave magnitude. (After Nowroozi (1986).) (b) A plot of local magnitude M_L against the logarithm of the seismic moment M_0 for 128 explosions (solid circles) and 173 earthquakes (open circles). The explosions were from the Nevada Test Site, U.S.A., and the earthquakes occurred in California and Nevada, U.S.A., and Baja California, Mexico. (After Woods and Helmberger, A new seismic discriminant for earthquakes and explosions, *EOS Trans. Am. Geophys. Un.*, **74** (8), 91. 1993. Copyright 1993 American Geophysical Union. Modified by permission of American Geophysical Union.)

shown in Fig. 4.12 even seems to extend down to very small earthquakes with log M_0 of 9 and source dimensions of only 10 m. Since the drop in stress, $\Delta\varsigma$, for earthquakes is constant over a large range of M_0, measurement of M_0 cannot yield an estimate for $\Delta\varsigma$. The total cumulative seismic moment of earthquakes each year is about 3×10^{28} N m. The lowest stress drops ($<$1.5 MPa) occur for thrust earthquakes at subduction zones and the highest stress drops ($>$3 MPa) occur for strike–slip earthquakes at mid-ocean-ridge transform faults.

Both the body-wave and the surface-wave magnitude scales saturate (do not give large enough values) for very large earthquakes and are strongly dependent on the frequency content of the seismic wave. Therefore, to estimate the size of large earthquakes, long-period (low-frequency) waves are used. A magnitude scale, the *moment magnitude* M_w, is obtained from the seismic moment and is a more reliable measure of the magnitude of an earthquake than is either m_b or M_S. The moment magnitude is defined as

$$M_w = \frac{2}{3}\log_{10} M_0 - 6.0 \tag{4.21}$$

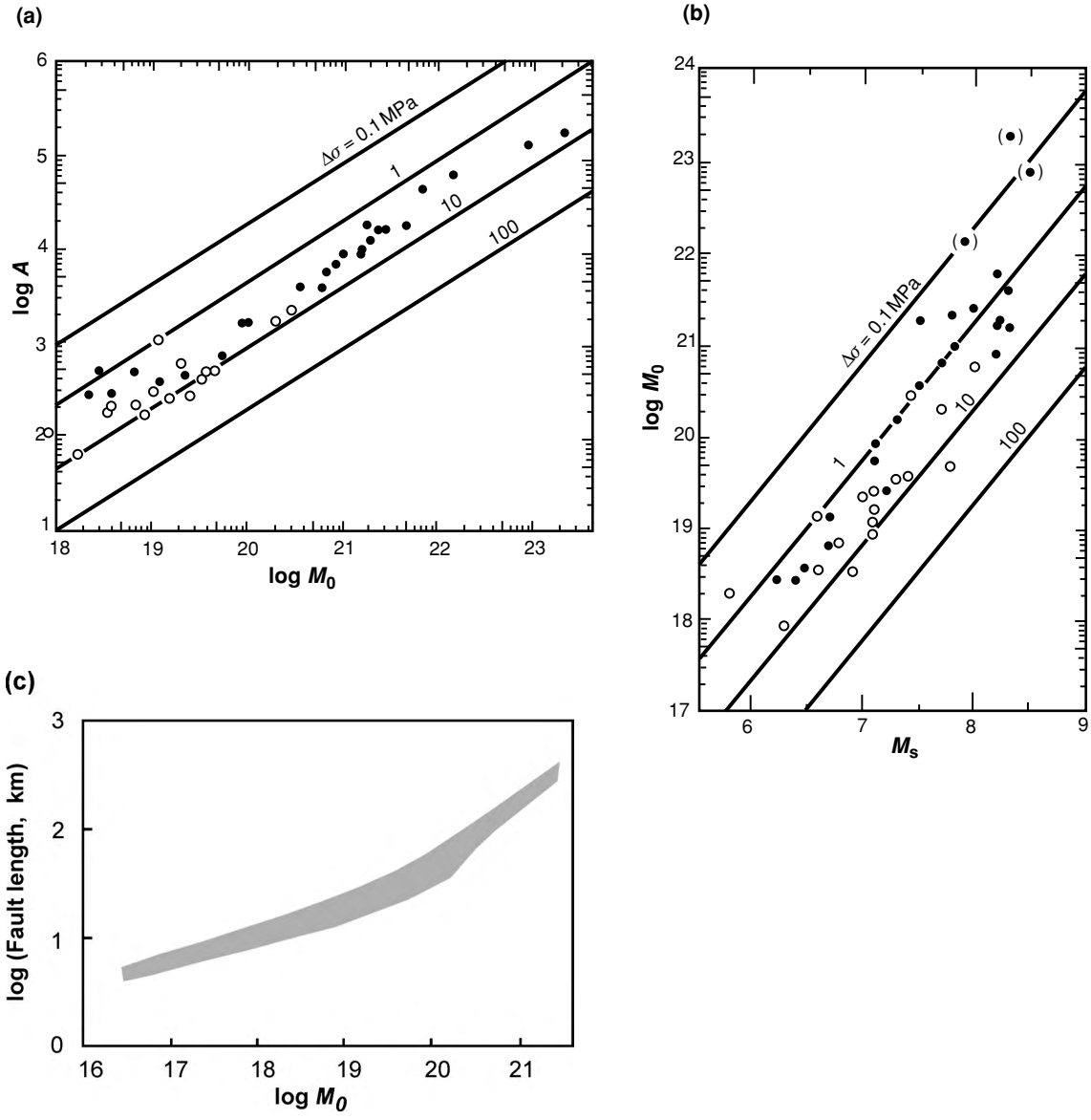


Figure 4.12. (a) The relationship between A (the surface area of the fault in km^2) and M_0 (the seismic moment in N m). (b) The relationship between M_0 and M_S (the surface-wave magnitude). Straight lines are for circular cracks with a constant stress drop $\Delta\sigma$ (in MPa). Both (a) and (b) show some difference in stress drop for interplate (•) and intraplate (o) earthquakes (After Kanamori and Anderson (1975).) (c) The relationship between the seismic moment of an earthquake and the length of the fault that it ruptures. There may be a difference between large- and small- thrust earthquakes: above some earthquake size, once the whole down-dip extent of the fault plane ruptures, the rupture can then continue only by extending along the length of the fault. This shows in the change of slope at $\sim 10^{20}$ N m: the seismic moment increases approximately as (fault length)^3 but for large earthquakes it increases as (fault length)^2. For strike–slip earthquakes, however, the seismic moment increases as (fault length)^2 up to 10^{21} N m. (After Henry and Das (2001).)

(with M_0 in N m) and has the important advantage over the other magnitude scales that it does not saturate towards the top end of the scale.

A few examples of the magnitudes of some famous earthquakes are as follows: the 'great' Lisbon earthquake of 1755, magnitude ~8.7, caused a tsunami and killed 60 000 people; the 1906 San Francisco, California, earthquake, $M_S = 7.8$–8.3 and $M_w = 7.9$ which ruptured about 300 km of the San Andreas Fault with ~6 m of displacement; the 22 May 1960 Chile earthquake, $M_S = 8.5$ and $M_w = 9.5$ which ruptured about 1000 km along the Peru–Chile subduction zone with ~20–40 m of displacement and caused a tsunami up to 15 m high across the Pacific; the 1963 earthquake in Škopje, Yugoslavia, $M_S = 6.0$; the 1964 Alaska earthquake, $M_S = 8.4$ and $M_w = 9.2$; the 1971 San Fernando, California, earthquake, $M_S = 6.6$ and $M_w = 6.7$; the 1976 Tangshan, China, earthquake, which killed 650 000 people, $M_S = 7.6$; the 1985 Mexico earthquake had $M_S = 8.1$ and $M_w = 8.0$; estimates of the magnitude of the 1988 Armenian earthquake range between 6.7 and 7.0; the 1989 Loma Prieta, California, earthquake had $M_w = 6.9$; and the 1995 Kobe, Japan earthquake which caused huge damage, left 300 000 homeless and killed 6300 people had a magnitude of 6.8. In 1999 the North Anatolian fault in Turkey was the site of two major earthquakes: the Izmit earthquake had $M_w = 7.4$, caused surface faulting for ~120 km and killed over 17 000 people; three months later the $M_w = 7.2$ Duzce earthquake extended the fault rupture eastwards (see Fig. 10.20(b)). The 2001 Gujarat, India, earthquake, which was a shallow, intraplate, thrust-faulting event with $M_w = 7.7$, caused immense damage and loss of life (the previous earthquake in this region had happened in 1819). The largest earthquake ever recorded in Japan occurred in 1933 and had $M_S = 8.5$ and $M_w = 8.4$.

The amount of damage done by an earthquake depends on factors such as population density, soil conditions and local building standards in the epicentral region, as well as on the focal depth, mechanism and magnitude of the earthquake. In general, a shallow earthquake with an M_S of 6 is very destructive, one with an M_S of 5 produces moderate damage, and earthquakes with magnitudes less than about 3 are usually not felt by many people, though they will, of course, be detected by local seismometers. The 1985 Mexico and 1989 Loma Prieta earthquakes were particularly damaging because of the enhanced shaking caused by shallow subsurface conditions – a drained lake-bed and young clayey-silt deposits beneath landfill, respectively. Had the buildings been on older, more solid rock, the damage would have been much less. The right-lateral strike–slip Kobe earthquake was not expected, although there was evidence to indicate that such earthquakes might occur in the complex fault system and regional stress field of Southwest Japan. *Earthquake hazard* is an estimate of the potential occurrence of damaging earthquakes. Figure 4.13, the earthquake-hazard map for the U.S.A., shows the estimate of the ground acceleration (as a percentage of gravity) that has a 10% chance of being exceeded during the next fifty years. Plate 3 shows the global earthquake hazard for a fifty-year window.

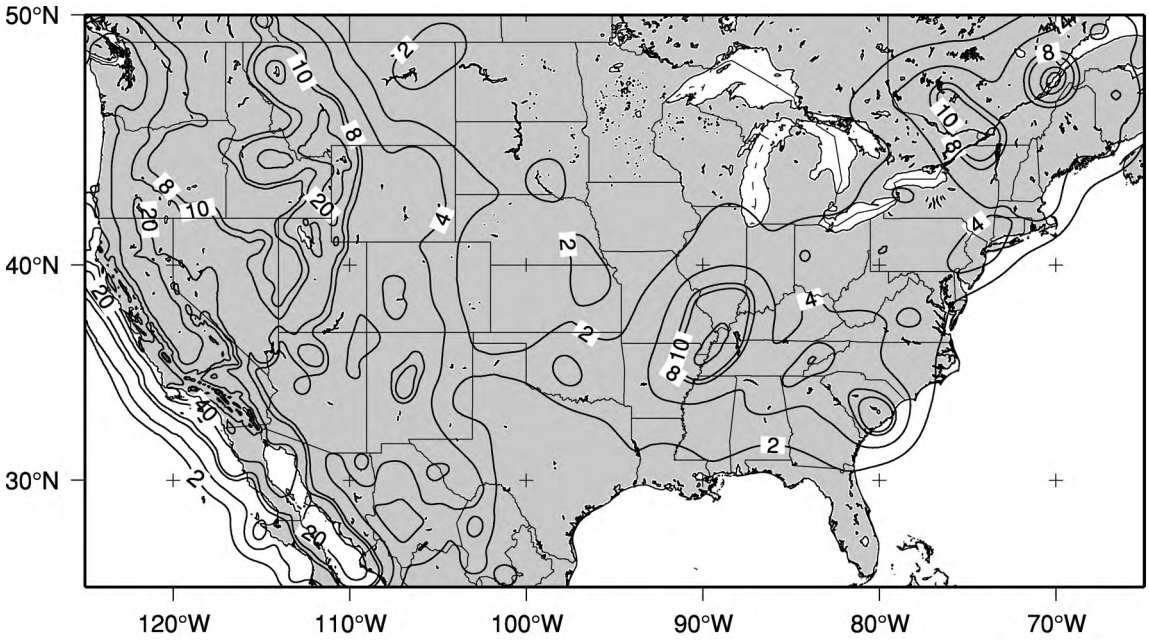


Figure 4.13. The earthquake-hazard map for the U.S.A. The ground acceleration (caused by earthquakes) that has a 10% probability of being exceeded in the next 50 years is shown as a percentage of gravity. The highest values are along the west coast with the San Andreas Fault system and the Cascadia subduction zone, but note that some areas in the east also have high values. Colour version Plate 4. See Plate 3 for the global seismic hazard map. (Data from Global Seismic Hazard Assessment Program: http://seismo.eth2.ch/GSHAP/ and USGS.)

Prediction of earthquakes is seen by some as the ultimate goal of earthquake research. However, prediction is of little value and is probably dangerous unless it is able to indicate with a reasonable probability both when and where an earthquake of a particular magnitude will occur. A predicted earthquake that then does not happen is a major social, political and economic problem. Present understanding is such that we have quite good knowledge of where earthquakes occur, although for very large infrequent earthquakes our knowledge of historical earthquakes is insufficient (Section 9.6.3), but reliable prediction of time and of magnitude is not yet attainable. Many approaches have been used to investigate and monitor the build-up of strain and its sudden release in an earthquake – these include monitoring of the deformation of the ground surface, recording microseismic activity, detecting changes in seismic P- and S-wave velocities, monitoring of changes in the local geomagnetic field and in the electrical resistivity of the upper crust and geochemical monitoring of groundwater and radon (a product of the decay of uranium in the crust). A section of the San Andreas Fault system near Parkfield, California, which has had a regular series of identical

devastasting-disastrous

magnitude-6 earthquakes has been subjected to intensive research. This resulted in the prediction of a magnitude-6 event between 1988 and 1992. Even though no earthquake occurred, one is still expected. A very successful prediction was made and evacuation warning given for the 1975 magnitude-7.3 Haicheng earthquake in China, but this success was not repeatable: the 1976 Tangshan earthquake was devastating. Despite concerted research effort in many fields over many years, the goal of earthquake prediction still remains distant.

4.2.4 Nuclear explosions

Nuclear explosions present particular problems: seismologists are required to monitor compliance with the ban on nuclear testing. To be certain that an event is identified correctly as an earthquake or an explosion, it is necessary to consider parameters such as location, depth and shape of the waveforms. Nuclear explosions have all been shallower than 2 km. Thus an event with a focal depth of 200 km is exceedingly unlikely to be a nuclear explosion, even assuming the maximum errors in depth calculation, whereas an event with a very shallow focal depth of 15 km would be a candidate for an explosion. The first P-wave from an explosion to arrive at any seismic station, irrespective of azimuth, should be compressional, whereas the first arrival from an earthquake will be either dilatational or compressional, depending upon the azimuth (Section 4.2.8).

As mentioned in Section 4.1.3, nuclear explosions are not as good as earthquakes at generating surface waves or S-waves. An earthquake will generate more long-period energy than will an explosion because it ruptures over a fault-plane that is large in comparison with the cavity produced by an explosion. In general, earthquakes have m_b values 1.0–1.5 units smaller than those for nuclear explosions with the same M_S values. The straight line in Fig. 4.11(a),

$$M_S = 1.68m_b - 4.82 \tag{4.22}$$

or

$$m_b = 2.87 + 0.60M_S \tag{4.23}$$

separates most of the explosions and earthquakes.

There is, however, a difficulty with small events. Calculations of m_b and M_S are generally performed with data recorded at distances greater than 2000 km from the epicentre but surface waves from small events ($m_b < 4.5$) cannot be reliably measured at such distances. A solution is to use M_L and M_0 rather than m_b and M_S. M_L, the local magnitude proposed by Richter, is based on measurements made close to the event. In Fig. 4.11(b) the earthquakes and explosions are separated over many orders of magnitude. The problems with this approach, however, are that an M_L scale is, as its name indicates, local and so has to be determined for each region; and local seismograph networks are required.

4.2.5 The magnitude–frequency relationship

Over 10 000 earthquakes with magnitude greater than about 4 occur each year.[4] As might be expected, small earthquakes are much more abundant than large ones. About twenty earthquakes with m_b or M_S greater than 7 occur worldwide each year and on average there is one 'great' earthquake with magnitude 8 or over per year, while there are over 1100 with m_b between 5.0 and 5.5, of which over 800 are shallower than 70 km. During the period 1968–1995 some 95 000 earthquakes with magnitude greater than 1 were recorded locally around San Francisco (Fig. 4.14).

It is typically found that there is a straight-line relationship between the surface-wave magnitude M_S and the logarithm of N, where N is the number of earthquakes with magnitude exceeding M_S occurring in an area per unit of time:

$$\log N = a - bM_S \tag{4.24a}$$

where a is a constant and the value of b is a measure of the relative abundance of large and small earthquakes. This statistical model is often referred to as the *Gutenberg–Richter 'law'*. The equivalent relationship between number of earthquakes and seismic moment is

$$\log N = \alpha - \beta \log M_0 \tag{4.24b}$$

A large b-value in a region indicates that small earthquakes occur frequently, whereas a small b-value indicates that small earthquakes are not so frequent and that large earthquakes are more likely to occur. An approximate worldwide average for b is 1 (the worldwide average for β is $2/3$); small b-values are generally those less than 1, and large b-values are greater than 1. If the b-value is known (or assumed), then knowledge of a (or α) means that Eq. (4.24) gives the maximum expected earthquake magnitude in the region. This is therefore useful in assessing areas for earthquake risk. Assume that a and b are known. Then Eq. (4.24a) gives N, the number of earthquakes exceeding magnitude M_S each year. If, therefore, N is 0.2, we could say that the recurrence time for an earthquake of this magnitude would be $1/N = 1/0.2 = 5$ yr. Alternatively, if $b = 1$, the value of a is the maximum expected earthquake magnitude per year for the region under consideration (since $\log_{10} 1 = 0$). Globally, since we have over 10 000 earthquakes each year with magnitude greater than 4, Eq. (4.24a) implies that we should expect one earthquake each year with magnitude greater

[4] The total number of earthquakes occurring each year is obviously variable. However, the sensitivity of instruments and extent of seismic networks, as well as the criteria used for confirmation of, and inclusion of, earthquakes in catalogues have changed over the years and continue to change. This makes it hard to present unambiguous statistics for global rates of earthquake occurrence. In 1931 there were about 350 instruments operating worldwide, now there are over ten times as many. The International Seismological Centre (ISC) data for 1905 show a total of 11 earthquakes; 84 in 1915; 617 in 1940; 3214 in 1960; 20 135 in 1980; 63 902 in 1990; and 48 596 in 1995. Most of the variability in these numbers is due to improvements in instrumentation.

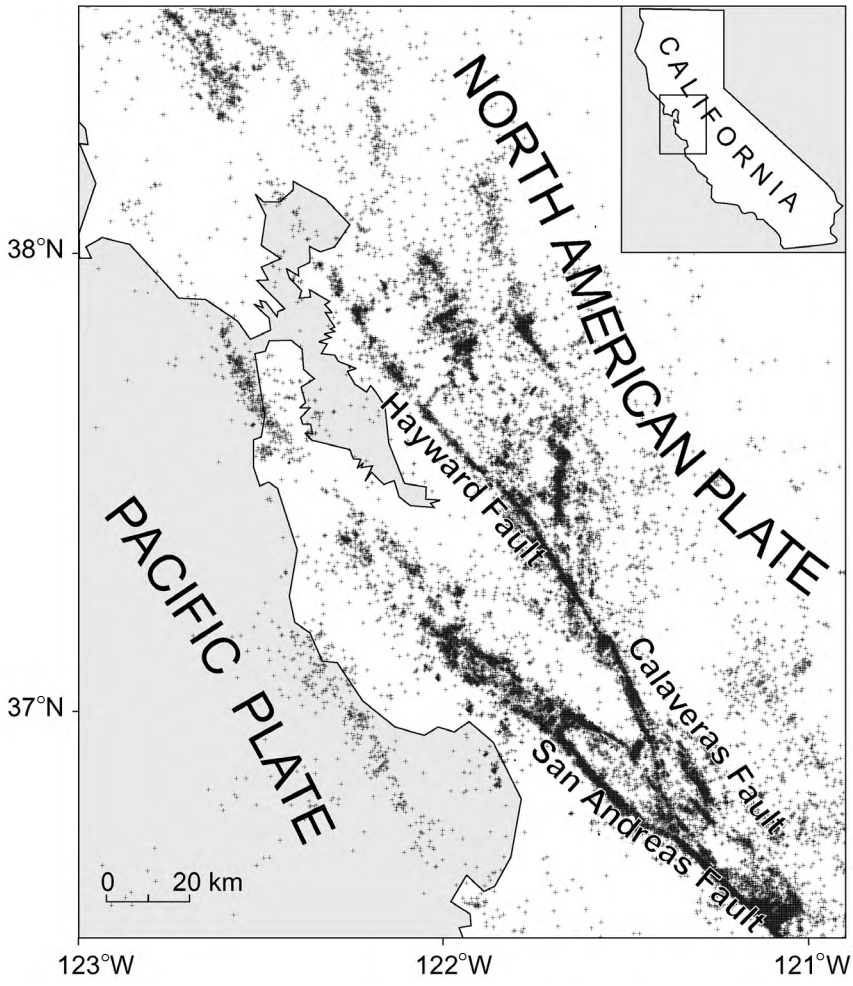


Figure 4.14. From 1968 to 1995 about 95,000 earthquakes with magnitude greater than 1 occurred in the San Francisco region. The main faults are well defined. (Reprinted with permission from *Nature* (Amelung and King, *Nature*, **386**, 702–5). Copyright 1997 Macmillan Magazines Ltd.)

than 8. These types of calculations, while straightforward to perform, should only be regarded as providing a broad indication of levels of earthquake activity.

Small b-values are usually characteristic of continental rifts and some regions with deep-focus earthquakes, whereas large b-values are typical of mid-ocean ridges and volcanic regions where volcanic swarms are common. Volcanic swarms can have b-values of up to 2.5, meaning that there are many small earthquakes and very few large events. Laboratory experiments indicate that high stress results in low b-values, and low stress results in high b-values. This observation is consistent with most of the regional determinations that have been made.

4.2.6 Energy released by earthquakes

Few determinations of the total energy released as seismic waves by earthquakes have been made, probably because the energy must be computed by an integration

over both time and space and must include all frequencies. To do this, the signal must be recorded with broad-band instruments[5] (instruments that record both high and low frequencies).

Consider a simple plane wave, a simple harmonic with displacement u:

$$u = A\ \cos(\kappa x - \omega t) \tag{4.25}$$

where ω is the angular frequency, κ the wavenumber (ω/wave-velocity), A the maximum amplitude and t time. The kinetic energy per unit volume associated with this wave is

$$\begin{aligned} \mathrm{KE} &= \frac{1}{2}\rho\left(\frac{\partial \mathrm{u}}{\partial \mathrm{t}}\right)^2 \\ &= \frac{1}{2}\rho A^2\omega^2 \sin^2(\kappa x - \omega t) \end{aligned} \tag{4.26}$$

where ρ is the density of the material. The kinetic energy varies from a maximum value of $\rho A^2\omega^2/2$ to a minimum value of zero during each cycle. For a perfectly elastic medium, energy is conserved (the sum of potential energy and kinetic energy is constant), so the total energy per unit volume, is equal to the maximum of the kinetic energy per unit volume, $\rho A^2\omega^2/2$. To calculate the total seismic energy released by an earthquake, it is necessary to integrate or sum an expression such as this over time, frequency and position around the epicentre. Because amplitude and magnitude are related logarithmically in Eq. (4.13), the expression of energy in terms of magnitude is also logarithmic (constants are determined from integrations as well as assumptions about density, etc.). Empirical relations between energy and magnitude are

$$\begin{aligned} \log_{10} E &= 4.8 + 1.5M_\mathrm{s} \\ \log_{10} E &= -1.2 + 2.4m_\mathrm{b} \end{aligned} \tag{4.27}$$

where E (measured in joules, J) is the total energy of the seismic waves. These relationships were determined for, and are therefore reliable for, earthquakes with magnitude >5 only. Note also that, as the m_b magnitude scale saturates, energy estimates will be too low at large magnitudes. An earthquake with $M_\mathrm{S} = 7.3$ would release 5.6×10^{15} J as seismic energy. Were this energy released by smaller earthquakes, some 75 magnitude-6 events, or 2000 magnitude-5 events, would need to occur. In comparison, the Hiroshima atomic bomb was approximately equivalent in terms of energy to an earthquake of magnitude 5.3. A one-megaton nuclear explosion would release roughly the same amount of energy as an earthquake of magnitude ~6.7. A very small magnitude-2 earthquake would release about as much energy as a bolt of lightning.

[5] Ideally, for local earthquakes frequencies in the range 0.1–500 Hz should be recorded, whereas for teleseismic recordings of surface waves 0.01–20 Hz would be appropriate. In contrast, for vibroseis (Section 4.4.1) deep reflection data in the range 5–200 Hz should be recorded and for explosion seismology (refraction lines, not nuclear explosion detection) 2–200 Hz. These frequency ranges are those suggested by the National Research Council (U.S.A.) in 1983.

We can use Eqs. (4.27) and (4.24) with Fig. 4.10 to estimate the energy released each year by earthquakes of various magnitudes and depth ranges. About three-quarters of the energy is released by shallow-focus events and the earthquakes with magnitude greater than 7 account for almost all the energy loss.

4.2.7 The observation of earthquakes: seismic phases

The various seismic-ray paths within the Earth are coded as follows:[6]

P	a P-wave in the mantle
S	an S-wave in the mantle
K	a P-wave through the outer core
I	a P-wave through the inner core
J	an S-wave through the inner core
c	a reflection from the mantle–outer-core boundary
i	a reflection from the outer-core–inner-core boundary
p	a P-wave reflected from the surface of the Earth close to the earthquake focus
s	an S-wave reflected from the surface of the Earth close to the earthquake focus
LR	a Rayleigh wave
LQ	a Love wave.

The first arrival at a seismometer is always P, closely followed by pP, the P-wave reflected once near the focus (Fig. 4.9). The S-waves arrive next and finally the surface waves.

When a body wave is incident on a discontinuity in the Earth, some of the energy is reflected as a P-wave, some is reflected as an S-wave, some is transmitted as a P-wave and some is transmitted as an S-wave. This is termed *mode*

Mode Conversion

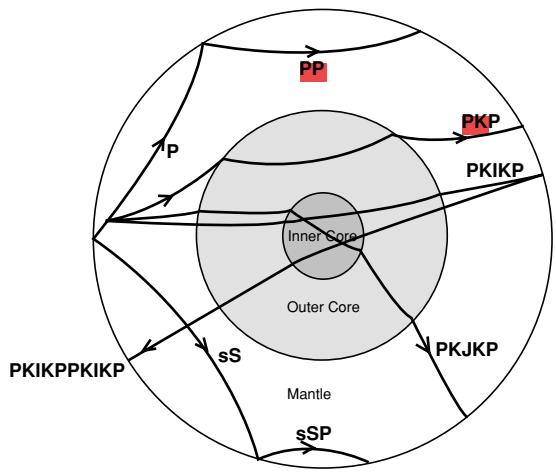


Figure 4.15. Some of the possible ray paths for seismic waves penetrating the Earth. In the mantle and core, the velocities increase with depth, so the ray paths bend away from the normal. The decrease in velocity at the mantle–core boundary causes those rays refracted into the core to bend towards the normal.

[6] The terminology P (primary), S (secondary) and L was introduced by von dem Borne (1904).

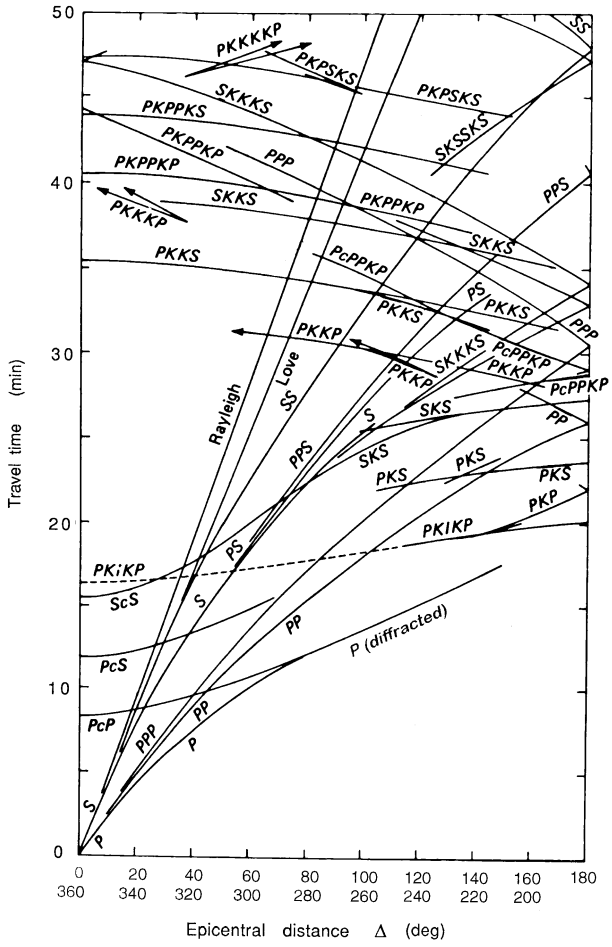


Figure 4.16. Jeffreys–Bullen (J–B) travel-time curves for an earthquake focus at the surface, with some modifications. Such curves are usually accurate to within a few seconds. The epicentral distance is the angle subtended by the earthquake epicentre and seismometer at the centre of the Earth. (After Bullen and Bolt (1985).)

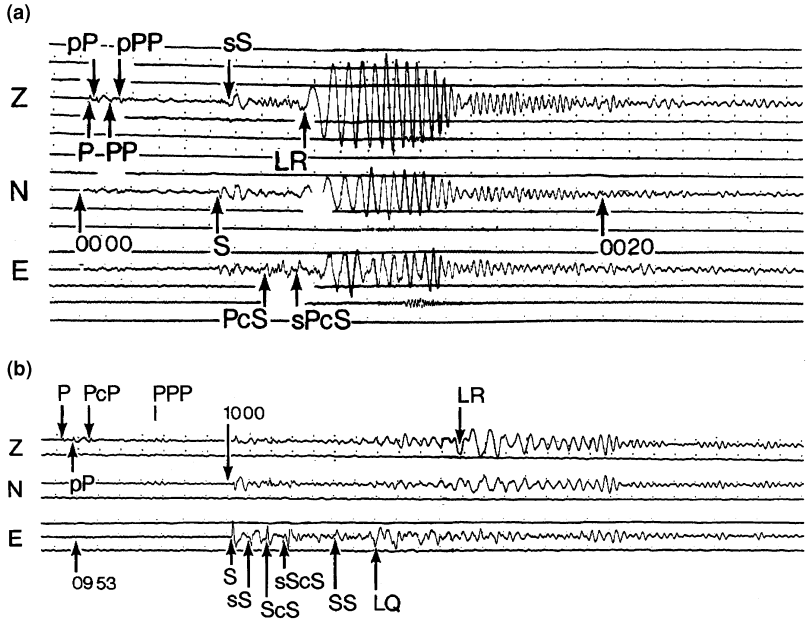


Figure 4.17. Examples of long-period earthquake records. Z is the vertical component and N (north) and E (east) are the two horizontal components. Time marks (dots) are every minute. The recordings are made in ink on a rotating drum – hence the other traces above and below the earthquake. (a) Offshore Nicaragua, focal depth 70 km, surface wave magnitude 5.7, recorded in New York. Note the clear dispersion of the Raleigh waves. (b) Northern Chile, focal depth 129 km, surface wave magnitude 6.5, recorded in New York. (From Simon (1981).)

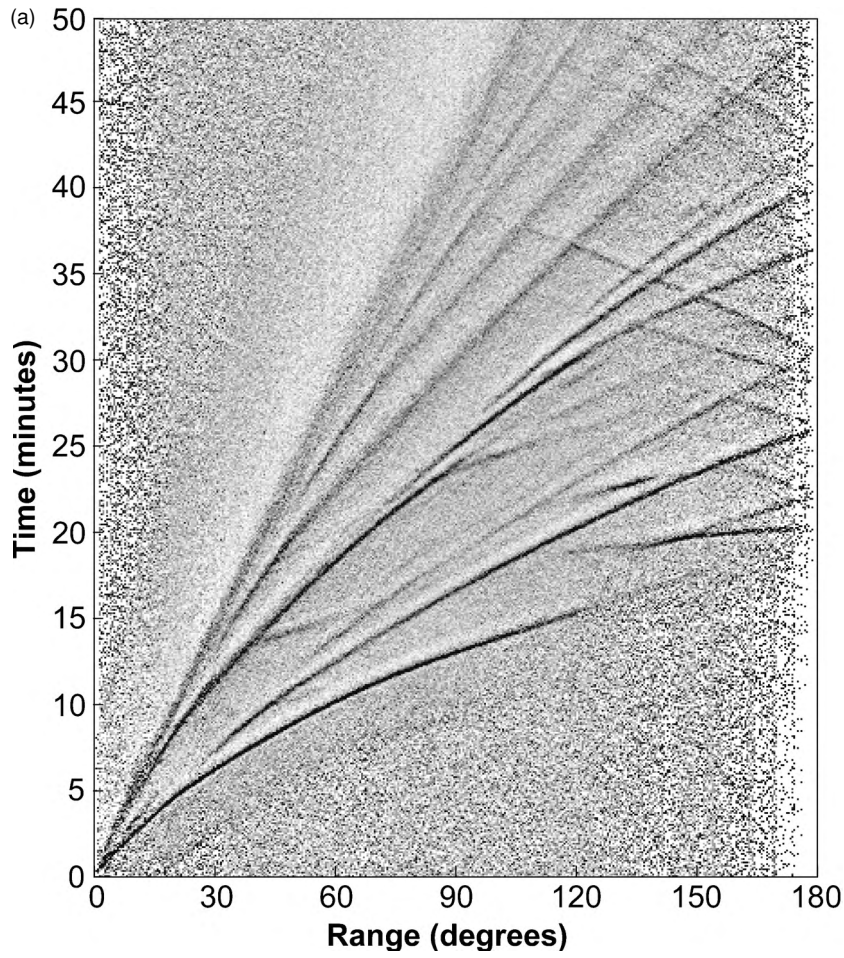


Figure 4.18. (a) Vertical-component long-period (∼20 s) seismogram travel-time curves. This plot has been generated using over 60,000 global records from 2995 earthquakes, which were recorded digitally between 1980 and 1987. The plot is dominated by P and S energy: P, PKP, PP, PPP, S, SKS, SKKS, etc. Faint reflections from the upper–lower–mantle discontinuity can also be seen. However, PcP and PcS reflections from the core–mantle boundary are not well imaged – shorter-period data are needed. (From Earle, P. S. and Shearer, P. M. Characterization of global seismograms using an automated-picking algorithm, *Bull. Seism. Soc. Am.*, **84**, 366–76, 1994 © Seismological Society of America.) (b) A stacked-record section of the Earth's seismic response at periods $>$ 60 s. The data were recorded between 1981 and 1991 by the IDA network, a global array of digital vertical-component instruments. Positive arrivals are black, negative arrivals white. The Rayleigh waves are imaged very well – note the clear dispersion (Section 4.3.1). (From Shearer, P. M. Imaging Earth's response at long periods, *EOS Trans. Am. Geophys. Un.*, **75**, 449, 1994. Copyright 1994 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

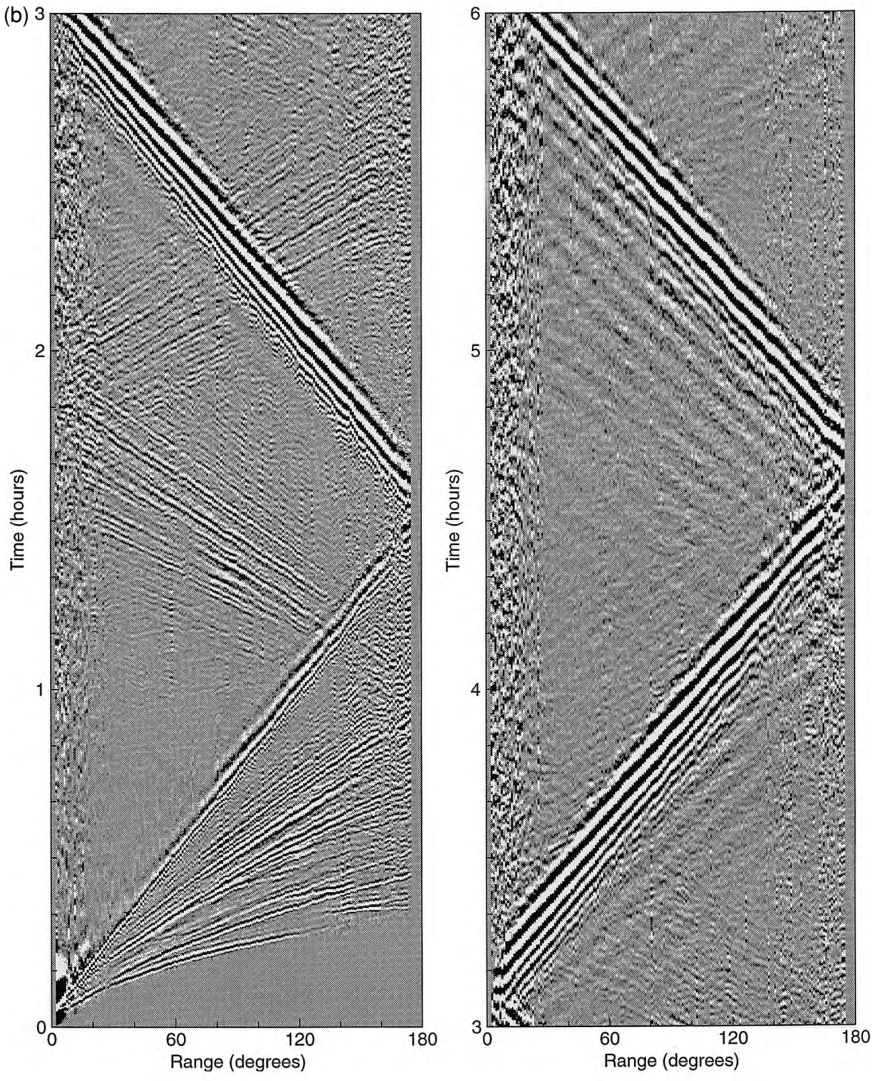


Figure 4.20. (*cont.*)

conversion. The amplitudes of the various waves are discussed in Section 4.3.5. Some examples of the coding of seismic waves follow:

PKP	a P-wave that passed down through the mantle and the outer core and then up through the mantle
PKIKP	a P-wave that passed down through the mantle and outer core and then through the inner core and up through the outer core and mantle

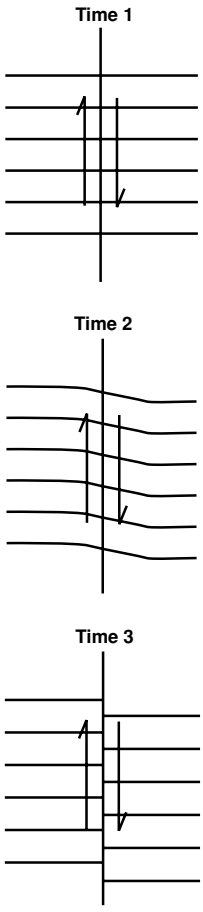


Figure 4.21. A simple model of the elastic-rebound theory of earthquake source mechanisms. Over time, one side of the fault is displaced relative to the other side (times 1 and 2). The deformation continues until the stresses on the fault are large enough to overcome the friction between the two blocks of material; then an earthquake (sudden displacement or rupture) occurs, and the strain is released (time 3). Thus, the size of the earthquake is directly related to the friction in the fault.

PKJKP	a wave that passes through the mantle and outer core as a P-wave, then through the inner core as an S-wave and up through the outer core and mantle as a P-wave
PKiKP	a P-wave that travelled down through the mantle and outer core, was reflected at the outer-core–inner-core boundary and travelled back up through the outer core and mantle
sSP	a wave that travelled from the focus as an S-wave, was reflected at the Earth's surface close to the focus, then travelled through the mantle as an S-wave, was reflected for a second time at the Earth's surface, converted to a P-wave and travelled as a P-wave through the mantle.

Figure 4.15 shows some ray paths (see also Section 8.1.1 and Figs. 8.1–8.3). Figure 4.16 is a travel-time plot for all the main earthquake phases. Identifying and recognizing the phases on an earthquake record is a skilled task (Fig. 4.17). Combining earthquake records, to produce a record section greatly facilitates identification of phases (Fig. 4.18). In Fig. 4.18(a) the main body-wave phases from the mantle and also some from the core are immediately apparent even to the untrained eye. The surface waves, which have greater amplitude and arrive later than the main body-wave phases, do not show up well on Fig. 4.18(a) but are very clear on the longer-period section shown in Fig. 4.18(b).

4.2.8 Earthquake mechanisms

The relative motions between plates on the Earth's surface, although regular over geological time, are not continuous over a daily or yearly period. Stress builds up along a fault, or in a region, over a period of years before it reaches some critical level; then an earthquake and, perhaps, aftershocks relieve it (Fig. 4.21). The length of the *fault plane*, along which the rocks are displaced, varies from metres, for a very small earthquake, to about 1000 km, for a very large earthquake. The 1960 Chile earthquake had a fault plane 1000 km in length, and the aftershock zone of the 1957 Aleutian earthquake ($M_w = 9.1$) was some 1200 km long, the longest aftershock zone known. Figure 4.22 illustrates the basic types of faults and the various names by which they are known. Most earthquakes occur along plate boundaries as a direct result of plate motions; these are *interplate* (between-plates) earthquakes. *Intraplate* (within-plate) earthquakes comprise only a small proportion of the total number occurring, but they can be large and can produce considerable damage. Examples of these include the sequence of great 1811 and 1812 earthquakes in New Madrid, Missouri, U.S.A., two of which must have had magnitudes of 7.3–7.5. These earthquakes, which have repeated about every 500 yr, are thought to be due to a heterogeneous zone in the lower crust that focusses deformation.

To determine the location of any particular earthquake, the seismic waves it produces must be observed at a number of recording stations around the world.

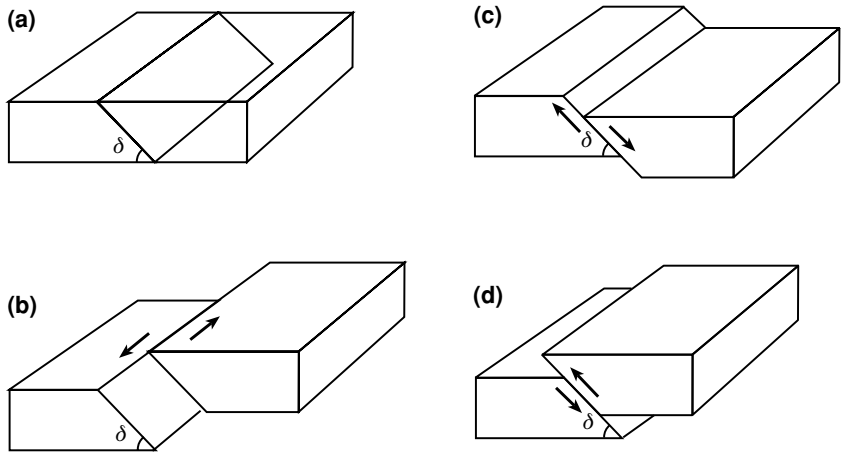


Figure 4.22. Fundamental types of faults and their various names. (a) A block model of a fault before any motion has taken place; δ is the angle of dip. (b) A strike–slip fault, transcurrent fault, lateral fault, tear fault, wrench fault; δ is often near 90°. The fault shown is left-handed or sinistral; a strike–slip fault with the opposite offset would be called right-handed or dextral. (c) A normal fault, dip–slip fault, normal-slip fault, tensional fault, gravity fault; generally, $45^\circ < \delta < 90^\circ$. (d) A thrust fault, dip–slip fault, reverse-slip fault, reverse fault, compressional fault; often, $0^\circ < \delta < 45^\circ$. Arrows show the slip vector, the relative motion between the two sides of the fault. The slip vector always lies in the fault plane. (From Båth (1979).)

The arrival times of the waves at the stations can then be used to locate the source (Section 4.2.1). In 1961, the United States Coast and Geodetic Survey established the World-Wide Standardized Seismograph Network (WWSSN). More recently a new Global Seismographic Network managed jointly by the United States Geological Survey (USGS) and Incorporated Research Institutions for Seismology (IRIS) has been established. Automatic digital recording by broad-band, low-gain and short-period instruments ensures that vibrations of all frequencies are recorded. Data are made available electronically to seismologists around the world for use in determination of the Earth's structure, mitigation of seismic risk and, importantly, as part of the international verification of the end of nuclear testing. Figure 2.1 shows the 1978–1989 distribution of epicentres. It is clear that seismicity accurately delineates many plate boundaries and that there are very few intraplate events.

Satellites are greatly improving our knowledge of the Earth. Synthetic-aperture radar interferometry (SAR) has provided a spectacular image of the displacement resulting from the 1992 Landers, California, earthquake. This $M_W = 7.3$, $M_S = 7.5$, right-lateral strike–slip earthquake ruptured some 85 km along a fault system, and resulted in a maximum slip of 6 m. By using two satellite radar images of the region, one taken before the earthquake, the other after, it was possible to calculate the small changes in the distance from the ground surface to the satellite

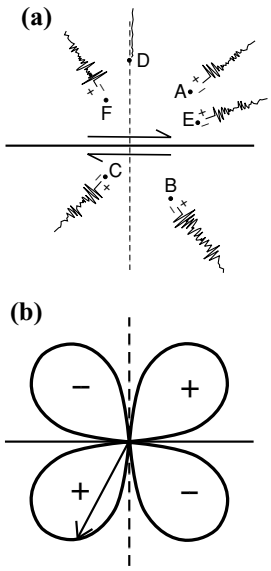


Figure 4.24. (a) A plane view of an earthquake strike–slip fault. A, B, C, D, E and F are six seismograph stations that recorded the earthquake. The first P-wave recorded at stations A, C and E would be compressional (positive, up); the first P-wave recorded at stations B and F would be dilatational (negative, down); station D would record no first P-wave arrival. (b) The distribution of polarity of the first P-wave motion falls into four quadrants. The lobes indicate the relative magnitude of the first motion at any location. The arrow shows the magnitude at location C.

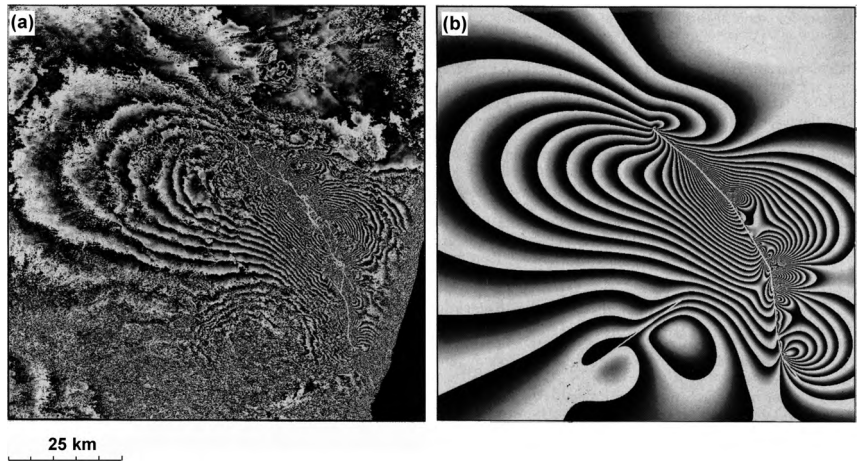


Figure 4.23. (a) A radar interferogram showing the change in distance from a satellite orbiting at an altitude of 785 km to the ground surface resulting from the Landers, California, earthquake. The two images used to construct the interferogram were taken on 24 April and 7 August 1992, the earthquake occurred on 28 June 1992. The $M_w = 7.3$, strike–slip earthquake ruptured 85 km along an almost vertical set of faults. (b) A synthetic interferogram calculated with the fault modelled as eight planar segments, which rupture from the surface to a depth of 15 km. Both images are 90 km × 110 km and one cycle of grey-scale represents a change in distance of 28 mm. Colour version Plate 5. (Reproduced with permission from *Nature* Massonnet *et al.*, *Nature*, **364**, 138–42) Copyright 1993 Macmillan Magazines Ltd.)

orbiting at an altitude of 785 km. The interferogram (Fig. 4.23, Plate 5), with each cycle of shading representing 2.8 cm of change in the ground-to-satellite distance, contours the spatial extent of the ground deformation very clearly. There are at least twenty fringes from the northern edge of the image to the cores of the lobes along the faults, representing a difference in distance of 56 cm – this agrees with the observed surface displacements. The area within 5–10 km of the rupture zone shows no organized fringe pattern, the signal is incoherent. This region was subject to large changes in satellite-to-ground distance, intense secondary faulting and block rotation. Space-geodetic studies will clearly become an increasingly powerful means for studying crustal deformation.

An earthquake occurring along the San Andreas Fault in California, U.S.A., is likely to have a different mechanism from one occurring on a subduction zone beneath Japan. By studying the direction of movement, or *polarity*, of the first seismic waves from an earthquake arriving at a number of seismograph stations distributed over the Earth's surface, both the type of earthquake and the geometry of the fault plane can usually be determined. To understand this, consider a simple strike–slip fault as shown in Fig. 4.24. Imagine that the world is flat and that seismograph stations A, B, C, D, E and F are located some distance from the fault where the earthquake occurred.

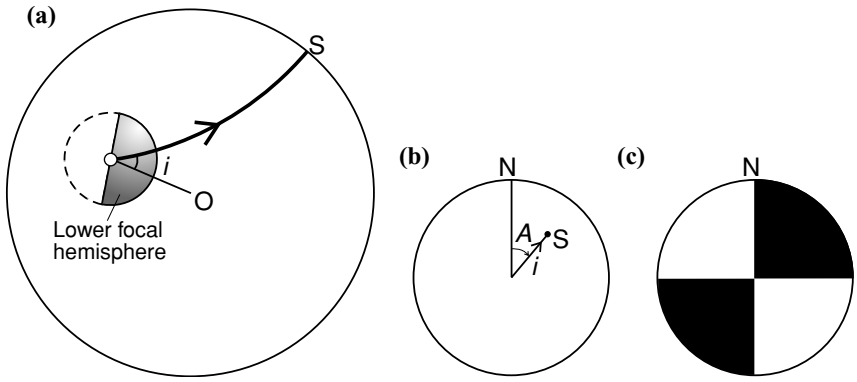


Figure 4.25. (a) A schematic cross section through the Earth, centre O. The focal sphere is an imaginary small sphere centred on the earthquake focus (○). The lower focal hemisphere is shaded. The ray path for the first arrival at seismograph S intersects the lower focal hemisphere at an angle *i* from the vertical. (b) Projection of the lower focal hemisphere onto a horizontal plane. N is north. Seismograph S plots at (*i*, *A*), where *i* is the angle of incidence shown in (a) and *A* is the geographic azimuth of S from the earthquake focus. The polarity of the first motion recorded at S is then plotted at (*i*, *A*). (c) The fault-plane solution for the strike–slip earthquake shown in Fig. 4.18(a). Compression, positive first motion, black; dilatation, negative first motion, white.

The first P-wave motion to arrive at any particular station from an earthquake on this fault will either be compressional (the first motion is a push) or dilatational (the first motion is a pull). At each of these six seismograph stations the polarity of the first-arriving energy is noted. Thus, the polarities at stations A, C and E are positive (compressional) and those at stations F and B are negative (dilatational), whereas station D receives no P-wave energy. If the distribution of compressional and dilatational first motions for many seismograph stations around the earthquake is plotted, it is seen that they fall into four quadrants (Fig. 4.22(b)), alternately positive and negative. The length of the arrow in the lobe represents the relative magnitude of the first P-wave at that location.

Since the Earth is spherical, in practice we must use spherical coordinates, which slightly complicates the geometry but not the results. Imagine a small sphere centred on and surrounding the focus of an earthquake. This sphere is known as the *focal sphere*. The ray that arrives first at each receiver intersects the *lower focal hemisphere* at an angle from the vertical *i* and at azimuth *A* (Fig. 4.23(a)). To calculate the angle *i* from the vertical at which the ray intersects the lower focal hemisphere, knowledge of the gross P-wave structure of the Earth is required in order to trace the ray back from receiver to source (see Section 8.1.1 and Appendix 3). However, in practice the calculations have already been done and the angle can be obtained from standard seismological tables. The lower focal hemisphere is then projected onto a horizontal plane using an equal-area

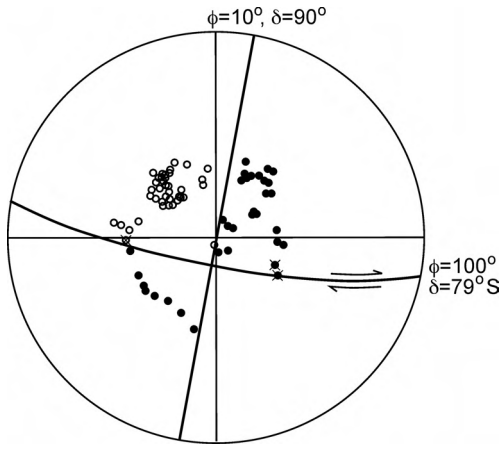


Figure 4.26. A fault-plane solution for a strike–slip earthquake. The strike and dip of the nodal planes are given by ϕ and δ. The east-striking nodal plane was chosen as the fault plane because the earthquake epicentre lies on an east-striking transform fault. The fault plane is not vertical in this instance; it is dipping at 79°. (After Sykes (1967).)

or a stereographic projection. It is obvious from Fig. 4.25 that the first P-wave arriving at a seismograph close to the earthquake focus must travel almost horizontally. Thus, this ray intersects the lower focal hemisphere almost at its equator, $i = 90^\circ$. In contrast, the direct P-wave to a station on the opposite side of the Earth to the earthquake travels almost vertically down from the focus and intersects the lower focal hemisphere near its centre (pole), $i = 0^\circ$. This means that close seismograph stations plot around the edge of the projection, and distant stations plot towards the centre. The azimuth of each seismograph station is easily measured geographically. The polarity (positive = compressional, or negative = dilatational) of the first motion recorded by each seismograph station is then plotted on the projection. In this way, data from seismograph stations around the world can be plotted on a graph that is a projection of the lower focal hemisphere. The right-handed or dextral strike–slip earthquake illustrated in Fig. 4.24 would have the *fault-plane solution* shown in Fig. 4.23(c). The four quadrants are separated by two orthogonal planes, or nodal surfaces, one of which is the *fault plane*. The other is called the *auxiliary plane*. There is no way of telling from the fault-plane solution alone which plane is the fault plane and which is the auxiliary plane. The radiation pattern shown in Fig. 4.23(c) could have been generated either by dextral strike–slip motion on a vertical fault plane striking 090° or by sinistral strike–slip motion on a vertical plane striking 000°.

If the fault plane is not vertical, the fault-plane solution for strike–slip motion still has four quadrants, but the fault plane and auxiliary plane do not plot as orthogonal straight lines passing through the origin. Instead, they plot as orthogonal great circles offset from the origin by $90^\circ - \delta$ where δ is their dip. An example is shown in Fig. 4.26. Again the fault-plane solution does not distinguish between the fault plane and the auxiliary plane. A normal-faulting earthquake (Fig. 4.25(a)) would have the fault-plane solution shown in Fig. 4.25(b). In contrast, a thrust-faulting earthquake and its fault-plane solution are shown in

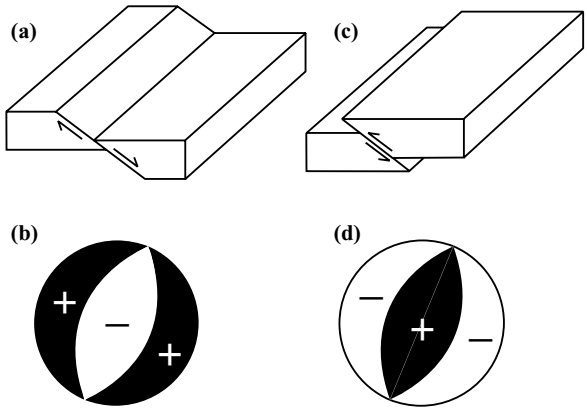


Figure 4.27. (a) A normal-faulting earthquake. (b) The fault-plane solution for the normal faulting earthquake in (a). The white region represents locations at which the first motion is dilatational (negative); black regions represent locations at which the first motion is compressional (positive). (c) A thrust-faulting earthquake. (d) The fault-plane solution for the thrust faulting earthquake in (c).

Figs. 4.25(c) and (d). In these cases, there is no ambiguity about the strike of the nodal planes but there is ambiguity about the dip of the fault plane. For a thrust fault, the first motion at the centre of the projection is always compressional, whereas for a normal fault the first motion at the centre of the projection is always dilatational. A vertical fault plane plots as a straight line passing through the origin, and the auxiliary plane plots around the circumference. Figure 4.26(a) shows an example of a predominantly normal-faulting event from the East African Rift. Figure 4.26(b) shows an example of a predominantly thrust-faulting event from the Macquarie Ridge (southwest of New Zealand). Both of these events had a small strike–slip component in addition to the main normal or thrust component, which means that there is ambiguity about the normal and auxiliary planes. Note that motion on a fault can have both normal and strike–slip components or both thrust and strike–slip components, but never normal and thrust components.

The *slip vector* **u** of the earthquake is the relative displacement which occurred between the two sides of the fault plane. It always lies in the fault plane (Fig. 4.22). The horizontal component of the slip vector u_h gives the azimuth of the relative horizontal motion occurring at the epicentre (Fig. 4.29). Although, for a pure strike–slip event such as that illustrated in Figs. 4.20(b) and 4.26, u_h is parallel to the fault plane, this is not the case in general. Consider the thrust-faulting earthquake shown in Fig. 4.25(a); in this case, u_h is perpendicular to the strike of the fault plane. For every fault-plane solution, however, the strike of u_h is *normal to the strike of the auxiliary plane*. Thus, the strike of u_h can be found by adding 90° to the strike of the auxiliary plane. The fault-plane solution cannot tell us anything about the magnitude of u_h, simply its direction.

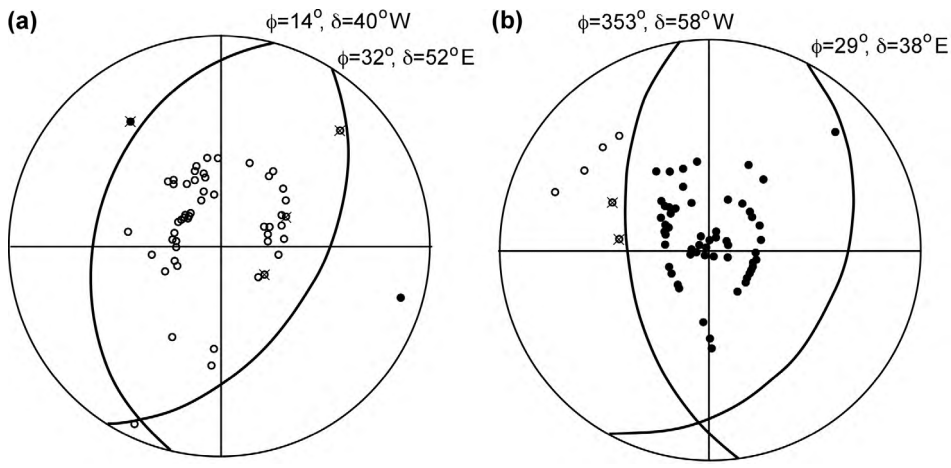


Figure 4.28. (a) A fault-plane solution for a predominantly normal-faulting earthquake that occurred on the East African Rift. (b) A fault-plane solution for a predominantly thrust-faulting earthquake that occurred on the Macquarie Ridge. Closed circles, first motion is compressional (positive); open circles, first motion is dilatational (negative); crosses, no clear first motion. (After Sykes (1967).)

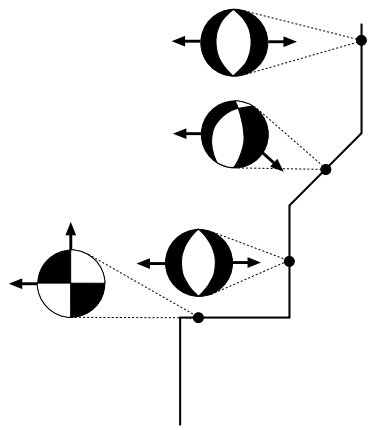


Figure 4.29. Fault-plane solutions for earthquakes (•) occurring along a plate boundary (solid line). Arrows indicate the azimuth of the possible horizontal components of the slip vector. The consistent arrows are the east–west set. The region is in extension, and the area is characterized by a mixture of strike–slip and normal faulting.

How to use earthquake first-motion directions to determine a fault-plane solution

To use earthquake first-motion directions you need to have tracing paper, pencil, eraser, pin and an equal-area projection net (also known as a Lambert or Schmidt equal-area projection net) (Figs. 4.28(a) and 4.54(c)). On that projection, great circles are the north–south lines of longitude; azimuth, or strike, is measured by counting lines of latitude clockwise around the edge of the projection.

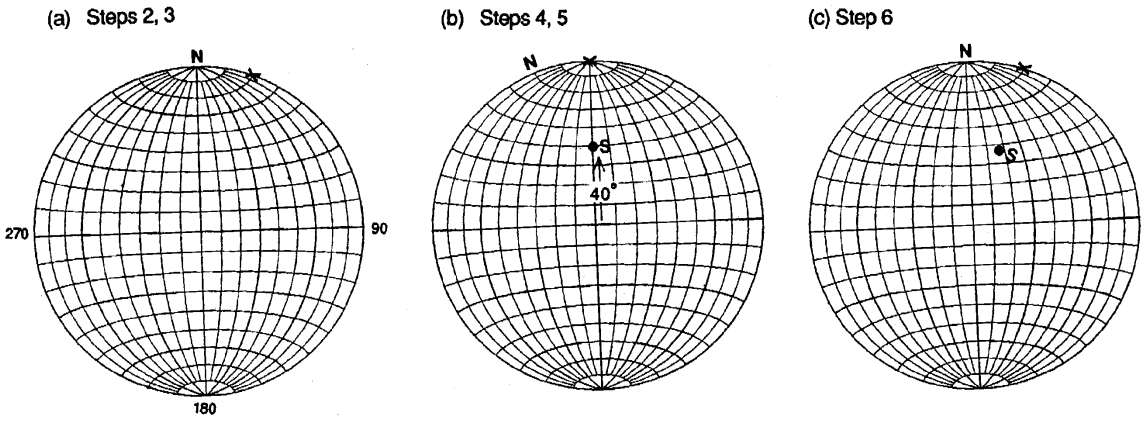


Figure 4.30. How to plot a seismograph station on an equal-area projection of the lower focal hemisphere.

Example 1

Plot seismograph station S (azimuth N20°E from the earthquake; angle of first-motion ray 40° from the vertical) on an equal-area projection.

1. Pin the tracing paper onto your base equal-area projection through its centre.
2. Mark N, the north axis.
3. Mark the azimuth N20°E as x at the edge of the projection (Fig. 4.28(a)).
4. Rotate the tracing paper so that x is at the top of the projection.
5. Mark the point S, 40° from the centre of the projection towards x (Fig. 4.28(b)).
6. Rotate the tracing paper so that N is again at the top of the projection (Fig. 4.28(c)).

S is now correctly plotted on the projection of the lower focal hemisphere.

Example 2

Find the great circle which joins the two points R and S.

1. Plot the points R and S as described in Example 1 (Fig. 4.29(a)).
2. Rotate the tracing paper until R and S both lie on the same great circle (i.e., they lie on the same line of longitude).
3. Trace that great circle (Fig. 4.29(b)).
4. Rotate the tracing paper so that N is again at the top of the projection (Fig. 4.29(c)).

The great circle which joins points R and S is now drawn.

Example 3

Determine the fault plane, auxiliary plane and slip vector for an earthquake, given the directions of first motion at several seismograph stations.

The fault plane and the auxiliary plane plot on the projection of the lower focal hemisphere as great circles. The fault plane and the auxiliary plane are

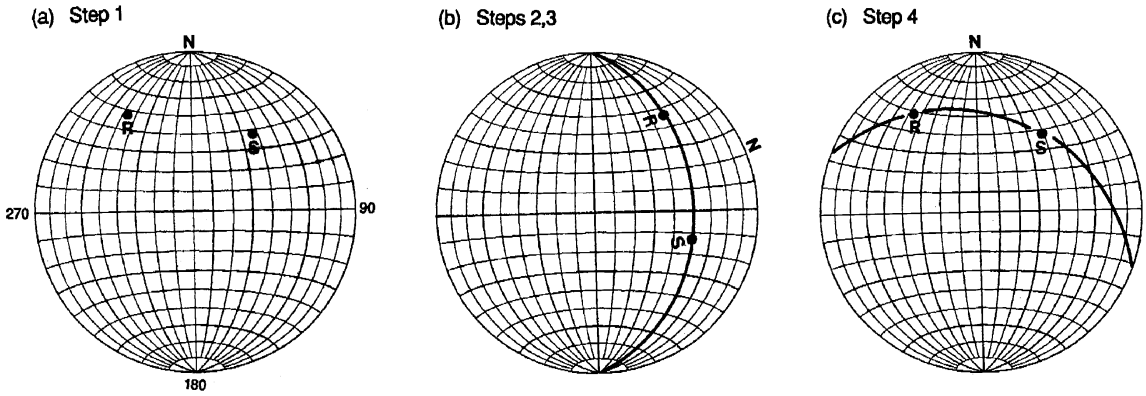


Figure 4.31. How to find the great circle which joins two points on the surface of the Earth.

orthogonal. This means that we need to find two orthogonal great circles that separate the positive first motions from the negative first motions.

1. Pin the tracing paper onto the base projection through its centre.
2. Put the first motions onto the tracing paper (Fig. 4.30(a)).
3. Rotate the tracing paper until you find a great circle that separates positive and negative first motions.
4. Trace that great circle; it is nodal plane 1.
5. The dip of nodal plane 1, δ_1, is measured along the equator of the base projection from the outside edge to the great circle. (A horizontal plane with zero dip plots around the edge of the projection. A vertical plane, with 90° dip, plots as a straight north–south line.)
6. Count 90° along the equator of the base projection from its intersection with nodal plane 1.
7. Mark that point P_1 on the tracing paper; it is normal to nodal plane 1 (Fig. 4.30(b)). (Sometimes the normal to a great circle is called the pole.)
8. Nodal plane 2 must separate the remaining positive and negative first motions, and, since it is also normal to nodal plane 1, point P_1 must lie on it. So, rotate the tracing paper until you find such a great circle.
9. Trace that great circle; it is nodal plane 2 (Fig. 4.30(c)). If you cannot find nodal plane 2, go back to step 3 and check that nodal plane 1 was correct.
10. Repeat steps 5–7 to find the dip of nodal plane 2, δ_2, and point P_2, the normal to nodal plane 2 (Fig. 4.30(c)).
11. Rotate the tracing paper so that N is again at the top of the projection.
12. The strike of the nodal planes is measured clockwise around the outside of the projection from N, 78° and 147° (Fig. 4.30(d)).
13. The slip vector is the normal to the auxiliary plane. Thus, if nodal plane 2 is the fault plane, point P_1 is the slip vector; and if nodal plane 1 is the fault plane, point P_2 is the slip vector. The strike of the horizontal component of the possible slip vector

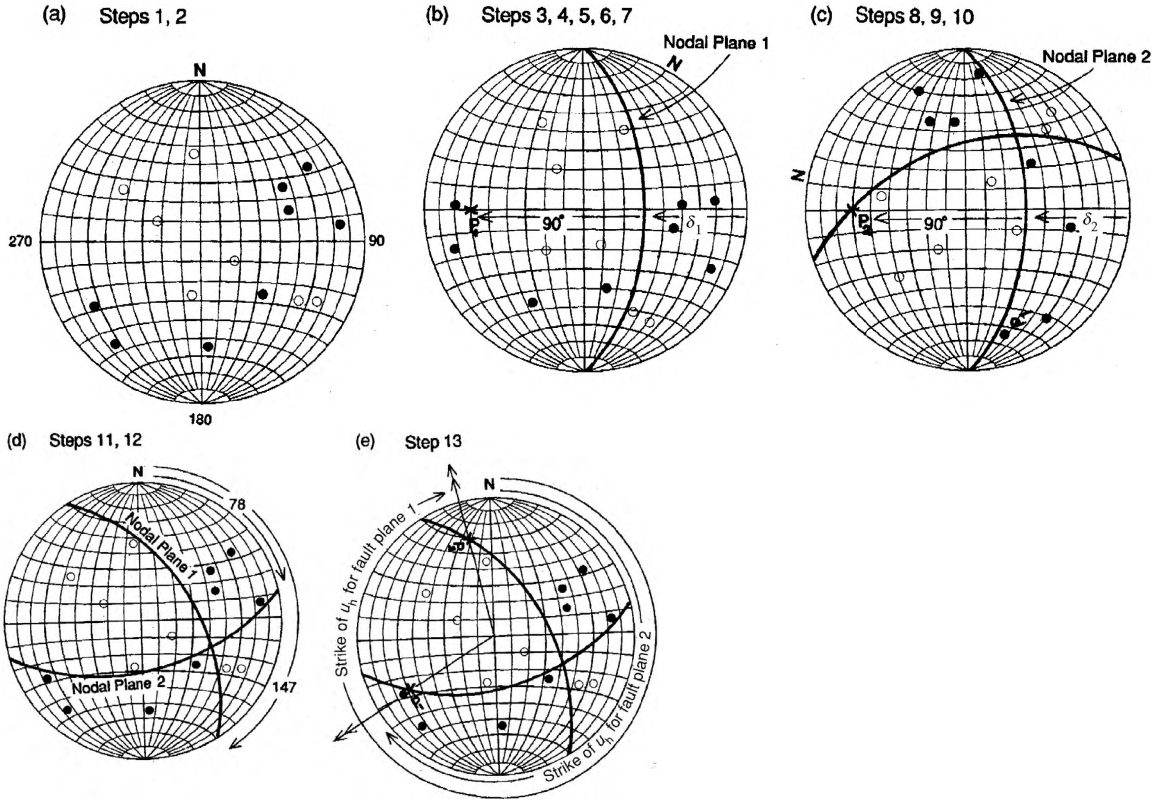


Figure 4.32. How to determine the fault plane, auxiliary plane and azimuth of the slip vector for an earthquake.

is measured clockwise around the outside of the projection from N, 238° and 348° (Fig. 4.30(e)).

This earthquake was therefore a combination of normal faulting and strike–slip on either a fault plane striking 78° and dipping at 60° with the strike of the horizontal component of slip 238°, or a fault plane striking 147° and dipping at 60° with the strike of the horizontal component of slip 348°.

For more details on using lower-focal-hemisphere projections, consult an introductory structural geology textbook or Cox and Hart (1986).

To distinguish between the nodal planes, it is often necessary to have some additional information. Sometimes local geology can be used to decide which nodal plane is most likely to have been the fault plane. On other occasions, the earthquake may show a surface break or fault, in which case the strike of that fault and the offset along it should agree with one of the nodal planes. Often though, this information is unavailable. McKenzie and Parker (1967) faced this problem in studying earthquakes in the Pacific. They plotted both of the possible horizontal components of the slip vector for a set of earthquakes that occurred

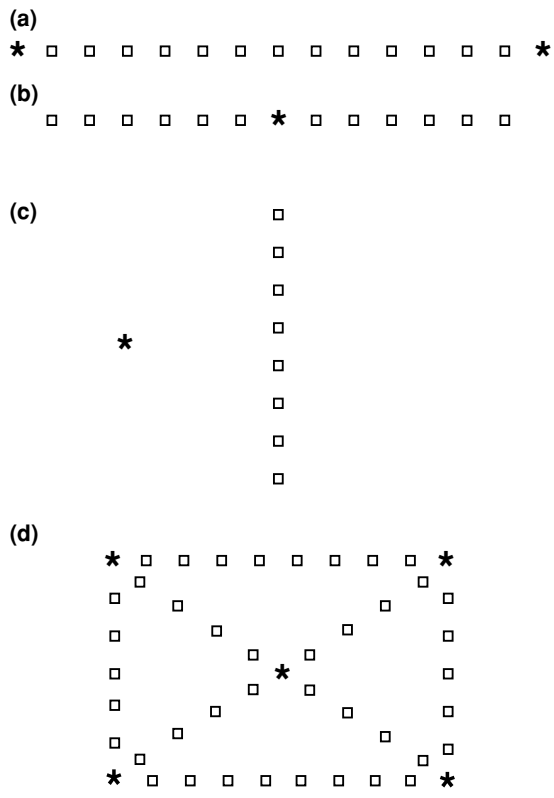


Figure 4.33. Examples of some layouts of energy sources (usually explosives) and seismometers used in refraction experiments: (a) a reversed refraction line; (b) a split profile; (c) fan shooting; and (d) a refraction experiment designed to determine the three-dimensional structure, including velocity anistropy. For refraction experiments on land, * is the explosive source and squares are the seismometer locations. For refraction experiments at sea, the source and receiver locations are exchanged: * is the seismometer/hydrophone location and squares are the locations of explosive sources.

along the northern boundary of the Pacific plate. They found that one set changed slowly and systematically in direction and the others showed no consistency. The correct horizontal components of slip were the consistent set, and so the fault and auxiliary planes were determined. An example of this method is shown in Fig. 4.29.

4.3 Refraction seismology

4.3.1 Refraction experiments

Earthquake seismology is able to reveal the broad details of the velocity and density structure of the Earth (Section 8.1), but to look in greater detail at the structure of the crust and uppermost mantle, especially in those regions well away from active seismic zones, it becomes impracticable to use earthquakes as the energy source. In refraction seismology, portable seismometers are deployed in the regions to be studied, and explosives are used as the energy source – careful logistical and environmental planning is necessary. Figure 4.33 shows some examples of the positioning of sources and receivers for such refraction experiments.

Refraction seismology on land is expensive and messy because tons of explosive are needed in order to make it possible to record energy at distances of

300 km or so, the distance necessary to determine a structure for the continental crust and uppermost mantle. It is necessary to have a seismometer at least every 5 km along each line, and preferable to have them very much closer together. Since the instruments are expensive, this can sometimes entail detonating one shot at the shotpoint, moving the seismometers along the line to new positions, and then detonating another shot at the shotpoint, and so on. Thus, large refraction experiments on land are usually carried out cooperatively by several universities or institutions so that enough people and recording instruments are available.

The situation at sea is different. In marine work, the source and receiver locations (Fig. 4.33) are usually exchanged so that a small number of seismometers or hydrophones (pressure recorders) is deployed and then a large number of shots is fired at ever increasing distances. Seismometers for use at sea are very expensive because they are necessarily enclosed in waterproof pressure vessels, have a homing detection device to ensure their subsequent recovery and, if laid on the seabed, must be equipped with a release device. Fewer people are needed for a marine experiment because the research ship can steam along, firing charges as it goes (no drilling needed here). At shorter ranges, an air gun (an underwater source that discharges a volume of very-high-pressure air) can be used as an energy source. Because the oceanic crust is much thinner than the continental crust, marine refraction lines need to be only 50 km or so in length to allow one to determine the crustal structure and an uppermost mantle velocity.

Most textbooks on exploration geophysics (e.g., Telford *et al.* 1990; Dobrin and Savit 1988) give the details of the field procedures and the corrections to be applied to seismic-refraction data. These are not discussed further here.

The coding used for crustal and uppermost-mantle phases is as follows:

P_g	P-wave through the upper continental crust[7]
S_g	S-wave through the upper continental crust
P_mP	P-wave reflection from the Moho
S_mS	S-wave reflection from the Moho
P_n	upper mantle P-head wave
S_n	upper mantle S-head wave.

4.3.2 A two-layered model

Let us assume that the crust beneath a refraction line consists of two horizontal layers, with distinct and constant P-wave velocities α_1 and α_2 such that $\alpha_2 > \alpha_1$ (Fig. 4.34). Energy from the source can then reach the seismometer by a variety of paths: directly through the top layer, by reflection from the interface between the two layers, by multiple reflections within the top layer or by travelling along the interface as a critically refracted wave or 'head' wave. The head wave, which is often called a refraction or refracted wave, has a travel time corresponding to

[7] The subscript g was first used to denote granite (Jeffreys 1926).

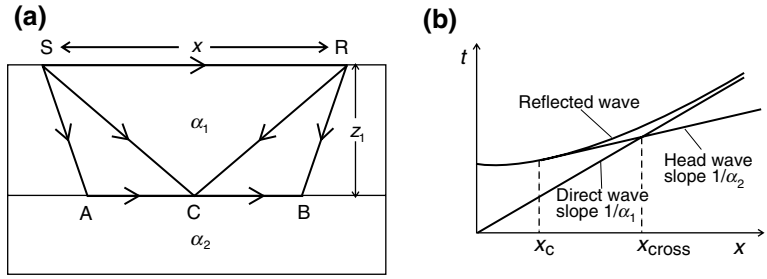


Figure 4.34. (a) Ray paths for seismic energy travelling from source S to receiver R in a simple two-layer model. The P-wave velocity is α_1 for the upper layer and α_2 for the lower layer, where $\alpha_2 > \alpha_1$. The direct wave takes ray path SR in the upper layer. The reflected wave takes ray path SCR in the upper layer. The head wave takes ray path SABR.
(b) Travel-time–distance plots for the model in (a).

a ray that travelled down to the interface at the *critical angle* i_c,[8] then along the interface with the velocity of the lower layer and then back up to the seismometer, again at the critical angle. (Head waves are second-order waves not predicted by geometrical ray theory.)

Direct wave

The time taken for energy to reach the receiver directly through the top layer is simply

$$t = \frac{x}{\alpha_1} \tag{4.28}$$

This is the equation of a straight line when time is plotted against distance.

Reflected wave

For the reflected-ray path, the travel time is

$$t = \frac{\mathrm{SC}}{\alpha_1} + \frac{\mathrm{CR}}{\alpha_1} \tag{4.29}$$

where SC and CR are the lengths of the ray paths shown in Fig. 4.34. Since the top layer is uniform, the reflection point C is midway between source S and seismometer R. Using Pythagoras' theorem, we can write

$$\mathrm{SC} = \mathrm{CR} = \sqrt{z_1^2 + \frac{x^2}{4}} \tag{4.30}$$

where z_1 is the thickness of the top layer.

[8] Snell's law for the interface between two media where i is the angle of incidence and r the angle of refraction is

$$\frac{\sin i}{\sin r} = \frac{\alpha_1}{\alpha_2}$$

When r is 90°, the angle of incidence $i = i_c = \sin^{-1}(\alpha_1/\alpha_2)$ is called the *critical angle*. For angles of incidence greater than the critical angle, no energy is refracted into the second layer.

The travel time t can then be expressed in terms of distance by substituting Eq. (4.30) into Eq. (4.29):

$$t = \frac{2}{\alpha_1}\sqrt{z_1^2 + \frac{x^2}{4}} \tag{4.31}$$

or

$$\alpha_1^2 t^2 = 4z_1^2 + x^2$$

which is the equation of a hyperbola.

Head wave or refracted wave

The travel time for the head wave is

$$t = \frac{SA}{\alpha_1} + \frac{AB}{\alpha_2} + \frac{BR}{\alpha_1} \tag{4.32}$$

By symmetry,

$$\begin{aligned} \mathrm{SA} &= \mathrm{BR} \\ &= \frac{z_1}{\cos i_c} \end{aligned} \tag{4.33}$$

and

$$\mathrm{AB} = x - 2z_1 \tan i_c \tag{4.34}$$

Substituting these expressions into Eq. (4.32) and using Snell's law ($\sin i_c = \alpha_1/\alpha_2$) gives

$$\begin{aligned} t &= \frac{2z_1}{\alpha_1 \cos i_c} + \frac{x}{\alpha_2} - \frac{2z_1}{\alpha_2}\tan i_c \\ &= \frac{2z_1}{\alpha_1 \cos i_c}\left(1 - \frac{\alpha_1}{\alpha_2}\sin i_c\right) + \frac{x}{\alpha_2} \\ &= \frac{2z_1}{\alpha_1}\cos i_c + \frac{x}{\alpha_2} \\ &= \frac{2z_1}{\alpha_1}\sqrt{1 - \frac{\alpha_1^2}{\alpha_2^2}} + \frac{x}{\alpha_2} \end{aligned} \tag{4.35}$$

On a time–distance (t–x) graph this is the equation of a straight line with slope $1/\alpha_2$ and intercept on the time axis $(2z_1/\alpha_1)\sqrt{1 - \alpha_1^2/\alpha_2^2}$.

Critical distance

Notice that the shortest range at which a head wave can be recorded is x_c, the critical distance, where

$$\begin{aligned} x_c &= 2z_1 \tan i_c \\ &= \frac{2z_1\alpha_1}{\sqrt{\alpha_2^2 - \alpha_1^2}} \end{aligned} \tag{4.36}$$

At ranges shorter than this, there is no head wave, only the precritical reflection from the interface. At the critical distance x_c, the travel time for the head wave is

the same as the travel time for the reflected wave (Fig. 4.34) since at this range the distance AB is zero, and these two ray paths are identical. The slope of the reflection hyperbola (Eq. (4.31)) is

$$\begin{aligned}\frac{\mathrm{d}t}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\alpha_1}\sqrt{4z_1^2 + x^2}\right) \\ &= \frac{x}{\alpha_1}\frac{1}{\sqrt{4z_1^2 + x^2}}\end{aligned} \tag{4.37}$$

At the critical distance ($x_c = 2z_1 \tan i_c$), this slope is

$$\begin{aligned}\left.\frac{\mathrm{d}t}{\mathrm{d}x}\right|_{x=x_c} &= \frac{2z_1 \tan i_c}{\alpha_1}\frac{1}{\sqrt{4z_1^2 + 4z_1^2 \tan^2 i_c}} \\ &= \frac{\sin i_c}{\alpha_1} \\ &= \frac{1}{\alpha_2}\end{aligned} \tag{4.38}$$

So, at the critical distance, the head wave is the tangent to the reflection hyperbola.

Crossover distance

The crossover distance x_{cross} is the range at which the direct wave and the head wave have the same travel time. It can be obtained from Eqs. (4.28) and (4.35):

$$\frac{x_{\text{cross}}}{\alpha_1} = \frac{x_{\text{cross}}}{\alpha_2} + \frac{2z_1}{\alpha_1}\sqrt{1 - \frac{\alpha_1^2}{\alpha_2^2}} \tag{4.39}$$

On rearranging, we obtain

$$x_{\text{cross}} = 2z_1\sqrt{\frac{\alpha_2 + \alpha_1}{\alpha_2 - \alpha_1}} \tag{4.40}$$

Time–distance graph

Figure 4.34 shows the travel time versus distance graph for this simple two-layer model. At short ranges, the first arrival is the direct wave, followed by the reflected wave; at long ranges, the first arrival is the head wave, followed by the direct wave, followed by the reflected wave.

To determine an initial velocity–depth structure from a refraction experiment, it is necessary to display the data on a time–distance graph. If we had a record section or first-arrival travel times from a seismic experiment shot over this model, we would determine α_1, α_2 and z_1, in that order, as follows.

1. α_1 is determined as the inverse of the slope of the direct-wave time–distance plot for distances less than x_{cross}.
2. α_2 is determined as the inverse of the slope of the head-wave time–distance plot for distances greater than x_{cross}.
3. z_1 is determined from the intercept of the head-wave time–distance line on the t axis (Eq. (4.35)); z_1 could also be calculated from the crossover distance (Eq. (4.40)), but this distance is not usually well enough defined to give an accurate value for z_1.

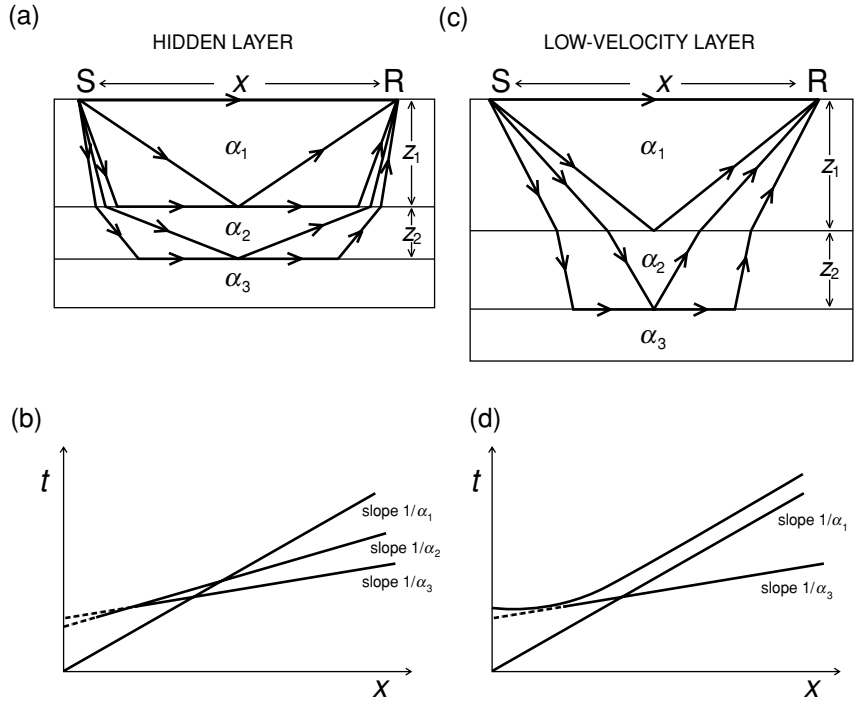


Figure 4.35. (a) A three-layer model in which $\alpha_3 > \alpha_2 > \alpha_1$. (b) The travel-time–distance plot for the model in (a). Only the direct wave (slope $1/\alpha_1$) and the two head waves (slopes $1/\alpha_2$ and $1/\alpha_3$) are shown. For distances less than the critical distance, the head-wave arrivals are extended back to the time axis by dashed lines. For this particular structure, the head wave from the second layer is always a second arrival; so this layer cannot be detected using first-arrival times alone. It is a hidden layer. (c) A three-layer model in which $\alpha_3 > \alpha_1 > \alpha_2$. (d) The travel-time–distance plot for the model in (c). Notice that the reflection from the interface at the base of the low-velocity layer is delayed with respect to the direct wave. When there is no low-velocity zone, this reflection and the direct wave are asymptotic (compare this with Fig. 4.32(b)).

Note that, although all of the equations have used P-wave velocities α_1 and α_2, they all hold for shear waves as well (velocities would then be β_1 and β_2). Shear waves are not usually studied in detail in crustal seismology, partly because of the extra expense of recording horizontal components as well as the vertical component but mainly because they are second arrivals; it is often very difficult to pick travel times accurately enough to make a good shear-wave velocity–depth model.

4.3.3 A multilayered model

The travel times for a model consisting of n uniform horizontal layers of thickness z_j and P-wave velocity α_j are determined in exactly the same way as for the two-layer model. The only extra matter to be remembered is that the rays bend according to Snell's law as they cross each interface ($\sin i/\alpha$ is a constant along each ray). The travel time for a wave refracted along the top of the mth layer is

$$t = \sum_{j=1}^{m-1} \left(\frac{2z_j}{\alpha_j} \sqrt{1 - \frac{\alpha_j^2}{\alpha_m^2}} \right) + \frac{x}{\alpha_m} \tag{4.41}$$

provided that $\alpha_j < \alpha_m$.

It is possible to determine correctly the velocities and thickness of all layers in an n-layer structure solely by using the first-arrival travel times if refracted arrivals from each interface are first arrivals over some distance range. If this is not the case and the refractions from a particular layer are always second arrivals, that layer is a *hidden layer*, and first arrivals alone will not give the correct structure (Figs. 4.33(a) and (b)). Another structure that cannot be determined uniquely by using first arrivals alone is illustrated in Figs. 4.33(c) and (d). A low-velocity layer cannot give rise to any head wave at its upper surface because the refracted ray bends towards the normal ($\alpha_{j-1} > \alpha_j$) rather than away from the normal. The only indications that a low-velocity layer is present come from the reflections from its upper surface, and the reflections and head wave from its lower surface. The critical distance for this lower interface is less than expected, and arrivals are larger in amplitude due to the large velocity contrast occurring there. This is exactly the same as the shadow-zone effect illustrated for the spherical Earth in Fig. 8.2.

Dipping layers

Real strata are far more complex than Fig. 4.32(a). If we allow the first interface to dip at an angle δ (Fig. 4.36) instead of being horizontal ($\delta = 0$) as in Fig. 4.34, then the travel time for the head wave is

$$t_{\mathrm{d}} = \frac{2z_{\mathrm{d}}}{\alpha_1}\sqrt{1 - \frac{\alpha_1^2}{\alpha_2^2}} + \frac{x}{\alpha_1}\sin(i_{\mathrm{c}} + \delta) \tag{4.42}$$

where z_{d} is the perpendicular distance from the shotpoint S to the interface. This is called shooting *down-dip*. Equation (4.42) is the equation of a straight line, but in this case the apparent head-wave velocity is α_{d}:

$$\alpha_{\mathrm{d}} = \frac{\alpha_1}{\sin(i_{\mathrm{c}} + \delta)} \tag{4.43}$$

α_{d} is less than α_2 ($=\alpha_1/\sin i_{\mathrm{c}}$). The fact that the interface is dipping cannot be determined from this time–distance graph alone. However, if the refraction line is 'reversed', that is, the shotpoint is placed at R and the receiver positions from R towards S, the travel times are

$$t_{\mathrm{u}} = \frac{2z_{\mathrm{u}}}{\alpha_1}\sqrt{1 - \frac{\alpha_1^2}{\alpha_2^2}} + \frac{x}{\alpha_1}\sin(i_{\mathrm{c}} - \delta) \tag{4.44}$$

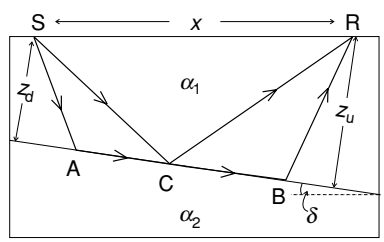


Figure 4.36. Ray paths for seismic energy travelling from source S to receiver R in a two-layer model in which the interface between the two layers dips at an angle δ. The P-wave velocity for the upper layer is α_1, and that for the lower layer is α_2, where $\alpha_2 > \alpha_1$.

where z_u is the perpendicular distance from the new shotpoint to the interface. This is called shooting *up-dip*. In this case, the apparent velocity is α_u:

$$\alpha_u = \frac{\alpha_1}{\sin(i_c - \delta)} \tag{4.45}$$

which is greater than the true velocity of the lower layer α_2. When a refraction line is reversed in this way, the true velocity of the dipping layer α_2 can be determined from Eqs. (4.43) and (4.45).

In one method, we rearrange Eqs. (4.43) and (4.45) and obtain

$$\sin(i_c + \delta) = \frac{\alpha_1}{\alpha_d}$$

and

$$\sin(i_c - \delta) = \frac{\alpha_1}{\alpha_u}$$

Hence

$$i_c + \delta = \sin^{-1}\left(\frac{\alpha_1}{\alpha_d}\right) \tag{4.46}$$

and

$$i_c - \delta = \sin^{-1}\left(\frac{\alpha_1}{\alpha_u}\right) \tag{4.47}$$

Therefore, i_c and δ are easily obtained by adding and subtracting Eqs. (4.46) and (4.47):

$$i_c = \frac{1}{2}\left[\sin^{-1}\left(\frac{\alpha_1}{\alpha_d}\right) + \sin^{-1}\left(\frac{\alpha_1}{\alpha_u}\right)\right] \tag{4.48}$$

$$\delta = \frac{1}{2}\left[\sin^{-1}\left(\frac{\alpha_1}{\alpha_d}\right) - \sin^{-1}\left(\frac{\alpha_1}{\alpha_u}\right)\right] \tag{4.49}$$

The velocity α_2 is known once i_c has been determined since i_c has been defined by the equation $\sin i_c = \alpha_1/\alpha_2$.

An alternative method that does not involve the use of inverse sines can be used in situations in which the dip angle δ is small enough for the approximations $\sin\delta = \delta$ and $\cos\delta = 1$ to be made (δ must be measured in radians, not degrees). In this case, by expanding $\sin(i_c + \delta)$ and $\sin(i_c - \delta)$ in Eqs. (4.43) and (4.45), we obtain

$$\begin{aligned}\frac{\alpha_1}{\alpha_d} &= \sin(i_c + \delta)\\ &= \sin i_c \cos\delta + \cos i_c \sin\delta\\ &= \sin i_c + \delta\cos i_c\end{aligned} \tag{4.50}$$

and

$$\begin{aligned}\frac{\alpha_1}{\alpha_u} &= \sin(i_c - \delta)\\ &= \sin i_c \cos\delta - \cos i_c \sin\delta\\ &= \sin i_c - \delta\cos i_c\end{aligned} \tag{4.51}$$

Now, i_c and δ can be obtained by adding and subtracting Eqs. (4.50) and (4.51):

$$\sin i_c = \frac{1}{2}\left(\frac{\alpha_1}{\alpha_d} + \frac{\alpha_1}{\alpha_u}\right) \tag{4.52}$$
$$= \frac{\alpha_1}{2}\left(\frac{1}{\alpha_d} + \frac{1}{\alpha_u}\right)$$

$$\delta = \frac{1}{2\cos i_c}\left(\frac{\alpha_1}{\alpha_d} - \frac{\alpha_1}{\alpha_u}\right) \tag{4.53}$$
$$= \frac{\alpha_1}{2\cos i_c}\left(\frac{1}{\alpha_d} - \frac{1}{\alpha_u}\right)$$

For small dip angles δ, α_2 is therefore given by

$$\frac{1}{\alpha_2} = \frac{\sin i_c}{\alpha_1}$$
$$= \frac{1}{2}\left(\frac{1}{\alpha_d} + \frac{1}{\alpha_u}\right) \tag{4.54}$$

If the dipping interface is not the first but a deeper interface, the velocities and thicknesses are determined exactly as above except that contributions from the overlying layers must be included in Eqs. (4.42) and (4.44). (The algebra can be somewhat tedious; see Slotnick (1959) for detailed derivations.)

4.3.4 Seismic record sections

When seismic-refraction travel-time data are plotted on a time versus distance graph, straight-line segments can be fitted to the points, and then velocities and layer thicknesses can be calculated. However, as has been mentioned, many structures are most unlikely to be detected by such a simple travel-time interpretation alone.

A good way to display all the data available is to plot a record section such as those shown in Figs. 4.18, 4.41, 9.6, 9.7, 9.27 and 10.12. The recording (*record*) from each seismometer is plotted at its appropriate range on a time–distance plot. The great advantage of plotting the data on a record section is that all of the amplitude and travel-time information is then displayed together. Phases can be easily correlated from trace to trace, which means that second arrivals and reflections, which often do not have a clear starting time and so would not necessarily be identified on a single record, stand out clearly. Hidden layers and low-velocity zones, which cannot be detected from the first-arrival travel times (Fig. 4.35), should be resolved when a record section is used in the interpretation of a refraction line.

To avoid having to plot record sections on exceedingly large pieces of paper or with a very small scale for time, *reduced record sections* are used. To do this, the time (t) axis of the plot is replaced by a reduced time ($t - x/v$) axis, where x is the offset distance (horizontal axis) and v is the *reduction velocity*. This means that phases arriving with a velocity v line up horizontally on the

reduced time–distance plot; phases with velocity less than v have a positive slope and phases with velocity greater than v have a negative slope. Reduced record sections greatly facilitate correlation of phases. Refraction data are generally plotted with a reduction velocity appropriate for the first-arrival velocities: long-range crustal/upper-mantle refraction lines would require a reduction velocity of perhaps 6.6 or 8 km s^{-1}, whereas a record section wth first arrivals from the deep mantle would need to be plotted with a reduction velocity of $\sim$14 km s^{-1}.

4.3.5 Amplitudes

Reflection and *refraction coefficients* Although we can determine a velocity–depth structure by using only the first-arrival travel times and the offset distances, much more information clearly is available from seismograms. Using amplitude information and the travel times of secondary phases can help immensely to minimize errors and to reduce the number of possible structures that could fit the first-arrival time–distance data. The examples of Fig. 4.35 showed common situations for which the first-arrival times do not tell the whole story.

To be able to use the information in the amplitudes and waveforms of seismic waves, it is necessary as a start to determine the amplitudes of reflected and refracted waves at interfaces. Figure 4.37 shows the rays reflected and transmitted at an interface between two media. Consider first the case of an incident P-wave, which gives rise to reflected and transmitted P-waves. This incident P-wave can also generate S-waves when it strikes the interface. Since the particle motion for P-waves is longitudinal, there is no motion in the direction perpendicular to the plane of the page, i.e., into or out of the page. This means that the S-waves generated by the incident P-wave cannot be SH-waves and must be SV-waves. Similarly, an incident SV-wave can generate reflected and refracted SV- and P-waves, but, in the case of an incident SH-wave, only SH-waves can be transmitted and reflected.

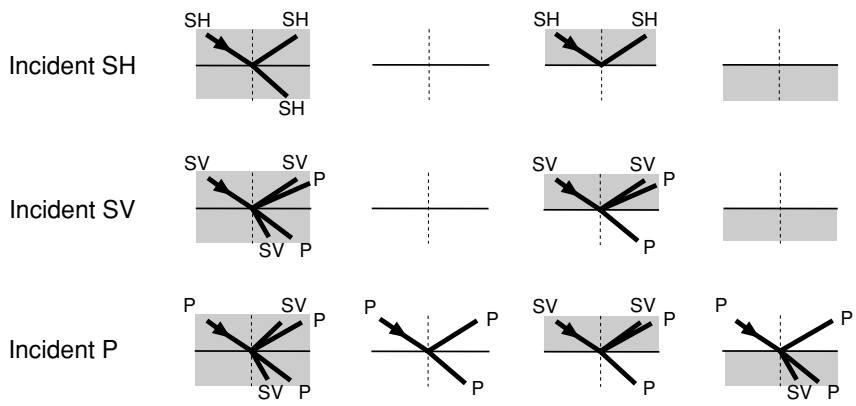


Figure 4.37. Reflection and transmission (refraction) of P-waves, vertically polarized shear waves (SV) and horizontally polarized shear waves (SH) incident from above on the boundary between two media (solid, shaded; liquid, white). Remember that S-waves cannot propagate in a liquid. Notice that incident P-waves can give rise only to reflected and transmitted P- and SV-waves and that incident SH-waves can give rise only to reflected and transmitted SH-waves. (After Båth (1979).)

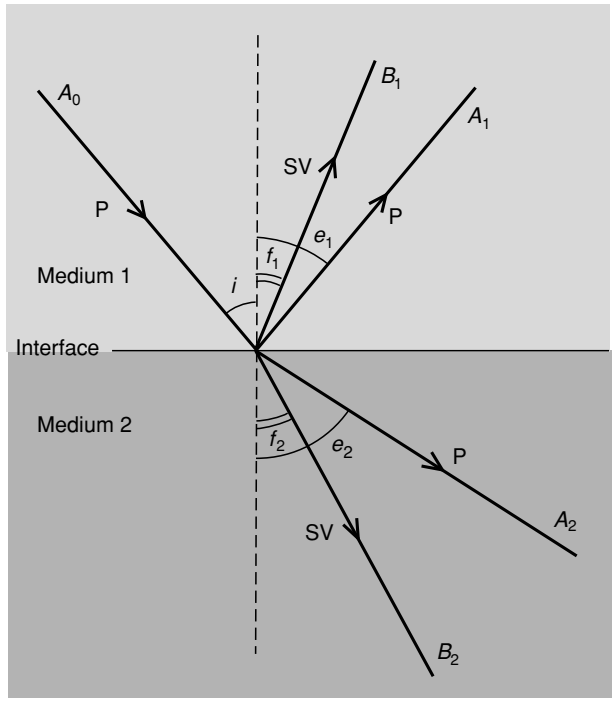


Figure 4.38. Waves generated at an interface between two elastic media by an incident P-wave. The incident P-wave has amplitude A_0 and angle of incidence i The reflected P- and SV-waves have angles of reflection e_1 and f_1 and amplitudes A_1 and B_1, respectively. The transmitted P- and SV-waves have angles of refraction e_2 and f_2 and amplitudes A_2 and B_2, respectively.

The angles made by the various rays with the normal to the interface are defined by Snell's law. The constant along each ray path, sin i/velocity, is often called p, the *ray parameter*. For the case of the incident P-wave (Fig. 4.38), the angles for the reflected and transmitted P- and SV-waves are therefore determined from

$$\frac{\sin i}{\alpha_1} = \frac{\sin e_1}{\alpha_1} = \frac{\sin f_1}{\beta_1} = \frac{\sin e_2}{\alpha_2} = \frac{\sin f_2}{\beta_2} = p \tag{4.55}$$

Clearly, for P-waves in the first layer, the angle of incidence is equal to the angle of reflection ($i = e_1$), and the other rays bend according to the seismic velocities of the media and the angle of incidence.

To determine the relative amplitudes of the reflected and transmitted waves, it is necessary to calculate the displacements and stresses resulting from the wave fields and to equate these values at the interface. Displacement and stress must both be continuous across the interface or else the two layers would move relative to each other. The following equations, which relate the amplitudes of the various waves illustrated in Fig. 4.38, are called the *Zoeppritz equations*:

$$A_1 \cos e_1 - B_1 \sin f_1 + A_2 \cos e_2 + B_2 \sin f_2 = A_0 \cos i \tag{4.56}$$

$$A_1 \sin e_1 + B_1 \cos f_1 - A_2 \sin e_2 + B_2 \cos f_2 = -A_0 \sin i \tag{4.57}$$

$$A_1 Z_1 \cos(2f_1) - B_1 W_1 \sin(2f_1) - A_2 Z_2 \cos(2f_2) - B_2 W_2 \sin(2f_2) = -A_0 Z_1 \cos(2f_1) \tag{4.58}$$

$$A_1 \gamma_1 W_1 \sin(2e_1) + B_1 W_1 \cos(2f_1) + A_2 \gamma_2 W_2 \sin(2e_2) - B_2 W_2 \cos(2f_2) = A_0 \gamma_1 W_1 \sin(2i) \tag{4.59}$$

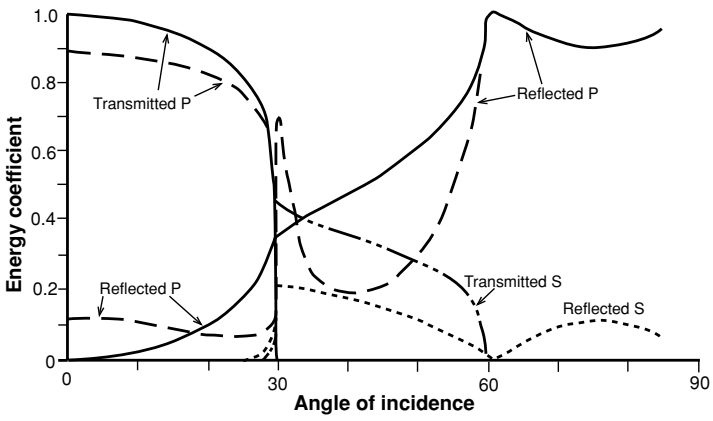


Figure 4.39. Reflected and transmitted energy (energy/incident energy) for the situation illustrated in Fig. 4.34: a P-wave incident on an interface between two solid media. For this example, $\alpha_1/\alpha_2 = 0.5$, $\rho_1/\rho_2 = 2.0$, $\alpha_1/\beta_1 = 1.87$ and $\alpha_2/\beta_2 = 1.73$. Solid lines, reflected and transmitted P-wave; short-dashed line, reflected S-wave; dashed and dotted line, transmitted S-wave; and long-dashed lines, reflected and transmitted P-wave when $\rho_1/\rho_2 = 1.0$. (After Tooley *et al.* (1965).)

where

$$Z_1 = \rho_1\alpha_1, \qquad Z_2 = \rho_2\alpha_2$$
$$W_1 = \rho_1\beta_1, \qquad W_2 = \rho_2\beta_2$$
$$\gamma_1 = \frac{\beta_1}{\alpha_1}, \qquad \gamma_2 = \frac{\beta_2}{\alpha_2}$$

Equations (4.56) and (4.57) come from requiring continuity of the vertical and horizontal displacements, and Eqs. (4.58) and (4.59) are required by continuity of the vertical and horizontal stresses on the boundary. Z and W, the product of the density and seismic velocity, are termed the impedance.

The ratio of the reflected or transmitted amplitude to the incident amplitude is called the *reflection* or *transmission coefficient*. (Unfortunately, and confusingly, the fractions of the incident energy which are reflected or transmitted are also sometimes referred to as the reflection and transmission coefficients.) Calculation of reflection and transmission coefficients for anything other than normal incidence is lengthy because the equations have to be solved numerically.

Figure 4.39 illustrates the reflection and transmission of energy for a P-wave incident on a solid–solid boundary. For this example there are two critical angles, one for P-waves and the other for S-waves, at 30° and 60°. Beyond the first critical angle, that for P-waves ($\sin 30^\circ = \alpha_1/\alpha_2 = 0.5$), there are no transmitted P-waves. The reflected P-wave energy increases greatly as the angle of incidence increases towards the first critical angle; these are the *wide-angle reflections* which are used extensively in seismic-refraction work to determine critical distances. Similarly, beyond the critical angle ($\sin^{-1}(\alpha_1/\beta_2)$) for S-waves at 60° there are no transmitted S-waves.

Reflection and transmission coefficients take a very much simpler form for the case of normal incidence on the boundary ($i = 0$). In this case,

$$B_1 = B_2 = 0$$

Since the tangential displacements and stresses on the boundary are zero for a vertically travelling P-wave, no SV-wave can be produced. Equations (4.56)–(4.59) reduce to

$$A_1 + A_2 = A_0 \tag{4.60}$$

$$Z_1 A_1 - Z_2 A_2 = -Z_1 A_0 \tag{4.61}$$

The amplitude ratios A_1/A_0 and A_2/A_0 are the reflection and transmission coefficients. Equations (4.60) and (4.61) are solved for the amplitude ratios:

$$\begin{aligned} \frac{A_1}{A_0} &= \frac{Z_2 - Z_1}{Z_2 + Z_1} \\ &= \frac{\rho_2\alpha_2 - \rho_1\alpha_1}{\rho_2\alpha_2 + \rho_1\alpha_1} \end{aligned} \tag{4.62}$$

$$\begin{aligned} \frac{A_2}{A_0} &= \frac{2Z_1}{Z_2 + Z_1} \\ &= \frac{2\rho_1\alpha_1}{\rho_2\alpha_2 + \rho_1\alpha_1} \end{aligned} \tag{4.63}$$

If Fig. 4.35 is studied again in the light of the foregoing discussion of reflection and transmission coefficients, it should be clear that, although the first arrivals should be the easiest phases to pick accurately, from the critical distance onwards the largest-amplitude events are the reflections. Thus, in attempts to determine a crustal structure, the reflections from major discontinuities in the crust are very important. Normally, the main reflection seen on refraction profiles is P_mP, the P-wave reflection from the Moho (the crust–mantle boundary). At the critical distance for P_n (the mantle head wave), the amplitude of P_mP is large, which often helps to constrain the mantle velocity since P_n does not usually become a first arrival on continental lines until perhaps 200 km, at which distance from the source low amplitudes relative to background noise may make accurate picking of arrival times difficult. At long range, the reflection P_mP is asymptotic to the head wave from the layer above the mantle (Fig. 4.32(b)), so the large amplitude of P_mP at long ranges is very useful in determining the lower-crustal velocity.

Amplitude–distance relations

In addition to reflection and transmission coefficients, it is useful to understand how the amplitude of waves decreases with increasing distance from the source. We have already seen in Section 4.1.3 that the amplitude of body waves varies as x^{-1} with distance and that the amplitude of surface waves varies only as $x^{-1/2}$.

For an interface between two uniform elastic media, the amplitude of a head wave decreases rapidly with increasing distance as $L^{-3/2}x^{-1/2}$ where $L = x - x_c$,

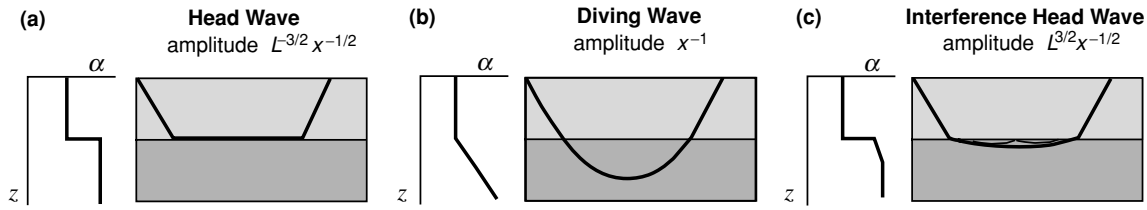


Figure 4.40. Amplitude–distance behaviour for head waves, diving waves and interference head waves. (From Kennett (1977).)

x is the distance and x_c is the critical distance (Fig. 4.38(a)). This relationship holds for values of L greater than five or six times the predominant wavelength of the seismic signal (i.e., for 8-km s^{-1} material and a 5-Hz signal, $L > 10$ km). At distances much greater than the critical distance, the amplitude of the head wave varies as x^{-2}.

In the case of continuous refraction in a velocity gradient (Fig. 4.38(b)), the refracted wave is a body wave, so its amplitude varies as x^{-1}. This type of arrival therefore dominates head waves at long distances.

However, if the velocity beneath the interface $\alpha(z)$ continues to increase with depth z as

$$\alpha(z) = \alpha_0(1 + bz) \tag{4.64}$$

the resultant head wave has an interference character (Fig. 4.38(c)). The first arrival is no longer the head wave but a wave continuously refracted below the interface. This is closely followed by a succession of continuously refracted waves, which have been reflected from the undersurface of the interface. Those that arrive within ΔT of each other, where ΔT is the duration of the incident signal, interfere and are termed the *interference head wave*. For $Lb^{2/3}\lambda^{-1/3} > 1$, where λ is the wavelength just below the interface, the amplitude of this interference head wave increases with range as $L^{3/2}x^{-1/2}$. At long ranges, when the first arrival arrives more than ΔT before the interference packet, the amplitude of the first arrival varies as x^{-1}. It should be clear, therefore, that introduction of velocity gradients can have a considerable effect on amplitude–distance behaviour.

Synthetic refraction seismograms

The best way to determine a velocity structure from seismic-refraction data is to compute synthetic (theoretical) seismograms and to compare these with those recorded in the refraction experiment. The velocity model can then be adjusted,

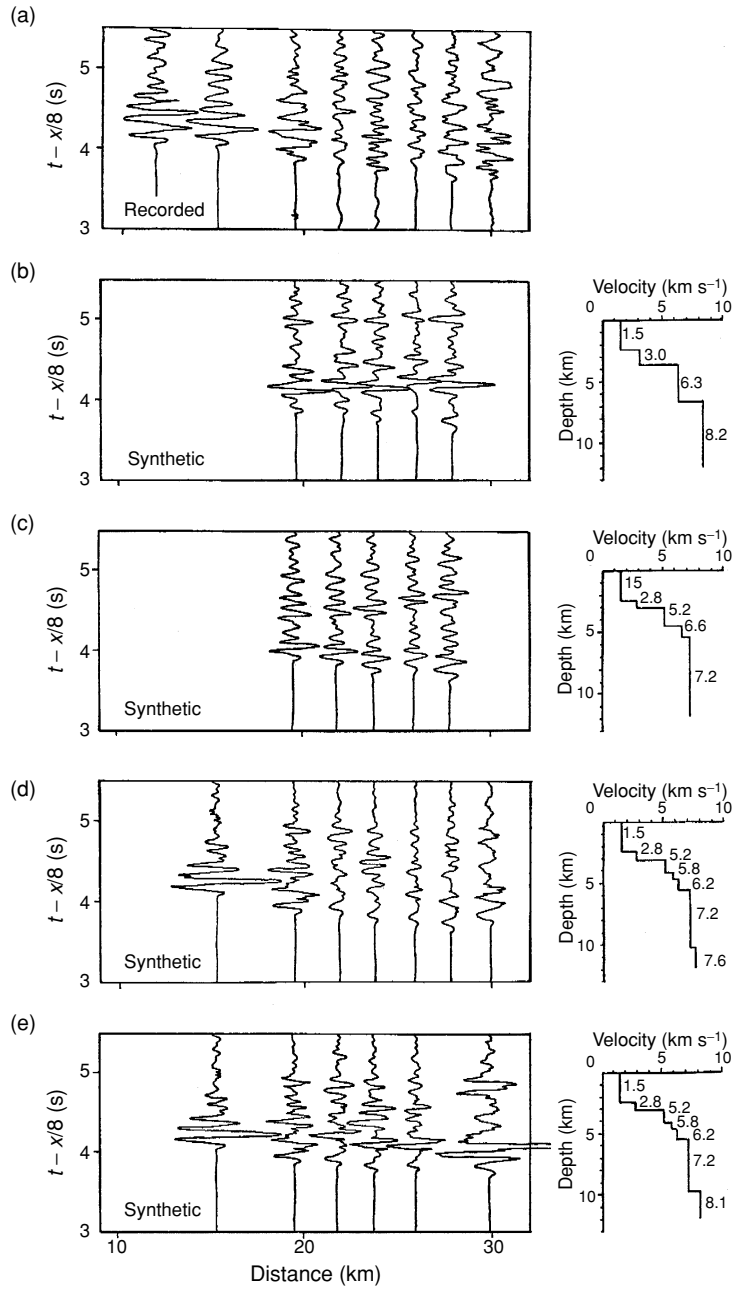


Figure 4.41. Use of synthetic seismograms in the determination of oceanic crustal structure. Synthetic seismograms were calculated using the 'reflectivity' method for four oceanic crustal models, all of which have first-arrival travel times that fit the observed times. Each seismogram is calculated for the charge size used in the original experiment – hence the variable amplitude–distance behaviour. All record sections are reduced to 6 km s^{-1}. All models include 2.6 km of sea water

velocities and thickness changed, or layers inserted or removed, and the computations repeated until the synthetic seismograms match the recorded seismograms to the desired degree. It is not usually possible to match every detail of a seismogram, but it is possible to match the arrival times and relative amplitudes of the main phases in a refraction record section and probably to match a few of the waveforms well. It should be stressed that this is still not an easy, quick or cheap process, even with increasing computer power.

Figure 4.41 shows examples of velocity models that agree with the first-arrival travel times but for which the synthetic seismograms do not match the recorded seismograms. For the final velocity model, both travel times and seismograms match well. The importance of synthetic seismograms in the determination of velocity structures should be evident from this example.

To match synthetic seismograms with recorded seismograms, one must compute the synthetics to include (a) the source function (explosion, earthquake, airgun or vibrator); (b) the effects of transmission through the Earth; and (c) the response of the detection systems (seismometers, amplifiers, filters, tape recorders, etc.).

The most difficult and time-consuming part of the computation of synthetic seismograms is the calculation of the effects on the seismic waveform caused by its passage through the Earth. There are undoubtedly lateral variations in the crust and upper mantle, but, provided that these are gradual and small compared with the wavelength of the incident seismic signal and the horizontal scale of the proposed or possible structure, they can often be approximated by a model in which the structure varies only with depth. This variation of α, β and ρ with depth can always be approximated by a large stack of uniform horizontal layers. Because computation of the total response, which means including all interconversions and reverberations of such a stack of layers, is costly, various approximations have been developed. Some of the approximations permit inclusion of lateral variations in the model.

($\alpha = 1.5\,\mathrm{km\,s^{-1}}$) and a thin seabed layer of fractured material ($\alpha = 2.8$–$3.0\,\mathrm{km\,s^{-1}}$). (a) The recorded record section is part of a refraction line shot along the median valley of the Mid-Atlantic Ridge at 37°N; these were some of the data used for the model in Fig. 9.22. (b) A model with a layer 3 and normal upper mantle. The large-amplitude arrivals at 4 s are the reflections from the 6.3/8.2-km s^{-1} interface. (c) The amplitude and complexity of the first arrivals have been increased. There are no large-amplitude reflections because the velocity contrast at the interfaces is much less than in (b). (d) These seismograms match the recorded record section (a) best; the model was obtained from model (c) by making slight changes in layer thicknesses and velocities and by including 7.6-km s^{-1} material at depth. (e) Replacing the 7.6-km s^{-1} material of model (d) with material with a normal upper-mantle velocity of 8.1 km s^{-1} dramatically changes the seismograms. The large-amplitude arrival is the reflection from the 7.2/8.1-km s^{-1} interface.

In the generalized (or exact) and WKBJ (Wenzel, Kramers, Brillouin and Jeffreys) ray methods (e.g., Helmberger 1968; Chapman 1978), the total response of a stack of layers is expressed as an infinite sum of the impulse responses of partial rays. Each partial ray is associated with a different propagation path through the structure, and together they include all interconversions and reverberations. The final synthetic seismogram is obtained by convolving the impulse response sum with the source and receiver functions. The problem with these methods comes from the termination of this infinite sum of impulse responses: rays that would make a significant contribution may inadvertently be excluded from the calculation. Often only primary reflections are considered, making this a cheap method.

The classical ray method allows the inclusion of lateral variation in velocity and of curved interfaces (e.g., Cerveny 1979). In this method, the particle displacement is expressed as an infinite sum, known as a ray series. For seismic body waves (reflected and refracted P- and S-waves), it is often sufficient to consider only the first term in the ray series: this is the solution obtained according to the principles of geometrical optics. This approximation is not valid near critical points, caustics or shadow zones; corrections must be applied there. Ray paths to be included in the computation must be specified (in practice, frequently only primary rays are considered, multiples being neglected). This method is not as accurate as the generalized ray method but has the significant advantage of including the effects of lateral variation.

The Maslov integral method (Chapman and Drummond 1982), an asymptotic high-frequency extension of ray theory, is useful in that it includes the signals in shadow zones and at caustics, but has the disadvantage of serious problems near pseudo-caustics (Kendall and Thompson 1993). However, an extension of the Maslov method in which a Kirchoff integration is used to reduce the problems associated with pseudo-caustics (Huang *et al*. 1998) is a considerable improvement.

The most comprehensive method for simple structures, although also the most costly, is the reflectivity method (Fuchs and Müller 1971; Müller 1985). In this technique, all the calculations are performed in the frequency domain, and all reflections and interconversions are included by using a matrix formulation for the reflection and transmission coefficients. The attenuation Q can also be specified for each layer (Kennett 1975) as well as α, β and ρ. The computation is reduced by calculating only arrivals within a specified phase-velocity window, and by including only the multiples and interconversions for the 'reflection zone', usually the lower part of the model. The layers above the reflection zone, which need not be the same beneath source and receiver, are assumed to introduce only transmission losses; no reflections are calculated for these upper layers. The source and receiver responses are included in the frequency domain by multiplication, and Fourier transformation then gives the synthetic seismogram as a time series. The disadvantage of this method is the amount of computing time

required and the fact that lateral variations cannot be included in the reflection zone.

4.4 Reflection seismology

4.4.1 The reflection method

Although earthquake seismology and refraction seismology enable scientists to determine gross Earth structures and crustal and upper-mantle structures, reflection seismology is the method used to determine fine details of the shallow structures, usually over small areas. The resolution obtainable with reflection seismology makes it the main method used by oil-exploration companies to map subsurface sedimentary structures. The method has also increasingly been used to obtain new information on the fine structures within the crust and at the crust–mantle boundary.

For land profiles, explosives can be used as a source. Other sources include the *gas exploder*, in which a gas mixture is exploded in a chamber that has a movable bottom plate resting on the ground, and the *vibrator*, in which a steel plate pressed against the ground is vibrated at increasing frequency (in the range 5–60 Hz) for several seconds (up to 30 s for deep crustal reflection profiling). Vibrators require an additional step in the data processing to extract the reflections from the recordings: the cross-correlation of the recordings with the source signal.

Of the many marine sources, the two most frequently used for deep reflection profiling are the *air gun*, in which a bubble of very-high-pressure air is released into the water, and the *explosive cord*. Many air guns are usually used in an array towed behind the shooting ship.

Deconvolution is the process which removes the effects of the source and receiver from the recorded seismograms and allows direct comparison of data recorded with different sources and/or receivers. For the details of the methods of obtaining and correcting seismic-reflection profiles, the reader is again referred to the textbooks on exploration geophysics (e.g., Telford *et al*. 1990; Dobrin and Savit 1988; Yilmaz 2001; Claerbout 1985).

The basic assumption of seismic reflection is that there is a stack of horizontal layers in the crust and mantle, each with a distinct seismic P-wave velocity. Dipping layers, faults and so forth can be included in the method (see Section 4.4.4). P-waves from a surface energy source, which are almost normally incident on the interfaces between these layers, are reflected and can be recorded by *geophones* (vertical-component seismometers) close to the source. Because the rays are close to normal incidence, effectively no S-waves are generated (Fig. 4.39). The P-waves reflected at almost normal incidence are very much smaller in amplitude than the wide-angle reflections near to, and beyond, the critical distance. This fact means that normal-incidence reflections are less easy to recognize than wide-angle reflections and more likely to be obscured by background

noise, and that sophisticated averaging and enhancement techniques must be used to detect reflecting horizons.

4.4.2 A two-layer model

Consider the two-layer model in Fig. 4.34. By application of Pythagoras' theorem, the travel time t for the reflection path SCR is given by Eqs. (4.29) and (4.30) as

$$t = \frac{\mathrm{SC}}{\alpha_1} + \frac{\mathrm{CR}}{\alpha_1}$$
$$= \frac{2}{\alpha_1}\sqrt{z_1^2 + \frac{x^2}{4}}$$

or

$$t^2 = \frac{4z_1^2}{\alpha_1^2} + \frac{x^2}{\alpha_1^2} \tag{4.65}$$

which is the equation of a hyperbola. At normal incidence ($x = 0$), the travel time is $t = t_0$, where

$$t_0 = \frac{2z_1}{\alpha_1} \tag{4.66}$$

This is the two-way normal-incidence time. At large distances ($x \gg z_1$) the travel time (Eq. (4.65)) can be approximated by

$$t \approx \frac{x}{\alpha_1} \tag{4.67}$$

This means that, at large distances, the travel-time curve is asymptotic to the travel time for the direct wave, as illustrated in Fig. 4.32(b). In reflection profiling, since we are dealing with distances much shorter than the critical distance, the travel-time–distance plot is still curved. Notice that, with increasing values of α_1, the hyperbola (Eq. (4.65)) becomes flatter. If travel-time–distance data were obtained from a reflection profile shot over such a model, one way to determine α_1 and z_1 would be to plot not t against x, but t^2 against x^2. Equation (4.65) is then the equation of a straight line with slope $1/\alpha_1^2$ and an intercept on the t^2 axis of $t_0^2 = 4z_1^2/\alpha_1^2$.

The normal-incidence reflection coefficient for P-waves is given by Eq. (4.62). Since normal-incidence reflections have small amplitudes, it is advantageous to average the signals from nearby receivers to enhance the reflections and reduce the background noise. This averaging process is called *stacking*. *Common-depth-point* (CDP) *stacking*, which combines all the recordings of reflections from each subsurface point, is the method usually used. Common-offset stacking, which combines all the recordings with a common offset distance, is less popular. Figure 4.40 shows the layout of shots (or vibrators) and receivers used in CDP reflection profiling. The coverage obtained by any profile is

$$\text{coverage} = \frac{\text{number of receivers}}{\text{twice the shot spacing}} \tag{4.68}$$

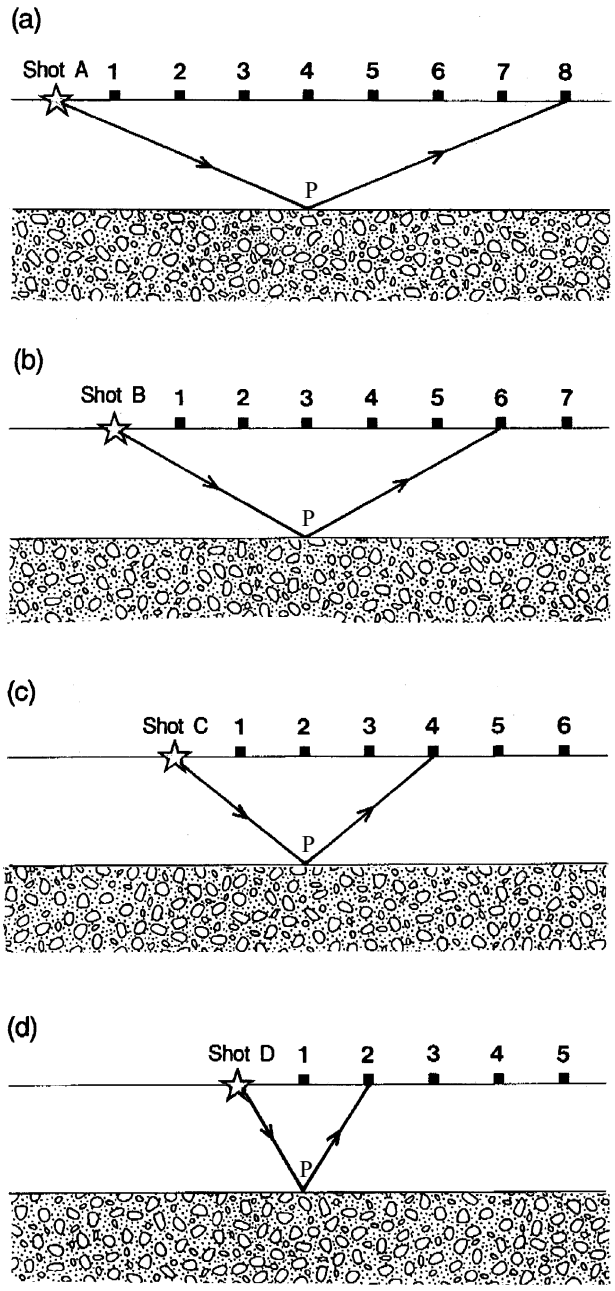


Figure 4.40. Common-depth-point (CDP) reflection profiling. In this example, eight geophones (■) record each shot (★). In (a), shot A is fired and a reflection from a particular point P on the reflector (the interface between the two layers) is recorded by geophone 8. In (b), all the geophones and the shotpoint have been moved one step to the right, and shot B is fired; the reflection from point P is recorded by geophone 6. Similarly, a reflection from P is recorded (c) by geophone 4 when shot C is fired and (d) by geophone 2 when shot D is fired. The four reflections from point P can be stacked (added together after time corrections have been made). In this example, because there are four reflections from each reflecting point on the interface (fewer at the two ends of the profile), there is said to be four-fold coverage. Alternatively, the reflection profile can be described as a four-fold CDP profile.

where the shot spacing is in units of receiver spacing. In the example of Fig. 4.40, the number of receivers is eight and the shot spacing is one. This results in a four-fold coverage. Reflection-profiling systems usually have 48 or 96 recording channels (and hence receivers), which means that 24-, 48-, or 96-fold coverage is possible. The greater the multiplicity of coverage, the better the system is for imaging weak and deep reflectors and the better the final quality of the record section. In practice, receiver spacing of a few tens to hundreds of metres is used, in contrast to the kilometre spacing of refraction surveys.

In order to be able to add all these recordings together to produce a signal reflected from the common depth point, one must first correct them for their different travel times, which are due to their different offset distances. This correction to the travel times is called the *normal-moveout* (NMO) *correction*.

The travel time for the reflected ray in the simple two-layer model of Fig. 4.34 is given by Eq. (4.65). The difference between values of the travel time t at two distances is called the *moveout*, Δt. The moveout can be written

$$\Delta t = \frac{2}{\alpha_1}\sqrt{z_1^2 + \frac{x_a^2}{4}} - \frac{2}{\alpha_1}\sqrt{z_1^2 + \frac{x_b^2}{4}} \tag{4.69}$$

where x_a and x_b ($x_a > x_b$) are the distances of the two geophones a and b from the shotpoint. The normal moveout Δt_{NMO} is the moveout for the special case when geophone b is at the shotpoint (i.e., $x_b = 0$). In this case, and dropping the subscript a, Eq. (4.69) becomes

$$\Delta t_{NMO} = \frac{2}{\alpha_1}\sqrt{z_1^2 + \frac{x^2}{4}} - \frac{2z_1}{\alpha_1} \tag{4.70}$$

If we make the assumption that $2z_1 \gg x$, which is generally appropriate for reflection profiling, we can use a binomial expansion for $\sqrt{z_1^2 + x^2/4}$:

$$\begin{aligned}\sqrt{z_1^2 + \frac{x^2}{4}} &= z_1\sqrt{1 + \frac{x^2}{4z_1^2}} \\ &= z_1\left[1 + \left(\frac{x}{2z_1}\right)^2\right]^{1/2} \\ &= z_1\left[1 + \frac{1}{2}\left(\frac{x}{2z_1}\right)^2 - \frac{1}{8}\left(\frac{x}{2z_1}\right)^4 + \frac{1}{16}\left(\frac{x}{2z_1}\right)^6 + \cdots\right]\end{aligned} \tag{4.71}$$

To a first approximation, therefore,

$$\sqrt{z_1^2 + \frac{x^2}{4}} = z_1\left[1 + \frac{1}{2}\left(\frac{x}{2z_1}\right)^2\right] \tag{4.72}$$

Substituting this value into Eq. (4.70) gives a first approximation for the normal moveout Δt_{NMO}.

$$\begin{aligned}\Delta t_{NMO} &= \frac{2z_1}{\alpha_1}\left[1 + \frac{1}{2}\left(\frac{x}{2z_1}\right)^2\right] - \frac{2z_1}{\alpha_1} \\ &= \frac{x^2}{4\alpha_1 z_1}\end{aligned} \tag{4.73}$$

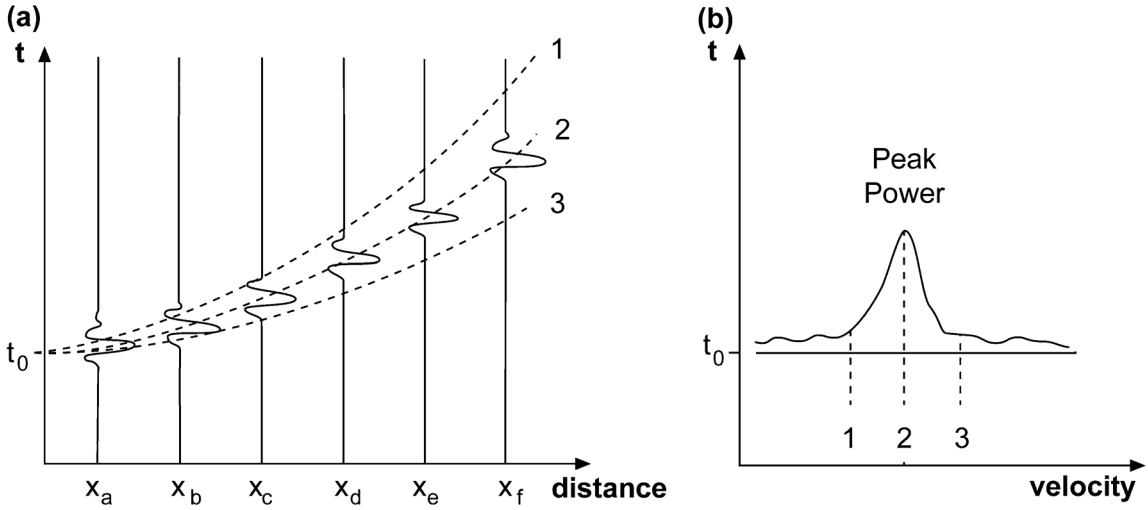


Figure 4.41. (a) Reflections from an interface, recorded at distances x_a, x_b, x_c, x_d, x_e and x_f. Three travel-time curves (1, 2 and 3) are shown for two-way normal-incidence time t_0 and increasing values of velocity. Clearly, curve 2 is the best fit to the reflections. To stack these traces, the NMO correction (Eq. (4.74)) for curve 2 is subtracted from each trace so that the reflections line up with a constant arrival time of t_0. Then the traces can be added to yield a final trace with increased signal-to-noise ratio. (b) The power in the stacked signal is calculated for each value of the stacking velocity and displayed on a time–velocity plot. The velocity which gives the peak value for the power in the stacked signal is then the best stacking velocity for that particular value of t_0. For (a), velocity 2 is best; velocity 1 is too low and velocity 3 too high. (After Taner and Koehler (1969).)

Using Eq. (4.66) for t_0, the two-way normal-incidence time, we obtain an alternative expression for Δt_{NMO}:

$$\Delta t_{\mathrm{NMO}} = \frac{x^2}{2\alpha_1^2 t_0} \tag{4.74}$$

This illustrates again the fact that the reflection time–distance curve is flatter (Δt_{NMO} is smaller) for large velocities and large normal-incidence times. This NMO correction must be subtracted from the travel times for the common-depth-point recordings. The effect of this correction is to line up all the reflections from each point P with the same arrival time t_0 so that they can be stacked (added together) to produce one trace. This procedure works well when we are using a model for which α_1 and z_1 are known, but in practice we do not know them: they are precisely the unknowns which we would like to determine from the reflections! This difficulty is overcome by the bootstrap technique illustrated in Fig. 4.41. A set of arrivals is identified as reflections from point P if their travel times fall on a hyperbola. Successive values of α_1 and t_0 are tried until a combination defining a hyperbola that gives a good fit to the travel times is found. These values α_1 and

t_0 then define the model. In the example of Fig. 4.41, it is obvious that curve 2 is correct, but with real data it is not sufficient to rely on the eye alone to determine velocities – a numerical criterion must be used. In practice, the reflections are stacked using a range of values for α_1 and t_0. The power (or some similar entity) in the stacked signal is calculated for each value of α_1 and t_0. For each value of t_0 the maximum value of the power is then used to determine the best velocity. A plot of power against both velocity and time, as in Fig. 4.41(b), is usually called a *velocity-spectrum* display.

Once the records have been stacked, the common-depth-point record section shows the travel times as if shots and receivers were coincident. The stacking process necessarily involves some averaging over fairly short horizontal distances, but it has the considerable advantage that the signal-to-noise ratio of the stacked traces is increased by a factor of $\sqrt{n}$ over the signal-to-noise ratio of the n individual traces.

4.4.3 A multilayered model

A two-layer model is obviously not a realistic approximation to a pile of sediments or to the Earth's crust, but it serves to illustrate the principle of the reflection method. A multiple stack of layers is a much better model than a single layer. Travel times through a stack of multiple layers are calculated in the same way as for two layers, with the additional constraint that Snell's law ($\sin i/\alpha = p$, a constant for each ray) must be applied at each interface (Eq. (4.55)). The travel times and distances for a model with n layers, each with thickness z_j and velocity α_j, are best expressed in the parametric form

$$\begin{aligned} x &= 2\sum_{j=1}^{n} \frac{z_j p \alpha_j}{\sqrt{1-\rho^2\alpha_j^2}} \\ t &= 2\sum_{j=1}^{n} \frac{z_j}{\alpha_j\sqrt{1-\rho^2\alpha_j^2}} \end{aligned} \tag{4.75}$$

where $p = \sin i_j/\alpha_{\mathrm{j}}$. It is unfortunately not usually possible to eliminate p from these equations in order to express the distance curve as one equation. In this multilayered case, the time–distance curve is *not* a hyperbola as it is when $n = 1$ (Eq. (4.65)). However, it can be shown that the square of the travel time, t^2, can be expressed as an infinite series in x^2:

$$t^2 = c_0 + c_1x^2 + c_2x^4 + c_3x^6 + \cdots \tag{4.76}$$

where the coefficients c_0, c_1, c_2, . . . are constants dependent on the layer thicknesses z_j and velocities α_j. In practice, it has been shown that use of just the first two terms of Eq. (4.76) (c_0 and c_1) gives travel times to an accuracy of within about 2%, which is good enough for most seismic-reflection work. This means that Eq. (4.76) can be simplified to

$$t^2 = c_0 + c_1x^2 \tag{4.77}$$

which is, after all, like Eq. (4.65), a hyperbola. The value of the constant c_0 is given by

$$c_0 = \left\{ \sum_{m=1}^{n} t_m \right\}^2 \tag{4.78}$$

where $t_m = 2z_m/\alpha_m$ is the two-way vertical travel time for a ray in the mth layer. The two-way normal-incidence travel time from the nth interface, $t_{0,n}$, is the sum of all the t_m:

$$t_{0,n} = \sum_{m=1}^{n} t_m = \sum_{m=1}^{n} \frac{2z_m}{\alpha_m} \tag{4.79}$$

Equation (4.78) can therefore be more simply written as

$$c_0 = (t_{0,n})^2 \tag{4.80}$$

The second constant of Eq. (4.77), c_1, is given by

$$c_1 = \frac{\sum_{m=1}^{n} t_m}{\sum_{m=1}^{n} \alpha_m^2 t_m} \tag{4.81}$$

We can define a time-weighted root-mean-square (RMS) velocity $\overline{\alpha}_n^2$ as

$$\overline{\alpha}_n^2 = \frac{\sum_{m=1}^{n} \alpha_m^2 t_m}{\sum_{m=1}^{n} t_m} \tag{4.82}$$

With these expressions for c_0 and c_1, Eq. (4.77) becomes

$$t^2 = (t_{0,n})^2 + \frac{x^2}{\overline{\alpha}_n^2} \tag{4.83}$$

This now has exactly the same form as the equation for the two-layer case (Eq. (4.65)), but instead of t_0 we have $t_{0,n}$, and instead of the constant velocity above the reflecting interface α_1, we now have $\overline{\alpha}_n$, the time-weighted RMS velocity above the nth interface. This means that NMO corrections can be calculated and traces stacked as described in Section 4.4.2. The only difference is that, in this multilayered case, the velocities determined are not the velocity above the interface but the RMS velocity above the interface.

Figure 4.42 shows a typical velocity-spectrum display for real data. By stacking reflection records and using such a velocity-spectrum display, one can estimate both $t_{0,n}$ and $\overline{\alpha}_n$ for each reflector. However, determining $t_{0,n}$ and $\overline{\alpha}_n$ is not enough; to relate these values to the rock structure over which the reflection line was shot, we have to be able to calculate the thicknesses and seismic velocities for each layer and to obtain an estimate of the accuracy of such values.

Let us suppose that the RMS velocity and normal-incidence times have been determined for each of two successive parallel interfaces (i.e., $t_{0,n-1}$, $t_{0,n}$, $\overline{\alpha}_{n-1}$

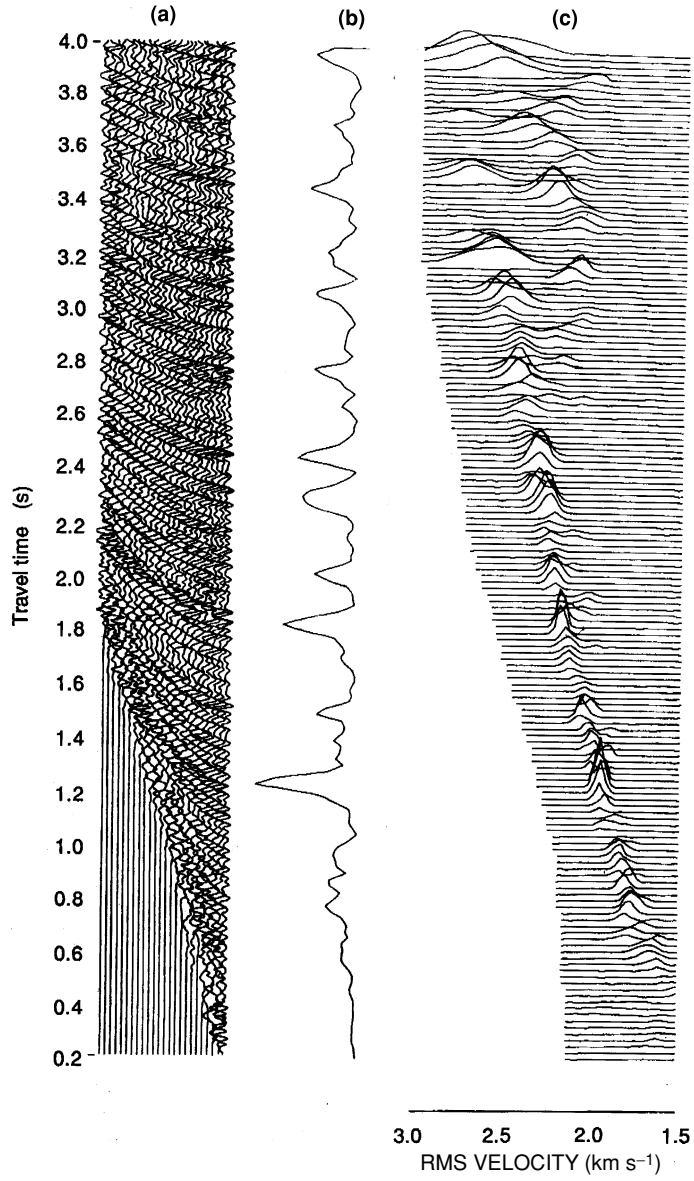


Figure 4.42. Steps in the computation of a velocity analysis. (a) The twenty-four individual reflection records used in the velocity analysis. (b) The maximum amplitude of the stacked trace shown as a function of t_0, two-way time along the trace. Notice that the reflections are enhanced compared with the original traces. Although the main reflections at $t_0 = 1.2$ and 1.8 s do stand out on the original traces, these intercept times are clearly defined by the stacked trace, and subsequent deeper reflections that were not clear on the original traces can now be identified with some confidence on the stacked trace. (c) The velocity spectrum for the traces in (a). The peak power at each time served to identify the velocity which would best stack the data. The stacking velocity clearly increases steadily with depth down to about 3 s. After this, some strong multiples (rays that have bounced twice or more in the upper layers and therefore need a smaller stacking velocity) confuse the velocity display. (From Taner and Koehler (1969).)

and $\overline{\alpha}_n$ are known). Then, by using Eqs. (4.79) and (4.82), we can determine the velocity of the nth layer α_n:

$$\begin{aligned}\overline{\alpha}_n^2 \sum_{m=1}^{n} t_m &= \sum_{m=1}^{n} \alpha_m^2 t_m \\ &= \sum_{m=1}^{n-1} \alpha_m^2 t_m + \alpha_n^2 t_n \end{aligned} \tag{4.84}$$

and

$$\overline{\alpha}_{n-1}^2 \sum_{m=1}^{n-1} t_m = \sum_{m=1}^{n-1} \alpha_m^2 t_m \tag{4.85}$$

Subtracting Eq. (4.85) from Eq. (4.84) gives

$$\overline{\alpha}_n^2 \sum_{m=1}^{n} t_m - \overline{\alpha}_{n-1}^2 \sum_{m=1}^{n-1} t_m = \alpha_n^2 t_m \tag{4.86}$$

or

$$\overline{\alpha}_n^2 t_{0,n} - \overline{\alpha}_{n-1}^2 t_{0,n-1} = \alpha_n^2 (t_{0,n} - t_{0,n-1}) \tag{4.87}$$

Rearranging this equation gives α_n, the velocity of the nth layer (also known as the *interval velocity*), in terms of the RMS velocities:

$$\alpha_n = \sqrt{\frac{\overline{\alpha}_n^2 t_{0,n} - \overline{\alpha}_{n-1}^2 t_{0,n-1}}{t_{0,n} - t_{0,n-1}}} \tag{4.88}$$

After α_n has been determined, z_n can be calculated from Eq. (4.79):

$$\begin{aligned} t_{0,n} &= \sum_{m=1}^{n} \frac{2z_m}{\alpha_m} \\ &= \sum_{m=1}^{n-1} \frac{2z_m}{\alpha_m} + \frac{2z_n}{\alpha_n} \\ &= t_{0,n-1} + \frac{2z_n}{\alpha_n} \end{aligned} \tag{4.89}$$

Thus,

$$z_n = \frac{\alpha_n}{2}(t_{0,n} - t_{0,n-1}) \tag{4.90}$$

Therefore, given the two-way normal-intercept times and corresponding stacking (RMS) velocities from a velocity analysis, the velocity–depth model can be determined layer by layer, starting at the top and working downwards.

Multiples are rays that have been reflected more than once at an interface. The most common multiple is the surface multiple, which corresponds to a ray that travels down and up through the layers twice. Reflections with multiple ray paths in one or more layers also occur. In marine work, the multiple which is reflected at the sea surface and seabed is very strong (Figs. 9.23 and 10.10). The periodicity of multiple reflections enables us to filter them out of the recorded data by deconvolution.

Example: calculation of layer thickness and seismic velocity from a normal-incidence reflection line

A velocity analysis of reflection data has

$$t_{0,1} = 1.0\,s, \qquad \overline{\alpha}_1 = 3.6 \text{ km s}^{-1}$$
$$t_{0,2} = 1.5\,s, \qquad \overline{\alpha}_2 = 4.0 \text{ km s}^{-1}$$

Calculate a velocity–depth model from these values, assuming that there is a constant velocity in each layer. Since $\overline{\alpha}_1$ must be equal to α_1, the velocity of the top layer, z_1, can be calculated from Eq. (4.90):

$$z_1 = \frac{3.6 \times 1.0}{2.0} = 1.8 \text{ km}$$

Now α_2 can be calculated from Eq. 4.88:

$$\alpha_2 = \sqrt{\frac{(4.0^2 \times 1.5) - (3.6^2 \times 1.0)}{1.5 - 1.0}}$$
$$= \sqrt{22.08} = 4.7 \text{ km s}^{-1}$$

Finally, z_2 is then calculated from Eq. (4.90):

$$z_2 = \frac{4.7 \times 0.5}{2.0} = 1.175 \text{ km}$$

Notice that the velocity α_2 is larger than the RMS velocity $\overline{\alpha}_2$ in this example.

Unfortunately, the interval velocities and depths frequently cannot be determined accurately by these methods (see Problem 21). In exploration work the inaccuracy of velocity–depth information can usually be made up for by detailed measurements from drill holes. However, when reflection profiling is used to investigate structures deep in the crust and upper mantle, such direct velocity information is not available. There is an added problem in this situation: when the depths of horizons are much larger than the maximum offset, the NMO correction is insensitive to velocity. This means that any velocity that is approximately correct will stack the reflections adequately, and so the interval velocities calculated from the stacking velocities will not be very accurate. As an example, consider a two-layer model:

$$\alpha_1 = 6 \text{ km s}^{-1}, \quad z_1 = 20 \text{ km}, \quad \alpha_2 = 8 \text{ km s}^{-1}$$

The reflection hyperbola (Eq. (4.65)) for this interface is

$$t = \frac{1}{3}\sqrt{400 + \frac{x^2}{4}}$$

and the two-way normal incidence time t_0 is 6.667 s. The NMO correction is (Eq. (4.73))

$$\Delta t_{\text{NMO}} = \frac{x^2}{480}$$

The maximum offset used in deep reflection profiling is rarely greater than about 5 km and often less. At this offset, the correct NMO correction for the reflection is 0.052 s. However, this correction has to be estimated from the data. Using the

correct value for t_0, $\Delta t_{\mathrm{NMO}} = 0.062$ s when $\alpha_1 = 5.5\,\mathrm{km\,s^{-1}}$ and $\Delta t_{\mathrm{NMO}} = 0.044$ s when $\alpha_1 = 6.5\,\mathrm{km\,s^{-1}}$. Clearly, a signal with predominant frequency of 20 Hz or less (i.e., time for one cycle >0.050 s) cannot give an accurate value for α_1. Higher-frequency signals give more accurate velocity analyses, as discussed in Section 4.4.4. The best way to obtain reliable interval-velocity measurements in such cases is to supplement the reflection profiles with wide-angle-reflection profiles. Such profiles, perhaps 50–80 km in length, allow deep wide-angle reflections to be recorded. Wide-angle reflections have larger amplitudes than those of normal-incidence reflections (Fig. 4.39), and their travel times (Eqs. (4.65) and (4.67)) give more reliable values for the interval velocities.

Example: amplitude of reflections

For normal-incidence reflections the P-wave reflection coefficient is given by Eq. (4.62). An increase in impedance will therefore give a positive reflection coefficient and a decrease in impedance will give a negative reflection coefficient. As a simple example, imagine a layer of sediment with an impedance of 4×10^6 kg $\mathrm{m^{-2}\,s^{-1}}$ sandwiched in sediment with an impedance of 3×10^6 kg $\mathrm{m^{-2}\,s^{-1}}$. The reflection coefficient at the top of the layer is

$$\frac{4-3}{4+3} = \frac{1}{7} = 0.143$$

while the reflection coefficient at the base of the layer is

$$\frac{3-4}{3+4} = \frac{-1}{7} = -0.143$$

To see a real example of the use of reflection coefficients, consider the *bottom-simulating reflector* (BSR) which is a common feature of marine seismic-reflection lines. The BSR arrives about 200–300 ms after the seafloor reflection, has the opposite polarity, follows the seabed reflection (hence its name) and so frequently cuts across reflections from any sedimentary stratigraphy. The BSR is a consequence of the presence of gas hydrates within the sediments. A gas hydrate is a rigid water-molecule cage that encloses and is stabilized by methane or other hydrocarbons. Hydrates are stable at temperatures over 0 °C at the elevated pressures reached in water over 300 m deep – effectively an 'ice' that is stable above 0 °C. These hydrates have a narrow stability field, the base (controlled by temperature) being only a few hundred metres below the seafloor. The base of the stability field marks the boundary between high-velocity hydrated sediments above and normal or gas-filled sediments below. Since there is a change in physical properties of the sediment at this boundary, it will yield a seismic reflection. Figure 4.43 shows a seismic-reflection line across the Cascadia margin where the Juan de Fuca plate is subducting beneath North America. The BSR is clearly visible, as is the deformation of the sediments filling this oceanic trench (see also Section 9.2.2). Figure 4.44 shows the detail of the seafloor and BSR reflections – the

BSR clearly has the opposite polarity to the seafloor reflection. For normal-incidence reflections the P-wave reflection coefficient is given by Eq. (4.62). There is an abrupt increase in impedance at the seafloor due primarily to the increase in density on going from sea water to the very porous sediment; this gives rise to a positive reflection. At the base of the hydrate stability field the impedance falls abruptly. This fall in impedance, due to the decrease in seismic velocity and partly density, gives rise to a negative-polarity reflection. Synthetic seismograms calculated from the impedance model match the reflection data well.

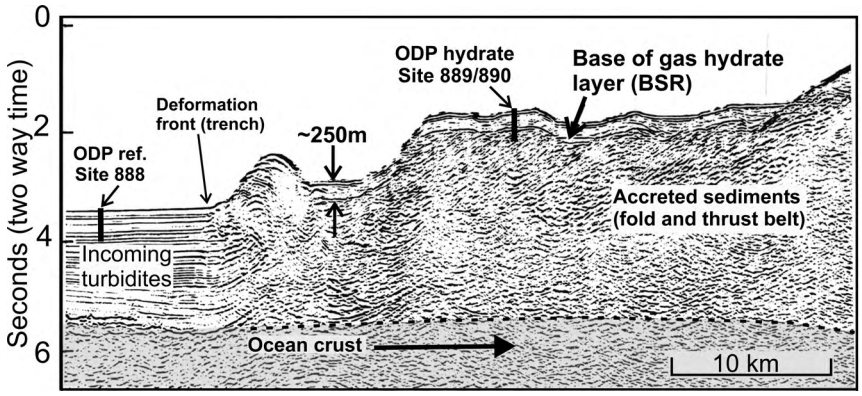


Figure 4.43. The bottom-simulating reflector (BSR) on a seismicreflection line on the continental slope west of Vancouver Island, Canada, part of the Cascadia margin where the Juan de Fuca plate is subducting beneath the North American plate. (From Hyndman *et al.* Geophysical studies of marine gas hydrates in northern Cascadia. Geophysical Monograph 124, 273–95, 2002. Copyright 2002 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

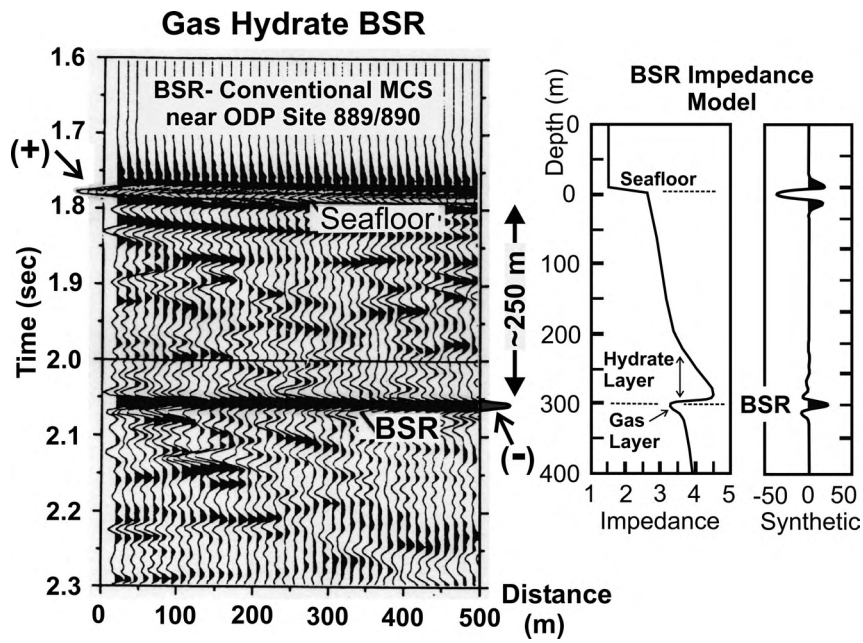


Figure 4.44. Detail of the seafloor and BSR (Bottom Simulating Reflector) reflections from Fig. 4.43. The seafloor reflection has positive polarity (+); the BSR has negative polarity (−). The impedance model (Eq. (4.87)) and synthetic reflections calculated from the impedance model match the data well. (From Hyndman *et al.* Geophysical studies of marine gas hydrates in northern Cascadia. Geophysical Monograph 124, 273–95, 2002. Copyright 2002 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

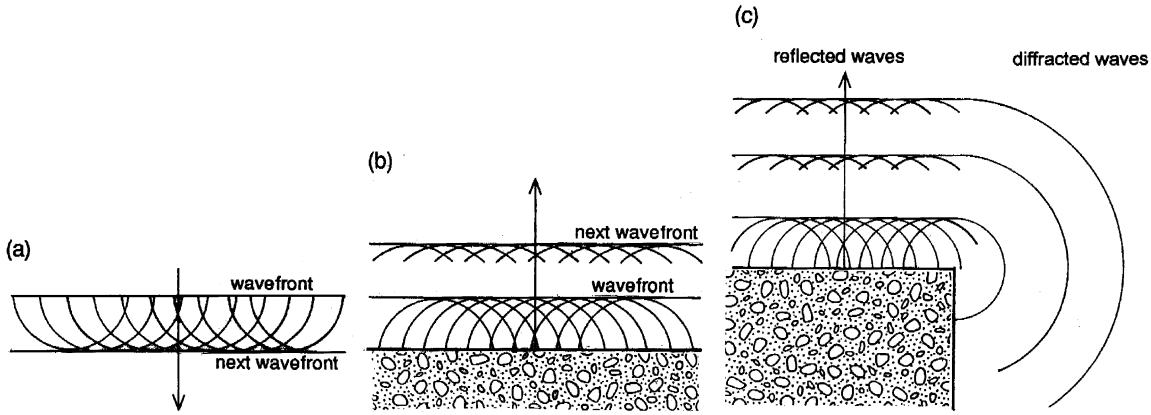


Figure 4.45. (a) A plane wave propagating vertically downwards. Each point on the wavefront acts as a point source and generates spherical waves. The next wavefront, which is the sum of all these spherical waves, is also a plane wave. All the lateral disturbances cancel each other out. (b) A plane wave travelling vertically downwards is reflected at an interface (lower material is shaded). Each point on the interface acts as a point source and generates spherical waves. Thus, the reflected wave is a plane wave travelling upwards (all lateral disturbances again cancel out). (c) A plane wave reflected from an interface that terminates at a corner, P. The reflected wave well to the left of P is unaffected. Since there are no point sources to the right of P, the reflected wave close to P is reduced in amplitude. In addition, the spherical wave generated by P is not cancelled out. It is called the diffracted wave.

4.4.4 Diffraction and migration

All the preceding sections have dealt with a horizontally layered model. Although some structures can be approximated by horizontal layers, structures such as faults, dykes and intrusions cannot. It is also necessary to allow for the effects of dipping layers. These complications bring into question the minimum size or contrast in physical properties which can be resolved by using the reflection-profiling method.

Diffraction

Huygens' principle states that every point on a wavefront acts as a point source generating spherical waves. The wavefront at a later time is then the sum of all the waves from these point sources. The case of a plane wave is illustrated in Fig. 4.45(a): the new wavefront is also planar because the lateral disturbances from the point sources cancel each other out. When a plane wave is reflected from an interface, the same principle applies: each point on the interface acts as a point source and generates spherical waves. When summed together, these spherical waves give a reflected plane wave (Fig. 4.45(b)). When an interface is not infinitely long but terminates at a point P (Fig. 4.45(c)), the reflected wave far to the left of P is not affected. However, since there are no more point sources

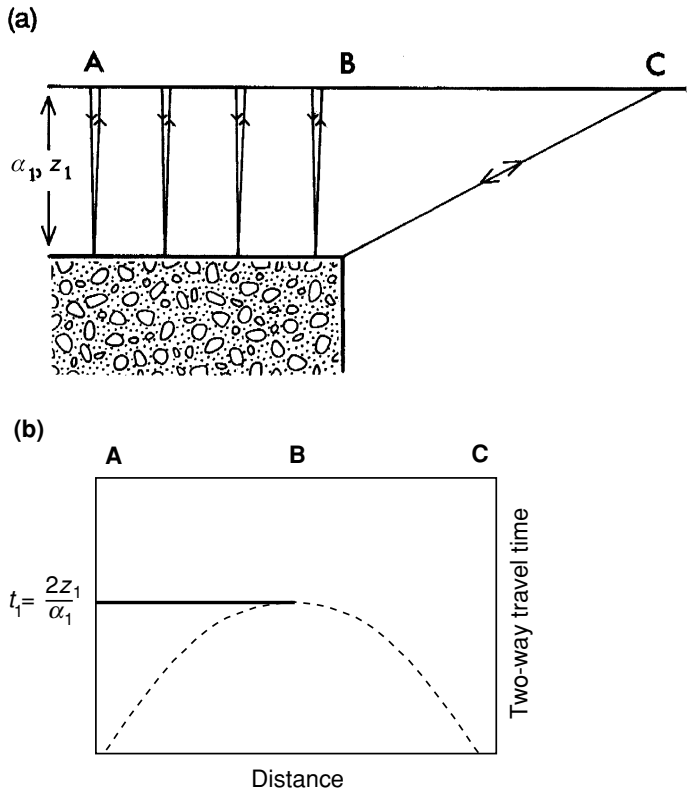


Figure 4.46. (a) A reflection profile over a horizontal reflector that terminates at point P. The overlying material has a constant P-wave velocity of α_1, and the reflector is at a depth z_1. Normal-incidence reflections are recorded from A to B, but no reflections can be recorded to the right of location B. A diffracted wave produced by the point P is recorded at all locations. The ray path for the diffracted wave arriving at location C is the straight line PC. The amplitude of diffracted waves is smaller than that of the reflections (amplitude being distance-dependent). (b) The reflection record section for the model in (a). The reflection is shown as a solid line and the diffraction as a dashed line.

to the right of point P to ensure that all but the plane wave cancels out, the wave field in the region around point P is affected. The spherical waves generated by point P which are not cancelled out combine to form the *diffracted* wave. This wave is detected both to the right and to the left of point P. Huygens' principle also applies to spherical waves: a spherical wave incident on the interface which terminates at point P also gives rise to a diffraction.

Diffractions are not restricted to the ends of reflectors; a diffracted wave is produced by any irregularity in a structure that is comparable in scale to the wavelength of the signal. (Recall the diffraction of light through a hole, the diffraction of water waves around the end of a breakwater and diffraction at the core–mantle boundary (Section 8.1.1).)

Consider the horizontal reflector which terminates abruptly at point P (Fig. 4.46(a)). Normal-incidence reflections are recorded from this reflector only at locations between A and B. However, the corner P acts as a point source when a seismic wave hits it and so radiates seismic waves in all directions in accordance with Huygens' principle. Very close to B, this diffracted wave arrives at the same time as the reflection, but with increasing distance from B (left or right) it arrives progressively later.

The travel time t for the diffracted wave recorded at location C, a horizontal distance x from B, is given by

$$t = \frac{2}{\alpha_1}\sqrt{x^2 + z_1^2}$$

or

$$\alpha_1^2 t^2 = 4\left(x^2 + z_1^2\right) \tag{4.91}$$

This is the equation of a hyperbola (with a greater curvature than the reflection hyperbola of Eq. (4.65)). The reflection and diffraction travel times are plotted on the record section in Fig. 4.46(b). The curvature of this diffracted arrival on the record section decreases with the depth z_1 of the diffracting point P; diffractions from a deep fault are flatter than those from a shallow fault.

Immediately to the left of point B, the amplitude of the reflection decreases as B is approached until at B the amplitude of the reflection is only half its value well away from the edge of the reflector. (See the discussion of the Fresnel zone later in this section.) The diffraction amplitude decreases smoothly with increasing distance from the edge of the reflector. Since the reflection is tangential to the diffraction hyperbola at point B, there is no discontinuity either in travel time or in amplitude to mark the edge of the reflector. However, the diffraction branch recorded to the left of B has opposite polarity to the branch recorded to the right of B (they are 180° out of phase). Detection of both diffraction branches can assist in the location of the edge of the reflector.

Dipping layers

On a CDP record section, the reflection point is shown as being vertically below the shot/receiver location. Dipping interfaces are therefore distorted on CDP record sections because the reflection points are not vertically below the shot/receiver location. The situation is illustrated in Fig. 4.47: the normal-incidence reflection recorded at location A is reflected from point A′ and that recorded at location B is reflected from point B′. If z_1 is the vertical depth of the reflector beneath A and δ is the dip of the reflector, then the length AA′ is given by

$$\mathrm{AA}' = z_1 \cos\delta \tag{4.92}$$

and the two-way travel time for this reflection is

$$t = \frac{2z_1}{\alpha_1}\cos\delta \tag{4.93}$$

Similarly, the length BB′ is given by

$$\mathrm{BB}' = z\cos\delta \tag{4.94}$$

and the two-way travel for this reflection is

$$t = \frac{2z}{\alpha_1}\cos\delta \tag{4.95}$$

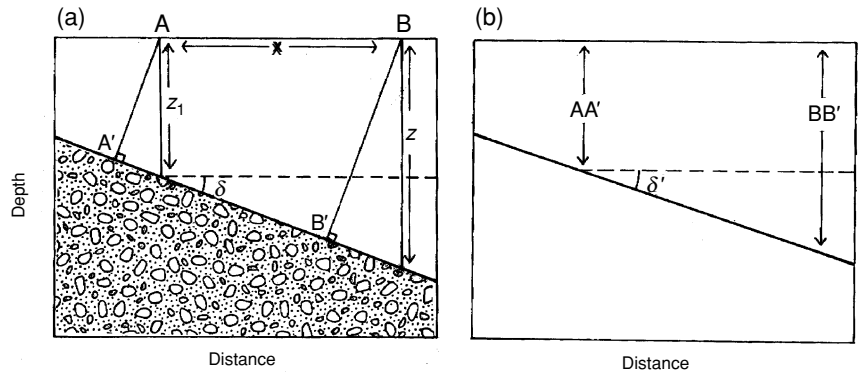


Figure 4.47. (a) Normal-incidence reflections from an interface with dip δ. With the source and receiver at A, the reflection point is A′ (likewise B and B′). The upper layer has a P-wave velocity of α_1. (b) The record section for the structure in (a). On a section with equal horizontal and vertical distance scales, the apparent dip δ' is less than the real dip δ (Eq. (4.99)).

where z is the vertical depth of the reflector below B. The depths z_1 and z, dip angle δ and horizontal distance x between A and B are related by

$$\tan\delta = \frac{z - z_1}{x} \tag{4.96}$$

Since the reflections from A′ and B′ plot on the record section as though they were vertically beneath A and B, respectively, the reflector appears to have a dip of δ', where

$$\tan\delta' = \frac{\mathrm{BB}' - \mathrm{AA}'}{x} \tag{4.97}$$

On substituting from Eqs. (4.92) and (4.94), this becomes

$$\tan\delta' = \frac{z\cos\delta - z_1\cos\delta}{x} \tag{4.98}$$

Further substitution from Eq. (4.96) yields

$$\tan\delta' = \sin\delta \tag{4.99}$$

Thus, the apparent dip δ' is less than the actual dip δ, except for very small angles, for which the approximation $\sin\delta \approx \tan\delta$ is valid. Reflection record sections are rarely plotted at true scale: distortion is also caused by differing horizontal and vertical scales as well as by the vertical axis frequently being two-way travel time.

Migration

A dipping reflector is not the only structure which is incorrectly located on a CDP record section. A reflection from any variation in structure in the third dimension, conveniently assumed to be nonexistent here, could also appear on the section.

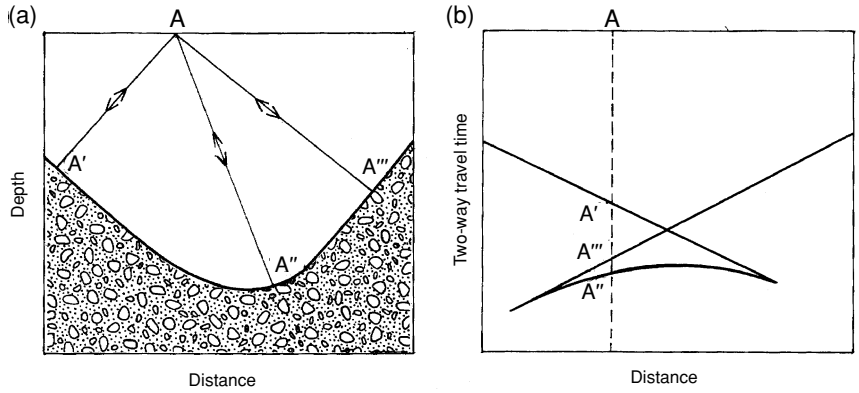


Figure 4.48. (a) A syncline with a centre of curvature (focus) that is below the Earth's surface gives three normal-incidence reflections for shot/receiver locations over the central part of the syncline. A shot/receiver location at A has reflections from points A′, A′ and A‴, one from each flank of the syncline and the third from the centre. The extreme example of a syncline is two planes intersecting – there is diffraction from the intersection. (b) Travel times for the syncline shown in (a). The deeper the centre of curvature (the tighter the syncline), the broader the convex reflection from the central portion of the syncline. Thus, on a record section, a syncline can look much like an anticline. The process of migration uses these three reflection branches to determine the shape of the reflecting surface, the syncline.

Diffractions can be mistaken for reflections. Curved interfaces can also produce some rather strange reflections (Fig. 4.48).

Migration is the name given to the process which attempts to deal with all of these problems and to move the reflectors into their correct spatial locations. Thus, diffractions are removed. They are recognized as being diffractions by a method similar to the method of estimating the best stacking velocity; that is, if they lie on a hyperbola, they are diffractions and are coalesced into a point at its apex. Dipping layers are plotted correctly, multiples (reflections bouncing twice or more in a layer) are removed and the reflections from curved interfaces are inverted. There are many methods by which migration is performed, all of them complex; the interested reader is referred to a reflection-processing text such as Robinson and Treitel (1980), Hatton *et al*. (1985) or Yilmaz (2001) for more details.

Resolution of structure

The resolution of structure obtainable with reflection profiling is controlled by the wavelength of the source signal. A wave of frequency 20 Hz travelling through material with a P-wave velocity of 4.8 km s^{-1} has a wavelength of 240 m. When

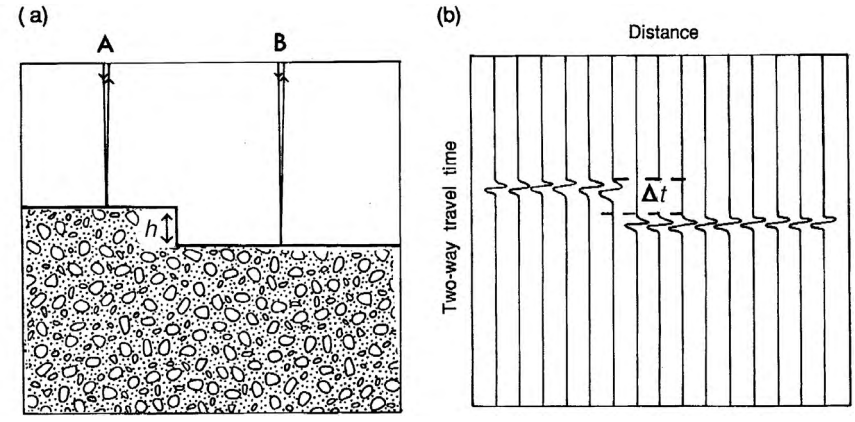


Figure 4.49. (a) The normal-incidence reflection profile over an interface that has a fault with a vertical throw of *h*. (b) The record section for (a). The time difference between the two-way travel times on either side of the fault is Δt. For the fault to be resolved, its throw *h* must be greater than about $\lambda/8$ to $\lambda/4$, where λ is the wavelength of the signal. (In terms of Δt, Δt must be greater than $1/(4 \times \text{frequency})$ to $1/(2 \times \text{frequency})$.)

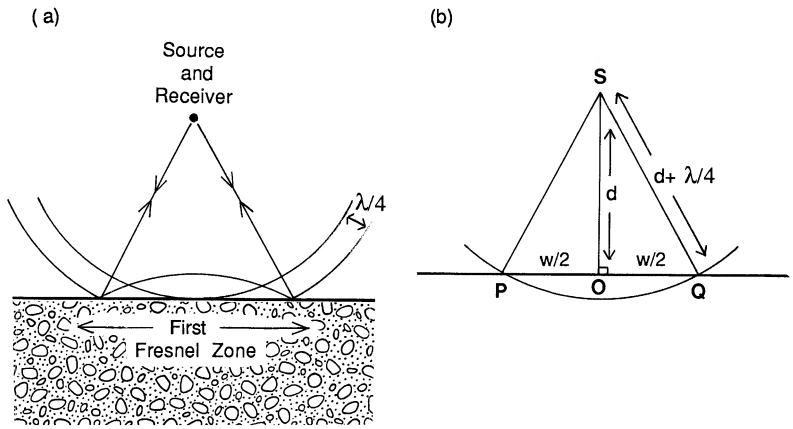


Figure 4.50. (a) Spherical waves from a point source reflected from a plane interface. The first Fresnel zone is that part of the interface which reflects energy back to the receiver within one-half a cycle of the first (normal-incidence) reflection. Because the wave must travel from source to interface and back to the receiver, energy from the wavefront one-quarter of a wavelength behind the first wavefront when reflected back to the receiver is delayed by half a cycle. (b) The geometry of the first Fresnel zone. The width *w* of this zone can be calculated by applying Pythagoras' theorem to the right triangle SOQ; *d* is the depth of the interface beneath the source and λ the wavelength of the signal. The upper material is assumed to have constant P-wave velocity α_1.

the same wave is travelling through material with a P-wave velocity of 7.2 km s^{-1}, the wavelength is increased to 360 m.

Consider a fault with a vertical throw of *h* (Fig. 4.49). For the fault to be detected, it must be greater than about one-eighth to one-quarter of the wavelength of the seismic wave. This means that the reflected wave from the down-faulted side is delayed one-quarter to one-half a wavelength. A smaller time delay than this is

Table 4.2 *The width of the Fresnel zone for specified depths and signal wavelengths*

	Wavelength (km)				
Depth (km)	0.10	0.20	0.30	0.40	0.50
2	0.63	0.90	1.11	1.28	1.44
5	1.00	1.42	1.74	2.01	2.25
10	1.42	2.00	2.45	2.84	3.17
30	2.45	3.46	4.24	4.90	5.48
50	3.16	4.47	5.48	6.33	7.08

unlikely to be resolved as being due to a fault since it would appear as a slight time error. Thus, a 20-Hz signal could resolve faults with throws in excess of about 30 m in 4.8-km s^{-1} material. However, in 6.4-km s^{-1} material, the same 20-Hz signal could only resolve faults with throws of more than 40 m; in 7.2-km s^{-1} material, the minimum resolvable throw would be 45 m. A 30-Hz signal would mean that faults of 20, 27 and 30 m, respectively, would be resolvable. Clearly, shorter wavelengths (higher frequencies) give better resolution of structures and greater detail than are obtainable from long wavelengths. Unfortunately, high-frequency vibrations do not travel as well through rock as low-frequency vibrations do (Section 8.1.3), so high frequencies are generally absent from reflections from deep structures.

The lateral extent that a reflector must have in order for it to be detectable is also controlled by wavelength. To understand the lateral resolution obtainable with a given source signal, we need to introduce the concept of the *Fresnel zone*. The first Fresnel zone is that part of a reflecting interface which returns energy back to the receiver within half a cycle of the first reflection (Fig. 4.50). In plan view, the first Fresnel zone is circular; the second Fresnel zone is a ring (concentric with the first Fresnel zone) from which the reflected energy is delayed by one-half to one cycle; and so on for the third, fourth and *n*th Fresnel zones. For an interface to be identified as such, it must be at least as wide as the first Fresnel zone (usually referred to as *the* Fresnel zone). If a reflecting zone is narrower than the first Fresnel zone, it effectively appears to be a diffractor.

The width of the Fresnel zone w can be calculated by applying Pythagoras' theorem to triangle SOQ of Fig. 4.50(b):

$$\left(d + \frac{\lambda}{4}\right)^2 = d^2 + \left(\frac{w}{2}\right)^2 \tag{4.100}$$

On rearranging this, we obtain

$$w^2 = 2d\lambda + \frac{\lambda^2}{4} \tag{4.101}$$

Table 4.2 gives values of w for various interface depths and wavelengths. Even at 2 km depth and with a wavelength of 0.10 km (which could correspond to $\alpha_1 = 3\ \text{km s}^{-1}$ and frequency 30 Hz, etc.) the width of the Freznel zone is 0.63 km. At depths appropriate for the middle or deep crust, the Fresnel zone is several kilometres in width, which is large on a geological scale. Deep-seismic-reflection profiling is therefore not able to give as clear a picture of the crust as one might at first hope.

4.4.5 Deep seismic-reflection profiling

The reflection data shown in Fig. 4.42(a) have two-way travel times of up to 4 s. Such recording times are used primarily when attempting to determine shallow sedimentary structures in oil exploration (seismic-reflection profiling is the main prospecting tool of the petroleum industry). However, it is also possible to use reflection profiles to probe the deeper structures in the Earth's crust. To do this, recording times of 15–30 s or more are necessary. In addition to long recording times, large energy sources are required for deep-crustal or uppermost-mantle reflections to be detected. As mentioned in Section 4.4.3, the best method of obtaining detailed velocity information on the lower crust and uppermost mantle is not normal-incidence reflection profiling but wide-angle reflection profiling. Recall from Eq. (4.65) that, since the curvature of the reflection hyperbola decreases with increasing velocity (and increasing depth), deep interval velocities can be only rather poorly determined. However, normal-incidence-reflection and wide-angle-reflection methods used together allow velocities and depths to be much better determined.

Deep seismic-reflection profiling in the U.S.A. started with the Consortium for Continental Reflection Profiling (COCORP) profile shot in Hardman County, Texas, in 1975. Since then, many profiles have been shot by many organizations around the world. The method is increasingly being used to relate surface and shallow geology to structures at depth, to trace thrusts and to map plutons. Deep seismic-reflection profiling has dramatically increased our knowledge of the structure of the continental crust. The Moho is frequently a reflector, although not unequivocably observed on every reflection line. The resolution obtainable with deep reflection profiling far exceeds that with seismic-refraction or wide-angle-reflection, or magnetic, electrical or gravitational methods. However, as with every method, the ultimate questions which remain unanswered concern the interpretation of the physical parameters (in this case, reflections and velocities) in terms of geology. For example, some interpretations ascribe lower-crustal reflections to lithology (rock type) and changes in metamorphic facies, whereas others ascribe reflections to physical factors, perhaps the presence of fluids trapped in the lower crust, or to sheared zones.

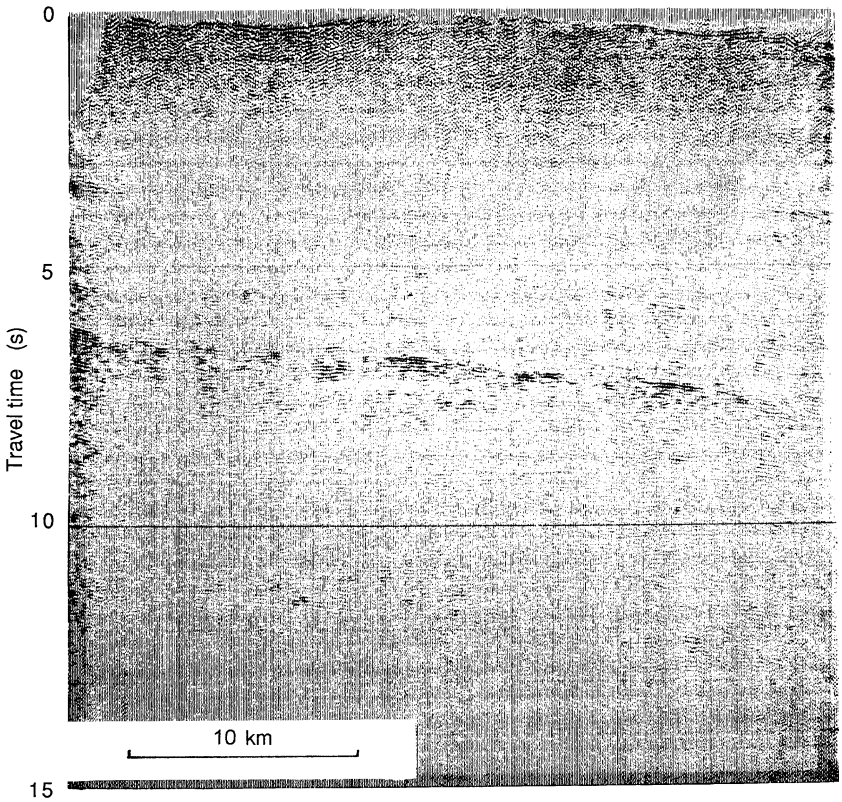


Figure 4.51. The Consortium for Continental Reflection Profiling (COCORP) has shot reflection profiles over many types of continental structures in the U.S.A.: accreted terranes, extensional rifts, thrust zones and so on. This record section is part of a series of profiles shot across the Rio Grande Rift, New Mexico, an active extensional rift (discussed in Section 10.4.3). The high-amplitude reflections at about 7 s two-way time in the midcrust are inferred to come from the lid of a magma chamber. The solid–liquid interface gives a large reflection coefficient. (From Brown (1986), reprinted by permission.)

As an example, a reflection coefficient of about 0.10 is necessary in order to explain the strong lower-crustal and upper-mantle reflections seen on some deep seismic-reflection profiles. Such values are not easily achieved by juxtaposition of proposed materials or measured velocities. A P-wave velocity contrast of 8.1–8.3 km s^{-1} in the mantle gives a reflection coefficient (Eq. (4.62)) of 0.012, assuming that there is no change in density. Even if somewhat larger velocity contrasts are considered and a change in density is included, the reflection coefficient could hardly increase by an order of magnitude to 0.1. Mantle reflections of the correct magnitude could, however, be produced by regions of partially hydrated peridotite (i.e., water acting on mantle peridotite to produce serpentine, which has a very much lower P-wave velocity), by bodies of gabbro or eclogite in a peridotitic mantle or by mylonite zones.

Details of the structures as revealed by reflection profiling are discussed in Chapter 10; here we point out two illustrations of the type of data which can be obtained: Fig. 4.51, which is a record section obtained by COCORP over the Rio Grande Rift in New Mexico, U.S.A.; and Fig. 4.52, which is part of a record section obtained by the British Institutions Reflection Profiling Syndicate (BIRPS) off the north coast of Scotland.

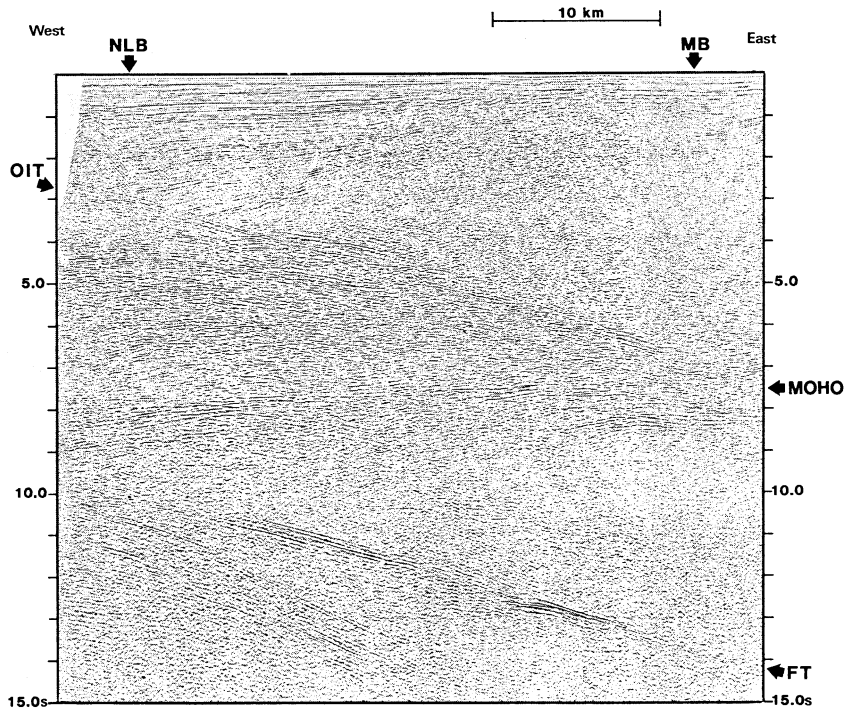


Figure 4.52. The British Institutions Reflection Profiling Syndicate (BIRPS), which started data acquisition in 1981, collected a large amount of data on the continental shelf around the British Isles. The short segment of unmigrated 30-fold data shown here (taken from the WINCH profile) is from an east–west line off the north coast of Scotland. NLB, North Lewis Basin; MB, Minch Basin; OIT, Outer Isles Thrust; and FT, Flannan Thrust. The Outer Isles Thrust can be traced from the surface down into the lower crust with a dip of approximately 25°. The basins are Mesozoic half-grabens. The North Lewis Basin and other basins apparently formed as a result of extension associated with the opening of the North Atlantic. The structure of the basins was controlled by the fabric of the existing Caledonian basement, in this instance the Outer Isles Thrust. The Moho on this section is a very clear reflector at 8–9 s two-way time (depth about 27 km). The Flannan Thrust cuts the Moho and extends into the upper mantle as a very clear reflector. Further data from this area are shown in Figs. 10.25 and 10.26. (From Brewer *et al.* (1983).)

Problems

1. How would you distinguish between Love waves and Rayleigh waves if you were given an earthquake record from a WWSSN station?
2. What is the wavelength of a surface wave with a period of (a) 10 s, (b) 100 s and (c) 200 s? Comment on the use of surface waves in resolving small-scale lateral inhomogeneities in the crust and mantle.
3. How much greater is a nuclear-explosion body-wave amplitude likely to be than an earthquake body-wave amplitude if both waves have the same M_S value and are recorded at the same distance?
4. (a) Which types of seismic wave can propagate in an unbounded solid medium?

(b) Which types of seismic wave can be detected by a vertical-component seismometer?

5. During a microearthquake survey in central Turkey, an earthquake was recorded by three seismometers (Fig. 4.53). Calculate the origin time for this earthquake and locate the epicentre relative to the seismometers by using information in the table that follows.

	Hours	Minutes	Seconds
Seismometer 1			
P-wave arrival time	13	19	58.9
S-wave arrival time	13	20	4.7
Seismometer 2			
P-wave arrival time	13	20	2.6
S-wave arrival time	13	20	10.8
Seismometer 3			
P-wave arrival time	13	19	54.5
S-wave arrival time	13	19	57.4

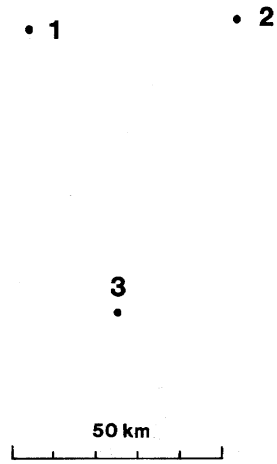


Figure 4.53. A map showing the locations of seismometers 1, 2 and 3.

Assume that the focus was at the surface and use P- and S-wave velocities of 5.6 and 3.4 km s^{-1}, respectively.

6. Calculate the amount by which the seismic energy released by an earthquake increases when the surface-wave magnitude increases by one unit. Repeat the calculation for an increase of one unit in the body-wave magnitude.
7. The daily electrical consumption of the U.S.A. in 2001 was about 9.9×10^9 kilowatt hours. If this energy were released by an earthquake, what would its magnitude be?
8. It is sometimes suggested that small earthquakes can act as safety valves by releasing energy in small amounts and so averting a damaging large earthquake. Assuming that an area (perhaps California) can statistically expect a large earthquake ($M_S > 8$) once every 50–100 years, calculate how many smaller earthquakes with (a) $M_S = 6$, (b) $M_S = 5$ or (c) $M_S = 4$ would be required during these years in order to release the same amount of energy. If it were possible to trigger these smaller earthquakes, would it help?
9. (a) Calculate the total amount of energy released each year by earthquakes. Suggest an estimate of the amount by which this figure may be in error.
   (b) Compare the total amount of energy released each year by earthquakes with the energy lost each year by heat flow (Chapter 7).
10. Using the earthquake travel-time curves in Fig. 4.16 and the two earthquake records in Fig. 4.17, determine for each earthquake (a) the epicentral distance (in degrees) and (b) the origin time.
11. Explain how the epicentre, focal mechanism and slip vector of an earthquake are determined. (Cambridge University Natural Sciences Tripos IB, 1974.)

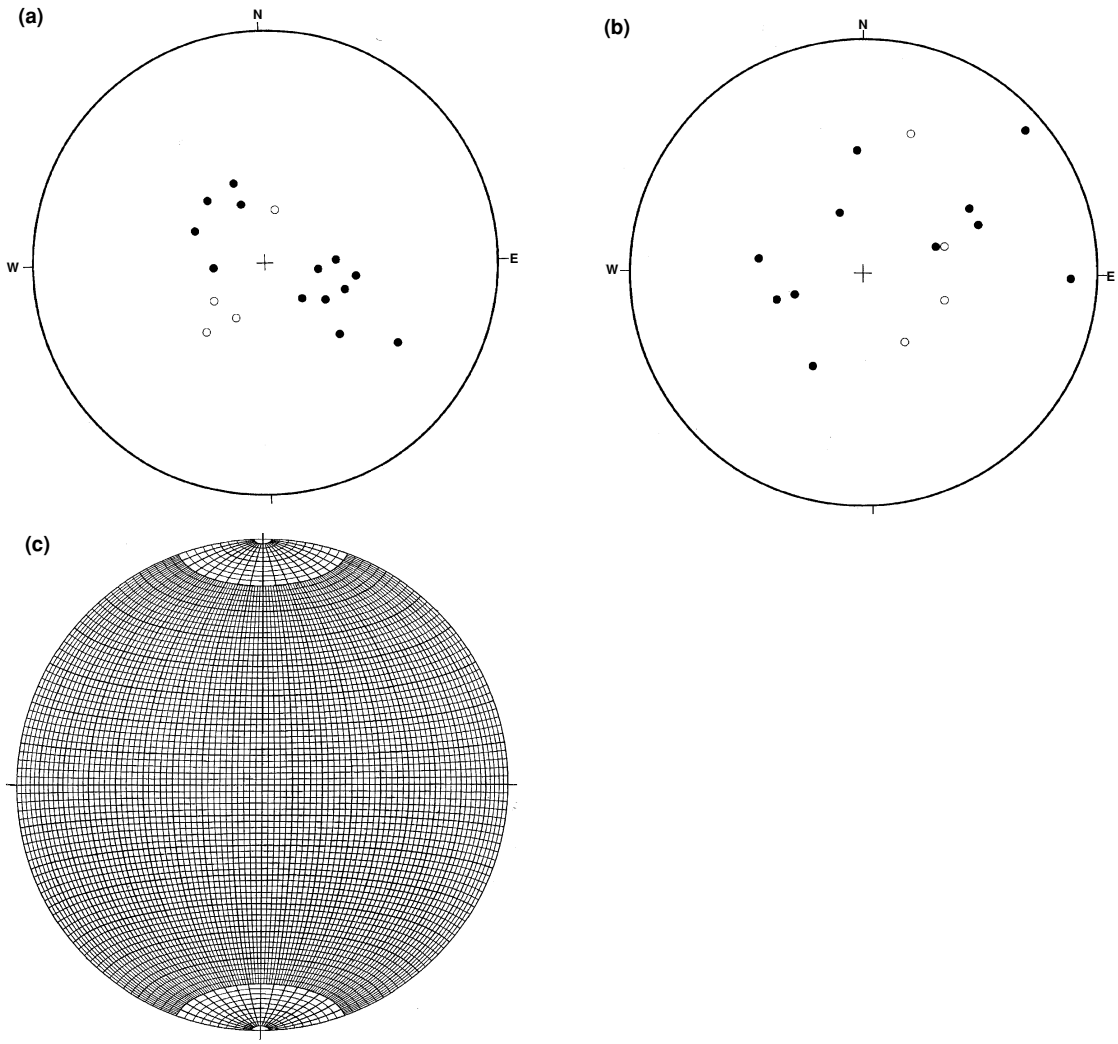


Figure 4.54. (a) Directions of first motion for an earthquake that occurred in Iran. •, compression (positive); ○, dilatation (negative). (b) Directions of first motion for an earthquake that occurred in the Eastern Atlantic. (c) An equal-area (Schmidt) net.

12. Figures 4.54(a) and (b) give the directions of first motion for two earthquakes, plotted on equal-area projections of the lower focal hemisphere. For each earthquake, do the following.
    (a) Draw the two perpendicular planes that divide the focal sphere into positive and negative regions.
    (b) Find the strike and dip of these planes.
    (c) Plot the horizontal component of strike.
    (d) Discuss the mechanism of the earthquake.

    N.B. You will need to photocopy and enlarge the equal-area projections.

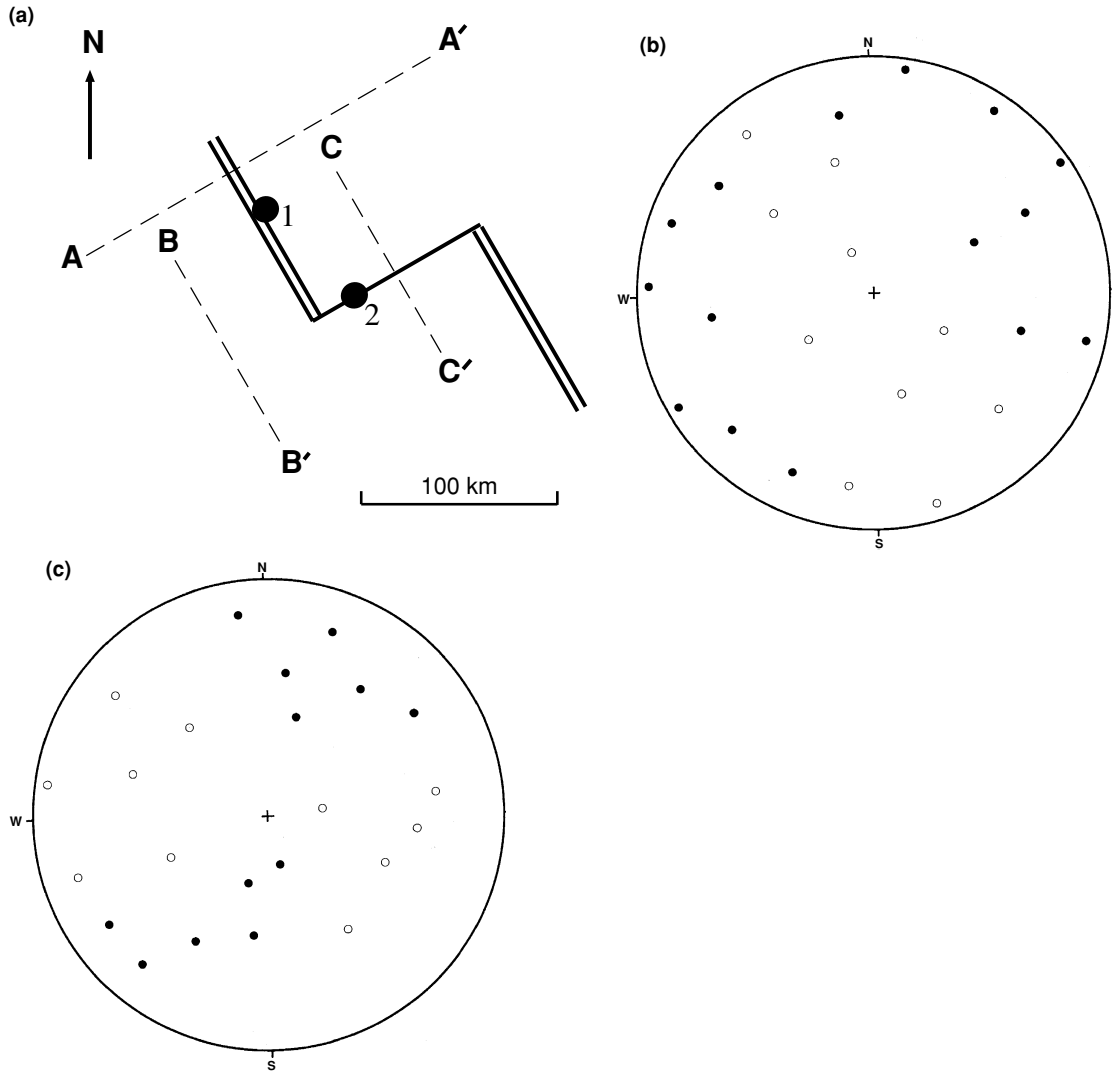


Figure 4.55. (a) Locations of earthquakes 1 and 2 on the Mid-Atlantic Ridge. (b) Directions of first motion for earthquake 1. (c) Directions of first motion for earthquake 2. Symbols denote compression and dilatation as in Fig. 4.54.

13. Contrast the distribution of earthquake foci in and around the Atlantic and Pacific Oceans and describe the mechanism of earthquakes on the mid-ocean ridges. What light do these facts throw on the plate motions? (Cambridge University Natural Sciences Tripos IB, 1976.)
14. The outline map in Fig. 4.55(a) shows the positions of two earthquakes recorded on the Mid-Atlantic ridge. Focal spheres of earthquakes 1 and 2 are also shown (Figs. 4.55(b) and (c)).

(a) Draw the fault-plane solutions on the focal spheres for earthquakes 1 and 2.
(b) Indicate the fault types which generated earthquakes 1 and 2. From the tectonic setting of each earthquake, explain how it is possible to infer which plane is likely to be the fault plane. List the strikes and dips of the probable fault planes and the probable slip vectors.
(c) Sketch the bathymetric profiles along sections A–A′, B–B′ and C–C′.
(d) In this tectonic setting, over what depth range would you expect the earthquakes to occur? (Cambridge University Natural Sciences Tripos IB, 1983.)

15. The boundary between the African and Eurasian plates, between the Azores Triple Junction and Gibraltar, strikes approximately east–west. What is the nature of this boundary? Sketch fault-plane solutions for the following:
(a) earthquakes along this plate boundary west of Gibraltar; and
(b) earthquakes east of Gibraltar in the Mediterranean.

16. A 200-m thick layer of sediment with P-wave velocity $2\,\mathrm{km\,s^{-1}}$ overlies material with velocity $3\,\mathrm{km\,s^{-1}}$. Calculate the critical distance for the refracted wave from the second layer. If you had to determine the velocities and thickness of a structure such as this by shooting a seismic-refraction experiment, indicate how many seismometers you would require and over what distances you would deploy them.

17. The first-arrival time–distance data from a seismic-refraction experiment are as follows:

refraction velocity $6.08\,\mathrm{km\,s^{-1}}$, intercept 1.00 s
refraction velocity $6.50\,\mathrm{km\,s^{-1}}$, intercept 1.90 s
refraction velocity $8.35\,\mathrm{km\,s^{-1}}$, intercept 8.70 s

(a) Calculate the velocity–depth structure.
(b) Now assume that there is a hidden layer: a refractor with velocity $7.36\,\mathrm{km\,s^{-1}}$ and intercept 5.50 s is a second arrival at all ranges. By how much is the total crustal thickness increased?

18. The first-arrival travel times and distances obtained from a refraction-experiment shot along a mid-ocean-ridge axis are tabulated below.

Offset distance (km)	First-arrival travel time (s)
2.0	4.10
3.0	4.45
4.0	4.85
5.0	5.20
6.0	5.40
7.0	5.60
8.0	5.80
9.0	5.95
10.0	6.15
12.0	6.50

(*cont.*)

14.0	6.75
16.0	7.00
18.0	7.35
20.0	7.50
22.0	7.90
24.0	8.10
26.0	8.40
28.0	8.70
30.0	8.90

(a) Plot the data on a time–distance graph and determine a crustal structure for the line, assuming that it can be approximated by uniform horizontal layers. The water is 3 km deep and has a P-wave velocity of 1.5 km s^{-1}.

(b) If the upper mantle has a normal P-wave velocity of 8.1 km s^{-1} in this region, what is its minimum depth?

(c) Repeat the calculation assuming that the upper mantle has a velocity of only 7.6 km s^{-1}.

(d) Ideally, how long should the refraction line have been to make a measurement of the upper-mantle velocity and depth?

(e) Explain how a record section would assist in determining the velocity structure.

19. Using the velocities for Love waves along continental ray paths (Fig. 4.6(a)), make a rough estimate of the minimum time over which Love waves arrive at a seismometer at epicentral distances of (a) 20° and (b) 100°. Assume that the seismometer records waves with periods 20–250 s.

20. Calculate a velocity–depth model for the data shown in Fig. 4.42. Assume that the structure is horizontally layered and that all layers have constant velocity. Use only the reflections at 1.2, 1.5, 1.8, 2.0, 2.3 and 2.45 s.

21. For the simple example used in Section 4.4.3,

$$t_{0,1} = 1.0\,\text{s}, \qquad \overline{\alpha}_1 = 3.6\ \text{km s}^{-1}$$
$$t_{0,2} = 1.5\,\text{s}, \qquad \overline{\alpha}_1 = 4.0\,\text{km s}^{-1}$$

calculate the errors in velocity and thickness which result in the following cases.

(a) The travel times have an error of ±0.05 s.

(b) The travel times have an error of ±0.01 s.

(c) The RMS velocities have an error of ±0.1 km s^{-1}.

Comment on the accuracy of interval velocity–depth models obtained from RMS velocities and two-way travel times.

22. Calculate the normal-incidence P-wave reflection coefficients for the following interfaces:

(a) sandstone ($\alpha = 3.0$ km s^{-1}, $\rho = 2.2 \times 10^3$ kg m^{-3}) above a limestone ($\alpha = 4.1$ km s^{-1}, $\rho = 2.2 \times 10^3$ kg m^{-3});

(b) a possible lower-crustal interface of basalt ($\alpha = 6.8$ km s^{-1}, $\rho = 2.8 \times 10^3$ kg m^{-3}) over granulite ($\alpha = 7.3$ km s^{-1}, $\rho = 3.2 \times 10^3$ kg m^{-3}); and

(c) a possible crust-mantle interface of granulite ($\alpha = 7.3\,\text{km s}^{-1}$, $\rho = 3.2 \times 10^3\,\text{kg m}^{-3}$) over peridotite ($\alpha = 8.1\,\text{km s}^{-1}$, $\rho = 3.3 \times 10^3\,\text{kg m}^{-3}$).

23. Calculate the wavelengths of 20-, 30- and 40-Hz seismic P-waves in these rocks:
    (a) sandstone ($\alpha = 3.0\,\text{km s}^{-1}$) and
    (b) gabbro ($\alpha = 7.0\,\text{km s}^{-1}$).
    What do these calculations indicate about the resolution of shallow and deep crustal structures by seismic-reflection profiling.
24. Given the following time–distance data from a seismic-reflection survey, calculate the depth of the reflector and a velocity for the material above it. What type is this velocity?

Distance (km)	Travel time (s)
0.5	1.01
1.0	1.03
1.5	1.07
2.0	1.11
2,5	1.17
3.0	1.25
3.5	1.33

25. If normal-incidence reflection lines were shot over the structure shown in Fig. 4.56, mark on record sections the arrival times of the phases that would be recorded. What would be the relative amplitudes of the phases and by what physical factors are they controlled?

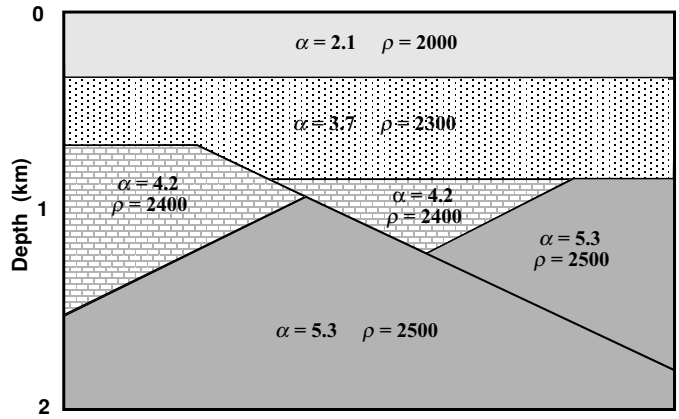


Figure 4.56. The vertical cross section for Problem 25, showing P-wave velocities in km s^{-1} and densities in kg m^{-3}.

26. Using the parametric form of the travel-time and epicentral-distance equations for earthquakes (see Appendix 3), obtain the following relation:

$$t = p\Delta + 2\int_{r_{\min}}^{r_{\max}} \frac{(\eta^2 - p^2)^{1/2}}{r}\,\mathrm{d}r$$

27. A region is believed to have the following horizontally layered structure.

Depth (km)	P-wave velocity (km s^{-1})
0.0–0.5	1.0
0.5–2.0	3.0
2.0–5.0	5.5
5.0–∞	7.0

(a) Plot the time–distance curve for the structure.
(b) How would you position shots and receivers to test whether the assumed structure is correct? Illustrate with simple layout diagrams and outline your reasoning.
(c) How can you exclude the possibility that the structure has dipping interfaces?
(d) Give two methods by which a velocity inversion in the structure could be detected.
(Cambridge University Natural Science Tripos II, 1986.)

28. Ewing Oil decides to shoot two marine seismic surveys. For the first survey, water guns are fired every 17 m, the hydrophone streamer is 2.45 km long and consists of 72 separate hydrophone sections and both the source and the streamers are towed at 5 m depth. In the second survey, air guns are fired every 50 m, the streamer is 5 km long and consists of 100 sections and both the source and the receivers are towed at 15 m depth.
(a) If the velocity of sea water is 1.5 km s^{-1}, what is the peak frequency for each survey?
(b) What is the 'fold' of each survey (i.e., the number of separate traces stacked together in the final section)?
(c) One survey was intended to be a high-resolution survey of a shallow sedimentary section, and the other was a deep survey to image the whole crust. Which survey is which? Give your reasoning.

29. The shallow survey of Problem 28 revealed a simple sedimentary sequence, shown in the seismic section of Fig. 4.57. The normal moveouts observed for each reflector when the source-to-receiver offset was 1000 m are listed below.

	Two-way time (s)	Δt_{NMO} (ms)
First reflection	1	222
Second reflection	2	77
Third reflection	3	27

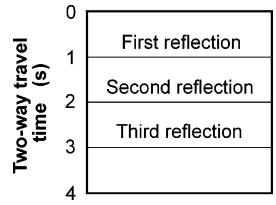


Figure 4.57. The seismic section for Problem 29.

(a) What is the appropriate stacking velocity for each reflector?
(b) Calculate the corresponding interval velocities, and from these determine the depth to the deepest reflector.
(c) Suggest a plausible stratigraphy for this area.

30. The deep survey of Problem 28 was run over the cross section shown in Fig. 4.58. Velocities and densities are given below.

	Sedimentary rocks	Basement 1	Basement 2
Velocity (km s^{-1})	3.0	6.0	6.0
Density (10^3 kg m^{-3})	2.2	2.6	2.7

(a) By considering rays perpendicular to the reflectors, draw the corresponding time section. Use a scale of 2 cm = 5 km and 1 cm = 1 s. Ignore diffractions.

(b) Draw a rough sketch of the shape of the Moho which would be obtained after migration if this section had mistakenly been migrated at a constant velocity of 6 km s^{-1}. Ignore diffractions.

31. What is the reflection coefficient of the fault shown in Fig. 4.58, (a) at the sediment–basement contact and (b) within the basement?

32. Desperate to find something of interest in the data of Problems 29 and 30, J. R. Ewing authorized heat-flow measurements over the sedimentary basin of Fig. 4.58. If the surface heat flow is 0.75 mW m^{-2} and the thermal conductivity is 2.5 W m^{-1} °C^{-1}, what is the temperature (a) at the base of the sediments and (b) at the Moho? State your assumptions and comment on your answer (see Chapter 7). (Problems 28–32 are from Cambridge Natural Science Tripos II, 1986.)

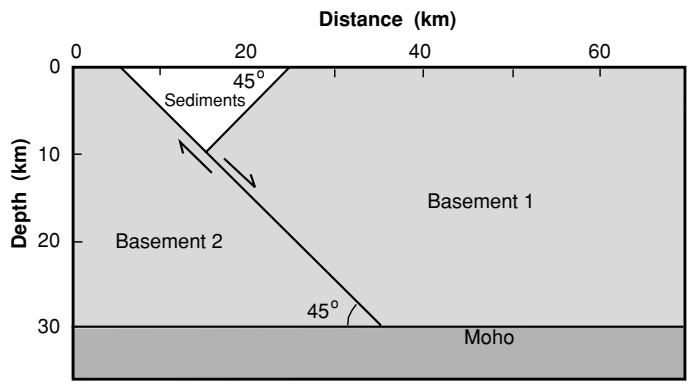


Figure 4.58. The vertical cross section for Problems 30 and 31.

33. Your foot is trapped under a railway sleeper. You feel the rail vibrate and then 20 s later hear a faint whistle. Estimate how much time you have to free yourself. State your assumptions.

References and bibliography

Abercrombie, R. and Leary, P. 1993. Source parameters of small earthquakes recorded at 2.5 km depth, Cajun Pass, Southern California: implications for earthquake scaling. *Geophys. Res. Lett.*, **20**, 1511–14.

Aki, K. and Richards, P. G. 2001. *Quantitative Seismology – Theory and Methods*, 2nd edn, New York: Columbia Press.

Amelung, F. and King, G. 1997. Large-scale tectonic deformation inferred from small earthquakes. *Nature*, **386**, 702–5.

Astiz, L., Shearer, P. and Earle, P. 1996. Global stacking of broadband seismograms, *Seismol. Res. Lett.*, **67**, 8–18.

Båth, M. 1966. Earthquake energy and magnitude. *Physics and Chemistry of the Earth*, Vol. 7. New York: Pergamon, pp. 115–65.

1977. Teleseismic magnitude relations. *Ann. Geofis.*, **30**, 299–327.

1979. *Introduction to Seismology*, 2nd edn. Basel: Birkhäuser-Verlag.

1981. Earthquake magnitude – recent research and current trends. *Earth Sci. Rev.*, **17**, 315–98.

Birch, F. 1964. Density and composition of the Earth's mantle and core. *J. Geophys. Res*., **69**, 4377–88.

1968. On the possibility of large changes in the Earth's volume. *Phys. Earth Planet. Int.*, **1**, 141–7.

Bolt, B. A. 1976. *Nuclear Explosions and Earthquakes: The Parted Veil*. San Francisco, California: Freeman.

1982. *Inside the Earth: Evidence from Earthquakes*. San Francisco, California: Freeman.

1999. *Earthquakes*, 4th edn. San Francisco, California: Freeman.

Bullen, K. E. 1947. *An Introduction to the Theory of Seismology*. Cambridge: Cambridge University Press.

Brewer, J. A., Matthews, D. H., Warner, M. R., Hall, J., Smythe, D. K. and Whittington, R. J. 1983. BIRPS deep seismic reflection studies of the British Caledonides. *Nature*, **305**, 206–10.

Brown, L. D. 1986. Aspects of COCORP deep seismic profiling. In M. Barazangi and L. D. Brown, eds., *Reflection Seismology: A Global Perspective*. Vol. 13 of Geodynamics Series. Washington: American Geophysical Union, pp. 209–22.

Brown, L. D., Kaufman, S. and Oliver, J. E. 1983. COCORP seismic traverse across the Rio Grande Rift. In A. W. Bally, ed., *Seismic Expression of Structural Styles*. American Association of Petroleum Geologists Studies in Geology No. 15. Tulsa, Oklahoma: American Association of Petroleum Geologists, Section 2.2.1, pp. 1–6.

Bullen, K. E. and Bolt, B. A. 1985. *An Introduction to the Theory of Seismology*, 4th edn. Cambridge: Cambridge University Press.

Cerveny, V. 1979. Ray theoretical seismograms for laterally inhomogeneous structures. *J. Geophys. Res.*, **46**, 335–42.

Cerveny, V., Popov, M. M. and Psencik, I. 1982. Computation of wave fields in inhomogeneous media – Gaussian beam approach. *Geophys. J. Roy. A str. Soc.*, **70**, 109–28.

Chapman, C. H. 1978. A new method for computing synthetic seismograms. *Geophys. J. Roy. Astr. Soc.*, **54**, 481–518.

1985. Ray theory and its extensions: WKBJ and Maslov seismograms. *J. Geophys.*, **58**, 27–43.

Chapman, C. H. and Drummond, R. 1982. Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory. *Bull. Seism. Soc. Am.*, **72**, S277–317.

Choy, G. L. and Boatwright, J. L. 1995. Global patterns of radiated seismic energy and apparent stress. *J. Geophys. Res.*, **100**, 18 205–28.

Claerbout, J. 1985. *Imaging the Earth's Interior*. Oxford: Blackwell.

Clark, S. P. 1966. *Handbook of Physical Constants*. Vol. 97 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America.

Cox, A. and Hart, R. B. 1986. *Plate Tectonics: How It Works*. Oxford: Blackwell Scientific.

Das, S. and Henry, C. 2003. Spatial relation between main earthquake slip and its aftershock distribution. *Rev. Geophys*., **41**, 3/doi 10.1029/2002 RG 000119.

Davies, D. and McKenzie, D. P. 1969. Seismic travel-time residuals and plates. *Geophys. J. Roy. Astr. Soc.*, **18**, 51–63.

Davis, J., Sykes, L., House, L. and Jacob, K. 1981. Shumagin seismic gap, Alaska peninsula: history of great earthquakes, tectonic setting and evidence for high seismic potential. *J. Geophys. Res.*, **86**, 3821–55.

Dey-Sarkar, S. K. and Chapman, C. H. 1978. A simple method for the computation of body wave seismograms. *Bull. Seism. Soc. Am.*, **68**, 1577–93.

Dobrin, M. B. and Savit, C. H. 1988. *Introduction to Geophysical Prospecting*, 4th edn. New York: McGraw-Hill.

Dziewonski, A. M. and Anderson, D. L. 1981. Preliminary reference earth model. *Phys. Earth Planet. Interior*, **25**, 297–356.

Earle, P. S. and Shearer, P. M. 1994. Characterization of global seismograms using an automatic-picking algorithm. *Bull. Seism. Soc. Am.*, **84**, 366–76.

Frohlich, C. 1989. Deep earthquakes. *Sci. Am.*, **259** (1), 48–55.

1989. The nature of deep-focus earthquakes. *Ann. Rev. Earth Planet. Sci.*, **17**, 227–54.

1998. Does maximum earthquake size depend upon focal depth? *Bull. Seism. Soc. Am.*, **88**, 329–36.

Frohlich, C. and Davis, S. D. 1993. Teleseismic *b* values; or much ado about 1.0. *J. Geophys. Res.*, **98**, 631–44.

Fuchs, K. and Müller, G. 1971. Computation of synthetic seismograms with the reflectivity method and comparison with observations. *Geophys. J. Roy. Astr. Soc.*, **23**, 417–33.

Garland, G. D. 1979. *Introduction to Geophysics – Mantle, Core and Crust*, 2nd edn. Philadelphia, Pennsylvania: Saunders.

Gebrande, H. 1982. In K. H. Hellwege, ed., *Landolt-Bornstein*, Group 5, vol. 16, *Physical Properties of Rocks*, ed. G. Angenheister. Berlin: Springer-Verlag, pp. 1–96.

Gubbins, D. 1989. *Seismology and Plate Tectonics*. Cambridge: Cambridge University Press.

Gutenberg, B. 1945. Amplitudes of P, PP and S and magnitude of shallow earthquakes. *Bull. Seism. Soc. Am.*, **35**, 57–69.

1945. Magnitude determination for deep focus earthquakes. *Bull. Seism. Soc. Am.*, **35**, 117–30.

1959. *Physics of the Earth's Interior*. New York: Academic Press.

Gutenberg, B. and Richter, C. F. 1954. *Seismicity of the Earth and Associated Phenomena*. Princeton, New Jersey: Princeton University Press.

1956. Magnitude and energy of earthquakes. *Ann. Geofis.*, **9**, 1–15.

Hanks, T. C. and Kanamori, H. 1979. A moment magnitude scale. *J. Geophys. Res.*, **84**, 2348–2350.

Hatton, L., Worthington, M. H. and Makin, J. 1985. *Seismic Data Processing: Theory and Practice*. Oxford: Blackwell Scientific.

Heki, K., Miyazaki, S. and Tsuji, H. 1997. Silent fault slip following an interplate thrust earthquake at the Japan Trench. *Nature*, **386**, 595–8.

Helmberger, D. W. 1968. The crust–mantle transition in the Bering Sea. *Bull. Seism. Soc. Am.*, **58**, 179–214.

Henry, C. and Das, S. 2001. Aftershock zones of large shallow earthquakes: fault dimensions, aftershock area expansion and scaling relations. *Geophys. J. Int.*, **147**, 272–93.

Herrin, E. 1968. Introduction to 1968 seismological tables for P phases. *Bull. Seism. Soc. Am.*, **58**, 1193–241.

Huang, X., West, G. F. and Kendall, J.-M. 1998. A Maslov–Kirchoff seismogram method. *J. Geophys. Int.*, **132**, 595–602.

Hyndman, R., Spence, G. D., Chapman, R., Riedel, M. and Edwards, R. N. 2002. Geophysical studies of marine gas hydrates in northern Cascadia. In *Natural Gas Hydrates: Occurence, Distribution, and Detection*. AGU Geophysical Monograph 124. Washington: American Geophysical Union, pp. 273–295.

Jeffreys, H. 1926. On near earthquakes. *Mon. Not. Roy. Astr. Soc. Geophys. Suppl.*, **1**, 385.

1939. The times of P, S and SKS and the velocities of P and S. *Mon. Not. Roy. Astr. Soc. Geophys. Suppl.*, **4**, 498–533.

1939. The times of PcP and ScS. *Mon. Not. Roy. Astr. Soc. Geophys. Suppl.*, **4**, 537–47.

1976. *The Earth*, 6th edn. Cambridge: Cambridge University Press.

Jeffreys, H. and Bullen, K. E. 1988. *Seismological Tables*. Cambridge: British Association Seismological Investigations Committee, Black Bear Press.

Johnston, A. C. and Schweig, E. S. 1996. The enigma of the New Madrid earthquakes of 1811–1812. *Ann. Rev. Earth Planet. Sci.*, **24**, 339–84.

Kagan, Y. Y. 1999. Universality of the seismic moment–frequency relation. *Pure Appl. Geophys*., **155**, 537–73.

Kanamori, H. 1994. Mechanics of earthquakes. *Ann. Rev. Earth Planet. Sci.*, **22**, 207–37.

Kanamori, H. and Anderson, D. L. 1975. Theoretical basis of some empirical relations in seismology. *Bull. Seism. Soc. Am.*, **65**, 1073–95.

Kendall, J.-M. and Thompson, C. J. 1993. Maslov ray summation, pseudo-caustics, Lagrangian equivalence and transient seismic waveforms. *Geophys. J. Int.*, **113**, 186–214.

Kennett, B. L. N. 1975. The effects of attenuation on seismograms. *Bull. Seism. Soc. Am.*, **65**, 1643–51.

1977. Towards a more detailed seismic picture of the oceanic crust and mantle. *Mar. Geophys. Res.*, **3**, 7–42.

1995. Seismic traveltime tables. In T. J. Ahrens, ed., *Global Earth Physics: A Handbook of Physical Constants.* AGU Reference Shelf 1. Washington:. American Geophysical Union, pp. 126–43.

2002. *The Seismic Wavefield 1: Introduction and Theoretical Development*. Cambridge: Cambridge University Press.

2003. *The Seismic Wavefield 2*: *Interpretation of Seismograms on Regional and Local Scales*. Cambridge: Cambridge University Press.

Kennett, B. L. N. and Engdahl, E. R. 1991. Travel times for global earthquake location and phase identification, *Geophys. J. Int.*, **105**, 429–65.

Kerr, R. A. 1979. Earthquake – Mexican earthquake shows the way to look for the big ones. *Science*, **203**, 860–2.

Knopoff, L. and Chang, F.-S. 1977. The inversion of surface wave dispersion data with random errors. *J. Geophys.*, **43**, 299–309.

Lay, T. and Wallace, T. 1995. *Modern Global Seismology*. San Diego, California: Academic Press, 521 pp.

Liu, L.-G. 1974. Birch's diagram: some new observations. *Phys. Earth Planet. Interior*, **8**, 56–62.

Love, A. E. H. 1911. *Some Problems of Geodynamics*. Cambridge: Cambridge University Press.

Lowrie, W. 1997. *Fundamentals of Geophysics*. Cambridge: Cambridge University Press.

Ludwig, W. J., Nafe, J. E. and Drake, C. L. 1970. Seismic refraction. In A. E. Maxwell, ed., *The Sea*, Vol. 4, Part 1. New York: Wiley-Interscience, pp. 53–84.

McGeary, S. and Warner, M. R. 1985. DRUM: seismic profiling of the continental lithosphere. *Nature*, **317**, 795–7.

McKenzie, D. P. 1983. The earth's mantle. *Sci. Am.*, **249** (3), 66–113.

McKenzie, D. P. and Parker, R. L. 1967. The North Pacific: an example of tectonics on a sphere. *Nature*, **216**, 1276–80.

Main, I. G. 1997. Long odds on prediction. *Nature*, **385**, 19–20.

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K. and Rabaute, T. 1993. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, **364**, 138–42.

Masters, T. G. and Shearer, P. M. 1995. Seismic models of the earth: elastic and anelastic. In T. J. Ahrens, ed., *Global Earth Physics: A Handbook of Physical Constants*. AGU Reference Shelf 1. Washington: American Geophysical Union, pp. 88–103.

Masters, T. G. and Widmer, R. 1995. Free oscillations: frequencies and attenuations. In T. J. Ahrens, ed., *Global Earth Physics: A Handbook of Physical Constants*. AGU Reference Shelf 1. Washinton: American Geophysical Union, pp. 104–25.

Meissner, R. 1986. *The Continental Crust, a Geophysical Approach*. Vol. 34 of International Geophysics Series. Orlando, Florida: Academic Press.

Mueller, K., Hough, S. E. and Bilham, R. 2004. Analysing the 1811–1812 New Madrid earthquakes with recent instrumentally recorded aftershocks. *Nature*, **429**, 284–8.

Müller, G. 1970. Exact ray theory and its applications to the reflection of elastic waves from vertically homogeneous media. *Geophys. J. Roy. Astr. Soc.*, **21**, 261–84.

1985. The reflectivity method: a tutorial. *J. Geophys.*, **58**, 153–74.

Müller, G. and Kind, Roy. 1976. Observed and computed seismogram sections for the whole earth. *Geophys. J. Roy. Astr. Soc.*, **44**, 699–716.

National Research Council (U.S.A.) 1983. *Workshop on Guidelines for Instrumentation Design in Support of a Proposed Lithospheric Seismology Program*. Washington: National Research Council.

Nowroozi, A. A. 1986. On the linear relation between m_b and M_s for discrimination between explosions and earthquakes. *Geophys. J. Roy. Astr. Soc.*, **86**, 687–99.

Nishenko, S. P. 1991. Circum-Pacific seismic potential: 1989–1991. *Pure Appl. Geophys.*, **135**, 169–259.

Officer, C. B. 1974. *Introduction to Theoretical Geophysics*. Berlin: Springer-Verlag.

Pacheco, J. F., Scholz, C. H. and Sykes, L. R. 1992. Changes in frequency–size relationship from small to large earthquakes. *Nature*, **355**, 71–3.

Panza, G. F. 1981. The resolving power of seismic surface waves with respect to crust and upper mantle structural models. In R. Cassinus, ed., *The Solution of the Inverse Problem in Geophysical Interpretation*. New York: Plenum, pp. 39–78.

Parsons, T., Toda, S., Stein, R. S., Barka, A. and Dieterich, J. H. 2000. Heightened odds of large earthquakes near Istanbul: and interaction-based probability calculation. *Science*, **288**, 661–5.

Pegler, G. and Das, S. 1996. Analysis of the relationship between seismic moment and fault length for large crustal strike–slip earthquakes between 1972–92. *Geophys. Res. Lett.*, **23**, 905–8.

Phillips, O. M. 1968. *The Heart of the Earth*. San Francisco: Freeman, Cooper & Co.

Press, F. 1966. Seismic velocities. In S. P. Clark, ed., *Handbook of Physical Constants*. Vol. 97 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America, pp. 195–218.

1975. Earthquake prediction. *Sci. Am.*, **232** (May), 14–23.

Rayleigh, Lord (J. W. Strutt) 1887. On waves propagated along the plane surface of an elastic solid. *London Math. Soc. Proc.*, **17**, 4–11.

Richards, P. G. and Zavales, J. 1990. Seismic discrimination of nuclear explosions. *Ann. Rev. Earth Planet. Sci.*, **18**, 257–86.

Richter, C. F. 1958. *Elementary Seismology*. New York: Freeman.

Robinson, E. A., and Treitel, S. 1980. *Geophysical Signal Analysis*. Englewood Cliffs, New Jersey: Prentice-Hall.

Scholz, C. H. 1998. Earthquakes and friction laws. *Nature*, **391**, 37–42.

1990. *The Mechanics of Earthquakes and Faulting*. Cambridge: Cambridge University Press.

Shearer, P. M. 1993. Imaging the Earth's upper mantle. *EOS Trans. Am. Geophys. Un*., **74**, 602.

1994. Imaging Earth's response at long periods. *EOS Trans. Am. Geophys. Un.*, **75**, 449.

1999. *Introduction to Seismology*. Cambridge: Cambridge University Press.

Shimazaki, K. 1986. Small and large earthquakes: the effects of the thickness of the seismogenic layer and the free surface. In Das, S., Boatwright, J. and Scholz, C. H., eds., *Earthquake Sources and Mechanics*. Maurice Ewing Series 6. Washington: American Geophysical Union, pp. 209–216.

Simon, R. B. 1981. *Earthquake Interpretations: A Manual for Reading Seismograms*. Los Altos, California: Kaufmann.

Sleep, N. H. and Fujita, K. 1997. *Principles of Geophysics*. Malden, Massachusetts: Blackwell Science.

Slotnick, M. M. 1959. *Lessons in Seismic Computing*. R. A. Geyer, ed. Tulsa, Oklahoma: Society of Exploration Geophysicists.

Stein, S. and Wysession, M. 2003. *Introduction to Seismology, Earthquakes and Earthstructure.* Oxford: Blackwell.

Sykes, L. R. 1967. Mechanism of earthquakes and faulting on the mid-ocean ridges. *J. Geophys. Res.*, **72**, 2131–53.

Sykes, L. R. and Evernden, J. F. 1982. The verification of a Comprehensive Nuclear Test Ban. *Sci. Am.*, **247**, 29–37.

Taner, T. G. and Koehler, F. 1969. Velocity spectra – digital computer derivation and applications of velocity functions. *Geophysics*, **34**, 859–81.

Telford, W. M., Geldart, L. P. and Sheriff, R. E. 1990. *Applied Geophysics*, 2nd edn. New York: Cambridge University Press.

Tooley, R. D., Spencer, T. W. and Sagoci, H. F. 1965. Reflection and transmission of plane compressional waves. *Geophysics*, **30**, 552–70.

Udías, A. 1999. *Principles of Seismology.* Cambridge: Cambridge University Press.

U.S. Geological Survey Staff 1990. The Loma Prieta, California, earthquake: an anticipated event. *Science*, **247**, 286–93.

von dem Borne, G. 1904. *Nachr. Ges. Wiss. Göttingen*, 1–25.

Woods, B. B. and Helmberger, D. V. 1993. A new seismic discriminant for earthquakes and explosions, *EOS Trans. Am. Geophys. Un.*, **74** (8), 91.

Yilmaz, O. 1987. *Seismic Data Processing*. Tulsa, Oklahoma: Society of Exploration Geophysicists.

2001. *Seismic Data Analysis*. Tulsa, Oklahoma: Society of Exploration Geophysicists.

Zoback, M. D. 1980. Recurrent intraplate tectonism in the New Madrid seismic zone. *Science*, **209**, 971–6.

Chapter 5
Gravity

5.1 Introduction

Seismic methods for determining the internal structure of the Earth require the recording and analysis of energy that has passed through the Earth, energy which was produced either by earthquakes or artificially by exploration teams. Gravimetric and magnetic methods are different in that they utilize measurements of existing potential fields, which are physical properties of the Earth itself.

5.2 Gravitational potential and acceleration

Two point masses m_1 and m_2 at distance r apart attract each other with a force F,

$$F = \frac{Gm_1m_2}{r^2} \tag{5.1}$$

where G is the gravitational or Newtonian constant (Fig. 5.1). In the SI system of units, G has a value of $6.67 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$. This inverse square law of gravitational attraction was deduced by Sir Isaac Newton in 1666. The legend is that an apple falling from a tree gave him the revolutionary idea that the force that attracted the apple downwards could also account for the Moon's orbit of the Earth.

The acceleration of the mass m_1 due to the presence of the mass m_2 is Gm_2/r^2 towards m_2 (since acceleration equals force divided by mass), and the acceleration of the mass m_2 is Gm_2/r^2 towards m_1.

The *gravitational potential* V due to mass m_1 can be defined[1] as

$$V = -\frac{Gm_1}{r} \tag{5.2}$$

The gravitational potential energy of a mass m_2 at distance r from mass m_1 is $-Gm_1m_2/r$. If mass m_2 moves away from mass m_1 to a new position, r', the gravitational energy that is released is

$$\frac{Gm_1m_2}{r'} - \frac{Gm_1m_2}{r}$$

[1] The standard convention is to define $V = Gm_1/r$.

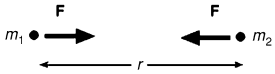


Figure 5.1. Two masses m_1 and m_2, at distance r apart, attract each other with force $F = Gm_1m_2/r^2$.

The *gravitational acceleration*, a, of mass m_2 towards m_1 (with the acceleration taken to be positive in the outward radial direction) can be written in terms of the potential V by using Eq. (5.2):

$$\begin{aligned} a &= -\frac{Gm_1}{r^2} \\ &= -\frac{\partial}{\partial r}\left(-\frac{Gm_1}{r}\right) \\ &= -\frac{\partial V}{\partial r} \end{aligned} \tag{5.3}$$

If we generalize Eq. (5.3) from one to three dimensions, it is written as

$$a = -\text{grad}\, V \tag{5.4}$$

or

$$a = -\nabla V$$

(See Appendix 1 for discussion of grad or ∇.)

If, instead of just one mass m_1, we imagine a distribution of masses, we can then define a potential V as

$$V = -G\sum_i \frac{m_i}{r_i} \tag{5.5}$$

or

$$V = -G\int_m \frac{\mathrm{d}m}{r} \tag{5.6}$$

where (in Eq. (5.5)) each mass m_i is at position r_i and (in Eq. (5.6)) the integral over m is summing all the infinitesimal masses $\mathrm{d}m$, each at its position r. The gravitational acceleration due to either of these distributions is then again given by Eq. (5.4).

We can now use Eq. (5.6) to calculate the potential of a spherical shell (Fig. 5.2). Let us calculate the potential at a point P at distance r from the centre O of the shell. Consider the thin strip of shell, half of which is shown (stippled) in Fig. 5.2. This circular strip has an area of

$$(2\pi b\, \sin\theta)(b\, \mathrm{d}\theta)$$

If we assume the shell to be of thickness t (and very thin) and to be of uniform density ρ, the total mass of the strip is

$$\rho t 2\pi b^2 \sin\theta\, \mathrm{d}\theta$$

Because every point on the strip is the same distance D from point P, Eq. (5.2) gives the potential at P due to the strip as

$$-\frac{G\rho t 2\pi b^2 \sin\theta\, \mathrm{d}\theta}{D}$$

Applying the cosine formula to triangle OQP gives D in terms of r, b and θ as

$$D^2 = r^2 + b^2 - 2br\cos\theta \tag{5.7}$$

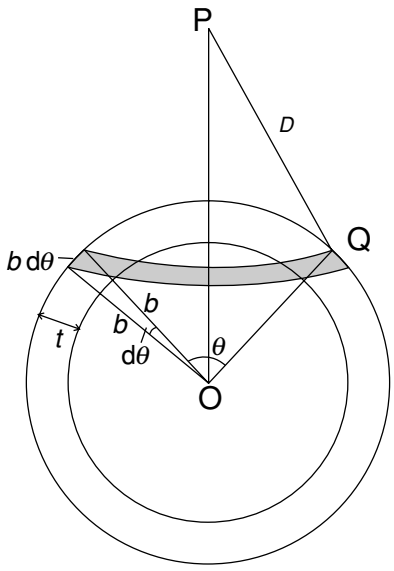


Figure 5.2. A strip of a thin, spherical shell with centre O. P is an arbitrary point at distance r from O.

The potential of the entire spherical shell can now be evaluated from Eq. (5.6) by integrating the potential of the strip:

$$V = -G\rho t 2\pi b^2 \int \frac{\sin\theta \, \mathrm{d}\theta}{(r^2 + b^2 - 2br\cos\theta)^{1/2}} \tag{5.8}$$

To evaluate this integral, we need to change it from an integral over θ to an integral over D. Differentiation of Eq. (5.7) gives

$$D \, \mathrm{d}D = br \sin\theta \, \mathrm{d}\theta \tag{5.9}$$

Equation (5.8) is then written

$$V = -G\rho t 2\pi b^2 \int_{D_{\min}}^{D_{\max}} \frac{\mathrm{d}D}{br} \tag{5.10}$$

To evaluate this integral, it is necessary to consider two situations: (i) when P is outside the spherical shell ($r > b$) and (ii) when P is inside the spherical shell ($r < b$).

(i) When the point P is *external* to the spherical shell, as in Fig. 5.2, the limits $D_{\min}$ and $D_{\max}$ are $r - b$ and $r + b$, respectively, and the potential at point P is

$$\begin{aligned} V &= -G\rho t 2\pi b^2 \left[\frac{D}{br}\right]_{r-b}^{r+b} \\ &= -\frac{G\rho t 4\pi b^2}{r} \end{aligned} \tag{5.11}$$

The total mass of the spherical shell is $4\pi b^2 \rho t$, and so, at the point P the potential of the shell is the same as that due to an equal mass placed at O, the centre of the shell. The gravitational acceleration a due to the spherical shell can be calculated from Eq. (5.4) (in spherical polar coordinates) as

$$a = -\frac{\partial V}{\partial r} = -\frac{G\rho t 4\pi b^2}{r^2} \tag{5.12}$$

This acceleration is the same as the acceleration for the situation when the entire mass of the shell $4\pi b^2 \rho t$ is concentrated at O, the centre of the shell.

(ii) When the point P is *inside* the spherical shell, the limits $D_{\min}$ and $D_{\max}$ in Eq. (5.10) are $b - r$ and $b + r$, respectively. In this case, the potential at P is

$$\begin{aligned} V &= -G\rho t 2\pi b^2 \left[\frac{D}{br}\right]_{b-r}^{b+r} \\ &= -G\rho t 4\pi b \end{aligned} \tag{5.13}$$

This potential is a constant, which is independent of the position of point P inside the shell. The gravitational acceleration, being the negative gradient of the potential, is therefore zero inside the shell.

Using these results for the gravitational potential and acceleration of a spherical shell, we can immediately see that, at any external point, the gravitational potential and the acceleration due to a sphere are the same as those values due to an equal mass placed at the centre of the sphere. In addition, at any point within a sphere the gravitational acceleration is that due to all the material closer to the centre than the point itself. The contribution from all the material outside the point is zero. Imagine the sphere being made of a series of thin uniform spherical shells; then, from Eq. (5.13), we can see that none of the shells surrounding the point makes any contribution to the acceleration. The gravitational acceleration a at a distance r from a sphere of radius b ($b < r$) and density ρ is, therefore,

$$a = -\frac{G\rho \frac{4}{3}\pi b^3}{r^2} = -\frac{GM}{r^2} \tag{5.14}$$

where M is the mass of the sphere. The minus sign in Eqs. (5.12) and (5.14) arises because gravitational acceleration is positive inwards, whereas r is positive outwards. A radial variation of density within the sphere does not affect these results, but any lateral variations within each spherical shell render them invalid.

5.3 Gravity of the Earth

5.3.1 The reference gravity formula

We can apply Eq. (5.14) to the Earth if we assume it to be perfectly spherical. The gravitational acceleration towards the Earth is then given by

$$-a = \frac{GM_{\mathrm{E}}}{r^2} \tag{5.15}$$

where M_{E} is the mass of the Earth. The value of the gravitational acceleration at the surface (denoted by g, which is taken to be positive inwards) of a spherical Earth is, therefore, GM_{E}/r^2, where R is the radius of the Earth. At the Earth's surface, the acceleration due to gravity has a value of about 9.81 m s^{-2}.

The first person to measure the Earth's gravity was Galileo. A celebrated legend is that he conducted his experiments by dropping objects from the top of the leaning tower in Pisa and timing their falls to the ground. (In fact, he slid

objects down inclined planes, which reduced their acceleration from g to $g \sin \delta$, where δ is the dip angle of the plane; thus, they moved more slowly and he could time them more accurately.) In his honour a gravitational unit, the *gal*, was named: $1\ \text{gal} = 10^{-2}\ \text{m s}^{-2}$. The gravitational acceleration at the Earth's surface is therefore about 981 gal. Nowadays gravity can be measured accurately by using vibrating strings, spring balances and lasers (for details see a text such as Telford *et al*. (1990)).

If the Earth were perfectly spherical and not rotating, the gravitational acceleration would have the same value at every point on its surface. However, the Earth is not a perfect sphere (it bulges at the equator and is flattened at the poles) and it is rotating. The Earth's shape can be approximated by an *oblate spheroid*, the surface that is generated by revolving an ellipse about its minor (shorter) axis. The *ellipticity*,[2] or *polar flattening*, f of an ellipse is defined as

$$f = \frac{R_\text{e} - R_\text{p}}{R_\text{e}} \tag{5.16}$$

where R_e and R_p are the equatorial (longer) and polar (shorter) radii, respectively. The oblate spheroid that best approximates the Earth's shape has an ellipticity of 1/298.247. The 'radius' of an oblate spheroid is given, to first order in f, by

$$r = R_\text{e}(1 - f \sin^2 \lambda) \tag{5.17}$$

where f is the ellipticity, λ the latitude and R_e the equatorial radius.

The fact that the Earth is spinning about its axis means that the value of the gravitational acceleration on its surface is reduced. *Centrifugal acceleration* means that the gravitational acceleration on a sphere rotating with angular frequency ω, g_rot, is less than that on a non-rotating sphere, g, and is dependent on latitude, λ:

$$g_\text{rot} = g - \omega^2 R_\text{e} \cos^2 \lambda \tag{5.18}$$

The gravitational acceleration on a rotating oblate spheroid can also be calculated mathematically. The *reference gravity formula* adopted by the International Association of Geodesy in 1967 is

$$g(\lambda) = g_\text{e}(1 + \alpha \sin^2 \lambda + \beta \sin^4 \lambda) \tag{5.19}$$

where the gravitational acceleration at the equator g_e is $9.780\,318\,5\ \text{m s}^{-2}$ and the constants are $\alpha = 5.278\,895 \times 10^{-3}$ and $\beta = 2.3462 \times 10^{-5}$. (There are relations between α and β and between f and ω.) About 40% of this variation of gravity with latitude λ is a result of the difference in shape between the spheroid with the best-fitting ellipticity and a perfect sphere; the remaining 60% of the variation is due to the Earth's rotation. Gravity observations are expressed as deviations from Eq. (5.19).

[2] The ellipticity of an ellipse should not be confused with the *eccentricity, e*, of an ellipse, which is defined as $e = \sqrt{R_\text{e}^2 - R_\text{p}^2} \Big/ R_\text{e}$.

5.3.2 Orbits of satellites

The orbit of a small object about a point mass is an ellipse. Consider a satellite orbiting the Earth, and for simplicity consider a circular orbit with radius r (many satellite orbits are almost circular; a circle is just an ellipse with an ellipticity of zero). The gravitational force of the Earth acting on the satellite is $GM_{\mathrm{E}}m/r^2$ (from Eq. (5.15)), and this is balanced by the outward centrifugal force $m\omega^2 r$ (where ω is the angular velocity and m the mass of the satellite). Thus,

$$\frac{GM_{\mathrm{E}}m}{r^2} = m\omega^2 r \tag{5.20}$$

Rearranging this equation gives the angular velocity ω:

$$\omega = \left(\frac{GM_{\mathrm{E}}}{r^3}\right)^{1/2} \tag{5.21}$$

Alternatively, the period of the satellite orbit, $T = 2\pi/\omega$, is given by

$$T = \left(\frac{4\pi^2 r^3}{GM_{\mathrm{E}}}\right)^{1/2} \tag{5.22}$$

Equations (5.21) and (5.22) are Kepler's third law. Johann Kepler (1571–1630) spent years studying the motions of the planets and discovered that the square of the period of the planet's orbit was proportional to the cube of the orbit's radius. This relationship was explained later when Isaac Newton developed his laws of motion and gravitation, which led to the exact relation derived here and given in Eq. (5.22).

5.4 The shape of the Earth

The Earth is neither a perfect sphere nor a perfect oblate spheroid. Clearly, mountains and deep oceanic trenches are deviations of several kilometres. Geodesists use the surface of the oceans as the reference surface, which is sensible since a liquid surface is necessarily an equipotential[3] (if it were not, the liquid would adjust until the surface became an equipotential). The Earth's reference surface is called the *geoid*. Over the oceans the geoid is the mean sea level, and over the continents it can be visualized as the level at which water would lie if imaginary canals were cut through the continents. All navigation and all surveying are referenced to the geoid. The surveyor's plumb bob, for example, does not point 'down', it points perpendicular to the local equipotential surface which, if not too far above sea level, means perpendicular to the geoid.

[3] An equipotential is a surface over which the potential has a constant value. The gradient of the potential is therefore perpendicular to this surface (see Appendix 1). Thus, gravity ($-\mathrm{grad}\, V$) is always normal to the mean sea-level surface.

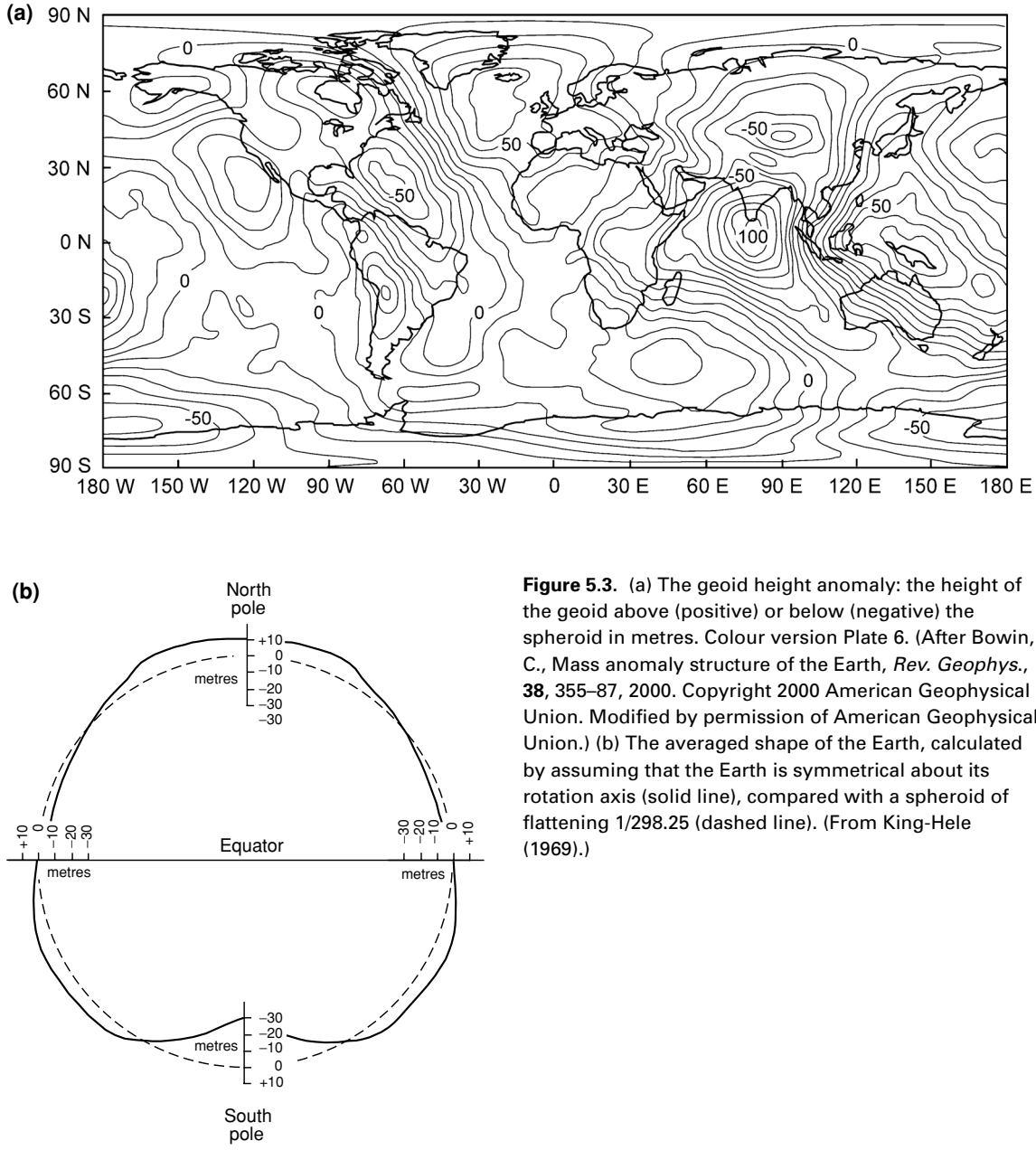


Figure 5.3. (a) The geoid height anomaly: the height of the geoid above (positive) or below (negative) the spheroid in metres. Colour version Plate 6. (After Bowin, C., Mass anomaly structure of the Earth, *Rev. Geophys.*, **38**, 355–87, 2000. Copyright 2000 American Geophysical Union. Modified by permission of American Geophysical Union.) (b) The averaged shape of the Earth, calculated by assuming that the Earth is symmetrical about its rotation axis (solid line), compared with a spheroid of flattening 1/298.25 (dashed line). (From King-Hele (1969).)

The oblate spheroid (discussed in the previous sectionprevious section) approximates the geoid. This reference oblate spheroid, known as the *reference spheroid* or *reference ellipsoid*, is a mathematical figure whose surface is an equipotential of the theoretical gravity field of a symmetrical spheroidal Earth model with realistic radial variations in density, plus the centrifugal potential. The

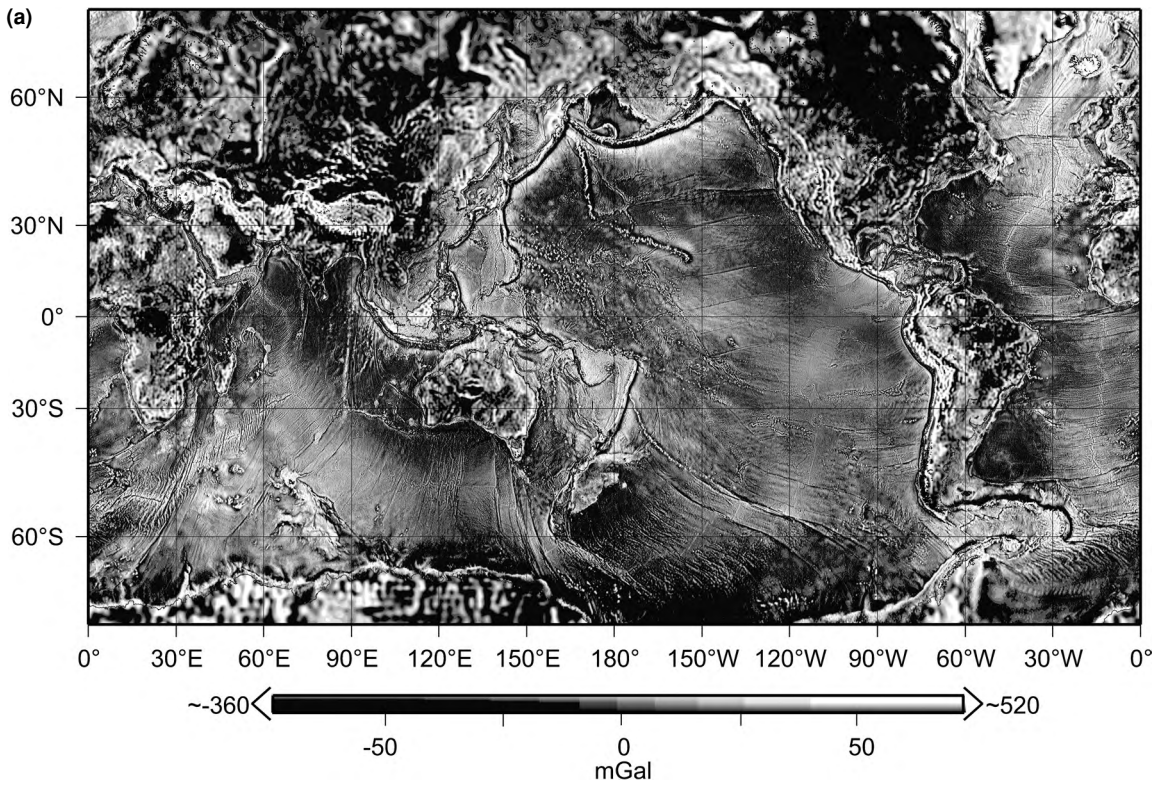


Figure 5.4. (a) A colour-shaded relief image of Earth's gravity anomaly field. (Continents, EGM96; oceans, Smith and Sandwell 1997). Colour version Plate 7. (b) Global topography from 2-min latitude/longitude grid ETOP02 (Smith and Sandwell 1997). Colour version Plate 8.

international gravity formula (Eq. (5.19)) gives the value of g on this spheroid. Figure 5.3(a) shows the deviation of the geoid from the reference spheroid. The largest feature is the 'hole' south of India. A ship sailing across that hole would drop by almost 100 m and then rise again by the same amount, all without doing work against gravity! Figure 5.3(b) shows the 'average' shape of the Earth (it is roughly pear-shaped) compared with the reference spheroid.

The fine details of the shape of the Earth and its gravity field have been determined from artificial satellites (Fig. 5.4). The results have come from studies of small changes in their orbital parameters and from direct radar-altimetry measurements, in which a radar pulse from the satellite is reflected at the ocean surface and its time of arrival back at the satellite measured. Satellites have enabled significant advances in radar altimetry to measure the geoid. GEOS3 and SEASAT, launched in 1975 and 1978, respectively, were so successful that other altimetric satellites

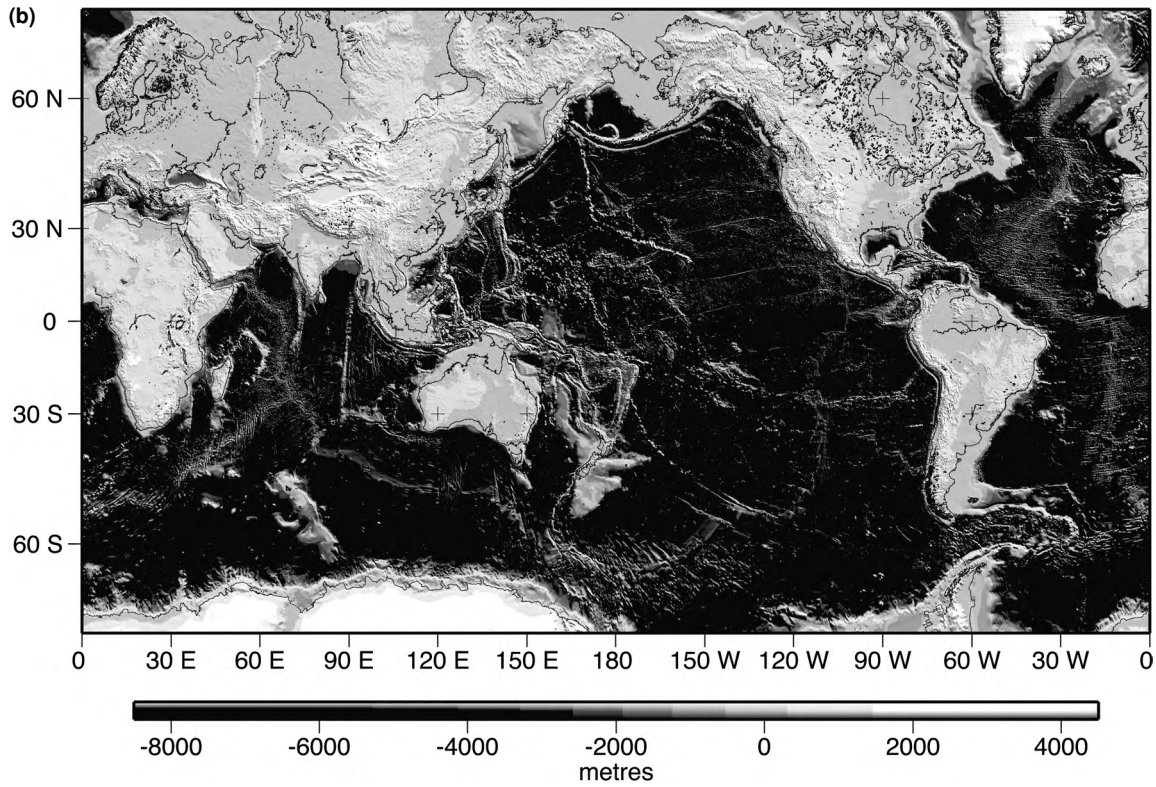


Figure 5.4. (*cont.*)

were launched – GEOSAT in 1985, ERS-1 in 1991, TOPEX/POSEISDON in 1992 and ERS-2 in 1995. This radar altimetry has defined the marine geoid to within better than 10 cm. The gravimetrically defined geoid and the geoid defined by radar altimetry agree well, indicating that the theories and assumptions are correct.

The shape of the geoid can be determined from gravity observations, but, to determine the geoid at a point, we effectively need worldwide gravity observations. Since it is difficult to measure gravity at sea due to accelerations of the ship, prior to the satellite era we had a relatively poor knowledge of the geoid, particularly over the oceans. Within weeks of the launch of the first Earth-orbiting satellite by the U.S.S.R. in 1957, geophysicists had begun to re-evaluate the geoid.

Altimetry satellites, of course, measure the geoid only over the oceans; thus, within months of the launch, we had much more detailed information on the geoid over oceans than over land, a reversal of the previous situation. Satellites now measure the geoid so accurately that we can detect the slow rebound of the

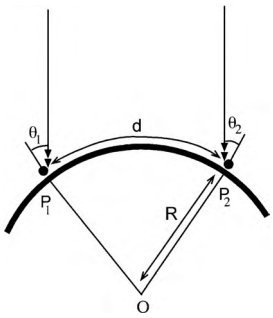


Figure 5.5. A schematic diagram showing use of a plumb line (solid circle with line) to measure the distance between two points on the Earth's surface (heavy arc). Dashed line, light from distant source; *d*, distance between two points P_1 and P_2; *R*, radius of the Earth; and O, centre of the Earth. Geometry gives $d = R(\theta_1 + \theta_2)$. In India, Everest found, on measuring 600 km along a line of longitude, that the value obtained using the plumb-line method was 162 m different (the normal error was ± 5 m). He thought that this difference was due to deflection of the plumb line from the vertical by the attraction of the Himalayas.

Earth to a more spherical shape following the removal of the Pleistocene glaciers. This represents a relative change in moment of inertia of about $10^{-8}\,\mathrm{yr}^{-1}$.

5.5 Gravity anomalies

5.5.1 Introduction

Measurements of the gravitational attraction of the Earth are not only useful for discovering the exact shape of the Earth and its rotational properties; they also provide information about the structure and properties of the lithosphere and mantle. Gravity anomalies are very small compared with the mean surface gravity value of $9.81\,\mathrm{m\,s}^{-2}$ and are therefore often quoted in a more convenient unit, the *milligal*, which is $10^{-5}\,\mathrm{m\,s}^{-2}$ (10^{-3} gal). Another unit that is used is the *gravity unit*, gu, which is $10^{-6}\,\mathrm{m\,s}^{-2}$ ($1\ \mathrm{gu} = 10^{-1}$ mgal). Gravimeters have a sensitivity of about 10^{-5} gal (10^{-2} mgal), or about 10^{-8} of the surface gravitational acceleration. John Milne (1906) gave an interesting example of an anomaly:

> When a squad of 76 men marched to within 16 or 20 feet of the Oxford University Observatory it was found that a horizontal pendulum inside the building measured a deflection in the direction of the advancing load . . .

5.5.2 Isostasy

Between 1735 and 1745 a French expedition under the leadership of Pierre Bouguer made measurements of the length of a degree of latitude in Peru and near Paris in order to determine the shape of the Earth. They realized that the mass of the Andes mountains would attract their plumb line but were surprised that this deflection was much less than they had estimated. In the nineteenth century, the Indian survey under Sir George Everest (after whom the mountain is named) found the same reduced deflection near the Himalayas (Fig. 5.5). Archdeacon J. H. Pratt estimated the deflection of a plumb line by the mountains and found that it should be three times greater than observed. This is now known to be a fairly common attribute of surface features. In 1855, Pratt and Sir George Airy proposed two separate hypotheses to explain these observations; and, in 1889, the term *isostasy* was first used to describe them. Both for the Himalayas and for the Andes, the lack of mass or mass deficiency beneath the mountain chains, which is required in order to account for the reduced deflection of the plumb lines, was found to be approximately equal to the mass of the mountains themselves. This is an alternative statement of Archimedes' principle of hydrostatic equilibrium: a floating body displaces its own weight of water. A mountain chain can therefore be compared with an iceberg or cork floating in water. Thus, isostasy requires the surface layers of the Earth to be rigid and to 'float' on, or in, a denser substratum. The rigid surface layer is termed the *lithosphere* and the region beneath, the *asthenosphere* (Section 2.1); recall that these layers are distinct

from the compositional layers, crust and mantle, which have been discussed previously.

The depth below which all pressures are hydrostatic is termed the compensation depth. At or below the compensation depth, the weight of (imaginary) vertical columns with the same cross-sectional area must be the same. A mountain in *isostatic equilibrium* is therefore *compensated* by a mass deficiency between the surface and the compensation depth. In contrast, an ocean basin in isostatic equilibrium is compensated by extra mass between the seabed and the compensation depth.

Airy's hypothesis

In this hypothesis the rigid upper layer and the substratum are assumed to have constant densities, ρ_u and ρ_s, respectively. Isostatic compensation is achieved by mountains having deep roots (exactly like an iceberg). Figure 5.6(a) illustrates this hypothesis. Taking an arbitrary compensation depth that is deeper than the deepest mountain root in the substratum and equating the masses above that depth in each vertical column of unit cross-sectional area, one obtains

$$\begin{aligned} t\rho_u + r_1\rho_s &= (h_1 + t + r_1)\rho_u \\ &= (h_2 + t + r_2)\rho_u + (r_1 - r_2)\rho_s \\ &= d\rho_\omega + (t - d - r_3)\rho_u + (r_1 + r_3)\rho_s \end{aligned} \tag{5.23}$$

A mountain of height h_1 would therefore have a root r_1 given by

$$r_1 = \frac{h_1\rho_u}{\rho_s - \rho_u} \tag{5.24}$$

Similarly, a feature at a depth d beneath sea level would have an anti-root r_3 given by

$$r_3 = \frac{d(\rho_u - \rho_w)}{\rho_s - \rho_u} \tag{5.25}$$

The rigid upper layer (lithosphere) has density ρ_u, but Eqs. (5.23)–(5.25) apply equally well when ρ_u is replaced by ρ_c (the density of the crust) and ρ_s is replaced by ρ_m (the density of the mantle). This is because the crust–mantle boundary is embedded in and is part of the lithosphere, so loading at the surface and subsequent deflection of the base of the lithosphere deflects the crust–mantle boundary. Furthermore, the difference between the density of the mantle at the crust–mantle boundary and the density of the mantle at the lithosphere–asthenosphere boundary may be very small (see Section 8.1.2). Therefore, Eqs. (5.24) and (5.25) are often applied in the following forms:

$$r_1 = \frac{h_1\rho_c}{\rho_m - \rho_c} \tag{5.26}$$

and

$$r_3 = \frac{d(\rho_c - \rho_w)}{\rho_m - \rho_c} \tag{5.27}$$

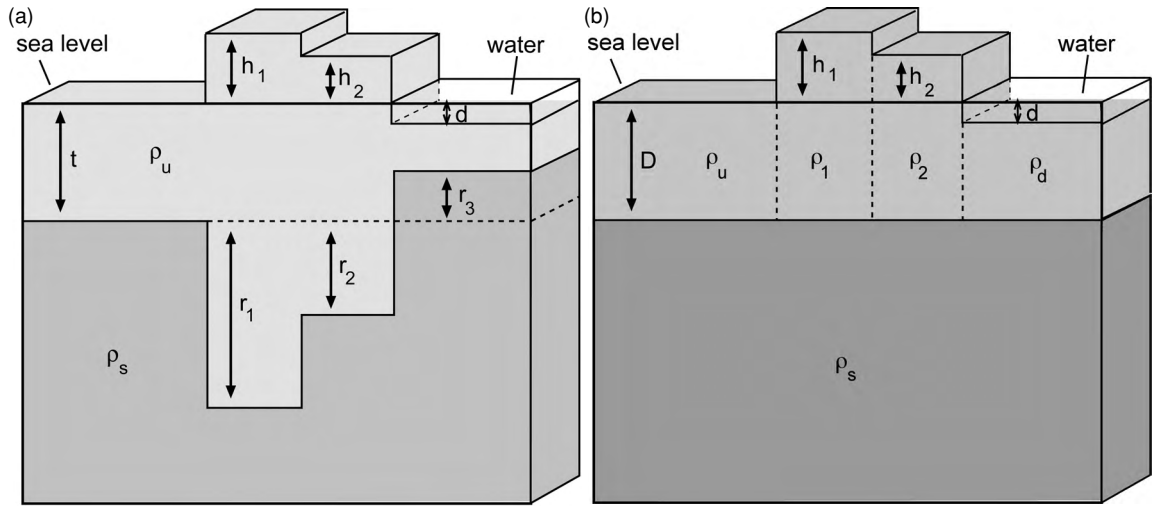


Figure 5.6. (a) An illustration of Airy's hypothesis of isostasy. The upper layer has density ρ_u, and the substratum has density ρ_s. Isostatic compensation is achieved by variation in the thickness of the upper layer: mountains have deep roots and ocean basins have anti-roots. (b) An illustration of Pratt's hypothesis of isostasy. Density of sea water, ρ_w; density of substratum, ρ_s; densities of upper layer beneath mountains of heights h_2 and h_1, ρ_2 and ρ_1; density of upper layer beneath ocean of depth d, ρ_d; density of upper layer beneath land at sea level, ρ_u; and compensation depth, D. Isostatic compensation is achieved by a lateral variation of density beneath surface features. Material beneath mountains is less dense; material beneath ocean basins is more dense.

Pratt's hypothesis

Pratt assumed that the depth of the base of the upper layer is level and that isostatic equilibrium is achieved by allowing this upper layer to be composed of columns of constant density (Fig. 5.6(b)). Taking the base of the upper layer as the compensation depth and equating the masses above this level in each column of unit cross-sectional area gives

$$\begin{aligned} \rho_u D &= (h_1 + D)\rho_1 \\ &= (h_2 + D)\rho_2 \\ &= \rho_w d + \rho_d (D - d) \end{aligned} \tag{5.28}$$

Thus, in this model, compensation is achieved by mountains consisting of and being underlain by material of low density,

$$\rho_1 = \rho_u \left(\frac{D}{h_1 + D} \right) \tag{5.29}$$

and the oceans being underlain by material of higher density,

$$\rho_d = \frac{\rho_u D - \rho_w d}{D - d} \tag{5.30}$$

These two hypotheses of Pratt and Airy are very different; but determining whether one, the other or a combination of them operates in a particular part of the Earth is not a simple matter – compensation can be and is achieved by both methods. Determining whether or not a surface feature is in isostatic equilibrium (compensated) is often easier (Section 5.5.4).

5.5.3 Calculation of gravity anomalies

Before any gravity measurements can be used, certain corrections have to be made. First, allowance must be made for the fact that the Earth is not a perfect sphere but is flattened at the poles and is rotating. The reference gravity formula of 1967 (Eq. (5.19)) includes these effects and expresses gravity g as a function of latitude λ. This enables a correction for the latitude of the measurement point to be made by subtracting the reference value $g(\lambda)$ (Eq. (5.19)) from the actual gravity measurement.

The second correction which must be made to any gravity measurement allows for the fact that the point at which the measurement was made was at an elevation h rather than at sea level on the spheroid. This correction, known as the *free-air correction*, makes no allowance for any material between the measurement point and sea level: it is assumed to be air. Therefore, using the inverse-square law and assuming that the Earth is a perfect sphere, we find that gravity at elevation h is

$$g(h) = g_0\left(\frac{R}{R+h}\right)^2 \tag{5.31}$$

where R is the radius of the Earth and g_0 is gravity at sea level. Since $h \ll R$, Eq. (5.31) can be written as

$$g(h) \simeq g_0\left(1 - \frac{2h}{R}\right) \tag{5.32}$$

The measured value of gravity at height h is therefore less than the value of gravity at sea level. The free-air correction, δg_F, is the amount that has to be added to the measured value to correct it to a sea-level value:

$$\begin{aligned} \delta g_F &= g_0 - g(h) \\ &= \frac{2h}{R} g_0 \end{aligned} \tag{5.33}$$

Since gravity decreases with height above the surface, points above sea level are therefore corrected to sea level by adding $(2h/R)g_0$. This correction amounts to 3.1×10^{-6} m s^{-2} per metre of elevation. A more accurate value for this correction can be obtained by using McCullagh's formula for the gravitational attraction of a rotating spheroid (e.g., Cook 1973, pp. 280–2).

The *free-air anomaly* g_F is then the measured gravity value g_{obs} with these two corrections applied:

$$\begin{aligned} g_F &= g_{obs} - g(\lambda) + \delta g_F \\ &= g_{obs} - g(\lambda)\left(1 - \frac{2h}{R}\right) \end{aligned} \tag{5.34}$$

Two other corrections are frequently applied to gravity measurements.

1. The first is the *Bouguer correction*, which allows for the gravitational attraction of the rocks between the measurement point and sea level, assuming that these rocks are of infinite horizontal extent. The Bouguer correction is given by

$$\delta g_B = 2\pi G\rho h \tag{5.35}$$

   where G is the constant of gravitation, ρ the density of the material between the measurement point and sea level, and h the height of the measurement point above sea level. Taking G to be $6.67 \times 10^{-11}\,m^3\,kg^{-1}\,s^{-2}$ and assuming an average crustal density of $2.7 \times 10^3\,kg\,m^{-3}$, we obtain a Bouguer correction of $1.1 \times 10^{-6}\,m\,s^{-2}$ per metre of elevation. Alternatively, the Bouguer correction for a layer 1 km thick with density $10^3\,kg\,m^{-3}$ is 42 mgal.
2. The final correction is a *terrain correction*, δg_T, which allows for deviations of the surface from an infinite horizontal plane. This correction can be calculated graphically by using a set of templates and a topographic map. The terrain correction is small and, except for areas of mountainous terrain, can often be ignored in crustal studies.

The *Bouguer anomaly* g_B is the free-air anomaly with these two extra corrections applied:

$$\begin{aligned} g_B &= g_F - \delta g_B + \delta g_T \\ &= g_{obs} - g(\lambda) + \delta g_F - \delta g_B + \delta g_T \end{aligned} \tag{5.36}$$

The Bouguer anomaly is therefore the observed value of gravity minus the theoretical value at the latitude and elevation of the observation point. Since we have allowed for the attraction of all the rock above sea level, the Bouguer anomaly represents the gravitational attraction of the material below sea level. The *mantle Bouguer anomaly* is the Bouguer anomaly as calculated above but with an additional correction made for the gravitational attraction of the crust. It thus represents the anomalous attraction of material in the mantle (e.g., Fig. 9.16).

The free-air anomaly is usually used for gravity measurements at sea. It is comparable to the Bouguer anomaly over continents since the measurements are then all corrected to the sea-level datum. If a Bouguer anomaly is required for oceanic gravity measurements, it must be calculated by 'replacing' the sea water with rocks of average crustal density. A terrain correction must then also be applied to account for the seabed topography.

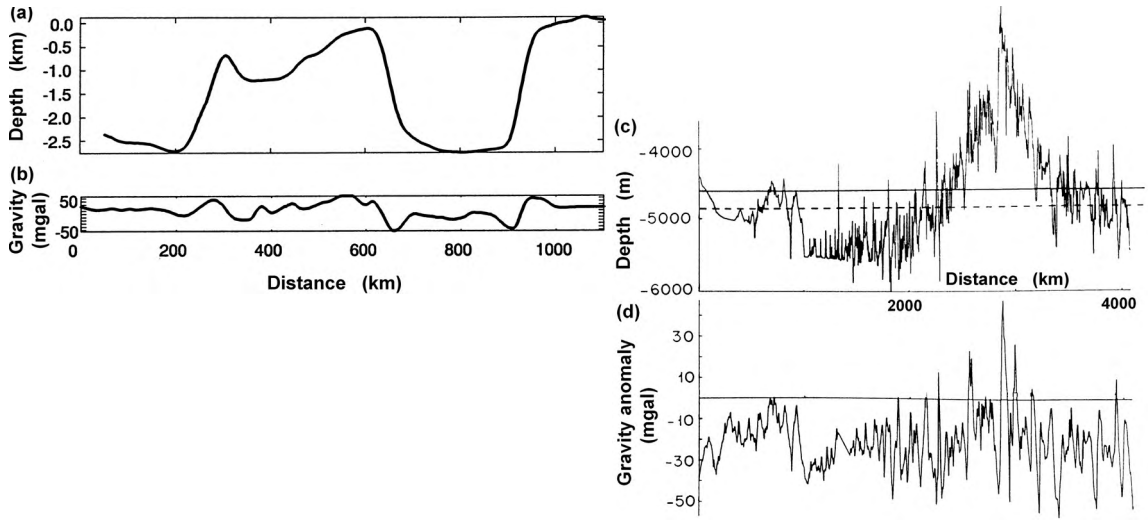


Figure 5.7. Left, profiles across the U.K. continental margin including Hatton Bank, (a) bathymetry and (b) free-air gravity anomaly (D. McKenzie, personal communication 2004); and right, profiles across the Mid-Atlantic Ridge, (c) bathymetry and (d) free-air gravity anomaly. Since the large changes in bathymetry are not accompanied by large changes in the gravity anomaly, both structures must be compensated. (After McKenzie, D. and Bowin, C., The relationship between bathymetry and gravity in the Atlantic Ocean, *J. Geophys. Res.*, **81**, 1903–15, published 1976 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

5.5.4 Use of gravity anomalies to investigate isostasy

Gravity measurements are used to determine whether an area is in isostatic equilibrium. If a region is in isostatic equilibrium, there should be no excess or lack of mass above the compensation depth and hence almost no gravity anomaly.

To a first approximation this can be illustrated by using Eq. (5.35), the Bouguer correction, of 42 mgal for a layer 1 km thick with a density of 1000 kg m^{-3}. Figures 5.7(a) and (b) show the bathymetry and free-air gravity anomaly across part of the U.K. continental margin. The water depth decreases by $\sim$2 km across the margin with water (density 1000 kg m^{-3}) 'replaced' by crust (density $\sim$2700 kg m^{-3}). If this structure were uncompensated then gravity should increase from ocean towards the continent by

$$42 \times 2 \times \frac{2700 - 1000}{1000} \approx 140 \text{ mgal}$$

It is clear from Fig. 5.7(b) that the change in gravity anomaly is much smaller than this, indicating that the structure must be compensated. Figures 5.7(c) and (d) show the Mid-Atlantic Ridge over which there is a change in water depth

of ∼3 km. If the ridge were uncompensated, the gravity anomaly should increase by

$$42 \times 3 \times \frac{2700 - 1000}{1000} \approx 200\,\text{mgal}$$

Since the observed anomaly increases by only∼20 mgal, the ridge must be compensated. In contrast, Figs. 5.13 and 9.64 show the gravity anomaly across Hawaii. An uncompensated decrease in water depth of ∼4 km would cause an increase in gravity of

$$42 \times 4 \times \frac{2700 - 1000}{1000} \approx 300\,\text{mgal}$$

This matches the observed gravity anomaly and so confirms that Hawaii is not compensated. The excess mass must therefore be supported by the Pacific plate upon which it rests (see Section 5.7.1).

The simplest way to determine whether a large-scale structure such as a mountain chain or large sedimentary basin is in isostatic equilibrium is therefore to use the free-air anomaly. If a structure or region is totally compensated, the free-air anomaly is very small away from the edges of the structure, provided that the structure is at least about ten times wider than the compensation depth. If the structure is only partially compensated, or not compensated at all, then the free-air anomaly is positive, perhaps up to several hundred milligals in magnitude, depending on the structure and degree of compensation. For a totally or partially compensated structure the Bouguer anomaly is negative, whereas for an uncompensated structure the Bouguer anomaly is zero. Free-air anomalies are almost isostatic anomalies. They do not imply any specific mechanism for compensation but are small if compensation is complete.

Take, for example, the mountains illustrated in Fig. 5.6(a) and assume them to be in isostatic equilibrium with the left-hand column of crust (of thickness t). The Bouguer anomaly across these mountains will be negative since below sea level there is a mass deficiency beneath the mountains. The attraction of the excess mass of the mountains that balanced this mass deficiency has been removed in calculating this anomaly. In contrast, the free-air anomaly over the mountains is positive and much smaller in magnitude. The free-air anomaly is positive over the mountains rather than being zero, because the mountains are closer to the measurement point than is their deep compensating structure. Even though the mass deficit of the compensating structure is the same as the mass excess of the mountains, its negative contribution to the gravity at the measurement point is smaller because it is further from the measurement point.

Another way to determine whether a structure or region is in isostatic equilibrium is to propose a series of density models and then calculate the Bouguer anomaly that each would give. The isostatic anomaly for the region is then the actual Bouguer anomaly minus the computed Bouguer anomaly for the proposed density model. Thus, each density model for the region has a different isostatic anomaly.

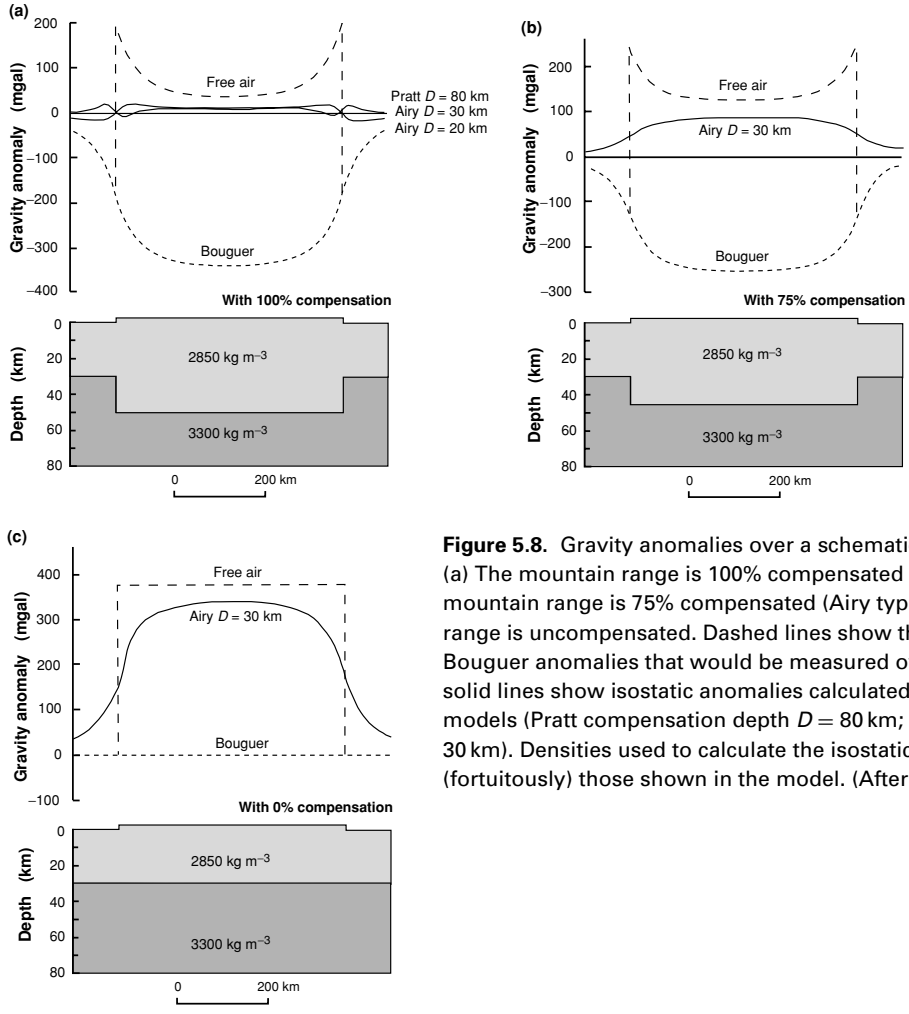


Figure 5.8. Gravity anomalies over a schematic mountain range. (a) The mountain range is 100% compensated (Airy type). (b) The mountain range is 75% compensated (Airy type). (c) The mountain range is uncompensated. Dashed lines show the free-air and Bouguer anomalies that would be measured over the mountain; solid lines show isostatic anomalies calculated for the density models (Pratt compensation depth $D = 80$ km; Airy $D = 20$ km and 30 km). Densities used to calculate the isostatic anomalies are (fortuitously) those shown in the model. (After Bott (1982).)

The effect that isostatic compensation has on the gravity anomalies is illustrated in Fig. 5.8. Figure 5.8(a) shows a schematic wide mountain range that is totally compensated. The Bouguer anomaly across this model is therefore very large and negative, whereas the free-air anomaly is small and positive in the centre of the model and large and positive at the edge of the mountains. Also shown in Fig. 5.8(a) are isostatic anomalies for three possible density models that have been formulated in order to test whether the structure is in isostatic equilibrium. All three isostatic anomalies are very close to zero and the anomaly calculated for Airy-type compensation with $D = 30$ km is exactly zero. The fact that the other two anomalies are almost zero indicates that the structure is in isostatic equilibrium.

Figure 5.8(b) shows the same mountain range, but this time it is only 75% compensated. Now the free-air anomaly is large and positive, whereas the Bouguer

anomaly is large and negative; all the indications are that the mountains are not in isostatic equilibrium. The isostatic anomaly calculated for Airy-type compensation for $D = 30$ km confirms this.

Figure 5.8(c) shows the case in which the mountains are totally uncompensated. Then the Bouguer anomaly is exactly zero since all the excess gravitational attraction is provided by the material above sea level, whereas the free-air anomaly is very large and positive. The isostatic anomaly is, in this case, also very large and positive.

To determine what form the compensation takes, the gravity anomaly must be calculated for a number of possible subsurface density structures and various compensation depths. A zero isostatic anomaly indicates that a correct density distribution and compensation depth have been determined (as in Fig. 5.8(a)). Unfortunately, it is often not possible to distinguish unequivocably amongst the various hypotheses or compensation depths because gravity is insensitive to minor changes in the deep density structure. In addition, small shallow structures can easily hide the effects of the deeper structure. To determine the extent and shape that any compensating mass takes, it is helpful to have additional information on the structure of the crust, such as that given by seismic-refraction and -reflection experiments.

The continents and oceans are in broad isostatic equilibrium. This is mainly achieved by variations in crustal thickness (Airy's hypothesis), although, for example, the mid-ocean ridges are partially compensated by lower-density rocks in the mantle (Pratt's hypothesis). Gabbro, a typical oceanic rock, is denser than granite, a typical continental rock, so Pratt's hypothesis also plays a role in this isostatic balancing of oceanic and continental crust.

Example: isostasy and seismic reflection

The many deep continental reflection lines from around the British Isles produced a result that was initially very surprising. The results of the lines showed that, although the structure of the upper crust varies from sedimentary basins to outcropping crystalline rocks, the clear reflection identified as the Moho is nevertheless generally horizontal and arrives at about 10 s two-way time on unmigrated sections. The reflection typically shows neither any depression (pull down) beneath basins nor elevation (pull up) beneath the crystalline rocks. It is possible that this feature may be just a desire of the eye of the seismic interpreter to find a suitable candidate for the Moho reflection at about this depth, but it is more likely that the feature is real and a consequence of isostasy.

Let us modify the Airy and Pratt isostatic compensation models to include a depth-dependent density $\rho(z)$ (Warner 1987). For the British reflection lines, which were shot in relatively shallow water, Airy-type isostatic compensation (Eq. (5.23)) requires

$$\int_0^{t_1} \rho_1(z)\,\mathrm{d}z = \int_0^{t_2} \rho_2(z)\,\mathrm{d}z + (t_1 - t_2)\rho_\mathrm{m} \tag{5.37}$$

where $\rho_1(z)$ and $\rho_2(z)$ are the density functions for crust with total thickness t_1 and t_2, respectively, and ρ_m is the density of the mantle. When the density and seismic velocities, v_1 and v_2, are related by

$$\begin{aligned} \rho_1(z) &= \rho_m - \frac{k}{v_1(z)} \\ \rho_2(z) &= \rho_m - \frac{k}{v_2(z)} \end{aligned} \tag{5.38}$$

where k is a constant, the two-way normal-incidence travel times T_1 and T_2 for the two structures, given by

$$\begin{aligned} T_1 &= \int_0^{t_1} \frac{2}{v_1(z)}\,\mathrm{d}z \\ T_2 &= \int_0^{t_2} \frac{2}{v_2(z)}\,\mathrm{d}z \end{aligned} \tag{5.39}$$

are equal. Although the density–velocity relationship does not exactly fit Eq. (5.38), values of k between 3×10^6 and $4 \times 10^6\ \mathrm{kg\,m^{-2}\,s^{-1}}$ approximate the Nafe–Drake density–velocity curve (Fig. 4.2(d)).

Pratt-type isostatic compensation (Eq. (5.28)) with a depth-dependent density requires that

$$\int_0^{t_1} \rho_1(z)\,\mathrm{d}z = \int_0^{t_2} \rho_2(z)\,\mathrm{d}z \tag{5.40}$$

In this case, the two two-way travel times (Eqs. (5.39)) are equal if the density and seismic velocity are related by

$$\begin{aligned} \rho_1(z) &= \rho_k - \frac{k}{v_1(z)} \\ \rho_2(z) &= \rho_k - \frac{k}{v_2(z)} \end{aligned} \tag{5.41}$$

where k and ρ_k are arbitrary constants. Thus, the Airy restriction on the density–velocity relationship (Eq. (5.38)) is just a special case of Eq. (5.41).

It is therefore possible that observation of a nearly horizontal Moho on time sections may just be an indication that the observed structures are isostatically compensated. In an isostatically compensated region, if the density–velocity relationship approximates Eq. (5.38) or Eq.(5.41), a structure on the Moho would not be seen on an unmigrated seismic section. Structures on the Moho would, however, be seen after corrections for velocity were made.

5.5.5 Gravity anomalies due to some buried bodies

The gravitational attraction of some simple shapes can be calculated analytically by using Eqs. (5.4) and (5.6). The attraction due to more complex shapes, however, must be calculated numerically by computer. To illustrate the magnitude and type of gravitational anomaly caused by subsurface bodies, we consider a few simple examples.

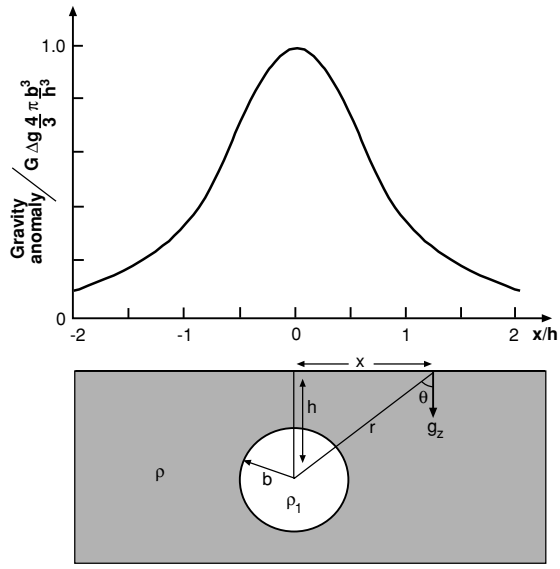


Figure 5.9. A gravity anomaly due to a sphere of radius b buried at a depth h. The density of surrounding medium is ρ; the density of the sphere is ρ_1. The density contrast is $\Delta\rho = \rho_1 - \rho$.

Figure 5.9 shows one problem. A sphere of density ρ_1 and radius b is buried with its centre at depth h in a medium with density ρ. The density contrast of the sphere with respect to the surrounding medium, $\Delta\rho$, is given by

$$\Delta\rho = \rho_1 - \rho \tag{5.42}$$

From the calculations in Section 5.2 we know that the gravitational acceleration g due to a sphere of mass m is Gm/r^2. However, that is the acceleration at point P in the radial direction r, and in this particular case we need to determine the vertical component of gravity, g_z:

$$\begin{aligned} g_z &= g\cos\theta \\ &= \frac{Gm}{r^2}\cos\theta \\ &= \frac{Gm}{r^2}\frac{h}{r} \\ &= \frac{Gmh}{(x^2+h^2)^{3/2}} \end{aligned} \tag{5.43}$$

The gravity anomaly δg_z is therefore given by

$$\delta g_z = \frac{4G\,\Delta\rho\,\pi b^3 h}{3(x^2+h^2)^{3/2}} \tag{5.44}$$

In SI units, Eq. (5.44) is

$$\delta g_z = 2.79\times10^{-10}\frac{\Delta\rho\, b^3 h}{(x^2+h^2)^{3/2}} \tag{5.45}$$

The anomaly due to this buried sphere is therefore symmetrical about the centre of the sphere and essentially confined to a width of about two to three times the depth of the sphere (Fig. 5.9).

This is the simplest gravity anomaly to calculate. All the others involve more tedious algebra (for many detailed examples the reader is referred to Telford *et al*. 1990). Here we merely note that the gravity anomaly for an infinitely long cylinder with anomalous density $\Delta\rho$ and radius b, buried at a depth d beneath the surface, is

$$\delta g_x = \frac{G\,\Delta\rho\,2\pi b^2 d}{x^2 + d^2} \tag{5.46}$$

and the gravity anomaly for a semi-infinite (extending to infinity in the positive x direction) horizontal sheet with anomalous density $\Delta\rho$ and thickness t, buried at depth d beneath the surface, is

$$\delta g_z = 2G\,\Delta\rho\,t\left[\frac{\pi}{2} + \tan^{-1}\left(\frac{x}{d}\right)\right] \tag{5.47}$$

Each particular buried body gives rise to its own anomaly. In many cases, the shape of the anomalous body can be determined from the shape of the gravity anomaly (e.g., the gravity anomaly due to a sphere is narrower than that due to an infinite horizontal cylinder and not stretched out in the y direction). Gravity models are, however, unfortunately *not* unique.

These three examples illustrate the fact that, to determine a shallow density structure, one must perform a detailed local survey. Anomalies due to bodies buried at depth d can be detected only at distances out to perhaps $2d$ from the body. To resolve details of the density structure of the lower crust (at say 20–40 km), gravity measurements must be made over an extensive area, because the widths of these anomalies are much greater (say, five to ten times greater) than those of the anomalies due to the shallow crustal structure. Likewise, gravity anomalies due to anomalous bodies in the mantle are of a much longer wavelength (hundreds of kilometres). Thus, although at first sight it might seem impossible to extract mantle density information from surface gravity measurements (which cannot fail to be affected by near-surface density anomalies), application of a wavelength filter or smoothing of the gravity data does allow such deep structures to be studied. The opposite is, of course, also true: to study the shallow structure, one must remove the regional long-wavelength anomaly which is of deeper origin. However, no amount of processing can always ensure a complete separation.

5.6 Observed gravity and geoid anomalies

5.6.1 Gravity anomalies

A gravity profile across the Mid-Atlantic Ridge is shown in Fig. 9.11. The maximum free-air anomaly on this profile is about 100 mgal. The four density models in Fig. 9.11 yield anomalies that adequately match the measurements, despite being very different. Note that, as is to be expected, model (c), the deepest model,

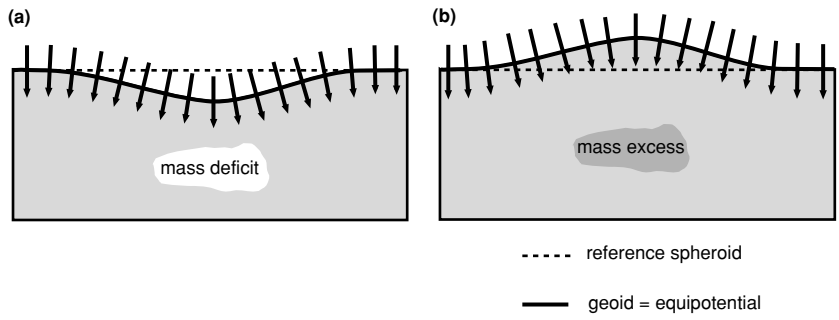


Figure 5.10. (a) A trough in the geoid, or negative geoid height anomaly, occurs over a region of mass deficit (such as as a depression in the seabed). A negative free-air gravity anomaly also occurs over such a mass deficit. (b) A bulge in the geoid, or positive geoid height anomaly, occurs over regions of excess mass (such an elevated region of the seabed). A positive free-air gravity anomaly also occurs over such a mass excess.

has only a very small, 40 kg m^{-3}, density anomaly, whereas the shallower models (e), (f) and (g) have density anomalies of 300 kg m^{-3} and more. Gravity profiles across oceanic trenches (Figs. 5.4, 9.57, 9.58 and 9.59) show very large magnitude anomalies. Gravity measurements over the continents and their implications are discussed in Chapter 10.

5.6.2 Geoid height anomalies

The lateral variations in density distribution within the Earth, although resulting in gravity anomalies, also result in deviations of the geoid from the theoretical spheroid. The *geoid height anomaly* Δh (geoid radius minus spheroid radius) is directly related to the anomaly in the gravitational potential ΔV as measured on the spheroid:

$$g\,\Delta h = -\Delta V \tag{5.48}$$

where g is the gravitational acceleration due to the spheroid (Eq. (5.19)). Since the gravitational acceleration is normal to the equipotential surface (which is by definition the geoid), a trough in the geoid is present wherever there is a negative gravity anomaly (mass deficit), and likewise there is a bulge in the geoid wherever there is a positive gravity anomaly (mass excess) (Fig. 5.10). Thus, a trough in the geoid, or negative geoid height anomaly, is the result of a negative gravity anomaly or a positive potential anomaly; and a bulge in the geoid, or positive geoid height anomaly, is the result of a positive gravity anomaly or a negative potential anomaly.

Overall, the deviations from the geoid are on a small scale, indicating that isostatic equilibrium is a normal attribute of large-scale surface features. In turn, this indicates that the mantle is, on a long timescale, not particularly strong; it must have a finite viscosity for it to be able to flow and adjust to changing surface loads. In contrast, we find that small-scale surface topography need not be compensated.

This observation is explained by the notion of the strong lithosphere, which is able to support small mass anomalies but bends or flexes under very-large-scale mass anomalies, causing flow and readjustment in the weaker underlying asthenosphere. Erosion also occurs. Isostatic balance is thus achieved.

The geoid height anomaly resulting from an isostatic density distribution is not zero and can be calculated. It can be shown (see Turcotte and Schubert 2002, p. 217) that the geoid height anomaly at any point P is given by

$$\Delta h = -\frac{2\pi G}{g}\int_0^D \Delta\rho(z)\, z\, \mathrm{d}z \tag{5.49}$$

where y is the reference gravity value, $\Delta\rho(z)$ the anomalous density at depth z beneath point P and D the compensation depth. Depth z is measured positively downwards with $z = 0$ corresponding to the spheroid. Equation (5.49) therefore gives the geoid height anomaly due to long-wavelength isostatic density anomalies. Geoid anomalies can be used to estimate the variation of density with depth. In practice, we need to work in reverse, by first calculating the geoid height anomaly for an isostatic density model and then comparing the calculations with the observed measurements.

As an example, consider the Airy compensation model illustrated in Fig. 5.6(a). The reference structure is an upper layer of density ρ_u and thickness t and a substratum of density ρ_s. All density anomalies are *with respect to this reference structure*. The geoid height anomaly over a mountain range of height h_1, calculated by using Eq. (5.49), is

$$\begin{aligned}\Delta h &= -\frac{2\pi G}{g}\left(\int_{-h_1}^{0}\rho_\mathrm{u} z\, \mathrm{d}z + \int_t^{t+r_1}(\rho_\mathrm{u}-\rho_\mathrm{s})\, z\, \mathrm{d}z\right)\\ &= -\frac{\pi G}{g}\left[-h_1^2\rho_\mathrm{u} + (\rho_\mathrm{u}-\rho_\mathrm{s})\left(2tr_1 + r_1^2\right)\right]\end{aligned} \tag{5.50}$$

After substituting for r_1 from Eq. (5.24) and rearranging terms, we finally obtain

$$\Delta h = \frac{\pi G}{g}\rho_\mathrm{u} h_1\left(2t + \frac{\rho_\mathrm{s} h_1}{\rho_\mathrm{s}-\rho_\mathrm{u}}\right) \tag{5.51}$$

Therefore, for crustal and mantle densities of 2.8×10^3 and 3.3×10^3 kg m^{-3}, respectively, and a reference crust 35 km thick, the geoid height anomaly is

$$\Delta h \simeq 6h_1(0.7 + 0.066h_1)\ \mathrm{m} \tag{5.52}$$

where h_1 is in kilometres. Thus, a compensated mountain range 3 km high would result in a positive geoid height anomaly of about 16 m.

Likewise, the geoid height anomaly for an Airy-compensated ocean basin of depth d (Fig. 5.6(a)) is

$$\begin{aligned}\Delta h &= -\frac{2\pi G}{g}\left(\int_0^d (\rho_w - \rho_u)z \, dz + \int_{t-r_3}^{t} (\rho_s - \rho_u)z \, dz\right) \\ &= -\frac{\pi G}{g}\left[(\rho_w - \rho_u)d^2 + (\rho_s - \rho_u)\left(2tr_3 - r_3^2\right)\right] \end{aligned} \tag{5.53}$$

On substituting for r_3 from Eq. (5.25) and rearranging terms, we obtain

$$\Delta h = -\frac{\pi G}{g}(\rho_u - \rho_w)\, d\left[2t - d\left(\frac{\rho_s - \rho_w}{\rho_s - \rho_u}\right)\right] \tag{5.54}$$

Thus, with the numerical values given for Eq. (5.52), the geoid height anomaly is

$$\Delta h \simeq 3.85d(0.7 - 0.046d) \text{ m} \tag{5.55}$$

where d is the ocean depth in kilometres. A compensated ocean basin 5 km deep would result in a negative geoid height anomaly of about 9 m.

Geoid height anomalies calculated for the Pratt compensation model are larger than those for the Airy model. The difference is particularly significant for seabed topography: the Pratt model gives a geoid height anomaly about twice that of the Airy model. This difference can be used to estimate the type and depth of compensation for major features.

In Chapter 8, the topic of convection in the mantle is discussed in some detail. Here, we assume that this process occurs and see what can be discovered about it by using measurements of the geoid and the Earth's gravity field.

The complex flow of material within the mantle gives rise to gravity and geoid height anomalies just as mountains and oceanic trenches do. Figure 5.11 illustrates the two contributions to the gravity anomaly due to a region of hot rising mantle. The surface is elevated over the hot material and depressed over cold regions. There is a negative contribution to the gravity anomaly from the hot mantle due to its reduced density. However, there is also a positive contribution to gravity due to the elevation of the surface. Since this positive contribution is larger than the negative contribution due to the reduced density, overall there is a positive gravity anomaly over the hot region. Figure 5.12 shows a numerical simulation of convection in the mantle (refer to Section 8.2 for details of these models). The columns of rising material are hotter and therefore less dense than the columns of sinking mantle material. Calculations indicate that there are small, positive gravity seabed and geoid height anomalies above the rising hot regions. These anomalies occur because the deflection of the surface by the rising current produces a larger anomaly than the negative anomaly which results from the density deficit. Similarly, there are negative anomalies above the sinking columns. Detailed studies of long-wavelength gravity, geoid and bathymetric

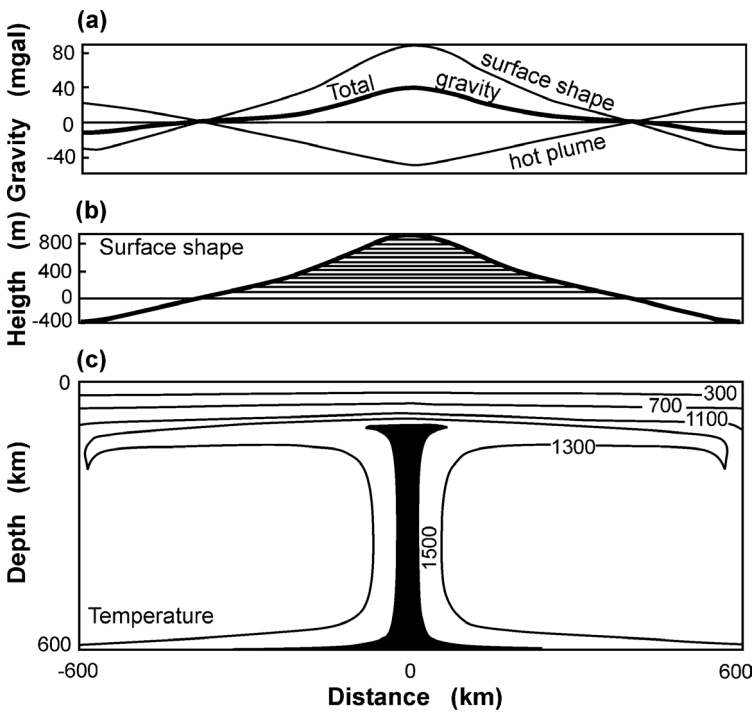


Figure 5.11. Contributions to the gravity anomaly over a hot region in the upper mantle: (a) gravity anomaly, (b) surface shape and (c) mantle temperature (°C). (From McKenzie (1998).)

anomalies can therefore give insights into the planform of convection in the mantle. Comparision of the geoid with bathymetry for the Pacific does not reveal linear anomalies, indicating that the convection cannot be simple rolls aligned with the direction of plate motion (e.g., Fig. 8.15). The planform of the convection seems to be three-dimensional and on a smaller scale than the width of the Pacific plate.

Figure 5.13 shows the bathymetry and the free-air gravity anomaly north and south of the Hawaiian island of Oahu. The central 200-km-wide Hawaiian ridge 'mountain' whose pinnacle is Oahu acts as a load on the Pacific plate and bends it, resulting in the depression on either side of the mountain. The bending of the plate and the associated gravity anomaly are confined to distances within about 200 km of the load. The plates are too thin and too flexible for local loads to result in any long-wavelength bathymetric or gravity anomalies. The long-wavelength bulge evident both in the bathymetry and in the gravity is, however, thought to be the surface expression of a hot upwelling region in the mantle as discussed previously. Again, the fortunate difference in wavelength between the elastic deformation of the plate in response to a surface load and the apparent wavelength of mantle convection allows these two deviations to be distinguished.

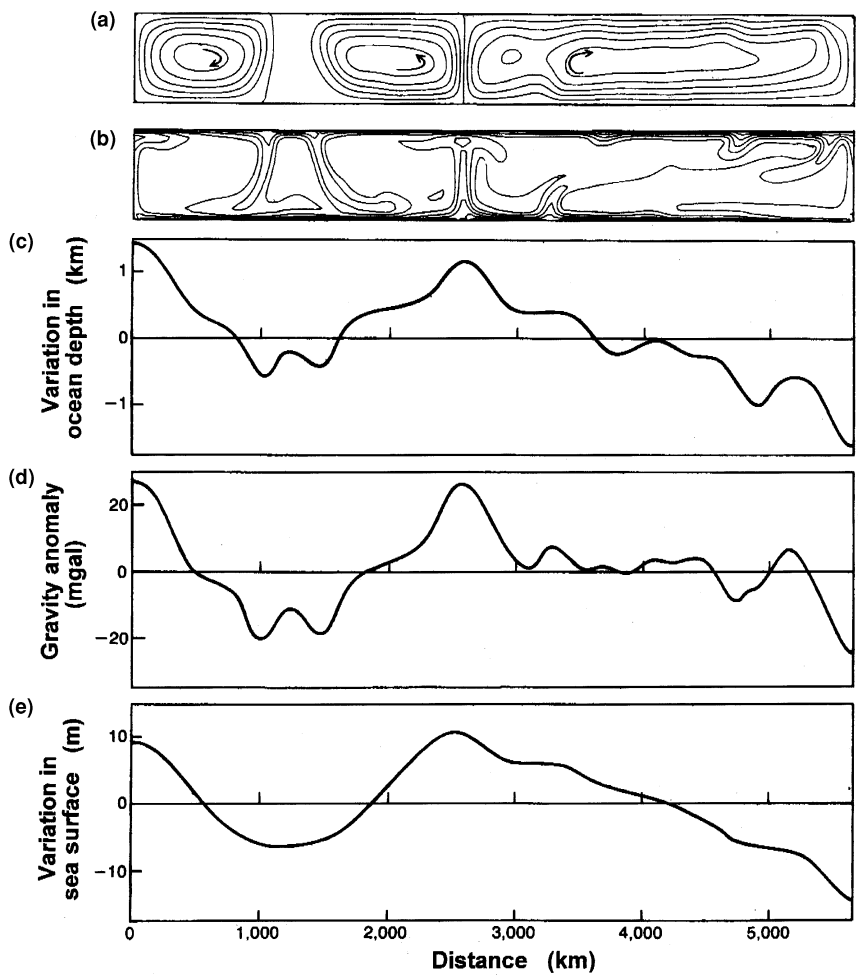


Figure 5.12. Computer modelling of convection in the upper mantle. This model assumes that the mantle has constant viscosity and is heated from below and cooled from above. (a) Circulation of the mantle material in convection cells. (b) Temperature in the model mantle contoured at 100-°C intervals. There are two rising limbs at the left-hand edge and the centre and two regions of sinking material at centre left and right. (c) Variation in the depth of the ocean caused by the convection. (d) Variation in the gravitational acceleration (gravity anomaly) caused by the convection. (e) Variation in the height of the sea surface (geoid height anomaly) caused by the convection. (From McKenzie *et al.* (1980).)

5.7 Flexure of the lithosphere and the viscosity of the mantle

5.7.1 The lithosphere as an elastic plate

In the theory of plate tectonics the thin lithospheric plates are assumed to be rigid and to float on the underlying mantle. On a geological timescale the lithosphere behaves elastically and the mantle behaves as a viscous fluid, whereas on the very short seismic timescale both behave as elastic solids. The study of the bending or flexure of the lithosphere which results from its loading by mountain chains, volcanoes and so on enables us to estimate the elastic properties of the lithosphere. Additionally, the rate of recovery or rebound which occurs when a load is removed is dependent on the viscosity of the underlying mantle as well as

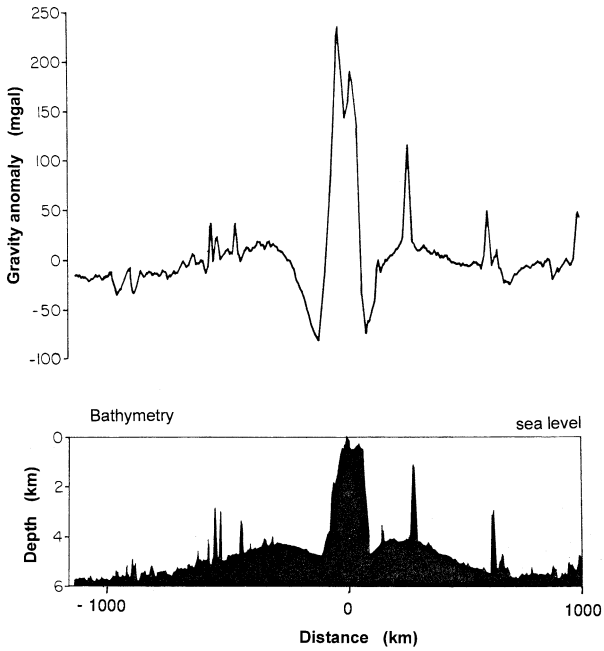


Figure 5.13. Free-air gravity and bathymetry anomalies along a north–south line centred on the Hawaiian island of Oahu. The volcanic islands act as a load on the Pacific plate, bending it downwards and resulting in the symmetrical trough both in the bathymetry and in the gravity anomaly. (After Watts and Daly (1981). Reproduced with permission from the *Annual Review and Planetary Sciences*, vol. 9, © 1981 by Annual Reviews Inc.)

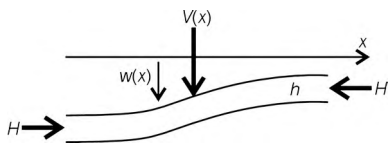


Figure 5.14. A thin plate of thickness h is deflected by $w(x)$ as a result of an imposed variable vertical force per unit area $V(x)$ and a constant horizontal force H per unit length.

the elastic properties of the lithosphere. Thus, given suitable loads, we can make estimates of mantle viscosity.

The general fourth-order differential equation governing the equilibrium deflection of an elastic plate as a function of horizontal distance x is well known in engineering:

$$D\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} = V(x) - H\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} \tag{5.56}$$

where $w(x)$ is the deflection of the plate, $V(x)$ a vertical force per unit length applied to the plate, H a constant horizontal force per unit length applied to the plate and D the *flexural rigidity* of the plate (Fig. 5.14). The flexural rigidity is defined by

$$D = \frac{Eh^3}{12(1-\sigma^2)} \tag{5.57}$$

where E is Young's modulus (see Appendix 2), h the thickness of the plate and σ Poisson's ratio (see Appendix 2).

Consider the oceanic lithosphere deforming under an applied vertical load $V(x)$ and no horizontal force. Water fills the resulting depression in the seabed. However, there is a net hydrostatic restoring force of $(\rho_m - \rho_w)gw$ per unit area. This restoring force acts because the deformed lithosphere is not in isostatic equilibrium: a thickness w of mantle with density ρ_m has effectively been replaced by water with density ρ_w. Thus, for the oceanic lithosphere, Eq. (5.56) is

$$D\frac{d^4w}{dx^4} = V(x) - (\rho_m - \rho_w)gw \tag{5.58}$$

In the case of the deformation of continental lithosphere, when the depression is filled with sediment, the hydrostatic restoring force is $(\rho_m - \rho_c)gw$, since mantle with density ρ_m has been replaced by crust with density ρ_c. For continental lithosphere Eq. (5.56) is

$$D\frac{d^4w}{dx^4} = V(x) - (\rho_m - \rho_c)gw \tag{5.59}$$

These differential equations must be solved for given loads and boundary conditions to give the deflection of the plate as a function of horizontal distance. In the particular case in which the load is an island chain (assumed to be at $x = 0$), Eq. 5.58 is

$$D\frac{d^4w}{dx^4} + (\rho_m - \rho_w)gw = 0 \tag{5.60}$$

The solution to this equation for a line load V at $x = 0$ is

$$w(x) = w_0 e^{-x/\alpha}[\cos(x/\alpha) + \sin(x/\alpha)], \quad x \geq 0 \tag{5.61}$$

where the deflection at $x = 0$ is

$$w_0 = \frac{V\alpha^3}{8D} \tag{5.62}$$

and the parameter α, called the *flexural parameter*, is given by

$$\alpha = \left(\frac{4D}{(\rho_m - \rho_w)g}\right)^{1/4} \tag{5.63}$$

Figure 5.15 shows the deflection given by Eq. (5.61) as a function of x. Notice the clear arch, or forebulge, on either side of the central depression. At the top

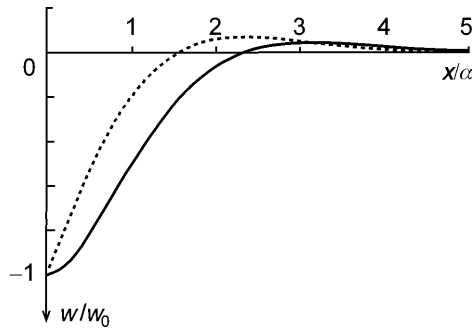


Figure 5.15. Deflection of an elastic plate by a line load at $x = 0$. The deflection is normalized with respect to w_0, the deflection at $x = 0$, which is determined by the load and the physical properties of the plate. Deflection is symmetrical about $x = 0$. Solid line, intact plate; dashed line, plate broken at $x = 0$.

Table 5.1 *Deflection on an elastic plate by a line load V at* $x = 0$

	Intact plate	Broken plate
w_0, deflection at $x = 0$	$\frac{V\alpha^3}{8D}$	$\frac{V\alpha^3}{4D}$
w_b, deflection of forebulge	$-0.0432w_0$	$-0.0670w_0$
$x_b\left(\frac{dw(x)}{dx} = 0 \text{ at } x = x_b\right)$	$\pi\alpha$	$\frac{3\pi\alpha}{4}$
$x_0\ (w(x) = 0 \text{ at } x = x_0)$	$\frac{3\pi\alpha}{4}$	$\frac{\pi\alpha}{2}$

of the forebulge $dw/dx = 0$, so, by differentiating Eq. (5.61), the horizontal distance from the origin ($x = 0$) to the top of the forebulge, $x = x_b$, can be calculated. Once x_b is known, Eq. (5.61) is used to obtain the height of the forebulge $w(x_b)$. The value x_b provides an estimate for the half-width of the depression. An alternative smaller estimate may be provided by finding x_0, the value of x for which $w(x_0) = 0$ (Table 5.1). It is important to note that the magnitude of the load, V, controls only the depth of the depression, not the width. The width of the depression depends upon the flexural rigidity of the plate. This fact provides a way to use the widths of basins to calculate the elastic thicknesses of the plates upon which they lie. Three steps are required:

1. the half-width, or width, of the depression (x_0, x_b) is used to obtain a value for α, the flexural parameter (Table 5.1, Fig. 5.15);
2. the value for α is used to calculate the flexural rigidity D (Eq. (5.63)); and
3. the value for D is used to calculate the elastic thickness h (Eq. (5.57)).

Consider, for example, a depression with a half-width of 150 km. Using Fig. 5.15 and Table 5.1, we can estimate the flexural parameter α as 64 km. Then, with Eqs. (5.63) and (5.57), the elastic thickness of the lithosphere is calculated as 24 km when $\rho_m = 3.3 \times 10^3\,\mathrm{kg\,m^{-3}}$, $\rho_w = 1.0 \times 10^3\,\mathrm{kg\,m^{-3}}$, $g = 10\,\mathrm{m\,s^{-2}}$, $D = 9.6 \times 10^{22}\,\mathrm{N\,m}$, $E = 70\,\mathrm{GPa}$ and $\sigma = 0.25$.

If the elastic plate is assumed to be broken at $x = 0$ rather than being intact, the solution to Eq. (5.60) is $w(x) = w_0 e^{-x/\alpha}\cos(x/\alpha)$, with $w_0 = V\alpha^3/(4D)$. This deflection of a broken plate is narrower and deeper than the deflection of an intact plate with the same rigidity (Table 5.1).

Similar but more complex analyses have been used to estimate the elastic thickness of the Pacific plate under the Hawaiian–Emperor island chain (Fig. 5.13). (The problem is more complex than the simple solution given here because the island chain has a finite width and so cannot be treated as a line force acting at $x = 0$. Also, the age of the Pacific plate and therefore its thickness change along the length of the island chain.)

The bending of the oceanic lithosphere at a subduction zone can also be modelled by Eq. (5.56). In this case, it is necessary to include a load V at one end ($x = 0$) of the plate and a horizontal bending moment M per unit length. The

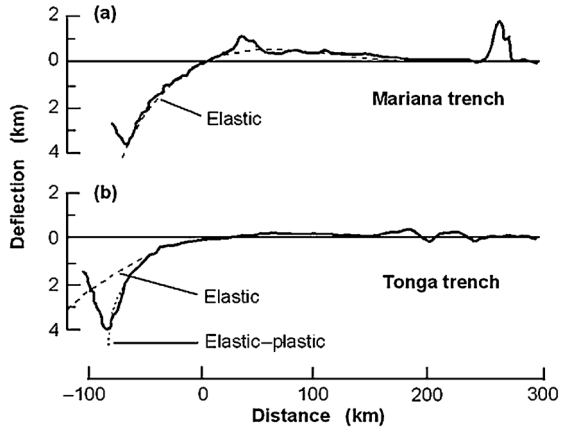


Figure 5.16. Observed (solid line) and theoretical (dashed line) bathymetric profiles across two subduction zones. (a) The bathymetry of the Mariana Trench is fitted by the flexure of a 28-km-thick elastic plate. (b) The bathymetry of the Tonga Trench is not fitted well by flexure of an elastic plate but is fitted by flexure of an elastic and then perfectly plastic plate 32 km thick with a yield stress of 1000 MPa. (After Turcotte *et al.* (1978).)

deflection of the plate is then given by

$$w(x) = \frac{\alpha^2}{2D} \exp\left(-\frac{x}{\alpha}\right)\left[-M \sin\left(\frac{x}{\alpha}\right) + (V\alpha + M)\cos\left(\frac{x}{\alpha}\right)\right], \quad x \geq 0 \qquad (5.64)$$

This topographic profile has, like the island-chain profile, a pronounced forebulge, as is observed for subduction zones. The parameters M and V cannot be reliably estimated in practice, but fortunately the width and height of the forebulge can be, and these can be used to calculate α and hence the elastic thickness of the lithosphere. Figure 5.16(a) shows a topographic profile across the Mariana Trench and a theoretical profile for $\alpha = 70$ km. The two profiles are clearly in good agreement. Using Eqs. (5.63) and (5.57) and the same values as in the previous example, the elastic thickness of the subducting oceanic plate is then 28 km.

Although the topography of many trenches can be explained simply as the bending of an elastic plate, there are exceptions. The Tonga Trench, for example, bends more steeply than can be explained by an elastic model of the plate. However, if the plate is assumed to behave elastically up to some critical yield stress (e.g., 1000 MPa) and then to behave perfectly plastically[4] above that stress, the observed deformation of the Tonga Trench can be explained. Figure 5.16(b) shows observed and theoretical topographic profiles across the Tonga Trench,

[4] If a material deforms plastically (or anelastically), the deformation is irreversible, unlike elastic deformation, which is reversible.

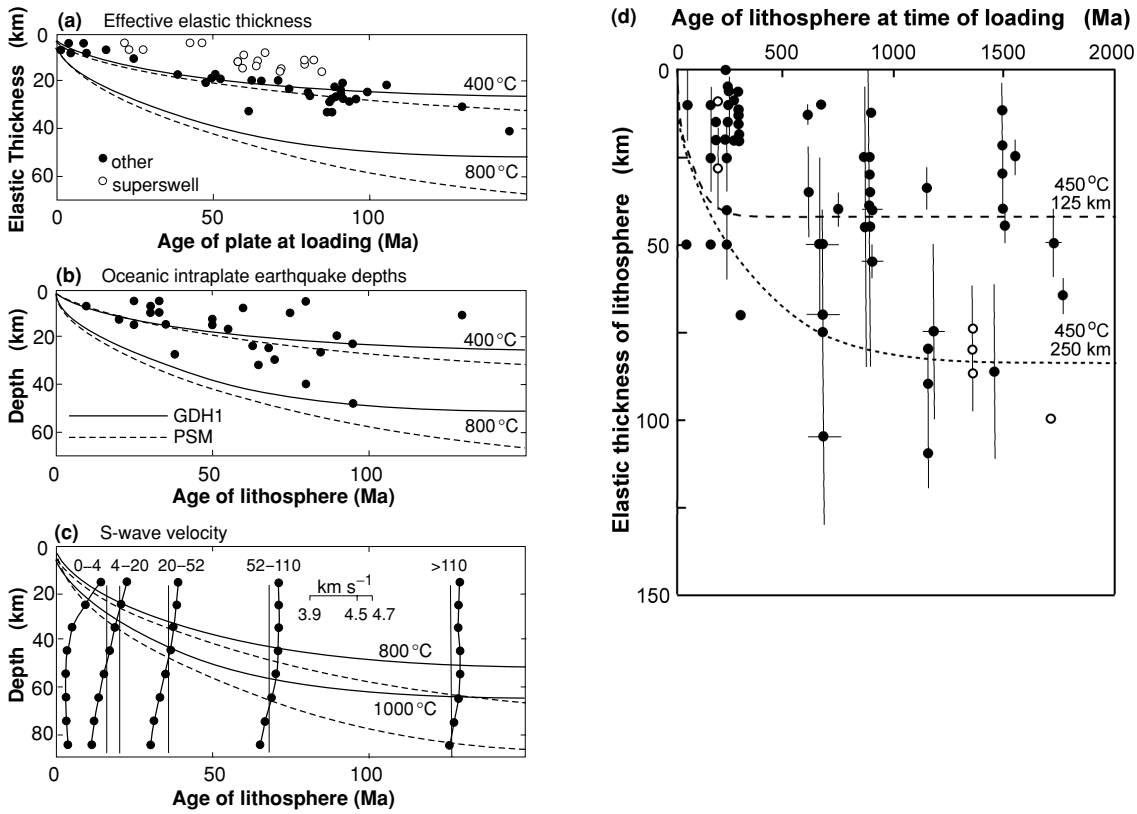


Figure 5.17. The thermal structure of the oceanic lithosphere is consistent with other geophysical data. Isotherms for two different plate models (Section 7.5.2) are shown: PSM (dashed line) and GDH1 (solid line). (a) The effective elastic thickness of the oceanic lithosphere, estimated from the way the lithosphere supports loads, increases approximately as the 400-°C isotherm. Open circles, normal oceanic lithosphere; filled circles, oceanic lithosphere in the vicinity of superswells. (b) The maximum depth of intraplate (within-plate) seismicity also increases with age. (c) The shear-wave velocity structure of the oceanic lithosphere as determined from surface wave dispersion in five age ranges. The vertical line corresponds to 4.5 km s^{-1}. The depth to the low-velocity zone increases with lithospheric age. (After Stein and Stein, Thermomechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes (overview), Geophysical Monograph 96, 1–17, 1996. Copyright 1996 American Geophysical Union. Reproduced by permission of American Geophysical Union.) (d) The elastic thickness of the continental lithosphere plotted against the age of the lithosphere at the time of loading. Dashed lines show 450-°C isotherms for lithosphere thickness 125 and 250 km. Open circles, glacial loading, black circles, foreland basins. (Redrawn from Watts, *Isostasy and Flexure of the Lithosphere*, 2001, Cambridge University Press.)

which indicate that an elastic and perfectly plastic lithosphere is the more appropriate model there.

The values for the elastic thickness of the oceanic plates determined in this section are considerably less than the values determined from seismic and thermal data (Sections 4.1.3 and 7.5). A 100-km-thick elastic plate could not deform as the oceanic lithosphere is observed to bend, bending would be much more gradual. These apparent contradictions are a consequence of the thermal structure of the lithosphere (discussed in Section 7.5). Figure 5.17 shows that the long-term elastic thickness of the Pacific plate apparently increases with age and approximately corresponds to the 450-°C isotherm. The elastic thickness of the continental lithosphere is frequently considerably greater than that of the oceanic lithosphere; however, there is a wide range of values and there is no simple relationship between elastic thickness and age of the lithosphere.

5.7.2 Isostatic rebound

The examples just discussed assume an equilibrium situation in which the load has been in place for a long time and deformation has occurred. A study of the rate of deformation after the application or removal of a load, however, shows that the rate of deformation is dependent both on the flexural rigidity of the lithosphere and on the viscosity[5] of the mantle. Mountain building and subsequent erosion can be so slow that the viscosity of the underlying mantle is not important; the mantle can be assumed to be in equilibrium at all times. However, the ice caps – which during the late Pleistocene covered much of Greenland, northern North America and Scandinavia – provide loads of both the right magnitude and the right age to enable the viscosity of the mantle to be estimated. Figure 5.18 illustrates the deformation and rebound which occur as the lithosphere is first loaded and then unloaded.

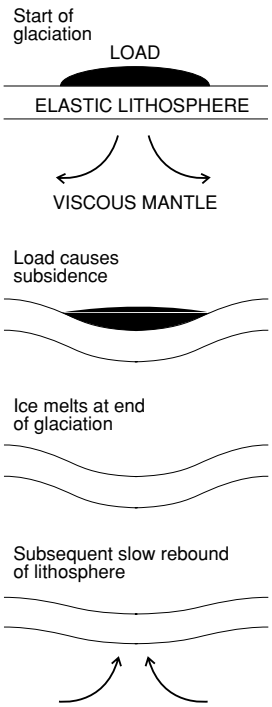


Figure 5.18. The deformation and uplift which occur as a result of loading and unloading of an elastic lithospheric plate overlying a viscous mantle.

To determine the viscosity of the uppermost mantle, one must find a narrow load. An example of such a load was the water of Lake Bonneville in Utah, U.S.A., the ancestor of the present Great Salt Lake. The old lake, which existed during the Pleistocene and had a radius of about 100 km and a central depth of about

[5] Newtonian viscosity is defined as the ratio of shear stress to strain rate and is therefore essentially a measure of the internal friction of a fluid. The viscosity of many fluids is Newtonian, that is independent of the strain rate (e.g. water, most gases, oils, glycerine). Fluids for which the viscosity varies with the strain rate are termed non-Newtonian (e.g. paints, egg white, play putty, cornflour/starch in water). There is debate about the exact properties of the mantle. However, here we simply assume it to be a Newtonian fluid. The unit of viscosity is the pascal second (Pa s), 1 Pa s $\equiv$ 1 N m^{-2} s. The viscosity of water at 20 °C is 10^{-3} Pa s; at 100 °C, it is 0.3×10^{-3} Pa s. Castor oil at 0 °C has a viscosity of 5.3 Pa s; at 20 °C its viscosity is 1 Pa s; and at 100 °C its viscosity is 2×10^{-2} Pa s. The viscosity of most fluids and of rock decreases rapidly with increasing temperature. However, within the Earth, pressure tends to counteract the effects of temperature.

300 m, dried up about 10 000 years ago. As a result of the drying of this lake, the ground is now domed: the centre of the old lake has risen 65 m relative to the margins. This doming is a result of the isostatic adjustment which took place after the water load had been removed. Estimates of the viscosity of the asthenosphere which would permit this amount of uplift to occur in the time available range from 10^{20} Pa s for a 250-km-thick asthenosphere to 4×10^{19} Pa s for a 75-km-thick asthenosphere. Postglacial uplift of a small region of northeastern Greenland and of some of the Arctic islands indicates that the lower value is a better estimate. This low-viscosity channel appears to correspond to the low-velocity zone in the upper mantle (Fig. 8.1), but, since the geographic distribution of the viscosity determination is very limited, the correlation is speculative.

Wider and much more extensive loads must be used to determine the viscosity of the upper and lower mantle. The last Fennoscandian (Finland plus Scandinavia) ice sheet, which melted some 10 000 years ago, was centred on the Gulf of Bothnia. It covered an area of approximately 4×10^6 km^2 with a maximum average thickness of ~2.5 km. The maximum present-day rate of uplift, which is more than 0.9 cm yr^{-1}, is occurring in the Gulf of Bothnia close to the centre of the ancient ice sheet (Fig. 5.19). Figure 5.20 shows the uplift calculated to occur in the northern Gulf of Bothnia for a model with an asthenosphere 75 km thick, of viscosity 4×10^{19} Pa s, overlying a mantle of viscosity 10^{21} Pa s. This predicted uplift is in good agreement with observations. It is predicted that 30 m of uplift remains. When a load is removed suddenly, rupture can occur. As the last remnants of the Fennoscandian ice sheet were lost, very large earthquakes (M_w up to 8.2) occurred, with fault lengths up to 160 km and average displacements up to 15 m. Breaks of this size probably ruptured the whole crust (Fig. 10.2).

Determination of the viscosity of the lower mantle requires loads of very large extent. The Wisconsin ice sheet, formed during the most recent Pleistocene glaciation in North America, was perhaps 3.5 km thick and covered much of Canada as well as part of the northern U.S.A. Melting of the ice resulted in rebound of the continent but also loading of the oceans. Thus, melting of an extensive ice sheet provides both loading and unloading data for study. Figure 5.21 illustrates the calculated uplift at various times after removal of a model Wisconsin ice sheet that had attained isostatic equilibrium on an elastic lithosphere plate underlain by a constant-viscosity mantle. Uplift is greatest at the centre of the load, and the rate of uplift decreases with time as expected. Regions peripheral to the load also undergo initial uplift before undergoing subsidence. Such uplift, followed by subsidence, is documented in ancient sea-level changes along the east coast of the U.S.A. (The melting of the ice sheet immediately causes sea level to rise, but the oceanic plate also deforms in response to the increased water load, thus resulting in a further change in sea level.) The uplift documented in Canada (present-day rates around Hudson Bay are about 1 cm yr^{-1}) and the sea-level changes along the east coast of North America are all in reasonable agreement with a

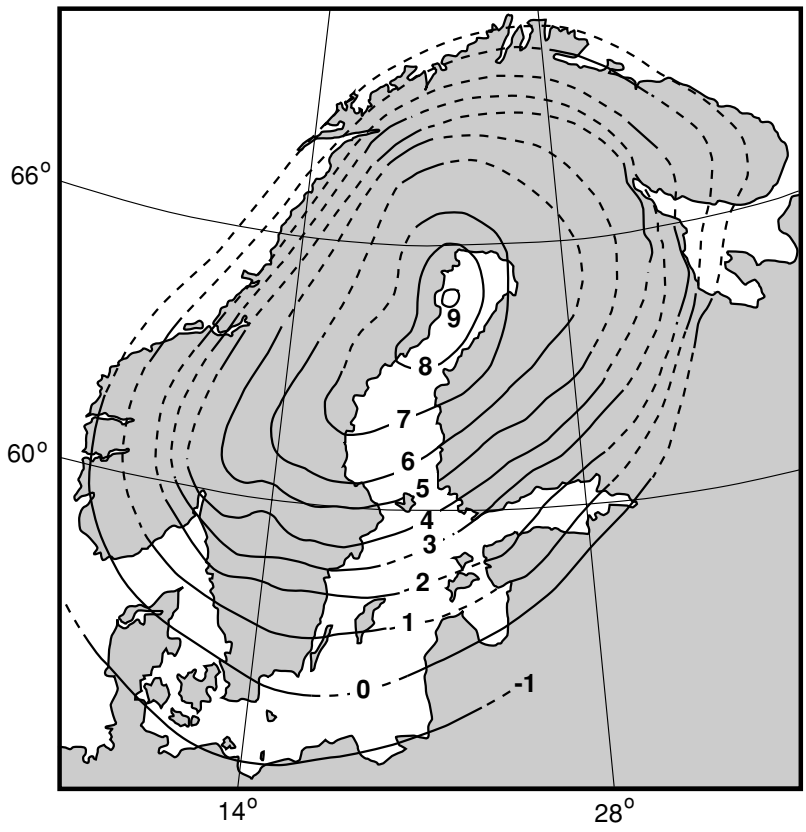


Figure 5.19. The present-day rate of uplift occurring in Fennoscandia. Contours are in millimetres per year. (After Ekman and Mäkinen (1996).)

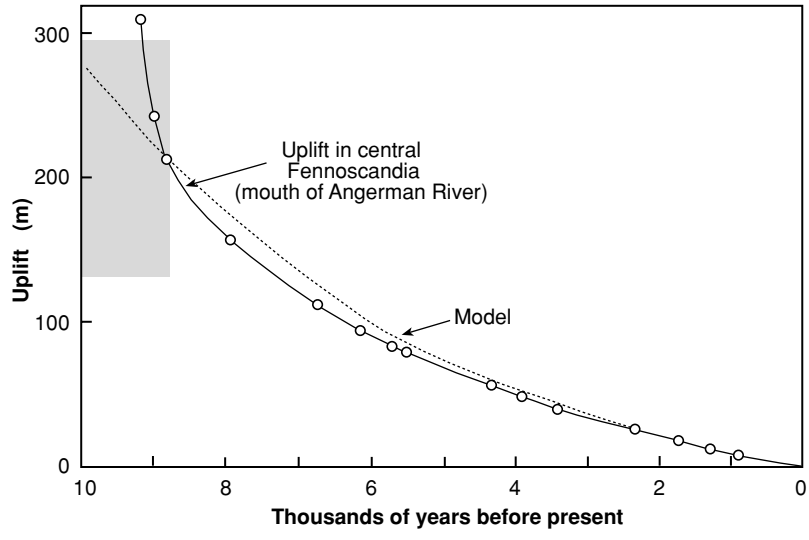
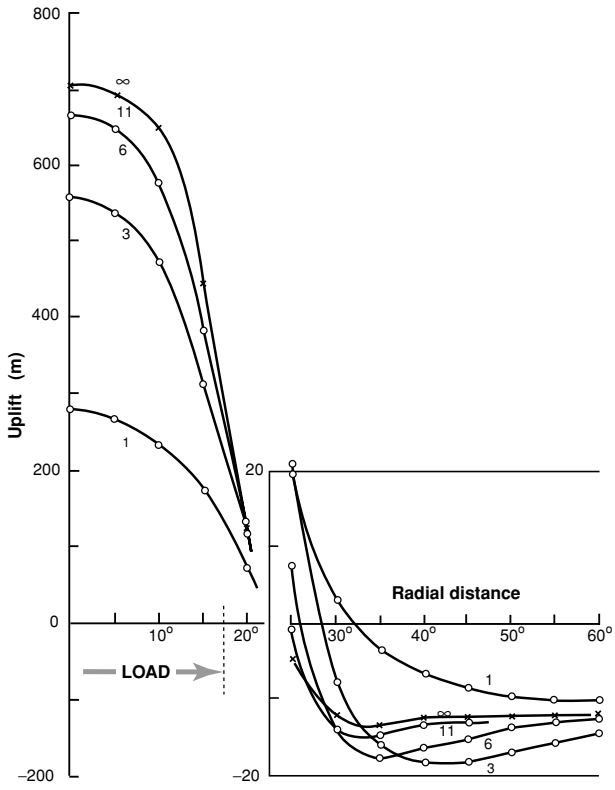


Figure 5.20. Uplift in central Fennoscandia calculated for a constant-viscosity (10^{21} Pa s) mantle (dashed line) and geological observations (o) from the northern Gulf of Bothnia. The shading indicates the time of melting of the ice sheet. (After Cathles (1975).)



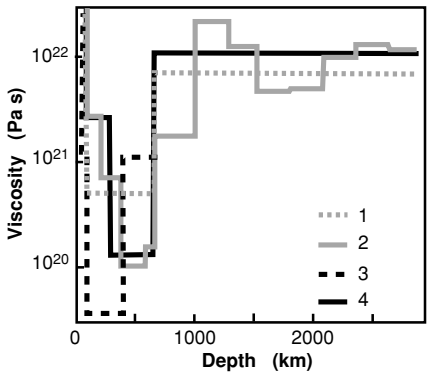


Figure 5.21. The uplift one, three, six, eleven (circles) and ∞ (crosses) thousand years after the removal of a 2.5-km-thick model Wisconsin ice sheet from an elastic lithospheric plate overlying a viscosity mantle of uniform 10^{21} Pa s. Note that, for clarity, the vertical scale at distances greater than 25° (1° = 110 km) is magnified by a factor of ten. (From Cathles (1975).)

Figure 5.22. Estimates of the viscosity structure of the mantle: 1, based on relative sea level in northwest Europe; 2, based on relative sea level in Fennoscandia and Hudson Bay together with upper- and lower-mantle seismic tomography; 3, based on a mantle-flow model and lower-mantle seismic tomography; and 4, based on a spectral study of the gravity field over Hudson Bay. (Reprinted with permission from *Nature* (Simons and Hager, *Nature*, **390**, 500–4) Copyright 1997 Macmillan Magazines Ltd.)

mantle of reasonably constant viscosity: the models imply an upper mantle with viscosity $\sim 10^{21}$ Pa s and a lower mantle with viscosity in the range 10^{22}–10^{23} Pa s. Figure 5.22 shows a series of estimates of detailed mantle viscosity structure based on sea-level and gravity data from Europe and Hudson Bay. It appears that the viscosity of the mantle increases somewhat with depth and that a low-viscosity asthenosphere is not a universal feature – beneath old continental regions the viscosity of the upper mantle seems to be fairly high.

Problems

1. Calculate the mass of the Earth.
2. Calculate the maximum gravitational attraction between Newton and the apple, assuming that (a) it passed 30 cm from his centre of mass and (b) it hit him.
3. Calculate the weight of a 60-kg astronaut on the Moon. How much less is this than the astronaut's weight on the Earth? (The Moon's mass is 7.4×10^{22} kg and its radius is 1738 km.)
4. (a) Calculate the radius of orbit of a geostationary satellite (a satellite whose orbit is such that it remains above the same point on the Earth's surface).
   (b) Calculate the period of a satellite orbiting 200 km above the Earth.
   (c) Calculate the period of a satellite orbiting 200 km above the Moon.
5. What is the radius of the Moon's orbit about the Earth? Use the mass of the Earth as calculated in Problem 1 and assume that the period of rotation is 28 days. Explain why it is not necessary to know the mass of the Moon to perform the calculation.
6. Calculate the mass of the Sun, assuming the Earth's orbit to be circular with a radius of 1.5×10^8 km.
7. A small moon of mass m and radius a orbits a planet of mass M, keeping the same face towards the planet. Show that, if the moon approaches the planet closer than $r_c = a(3M/m)^{1/3}$, loose rocks lying on the surface of the moon will be lifted off. (From Thompson (1987).)
8. What is the difference in length, in kilometres, between a degree of latitude at the equator and at the pole?
9. Explain (qualitatively and quantitatively) how the source of the Mississippi River can be about 5 km closer to the centre of the Earth than its mouth. (From Thompson (1987).)
10. It is observed that the acceleration g due to gravity is greater down a mine than it is at the Earth's surface. Show that this result can be explained if the Earth's density increases sufficiently rapidly with depth. (From Thompson (1987).)
11. What is the difference between the value of the Earth's gravitational acceleration at the equator and the poles? In what units would this difference usually be quoted?
12. (a) Verify the isostatic-equilibrium calculation for continents and ocean basins, using densities of sea water, crust and mantle of 1.03×10^3, 2.9×10^3 and 3.3×10^3 kg m^{-3}, respectively, an ocean-basin depth of 5 km and an oceanic crustal thickness of 6.6 km.

(b) What would happen if the ocean-basin depth were to change to (i) 1 km, (ii) 3 km or (iii) 8 km?

13. Calculate the depths and densities beneath a 5-km-high mountain chain in isostatic equilibrium with a 35-km-thick continental crust of density 2.8×10^3 kg m^{-3} and a mantle of density 3.3×10^3 kg m^{-3} by using the hypotheses of (a) Pratt and (b) Airy.
14. A mountain range 4 km high is in isostatic equilibrium.
    (a) During a period of erosion, a 2 km thickness of material is removed from the mountains. When the new isostatic equilibrium is achieved, how high are the mountains?
    (b) How high would they be if 10 km of material were eroded away?
    (c) How much material must be eroded to bring the mountains down to sea level? (Use crustal and mantle densities of 2.8×10^3 kg m^{-3} and 3.3×10^3 kg m^{-3}, respectively.)
15. A 500-m-deep depression on the Earth's surface fills with (a) sandstone of density 2.2×10^3 kg m^{-3} and (b) ironstone of density 3.4×10^3 kg m^{-3}. Assuming that isostatic equilibrium is attained, calculate the thicknesses of sediment that will be deposited in these two cases. (Use crustal and mantle densities of 2.8×10^3 and 3.3×10^3 kg m^{-3}, respectively.)
16. If subduction doubled the thickness of the continental crust, calculate the elevation of the resulting plateau. Assuming that all plateaux are eroded down to sea level, calculate the total thickness of material that would be eroded.
17. Assume that the oceanic regions are in Airy-type isostatic equilibrium. If the lithosphere has a uniform density, show that the depth of the seabed is given by

$$d + (L(t) - L(0))\frac{\rho_1 - \rho_a}{\rho_a - \rho_w}$$

    where $L(t)$ is the thickness of the lithosphere of age t, d is the depth of water at the ridge axis and ρ_w, ρ_1 and ρ_a are the densities of water, lithosphere and asthenosphere, respectively.
18. (a) Calculate the maximum gravity anomaly due to a sphere of radius 1 km with density contrast 300 kg m^{-3} buried at a depth of (i) 1 km, (ii) 2 km and (iii) 15 km.
    (b) Calculate the gravity anomalies at distances of 1, 5 and 10 km from these spheres.
19. (a) What is the maximum gravity anomaly due to a cylinder of radius 1 km with density contrast 200 kg m^{-3} buried at a depth of 1 km.
    (b) A cylinder, of radius 50 km and buried at a depth of 100 km, yields the same maximum gravity anomaly as that in (a). Calculate the density contrast of this deep cylinder.
    (c) Can gravity measurements resolve deep anomalous mantle density structures?
20. (a) Calculate the geoid height anomaly due to a mountain of height h for Pratt-type compensation. What is this anomaly for $h = 2$ km?
    (b) Calculate the same anomaly for an ocean depth d. What is this anomaly for $d =$ 5 km?
21. What is the anomaly in the gravitational potential as measured on the spheroid if the geoid height anomaly is (a) 3 m, (b) -5 m and (c) 8 m?

22. Calculate the depression of the land surfaces beneath (a) the Wisconsin ice sheet and (b) the Fennoscandia ice sheet, assuming that, prior to and after the emplacement of the ice sheets, they were in isostatic equilibrium.
23. What gravitational arguments can you put forward to counter the proposal that the oceans formed when huge continental areas sank beneath the sea?
24. Estimate the gravitational effect of the 76 men mentioned in Section 5.2.1.

References and bibliography

Airy, G. B. 1855. On the computations of the effect of the attraction of the mountain masses as disturbing the apparent astronomical latitude of stations in geodetic surveys. *Phil. Trans. Roy. Soc. Lond.*, **145**, 101–4.

Arvidsson, R. 1996. Fennoscandian earthquakes: whole crust rupturing related to postglacial rebound, *Science*, **274**, 744–6.

Beaumont, C. 1978. The evolution of sedimentary basins on a viscoelastic lithosphere: theory and examples. *Geophys. J. Roy. Astr. Soc.*, **55**, 471–98.

Bott, M. H. P. 1982. *The Interior of the Earth*, 2nd edn. Amsterdam: Elsevier.

Bowin, C. 2000. Mass anomaly structure of the Earth. *Rev. Geophys.*, **38**, 355–87.

Caldwell, J. E., Haxby, W. F. Kang, D. E. and Turcotte, D. L. 1976. On the applicability of a universal elastic trench profile. *Earth Planet. Sci. Lett.*, **31**, 239–46.

Cathles III, L. M. 1975. *The Viscosity of the Earth's Mantle*. Princeton, New Jersey: Princeton University Press.

Chapman, M. E. and Talwani, M. 1979. Comparison of gravimetric geoids with Geos 3 altimetric geoid. *J. Geophys. Res.*, **84**, 3803–16.

Cook, A. H. 1973. *Physics of the Earth and Planets*. New York: Wiley.

De Bremaecker, J.-C. 1985. *Geophysics: The Earth's Interior*. New York: Wiley.

Dobrin, M. B. and Savit, C. H. 1988. *Introduction to Geophysical Prospecting*, 4th edn. New York: McGraw-Hill.

Ekman, M. and Mäkinen, J. 1996. Recent postglacial rebound, gravity change and mantle flow in Fennoscandia. *Geophys. J. Int.*, **126**, 229–34.

Flint, R. F. 1971. *Glacial and Quarternary Geology*. New York: Wiley.

Forte, A. M. and Mitrovica, J. X. 1996. New inferences of mantle viscosity from joint inversion of long-wavelength mantle convection and post-glacial rebound data. *Geophys. Res. Lett.*, **23**, 1147–50.

Garland, G. D. 1979. *Introduction to Geophysics: Mantle, Core and Crust*, 2nd edn. Philadelphia, Pennsylvania: Saunders.

Hoffman, N. R. A. and McKenzie, D. P. 1985. The destruction of geochemical heterogeneities by differential fluid motions during mantle convection. *Geophys. J. Roy. Astr. Soc.*, **82**, 163–206.

King-Hele, D. G. 1969. Royal Aircraft Establishment Technical Memorandum Space 130, Farnborough, U.K.

King-Hele, D. G., Brookes, C. J. and Cook, G. E. 1981. Odd zonal harmonics in the geopotential from analysis of 28 satellite orbits. *Geophys. J. Roy. Astr. Soc.*, **64**, 3–30.

Lambeck, K. and Johnston, P. 1998. The viscosity of the mantle from analyses of glacial rebound. In I. Jackson ed., *The Earth's Mantle*. Cambridge: Cambridge University Press, pp. 461–502.

Lerch, F. J., Klosko, S. M., Laubscher, R. E. and Wagner, C. A. 1979. Gravity model improvement using GEOS 3 (GEM 9 and 10). *J. Geophys. Res.*, **84**, 3897–916.

McKenzie, D. P. 1977. Surface deformation, gravity anomalies and convection. *Geophys. J. Roy. Astr. Soc.*, **48**, 211–38.

1983. The Earth's mantle. *Sci. Am.*, **249** (3), 66–113.

McKenzie, D. 1998. Magellan looks at Venus. *Proc. Roy. Inst. Great Britain*, **69**, 15–48.

McKenzie, D. and Bowin, C. 1976. The relationship between bathhymetry and gravity in the Atlantic Ocean. *J. Geophs. Res.*, **81**, 1903–15.

McKenzie, D. P., Roberts, J. M. and Weiss, N. O. 1974. Convection in the Earth's mantle: towards a numerical simulation. *J. Fluid Mech.*, **62**, 465–538.

McKenzie, D. P., Watts, A. B., Parsons, B. and Roufosse, M. 1980. Planform of mantle convection beneath the Pacific Ocean. *Nature*, **288**, 442–6.

Milne, J. 1906. Bakerian Lecture: Recent advances in seismology. *Proc. Roy. Soc.* A, **77**, 365–76.

Morner, N. A., ed., 1980. *Earth Rheology, Isostacy and Eustacy*. New York: Wiley.

Nakada, M. and Lambeck, K. 1987. Glacial rebound and relative sea-level variations: a new appraisal. *Geophys. J. Roy. Astr. Soc.*, **90**, 171–224.

Peltier, W. R. 1983. Constraint on deep mantle viscosity from Lageos acceleration data. *Nature*, **304**, 434–6.

1984. The thickness of the continental lithosphere. *J. Geophys. Res.*, **89**, 11303–16.

1985. New constraint on transient lower mantle rheology and internal mantle buoyancy from glacial rebound data. *Nature*, **318**, 614–17.

Pratt, J. H. 1855. On the attraction of the Himalaya Mountains, and of the elevated regions beyond them, upon the plumb line in India. *Phil. Trans. Roy. Soc. Lond.*, **145**, 53–100.

Ranalli, G. 1995. *Rheology of the Earth*, 2nd edn. London: Chapman and Hall.

Sandwell, D. T. and Renkin, M. L. 1988. Compensation of swells and plateaus in the North Pacific: no direct evidence for mantle convection. *J. Geophys. Res.*, **93**, 2775–83.

Sandwell, D. T. and Smith, W. H. F. 1997. Marine gravity anomaly from Geosat and ERS 1 satellite altimetry. *J. Geophys. Res.*, **102**, 10 039–54.

Simons, M. and Hager, B. H. 1997. Localization of the gravity field and the signature of glacial rebound. *Nature*, **390**, 500–4.

Smith W. H. F. and Sandwell, D. T. 1997. Global sea floor topography from satellite altimetry and ship depth soundings. *Science*, **277**, 1956–62.

Stein, C. A. and Stein, S. 1996. Thermo-mechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes (overview). In G. E. Bebout, D. W. Scholl, S. H. Kirkby and J. P. Platt, *Subduction: Top to Bottom*. Geophysical Monograph **96**. Washington: American Geophyical Union, pp. 1–17.

Telford, W. M., Geldart, L. P. and Sheriff, R. E. 1990. *Applied Geophysics*, 2nd edn. New York: Cambridge University Press.

Thompson, N., ed., 1987. *Thinking like a Physicist*. Bristol: Adam Hilger.

Turcotte, D. L. 1979. Flexure. *Adv. Geophys.*, **21**, 51–86.

Turcotte, D. L., McAdoo, D. C. and Caldwell, J. G. 1978. An elastic–perfectly plastic analysis of the bending of the lithosphere at a trench. *Tectonophysics*, **47**, 193–205.

Turcotte, D. L. and Schubert, G. 2002. *Geodynamics*, 2nd edn. Cambridge: Cambridge University Press.

Warner, M. R. 1987. Seismic reflections from the Moho: the effect of isostasy. *Geophys. J. Roy. Astr. Soc.*, **88**, 425–35.

Watts, A. B. 1978. An analysis of isostasy in the world's oceans 1. Hawaiian–Emperor seamount chain. *J. Geophys. Res.*, **83**, 5989–6004.

1979. On geoid heights derived from Geos 3 altimeter data along the Hawaiian–Emperor seamount chain. *J. Geophys. Res.*, **84**, 3817–26.

2001. *Isostasy and Flexure of the Lithosphere*. Cambridge: Cambridge University Press.

Watts, A. B., Bodine, J. H. and Ribe, N. M. 1980. Observations of flexure and the geological evolution of the Pacific Ocean. *Nature*, **283**, 532–7.

Watts, A. B. and Daly, S. F. 1981. Long wavelength gravity and topography anomalies. *Ann. Rev. Earth Planet. Sci.*, **9**, 415–48.

Watts, A. B., McKenzie, D. P., Parsons, B. E. and Roufosse, M. 1985. The relationship between gravity and bathymetry in the Pacific Ocean. *Geophys. J. Roy. Astr. Soc.*, **83**, 263–98.

Chapter 6
Geochronology

6.1 Introduction

Radioactivity was discovered in 1896 by Henri Becquerel. The new understanding of the atom that came from the work of Rutherford, Soddy, Boltwood and others over the following decades had a major impact on geology. Before this work, the age of the Earth was unknown. In the nineteenth century, Lord Kelvin (William Thompson) attempted to calculate the age of the Earth by assuming that the planet was a hot body cooling by conduction. He obtained a young age that conflicted with the observations of geologists, who had concluded that the Earth must be at least several hundred million years old. The geological reasoning was based on rather qualitative evidence such as the observation of sedimentary deposition rates, calculations about the amount of salt in the sea[1] and guesses of evolutionary rates. Not unnaturally, Kelvin's apparently more rigorous and quantitative physical calculation was regarded as much sounder by most scientists. In 1904, while at McGill University in Montreal, Ernest Rutherford realized that radioactive heat could account for some of the apparent discrepancy. Kelvin was sceptical to the extent that he bet the younger Rayleigh (Hon. R. J. Strutt) five shillings on the matter, but later he paid up. Rutherford gave a lecture at the Royal Institution in London in 1904 about which he wrote the following:

> To my relief Kelvin fell fast asleep but as I came to the important point, I saw the old bird sit up, open an eye and cock a baleful glance at me. Then a sudden inspiration came and I said Lord Kelvin had limited the age of the Earth, provided no new source (of heat) was discovered . . . Behold! the old boy beamed upon me.

In fact, Kelvin's calculation is now known to be too simple; not only does radioactive decay provide an extra source of heat, but also, much more importantly, heat is transferred within the mantle by convection, not by conduction (see Section 7.4). Had Kelvin carried out a convection calculation, as he did for the Sun, he would have obtained a more reasonable answer. However, even had he done this, the fact that estimates for the age of the Sun (based on its energy output) were also only

[1] Interestingly, the sea-salt puzzle was not solved until many years later when, with plate tectonics, the significance of geochemical exchange in hydrothermal convection at ridges was understood.

about twenty million years[2] would have made any such calculations seem quite unrealistic to those involved in the debate. Soon after this lecture, Rutherford and his colleague Boltwood developed radioactive dating, which was first applied to show that even the geological estimates of the age of the Earth were too modest. Rutherford's first dating method was to measure the accumulation of helium (α particles) from the radioactive decay of uranium but he later suggested that measuring the accumulation of lead would be a better method. In 1907, Boltwood published the first uranium–lead dates.

The geological tools of stratigraphy and palaeontology provide a very *precise*[3] method of measuring *relative ages*. In good cases, a palaeontological resolution of better than 0.25 Ma can be attained. Stratigraphy and palaeontology enable us to determine the order in which rocks were laid down and the order in which tectonic events and sea transgressions occurred. The geological timescale is made up of stratigraphic divisions based on observed rock sequences. However, palaeontology is unable to give any *accurate*[4] estimate of the *absolute ages* of geological events. The use of radioactive isotopes to date rocks, being a measurement of physical properties, is quite separate from any intuitive method and thus can provide an independent, fairly accurate and sometimes very precise date. There are, of course, many assumptions and inherent problems in these dating methods, as in any other, of which anyone using such dates needs to be aware. For details of the practical uses of the various dating methods the reader is referred to textbooks on geochronology (e.g., Faure 1986).

Table 6.1 shows a standard geological timescale, which is based on stratigraphy, palaeontology, geochronology and reversals of geomagnetic polarity.

6.2 General theory

The disintegration of any radioactive atom is a random event, which occurs independently of neighbouring atoms, physical conditions and the chemical state of the atom. Disintegration depends only on the structure of the nucleus. This means that every atom of a given type has the same probability of disintegrating in unit time. This probability is called the *decay constant*, λ, a different constant for each isotope. Suppose that at time t there are P atoms and that at a time δt later, δP of these atoms have disintegrated. Then δP is the product of the probability that any one of these atoms will decay in a unit of time (λ) multiplied by the number of atoms present (P) multiplied by the length of time (δt):

$$\delta P = -\lambda P \, \delta t \tag{6.1}$$

[2] The estimate of the Sun's age was based on assumptions that its energy was of gravitational origin. It was not until the 1930s that thermonuclear fusion reactions were discovered and understood to be fuelling the Sun. Only then were the apparently young ages for the Sun shown to be unfounded.

[3] *Precision* describes the reproducibility with which measurements are made.

[4] *Accuracy* describes the truth of these measurements.

Table 6.1 *A Geological time scale*

Eon	Era	Period, subera	Epoch, subperiod	Age (Ma)
Phanerozoic	Cenozoic	Quaternary Q	Holocene	0.01
			Pleistocene	1.8
		Tertiary TT	Pliocene	5.3
			Miocene	23.8
			Oligocene	33.7
			Eocene	54.8
			Palaeocene	65.0
	Mesozoic	Cretaceous K	Late	99.0
			Early	144
		Jurassic J	Late	159
			Middle	180
			Early	206
		Triassic Tr	Late	227
			Middle	242
			Early	248
	Palaeozoic	Permian P	Late	256
			Early	290
		Carboniferous C	Pennsylvanian	323
			Mississippian	354
		Devonian D	Late	370
			Middle	391
			Early	417
		Silurian S	Late	423
			Early	443
		Ordovician O	Late	458
			Middle	470
			Early	490
		Cambrian ϵ	Merioneth	
			St David's	
			Caerfai	543
Proterozoic	Late Middle Early	Hadrynian Helikian Aphebian *(Canada)*		~1000 1800 2500
Archaean	Late Middle Early	Kenoran Pilbara *(Canada)*	Transvaal Witwatersrand Pongola Barberton *(S. Africa)*	Shamvaian Bulawayan Belingwean Sebakwian *(Zimbabwe)* Isua *(Greenland)* ~4000
		Zircons in Jack Hills (Australia)		
		Origin of Earth		~4560

Note: There is continuing debate about the ages assigned to various boundaries (e.g. see Fig. 3.14). Precambrian names have no international status and are illustrative examples only.

Source: Based on Gradstein and Ogg (1996) and Palmer and Geissman (1999).

In the limit as δP and $\delta t \to 0$, Eq. (6.1) can be written as a differential equation:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\lambda P \tag{6.2}$$

Thus, the rate of disintegration $\mathrm{d}P/\mathrm{d}t$ is proportional to the number of atoms present, P. Integration of Eq. (6.2) gives

$$\log_{\mathrm{e}} P = -\lambda t + c \tag{6.3}$$

where c is a constant. If at some starting time $t = 0$ there are P_0 atoms present, then

$$c = \log_{\mathrm{e}} P_0 \tag{6.4}$$

Thus, Eq. (6.3) can be written

$$P = P_0 \mathrm{e}^{-\lambda t} \tag{6.5}$$

The rate of disintegration $\mathrm{d}P/\mathrm{d}t$ is sometimes called the activity A. Equation (6.2) can be rewritten in terms of A instead of P, giving

$$A = A_0 \mathrm{e}^{-\lambda t} \tag{6.6}$$

which is then an alternative expression for Eq. (6.5).

Suppose that at some starting time there are P_0 radioactive atoms. The *half-life*, $T_{1/2}$, is then the length of time required for one-half of these original atoms to undergo disintegration. Putting $P = P_0/2$ into Eq. (6.5) gives

$$\frac{P_0}{2} = P_0 \mathrm{e}^{-\lambda T_{1/2}} \tag{6.7}$$

or, after reorganizing the equation,

$$\begin{aligned} T_{1/2} &= \frac{1}{\lambda} \log_{\mathrm{e}} 2 \\ &= \frac{0.693}{\lambda} \end{aligned} \tag{6.8}$$

Consider the case of a radioactive *parent* atom disintegrating to a stable atom called the *daughter*. At time $t = 0$ there are P_0 parent atoms and no daughter atoms. After a time t, there are D daughter atoms and $P = P_0 - D$ parent atoms since all the daughter atoms have come from disintegration of the radioactive parent atoms. Thus Eq. (6.5) can be written

$$P_0 - D = P_0 \mathrm{e}^{-\lambda t} \tag{6.9}$$

or

$$D = P_0(1 - \mathrm{e}^{-\lambda t}) \tag{6.10}$$

However, since it is not possible to measure P_0, but only P, Eq. (6.5) must be used to rewrite Eq. (6.10):

$$D = P(\mathrm{e}^{\lambda t} - 1) \tag{6.11}$$

This equation expresses the number of daughter atoms D in terms of the number of parent atoms P, both measured at time t. Taking the natural logarithm of Eq. (6.11) allows the time t that has elapsed to be expressed in terms of the numbers of the parent P and daughter D atoms present:

$$t = \frac{1}{\lambda}\log_e\left(1 + \frac{D}{P}\right) \tag{6.12}$$

This is therefore a clock based on the gradual reduction in the number of parent atoms and build-up in the number of daughter atoms. The equations use the probability that any atom will survive a time t. The actual proportion of parent atoms that survive in any particular case is subject to statistical fluctuations, but these are usually very small because P_0 and P are very large numbers. In practice, measurements of the ratio D/P would be made with a mass spectrometer.[5] The half-life of the radioactive isotope used to estimate the age of a rock should ideally be comparable to the age of the sample because, if they are very different, the measurement of D/P in Eq. (6.12) may become impracticable.

Age determination is rarely as simple as applying Eq. (6.12). In most cases there will have been an initial concentration of the daughter in the sample, and thus the assumption that all the measured daughter atoms are a product of the parent is not necessarily valid. Also, all systems are not *closed*, as has been assumed here; in other words, over time there may have been some exchange of the parent and/or daughter atoms with surrounding material (in other words, the parent and daughter atoms have not been kept in a 'closed' box). It is important to remember that the parent atoms keep on decaying to their daughters regardless of what is happening to the rock. The numbers of parent and daughter atoms can be used to act as a simple clock only if the rock system is closed so that none of them escape. In other words, the stop watch starts ticking only when the system closes (the box shuts). If, when more than one dating method is used to estimate t, the resulting ages are within analytical error, then these potential problems can be neglected; in this case the various dates are *concordant*. In the case of *discordant* ages, the possibility of a non-closed or *open* system or the initial presence of the daughter must be considered and, when possible, appropriate corrections made. (See 'Other important factors' later in this section.)

The value for t that is estimated does not necessarily establish the time that has passed since the formation of a rock; it may represent the time since the rock crystallized or since cooling following a metamorphic event that heated the rock such that chemical changes took place (see Section 7.8.5 for discussion).

Table 6.2 shows the decay products, half-lives and decay constants for those radioactive isotopes most often used in dating geological samples. Also given are the *heat-generation rates* for the four isotopes that are responsible for effectively

[5] A mass spectrometer is an instrument used to establish the ratio of numbers of isotopes with different masses. It depends for its operation on the fact that moving ions with different masses can be separated by magnetic fields.

Table 6.2 *Decay schemes for radioactive isotopes primarily used in geochronology*

Parent isotope	Daughter isotope	Decay products	Decay constant (a^{-1})	Half-life (a)	Present rate of heat generation ($W\,kg^{-1}$)
^{238}U	^{206}Pb	$8\alpha + 6\beta$	1.55×10^{-10}	$4\,468 \times 10^6$	9.4×10^{-5}
^{235}U	^{207}Pb	$7\alpha + 4\beta$	9.85×10^{-10}	704×10^6	5.7×10^{-4}
^{232}U	^{208}Pb	$6\alpha + 4\beta$	4.95×10^{-11}	$14\,010 \times 10^6$	2.7×10^{-5}
^{87}Rb	^{87}Sr	β	1.42×10^{-11}	$48\,800 \times 10^6$	
^{147}Sm	^{143}Nd	α	6.54×10^{-12}	$106\,000 \times 10^6$	
^{40}K	^{40}Ca	β	4.96×10^{-10}	$1\,400 \times 10^{10}$	2.8×10^{-5}
	^{40}Ca	Electron capture	5.81×10^{-11}	$11\,900 \times 10^6$	$1\,250 \times 10^6$
^{39}Ar	^{39}K	β	2.57×10^{-3}	269	
^{176}Lu	^{176}Hf	β	1.94×10^{-11}	$35\,000 \times 10^6$	
^{187}Re	^{187}Os	β	1.52×10^{-11}	$45\,600 \times 10^6$	
^{14}C	^{14}N	β	1.21×10^{-4}	5 730	

Note: Annum (a) is the SI unit for year: $a^{-1} = yr^{-1}$.

Source: Decay constants and half-lives are based on Steiger and Jaeger (1977).

all the radioactive heating in the Earth (see Chapter 7). All the radiations produced (except the neutrinos[6] that accompany every decay) contribute heat as they are absorbed within about 30 cm of their origin. The neutrinos interact so little with matter that they penetrate the whole Earth and escape.

Radioactive decay schemes are not always simple. A series of disintegrations may take place before a stable daughter nuclide is formed. Of those isotopes that are commonly used for geological dating, ^{87}Rb and ^{147}Sm undergo a simple decay to stable daughters (β decay and α decay, respectively); some of the others undergo a sequence of decays (Table 6.2).

Decay series

The long series of decays undergone by the isotopes of uranium and thorium can be described by a series of equations like Eq. (6.2). Suppose that the radioactive isotope decays in series to $X_1, X_2, \ldots, X_n$ and the isotope X_n is the stable daughter. Let the number of atoms of each isotope present at time t be $P, P_1, P_2, \ldots, P_{n-1}$, D and let the decay constants be $\lambda, \lambda_1, \lambda_2, \ldots, \lambda_{n-1}$.

The number of atoms of each isotope in the series increases due to the decay of its parent and is reduced as the result of its own radioactive decay. Equation (6.2) can be applied to each disintegration in the series:

$$\begin{aligned}
\frac{\mathrm{d}P}{\mathrm{d}t} &= -\lambda P \\
\frac{\mathrm{d}P_1}{\mathrm{d}t} &= -\lambda_1 P_1 + \lambda P \\
\frac{\mathrm{d}P_2}{\mathrm{d}t} &= -\lambda_2 P_2 + \lambda_1 P_1 \\
&\vdots \\
\frac{\mathrm{d}P_{n-1}}{\mathrm{d}t} &= -\lambda_{n-1} P_{n-1} + \lambda_{n-2} P_{n-2} \\
\frac{\mathrm{d}D}{\mathrm{d}t} &= \lambda_{n-1} P_{n-1}
\end{aligned} \tag{6.13}$$

The general solution to these equations is more complicated than Eq. (6.5). The solution for, say, the ith isotope in the chain involves exponentials for all the decays from the parent to isotope i inclusive. (In physics, these equations are known as the *Bateman relations*.)

For the decay schemes of uranium and thorium, the half-life of the first decay is many orders of magnitude greater than those of the subsequent disintegrations

[6] A neutrino is a particle produced during nuclear reactions, which has no charge and little interaction with matter. A remarkable illustration of the neutrino's penetrating ability is that the burst of neutrinos detected on 23 February 1987, which came from the supernova explosion in the Greater Magellanic Cloud high in the southern sky, was detected in the northern hemisphere, in Japan and the U.S.A., after they had passed through the Earth.

($\lambda \ll \lambda_1, \ldots, \lambda_{n-1}$). In this special case, the solution to Eqs. (6.13) is much simplified:

$$\begin{aligned} P &= P_0 e^{-\lambda t} \\ P_i &= \frac{\lambda}{\lambda_i} P \quad (i = 1, \ldots, n-1) \\ D &= P_0(1 - e^{-\lambda t}) \end{aligned} \tag{6.14}$$

where P_0 is the number of atoms of the parent isotope at the starting time $t = 0$. This simplified solution in which the rate of decay of the daughter is equal to the rate of decay of the parent is termed *secular equilibrium* ($\lambda_i P_i = \lambda P$) and is established after a time comparable to the half-life of the radioactive daughters. Then, just as for the simple decay scheme, the time t is given by Eq. (6.12).

Branched decay

The decay of ^{40}K is also complex. It can decay either by β decay to ^{40}Ca or by electron capture to ^{40}Ar. At time t, let P be the number of ^{40}K atoms and D_C and D_A the numbers of ^{40}Ca and ^{40}Ar atoms, respectively. Then Eq. (6.2) can be written as

$$\frac{dP}{dt} = -(\lambda_A + \lambda_C)P \tag{6.15}$$

where λ_A and λ_C are the rates of decay to ^{40}Ar and ^{40}Ca, respectively. Equation (6.5) then becomes

$$P = P_0 e^{-(\lambda_A + \lambda_C)t} \tag{6.16}$$

The rate of increase in number of ^{40}Ar atoms dD_A/dt is given by

$$\frac{dD_A}{dt} = \lambda_A P \tag{6.17}$$

Similarly, the rate of increase in number of ^{40}Ca atoms is

$$\frac{dD_C}{dt} = \lambda_C P \tag{6.18}$$

Substituting the expression for P from Eq. (6.16) into Eq. (6.17) gives

$$\frac{dD_A}{dt} = \lambda_A P_0 e^{-(\lambda_A + \lambda_C)t} \tag{6.19}$$

and so

$$D_A = -\frac{\lambda_A P_0}{\lambda_A + \lambda_C} e^{-(\lambda_A + \lambda_C)t} + c \tag{6.20}$$

where c is a constant of integration.

If we assume that $D_A = 0$ at $t = 0$, then $c = \lambda_A P_0/(\lambda_A + \lambda_C)$, so

$$D_A = \frac{\lambda_A P_0}{\lambda_A + \lambda_C}\left(1 - e^{-(\lambda_A + \lambda_C)t}\right) \tag{6.21}$$

Using Eq. (6.16) to eliminate P_0, we can finally express the ratio of the numbers of atoms of ^{40}Ar to ^{40}K at time t in terms of the decay rates:

$$\frac{D_A}{P} = \frac{\lambda_A}{\lambda_A + \lambda_C}\left(e^{(\lambda_A+\lambda_C)t} - 1\right) \tag{6.22}$$

The age of the sample t is therefore given by

$$t = \frac{1}{\lambda_A + \lambda_C}\log_e\left[1 + \left(\frac{\lambda_A + \lambda_C}{\lambda_A}\right)\frac{D_A}{P}\right] \tag{6.23}$$

If the decay to ^{40}Ca were used, Eqs. (6.19)–(6.23) would have the same form but with D_C and the rate of decay to calcium, λ_C, replacing D_A and the rate of decay to argon λ_A. The decay of ^{40}K to ^{40}Ca is little used in geochronology because calcium is so common that the ^{40}Ca resulting from ^{40}K decays cannot be accurately determined.

Other important factors

The choice and application of these radioactive dating methods to actual samples require knowledge of a number of factors concerning both the methods and the samples.

The choice of a particular dating method depends on the probable age of the sample. Ideally, the decay scheme used should have a half-life of about the same order of magnitude as the age of the sample. This ensures that the number of daughter atoms is of the same magnitude as the number of parent atoms (see Eq. (6.11)). If the half-life is very much greater or smaller than the age of the sample, the ratio of daughter to parent atoms is either very small or very large, which, in either case, may be difficult to measure accurately. Hence, in this respect, the uranium–lead method appears to be the best radioactive dating method for younger rocks. Very small quantities of the daughter atoms have to be measured, and careful analysis is required in all isotope work. For recent samples, particularly archaeological samples, short-half-life methods involving isotopes such as carbon-14 are used.

Two other important factors affect the choice of dating method: the amounts of parent and daughter elements present in the rock; and whether or not the sample was a closed system. Clearly a sample that had never contained any rubidium could not be a candidate for rubidium–strontium dating (rubidium is a trace element). The advantage of potassium–argon dating is that it can be used on most rocks because potassium is a widespread element. Table 6.3 gives rough estimates of the concentrations of radioactive elements in average rocks.

In the formulation of Eqs. (6.1)–(6.23), we have assumed that there has been no loss or gain of parent or daughter atoms except by radioactive decay. In other words, we have assumed that our sample was a closed system. For most rocks this is clearly a false assumption. For example, radioactively derived lead (termed *radiogenic* lead) is sometimes deposited together with uranium in cooling rocks (both are transported in circulating hydrothermal water). The extra amount of

Table 6.3 *Broad estimates of concentrations of radioactive and common daughter elements in rocks*

	U (ppm)	Th (ppm)	Pb (ppm)	K (%)	Rb (ppm)	Sr (ppm)	Sm (ppm)	Nd (ppm)
Granitoid	4	15	20	3.5	200	300	8	44
Basalt	0.5	1	<4	0.8	30	470	10	40
Ultramafic	0.02	0.08	0.1	0.01	0.5	50	0.5	2
Shale	4	12	20	2.7	140	300	10	50

Source: After York and Farquhar (1972) and Faure (1986).

daughter lead present in the rocks would result in an overestimation of their age. As another example consider argon, a gas, which is the daughter product of ^{40}K decay. Although argon is not likely to have been present in the rock initially, it cannot be retained in minerals until they have cooled to below their *closure temperatures* (about 300 °C for biotite and 550 °C for hornblende). These are informal estimates; actual temperatures depend on local conditions. Thus, a potassium–argon date of, say, a granite is not the date of its intrusion but the time at which the minerals in the granite cooled below their particular closure temperatures. If, after cooling, the granite were reheated to temperatures above the closure temperatures, then argon would be lost. A potassium–argon date, therefore, dates the last time that the sample cooled below the closure temperature. Table 6.4 is an informal compilation of closure temperatures for various dating methods.

The relationship between the closure temperature of a mineral and its cooling history can be put onto a more rigorous footing by the use of thermodynamics. Since the diffusion of any species in a solid mineral is controlled by temperature, the *diffusion coefficient D* can be defined as

$$D = D_0 \mathrm{e}^{-E/(RT)} \tag{6.24}$$

where D_0 is the diffusion coefficient of the particular species and mineral involved at infinitely high temperature, E the activation energy of the diffusion process, R the gas constant and T the temperature. This is the *Arrhenius equation*. Clearly D, the diffusion coefficient, is very dependent on temperature; a small change in T can produce an order-of-magnitude change in D (Fig. 6.1). An expression for the closure temperature T_c (sometimes referred to as the blocking temperature) is

$$\frac{E}{RT_c} = \log_e \left(-\frac{AD_0RT_c^2}{\alpha^2 E\, \partial T/\partial t\big|_{T=T_c}} \right) \tag{6.25}$$

Table 6.4 *An informal compilation of closure temperatures*

Mineral	Closure temperature (°C)
Potassium–argon	
Hornblende	530 ± 40
Biotite	280 ± 40
Muscovite	~350
Microcline	130 ± 15 (plateau segment)
	110 (0% release intercept)
Uranium–lead	
Zircon	>1000
Monazite	>650
Sphene	>600
Allanite	>600
Apatite	~350
Rubidium–strontium	
Biotite	300
Muscovite	>400–500
Apatite, feldspar	~350
Fission tracks	
Zircon	175–225
Sphene	290 ± 40
Apatite	105 ± 10

Source: After Ghent *et al.* (1988).

where T_c is the closure temperature, α the characteristic dimension of the crystal, A a geometrical factor that depends on the geometry of the system and $\partial T/\partial t\big|_{T=T_c}$ the rate of change of temperature at the closure temperature (the slope of the upper graph at T_c). The solution of this equation can be found only by substitution; it cannot be inverted to give an explicit relation between T_c and the other variables.

To interpret the isotopic systematics of a mineral, all of these factors have to be taken into account, which is not always an easy matter. Isotope measurements can be made very accurately on very small samples. Advancing technology means that, whereas earliest measurements had to be made using 'whole-rock' samples, now individual mineral grains and even regions within mineral grains can be used.[7] In this way it has become possible to attempt to unravel the timing of the metamorphic history of a rock. In the following sections, we discuss ways to avoid some of the problems, and how to use some of them to obtain information

[7] The SHRIMP permits isotope measurements to be made on very small zones within individual minerals and so can aid determination of the metamorphic history of the mineral.

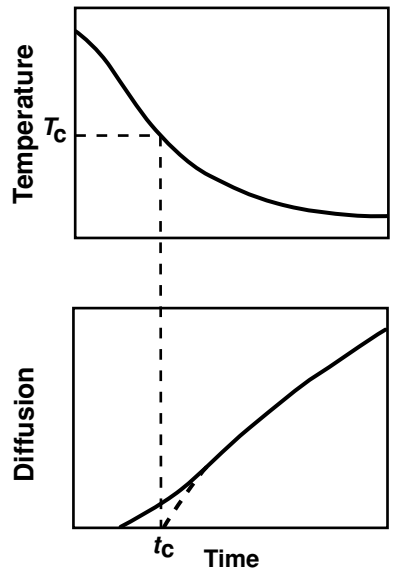


Figure 6.1. The relationship between geochronological closure of a mineral and its cooling history. The upper graph shows the cooling history of a mineral. The vertical axis in the lower graph, the diffusion coefficient D, approximates the rate of escape of the radiogenic daughter product. T_c is the closure temperature, which is attained at time t_c. Each mineral in a rock has a different set of diffusion coefficients, resulting in a variety of closure temperatures and times, as discussed in the text. (From Dodson (1973).)

on the chemical and thermal history of the crust and mantle. Rubidium–strontium dating is discussed first: it is simple and a good system to use to illustrate the problems and pitfalls of geochronology.

6.3 Rubidium–strontium

For the decay of ^{87}Rb to ^{87}Sr, Eq. (6.11) is

$$[^{87}\mathrm{Sr}]_{\mathrm{now}} = [^{87}\mathrm{Rb}]_{\mathrm{now}}(e^{\lambda t} - 1) \tag{6.26}$$

where $[^{87}\mathrm{Sr}]_{\mathrm{now}}$ is the number of ^{87}Sr atoms and $[^{87}\mathrm{Rb}]_{\mathrm{now}}$ is the number of ^{87}Rb atoms, both measured now. Since strontium occurs naturally in rocks independently of rubidium, it is not reasonable to assume that all the $[^{87}\mathrm{Sr}]_{\mathrm{now}}$ is a result of the decay of ^{87}Rb. Equation (6.26) must therefore be modified to include $[^{87}\mathrm{Sr}]_0$, the amount of originally occurring ^{87}Sr:

$$[^{87}\mathrm{Sr}]_{\mathrm{now}} = [^{87}\mathrm{Sr}]_0 + [^{87}\mathrm{Rb}]_{\mathrm{now}}(e^{\lambda t} - 1) \tag{6.27}$$

There are four isotopes of natural strontium with relative atomic masses 84, 86, 87 and 88, which have fractional abundances of about 0.6%, 10%, 7% and 83%, respectively. Rubidium has just two isotopes, ^{85}Rb and ^{87}Rb; ^{87}Rb has a fractional abundance of ~28%.

Since strontium-86 is not a product of any radioactive decay, the amount of strontium-86 present now should be the amount that was originally present:

$$[^{86}\mathrm{Sr}]_{\mathrm{now}} = [^{86}\mathrm{Sr}]_0 \tag{6.28}$$

Normalizing Eq. (6.27) by $[^{86}Sr]_{now}$ gives

$$\frac{[^{87}Sr]_{now}}{[^{86}Sr]_{now}} = \frac{[^{87}Sr]_0}{[^{86}Sr]_{now}} + \frac{[^{87}Rb]_{now}}{[^{86}Sr]_{now}}(e^{\lambda t} - 1)$$
$$\left[\frac{^{87}Sr}{^{86}Sr}\right]_{now} = \left[\frac{^{87}Sr}{^{86}Sr}\right]_0 + \left[\frac{^{87}Rb}{^{86}Sr}\right]_{now}(e^{\lambda t} - 1) \quad (6.29)$$

This is the equation of a straight line with intercept $[^{87}Sr/^{86}Sr]_0$ and slope $e^{\lambda t} - 1$. When the ratios $[^{87}Sr/^{86}Sr]_{now}$ and $[^{87}Rb/^{86}Sr]_{now}$ are measured for various minerals in a rock and plotted against each other, the slope of the resulting line[8] is $e^{\lambda t} - 1$, from which t can be determined. The intercept $[^{87}Sr/^{86}Sr]_0$ is called the *initial ratio*. The straight line is called an *isochron*. Alternatively, from any locality a set of rock samples selected to have varying ratios of rubidium to strontium can be used. An example of such a whole-rock isochron is shown in Fig. 6.2(a).

This method presumes that the initial strontium isotope ratio was the same for all the minerals in the rock or for all the rocks from the chosen locality and means that the age t can be determined without having to make any assumptions about $[^{87}Sr]_0$, the original amount of strontium-87. The rubidium–strontium method became popular in the 1960s because the measurements were accessible to early mass-spectrometer techniques. The main disadvantages are that rubidium and strontium are often mobile – that is, they can be transported in or out of the rock by geochemical processes – and that rubidium does not occur in abundance (in particular, not in limestones and ultramafic rocks); thus, not all rocks are suitable for dating by this method. In addition, because the half-life of ^{87}Rb is very long, it may sometimes be difficult to date very young rocks accurately.

The initial ratio $[^{87}Sr/^{86}Sr]_0$ is an indicator of whether the samples have been derived from the remelting of crustal rocks, or otherwise reset, or are of deeper origin. Since λ is small (1.42×10^{-11} yr^{-1}), λt is very much less than unity for all geological ages. Therefore, a good approximation to Eq. (6.29) is

$$\left[\frac{^{87}Sr}{^{86}Sr}\right]_{now} = \left[\frac{^{87}Sr}{^{86}Sr}\right]_0 + \left[\frac{^{87}Rb}{^{86}Sr}\right]_{now}\lambda t \quad (6.30)$$

When $[^{87}Sr/^{86}Sr]_{now}$ is plotted against t (Fig. 6.2(b)), Eq. (6.30) is the equation of a straight line with a slope $\lambda[^{87}Rb/^{86}Sr]_{now}$ and an intercept $[^{87}Sr/^{86}Sr]_0$, and describes the increase in the strontium ratio with time. Rocks derived from the Archaean mantle have an initial ratio of roughly 0.700, and it is believed that, when the Earth formed, the ratio was 0.699 (the value for some meteorites). The ratio for modern oceanic basalts (which are mantle-derived) is 0.704, and this increase of the initial ratio for strontium is attributed to the accumulation of

[8] The slope and intercept and their errors can be obtained by least–squares fitting of straight lines to data points as described in Appendix 4.

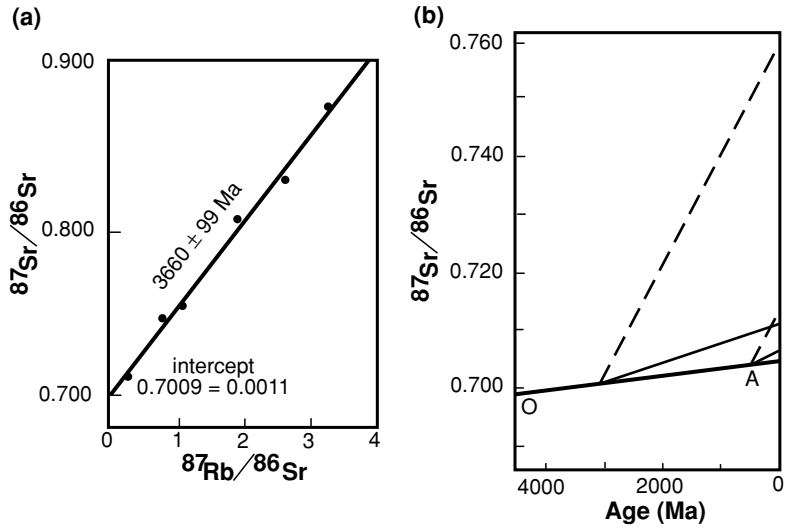


Figure 6.2. (a) A whole-rock rubidium–strontium isochron for the Amitsôq gneiss of the Godthaab district of western Greenland. These rocks are some of the oldest found on the Earth. They were originally granite before undergoing metamorphism to gneiss. The age of 3660 Ma determined by this method probably represents the time of metamorphism. The original granite was therefore older. (After Moorbath *et al.* (1972).) (b) Strontium isotope growth curves for the mantle and crust. The heavy line OA is the mantle growth curve. Crustal growth curves are illustrated for rocks extracted from the mantle 3000 and 500 Ma ago. Gross rubidium–strontium ratios for crustal rocks are 0.10 (solid lines) and 0.50 (dashed lines).

^{87}Sr from the decay of ^{87}Rb. These ratios are consistent with a gross rubidium–strontium ratio for the mantle of about 0.03. The line OA on Fig. 6.2(b) is known as the *mantle growth curve*. Melts, now forming crustal rocks, that were extracted from mantle can be assumed to have the same initial strontium isotope ratio as the mantle at the time of their extraction. Because the gross rubidium–strontium ratio for crustal rocks is greater than that for the mantle (generally between 0.05 and 1.0, compared with 0.03), the rate of increase of the strontium isotope ratio for crustal rocks, the *crustal growth curve*, is much steeper than the mantle growth curve. Equation (6.30) enables us to calculate crustal growth curves, given the gross rubidium–strontium ratio and assuming approximate isotopic proportions. Figure 6.2(b) shows crustal growth curves for rocks derived from the mantle at 3000 and 500 Ma with rubidium–strontium ratios of 0.10 and 0.5. The strontium isotope ratio measured today for these rocks would therefore be considerably more than 0.704.

If a crustal rock is remelted, then the subsequent rock has an initial $^{87}Sr/^{86}Sr$ ratio corresponding to the $^{87}Sr/^{86}Sr$ ratio of the source rock at the time of melting. Clearly this will almost always be considerably *greater* than 0.704, often greater than 0.710. However, rocks derived directly from the mantle have an initial ratio of

less than 0.704. Thus, the initial ratio provides an excellent method of determining the origin of some rocks.

6.4 Uranium–lead

There are three naturally occurring isotopes of uranium: ^{238}U, ^{235}U and ^{234}U. All are radioactive. They have abundances of 99.3%, 0.7% and 0.006%, respectively. For the decay[9] of ^{238}U to ^{206}Pb, Eq. (6.11) is

$$[^{206}\mathrm{Pb}]_{\mathrm{now}} = [^{238}\mathrm{U}]_{\mathrm{now}}(\mathrm{e}^{\lambda_{238}t} - 1) \tag{6.31}$$

and from Eq. (6.12) the age t is given by

$$t = \frac{1}{\lambda_{238}}\log_{\mathrm{e}}\left(1 + \left[\frac{^{206}\mathrm{Pb}}{^{238}\mathrm{U}}\right]_{\mathrm{now}}\right) \tag{6.32}$$

For the decay[10] of ^{235}U to ^{207}Pb, these equations are

$$[^{207}\mathrm{Pb}]_{\mathrm{now}} = [^{235}\mathrm{U}]_{\mathrm{now}}(\mathrm{e}^{\lambda_{235}t} - 1) \tag{6.33}$$

and

$$t = \frac{1}{\lambda_{235}}\log_{\mathrm{e}}\left(1 + \left[\frac{^{207}\mathrm{Pb}}{^{235}\mathrm{U}}\right]_{\mathrm{now}}\right) \tag{6.34}$$

Radiogenic:- a stable product produced by radioactivity

Therefore, provided that the samples come from closed systems that had no initial lead, measurement of the lead–uranium ratios will yield an age. An alternative method of calculating the age is to take the ratio of Eqs. (6.33) and (6.31),

$$\left[\frac{^{207}\mathrm{Pb}}{^{206}\mathrm{Pb}}\right]_{\mathrm{now}} = \left[\frac{^{235}\mathrm{U}}{^{238}\mathrm{U}}\right]_{\mathrm{now}} \frac{\mathrm{e}^{\lambda_{235}t} - 1}{\mathrm{e}^{\lambda_{238}t} - 1} \tag{6.35}$$

where $[^{207}\mathrm{Pb}/^{206}\mathrm{Pb}]_{\mathrm{now}}$ is the present-day ratio of radiogenic ^{207}Pb to radiogenic ^{206}Pb. The present-day universal ratio $[^{235}\mathrm{U}/^{238}\mathrm{U}]_{\mathrm{now}}$ is 1/137.88 and is independent of the age and history of the sample. Equation (6.35) can therefore be used to estimate t from just one sample if the lead ratio is measured. The mineral zircon is the best candidate for dating by this method because it is often retentive of uranium and its daughter products, crystallizes with almost no initial lead and is widely distributed.

6.4.1 Concordia diagrams

Many samples are not ideal and either have initial lead and/or have lost or gained lead or uranium and so require alternative dating methods. An ideal plot of $[^{206}\mathrm{Pb}/^{238}\mathrm{U}]_{\mathrm{now}}$ against $[^{207}\mathrm{Pb}/^{235}\mathrm{U}]_{\mathrm{now}}$ from Eq. (6.35) is known as

[9] The decay series from ^{238}U to stable ^{206}Pb is the 'uranium series'. It involves a chain of fourteen decays, with ^{234}U being one of the intermediate daughters.

[10] The decay series from ^{235}U to stable ^{207}Pb is the 'actinium series'. It involves a chain of eleven decays, with actinium being one of the intermediate daughters.

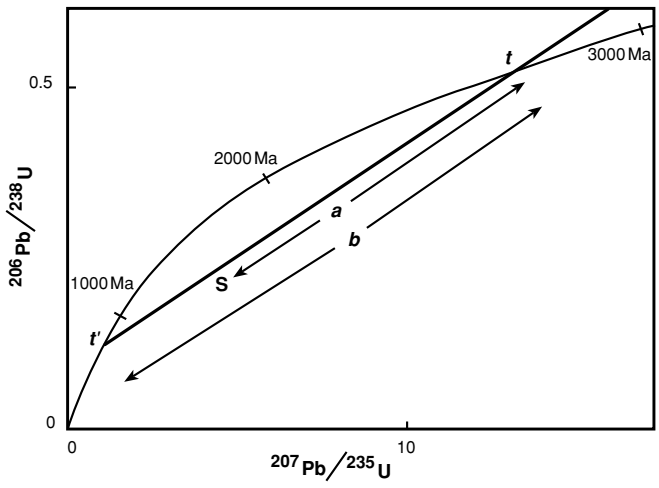


Figure 6.3. The concordia curve as defined by Eq. (6.35). The solid straight line joining points t and t' is the discordia. A sample that crystallized at time t ago and then lost lead and closed at time t' ago will have lead–uranium ratios that plot on this line. A sample that plots at point S must have either lost a fraction a/b of its lead or gained uranium in the disturbing event.

the *concordia* curve (Fig. 6.3) because all the points along the curve have concordant U–Pb ages (i.e., the ^{238}U–^{206}Pb age is the same as the ^{235}U–^{207}Pb age). If a system has been closed to uranium and lead, and the other elements in the decay scheme, since its formation, and if correction is made for initial lead, then the ^{238}U–^{206}Pb and ^{235}U–^{207}Pb methods give concordant ages and the isotope ratios will plot on the concordia. Using this dating method on individual zircon grains (e.g., using ion-probe techniques) often means that concordant results can be obtained and the metamorphic history of the rock determined.

In the event of discordant ages, a concordia plot may be able to yield good estimates both of the ages of the samples and of the time at which the disturbance occurred. Suppose that crystallization of the samples occurred at time t ago and that at time t' ago a disturbing event caused the system to lose lead. Since time t', the system has been closed. Then, if this disturbing event removed ^{207}Pb and ^{206}Pb in the same proportion as that in which they were present in the sample, the $[^{206}Pb/^{238}U]_{now}$ and $[^{207}Pb/^{235}U]_{now}$ ratios will lie on a straight line called the *discordia*, joining points t' and t on the theoretical concordia curve. The position of any particular sample on this straight line depends on the amount of lead the sample lost. Any loss or gain of uranium by the sample during the disturbing event changes the position of the sample on the line (uranium loss plots on the extension of the line above the concordia curve to the right of t, and uranium gain plots below to the left of t) but *does not* alter the intersection of the line with the concordia curve and therefore does not alter the values t' and t. If lead is lost continuously (e.g., by diffusion) as opposed to loss in a single disturbing event, the isotope ratios still plot on a straight line. However, in this case, while t represents the time of crystallization, t' is fictitious and has no meaning. Thus, the stratigraphic and erosional/depositional setting of samples must be known fully

in order to understand their history. This concordia/discordia method provides a way of dating altered rocks.

6.4.2 Isochrons: U–Pb and Pb–Pb

Lead isotopes 204, 206, 207 and 208 all occur naturally, but only ^{204}Pb is non-radiogenic. Equations (6.31)–(6.34) assume that all the lead present is a decay product of uranium. However, we can normalize the equations and correct for an initial, unknown amount of lead, in the same manner as for initial strontium in the rubidium–strontium method:

$$\begin{aligned}\left[\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right]_{\text{now}} &= \left[\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right]_0 + \left[\frac{^{238}\text{U}}{^{204}\text{Pb}}\right]_{\text{now}} (\text{e}^{\lambda_{238}t} - 1)\\ \left[\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right]_{\text{now}} &= \left[\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right]_0 + \left[\frac{^{235}\text{U}}{^{204}\text{Pb}}\right]_{\text{now}} (\text{e}^{\lambda_{235}t} - 1)\end{aligned} \tag{6.36}$$

Thus, *uranium–lead isochrons* can be plotted in exactly the same manner as for rubidium–strontium. This method can fail because extensive losses of uranium can occur in certain geochemical settings (Fig. 6.4(b)).

Taking the ratio of Eqs. (6.36) gives

$$\begin{aligned}\frac{[^{207}\text{Pb}/^{204}\text{Pb}]_{\text{now}} - [^{207}\text{Pb}/^{204}\text{Pb}]_0}{[^{206}\text{Pb}/^{204}\text{Pb}]_{\text{now}} - [^{206}\text{Pb}/^{204}\text{Pb}]_0} &= \left[\frac{^{235}\text{U}}{^{238}\text{U}}\right]_{\text{now}} \frac{\text{e}^{\lambda_{235}t} - 1}{\text{e}^{\lambda_{238}t} - 1}\\ &= \frac{1}{137.88}\frac{\text{e}^{\lambda_{235}t} - 1}{\text{e}^{\lambda_{238}t} - 1}\end{aligned} \tag{6.37}$$

This is the equation of a straight line that passes through the points

$$\left[\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right]_0 \quad \text{and} \quad \left[\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right]_0$$

and has a slope of

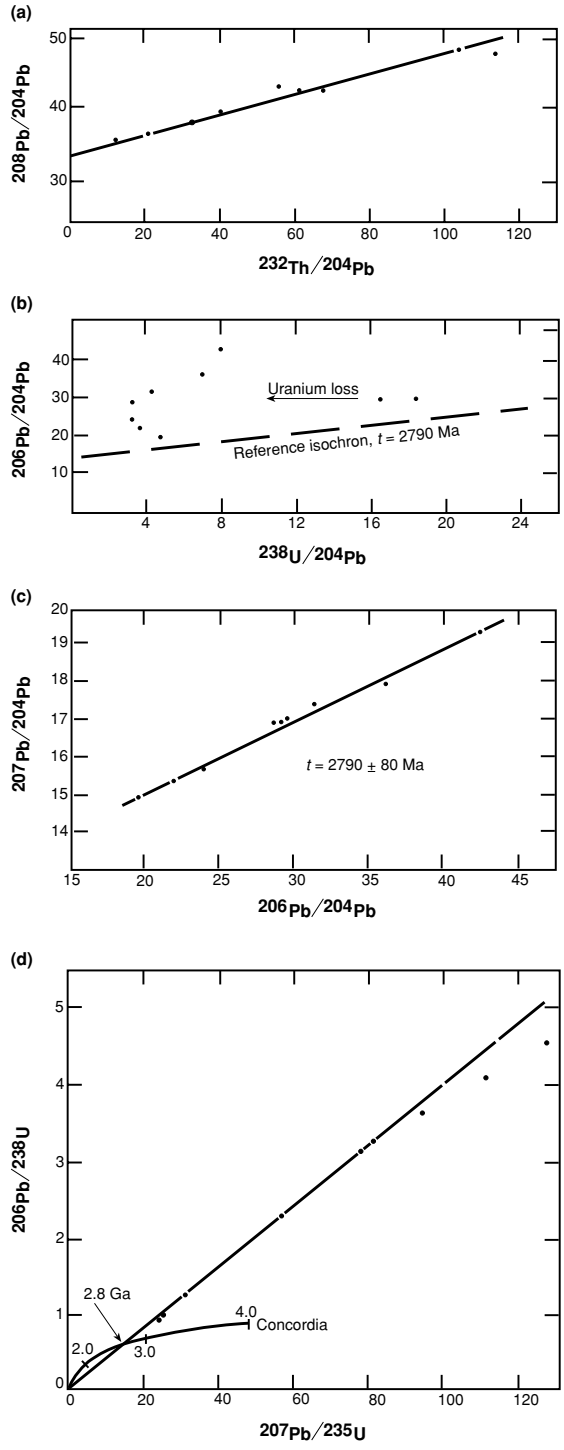
$$\frac{1}{137.88}\frac{\text{e}^{\lambda_{235}t} - 1}{\text{e}^{\lambda_{238}t} - 1}$$

Thus, plotting $[^{207}\text{Pb}/^{204}\text{Pb}]_{\text{now}}$ against $[^{206}\text{Pb}/^{204}\text{Pb}]_{\text{now}}$ gives a straight line, a *lead–lead isochron* (Fig. 6.4(c)). A lead–lead isochron yields a reliable value for t provided that all the samples have the same initial isotope ratios and were closed to uranium and lead at least until recent time. Lead–lead methods are commonly used for dating granulites and limestones.

6.5 Thorium–lead

Applying Eq. (6.11) to the thorium–lead decay (^{232}Th–^{208}Pb) gives

$$[^{208}\text{Pb}]_{\text{now}} = [^{232}\text{Th}]_{\text{now}}(\text{e}^{\lambda\text{t}} - 1) \tag{6.38}$$


(a)
$^{208}Pb/^{204}Pb$
$^{232}Th/^{204}Pb$
(b)
$^{206}Pb/^{204}Pb$
Uranium loss
Reference isochron, t = 2790 Ma
$^{238}U/^{204}Pb$
(c)
$^{207}Pb/^{204}Pb$
t = 2790 ± 80 Ma
$^{206}Pb/^{204}Pb$
(d)
$^{206}Pb/^{238}U$
2.8 Ga
4.0
Concordia
2.0
3.0
$^{207}Pb/^{235}U$

Using ^{204}Pb, the non-radiogenic isotope of lead, in the same way as for the uranium–lead method and including an initial amount of ^{208}Pb, we can write Eq. (6.38) as

$$\left[\frac{^{208}\text{Pb}}{^{204}\text{Pb}}\right]_{\text{now}} = \left[\frac{^{208}\text{Pb}}{^{204}\text{Pb}}\right]_0 + \left[\frac{^{232}\text{Th}}{^{204}\text{Pb}}\right]_{\text{now}} (e^{\lambda t} - 1) \tag{6.39}$$

The construction of an isochron therefore allows t and $[^{208}\text{Pb}/^{204}\text{Pb}]_0$ to be calculated (Fig. 6.4(a)). Alternatively, with an estimate for the initial lead–isotope ratio, Eq. (6.39) gives a value for t directly with one sample. Thorium–lead isochrons can be more successful than uranium–lead isochrons because thorium and lead tend to be less mobile and so less easily lost than uranium.

6.6 Potassium–argon

Potassium has three naturally occurring isotopes, ^{39}K, ^{40}K and ^{41}K. Their relative abundances are 93%, 0.012% and 6.7%, respectively. Equation (6.22) gives the amount of ^{40}Ar produced by ^{40}K decay:

$$[^{40}\text{Ar}]_{\text{now}} = [^{40}\text{K}]_{\text{now}} \frac{\lambda_\text{A}}{\lambda_\text{A} + \lambda_\text{C}} \left(e^{(\lambda_\text{A}+\lambda_\text{C})t} - 1\right) \tag{6.40}$$

Including an initial amount of ^{40}Ar means that the amount of ^{40}Ar measured now at time t is

$$[^{40}\text{Ar}]_{\text{now}} = [^{40}\text{Ar}]_0 + [^{40}\text{K}]_{\text{now}} \frac{\lambda_\text{A}}{\lambda_\text{A} + \lambda_\text{C}} \left(e^{(\lambda_\text{A}+\lambda_\text{C})t} - 1\right) \tag{6.41}$$

←

Figure 6.4. Lead–lead, thorium–lead and uranium–lead whole-rock isochrons and the concordia plot for a granite. (a) Thorium–lead isochron. The slight scatter in the data indicates that the rocks may have lost or gained variable amounts of thorium. However, since the points are not very scattered, the actual quantity of thorium lost or gained was small. (b) Uranium–lead isochron. These isotope ratios are very scattered due to loss of uranium. Had no uranium been lost, the points would fall on the 2790-Ma reference isochron. However, loss of uranium causes the data to plot to the left of their position on the reference isochron, as indicated by the arrow. The amount of uranium lost by each rock can be estimated by assuming that the $^{206}\text{Pb}/^{204}\text{Pb}$ ratio and the reference isochron are correct. The reference $^{238}\text{U}/^{204}\text{Pb}$ ratio can then be compared with the actual value: this gives the fraction of uranium retained by each rock. (c) Lead–lead isochron. The points fall on a straight line with slope 0.1911, which corresponds to an age of 2790 Ma. Plots (a)–(c) show some of the problems which are encountered when using the uranium and thorium decay schemes and some of the ways in which information can be obtained despite the mobility of uranium and thorium. (d) A concordia plot. The losses of uranium mean that the data plot above the concordia curve. The data points lie on a straight line (the discordia) which passes through the origin and intersects the concordia at 2800 Ma. This is consistent with the lead–lead isochron age and with the uranium loss having occurred as the result of recent weathering. (From Faure (1986), after Rosholt *et al.* (1973).)

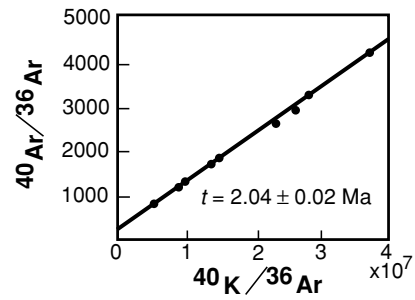


Figure 6.5. A potassium–argon whole-rock isochron for tuffs (volcanic ash) from the Olduvai Gorge in Tanzania. The age of these rocks is very important because it is used to date both recent geomagnetic reversals (as discussed in Section 3.2.1) and the early hominoid remains which have been discovered in the gorge by Leakey and co-workers. (After Fitch *et al.* (1976).)

The non-radiogenic isotope ^{36}Ar is used to normalize this equation:

$$\left[\frac{^{40}Ar}{^{36}Ar}\right]_{now} = \left[\frac{^{40}Ar}{^{36}Ar}\right]_0 + \left[\frac{^{40}K}{^{36}Ar}\right]_{now} \frac{\lambda_A}{\lambda_A + \lambda_C}\left(e^{(\lambda_A+\lambda_C)t} - 1\right) \qquad (6.42)$$

An isochron can be constructed using this equation, which allows t and the initial argon ratio to be estimated (Fig. 6.5). It is often possible to assume that the initial argon ratio $[^{40}Ar/^{36}Ar]_0$ was 295.5, its present-day value in the atmosphere, and so obtain an age from a single, whole rock or mineral, which makes this an attractive dating method. In addition, the shorter half-life of ^{40}K and its relative abundance compared with the previously discussed elements mean that this method is good for dating young rocks. However, problems can arise because argon is a gas and thus easily lost from the system and because potassium is also mobile. Another problem arises if the argon initially present in the sample was not totally of atmospheric origin but incorporated some argon from crustal or mantle outgassing: the calculated age would then be too great. Closure temperatures (see Sections 6.2 and 7.8.5) for argon depend on the mineral involved, ranging between approximately 100 and 500 °C; thus, each mineral gives a different age (see Fig. 7.26 for an example). These differences in closure temperatures lead to a potentially powerful method of unravelling the thermal history of the rocks (discussed in Section 7.8.5).

6.7 Argon–argon

The three naturally occurring isotopes of argon are ^{36}Ar, ^{38}Ar and ^{40}Ar, which are present in the atmosphere in proportions of 0.34%, 0.06% and 99.6%, respectively. The argon–argon dating method depends first on bombarding a sample with fast neutrons in a nuclear reactor, which converts some of the ^{39}K into ^{39}Ar. The reaction is

$$^{39}K + n \rightarrow {}^{39}Ar + p \qquad (6.43)$$

where n is a neutron and p a proton. Following irradiation, the argon must be extracted from the sample. Originally this was achieved by heating step by step in a vacuum. Nowadays a laser is focussed to melt individual mineral grains. This is termed the 'total-fusion' method. Alternatively 'laser-probe' dating releases

argon from spots within individual minerals. At each heating step, the $^{40}Ar/^{39}Ar$ ratio of the gas given off is measured in a mass spectrometer. Laser techniques have the advantage that, since individual minerals (or parts of minerals) can be dated, minerals of differing ages in a sample can be distinguished. Argon-39 is an unstable isotope and decays back to potassium-39, but its half-life of 269 yr is long enough (compared with the time taken to make the analysis) for it to be treated as though it were stable.

The amount of ^{39}Ar present in the sample after it has been irradiated is proportional to the amount of ^{39}K present before it was irradiated in the nuclear reactor:

$$[^{39}\mathrm{Ar}]_{\mathrm{now}} = c[^{39}\mathrm{K}]_{\mathrm{now}} \tag{6.44}$$

where c is the constant of proportionality (a function of the length of time for which the sample was irradiated, the neutron energies and the capture cross section for the reaction).

The number of radiogenic ^{40}Ar atoms present in the sample is given by Eq. (6.40). Therefore, by combining Eqs. (6.44) and (6.40), we can write the $^{40}Ar/^{39}Ar$ ratio as

$$\begin{aligned}\left[\frac{^{40}\mathrm{Ar}}{^{39}\mathrm{Ar}}\right]_{\mathrm{now}} &= \frac{1}{c}\left[\frac{^{40}\mathrm{K}}{^{39}\mathrm{K}}\right]_{\mathrm{now}}\frac{\lambda_{\mathrm{A}}}{\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}}}\left(\mathrm{e}^{(\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}})t}-1\right)\\ &= \frac{\mathrm{e}^{(\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}})t}-1}{J}\end{aligned} \tag{6.45}$$

where

$$J = c\left[\frac{^{39}K}{^{40}K}\right]_{\mathrm{now}}\frac{\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}}}{\lambda_{\mathrm{A}}}$$

If a standard sample of known age t_s is irradiated at the same time as the sample to be studied, the term J can be determined. Measurement of the $^{40}Ar/^{39}Ar$ ratio for this standard sample gives

$$\left[\frac{^{40}\mathrm{Ar}}{^{39}\mathrm{Ar}}\right]^{\mathrm{s}}_{\mathrm{now}} = \frac{\mathrm{e}^{(\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}})t_{\mathrm{s}}}-1}{J} \tag{6.46}$$

where the superscript s refers to measurements on the standard. Rearranging Eq. (6.46) gives

$$J = \left(\mathrm{e}^{(\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}})t_{\mathrm{s}}}-1\right)\left[\frac{^{39}\mathrm{Ar}}{^{40}\mathrm{Ar}}\right]^{\mathrm{s}}_{\mathrm{now}} \tag{6.47}$$

On substituting this value for J back into Eq. (6.45), we obtain

$$\left[\frac{^{40}\mathrm{Ar}}{^{39}\mathrm{Ar}}\right]_{\mathrm{now}} = \frac{\mathrm{e}^{(\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}})t}-1}{\mathrm{e}^{(\lambda_{\mathrm{A}}+\lambda_{\mathrm{C}})t_{\mathrm{s}}}-1}\left[\frac{^{40}\mathrm{Ar}}{^{39}\mathrm{Ar}}\right]^{\mathrm{s}}_{\mathrm{now}} \tag{6.48}$$

An age t can now be calculated from this equation. Corrections for the presence of atmospheric ^{40}Ar can be made to Eq. (6.48).

As already mentioned, the isotope ratios are measured at each step, so an age can be calculated for every temperature or power. If the sample has been closed

both to argon and to potassium since its formation, then the ages determined at every step should be the same. Loss of argon from some minerals or parts of minerals and not others results in a disturbed age spectrum which can sometimes be interpreted to yield the thermal history of the sample. An Ar–Ar isochron can be constructed by plotting $^{39}Ar/^{36}Ar$ against $^{40}Ar/^{36}Ar$. The slope of the isochron is then $(e^{(\lambda_A+\lambda_C)t} - 1)/J$ from Eq. (6.45) and the intercept on the $^{40}Ar/^{39}Ar$ axis is the $^{40}Ar/^{36}Ar$ ratio of non-radiogenic atmospheric argon, normally 295.5.

A problem with very-fine-grained samples is that ^{39}Ar can be lost during the irradiation as it recoils out of the sample during the reaction. This results in a value for t that is greater than the age of the sample.

6.8 Samarium–neodymium

Samarium and neodymium are rare-earth elements that occur in low abundance in most rocks. Although the decay constant for ^{147}Sm–^{143}Nd is very small (see Table 6.2), which makes the decay most useful for dating very old rocks, the system can also be used for younger samples. Applying Eq. (6.11) to this decay and including an initial concentration of ^{143}Nd gives

$$[^{143}Nd]_{now} = [^{143}Nd]_0 + [^{147}Sm]_{now}(e^{\lambda t} - 1) \tag{6.49}$$

One of the seven non-radiogenic isotopes of neodymium, ^{144}Nd, is used as a standard:

$$\left[\frac{^{143}Nd}{^{144}Nd}\right]_{now} = \left[\frac{^{143}Nd}{^{144}Nd}\right]_0 + \left[\frac{^{147}Sm}{^{144}Nd}\right]_{now}(e^{\lambda t} - 1) \tag{6.50}$$

An isochron can therefore give t and the initial Nd ratio. Instrumentation of high precision is required, since neodymium and samarium are frequently present in abundances of less than 10 ppm. Values of $[^{147}Sm/^{144}Nd]_{now}$ are typically 0.1–0.2 in rock samples. In the same manner as was discussed for the rubidium–strontium method, the $[^{143}Nd/^{144}Nd]_0$ initial ratio provides an indication of the origin of the samples. The advantage of this method is that samarium and neodymium are similar in their chemical characteristics and little affected by processes such as weathering and metamorphism; so, even when a system is not closed to rubidium and strontium, the Sm–Nd method can be successful. Since geochemical processes do not preferentially separate samarium and neodymium, large differences in their concentrations are rare. However, Sm–Nd results for some volcanic rocks can reflect contamination of the igneous melt as it ascended through country rock; such results can be very difficult to interpret. Very old rocks such as some meteorites and basalts that contain little K, U or Rb are best dated using the Sm–Nd method.

The abundance of terrestrial ^{143}Nd has increased with time since the Earth's formation because of the decay of ^{147}Sm to ^{143}Nd. The $^{143}Nd/^{144}Nd$ ratio has therefore also increased with time. We can model the time dependence of the $^{143}Nd/^{144}Nd$ ratio by using Eq. (6.50) and by assuming that the $^{147}Sm/^{144}Nd$ ratio of the Earth is the same as that of the chondritic meteorites (Section 6.10).

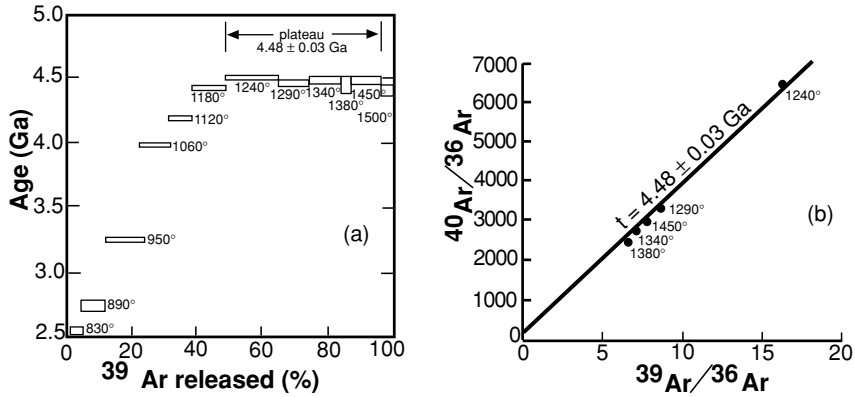


Figure 6.6. (a) The argon age spectrum (sometimes called an argon-release diagram) for the Menow meteorite. The ages from the incremental heating as calculated from Eq. (6.48) are plotted against the percentage of ^{39}Ar released up to that step (the width of the lines indicates the error). About 20% of the ^{39}Ar was released by the third temperature step which gave an 'age' of 3.3 Ga. However, not until the temperature reached 1240 °C and half the ^{39}Ar had been released was a plateau reached. Thereafter the 'age' at ∼4.5, Ga does not change with increasing temperature. Such an age spectrum can be most simply visualized as representing 'age' with increasing depth in the mineral. Any metamorphic event may have caused the mineral to lose argon. The incremental heating essentially samples from progressively deeper into the mineral. Low temperatures release argon from the surface of the mineral. These data will have low ^{40}Ar/^{39}Ar values (and low 'ages') due to the previous loss of ^{40}Ar during metamorphism. Subsequent incremental steps will yield increasingly higher values of ^{40}Ar/^{39}Ar (and higher 'ages') as argon is released from deeper locations that lost less ^{40}Ar during the metamorphism. Ultimately, ideally a 'plateau' is reached, which indicates the original time and 'age' of cooling of the mineral. The high-temperature 'plateau' data points all plot on an isochron (b) giving an age of 4.48 ± 0.06 Ga. The shape of the age spectrum for this meteorite implies that it subsequently lost 25% of its argon at about 2.5 Ga. (After Turner *et al.* (1978).)

For chondritic meteorites, the average present-day ^{147}Sm/^{144}Nd ratio is 0.1967, and the average present-day ^{143}Nd/^{144}Nd ratio is 0.512 638.[11] Therefore, we can rewrite Eq. (6.50) for such chondrites as

$$\left[\frac{^{143}\text{Nd}}{^{144}\text{Nd}}\right]_0 = \left[\frac{^{143}\text{Nd}}{^{144}\text{Nd}}\right]_{\text{now}} - \left[\frac{^{147}\text{Sm}}{^{144}\text{Nd}}\right]_{\text{now}} (e^{\lambda t} - 1)$$
$$= 0.512\,638 - 0.1967(e^{\lambda t} - 1) \qquad (6.51)$$

where the subscript 'now' represents measurements made now and the subscript 0 refers to values time t ago.

[11] This isotope ratio is normalized with respect to the reference isotope ratio ^{146}Nd/^{144}Nd = 0.7219. Various laboratories use slightly different values due to varying methods; this is a problem for isotope geologists using the Sm–Nd method.

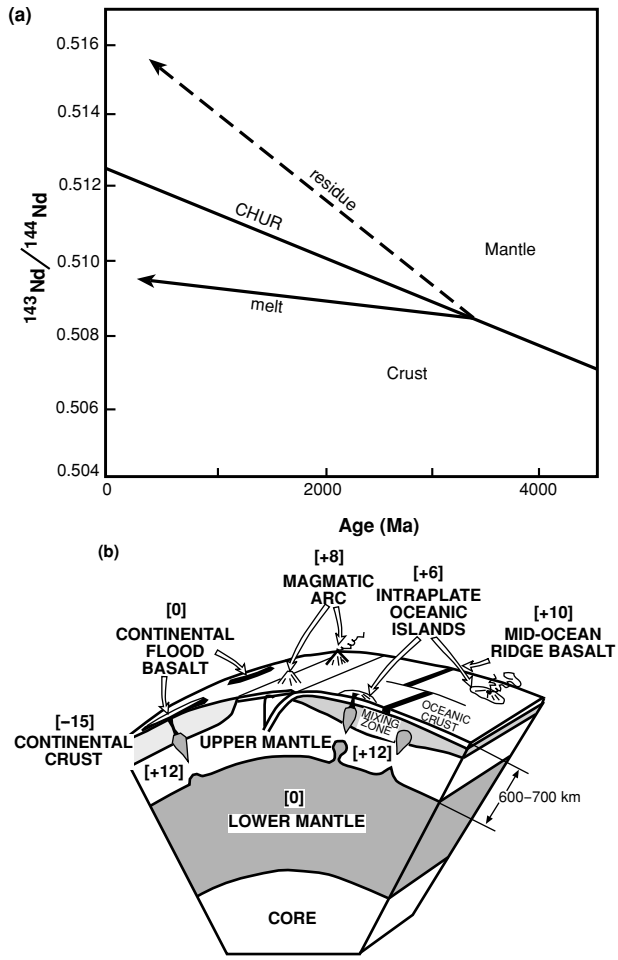


Figure 6.7. (a) The isotopic evolution of neodymium for the CHUR model for the primitive mantle. The straight-line CHUR is defined by Eq. (6.51). If the CHUR undergoes partial melting at, say, 3400 Ma, the melt fraction (solid line) has a lower Sm/Nd ratio (and hence a lower $^{143}Nd/^{144}Nd$ ratio) than that of the CHUR, whereas the solid residue (dashed line) has a higher Sm/Nd ratio (and hence a higher $^{143}Nd/^{144}Nd$ ratio) than that of the CHUR. Thus, the residual mantle plots above the CHUR (positive epsilon) as it evolves, whereas the crust plots below (negative epsilon). (b) A model of the Earth based on neodymium isotopic data. The present ε_{NA} value of the continental crust is about -15, and the value for the upper mantle is $+12$. The lower mantle may be chondritic (i.e., $\varepsilon_{Nd} = 0$ there). (From DePaolo (1981).

The model of the primitive mantle called the chondritic uniform reservoir (CHUR) is used as a standard for mantle isotopic geochemistry. The CHUR is described by Eq. (6.51), and, since the half-life of ^{147}Sm is so long, a plot of the $^{143}Nd/^{144}Nd$ ratio against time is nearly a straight line (Fig. 6.7(a)). Initial $^{143}Nd/^{144}Nd$ ratios are often quoted as deviations from the CHUR model, where

the parameter $\varepsilon_{\mathrm{Nd}}$ (epsilon Nd) is the deviation in parts per ten thousand:

$$\varepsilon_{\mathrm{Nd}} = \left(\frac{[^{143}\mathrm{Nd}/^{144}\mathrm{Nd}]_0}{[^{143}\mathrm{Nd}/^{144}\mathrm{Nd}]_0^{\mathrm{CHUR}}} - 1 \right) \times 10^4 \tag{6.52}$$

where $[^{143}\mathrm{Nd}/^{144}\mathrm{Nd}]_0$ is the initial ratio of a whole-rock sample and $[^{143}\mathrm{Nd}/^{144}\mathrm{Nd}]_0^{\mathrm{CHUR}}$ is the initial neodymium ratio of the CHUR model (calculated from Eq. (6.51)) at the same time t ago. Indeed, $\varepsilon_{\mathrm{Nd}}$ values can be calculated at any time from the formation of a system to the present day: they should always be accompanied by the time at which they were calculated.

As shown in Fig. 6.7, a rock that was derived directly from the CHUR would have an initial ratio equal to the CHUR ratio for that age and hence a zero epsilon value ($\varepsilon_{\mathrm{Nd}} = 0$). However, partial melts of the CHUR develop negative epsilon values as time passes, so a rock derived from the melting of older crustal rocks will have a negative epsilon value ($\varepsilon_{\mathrm{Nd}} < 0$). Likewise, rocks derived from the residue left after the CHUR had been partially molten and the melt withdrawn will have positive epsilon values ($\varepsilon_{\mathrm{Nd}} > 0$). Such a source is generally referred to as *depleted mantle* because it is depleted in the highly incompatible elements such as rubidium and the light rare-earth elements. Equations (6.50) and (6.51) can be used together to calculate a model age T, which represents the time that has elapsed since the neodymium in the rock sample had the same isotopic ratio as that of the CHUR model (i.e., since the rock separated from a chondritic reservoir such as primordial mantle). Such an age has a meaning only provided that the samarium–neodymium ratio of the rock has not been altered since the rock separated from the CHUR mantle. However, the Sm/Nd ratio is relatively insensitive to crustal processes such as metamorphism, weathering and diagenesis, so the model age T can be used, with caution, as a time estimate. Equation (6.50) gives the initial neodymium ratio for a rock as

$$\left[\frac{^{143}\mathrm{Nd}}{^{144}\mathrm{Nd}}\right]_0 = \left[\frac{^{143}\mathrm{Nd}}{^{144}\mathrm{Nd}}\right]_{\mathrm{now}} - \left[\frac{^{147}\mathrm{Sm}}{^{144}\mathrm{Nd}}\right]_{\mathrm{now}} (\mathrm{e}^{\lambda T} - 1) \tag{6.53}$$

From Eq. (6.51), the neodymium ratio in the CHUR at a time T ago was

$$\left[\frac{^{143}\mathrm{Nd}}{^{144}\mathrm{Nd}}\right]_0^{\mathrm{CHUR}} = 0.512\,638 - 0.1967(\mathrm{e}^{\lambda T} - 1) \tag{6.54}$$

Equating these two expressions and rearranging terms to obtain a value for T gives

$$T = \frac{1}{\lambda} \log_{\mathrm{e}} \left(\frac{[^{143}\mathrm{Nd}/^{144}\mathrm{Nd}]_{\mathrm{now}} - 0.512\,638}{[^{147}\mathrm{Sm}/^{144}\mathrm{Nd}]_{\mathrm{now}} - 0.1967} + 1 \right) \tag{6.55}$$

If the rock came from a depleted mantle source instead of the CHUR, such an age estimate would be too low. Since there is little evidence for the existence of any primordial mantle since ~4 Ga, it is generally conventional to calculate model ages with respect to depleted mantle rather than to the CHUR. Age estimates

such as these need to be made with caution and with as much knowledge of the origin and chemistry of the rocks as possible.

The beta decays of rhenium to osmium (^{187}Re to ^{187}Os) and lutetium to hafnium (^{176}Lu to ^{176}Hf) mean that these isotopes can be used as geochronometers in the same way that the decays ^{87}Rb to ^{87}Sr and ^{147}Sm to ^{143}Nd are used. Lutetium, a rare-earth element, occurs in most rocks as a trace element with an abundance of less than 1 ppm. Rhenium and osmium are metals – their abundance is generally very low, less than 10 ppb. However, rhenium occurs in some ores at a sufficient level to be used to date them and has been used to date 0.5-ppm-metallic meteorites. Isochrons and isotopic evolution diagrams are constructed in an analogous manner to the rubidium–strontium and samarium–neodymium methods. A notation similar to Eq. (6.52) can be used to express deviations of the Lu–Hf and Re–Os systems from their bulk reservoirs and these deviations then used in an analogous way to ε_{Nd}. These isotopes, though rarely used yet as geochronometers, are useful in the study of mantle evolution and the origin of magmas and will become more popular in the near future.

6.9 Fission-track dating

As well as undergoing a series of radioactive decays to stable lead-206, uranium-238 is also subject to *spontaneous fission*.[12] That is, the nucleus can disintegrate into two large but unequal parts, releasing two or three neutrons and considerable energy (about 150 MeV). The decay constant for this spontaneous fission of ^{238}U is 8.46×10^{-17} yr^{-1}, very much less than the decay constant for the decay to ^{206}Pb. Thus, fission of ^{238}U occurs only rarely; the ratio of spontaneous fission to α-particle emission is $(8.46 \times 10^{-17})/(1.55 \times 10^{-10}) = 5 \times 10^{-7}$ only. The decay products from the fission of ^{238}U are of such energy that they are able to travel through minerals for about 10 μm (1 μm $= 10^{-6}$ m).

The passage of a charged particle through a solid results in a damaged zone along its path. This is one of the ways by which *cosmic rays*[13] can be studied. If a singly charged particle passes through a photographic emulsion that is subsequently developed, the track of the particle can be seen under the microscope as

[12] Spontaneous fission occurs only in nuclei over a critical size (about atomic number 90). It is the principal decay method for some of the synthesized transuranium elements.

[13] Cosmic rays are very-high–velocity nuclei that constantly bombard the Earth. Their flux is about 1 particle cm^{-2} s^{-1}, about the same energy flux as starlight. Most cosmic rays are protons, and 10% are ^{3}He and ^{4}He. The remainder are heavier elements (with B, C, O, Mg, Si and Fe being prominent); in fact, iron has an abundance of $\sim 3 \times 10^{-4}$ that of the protons by number or $\sim$1.7% by mass. The abundance of the heaviest nuclei, with charge greater than 70, is $\sim 3 \times 10^{-9}$ that of the protons. The boron and similar nuclei are fragments of the original carbon and oxygen that underwent collision with interstellar hydrogen during their 10^{7}-yr journey to Earth at $\sim$90% of the speed of light.

a trail of grains of silver, but in the case of, say, a nucleus of iron (charge 26), one sees a hairy sausage-like cylinder penetrating the emulsion.

Another way of detecting particles uses solid-state track detectors. Tracks can be registered in many important mineral crystals and in a number of commercially available plastic sheets. A highly charged particle that passes through the plastic sheet produces sufficient damage for later etching (usually in NaOH) to reveal the damaged zone, which is dissolved more rapidly than the undamaged material. Under the microscope, one sees two cones, one on each side of the plastic, marking the entry and exit points of the particle. An important feature of all solid-state track detectors is that they have a threshold damage level below which no track is produced. This enables one to detect a minute number of, say, fission particles among a very large number of particles that leave no tracks. This is relevant in detection of spontaneous fission in uranium-containing mineral crystals. One of the most noteworthy results of this technique was the discovery of tracks of spontaneous fission of plutonium-244 in crystals in certain meteorites.[14]

When a surface of a rock or mineral is cut and polished and then etched in a suitable solvent, tracks of these fission products of ^{238}U, or *fission tracks*, are visible under a microscope because the very numerous α particles do not register. Thus, it is possible to use fission tracks to date geological samples. For dating meteorite samples, one has to be concerned about the now-extinct ^{244}Pu; in principle, a correction could be needed for Archaean terrestrial samples also.

Consider a small polished sample of a mineral and assume that at present it has $[^{238}\mathrm{U}]_{\mathrm{now}}$ atoms of ^{238}U distributed evenly throughout its volume. The number of radioactive decays of uranium 238, D_{R}, during time t is given by Eq. (6.11) as

$$D_{\mathrm{R}} = [^{238}U]_{\mathrm{now}}(\mathrm{e}^{\lambda t} - 1) \tag{6.56}$$

where λ is the decay constant for ^{238}U decay.

The number of decays of uranium-238 by spontaneous fission, D_{S}, occurring in time t is then

$$D_{\mathrm{S}} = \frac{\lambda_{\mathrm{S}}}{\lambda}[^{238}\mathrm{U}]_{\mathrm{now}}(\mathrm{e}^{\lambda t} - 1) \tag{6.57}$$

where λ_{S} is the decay constant for the spontaneous fission of ^{238}U. To use Eq. (6.57) to calculate a date t, we must count the visible fission tracks. In addition, we must estimate what proportion of the fission tracks produced in the sample crossed the polished surface and so became visible and therefore countable. We must also measure $[^{238}\mathrm{U}]_{\mathrm{now}}$.

[14] Plutonium was probably initially present on Earth with an abundance $\sim$10% that of uranium. It decays principally by α-particle emission, but in one in 10^4 cases it decays by spontaneous fission. The activity on Earth of plutonium-244 is now essentially extinct since it has a half-life of only 8×10^7 yr. Thus, the proportion remaining is $\sim 2^{-60} = 10^{-20}$.

Fortunately, it is not necessary to carry out the analysis in an absolute manner because another isotope of uranium, ^{235}U, can be made to undergo fission by the absorption of slow neutrons. (Such an *induced fission* is the heart of the generation of power in nuclear reactors and atomic bombs.) The induced fission of ^{235}U is achieved by putting the sample in a reactor and bombarding it with slow neutrons for a specified time (hours). This provides us with a standard against which to calibrate the number of tracks per unit area (track density).

The analysis that follows assumes that neither of the other two isotopes of uranium that occur naturally (^{234}U and ^{235}U) contributes significantly to the spontaneous fissions; we can make this assumption because their fission branching ratios and isotope abundances are very low relative to those of ^{238}U. The analysis also assumes that there has been no previous interaction with neutrons, which would have contributed neutron-induced fission tracks from ^{235}U. This is almost always a safe assumption unless the uranium has been associated with any of the natural thermal nuclear reactors that occurred in uranium mineral deposits in the early Precambrian (an example is the set of natural nuclear reactors at Oklo in Gabon). At that time, ^{235}U was relatively more abundant than it is now since it has a shorter half-life than ^{238}U.

The number of induced fissions of ^{235}U, D_I, is defined as

$$D_I = [^{235}U]_{now}\sigma n \tag{6.58}$$

where σ is the known neutron-capture cross section (the probability that capture of a neutron by ^{235}U will occur) and n is the neutron dose in the reactor (the number of neutrons crossing a square centimetre). Since the fission products of ^{235}U have almost exactly the same average kinetic energy as the fission products of ^{238}U, we can assume that, if uranium-235 is distributed throughout the sample in the same even way as uranium-238, the same proportion of both fission products will cross the sample's polished surface and be counted. This being the case, we can combine Eqs. (6.57) and (6.58) to obtain

$$\frac{\lambda_S}{\lambda}\left[\frac{^{238}U}{^{235}U}\right]_{now}\frac{e^{\lambda t}-1}{\sigma n} = \frac{D_S}{D_I} = \frac{N_S}{N_I} \tag{6.59}$$

where N_S and N_I are the numbers of spontaneous and induced fission tracks counted in a given area. The time t can then be determined by rearranging Eq. (6.59) and using the uranium isotopic ratio $[^{238}U/^{235}U]_{now} = 137.88$:

$$t = \frac{1}{\lambda}\log_e\left(1 + \frac{N_S}{N_I}\frac{\lambda}{\lambda_S}\frac{\sigma n}{137.88}\right) \tag{6.60}$$

In practice, after the number of spontaneous fission tracks N_S has been counted, the sample is placed in the reactor for a specified time and then etched again (which enlarges the original spontaneous fission tracks as well as etching the newly induced fission tracks) so that the number of induced ^{235}U fission tracks N_I can be counted.

One major advantage of using these ancient fission tracks as a geological dating method is that their stability is temperature-dependent (Eq. (6.25)). At

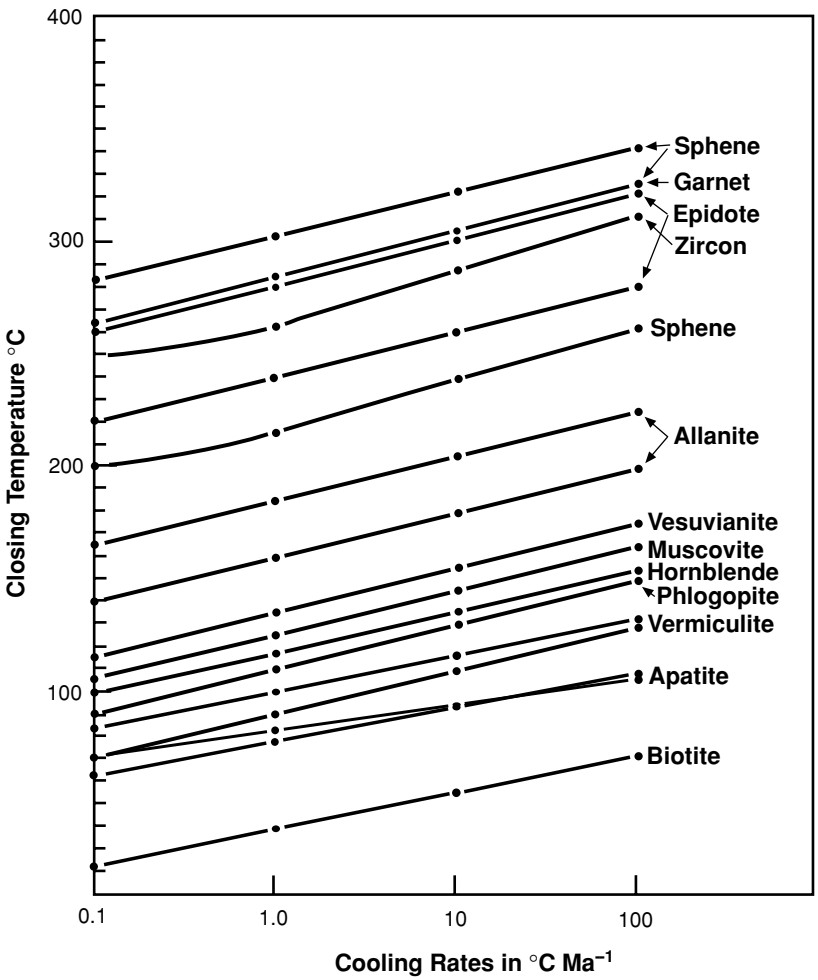


Figure 6.8. Closure temperatures for the retention of fission tracks as a function of cooling rate, for a variety of minerals. (From Faure (1986).)

high temperatures over geological periods of time, the damaged zones in the crystals along the particle track anneal (heal). The rate of annealing differs for every mineral and is temperature-dependent. At room temperatures the tracks are stable. Thus, two minerals of the same age that have been at the same high temperature for the same length of time can yield two different fission-track ages. For example, after 1 Ma at 50 °C a small number of fission tracks in apatite will have annealed, but to anneal all fission tracks within 1 Ma the apatite would need to be at 175 °C. If the heating time is only 10 000 yr, the temperatures required for annealing increase to 75 and 190 °C, respectively. Tracks in the mineral *sphene* can withstand much higher temperatures: annealing will start if the mineral has been at 250 °C for 1 Ma, but not all tracks will completely anneal unless the mineral has been at 420 °C for 1 Ma. The corresponding temperatures for 10 000 yr are 295 and 450 °C. This means that, although fission-track dates can be completely reset by heating, the temperature history of a

particular sample or set of samples can be determined by measuring dates in various minerals.

For the two minerals discussed above, ages determined from fission tracks in sphene are always greater than ages determined from fission tracks in apatite. The ages are interpreted as representing the last time the mineral cooled below its closure temperature (which depends on the cooling rate!). The differences among the closure temperatures of minerals do not, however, depend on the cooling rate. Thus, the difference between the sphene and apatite ages indicates the length of time taken for the sample to cool between the two closure temperatures, and so the cooling rate can be determined. Closure temperatures for fission tracks in a wide range of minerals cooled at various rates are shown in Fig. 6.8.

The temperature dependence of fission tracks provides an excellent method of determining the details of the cooling history of rock samples. This method has been used in the analysis of the erosional history of sedimentary basins.

6.10 The age of the Earth

Some of the early estimates of the age of the Earth were discussed in Section 6.1. Radioactivity provided the tool with which accurate estimates of the Earth's age could be made as well as providing an 'unknown' source of heat that helped to make sense of the early thermal models (see Section 7.4).

The first radioactive dating method used to limit the age of the Earth was the accumulation of α particles (helium nuclei) in minerals as the result of the decay of uranium. In 1905, Rutherford obtained ages of around 500 Ma for the uranium mineral he tested. Also in 1905, Boltwood, as a result of an idea of Rutherford, used the relative proportion of lead and uranium in a rock sample to obtain a date. Measurements on a variety of samples gave dates of between 92 and 570 Ma with the radioactive production rates then available. (This first attempt at U–Pb dating was hampered by the fact that in 1905 neither isotopes nor the thorium–lead decay were understood.) Now, almost a century later, we have a detailed knowledge of the age of rocks and of the Earth based on a variety of radiogenic methods.

The oldest rocks on the surface of the Earth are to be found in the ancient cratons which form the hearts of the continents. The oldest known rocks on earth are the Acasta gneisses in the Slave province of northwestern Canada. U–Pb measurements on zircon grains indicate that the original granitoid parent to this metamorphosed gneiss crystallized at 3962 ± 3 Ma. Amongst the most ancient rocks are the deformed and metamorphosed Isua supracrustal rocks in Greenland, for which the igneous activity has been dated at 3770 ± 42 Ma by Sm–Nd data and at 3769^{+11}_{-8} Ma by U–Pb work on zircons. Felsic volcanic rocks from the Duffer Formation of the Pilbara supergroup in Western Australia have been dated at 3452 ± 16 Ma by U–Pb work on zircons; and, most interesting of all, dates of 4408 ± 8 Ma have been obtained for some detrital zircons from the Jack Hills region of the Yilgarn block in the south of Western Australia (the source of the zircons is unknown). The conclusion to be drawn from all these ancient rocks is that, although they

are rare, they do indicate that, by 4400 Ma and without doubt by 4000 Ma, continents were in existence and the surface temperature was cool enough to have liquid water. The Earth itself is certainly older than these oldest rocks.

Our present knowledge of the age of the Earth comes from a study of the isotopes of lead and from meteorites. First, consider a general model of the lead evolution of the Earth, usually known as the *Holmes–Houtermans model*, after its two independent creators, who built on earlier work by Holmes and by Rutherford. They assumed that, when the Earth was formed, it was homogeneous with a uniform internal distribution of U, Pb and Th. Very soon afterwards, the Earth separated (*differentiated*) into a number of subsystems (e.g., mantle and core), each of which had its own characteristic U/Pb ratio. After this differentiation, the U/Pb ratio in each subsystem changed only as a result of the radioactive decay of uranium and thorium to lead (i.e., each subsystem was closed). Finally, when any lead mineral formed (a common one is galena), its lead separated from all uranium and thorium; so its lead isotopic ratios now are the same as they were at its formation. Applying Eq. (6.37) to this model gives

$$\frac{[^{207}\mathrm{Pb}/^{204}\mathrm{Pb}]_{\mathrm{now}} - [^{207}\mathrm{Pb}/^{204}\mathrm{Pb}]_0}{[^{206}\mathrm{Pb}/^{204}\mathrm{Pb}]_{\mathrm{now}} - [^{206}\mathrm{Pb}/^{204}\mathrm{Pb}]_0} = \frac{1}{137.88}\,\frac{e^{\lambda_{235}T} - e^{\lambda_{235}t}}{e^{\lambda_{238}T} - e^{\lambda_{238}t}} \tag{6.61}$$

where T is the age of the Earth, t is the time since the formation of the lead mineral, the subscript 'now' refers to the isotope ratio of the lead mineral measured now and the subscript 0 refers to the primordial isotope ratio of the Earth time T ago. This is the *Holmes–Houtermans equation*. There are three unknowns in the equation: T, $[^{207}\mathrm{Pb}/^{204}\mathrm{Pb}]_0$ and $[^{206}\mathrm{Pb}/^{204}\mathrm{Pb}]_0$.

Thus, having at least three lead minerals of known age and lead isotope ratios from different subsystems should enable us to determine the age of the Earth and the primordial isotope ratios from Eq. (6.61). Unfortunately, the complex history of the crust and the fact that rocks are frequently not closed to uranium means that in practice T cannot be determined satisfactorily using terrestrial samples. However, meteorites satisfy the criteria of the Holmes–Houtermans model. They are thought to have had a common origin with the planets and asteroids and to have remained a separate subsystem since their separation at the time of formation of the Earth.

Meteorites are fragments of comets and asteroids that hit the Earth. They vary widely in size from dust upwards and can be classified into three main types: *chondrites*, *achondrites* and *iron*. Chondrites are the most primitive and the most common, comprising about 90% of those meteorites observed to fall on Earth. Chondrites are characterized by chondrules (small glassy spheres of silicate), the presence of which indicates that the material was heated, then rapidly cooled and later coalesced into larger bodies. Achondrites are crystalline silicates containing no chondrules and almost no metal phases. Chondrites and achondrites together are termed *stony meteorites*. Some of the chondritic meteorites, termed *carbonaceous chondrites*, are the least metamorphosed of the meteorites and still retain significant amounts of water and other volatiles.

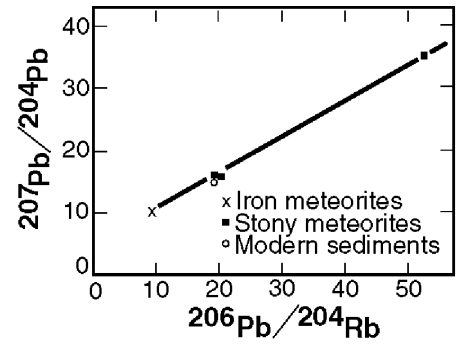


Figure 6.9. Lead–lead isochron for meteorites and recent oceanic sediments. The slope of the straight line gives an age T of 4540 ± 70 Ma for the meteorites. That lead isotopic ratios of recent oceanic sediments also fall on this line indicates that meteorites and the Earth are of the same age and initially contained lead of the same isotopic composition. (From Faure (1986), after Patterson (1956).)

Their composition is believed to be close to the original composition of the solar nebula from which the solar system formed, and thus they provide an initial composition to use for chemical models of the Earth, such as the CHUR (Section 6.8). An asteroid that had undergone partial melting and chemical differentiation into crust, silicate mantle and iron core could fragment into stony and iron meteorites. The stony meteorites, composed primarily of the silicate minerals olivine and pyroxene, are thus similar to the Earth's crust and mantle, whereas the iron meteorites are made up of alloys of iron and nickel, which have been postulated to be present in the core (see Section 8.1.5). Dating stony meteorites by the rubidium–strontium method gives ages of about 4550 Ma (their initial ratio is 0.699). Dates from iron meteorites are similar.

A particular iron sulphide (FeS) phase known as troilite is present in meteorites. Because troilite contains lead but almost no uranium or thorium, its present lead isotopic composition must be close to its original composition. Thus, lead isotope ratios of meteorites can be used as in Eq. (6.37) (Eq. (6.61) with $t = 0$ because meteorites are still closed systems) to construct a lead–lead isochron (as in Fig. 6.4(c)):

$$\frac{[^{207}Pb/^{204}Pb]_{now} - [^{207}Pb/^{204}Pb]_0}{[^{206}Pb/^{204}Pb]_{now} - [^{206}Pb/^{204}Pb]_0} = \frac{1}{137.88}\frac{e^{\lambda_{235}T} - 1}{e^{\lambda_{238}T} - 1} \tag{6.62}$$

This is the equation of a straight line passing through the point ($[^{206}Pb/^{204}Pb]_0$, $[^{207}Pb/^{204}Pb]_0$), with a slope of

$$\frac{1}{137.88}\frac{e^{\lambda_{235}T} - 1}{e^{\lambda_{238}T} - 1}$$

Therefore, plotting the lead isotope ratios of meteorites, $[^{207}Pb/^{204}Pb]_{now}$, against $[^{206}Pb/^{204}Pb]_{now}$ enables the time T to be determined from the slope of the best-fitting straight lines. The first determination (Patterson 1956) using three stony meteorites and two iron meteorites yielded a value for T of 4540 Ma (Fig. 6.9). Many subsequent measurements have been made, all giving an age for the meteorites, and by inference for the Earth, of between 4530 and 4570 Ma.

Using a value for the primordial lead isotope ratios obtained from meteorites (the lead from troilite in the iron meteorite from Canyon Diablo is frequently used because it is the least radiogenic of all the meteorite lead) means that Eq. (6.61) can be applied to terrestrial samples. If a sample's age t is known from other dating methods, T can then be obtained from the present-day lead isotope ratios. Such estimates of T, using the oldest lead ores known, yield values for T of 4520–4560 Ma, in agreement with meteorite ages.

If the meteorites and the Earth are of the same age and initially contained lead of the same isotopic composition, the isotope ratios of average terrestrial lead should lie on the meteorite-lead isochron. Average terrestrial lead is not a straightforward sample to obtain, but oceanic sediments, originating as they do from varied sources, provide an average upper-crustal estimate. In 1956 Patterson showed that the isotope ratios of oceanic sediments lay on the meteorite-lead isochron (Fig. 6.9). This implied that the Earth and the meteorites were the same age and that they had the same primordial lead isotope ratios.

This concept that the Earth and meteorites are of the same age indicates that the ages provided by the meteorite data represent not the age of the solid Earth but rather the time when the parts of the solar system had a uniform isotopic composition and became separate bodies, accreted. The best estimate of this age is 4550 Ma. The differentiation of the Earth into mantle and core (and the degassing of the atmosphere) probably then took place over the next 100 Ma.

Problems

1. Six samples of granodiorite from a pluton in British Columbia, Canada, have strontium and rubidium isotopic compositions as follows.

$^{87}Sr/^{86}Sr$	$^{87}Rb/^{86}Sr$
0.7117	3.65
0.7095	1.80
0.7092	1.48
0.7083	0.82
0.7083	0.66
0.7082	0.74

   (a) Find the age of the intrusion.
   (b) Find the initial $^{87}Sr/^{86}Sr$ ratio of the magma at the time of the intrusion.
   (c) Assuming an $^{87}Sr/^{86}Sr$ ratio of 0.699 and an $^{87}Rb/^{86}Sr$ ratio of 0.1 for the undifferentiated Earth 4550 Ma ago, comment on the possibility that this batholith originated in the mantle.

2. Whole-rock rubidium–strontium isochrons for two plutons are shown in Fig. 6.10. Calculate the age of each pluton and comment on the source of the magma.

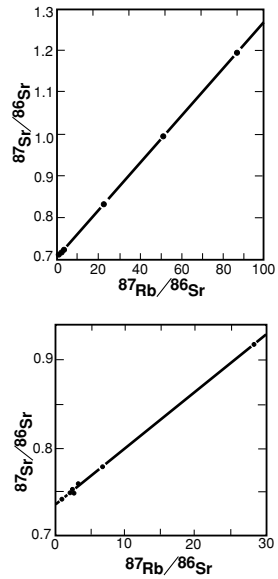


Figure 6.10. A Rubidium–strontium whole-rock isochrons for two plutons. (From Fullagar *et al.* (1971) and Gunner (1974).)

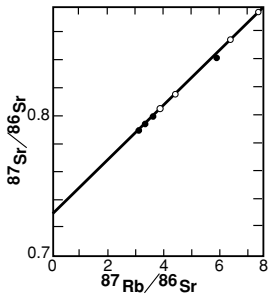


Figure 6.11. Rubidium–strontium isochron for an unmetamorphosed sediment: open circles, whole rock; and solid circles, illite and chlorite. (Data from Gorokhov *et al.* (1981).)

3. Figure 6.11 shows a rubidium–strontium isochron for some unmetamorphosed sediments and clay minerals, namely illite and chlorite, found in the sediment. Calculate an age from this isochron and comment on its meaning.
4. (a) Given that $N = N_0 e^{-\lambda t}$, where N is the number of surviving radioactive atoms at time t, N_0 the initial number and λ the decay constant, and given that today $^{235}U/^{238}U = 0.007\,257$ in the Earth, Moon and meteorites, estimate the date of sudden nucleosynthesis for the two estimated production ratios of $^{235}U/^{238}U = 1.5$ and 2.0. Show all steps in your work clearly. *Note*: $\lambda(^{235}U) = 9.8485 \times 10^{-10}$ yr^{-1} and $\lambda(^{238}U) = 1.55\,125 \times 10^{-10}$ yr^{-1}.
   (b) Given that the solar system had an $^{87}Sr/^{86}Sr$ ratio of 0.699 4.6 Ga ago, what should the initial $^{87}Sr/^{86}Sr$ ratio of the earliest continental crust that formed at 3.8 Ga be, given that the mantle had an $^{87}Rb/^{86}Sr$ ratio of 0.09 at that time and that $\lambda(^{87}Rb) = 1.42 \times 10^{-11}$ yr^{-1}? Show all steps in your work clearly. (Cambridge University Natural Science Tripos IB, 1980.)
5. A granite was extracted from the mantle 2500 Ma ago. Later, at 500 Ma, this granite was remelted. What is the present-day strontium isotope ratio ($^{87}Sr/^{86}Sr$) in the young pluton? Assume that the granites have gross rubidium–strontium ratios of 0.5 and 0.7, respectively.
6. A sample is being dated by the $^{40}Ar/^{39}Ar$-ratio method. Would a delay of one month between the irradiation of the sample and the spectrographic measurements adversely affect the results?
7. Discuss which radioactive dating methods are most appropriate for dating basalt, granite, shale and ultramafic samples (use Table 6.3).
8. Four mineral samples from a meteorite have neodymium and samarium–neodynium isotope ratios as follows.

$^{143}Nd/^{144}Nd$	$^{147}Sm/^{144}Nd$
0.5105	0.12
0.5122	0.18
0.5141	0.24
0.5153	0.28

   (a) Find the age of the meteorite.
   (b) Find the initial $^{143}Nd/^{144}Nd$ ratio for this meteorite.
   (c) Discuss the relevance of the initial $^{143}Nd/^{144}Nd$ ratio of meteorites.
9. Fission-track dating was performed on two minerals in a sample. The track date from garnet was 700 Ma and the date from muscovite was 540 Ma.
   (a) Determine the cooling rate for these minerals in °C Ma^{-1}.
   (b) Determine the closure temperatures for these minerals.
   (c) Assuming that cooling continued at this rate, calculate the track date that would be given by apatite and its closure temperature.
   (d) Would a biotite fission-track date be useful?

10. Fission-track dating on four minerals from one rock sequence yielded the following dates: zircon, 653 Ma; sphene, 646 Ma; hornblende, 309 Ma; and apatite, 299 Ma. Comment on the cooling history of this rock. If you could have one isotopic date determined for these rocks, which method would you choose and why?

References and bibliography

Allègre, C. J., Manhès, G. and Göpel, C. 1995. The age of the Earth. *Geochim. Cosmochim. Acta*, **59**, 1445–56.

Badash, L. 1968. Rutherford, Boltwood and the age of the Earth: the origin of radioactive dating techniques. *Proc. Am. Phil. Soc.*, **112**, 157–69.

1969. *Rutherford and Boltwood: Letters on Radioactivity*. New Haven, Connecticut: Yale University Press.

Curtis, G. H. and Hay, R. L. 1972. Further geological studies and potassium–argon dating at Olduvai Gorge and Ngorongoro Crater. In W. W. Bishop and J. A. Miller, eds., *Calibration of Hominoid Evolution*. Edinburgh: Scottish Academic Press, pp. 289–301.

Dalrymple, G. B. 1991. *The Age of the Earth*. Stanford, California: Stanford University Press.

DePaolo, D. J. 1981. Nd isotopic studies: some new perspectives on earth structure and evolution, *EOS Trans. Am. Geophys. Un*., **62**, 137–40.

Dodson, M. A. 1973. Closure temperature in cooling geochronological and petrological systems. *Contrib. Mineral. Petrol*., **40**, 259–74.

Eicher, D. L. 1976. *Geologic Time*, 2nd edn. Englewood Cliffs, New Jersey: Prentice-Hall.

Eve, A. S. 1939. *Rutherford, Being the Life and Letters of the Rt. Hon. Lord Rutherford, O.M*. Cambridge: Cambridge University Press.

Faul, H. 1966. *Ages of Rocks, Planets and Stars*. New York: McGraw-Hill.

Faure, G. 1986. *Principles of Isotope Geology*, 2nd edn. New York: Wiley.

Fitch, F. J., Miller, J. A. and Hooker, P. J. 1976. Single whole rock K–Ar isochrons. *Geol. Mag*., **113**, 1–10.

Fleischer, R. L., Price, B. and Walker, R. M. 1975. *Nuclear Tracks in Solids*. Berkeley, California: University of California Press.

Fullagar, P. D., Lemmon, R. E. and Ragland, P. C. 1971. Petrochemical and geochronological studies of plutonic rocks in the southern Appalachians: part I. The Salisbury pluton. *Geol. Soc. Am. Bull*., **82**, 409–16.

Ghent, E. D., Stout, M. Z. and Parrish, R. R. 1988. Determination of metamorphic pressure–temperature–time (*P–T–t*) paths. In E. G. Nisbet and C. M. R. Fowler, eds., *Heat, Metamorphism and Tectonics*, Mineralogical Association of Canada Short Course, 14. Toronto: Mineralogical Association of Canada, pp. 155–88.

Gorokhov, I. M., Clauer, N., Varshavskaya, S., Kutyavin, E. P. and Drannik, A. S. 1981. Rb–Sr Ages of Precambrian sediments from the Ovruch Mountain Range, northwestern Ukraine (USSR). *Precambrian Res.*, **16**, 55–65.

Gradstein, F. M. and Ogg, J. 1996. A Phanerozoic timescale. *Episodes*, **19**, nos 1 & 2.

Gradstein, F. M., Ogg, J. G. and Smith, A. G. 2004. *A Geologic Time Scale*. Cambridge: Cambridge University Press.

Gunner, J. D. 1974. Investigations of lower Paleozoic granites in the Beardmore Glacier region. *Antarct. J. U. S.*, **9**, 76–81.

Harland, W. B., Armstrong, R. L., Cox, A. V., Craig, L. F., Smith, A. G. and Smith, D. G. 1990. *A Geologic Time Scale 1989*. Cambridge: Cambridge University Press.

Harrison, T. M. 1987. Comment on 'Kelvin and the age of the Earth'. *J. Geol.*, 94, 725–7.

Jeffreys, H. 1976. *The Earth*, 6th edn. Cambridge: Cambridge University Press.

McNutt, R. H., Crocket, J. H., Clark, A. H., Caelles, J. C., Farrar, E., Haynes, S. J. and Zentilli, M. 1975. Initial $^{87}Sr/^{86}Sr$ ratios of plutonic and volcanic rocks of the central Andes between latitudes 26° and 29° South. *Earth Planet. Sci. Lett*., **27**, 305–13.

Moorbath, S., O'Nions, R. K., Pankhurst, R. J., Gale, N. H. and McGregor, V. R. 1972. Further rubidium–strontium age determinations on the very early Precambrian rocks of the Godthaab district, West Greenland. *Nature Phys. Sci*., **240**, 78–82.

Palmer, A. R. 1983. The decade of North American geology (DNAG) 1983 geologic time scale. *Geology*, **11**, 503–4.

Palmer, A. R. and Geissman, J. 1999. Geologic Time Scale. *Geol. Soc. Am*., CTS004.

Patchett, P. J., White, W. M., Feldmann, H., Kielinczuk, S. and Hofmann, A. W. 1984. Hafnium/rare earth element fractionation in the sedimentary system and crustal recycling into the earth's mantle. *Earth Planet. Sci. Lett*., **69**, 365–75.

Patterson, C. C. 1956. Age of meteorites and the Earth. *Geochem. Cosmochim. Acta*, **10**, 230–7.

Richter, F. M. 1986. Kelvin and the age of the Earth. *J. Geol*., **94**, 395–401.

Rosholt, J. N., Zartman, R. E. and Nkomo, I. T. 1973. Lead isotope systematics and uranium depletion in the Granite Mountains, Wyoming. *Geol. Soc. Am. Bull*., **84**, 989–1002.

Rutherford, E. 1907. Some cosmical aspects of radioactivity. *J. Roy. Astr. Soc. Canada*, May–June, 145–65.

1929. Origin of actinium and the age of the Earth. *Nature*, **123**, 313–14.

Stacey, F. D. 2000. Kelvin's age of the Earth paradox revisited. *J. Geophys. Res*., **105**, 13 155–8.

Steiger, R. H. and Jaeger, E. 1977. Subcommission on geochemistry: convention on the use of decay constants in geo- and cosmochronology. *Earth Planet. Sci. Lett*., **36**, 359–62.

Strutt, Hon. R. J. 1906. On the distribution of radium in the Earth's crust and on the Earth's internal heat. *Proc. Roy. Soc.* A, **77**, 472–85.

Tera, F. 1981. Aspects of isochronism in Pb isotope systematics – application to planetary evolution. *Geochim. Cosmochim. Acta*, **45**, 1439–48.

Thirlwall, M. F. 1991. Long-term reproducibility of multicollector Sr and Nd isotope ratio analysis. *Chem. Geol. (Isot. Geosci. Sect.)*, **94**, 85–104.

Thompson, W. (Kelvin) 1864. On the secular cooling of the Earth. *Trans. Roy. Soc. Edinburgh*, **23**, 157–69.

Turner, G., Enright, M. C. and Cadogan, P. H. 1978. The early history of chondrite parent bodies inferred from ^{40}Ar–^{39}Ar ages. *Proceedings of the 9th Lunar and Planetary Science Conference*. Houston, Texas: Lunar and Planetary Institute, pp. 989–1025.

Wasserburg, G. J. and DePaolo, D. J. 1979. Models of earth structure inferred from neodymium and strontium isotopic abundances. *Proc. Nat. Acad. Sci. U.S.A*., **76**, 3594–8.

Wilde, S. A., Valley, J. W., Peck, W. H. and Graham, C. M. 2001. Evidence from detrital zircons for the existence of continental crust and oceans on Earth 4.4 Gyr ago. *Nature*, **409**, 175–8.

York, D. 1984. Cooling histories from $^{40}Ar/^{39}Ar$ age spectra: implications for Precambrian plate tectonics. *Ann. Rev. Earth Planet. Sci*., **12**, 383–409.

York, D. and Farquhar R. M. 1972. *The Earth's Age and Geochronology.* Oxford: Pergamon.

Chapter 7
Heat

7.1 Introduction

Volcanoes, intrusions, earthquakes, mountain building and metamorphism are all controlled by the transfer and generation of heat. The Earth's thermal budget controls the activity of the lithosphere and asthenosphere as well as the development of the innermost structure of the Earth.

Heat arrives at the Earth's surface from its interior and from the Sun. Virtually all the heat comes from the Sun, as any sunbather knows, but is all eventually radiated back into space. The rate at which heat is received by the Earth, and re-radiated, is about 2×10^{17} W or, averaged over the surface, about 4×10^{2} W m^{-2}. Compare this value with the mean rate of loss of internal heat from the Earth, 4.4×10^{13} W (or 8.7×10^{-2} W m^{-2}); the approximate rate at which energy is released by earthquakes, 10^{11} W; and the rate at which heat is lost by a clothed human body on a very cold ($-30\,^{\circ}$C), windy (10 m s^{-1}) Canadian winter day, 2×10^{3} W m^{-2}. From a geological perspective, the Sun's heat is important because it drives the surface water cycle, the rainfall and, hence, erosion. However, the heat source for igneous intrusion, metamorphism and tectonics is within the Earth, and it is this internal source which accounts for most geological phenomena. The Sun and the biosphere have kept the surface temperature within the range of the stability of liquid water, probably 15–25 $^{\circ}$C averaged over geological time. Given that constraint, the movement of heat derived from the interior has governed the geological evolution of the Earth, controlling plate tectonics, igneous activity, metamorphism, the evolution of the core and hence the Earth's magnetic field.

Heat moves by conduction, convection, radiation and advection. *Conduction* is the transfer of heat through a material by atomic or molecular interaction within the material. In *convection*, heat transfer occurs because the molecules themselves are able to move from one location to another within the material; it is important in liquids and gases. In a room with a hot fire, air currents are set up, which move the light, hot air upwards and away from the fire while dense cold air moves in. Convection is a much faster way of transferring heat than conduction. As an example, when we boil a pan of water on the stove, the heat is transferred through the metal saucepan by conduction but through the water primarily by convection.

Radiation involves direct transfer of heat by electromagnetic radiation (e.g., from the Sun or an electric bar heater). Within the Earth, heat moves predominantly by conduction through the lithosphere (both oceanic and continental) and the solid inner core. Although convection cannot take place in rigid solids, over geological times the Earth's mantle appears to behave as a very-high-viscosity liquid, which means that slow convection is possible in the mantle (see Sections 6.1, 7.4 and 8.2); in fact, heat is generally thought to be transferred by convection through most of the mantle as well as through the liquid outer core. Although hot lava radiates heat, as do crystals at deep, hot levels in the mantle, radiation is a minor factor in the transfer of heat within the Earth. *Advection* is a special form of convection. When a hot region is uplifted by tectonic events or by erosion and isostatic rebound, heat (called advected heat) is physically lifted up with the rocks.

It is not possible to measure temperatures deep in the Earth. Temperatures and temperature gradients can be measured only close to the Earth's surface, usually in boreholes or mines or in oceanic sediments. The deeper thermal structure must be deduced by extrapolation, by inference from seismic observations, from knowledge of the behaviour of materials at high temperatures and pressures, from metamorphic rocks and from models of the distribution of heat production and of the Earth's thermal evolution.

7.2 Conductive heat flow

7.2.1 The heat-conduction equation

Heat, as everyone knows, flows from a hot body to a cold body, not the other way around. The *rate* at which heat is conducted through a solid is proportional to the temperature gradient (the difference in temperature per unit length). Heat is conducted faster when there is a large temperature gradient than when there is a small temperature gradient (all other things remaining constant). Imagine an infinitely long and wide solid plate, d in thickness, with its upper surface kept at temperature T_1 and its lower surface at temperature T_2 ($T_2 > T_1$). The rate of flow of heat per unit area *up* through the plate is proportional to

$$\frac{T_2 - T_1}{d} \tag{7.1}$$

The rate of flow of heat per unit area *down* through the plate, Q, is therefore

$$Q = -k\frac{T_2 - T_1}{d} \tag{7.2}$$

where k, the constant of proportionality, is called the *thermal conductivity*. The thermal conductivity is a physical property of the material of which the plate is made and is a measure of its physical ability to conduct heat. The rate of flow of heat per unit area Q is measured in units of watts per square metre ($\mathrm{W\,m^{-2}}$), and thermal conductivity k is in watts per metre per degree centigrade

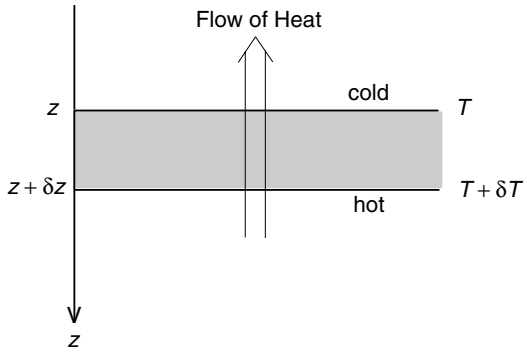


Figure 7.1. Conductive transfer of heat through an infinitely wide and long plate δz in thickness. Heat flows from the hot side of the slab to the cold side (in the negative z direction).

($\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$).[1] Thermal conductivities of solids vary widely: 418 $\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$ for silver; 159 $\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$ for magnesium; 1.2 $\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$ for glass; 1.7–3.3 $\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$ for rock; and 0.1 $\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$ for wood.

To express Eq. (7.2) as a differential equation, let us assume that the temperature of the upper surface (at z) is T and that the temperature of the lower surface (at $z + \delta z$) is $T + \delta T$ (Fig. 7.1). Substituting these values into Eq. (7.2) then gives

$$Q(z) = -k\frac{T + \delta T - T}{\delta z} \tag{7.3}$$

In the limit as $\delta z \to 0$, Eq. (7.3) is written

$$Q(z) = -k\frac{\partial T}{\partial z} \tag{7.4}$$

The minus sign in Eq. (7.4) arises because the temperature is increasing in the positive z direction (i.e., downwards); since heat flows from a hot region to a cold region, it flows in the negative z direction (i.e., upwards).

If we consider Eq. (7.4) in the context of the Earth, z denotes depth beneath the surface. Since z increases downwards, a positive temperature gradient (temperature increases with depth) means that there is a net flow of heat upwards out of the Earth. Measurement of temperature gradients and thermal conductivity in near-surface boreholes and mines can provide estimates of the rate of loss of heat from the Earth.

Consider a small volume of height δz and cross-sectional area a (Fig. 7.2). Any change in temperature δT of this small volume in time δt depends on

1. the flow of heat across the volume's surface (net flow is in or out),
2. the heat generated in the volume and
3. the thermal capacity (specific heat) of the material.

[1] Until fairly recently, the c.g.s. system was used in heat-flow work. In that system, 1 hgu (heat-generation unit) = $10^{-13}\,\mathrm{cal\,cm^{-3}\,s^{-1}} = 4.2 \times 10^{-7}\,\mathrm{W\,m^{-3}}$; 1 hfu (heat-flow unit) = $10^{-6}\,\mathrm{cal\,cm^{-2}\,s^{-1}} = 4.2 \times 10^{-2}\,\mathrm{W\,m^{-2}}$; and thermal conductivity, $0.006\,\mathrm{cal\,cm^{-1}\,s^{-1}\,{}^{\circ}C^{-1}} = 2.52\,\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$.

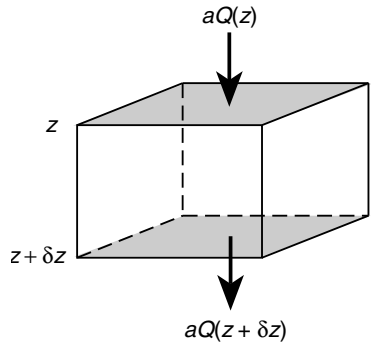


Figure 7.2. A volume element of height δz and cross-sectional area a. Heat is conducted into and out of the element across the shaded faces only. We assume that there is no heat transfer across the other four faces.

The heat per unit time entering the volume across its face at z is $aQ(z)$, whereas the heat per unit time leaving the element across its face at $z + \delta z$ is $aQ(z + \delta z)$. Expanding $Q(z + \delta z)$ in a Taylor series gives

$$Q(z + \delta z) = Q(z) + \delta z \frac{\partial Q}{\partial z} + \frac{(\delta z)^2}{2} \frac{\partial^2 Q}{\partial z^2} + \cdots \tag{7.5}$$

In the Taylor series, the $(\delta z)^2$ term and those of higher order are very small and can be ignored. From Eq. (7.5) the net gain of heat per unit time is

$$\begin{aligned} &\text{heat entering across z} - \text{heat leaving across } z + \delta z \\ &\quad = a\,Q(z) - a\,Q(z + \delta z) \\ &\quad = -a\,\delta z \frac{\partial Q}{\partial z} \end{aligned} \tag{7.6}$$

Suppose that heat is generated in this volume element at a rate A per unit volume per unit time. The total amount of heat generated per unit time is then

$$Aa\,\delta z \tag{7.7}$$

Radioactive heat is the main internal heat source for the Earth as a whole; however, local heat sources and sinks include radioactive heat generation (Section 7.2.2), latent heat, shear heating and endothermic and exothermic chemical reactions. Combining expressions (7.6) and (7.7) gives the total gain in heat per unit time to first order in δz as

$$Aa\,\delta z - a\,\delta z \frac{\partial Q}{\partial z} \tag{7.8}$$

The *specific heat* c_P of the material of which the volume is made determines the rise in temperature due to this gain in heat since specific heat is defined as *the amount of heat necessary to raise the temperature of 1 kg of the material by 1 °C*. Specific heat is measured in units of $\mathrm{W\,kg^{-1}\,{}^{\circ}C^{-1}}$.

If the material has density ρ and specific heat c_P, and undergoes a temperature increase δT in time δt, the rate at which heat is gained is

$$c_P a\,\delta z\,\rho \frac{\delta T}{\delta t} \tag{7.9}$$

Thus equating the expressions (7.8) and (7.9) for the rate at which heat is gained by the volume element gives

$$c_P a\,\delta z\,\rho\frac{\delta T}{\delta t} = Aa\,\delta z - a\,\delta z\frac{\partial Q}{\partial z}$$

$$c_P\rho\frac{\delta T}{\delta t} = A - \frac{\partial Q}{\partial z} \tag{7.10}$$

In the limiting case when δz, $\delta t \to 0$, Eq. (7.10) becomes

$$c_P\rho\frac{\partial T}{\partial t} = A - \frac{\partial Q}{\partial z} \tag{7.11}$$

Using Eq. (7.4) for Q (heat flow per unit area), we can write

$$c_P\rho\frac{\partial T}{\partial t} = A + k\frac{\partial^2 T}{\partial z^2} \tag{7.12}$$

or

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\frac{\partial^2 T}{\partial z^2} + \frac{A}{\rho c_P} \tag{7.13}$$

This is the one-dimensional heat-conduction equation.

In the derivation of this equation, temperature was assumed to be a function solely of time t and depth z. It was assumed not to vary in the x and y directions. If temperature were assumed to be a function of x, y, z and t, a three-dimensional heat-conduction equation could be derived in the same way as this one-dimensional equation. It is not necessary to go through the algebra again: we can generalize Eq. (7.13) to a three-dimensional Cartesian coordinate system as

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) + \frac{A}{\rho c_P} \tag{7.14}$$

Using differential-operator notation (see Appendix 1), we write Eq. (7.14) as

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\nabla^2 T + \frac{A}{\rho c_P} \tag{7.15}$$

Equations (7.14) and (7.15) are known as the *heat-conduction equation*. The term $k/(\rho c_P)$ is known as the *thermal diffusivity* κ. Thermal diffusivity expresses the ability of a material to lose heat by conduction. Although we have derived this equation for a Cartesian coordinate system, we can use it in any other coordinate system (e.g., cylindrical or spherical), provided that we remember to use the definition of the Laplacian operator, ∇^2 (Appendix 1), which is appropriate for the desired coordinate system.

For a steady-state situation when there is no change in temperature with time, Eq. (7.15) becomes

$$\nabla^2 T = -\frac{A}{k} \tag{7.16}$$

In the special situation when there is no heat generation, Eq. (7.15) becomes

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\nabla^2 T \tag{7.17}$$

This is the *diffusion equation* (Section 7.3.5).

So far we have assumed that there is no relative motion between the small volume of material and its immediate surroundings. Now consider how the temperature of the small volume changes with time if it is in relative motion through a region where the temperature varies with depth. This is an effect not previously considered, so Eq. (7.13) and its three-dimensional analogue, Eq. (7.15), must be modified. Assume that the volume element is moving with velocity u_z in the z direction. It is now no longer fixed at depth z; instead, at any time t, its depth is $z + u_z t$. The $\partial \mathrm{T}/\partial \mathrm{t}$ in Eq. (7.13) must therefore be replaced by

$$\frac{\partial T}{\partial t} + \frac{\mathrm{d}z}{\mathrm{d}t}\frac{\partial T}{\partial z}$$

The first term is the variation of temperature with time at a fixed depth z in the region. The second term

$$\frac{\mathrm{d}z}{\mathrm{d}t}\frac{\partial T}{\partial z}$$

is equal to $u_z \partial T/\partial z$ and accounts for the effect of the motion of the small volume of material through the region where the temperature varies with depth. Equations (7.13) and (7.15) become, respectively,

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\frac{\partial^2 T}{\partial z^2} + \frac{A}{\rho c_P} - u_z \frac{\partial T}{\partial z} \tag{7.18}$$

and

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\nabla^2 T + \frac{A}{\rho c_P} - \mathbf{u} \cdot \boldsymbol{\nabla} T \tag{7.19}$$

In Eq. (7.19), $\mathbf{u}$ is the three-dimensional velocity of the material. The term $\mathbf{u} \cdot \boldsymbol{\nabla} T$ is the *advective-transfer* term.

Relative motion between the small volume and its surroundings can occur for various reasons. The difficulty involved in solving Eqs. (7.18) and (7.19) depends on the cause of this relative motion. If material is being eroded from above the small volume or deposited on top of it, then the volume is becoming nearer to or further from the cool surface of the Earth. In these cases, u_z is the rate at which erosion or deposition is taking place. This is the process of advection referred to earlier. On the other hand, the volume element may form part of a thermal-convection cell driven by temperature-induced differences in density. In the latter case, the value of u_z depends on the temperature field itself rather than on an external factor such as erosion rates. The fact that, for convection, u_z is a function of temperature means that Eqs. (7.18) and (7.19) are nonlinear and hence significantly more difficult to solve (Section 8.2.2).

7.2.2 Radioactive heat generation

Heat is produced by the decay of radioactive isotopes (Table 6.2). Those radioactive elements which contribute most to the internal heat generation of the Earth are uranium, thorium and potassium. These elements are present in the crust in

very small quantities, parts per million for uranium and thorium and of the order of 1% for potassium; in the mantle they are some two orders of magnitude less abundant. Nevertheless, these radioactive elements are important in determining the temperature and tectonic history of the Earth. Other radioactive isotopes, such as aluminium-26 and plutonium-244, have been important in the earliest history of the planet.

The radioactive isotopes producing most of the heat generation in the crust are ^{238}U, ^{235}U, ^{232}Th and ^{40}K. The uranium in the crust can be considered to be ^{238}U and ^{235}U, with present-day relative abundances of 99.28% and 0.72%, respectively; but ^{40}K is present at a level of merely one in 10^4 of total potassium (Chapter 6). The radioactive heat generation for these elements in the Earth is therefore as follows: uranium, 9.8×10^{-5} W kg^{-1}; thorium, 2.6×10^{-5} W kg^{-1}; and potassium, 3.5×10^{-9} W kg^{-1}. Table 7.1 gives the radioactive heat generation of some average rock types. It is clear from this table that, on average, the contributions of uranium and thorium to heat production are larger than that of potassium. On average, granite has a greater internal heat generation than do mafic igneous rocks, and the heat generation of undepleted mantle is very low.

The heat generated by these radioactive isotopes when measured today can be used to calculate the heat generated at earlier times. At time t ago, a radioactive isotope with a decay constant λ would have been a factor $e^{\lambda t}$ more abundant than it is today (Eq. (6.5)). Table 7.2 shows the changes in abundance of isotopes and consequent higher heat generation in the past relative to the present.

Although the heat generation of the crust is some two orders of magnitude greater than that of the mantle, the rate at which the Earth as a whole produces heat is influenced by the mantle because the volume of the mantle is so much greater than the total crustal volume. About one-fifth of radioactive heat is generated in the crust. The mean abundances of potassium, thorium and uranium, for the crust and mantle taken together, are in the ranges 150–260 ppm, 80–100 ppb and 15–25 ppb, respectively. These abundances result in a total radioactive heat production for the crust and mantle of $(1.4–2.7) \times 10^{13}$ W, with a best-guess value of 2.1×10^{13} W.

7.3 Calculation of simple geotherms

7.3.1 Equilibrium geotherms

As can be seen from Eq. (7.18), the temperature in a column of rock is controlled by several parameters, some internal and some external to the rock column. The internal parameters are the conductivity, specific heat, density and radioactive heat generation. External factors include heat flow into the column, the surface temperature and the rate at which material is removed from or added to the top of the column (erosion or deposition). Temperature–depth profiles within the Earth are called *geotherms*. If we consider a one-dimensional column with no erosion

Table 7.1 *Typical concentrations of radioactive elements and heat production of some rock types*

	Granite	Tholeiitic basalt	Alkali basalt	Peridotite	Average continental upper crust	Average continental crust	Average oceanic crust	Undepleted mantle
Concentration by weight								
U (ppm)	4	0.1	0.8	0.006	2.8	1.1	0.9	0.02
Th (ppm)	15	0.4	2.5	0.04	10.7	4.2	2.7	0.10
K (%)	3.5	0.2	1.2	0.01	3.4	1.3	0.4	0.04
Heat generation (10^{-10} W kg^{-1})								
U	3.9	0.1	0.8	0.006	2.8	1.1	0.9	0.02
Th	4.1	0.1	0.7	0.010	3.0	1.2	0.7	0.03
K	1.3	0.1	0.4	0.004	1.2	0.5	0.1	0.007
Total	9.3	0.3	1.9	0.020	7.0	2.7	1.7	0.057
Density (10^3 kg m^{-3})	2.7	2.8	2.7	3.2	2.7	2.7	2.9	3.2
Heat generation (μW m^{-3})	2.5	0.08	0.5	0.006	1.8	0.7	0.5	0.02

Table 7.2 *Relative abundances of isotopes and crustal heat generation in the past relative to the present*

	Relative abundance					Heat generation	
Age (Ma)	^{238}U	^{235}U	U[a]	Th	K	Model A[b]	Model B[c]
Present	1.00	1.00	1.00	1.00	1.00	1.00	1.00
500	1.08	1.62	1.10	1.03	1.31	1.13	1.17
1000	1.17	2.64	1.23	1.05	1.70	1.28	1.37
1500	1.26	4.30	1.39	1.08	2.22	1.48	1.64
2000	1.36	6.99	1.59	1.10	2.91	1.74	1.98
2500	1.47	11.4	1.88	1.13	3.79	2.08	2.43
3000	1.59	18.5	2.29	1.16	4.90	2.52	3.01
3500	1.71	29.9	2.88	1.19	6.42	3.13	3.81

[a] This assumes a present-day isotopic composition of 99.2886% ^{238}U and 0.7114% ^{235}U.

[b] Model A, based on Th/U = 4 and K/U = 20 000.

[c] Model B, based on Th/U = 4 and K/U = 40 000.

Source: Jessop and Lewis (1978).

or deposition and a constant heat flow, the column may eventually reach a state of thermal equilibrium in which the temperature at any point is steady. In that case, the temperature–depth profile is called an *equilibrium geotherm*. In this equilibrium situation, $\partial T/\partial t = 0$ and Eq. (7.16) applies:

$$\frac{\partial^2 T}{\partial z^2} = -\frac{A}{k} \tag{7.20}$$

Since this is a second-order differential equation, it can be solved given two boundary conditions. Assume that the surface is at $z = 0$ and that z increases downwards. Let us consider two pairs of boundary conditions. One possible pair is

(i) temperature $T = 0$ at $z = 0$ and
(ii) surface heat flow $Q = -k\,\partial T/\partial z = -Q_0$ at $z = 0$.

The surface heat flow $Q = -Q_0$ is negative because heat is assumed to be flowing upwards out of the medium, which is in the negative z direction. Integrating Eq. (7.20) once gives

$$\frac{\partial T}{\partial z} = -\frac{Az}{k} + c_1 \tag{7.21}$$

where c_1 is the constant of integration. Because $\partial T/\partial z = Q_0/k$ at $z = 0$ is boundary condition (ii), the constant c_1 is given by

$$c_1 = \frac{Q_0}{k} \tag{7.22}$$

Substituting Eq. (7.22) into Eq. (7.21) and then integrating the second time gives

$$T = -\frac{A}{2k}z^2 + \frac{Q_0}{k}z + c_2 \tag{7.23}$$

where c_2 is the constant of integration. However, since $T = 0$ at $z = 0$ was specified as boundary condition (i), c_2 must equal zero. The temperature within the column is therefore given by

$$T = -\frac{A}{2k}z^2 + \frac{Q_0}{k}z \tag{7.24}$$

An alternative pair of boundary conditions could be

(i) temperature $T = 0$ at $z = 0$ and
(ii) heat flow $Q = -Q_d$ at $z = d$.

This could, for example, be used to estimate equilibrium crustal geotherms if d was the depth of the crust/mantle boundary and Q_d was the mantle heat flow into the base of the crust. For these boundary conditions, integrating Eq. (7.20) gives, as before,

$$\frac{\partial T}{\partial z} = -\frac{A}{k}z + c_1 \tag{7.25}$$

where c_1 is the constant of integration. Because $\partial T/\partial z = Q_d/k$ at $z = d$ is boundary condition (ii), c_1 is given by

$$c_1 = \frac{Q_d}{k} + \frac{Ad}{k} \tag{7.26}$$

Substituting Eq. (7.26) into Eq. (7.25) and then integrating again gives

$$T = -\frac{A}{2k}z^2 + \frac{Q_d + Ad}{k}z + c_2 \tag{7.27}$$

where c_2 is the constant of integration. Because $T = 0$ at $z = 0$ was boundary condition (i), c_2 must equal zero. The temperature in the column $0 \leq z \leq d$ is therefore given by

$$T = -\frac{A}{2k}z^2 + \frac{Q_d + Ad}{k}z \tag{7.28}$$

Comparison of the second term in Eq. (7.24) with that in Eq. (7.28) shows that a column of material of thickness d and radioactive heat generation A makes a contribution to the surface heat flow of Ad. Similarly, the mantle heat flow Q_d contributes $Q_d z/k$ to the temperature at depth z.

7.3.2 One-layer models

Figure 7.3 illustrates how the equilibrium geotherm for a model rock column changes when the conductivity, radioactive heat generation and basal heat flow

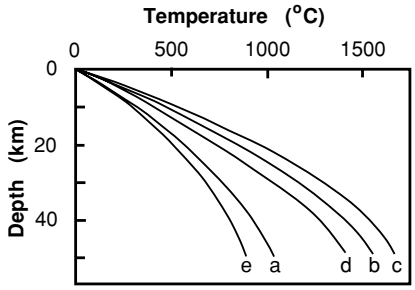


Figure 7.3. Equilibrium geotherms calculated from Eq. (7.28) for a 50-km-thick column of rock. Curve a: standard model with conductivity 2.5 W m^{-1} °C^{-1}, radioactive heat generation 1.25 μW m^{-3} and basal heat flow 21 × 10^{-3} W m^{-2}. Curve b: standard model with conductivity reduced to 1.7 W m^{-1} °C^{-1}. Curve c: standard model with radioactive heat generation increased to 2.5 μW m^{-3}. Curve d: standard model with basal heat flow increased to 42 × 10^{-3} W m^{-2}. Curve e: standard model with basal heat flow reduced to 10.5 × 10^{-3} W m^{-2}. (From Nisbet and Fowler (1982).)

are varied. This model column is 50 km thick, has conductivity 2.5 W m^{-1} °C^{-1}, radioactive heat generation 1.25 μW m^{-3} and a heat flow into the base of the column of 21 × 10^{-3} W m^{-2}. The equilibrium geotherm for this model column is given by Eq. (7.28) and is shown as curve a in Fig. 7.3; at shallow levels the gradient is approximately 30 °C km^{-1}, whereas at deep levels the gradient is 15 °C km^{-1} or less.

Conductivity

Reducing the conductivity of the whole column to 1.7 W m^{-1} °C^{-1} has the effect of increasing the shallow-level gradient to about 45 °C km^{-1} (see curve b in Fig. 7.3). Increasing the conductivity to 3.4 W m^{-1} °C^{-1} would have the opposite effect of reducing the gradient to about 23 °C km^{-1} at shallow levels.

Heat generation

Increasing the heat generation from 1.25 μW m^{-3} to 2.5 μW m^{-3} raises the shallow-level gradient to over 50 °C km^{-1} (curve c in Fig. 7.3); in contrast, reducing the heat generation to 0.4 μW m^{-3} reduces this shallow-level gradient to about 16 °C km^{-1}.

Basal heat flow

If the basal heat flow is doubled from 21 × 10^{-3} to 42 × 10^{-3} W m^{-2}, the gradient at shallow level is increased to about 40 °C km^{-1} (curve d in Fig. 7.3). If the basal heat flow is halved to 10.5 × 10^{-3} W m^{-2}, the shallow-level gradient is reduced to about 27 °C km^{-1} (curve e in Fig. 7.3).

7.3.3 Two-layer models

The models described so far have been very simple, with a 50-km-thick surface layer of uniform composition. This is not appropriate for the real Earth but is a mathematically simple illustration. More realistic models have a layered crust with the heat generation concentrated towards the top (see, e.g., Section 7.6.1). The equilibrium geotherm for such models is calculated exactly as described in Eqs. (7.20)–(7.28) except that each layer must be considered separately and temperature and temperature gradients must be matched across the boundaries.

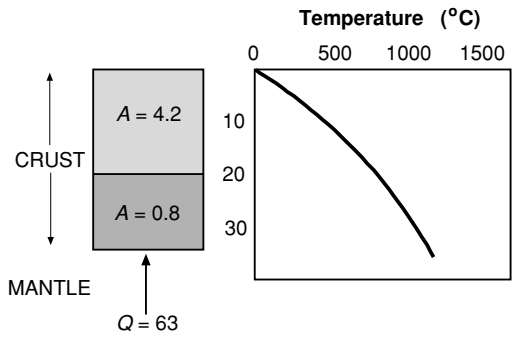


Figure 7.4. A two-layer model for the crust and equilibrium geotherm in the Archaean. Heat generation A is in $\mu W\,m^{-3}$; heat flow from the mantle Q is in $10^{-3}\,W\,m^{-2}$. Recall that, during the Archaean, heat generation was much greater than t is now (Table 7.2). (After Nisbet and Fowler (1982).)

Consider a two-layer model:

$$A = A_1 \quad \text{for } 0 \leq z < z_1$$
$$A = A_2 \quad \text{for } z_1 \leq z < z_2$$
$$T = 0 \quad \text{on } z = 0$$

with a basal heat flow $Q = -Q_2$ on $z = z_2$. In the first layer, $0 \leq z < z_1$, the equilibrium heat-conduction equation is

$$\frac{\partial^2 T}{\partial z^2} = -\frac{A_1}{k} \tag{7.29}$$

In the second layer, $z_1 \leq z < z_2$, the equilibrium heat-conduction equation is

$$\frac{\partial^2 T}{\partial z^2} = -\frac{A_2}{k} \tag{7.30}$$

The solution to these two differential equations, subject to the boundary conditions and matching both temperature, T, and temperature gradient, $\partial T/\partial z$, on the boundary $z = z_1$, is

$$T = -\frac{A_1}{2k}z^2 + \left(\frac{Q_2}{k} + \frac{A_2}{k}(z_2 - z_1) + \frac{A_1 z_1}{k}\right)z \quad \text{for } 0 \leq z < z_1 \tag{7.31}$$

$$T = -\frac{A_2}{2k}z^2 + \left(\frac{Q_2}{k} + \frac{A_2 z_2}{k}\right)z + \frac{A_1 - A_2}{2k}z_1^2 \quad \text{for } z_1 \leq z < z_2 \tag{7.32}$$

Figure 7.4 shows an equilibrium geotherm calculated for a model Archaean crust. The implication is that, during the Archaean, crustal temperatures may have been relatively high (compare with Fig. 7.3.).

7.3.4 The timescale of conductive heat flow

Geological structures such as young mountain belts are not usually in thermal equilibrium because the thermal conductivity of rock is so low that it takes many millions of years to attain equilibrium. For example, consider the model rock column with the geotherm shown as curve a in Fig. 7.3. If the basal heat flow were suddenly increased from 21×10^{-3} to 42×10^{-3} $\mu W\,m^{-2}$, the temperature of the column would increase until the new equilibrium temperatures were attained (curve d in Fig. 7.3). That this process is very slow can be illustrated by considering a rock at depth 20 km. The initial temperature at 20 km would be 567 °C, and, 20 Ma after the basal heat flow increased, conduction would have raised the temperature at 20 km to about 580 °C. Only after 100 Ma would the temperature at 20 km be over 700 °C and close to the new equilibrium value of 734 °C. This can be estimated quantitatively from Eq. (7.17):

$$\frac{\partial \mathrm{T}}{\partial \mathrm{t}} = \kappa \frac{\partial^2 \mathrm{T}}{\partial \mathrm{z}^2}$$

The *characteristic time* $\tau = l^2/\kappa$ gives an indication of the amount of time necessary for a change in temperature to propagate a distance of l in a medium having thermal diffusivity κ. Likewise, the *characteristic thermal diffusion distance*, $l = \sqrt{\kappa\tau}$, gives an indication of the distance that changes in temperature propagate during a time τ. To give a geological example, it would take many tens of millions of years for thermal transfer from a subduction zone at 100 km depth to have a significant effect on the temperatures at shallow depth if all heat transfer were by conduction alone. Hence, melting and intrusion are important mechanisms for heat transfer above subduction zones. As a second example, a metamorphic belt caused by a deep-seated heat source is characterized by abundant intrusions, often of mantle-derived material; this is the dominant factor in transfer of heat to the surface. Magmatism occurs because large increases in the deep heat flow cause large-scale melting at depth long before the heat can penetrate very far towards the surface by conduction.

When a rock column is assembled by some process such as sedimentation, overthrusting or intrusion, the initial temperature gradient is likely to be very different from the equilibrium gradient. This should always be borne in mind when evaluating thermal problems.

7.3.5 Instantaneous cooling or heating

Assume that there is a semi-infinite solid with an upper surface at $z = 0$, no heat generation ($A = 0$) and an initial temperature throughout the solid of $T = T_0$. For $t > 0$, let the surface be kept at temperature $T = 0$. We want to determine how the interior of the solid cools with time.

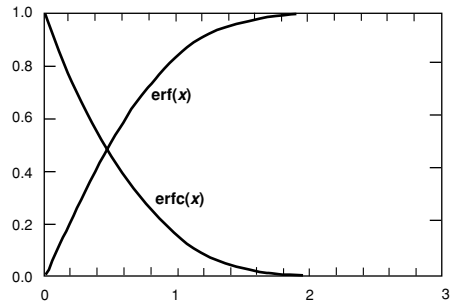


Figure 7.5. The error function erf(x) and complementary error function erfc(x).

The differential equation to be solved is Eq. (7.13) with $A = 0$, the *diffusion equation*:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} \tag{7.33}$$

where $\kappa = k/(\rho c_P)$ is the thermal diffusivity.

Derivation of the solution to this problem is beyond the scope of this book, and the interested reader is referred to Carslaw and Jaeger (1959), Chapter 2, or Turcotte and Schubert (2002), Chapter 4. Here we merely state that the solution of this equation which satisfies the boundary conditions is given by an *error function* (Fig. 7.5 and Appendix 5):

$$T = T_0 \operatorname{erf}\left(\frac{z}{2\sqrt{\kappa t}}\right) \tag{7.34}$$

The error function is defined by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-y^2} \mathrm{d}y \tag{7.35}$$

You can check that Eq. (7.34) is a solution to Eq. (7.33) by differentiating with respect to t and z. Equation (7.34) shows that the time taken to reach a given temperature is proportional to z^2 and inversely proportional to κ.

The temperature gradient is given by differentiating Eq. (7.34) with respect to z:

$$\begin{aligned} \frac{\partial T}{\partial z} &= \frac{\partial}{\partial z}\left[T_0 \operatorname{erf}\left(\frac{z}{2\sqrt{\kappa t}}\right)\right] \\ &= T_0 \frac{2}{\sqrt{\pi}} \frac{1}{2\sqrt{\kappa t}} \mathrm{e}^{-z^2/(4\kappa t)} \\ &= \frac{T_0}{\sqrt{\pi \kappa t}} \mathrm{e}^{-z^2/(4\kappa t)} \end{aligned} \tag{7.36}$$

This error-function solution to the heat-conduction equation can be applied to many geological situations. For solutions to these problems, and numerous others, the reader is again referred to Carslaw and Jaeger (1959).

For example, imagine a dyke of width $2w$ and of infinite extent in the y and z directions. If we assume that there is no heat generation and that the dyke has an initial temperature of T_0, and if we ignore latent heat of solidification, then the differential equation to be solved is

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2}$$

with initial conditions

(i) $T = T_0$ at $t = 0$ for $-w \leq x \leq w$ and
(ii) $T = 0$ at $t = 0$ for $|x| > w$.

The solution of this equation which satisfies the initial conditions is

$$T(x, t) = \frac{T_0}{2}\left[\text{erf}\left(\frac{w - x}{2\sqrt{\kappa t}}\right) + \text{erf}\left(\frac{w + x}{2\sqrt{\kappa t}}\right)\right] \tag{7.37}$$

If the dyke were 2 m in width ($w = 1$ m) and intruded at a temperature of 1000 °C and if κ were 10^{-6} m^2 s^{-1}, then the temperature at the centre of the dyke would be about 640 °C after one week, 340 °C after one month and only about 100 °C after one year! Clearly, a small dyke cools very rapidly.

For the general case, the temperature in the dyke is about $T_0/2$ when $t = w^2/\kappa$ and about $T_0/4$ when $t = 5\text{w}^2/\kappa$. High temperatures outside the dyke are confined to a narrow contact zone: at a distance w away from the edge of the dyke the highest temperature reached is only about $T_0/4$. Temperatures close to $T_0/2$ are reached only within about $w/4$ of the edge of the dyke.

Example: periodic variation of surface temperature

Because the Earth's surface temperature is not constant but varies periodically (daily, annually, ice ages), it is necessary to ensure that temperature measurements are made deep enough that distortion due to these surface periodicities is minimal. We can model this periodic contribution to the surface temperature as $T_0 e^{i\omega t}$, where ω is 2π multiplied by the frequency of the temperature variation, i is the square root of -1 and T_0 is the maximum variation of the mean surface temperature. The temperature $T(z, t)$ is then given by Eq. (7.13) (with $A = 0$) subject to the following two boundary conditions:

(i) $T(0, t) = T_0 e^{i\omega t}$ and
(ii) $T(z, t) \to 0$ as $z \to \infty$.

We can use the separation-of-variables technique to solve this problem. Let us assume that the variables z and t can be separated and that the temperature can be written as

$$T(z, t) = V(z)W(t) \tag{7.38}$$

This supposes that the periodic nature of the temperature variation is the same at all depths as it is at the surface, but it allows the magnitude and phase of the variation

to be depth-dependent, which seems reasonable. Substitution into Eq. (7.13) (with $A = 0$) then yields

$$V\,\frac{\mathrm{d}W}{\mathrm{d}t} = \frac{k}{\rho c_P} W\,\frac{\mathrm{d}^2 V}{\mathrm{d}z^2} \tag{7.39}$$

which, upon rearranging, becomes

$$\frac{1}{W}\frac{\mathrm{d}W}{\mathrm{d}t} = \frac{k}{\rho c_P}\frac{1}{V}\frac{\mathrm{d}^2 V}{\mathrm{d}z^2} \tag{7.40}$$

Because the left-hand side of this equation is a function of z alone and the right-hand side is a function of t alone, it follows that each must equal a constant, say, c_1. However, substitution of Eq. (7.38) into the boundary conditions (i) and (ii) yields, respectively,

$$W(t) = \mathrm{e}^{\mathrm{i}\omega t} \tag{7.41}$$

and

$$V(z) \to 0 \quad \text{as } z \to \infty \tag{7.42}$$

Boundary condition (i) therefore means that the constant c_1 must be equal to $\mathrm{i}\omega$ (differentiate Eq. (7.41) to check this). Substituting Eq. (7.41) into Eq. (7.40) gives the equation to be solved for $V(z)$:

$$\frac{\mathrm{d}^2 V}{\mathrm{d}z^2} = \frac{\mathrm{i}\omega\rho c_P V}{k} \tag{7.43}$$

This has the solution

$$V(z) = c_2\mathrm{e}^{-qz} + c_3\mathrm{e}^{qz} \tag{7.44}$$

where $q = (1+\mathrm{i})\sqrt{\omega\rho c_P/(2k)}$ (remember that $\sqrt{\mathrm{i}} = (1+\mathrm{i})/\sqrt{2}$) and c_2 and c_3 are constants. Equation (7.37), boundary condition (ii), indicates that the positive exponential solution is not allowed; the constant c_3 must be zero. Boundary condition (i) indicates that the constant c_2 is T_0; so, finally, $T(z, t)$ is given by

$$\begin{aligned} T(z,t) &= T_0 \exp(\mathrm{i}\omega t)\exp\left(-(1+\mathrm{i})\sqrt{\frac{\omega\rho c_P}{2k}}z\right) \\ &= T_0 \exp\left(-\sqrt{\frac{\omega\rho c_P}{2k}}z\right)\exp\left[\mathrm{i}\left(\omega t - \sqrt{\frac{\omega\rho c_P}{2k}}z\right)\right] \end{aligned} \tag{7.45}$$

For large z this periodic variation dies out. Thus, temperatures at great depth are unaffected by the variations in surface temperatures, as required by boundary condition (ii).

At a depth of

$$L = \sqrt{\frac{2k}{\omega\rho c_P}} \tag{7.46}$$

the periodic disturbance has an amplitude $1/\mathrm{e}$ of the amplitude at the surface. This depth L is called the skin depth. Taking $k = 2.5\,\mathrm{W\,m^{-1}\,{}^\circ C^{-1}}$, $c_P = 10^3\,\mathrm{J\,kg^{-1}\,{}^\circ C^{-1}}$ and $\rho = 2.3 \times 10^3\,\mathrm{kg\,m^{-3}}$, which are reasonable values for a sandstone, then for the daily variation ($\omega = 7.27 \times 10^{-5}\,\mathrm{s^{-1}}$), L is approximately 17 cm; for the annual

variation ($\omega = 2 \times 10^{-7}\ \mathrm{s}^{-1}$), L is 3.3 m; and for an ice age (with period of the order of 100 000 yr), L is greater than 1 km. Therefore, provided that temperature measurements are made at depths greater than 10–20 m, the effects of the daily and annual surface temperature variation are negligible. The effects of ice ages cannot be so easily ignored and must be considered when borehole measurements are made. Measurement of temperatures in ocean sediments is not usually subject to these constraints, the ocean-bottom temperature being comparatively constant.

Equation (7.45) shows that there is a phase difference ϕ between the surface temperature variation and that at depth z,

$$\phi = \sqrt{\frac{\omega \rho c_P}{2k}}\, z \tag{7.47}$$

At the skin depth, this phase difference is one radian. When the phase difference is π, the temperature at depth z is exactly half a cycle out of phase with the surface temperature.

7.4 Worldwide heat flow: total heat loss from the Earth

The total present-day worldwide rate of heat loss by the Earth is estimated to be (4.2–4.4) $\times$ 10^{13} W. Table 7.3 shows how this heat loss is distributed by area: 71% of this heat loss occurs through the oceans (which cover 60% of the Earth's surface). Thus, most of the heat loss results from the creation and cooling of oceanic lithosphere as it moves away from the mid-ocean ridges. Plate tectonics is a primary consequence of a cooling Earth. Conversely, it seems clear that the mean rate of plate generation is determined by some balance between the total rate at which heat is generated within the Earth and the rate of heat loss at the surface. Some models of the thermal behaviour of the Earth during the Archaean (before 2500 Ma) suggest that the plates were moving around the surface of the Earth an order of magnitude faster then than they are today. Other models suggest less marked differences from the present. The heat generated within the Archaean Earth by long-lived radioactive isotopes was probably three-to-four times greater than that generated now (see Table 7.2). A large amount of heat also has been left over from the gravitational energy that was dissipated during accretion of the Earth (see Problem 23) and from short-lived but energetic isotopes such as ^{26}Al, which decayed during the first few million years of the Earth's history.

Evidence from Archaean lavas that were derived from the mantle suggests that the Earth has probably cooled by several hundred degrees since the Archaean as the original inventory of heat has dissipated. The Earth is gradually cooling, and the plates and rates of plate generation may be slowing to match. Presumably, after many billion years all plate motion will cease.

Measured values of heat flow depend on the age of the underlying crust, be it oceanic or continental (Figs. 7.6 and 7.11). Over the oceanic crust the heat flow generally decreases with age: the highest and very variable measurements occur

Table 7.3 *Heat loss and heat flow from the Earth*

	Area (10^6 km^2)	Mean heat flow (10^3 W m^{-2})	Heat loss (10^{12} W)
Continents (post-Archaean)	142	63	9.0
Archaean	13	52	0.7
Continental shelves	46	78	3.5
Total continental area	201	65 ± 1.6	13.1 ± 0.3
Oceans (including marginal basins)	309	101 ± 2.2	31.2 ± 0.7
Worldwide total	510	87 ± 2.0	44.2 ± 1.0

Note: The estimate of convective heat transport by plates is ~65% of the total heat loss; this includes lithospheric creation in oceans and magmatic activity on continents. The estimate of heat loss as a result of radioactive decay in the crust is ~17% of the total heat loss. Although oceanic regions younger than 66 Ma amount to one-third of the Earth's surface area, they account for over half the total global heat loss. About one-third of the heat loss in oceanic regions is by hydrothermal flow. The estimate of the heat loss of the core is 10^{12}–10^{13} W; this is a major heat source for the mantle.
Source: Pollack *et al.* (1993).

over the mid-ocean ridges and young crust, and the lowest values are found over the deep ocean basins. In continental regions the highest heat flows are measured in regions that are subject to the most recent tectonic activity, and the lowest values occur over the oldest, most stable regions of the Precambrian Shield. These oceanic and continental heat-flow observations and their implications are discussed in Sections 7.5 and 7.6.

To apply the heat-conduction equation (7.15) to the Earth as a whole, we need to use spherical polar coordinates (r, θ, ϕ) (refer to Appendix 1). If temperature is not a function of θ or ϕ but only of radius r, Eq. (7.15) is

$$\frac{\partial T}{\mathrm{d}t} = \frac{k}{\rho c_P}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\mathrm{d}r}\right) + \frac{A}{\rho c_P} \tag{7.48}$$

First let us assume that there is no internal heat generation. The equilibrium temperature is then the solution to

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) = 0 \tag{7.49}$$

On integrating once, we obtain

$$r^2\frac{\partial T}{\partial r} = c_1 \tag{7.50}$$

and integrating the second time gives

$$T = -\frac{c_1}{r} + c_2 \tag{7.51}$$

where c_1 and c_2 are the constants of integration.

Now impose the boundary conditions for a *hollow sphere* $b < r < a$:

(i) zero temperature $T = 0$ at the surface $r = a$ and
(ii) constant heat flow $Q = -k\partial T/\partial r = Q_b$ at $r = b$.

The temperature in the spherical shell $b < r < a$ is then given by

$$T = -\frac{Q_b b^2}{k}\left(\frac{1}{a} - \frac{1}{r}\right) \tag{7.52}$$

An expression such as this could be used to estimate a steady temperature for the lithosphere. However, since the thickness of the lithosphere is very small compared with the radius of the Earth, $(a - b)/a \ll 1$, this solution is the same as the solution to the one-dimensional equation (7.28) with $A = 0$.

There is no non-zero solution to Eq. (7.49) for the whole sphere, which has a finite temperature at the origin ($r = 0$). However, there is a steady-state solution to Eq. (7.48) with constant internal heat generation A within the sphere:

$$0 = \frac{k}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) + A$$

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) = -\frac{Ar^2}{k} \tag{7.53}$$

On integrating twice, the temperature is given by

$$T = -\frac{Ar^2}{6k} - \frac{c_1}{r} + c_2 \tag{7.54}$$

where c_1 and c_2 are the constants of integration.

Let us impose the following two boundary conditions:

(i) T finite at $r = 0$ and
(ii) $T = 0$ at $r = a$.

Then Eq. (7.54) becomes

$$T = \frac{A}{6k}(a^2 - r^2) \tag{7.55}$$

and the heat flux is given by

$$-k\frac{dT}{dr} = \frac{Ar}{3} \tag{7.56}$$

The surface heat flow (at $r = a$) is therefore equal to $Aa/3$.

If we model the Earth as a solid sphere with constant thermal properties and uniform heat generation, Eqs. (7.55) and (7.56) yield the temperature at the centre of this model Earth, given a value for the surface heat flow. Assuming values of surface heat flow $80 \times 10^{-3}\,\mathrm{W\,m^{-2}}$, $a = 6370\,\mathrm{km}$ and $k = 4\,\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$, we obtain a temperature at the centre of this model solid Earth of

$$T = \frac{80 \times 10^{-3} \times 6370 \times 10^3}{2 \times 4}$$
$$= 63\,700\ {}^{\circ}\mathrm{C}$$

This temperature is clearly too high for the real Earth because the temperature at the visible surface of the Sun is only about 5700 °C – the Earth is not a star. The model is unrealistic since, in fact, heat is not conducted but convected through the mantle, and the heat-generating elements are concentrated in the upper crust rather than being uniformly distributed throughout the Earth. These facts mean that the actual temperature at the centre of the Earth is much lower than this estimate. Convection is important because it allows the surface heat flow to exploit the entire internal heat of the Earth, instead of just the surface portions of a conductive Earth (see Section 6.1 for Kelvin's conduction calculation of the age of the Earth).

Another fact that we have neglected to consider is the decrease of the radioactive heat generation with time. Equation (7.48) can be solved for an exponential time decay and a non-uniform distribution of the internal heat generation; the temperature solutions are rather complicated (see Carslaw and Jaeger (1959), Section 9.8, for some examples) and still are not applicable to the Earth because heat is convected through mantle and outer core rather than conducted.

It is thought that the actual temperature at the centre of the Earth is about 7000 °C, on the basis of available evidence: thermal and seismic data, laboratory behaviour of solids at high temperatures and pressures and laboratory melting of iron-rich systems at high pressures.

7.5 Oceanic heat flow

7.5.1 Heat flow and the depth of the oceans

Figure 7.6 shows the mean heat flow as a function of age for the three major oceans. The average heat flow is higher over young oceanic crust but exhibits a much greater standard deviation than does that over the older ocean basins. This decrease of heat flow with increasing age is to be expected if we consider hot volcanic material rising along the axes of the mid-ocean ridges and plates cooling as they move away from the spreading centres. The very scattered heat-flow values measured over young oceanic crust are a consequence of the hydrothermal circulation of sea water through the crust (which is also discussed in Section 9.4.4). Heat flow is locally high in the vicinity of hot-water vents and low where cold sea water enters the crust. Water temperatures approaching 400 °C have been measured at the axes of spreading centres by submersibles, and the presence of hot springs on Iceland (which is located on the Reykjanes Ridge) and other islands or regions proximal to spreading centres is well known. As will be discussed in Chapter 9, the oceanic crust is formed by the intrusion of basaltic magma from below. Contact with sea water causes rapid cooling of the new crust, and many cracks form in the lava flows and dykes. Convection of sea water through the cracked crust occurs, and it is probable that this circulation penetrates through most of the crust, providing an efficient cooling mechanism (unless you drive a Volkswagen, your car's engine is cooled in the same manner). As the newly formed

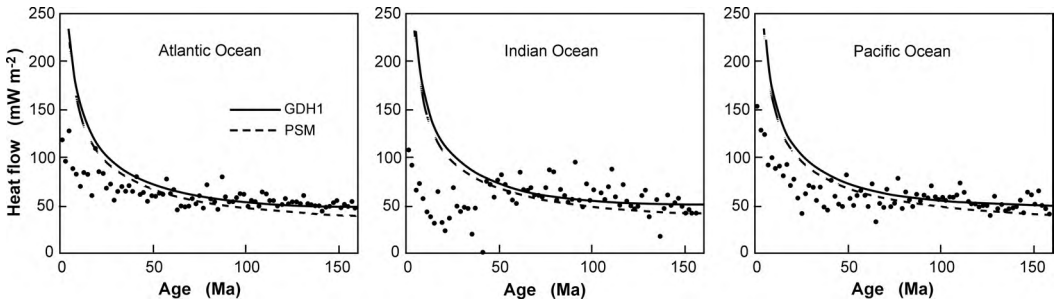


Figure 7.6. Observed heat flows for the Atlantic, Indian and Pacific Oceans. Heat flow predicted by the plate models: solid line, GDH1; and dashed line, PSM. Heat flow predicted by the half-space model HS is not shown – it is almost coincident with PSM (Table 7.5). (After Stein and Stein, Constraints on hydrothermal heat flux through the oceanic lithosphere from global heat data, *J. Geophys. Res.*, **99**, 3881–95, 1994. Copyright 1994 American Geophysical Union. Modified by permission of American Geophysical Union.)

plate moves away from the ridge, sedimentation occurs. Deep-sea sediments have a low permeability[2] and, in sufficient thickness, are impermeable to sea water. In well-sedimented and therefore generally older crust, measurements of conductive heat flow yield reliable estimates of the actual heat flow. Another important factor affecting cessation of hydrothermal circulation is that, in older crust, pores and cracks will become plugged with mineral deposits. As a result, hydrothermal circulation will largely cease. Loss of heat due to hydrothermal circulation is difficult to measure, so heat-flow estimates for young crust are generally very scattered and also significantly lower than the theoretical estimates of heat loss (Fig. 7.6). That heat-flow measurements are generally less than the predicted values for oceanic lithosphere younger than 65 ± 10 Ma indicates that this is the 'sealing age'.

As an oceanic plate moves away from the ridge axis and cools, it contracts and thus increases in density. If we assume the oceanic regions to be compensated (see Section 5.5.2), the depth of the oceans should increase with increasing age (and thus plate density). For any model of the cooling lithosphere, the expected ocean depth can be calculated simply (see Section 7.5.2).

Figure 7.7(a) shows the mean depth of the oceans plotted against age. For ages less than 20 Ma a simple relation between bathymetric depth d (km) and lithosphere age t (Ma) is observed:

$$d = 2.6 + 0.365t^{1/2} \tag{7.57a}$$

Depth increases linearly with the square root of age. For ages greater than 20 Ma this simple relation does not hold; depth increases more slowly with increasing

[2] Permeability and porosity are not the same. Sediments have a higher porosity than that of crustal rocks but lack the connected pore spaces needed for high permeability.

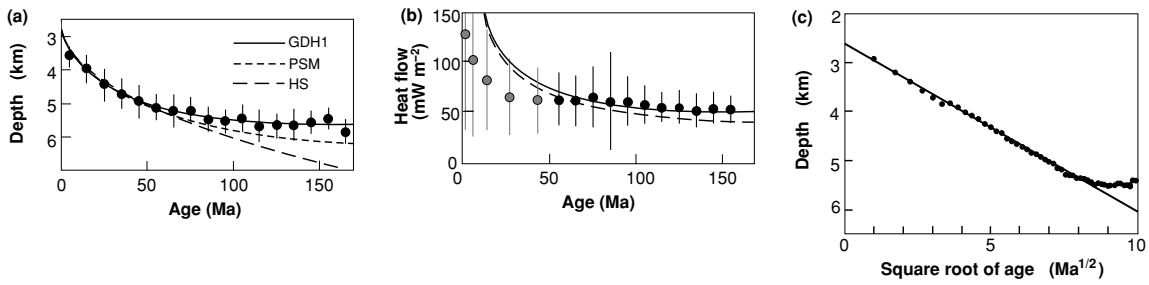


Figure 7.7. Mean oceanic depth (a) and oceanic heat flow (b) with standard deviations plotted every 10 Ma against age. The data are from the north Pacific and northwest Atlantic. These global depths exclude data from the hotspot swells. The three model predictions for ocean depth and heat flow are shown as solid and dashed lines. The plate model GDH1 fits both data sets overall better than does either the half-space model HS or the alternative plate model PSM. Data shown in black were used to determine GDH1. Heat flow data at <50 Ma are shown in grey – these are affected by hydrothermal circulation and were not used to determine GDH1.(c) Mean oceanic depth plotted against the square root of the age of the lithosphere ($\sqrt{t}$). The solid line is the best-fitting half-space model: $d = 2.607 + 0.344t^{1/2}$.(After Stein and Stein (1992) Thermo-mechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes (overview), Geophysical Monograph 96, 1–17, 1996. Copyright 1996 American Geophysical Union. Modified by Permission of American Geophysical Union; and Carlson and Johnson, On modeling the thermal evolution of the oceanic upper mantle: an assessment of the cooling plate model, *J. Geophys. Res.*, **99**, 3201–14, 1994. Copyright 1994 American Geophysical Union. Modified by permission of American Geophysical Union.)

age and approximates a negative exponential:

$$d = 5.65 - 2.47\mathrm{e}^{-t/36} \tag{7.57b}$$

Figure 7.7(b) shows the measured heat flow plotted against age. A simple relationship, linked to that for ocean depth, between heat flow Q (10^{-3} W m^{-2}) and lithosphere age t (Ma) is predicted for crust younger than 55 Ma:

$$Q = 510t^{-1/2} \tag{7.58a}$$

Heat flow decreases linearly with the inverse square root of age. For ages greater than 55 Ma this simple relation does not hold; heat flow decreases more slowly with increasing age and follows a negative exponential:

$$Q = 48 + 96\mathrm{e}^{-t/36} \tag{7.58b}$$

7.5.2 Models of plate formation and cooling

The creation of a lithospheric plate at the axis of a mid-ocean ridge and the subsequent cooling of the plate as it moves away from the ridge axis give rise to

the type of problem that can be solved by using the two-dimensional version of the heat-conduction equation in a moving medium (Eq. (7.19)). The boundary conditions can be specified in a number of ways: these necessarily lead to different solutions and thus to different estimates of heat flow and bathymetric depth. The bathymetric depth is calculated from the temperature by assuming that the plate is in isostatic equilibrium, and the heat flow is calculated from the temperature gradient at the surface of the lithosphere. In this way, an understanding of the thermal structure and formation of the plates has been built up. As in all scientific work, the best model is the one which best fits the observations, in this case variations of bathymetric depth and heat flow with age.

A simple model

The simplest thermal model of the lithosphere is to assume that the lithosphere is cooled asthenospheric material, which, at the ridge axis, had a constant temperature T_a and no heat generation. If we assume the ridge to be infinite in the y direction and the temperature field to be in equilibrium, then the differential equation to be solved is

$$\frac{k}{\rho c_P}\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial z^2}\right)=u\frac{\partial T}{\partial x} \tag{7.59}$$

where u is the horizontal velocity of the plate and the term on the right-hand side of the equation is due to advection of heat with the moving plate. A further simplification can be introduced by the assumption that horizontal conduction of heat is insignificant in comparison with horizontal advection and vertical conduction of heat. In this case, we can disregard the $\partial^2 T/\partial x^2$ term, leaving the equation to be solved as

$$\frac{k}{\rho c_P}\frac{\partial^2 T}{\partial z^2}=u\frac{\partial T}{\partial x} \tag{7.60}$$

This equation, however, is identical to Eq. (7.43) if we write $t = x/u$, which means that we reintroduce time through the spreading of the ridge. Approximate initial and boundary conditions are $T = T_a$ at $x = 0$ and $T = 0$ at $z = 0$. According to Eq. (7.44), the solution to Eq. (7.60) is

$$T(z,t)=T_a\,\mathrm{erf}\left(\frac{z}{2\sqrt{\kappa t}}\right) \tag{7.61}$$

The surface heat flow at any distance (age) from the ridge axis is then obtained by differentiating Eq. (7.61):

$$\begin{aligned} Q(t) &= -k\left.\frac{\partial T}{\partial z}\right|_{z=0} \\ &= -\frac{kT_a}{\sqrt{\pi\kappa t}} \end{aligned} \tag{7.62}$$

The observed $t^{1/2}$ relationship between heat flow and age is thus a feature of this model which is called a half-space cooling model.

We can estimate the lithospheric thickness L from Eq. (7.61) by specifying a temperature for the base of the lithosphere. For example, if we assume the temperature of the asthenosphere at the ridge axis to be 1300 °C and the temperature at the base of the lithosphere to be 1100 °C, then we need to find the combination of L and t such that

$$1100 = 1300 \operatorname{erf}\left(\frac{L}{2\sqrt{\kappa t}}\right) \tag{7.63}$$

In other words, we need the inverse error function of 0.846. Using Fig. 7.5 (or Appendix 5), we can write

$$1.008 = \frac{L}{2\sqrt{\kappa t}} \tag{7.64}$$

Thus, if $\kappa = 10^{-6}\,\mathrm{m^2\,s^{-1}}$,

$$L = 2.016 \times 10^{-3}\sqrt{t} \tag{7.65a}$$

when L is in metres and t in seconds, or

$$L = 11\sqrt{t} \tag{7.65b}$$

when L is in kilometres and t in millions of years (Ma). At 10 Ma this lithosphere would be 35 km thick, and at 80 Ma it would be 98 km thick. Different choices for the temperature at the ridge axis and at the base of the lithosphere will yield slightly different values for the numerical constants in Eq. (7.65) but the 'root t' dependence of lithosphere thickness on age will not change.

The depth of the seabed at any given age can be calculated by using the principle of isostasy (see Section 5.5.2) and the gradual increase in density of the lithosphere as it cools. If we take the compensation depth D to be in the mantle beneath the base of the lithosphere, the total mass in a vertical column extending down to D is

$$\int_0^D \rho(z)\,\mathrm{d}z$$

Isostatic compensation requires that this mass be constant for all columns whatever their age. At the ridge axis the lithosphere has zero thickness, and so, taking $z = 0$ to be at sea level, the mass of the column is

$$\int_0^{d_\mathrm{r}} \rho_\mathrm{w}\,\mathrm{d}z + \int_{d_\mathrm{r}}^D \rho_\mathrm{a}\,\mathrm{d}z$$

where ρ_w is the density of sea water, ρ_a the density of the asthenosphere (at temperature T_a) and d_r the depth of the water over the ridge axis. The mass of a column aged t is then

$$\int_0^d \rho_\mathrm{w}\,\mathrm{d}z + \int_d^{d+L} \rho(z)\,\mathrm{d}z + \int_{d+L}^D \rho_\mathrm{a}\,\mathrm{d}z$$

where d is the water depth and L the thickness of the lithosphere. Because the mass in this column must be the same as the mass in the column at the ridge axis,

we obtain the equation

$$\int_0^{d_r} \rho_w \, dz + \int_{d_r}^{D} \rho_a \, dz = \int_0^{d} \rho_w \, dz + \int_{d}^{d+L} \rho(z) \, dz + \int_{d+L}^{D} \rho_a \, dz \tag{7.66}$$

Rearranging yields

$$(d - d_r)(\rho_a - \rho_w) = \int_0^{L} (\rho(z) - \rho_a) \, dz \tag{7.67}$$

To determine $\rho(z)$, we must use the expression for density as a function of temperature and α the coefficient of thermal expansion,

$$\rho(T) = \rho_a[1 - \alpha(T - T_a)] \tag{7.68}$$

and Eq. (7.61) for the temperature structure of the lithosphere. Substituting these two equations into Eq. (7.67) gives

$$(d - d_r)(\rho_a - \rho_w) = \rho_a \alpha T_a \int_0^{L} \operatorname{erfc}\left(\frac{z}{2\sqrt{\kappa t}}\right) dz \tag{7.69}$$

where erfc is the complementary error function (see Appendix 5). The integral on the right-hand side of this equation can easily be calculated or looked up in a set of mathematical tables. However, for our purposes it is sufficient to change the upper limit of integration from L to ∞ (the error introduced by this approximation is about 5%). This integral of erfc(x) between $x = 0$ and infinity is $1/\sqrt{\pi}$ (Appendix 5). When this approximation is made, Eq. (7.69) becomes

$$(d - d_r)(\rho_a - \rho_w) = 2\rho_a \alpha T_a \sqrt{\frac{\kappa t}{\pi}} \tag{7.70}$$

Rearranging Eq. (7.70) gives

$$d = d_r + \frac{2\rho_a \alpha T_a}{\rho_a - \rho_w} \sqrt{\frac{\kappa t}{\pi}} \tag{7.71}$$

If we assume values for ρ_a and for ρ_w of 3.3×10^3 and $1.0 \times 10^3 \, kg \, m^{-3}$, respectively; for α, $3 \times 10^{-5} \, ^\circ C^{-1}$; for κ, $10^{-6} \, m^2 \, s^{-1}$; for T_a, $1200 \, ^\circ C$; and for d_r, 2.6 km; with t in millions of years and d in kilometres, then Eq. (7.71) is

$$d = 2.6 + 0.33\sqrt{t} \tag{7.72}$$

When T_a is taken to be $1300 \, ^\circ C$, Eq. (7.71) is

$$d = 2.6 + 0.36\sqrt{t} \tag{7.73}$$

Such dependence of ocean depth on age is in broad agreement with the depths observed for oceanic plates less than 70 Ma old (Eq. (7.57a)). Thus, as the lithosphere moves away from the ridge axis, it cools, contracts and hence subsides. Table 7.4 gives the details of the physical parameters for oceanic-lithosphere thermal models.

The boundary-layer or half-space models

The boundary-layer model is a modification of the simple half-space model: the rigid lithosphere is assumed to be cooled asthenosphere, and the base

Table 7.4 *Thermal parameters for oceanic-lithosphere models*

		GDH1	PSM	HS
L,	plate thickness (km)	95 ± 10	125 ± 10	—
T_a,	temperature at base of plate (°C)	1450 ± 100	1350 ± 275	1365 ± 10
α,	coefficient of thermal expansion ($°C^{-1}$)	3.1×10^{-5}	3.28×10^{-5}	3.1×10^{-5}
k,	thermal conductivity ($W\,m^{-1}$)	3.138	3.138	3.138
c_P,	specific heat ($kJ\,kg^{-1}$)	1.171	1.171	1.171
κ,	thermal diffusivity ($m^2\,s^{-1}$)	0.804×10^{-6}	0.804×10^{-6}	0.804×10^{-6}
ρ_m,	mantle density ($kg\,m^{-3}$)	3330	3330	3330
ρ_w,	water density ($kg\,m^{-3}$)	1000	1000	1000
d_r,	ridge depth (km)	2.6	2.5	2.6

Abbreviations: GDH, global depth and heat; PSM, Parsons, Sclater and McKenzie; HS, halfspace.
Sources: Stein and Stein (1992), Parsons and Sclater (1977) and Carlson and Johnson (1994).

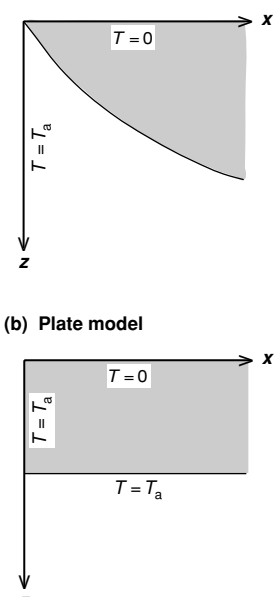


Figure 7.8. Schematic diagrams of (a) the half-space cooling and boundary-layer models and (b) the plate model of the oceanic lithosphere. Lithosphere is shaded.

of the lithosphere is defined by an isotherm (Fig. 7.8). The boundary condition at the base of the lithosphere is that the heat flux from the mantle is specified. The bathymetric depth predicted by these models increases as $t^{1/2}$ for all t. It is apparent from Fig. 7.7(a) that, although this model fits the observed depths out to about 60–70 Ma, for greater ages the predicted depth is too great. The heat-flow values predicted by the model decrease as $t^{-1/2}$ for all t. These predicted heat flows are close to, but lower than, the observed values (Fig. 7.7(b)). For ages greater than a few million years, the temperature and lithospheric thickness predicted by the simple model are very close to the values for the boundary-layer model.

The plate model

In the plate models, the oceanic lithosphere is taken to be of constant thickness L, the base of the lithosphere is at the same constant temperature T_a as the vertical ridge axis and the top surface of the lithosphere and the seabed is another isotherm, usually put at 0 °C (Fig. 7.8). The solution to Eq. (7.59) with these boundary conditions, $T = T_a$ on $z = L$, $T = 0$ on $z = 0$ and $T = T_a$ on $t = 0$, is

$$T = T_a \left\{ \frac{z}{L} + \sum_{n=1}^{\infty} \frac{2}{n\pi} \sin\left(\frac{n\pi z}{L}\right) \exp\left[\left(\frac{uL}{2\kappa} - \sqrt{\frac{u^2 L^2}{4\kappa^2} + n^2\pi^2} \right) \frac{ut}{L} \right] \right\} \qquad (7.74)$$

When the horizontal conduction of heat is ignored, this simplifies to

$$T = T_{\mathrm{a}}\left[\frac{z}{L} + \sum_{n=1}^{\infty}\frac{2}{n\pi}\sin\left(\frac{n\pi z}{L}\right)\exp\left(-\frac{n^2\pi^2\kappa t}{L^2}\right)\right]$$

A thermal time constant can be defined as $t_0 = L^2/(\pi^2\kappa)$.

The surface heat flow as a function of plate age, $Q\,(t)$, for this model is

$$Q(t) = \frac{kT_{\mathrm{a}}}{L}\left[1 + 2\sum_{n=1}^{\infty}\exp\left(-\frac{n^2\pi^2 t}{L^2}\right)\right]$$

The asymptotic value for the heat flow over old lithosphere is therefore kT_{a}/L.

The depth of the seabed as a function of plate age, $d(t)$, for this model is calculated in exactly the same way as for the simple half-space model:

$$d(t) = d_r + \frac{\rho_{\mathrm{a}}\alpha T_{\mathrm{a}} L}{2(\rho_{\mathrm{a}} - \rho_{\mathrm{w}})}\left[1 - \frac{8}{\pi^2}\sum_{n=1}^{\infty} n^2\left(-\frac{n^2\pi^2\kappa t}{L^2}\right)\right]$$

The asymptotic value for the ocean depth over old lithosphere is therefore

$$d_{\mathrm{r}} + \frac{\rho_{\mathrm{a}}\alpha T_{\mathrm{a}} L}{2(\rho_{\mathrm{a}} - \rho_{\mathrm{w}})}$$

Isotherms for the two plate models, PSM and GDH1, with temperatures of 1350 and 1450 °C, respectively, at the base of the lithosphere and at the ridge axis are shown in Fig. 7.9. As the lithosphere ages and moves away from the ridge axis, the isotherms descend until, far from the ridge, they essentially reach equilibrium. Figures 7.6 and 7.7 show that the heat flow and bathymetric depths predicted by these models are in good agreement with observations. Note that there is effectively no difference between the heat flows predicted by the HS and PSM models, compared with the scatter in the data (Figs. 7.6, 7.7(b) and Table 7.5). The differences between ocean depths predicted by the plate models and by the boundary-layer model become apparent beyond about 60–70 Ma and the deviation of the ocean depth from the $\sqrt{t}$ curve predicted by the half-space model shows up clearly in Fig. 7.7(c). Since there is no limit to how cool the upper regions of the boundary-layer model can become, there is no limit to its predicted ocean depths. The plate model has a uniformly thick lithosphere, so temperatures in the lithosphere, as well as ocean depths, predicted by that model approach equilibrium as age increases. For the same reason, differences between the surface heat flows predicted by the two types of model begin to become apparent for ages greater than about 100 Ma – the boundary-layer model keeps on cooling whereas the plate models approach equilibrium.

Any thermal model must account for the $\sqrt{t}$ dependence of ocean depth on age for young oceanic lithosphere. Additionally the model must account for the asymptotic behaviour both of ocean depth and of heat flow on old lithosphere. The GDH1 plate model fits the whole dataset best (with the lowest residuals), but a half-space model fits the ocean-depth dataset for young lithosphere best.

Table 7.5 *Variation of depth and heat flow with age for oceanic-lithosphere models*

	Ocean depth (km)	Heat flow ($mW\,m^{-2}$)
Half-space	$2.6 + 0.345t^{1/2}$	$480t^{-1/2}$
PSM	$2.5 + 0.350t^{1/2}$, $t < 70$ Ma	$473t^{-1/2}$, $t < 120$ Ma
	$6.4 - 3.2e^{-t/62.8}$, $t > 70$ Ma	$33.5 + 67e^{t/62.8}$, $t > 120$ Ma
GDH1	$2.6 + 0.365t^{1/2}$, $t < 20$ Ma	$510t^{-1/2}$, $t < 55$ Ma
	$5.65 - 2.47e^{t/36}$ $t > 20$ Ma	$49 + 96e^{-t/36}$, $t > 55$ Ma

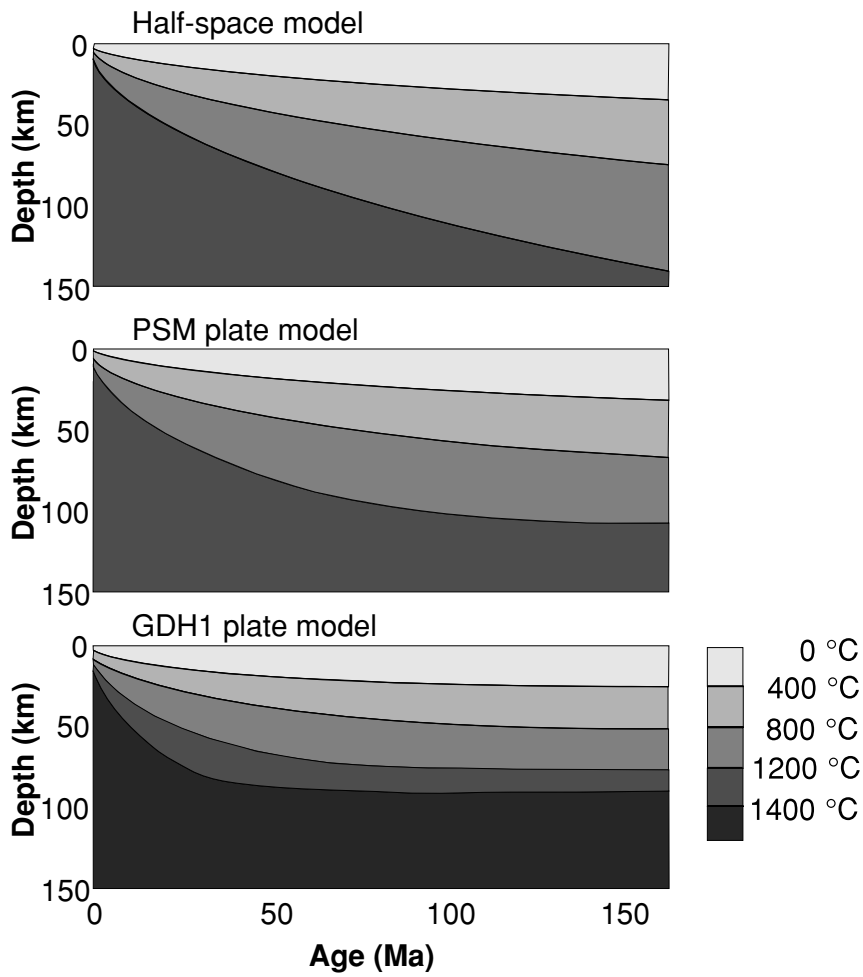


Figure 7.9. Temperature contours for three thermal models of the oceanic lithosphere. The half-space model and the plate models PSM (Parsons, Sclater and McKenzie) and GDH1 (global depth and heat) are shown schematically in Fig. 7.8. The GDH1 model has been constrained so that it fits both the oceanic depth and heat-flow measurements. Note that GDH1 has a thinner plate and higher temperatures than the other models. (After Stein and Stein, Thermo-mechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes (overview), Geophysical Monograph 96, 1–17, 1996. Copyright 1996 American Geophysical Union. Modified by permission of American Geophysical Union.)

The differences between the predictions of these models are small, but their values for mantle temperature and plate thickness are rather different (Table 7.4). The variability in the depth and heat-flow data resulting from hotspot proximity, mantle thermal structure and hydrothermal circulation means that it is not possible to establish an unequivocal global thermal model that can simultaneously account for all the depth and heat-flow data at every age. The variations of depth and heat flow with age for the half-space, PSM and GDH1 models are summarized in Table 7.5.

Thermal structure of the oceanic lithosphere

Both plate and boundary-layer models of the lithosphere provide heat-flow values that are in reasonable agreement with the measured values, but the ocean depths predicted by the plate model and boundary-layer models differ, with the plate-model predictions being overall in much better agreement with observed ocean depths. Other geophysical evidence on the structure of the oceanic lithosphere also shows that the oceanic lithosphere thickens with age, but they cannot distinguish amongst the thermal models (Fig. 5.17). The effective elastic thickness (determined from studies of loading and a measure of the long-term strength of the lithosphere) increases with age approximately as the 400-°C isotherm. The maximum focal depth of intraplate earthquakes (a measure of the short-term strength of the lithosphere) increases with age approximately as the 600–700-°C isotherm. Results of surface-wave-dispersion studies show that the depth to the low-velocity zone (the top of the asthenosphere) also deepens with age with plate-model isotherms. However, while all these parameters clearly increase with lithospheric age and broadly follow isotherms for the plate models, they are not well enough determined to allow one to distinguish amongst the various thermal models.

The observations could be reconciled with the boundary-layer model if some mechanism to slow the cooling of the boundary layer model for ages greater than ∼70 Ma were found, so that it would resemble the plate model. Two mechanisms for maintaining the heat flux at the base of the lithosphere have been proposed: shear-stress heating caused by a differential motion between lithosphere and asthenosphere; and an increasing rate of heat production in the upper mantle. These mechanisms are both somewhat unlikely; perhaps a better proposal is that small-scale convection occurs in the asthenosphere at the base of the older lithosphere. This would increase the heat flux into the base of the rigid lithosphere and maintain a more constant lithospheric thickness.

The lithospheric plate is thought to consist of two parts: an upper rigid layer and a lower viscous thermal boundary layer (Fig. 7.10). At about 60 Ma this thermal boundary layer becomes unstable; hence small-scale convection develops within it (see Section 8.2), resulting in an increase in heat flow to the base of the rigid layer and a thermal structure similar to that predicted by the plate model. Very detailed, accurate measurements of heat flow, bathymetry and the geoid on old

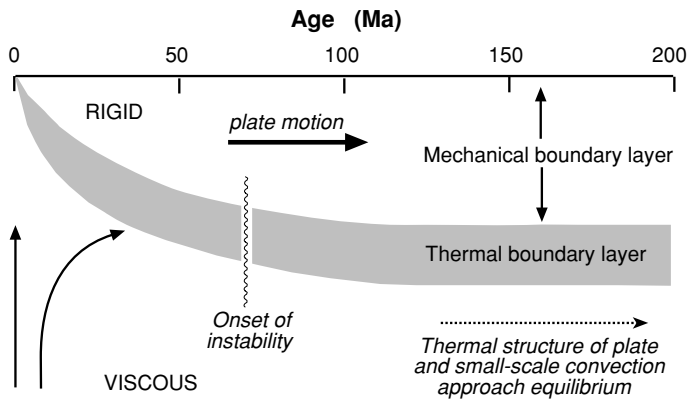


Figure 7.10. A schematic diagram of the oceanic lithosphere, showing the proposed division of the lithospheric plate. The base of the mechanical boundary layer is the isotherm chosen to represent the transition between rigid and viscous behaviour. The base of the thermal boundary layer is another isotherm, chosen to represent correctly the temperature gradient immediately beneath the base of the rigid plate. In the upper mantle beneath these boundary layers, the temperature gradient is approximately adiabatic. At about 60–70 Ma the thermal boundary layer becomes unstable, and small-scale convection starts to occur. With a mantle heat flow of about $38 \times 10^{-3}\,\mathrm{W\,m^{-2}}$ the equilibrium thickness of the mechanical boundary layer is approximately 90 km. (From Parsons and McKenzie (1978).)

oceanic crust and across fracture zones may improve our knowledge of the thermal structure of the lithosphere.

7.6 Continental heat flow

7.6.1 The mantle contribution to continental heat flow

Continental heat flow is harder to understand than oceanic heat flow and harder to fit into a general theory of thermal evolution of the continents or of the Earth. Continental heat-flow values are affected by many factors, including erosion, deposition, glaciation, the length of time since any tectonic events, local concentrations of heat-generating elements in the crust, the presence or absence of aquifers and the drilling of the hole in which the measurements were made. Nevertheless, it is clear that the measured heat-flow values decrease with increasing age (Fig. 7.11). This suggests that, like the oceanic lithosphere, the continental lithosphere is cooling and slowly thickening with time. The mean surface heat flow for the continents is $\sim 65\,\mathrm{mW\,m^{-2}}$. The mean surface heat flow in non-reactivated Archaean cratons is $41 \pm 11\,\mathrm{mW\,m^{-2}}$, which is significantly lower than the mean value of $55 \pm 17\,\mathrm{mW\,m^{-2}}$ for stable Proterozoic crust well away from Archaean craton boundaries.

That all erosional, depositional, tectonic and magmatic processes occurring in the continental crust affect the measured surface heat-flow values is shown in

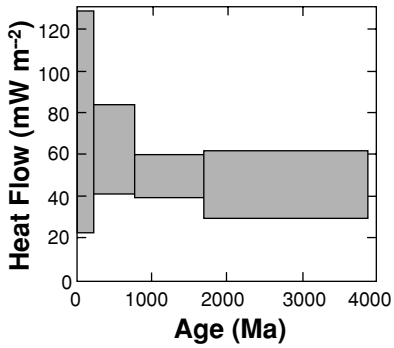


Figure 7.11. Heat flow versus crustal age for the continents. The heights of the boxes indicate the standard deviation about the mean heat flow, and the widths indicate the age ranges. (After Sclater *et al.* (1980).)

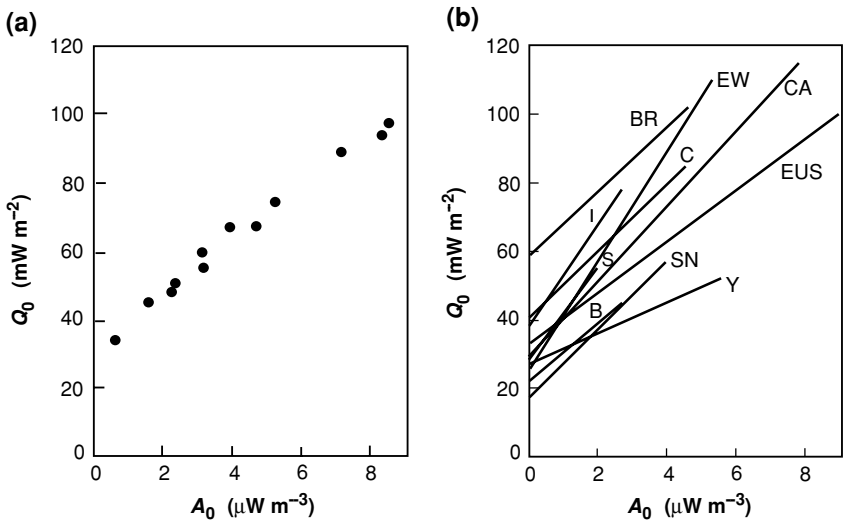


Figure 7.12. Measured heat flow Q_0 plotted against internal heat generation A_0 for (a) the eastern-U.S.A. heat-flow province. The straight line $Q_0 = Q_r + DA_0$ that can be fitted to these measurements has $Q_r = 33 \times 10^{-3}$ W m^{-2} and $D = 7.5$ km. (After Roy *et al.* (1968).) (b) Best-fitting straight lines for other heat-flow provinces: CA, central Australia; B, Baltic shield; BR, Basin and Range; C, Atlantic Canada; EW, England and Wales; EUS, eastern USA; I, India; S, Superior Province; SN, Sierra Nevada; and Y, Yilgarn block, Australia. (After Jessop (1990).)

the examples of Sections 7.3 and 7.8. The particularly scattered heat-flow values measured at ages less than about 800 Ma are evidence of strong influence of these transient processes and are therefore very difficult to interpret in terms of the deeper thermal structure of the continents.

In some specific areas known as heat-flow provinces, there is a linear relationship between surface heat flow and surface radioactive heat generation (Fig. 7.12). Using this relationship, one can make an approximate estimate of the contribution of the heat-generating elements in the continental crust to the surface heat flow. In these heat-flow provinces, some of which are listed in Table 7.6, the surface heat flow Q_0 can be expressed in terms of the measured surface radioactive heat generation A_0 as

$$Q_0 = Q_r + A_0 D \tag{7.75}$$

where Q_r and D are constants for each heat-flow province.

Table 7.6 *Some continental heat-flow provinces*

Province	Mean Q_0 (10^{-3} W m^{-2})	Q_r (10^{-3} W m^{-2})	D (km)
Basin and Range (U.S.A.)	92	59	9.4
Sierra Nevada (U.S.A.)	39	17	10.1
Eastern U.S.A.	57	33	7.5
Superior Province (Canadian Shield)	39	21	14.4
U.K.	59	24	16.0
Western Australia	39	26	4.5
Central Australia	83	27	11.1
Ukrainian Shield	37	25	7.1

Source: Sclater *et al.* (1980).

We consider here two extreme models of the distribution of the radioactive heat generation in the crust, both of which yield a surface heat flow in agreement with this observed linear observation.

1. *Heat generation is uniformly concentrated within a slab with thickness D.*

In this case, using Eq. (7.16), we obtain

$$\frac{\partial^2 T}{\partial z^2} = -\frac{A_0}{k} \quad \text{for } 0 \le z \le D$$

Integrating once gives

$$\frac{\partial T}{\partial z} = -\frac{A_0}{k} z + c \tag{7.76}$$

where c is the constant of integration. At the surface, $z = 0$, the upward heat flow $Q(0)$ is

$$\begin{aligned} Q(0) = Q_0 &= k \left.\frac{\partial T}{\partial z}\right|_{z=0} \\ &= kc \end{aligned} \tag{7.77}$$

Therefore, the constant c is given by

$$c = \frac{Q_0}{k}$$

At depth D, the upward heat flow is

$$\begin{aligned} Q(D) &= k\left(-\frac{A_0 D}{k} + \frac{Q_0}{k}\right) \\ &= -A_0 D + Q_0 \\ &= Q_r \end{aligned} \tag{7.78}$$

Thus, in this case, the heat flow $Q(D)$ into the base of the uniform slab (and the base of the crust, since all the heat generation is assumed to be concentrated in the slab) is the Q_r of Eq. (7.75).

2. *Heat generation is an exponentially decreasing function of depth within a slab of thickness z**. Equation (7.16) then becomes

$$\frac{\partial^2 T}{\partial z^2} = -\frac{A(z)}{k} \tag{7.79}$$

where

$$A(z) = A_0 e^{-z/D} \quad \text{for } 0 \le z \le z^*$$

Integrating Eq. (7.79) once gives

$$\frac{\partial T}{\partial z} = \frac{A_0}{k} D e^{-z/D} + c \tag{7.80}$$

where c is the constant of integration. At the surface, $z = 0$, the heat flow is $Q(0)$

$$\begin{aligned} Q(0) = Q_0 &= k\left(\frac{A_0 D}{k} + c\right) \\ &= A_0 D + kc \end{aligned} \tag{7.81}$$

The constant c is given by

$$c = \frac{Q_0 - A_0 D}{k} \tag{7.82}$$

At depth z^* (which need not be uniform throughout the heat-flow province), the heat flow is

$$\begin{aligned} Q(z^*) &= k\left(\frac{A_0 D}{k} e^{-z^*/D} + \frac{Q_0 - A_0 D}{k}\right) \\ &= A_0 D e^{-z^*/D} + Q_0 - A_0 D \end{aligned} \tag{7.83}$$

Thus, by rearranging, we obtain

$$Q_0 = Q(z^*) + A_0 D - A_0 D e^{-z^*/D} \tag{7.84}$$

Equation (7.84) is the same as Eq. (7.75) if we write

$$\begin{aligned} Q_r &= Q(z^*) - A_0 D e^{-z^*/D} \\ &= Q(z^*) - A(z^*) D \end{aligned} \tag{7.85}$$

Thus, the linear relation is valid for this model if the heat generation $A(z^*)$ at depth z^* is constant throughout the heat-flow province. Unless $A(z^*)D$ is small, the observed value of Q_r may be very different from the actual heat flow $Q(z^*)$ into the base of the layer of thickness z^*. However, it can be shown (for details see Lachenbruch (1970) that, for some heat-flow provinces, $A(z^*)D$ is small, and thus Q_r is a reasonable estimate of $Q(z^*)$. This removes the constraint that $A(z^*)$ must be the same throughout the heat-flow province. Additionally, for those provinces in which $A(z^*)D$ is small, it can be shown that z^* must be substantially greater than D. Thus, the exponential distribution of heat production satisfies the observed linear relationship between surface heat flow and heat generation and does so even in cases of differential erosion. In this model, D is a measure of the upward migration of the heat-producing radioactive isotopes (which can be justified on geochemical grounds), and Q_r is approximately the heat flow into

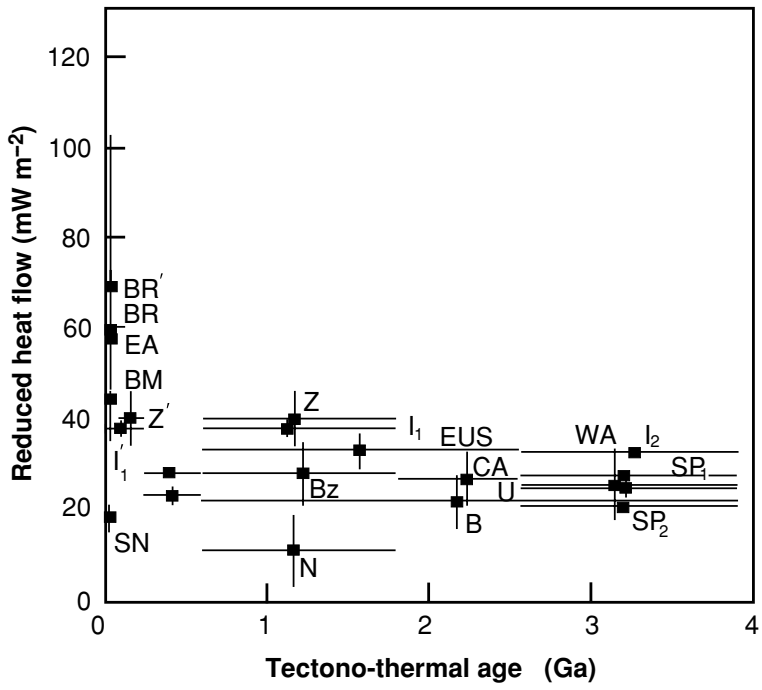


Figure 7.13. Reduced heat flow Q_r versus time since the last tectono-thermal event for the continental heat flow provinces. The error bars represent the uncertainties in the data. The solid lines show the reduced heat flows predicted by the plate model. BR and BR′, Basin and Range; SEA, southeast Appalachians; SN, Sierra Nevada; EUS, eastern U.S.A.; SP_1 and SP_2, Superior Province; Bz, Brazilian coastal shield; B, Baltic shield; BM, Bohemian massif; U, the Ukraine; EW, England and Wales; N, Niger; Z and Z′, Zambia; WA, western Australia; CA, central Australia; EA, eastern Australia; I_1 and I'_1 Indian shield; and I_2, Archaean Indian shield. (After Morgan (1984) and Stein (1995), *Heat Flow of the Earth*, AGU Reference Shelf 1, 144–58, 1995. Copyright 1995 American Geophysical Union. Modified by permission of American Geophysical Union.)

the base of the crust (because z^* is probably approximately the thickness of the crust).

Neither of these models of the distribution of heat generation within the crust allows for different vertical distributions among the various radioactive isotopes. There is some evidence for such variation. Nevertheless, it is clear that much of the variation in measured surface heat flow is caused by the radioactive heat generation in the crust and that the reduced heat flow Q_r is a reasonable estimate of the heat flow into the base of the crust. Figure 7.13 shows this reduced heat flow plotted against age. After about 300 Ma since the last tectonic/thermal event, the reduced heat flow exhibits no variation and attains a value of $(25 \pm 8) \times 10^{-3}$ W m^{-2}. This is within experimental error of the value predicted by the plate model of the oceanic lithosphere and suggests that there should be no significant difference between models of the thermal structures of the oceanic and continental lithospheres. The present-day thermal differences are primarily a consequence of the age disparity between oceanic and continental lithospheres.

7.6.2 The temperature structure of the continental lithosphere

Figure 7.14 shows two extreme temperature models of the equilibrium oceanic lithosphere, O_1 and O_2, and two extreme models of the old stable continental

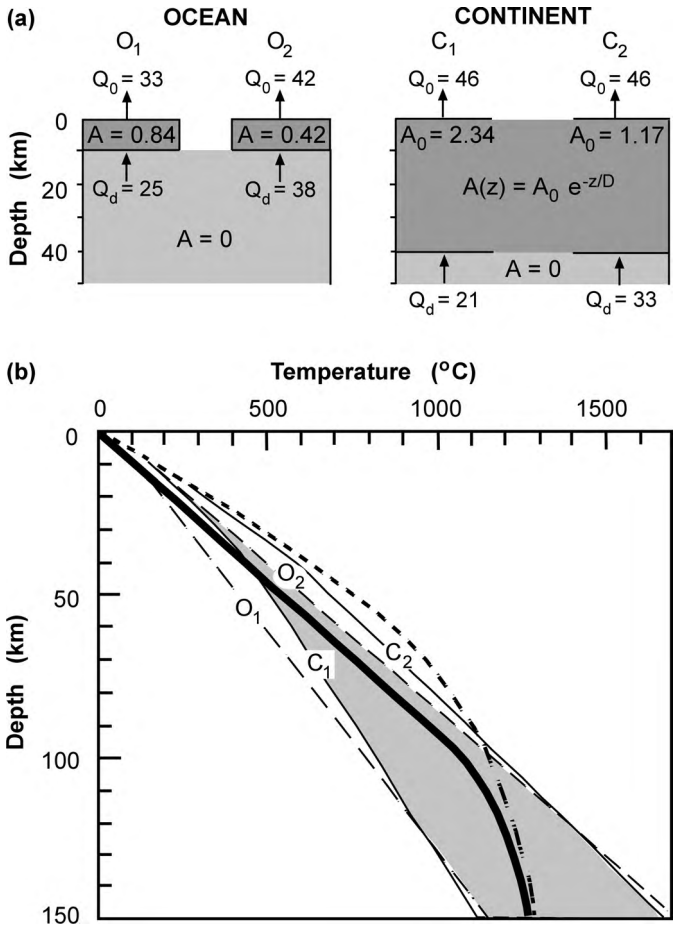


Figure 7.14. (a) Extreme thermal models used to calculate equilibrium geotherms beneath an ocean, O_1 and O_2, and beneath an old stable continent, C_1 and C_2. Heat flows Q_0 and Q_d are in mW m^{-2}; heat generation A_0 is in μW m^{-3}. (b) Predicted geotherms for these models. Thin dashed lines, oceanic geotherms; thin solid lines, continental geotherms; heavy solid line, equilibrium geotherm for the PSM plate model, taking into account the small-scale convection occurring in the thermal boundary layer (see Fig. 7.10). Grey shading, region of overlap. The heavy dashed line is an error function for the geotherm of age 70 Ma (see Section 7.5.2). The mantle temperature T_a is taken as 1300 °C. (After Parsons and McKenzie (1978) and Sclater *et al.* (1981b).)

lithosphere, C_1 and C_2. These have been calculated by using the one-dimensional heat-conduction equation. The extensive region of overlap of these four geotherms indicates that, on the basis of surface measurements, for depths greater than about 80 km there need be little difference in equilibrium temperature structure beneath oceans and continents. All the proposed oceanic thermal models (Section 7.5) fall within the shaded region of overlap. The solid line is the geotherm for the oceanic-plate model in Fig. 7.10. The heavy dashed line is the geotherm for the simple error-function model of Section 7.5.2. Figure 7.15 shows thermal models of oceanic and old continental lithospheres.

7.7 The adiabat and melting in the mantle

The previous sections have dealt in some detail with the temperatures in the continental and oceanic lithosphere and with attempts to estimate the temperatures in the mantle and core, assuming that heat is transferred by conduction. For the

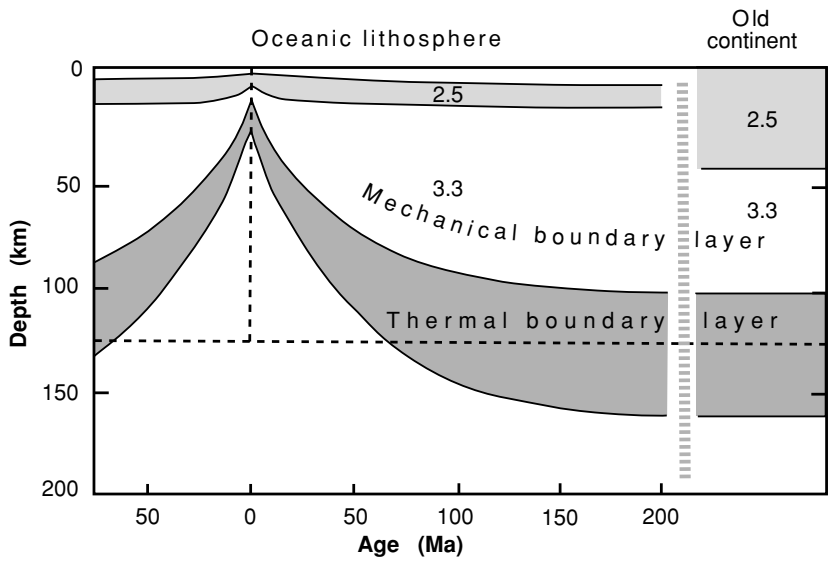


Figure 7.15. Thermal models of the lithospheric plates beneath oceans and continents. The dashed line is the plate thickness predicted by the PSM plate model; k (values of 2.5 and 3.3) is the conductivity in $\mathrm{W\,m^{-1}\,{}^{\circ}C^{-1}}$. (After Sclater *et al.* (1981b).)

mantle and outer core, however, where conduction is not the primary method of heat transfer, the methods and estimates of the previous sections are not appropriate. In the mantle and the outer core, where convection is believed to be occurring, heat is transported as hot material moves; thus, the rate of heat transfer is much greater than that by conduction alone, and as a result the temperature gradient and temperatures are much lower. In the interior of a vigorously convecting fluid, the mean temperature gradient is approximately adiabatic. Hence, the temperature gradient in the mantle is approximately adiabatic.

To estimate the adiabatic temperature gradient in the mantle, we need to use some thermodynamics. Consider *adiabatic expansion*, an expansion in which entropy is constant for the system (the system can be imagined to be in a sealed and perfectly insulating rubber bag). Imagine that a rock unit that is initially at depth z and temperature T is suddenly raised up to depth z'. Assume that the rock unit is a closed system, and let us consider the change in temperature that the unit undergoes. When it reaches its new position z', it is hotter than the surrounding rocks; but, because it was previously at a higher pressure, it expands and, in so doing, cools. If the temperature to which it cools as a result of this expansion is the temperature of the surrounding rocks, then the temperature gradient in the rock pile is *adiabatic*. Thus, an adiabatic gradient is essentially the temperature analogue of the self-compression density model discussed in Section 8.1.2. Temperature gradients in a convecting system are close to adiabatic.

To determine the adiabatic gradient, we need to determine the rate of change of temperature T (in K not °C) with pressure P at constant entropy S. Using the reciprocal theorem (a mathematical trick), we can write this as

$$\left(\frac{\partial T}{\partial P}\right)_S = -\left(\frac{\partial T}{\partial S}\right)_P \left(\frac{\partial S}{\partial P}\right)_T \tag{7.86}$$

However, we know from Maxwell's thermodynamic relations that

$$\left(\frac{\partial S}{\partial P}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_P \tag{7.87}$$

where V is volume. Thus, Eq. (7.86) becomes

$$\left(\frac{\partial T}{\partial P}\right)_S = \left(\frac{\partial T}{\partial S}\right)_P \left(\frac{\partial V}{\partial T}\right)_P \tag{7.88}$$

The definition of α, the coefficient of thermal expansion, is

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P \tag{7.89}$$

The definition of specific heat at constant pressure c_P is

$$mc_P = T\left(\frac{\partial S}{\partial T}\right)_P \tag{7.90}$$

where m is the mass of the material.

Using Eqs. (7.89) and (7.90), we can finally write Eq. (7.88) as

$$\left(\frac{\partial T}{\partial P}\right)_S = \frac{T\alpha V}{mc_P} \tag{7.91}$$

Since $m/v = \rho$, density, Eq. (7.91) further simplifies to

$$\left(\frac{\partial T}{\partial P}\right)_S = \frac{T\alpha}{\rho c_P} \tag{7.92}$$

For the Earth, we can write

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -g\rho \tag{7.93}$$

where g is the acceleration due to gravity. The change in temperature with radius r is therefore given by

$$\begin{aligned}\left(\frac{\partial T}{\partial r}\right)_S &= \left(\frac{\partial T}{\partial P}\right)_S \frac{\mathrm{d}P}{\mathrm{d}r} \\ &= -\frac{T\alpha}{\rho c_P} g\rho \\ &= -\frac{T\alpha g}{c_P}\end{aligned} \tag{7.94}$$

For the uppermost mantle, the adiabatic temperature gradient given by Eq. (7.94) is about $4 \times 10^{-4}\,^{\circ}\mathrm{C\,m^{-1}}$ ($0.4\,^{\circ}\mathrm{C\,km^{-1}}$) assuming the following values: T, 1700 K (1427 °C); α, $3 \times 10^{-5}\,^{\circ}\mathrm{C^{-1}}$; g, $9.8\,\mathrm{m\,s^{-2}}$; and c_P, $1.25 \times 10^3\,\mathrm{J\,kg\,^{\circ}C^{-1}}$. At greater depths in the mantle, where the coefficient of thermal expansion is somewhat less, the adiabatic gradient is reduced to about $3 \times 10^{-4}\,^{\circ}\mathrm{C\,m^{-1}}$ ($0.3\,^{\circ}\mathrm{C\,km^{-1}}$). Figure 7.16 illustrates a range of possible models for the temperature through the mantle. Many estimates of the increase of temperature with depth in the mantle have been made: all agree that the temperature gradient though the upper mantle will be approximately adiabatic. If the upper and lower mantle are separately convecting systems, the temperature will increase by several hundred degrees on passing through the boundary layer at 670 km. In the outer part of

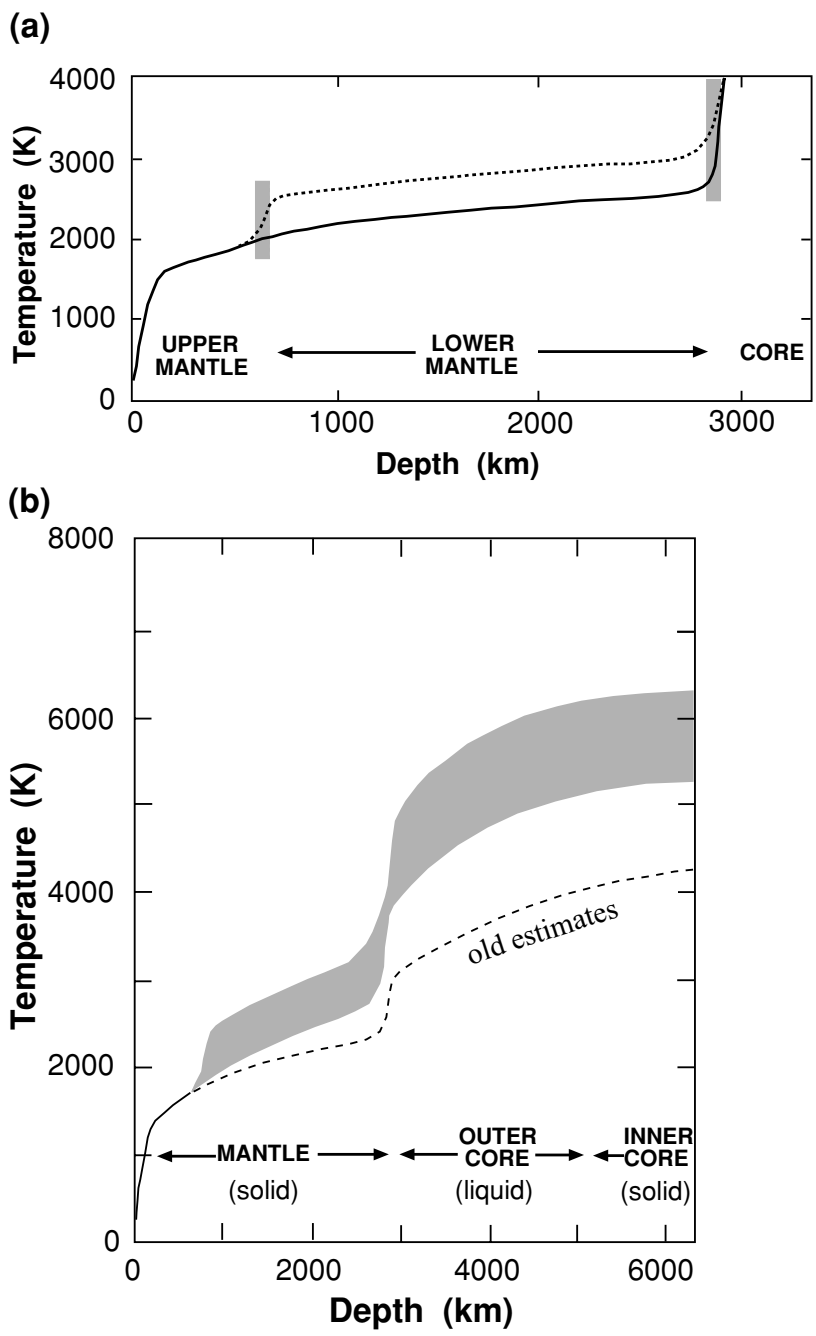


Figure 7.16. Models of temperature profiles in the Earth. (a) Solid black line, mantle adiabat with a thermal boundary layer at the surface and at the core–mantle boundary. Dashed black line, mantle adiabat for a thermal boundary layer both at the top and at the bottom of the lower mantle. The dashed-line adiabat indicates that there is a chemical and dynamic boundary between the upper and lower mantle, which are assumed to be separate systems. Grey lines, location of boundary layers. The temperature at the core–mantle boundary is assumed to be in the range 2900–3200 °C. (b) An estimate of temperatures in the Earth based on high-pressure (over 1000 GPa) and high-temperature (up to 6700 °C) experiments on iron and on theoretical calculations. The shading reflects the uncertainty in temperatures at depth. The 'old estimate' is typical of temperature profiles proposed prior to 1987. The dashed line is a theoretical superadiabatic gradient in the deep lower mantle. (Based on Jeanloz and Richter (1979), Jeanloz (1988), Bukowinski (1999) and da Silva (2000).)

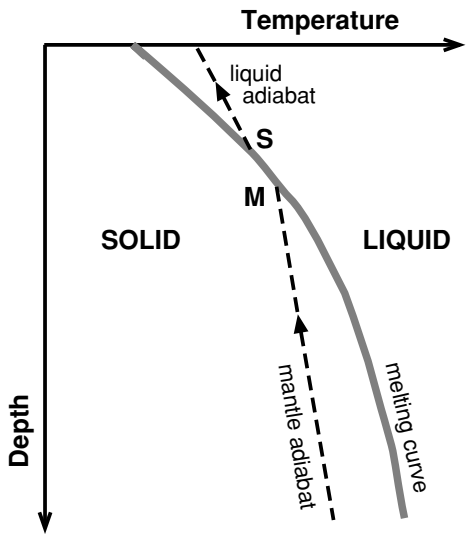


Figure 7.17. Since temperatures in the mantle are approximately adiabatic, the temperature of rising mantle material follows the adiabat. The rising mantle material is initially too cool to melt. At point M, where the mantle adiabat and the melting curve intersect, melting starts. The melting rock then follows the melting curve until the melt separates from the residue S. The melt then rises to the surface as a liquid along a liquid adiabat.

the lower mantle the temperature gradient is also likely to be adiabatic and may be adiabatic down to the boundary layer at the base of the mantle. However, some calculations of the temperature for the basal 500–1000 km of the lower mantle give a gradient significantly greater than that of the adiabat. Calculated values for the mantle temperature close to the core–mantle boundary consequently vary widely, ranging from 2500 K to ~4000 K (Fig. 7.16).

The adiabatic gradient can also tell us much about melting in the mantle. For most rocks, the melting curve is very different from the adiabatic gradient (Fig. 7.17), and the two curves intersect at some depth. Imagine a mantle rock rising along an adiabat. At the depth at which the two curves intersect, the rock will begin to melt and then rises along the melting curve. At some point, the melted material separates from the solid residue and, being less dense, rises to the surface. Since melt is liquid, it has a coefficient of thermal expansion α greater than that of the solid rock. The adiabat along which the melt rises is therefore considerably different from the mantle adiabat (perhaps $1\,^{\circ}\mathrm{C\,km^{-1}}$ instead of $0.4\,^{\circ}\mathrm{C\,km^{-1}}$). One way of comparing the thermal states of rising melts is to define a *potential temperature* T_{p}, which is the temperature an adiabatically rising melt would have at the Earth's surface. T_{p} is therefore the temperature at the theoretical intersection of the adiabat with $z = 0$, the surface. Integrating Eq. (7.94) gives the potential temperature T_{p} for a melt at depth z and temperature T:

$$T_{\mathrm{p}} = T\mathrm{e}^{-\alpha g z/c_P} \tag{7.95}$$

The potential temperature T_{p} is a constant for that melt and so is unaffected by adiabatic upwelling.

Equation (7.94) can also be used to estimate temperature gradients in the outer core, where temperatures are constrained by the solidus of iron (Section 8.3).

Using the values T perhaps 5773 K (5500 °C), estimated from high-pressure melting experiments on iron compounds (see Fig. 7.16(b)); α perhaps $10^{-5}\,^{\circ}\mathrm{C}^{-1}$; g, $6\,\mathrm{m\,s^{-2}}$; and $c_P \sim 7 \times 10^2\,\mathrm{J\,kg^{-1}\,^{\circ}C^{-1}}$ gives an adiabatic gradient of $5 \times 10^{-4}\,^{\circ}\mathrm{C\,m^{-1}}$ ($0.5\,^{\circ}\mathrm{C\,km^{-1}}$). However, because estimates of the ratio α/c_P in the core decrease with depth, the adiabatic gradient in the outer core decreases with depth from perhaps 0.8 to $0.2\,^{\circ}\mathrm{C\,km^{-1}}$. These estimates are just that, being reliable perhaps to within $\pm 0.3\,^{\circ}\mathrm{C\,km^{-1}}$; such is the uncertainty in physical properties of the outer core.

7.8 Metamorphism: geotherms in the continental crust

7.8.1 Introduction

Metamorphism is yet another process that is controlled by the transfer and generation of heat, and understanding of the thermal constraints on metamorphism is important in attempts to deduce past tectonic and thermal settings from the metamorphic evidence available to geologists today. Thus, in this section, considering heat to be transferred by conduction, we study the thermal evolution of some two-dimensional models of the crust.

Two-dimensional thermal models are conceptually easier to understand than one-dimensional models, but, except for a few limited cases, simple analytical solutions to the differential equations are not possible. For the examples shown here, the two-dimensional heat-conduction equation with erosion or sedimentation (Eq. (7.19)) has been solved numerically by finite-difference methods.

Three models are illustrated: a model of burial metamorphism, a model of intrusion and a model of overthrusting. These have been chosen to demonstrate a variety of possible metamorphic environments and by no means represent the possible range existing in the Earth. No metamorphic rock is exposed at the surface without erosion or tectonic accident; but, initially, we discuss hypothetical cases with no erosion or sedimentation.

7.8.2 Two-dimensional conductive models

Burial metamorphism

A model of a typical burial terrain consists of a granitic country rock in which a rectangular trough of sediment has been deposited. Beneath both granite and sediment is a gneissic continental crust overlying the mantle (Fig. 7.18(a)). The initial temperature gradient in the country rock is the equilibrium gradient. Initially, we arbitrarily assume the sediment to be at 100 °C throughout and to have radioactive heat generation of $0.84\,\mu\mathrm{W\,m^{-3}}$. A model such as this could be similar to a sedimentary trough formed on a continent above a subduction zone or to an Archaean greenstone belt filled with thick sediment and set in a granitic terrain. Figure 7.18(b) shows how the model evolves after 20 Ma. The sediment

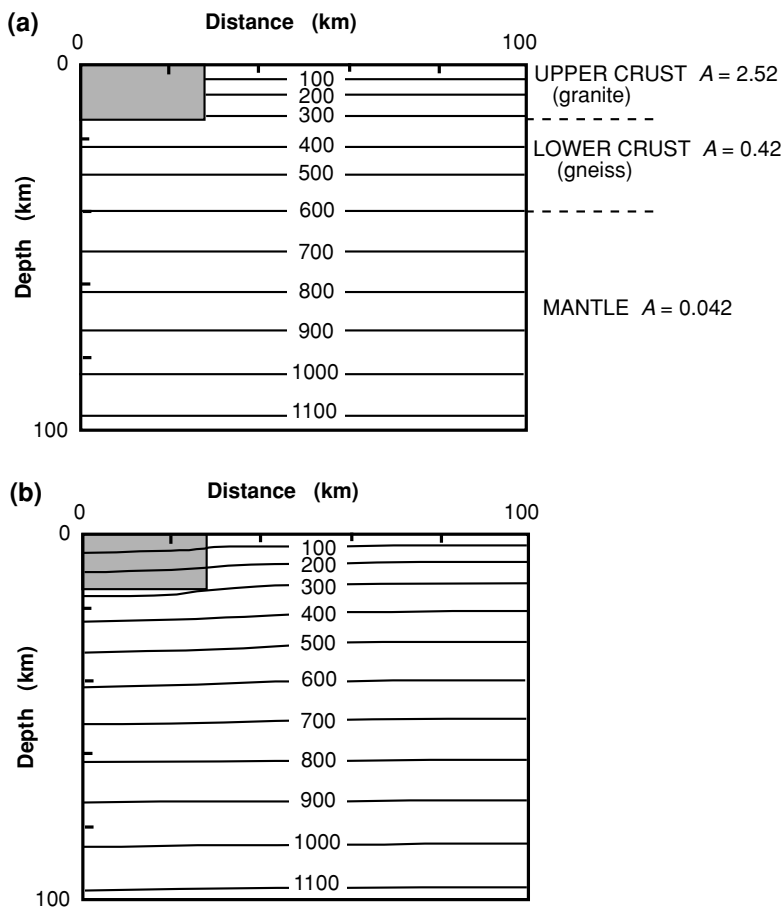


Figure 7.18.
(a) Dimensions and physical parameters of the two-dimensional burial and intrusion models. Initial temperature, equilibrium. The heat generation, A, is in $\mu W\,m^{-3}$. Because the model is symmetrical about the left-hand edge, only half is shown. The shaded region denotes sediment or intrusion. (b) The burial model after 20 Ma. Sediment had an initial temperature of 100 °C and has a radioactive heat generation of $0.84\,\mu W\,m^{-3}$. The sediment has very little effect on the crustal temperatures. (From Fowler and Nisbet (1982).)

rapidly equilibrates towards the temperature of the surrounding country rock and is strongly influenced by the heat production in the surrounding granite. If the country rock had been mafic rather than granitic (see Table 7.1), the equilibrium gradient would have been lower and thus the final temperature at the base of the sediment would be some 50–100 °C lower.

Intrusion metamorphism

Figure 7.18(a) also illustrates a family of models in which a large igneous body is intruded into the country rock. During the period immediately after the intrusion, hydrothermal convection cells occur around the hot body, especially if the intrusion is in a relatively wet country rock. These cells dominate the heat-transfer process, so the simple conductive models considered here should be regarded only as rough guides to the real pattern of metamorphism. Convective heat transfer, which moves heat more quickly than conduction alone, tends to speed up the cooling of the intrusion. Furthermore, it tends to concentrate the metamorphic

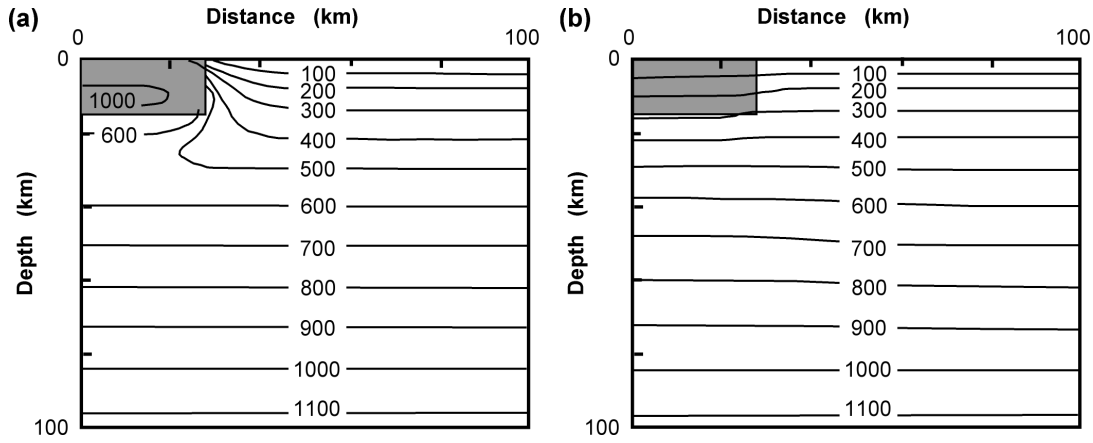


Figure 7.19. The shaded region denotes a large basalt body, which intruded at 1100 °C: (a) 1 Ma after intrusion the basalt body is cooling, and the country rock is being heated; (b) 20 Ma after intrusion the basalt body has solidified and cooled. (From Fowler and Nisbet (1982).)

effects near the source of heat because that is where convection is most active. In granites, radioactive heat generation may prolong the action of convection cells. The presence of water also has profound effects on the mineralogical course of the metamorphism. Nevertheless, simple conductive models are useful for a general understanding of metamorphism around plutons.

Basic intrusion

The igneous body is assumed to intrude at 1100 °C and to have radioactive heat generation of 0.42 μW m^{-3}; the latent heat of crystallization, 4.2×10^5 J kg^{-1}, is released over a 1-Ma cooling interval. Figure 7.19 shows the temperature field after 1 Ma and after 20 Ma. Contact metamorphism of the country rock is an important transient phenomenon, but, if such basic intrusions are to be a major cause of regional as opposed to *local* metamorphism, the intrusions must form a large proportion of the total rock pile.

Granitic intrusion

The granite is assumed to intrude at 700 °C and to have radioactive heat generation of 4.2 μW m^{-3}; the latent heat of crystallization, 4.2×10^5 J kg^{-1}, is released over 2 Ma. Figure 7.20 shows the thermal evolution of this model. It is clear that there is less contact metamorphism than for the basic intrusion, but there is extensive deep-level or regional metamorphism. Indeed, massive lensoid granitic bodies are common in calc-alkaline mountain chains such as the Andes, and may be an

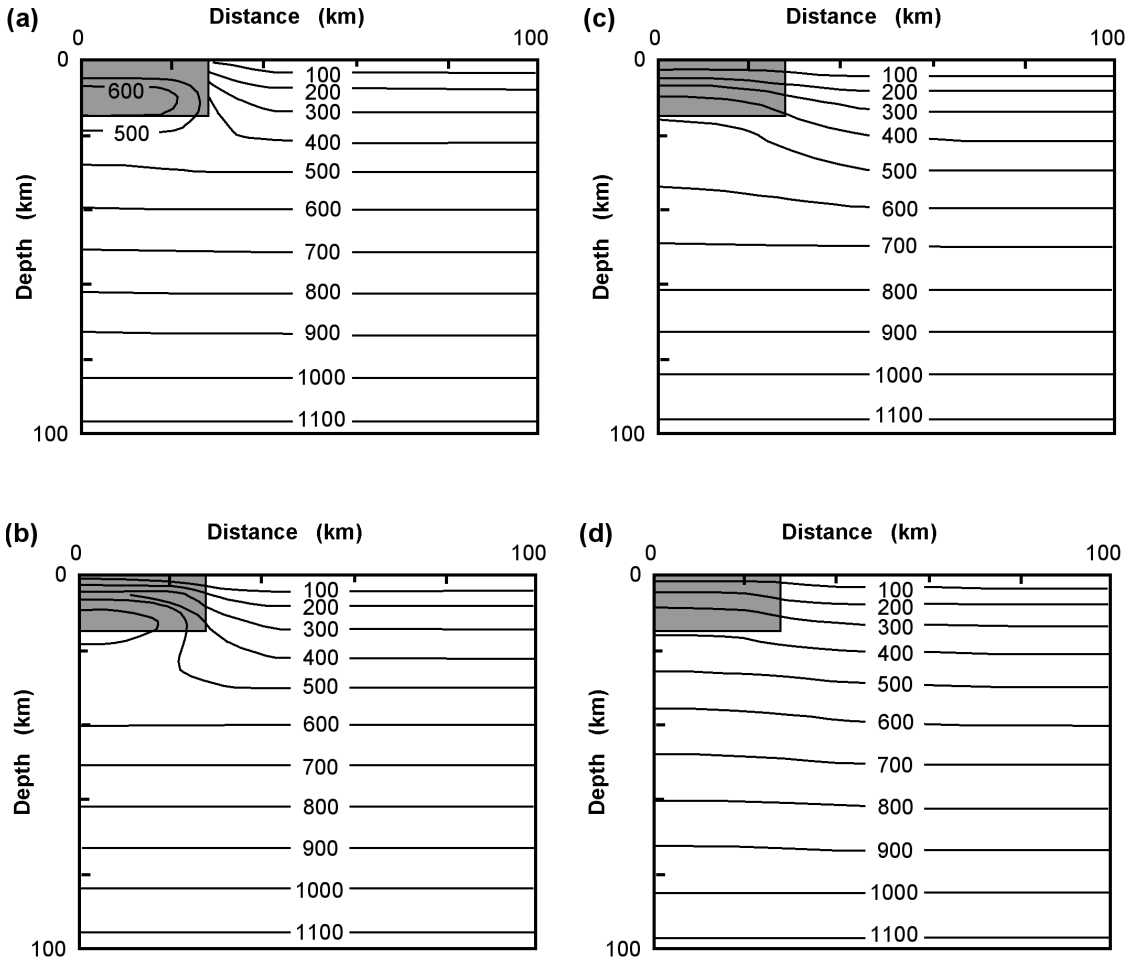


Figure 7.20. The shaded region denotes a large granite body, which intruded at 700 °C: (a) 1 Ma, (b) 2 Ma and (c) 5 Ma after intrusion, the granite is cooling and solidifying and heating the country rock; (d) 20 Ma after intrusion, the high heat generation of the granite means that temperatures in and around it are elevated. (From Fowler and Nisbet (1982).)

important cause of regional metamorphism. Beneath and around the intrusion at depth, temperatures may be raised to such an extent that some local partial melting takes place in the country rock, which may cause further intrusion. If this occurs, then low $^{87}Sr/^{86}Sr$ initial ratios will be measured at the top of the original intrusion; younger, high $^{87}Sr/^{86}Sr$ initial ratios and young partial-melting textures will be present at the base of the pluton and below. Unravelling the history of the intrusion would be very difficult.

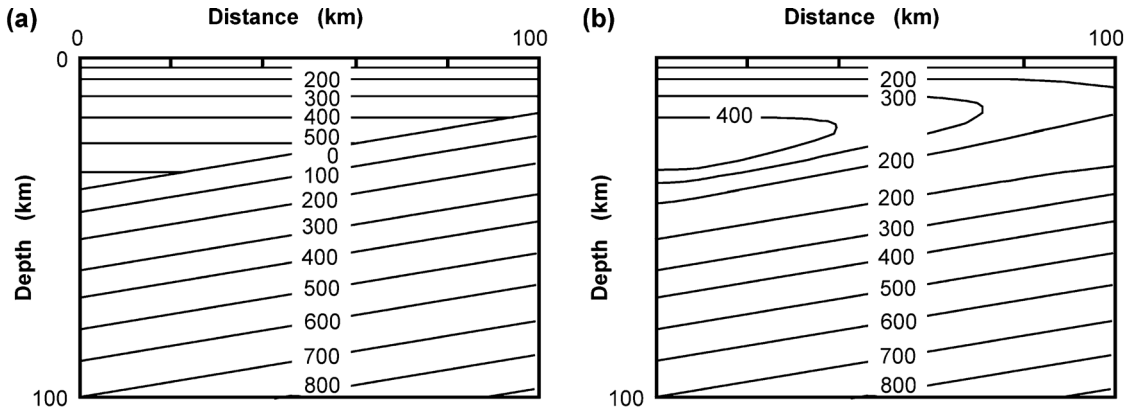


Figure 7.21. (a) Overthrust-model initial temperatures. (b) The overthrust model after 1 Ma of cooling. (From Fowler and Nisbet (1982).)

Overthrusts

A wholly different type of metamorphism is produced by overthrusts. Figure 7.21(a) illustrates an example with a large overthrust slice of granite gneiss material emplaced over mafic rock. Real parallels include subduction zones or an area such as the eastern Alps, where a thick overthrust crystalline block has produced metamorphism below it. In this simple model, thrusting is assumed to be instantaneous. The most interesting feature of this model is that one thrusting event necessarily leads to two very distinct metamorphic events:

1. very early and rapid *retrogression* (cooling) in the upper block and *progression* (heating) beneath, followed by
2. slow progression (heating) throughout and finally partial melting and 'late' intrusion to high levels.

Immediately after thrusting (Fig. 7.21(b)), the hot base of the overthrust block thermally re-equilibrates with the cool underthrust rocks beneath. This initial thermal re-equilibration is very rapid, and inverted thermal gradients are probably very short-lived. The resulting geotherm (temperature–depth curve) in the thrust zone and below is of the order of a few degrees Celsius per kilometre. If the cool lower slab is rich in volatiles, rapid retrograde metamorphism takes place in the upper block. At the same time, equally rapid prograde high-pressure metamorphism occurs in the lower slab. At the deeper end of the overthrust block, local partial melting may take place if large amounts of volatiles move from the lower slab into the hot crust of the upper block. This can produce shallow granitic intrusions.

After this initial re-equilibration comes a long period (perhaps 30–50 Ma, or more, depending on the size of the pile) in which a slow build-up of the geotherm, which in reality would be affected by uplift and erosion, takes place. This is a

period of prograde metamorphism throughout the pile, with the removal of water to higher levels during recrystallization. Finally, partial melting takes place at the base of the pile, and the radioactive heat production is redistributed until thermal stability is reached. This upward redistribution of the radioactive elements is an episodic process. When a partial melt forms, it tends to be rich in potassium, thorium and uranium. Thus, over time this process of melting and intrusion effectively scours the deep crust of heat-producing elements and leads to their concentration in shallow-level intrusions (which would be recognized as 'late' or 'post-tectonic') and in pegmatites. Eventually, after erosion, they tend to be concentrated in sediments and sea water. The net effect is a marked concentration of heat production in the upper crust; whatever the initial distribution of heat production, this leads to the stabilization of the rock pile to a non-melting equilibrium.

7.8.3 Erosion and deposition

Erosion and deposition are two processes that are able to change a geotherm rapidly. They are also interesting to geologists because no sedimentary rock can exist without deposition; neither can any metamorphic rock become exposed at today's surface without erosion. Erosion represents the solar input to the geological machine, and the volcanism and deformation that provide the material to be eroded are driven from the interior. Figure 7.22(a) shows the effect of eroding the model rock column of Fig. 7.3(a) at a rate of 1 km Ma^{-1} for 25 Ma. The shallow-level geotherm is raised to 50 °C km^{-1} after the 25 Ma of erosion, after which it slowly relaxes towards the new equilibrium. If, instead of erosion, the model column is subject to sedimentation, then the shallow-level geotherm is depressed. Sedimentation at 0.5 km Ma^{-1} for 25 Ma depresses the shallow-level geotherm to about 23 °C km^{-1} (Fig. 7.22(b)). After sedimentation ceases, the temperatures slowly relax towards the new equilibrium.

Alternatively, instead of considering the effects of erosion and deposition on the geotherm, we could trace the temperature history of a particular rock (e.g., the rock originally at 30 km depth). In the erosion example, the temperature of this rock decreases, dropping some 500 °C during erosion and a further 200 °C during the re-equilibration. In the depositional example, the temperature of this rock is not affected much during the deposition, but it increases some 400 °C during the subsequent slow re-equilibration.

7.8.4 Erosional models: the development of a metamorphic geotherm

Erosion is essential in the formation of a metamorphic belt since without it no metamorphic rocks would be exposed at the surface. However, as illustrated with the simple one-dimensional model, the process of erosion itself has a

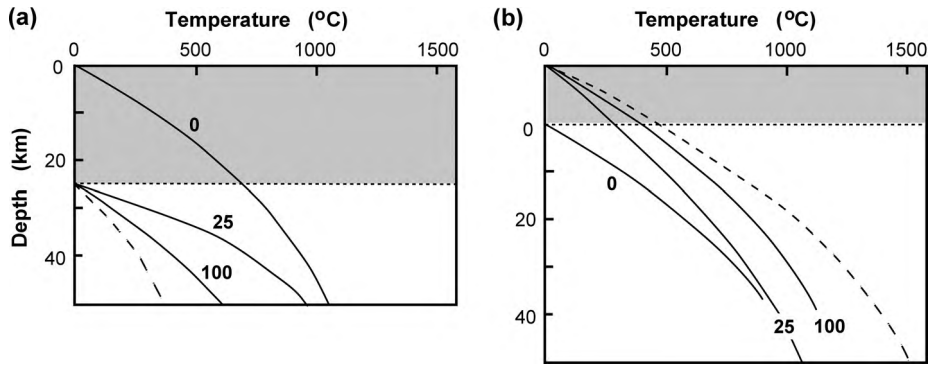


Figure 7.22. Geotherms for the standard model shown in Fig. 7.3(a) for two cases. (a) Erosion at 1 km Ma^{-1} for 25 Ma, and then no further erosion. The dotted horizontal line is the surface after erosion; grey shading shows material eroded. (b) Deposition at 0.5 km Ma^{-1} for 25 Ma. The dotted horizontal line is the original surface before deposition; grey shading shows material deposited. Curve 0, the standard equilibrium geotherm of Fig. 7.3(a); curve 25, the geotherm immediately after erosion/deposition for 25 Ma; curve 100, the geotherm after 100 Ma; and unlabelled dashed line, the final equilibrium geotherm. (After Nisbet and Fowler (1982).)

profound effect on the geotherm (Fig. 7.22) and on the pressure–temperature (P–T) path through which any metamorphic rock passes. The shape of the metamorphic geotherm, that is, the P–T trajectory inferred from the metamorphic rocks exposed at the surface, is also strongly influenced by erosion and, as is shown later, often does not at any time represent the actual geotherm.

The intrusion, burial and overthrust models of the previous sections are now subjected to erosion and deposition to illustrate the effects of these processes. Two erosion models are used for the burial and intrusion models: in space across the model, the first has strong erosion of the country rock and deposition on the trough, whereas the second has strong erosion of the trough and deposition on the country rock. For the overthrust, the erosion is taken to be constant across the model. All erosion rates decay with time.

The burial model with erosion

In the first erosional model, deposition occurs in the centre of the trough at an initial rate of 1.1 km Ma^{-1} while erosion occurs at the edges of the model at an initial rate of 3.3 km Ma^{-1}. Figure 7.23 shows this burial model after 20 Ma when the sedimentary trough has been further covered by sediment and deep erosion has taken place in the country rock. Figure 7.23 also indicates the maximum temperatures attained during the 20 Ma period by the rocks finally exposed at the surface. These are the rocks available to a field geologist. This maximum temperature is not necessarily preserved by the highest-grade minerals, but it is sufficient here to assume that the mineral assemblages exposed at the surface are

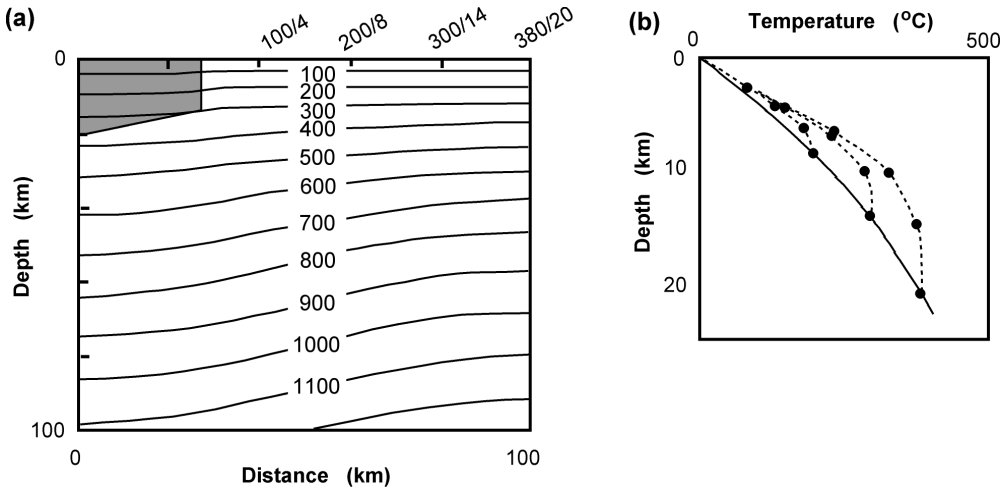


Figure 7.23. (a) The burial model after 20 Ma of erosion of the country rock and deposition on the sediment. Numbers along the surface are the maximum temperature/depth (°C/km) attained by rocks exposed at the surface. (b) Temperature–depth paths followed by rocks originally at 8, 14 and 20 km depth and finally exposed at the surface. The solid line shows the initial equilibrium geotherm in the country rock; dots indicate temperatures every 2 Ma from the start of erosion. (From Fowler and Nisbet (1982).)

those formed when the rock reaches its highest temperature. As it cools, the rock equilibrates to a lower temperature, and the mineral composition alters. However, the reaction kinetics become markedly slower as the rock cools, and there is a good chance of preserving some of the higher-grade minerals if erosion is fast enough.

In this simple model of burial metamorphism, the trough of buried sediment has no heating effect on the country rock. Therefore, the *P–T* curve (which can be plotted from the highest-grade minerals in the exposed rocks) is simply that of the initial, equilibrium thermal gradient in the rock. In this case, the metamorphic geotherm is identical to the equilibrium geotherm, no matter what the rate of erosion (provided that erosion is fast enough to 'quench' the mineral compositions at their highest temperatures). The *metamorphic facies series* (temperatures and pressures recorded in the rocks) produced by the event is that of a normal equilibrium geotherm in the country rock (facies series I in Fig. 7.24).

Intrusion models with erosion

If the country rock is eroded, little metamorphic effect is seen even from these very large intrusions. With the exception of a localized contact zone (of the order of 5 km across) in both cases, the country rock gives a metamorphic facies series identical to the equilibrium facies series 1, and the net result is similar to that in Fig. 7.23(a). A real example of this could be the Great Dyke of Zimbabwe, which

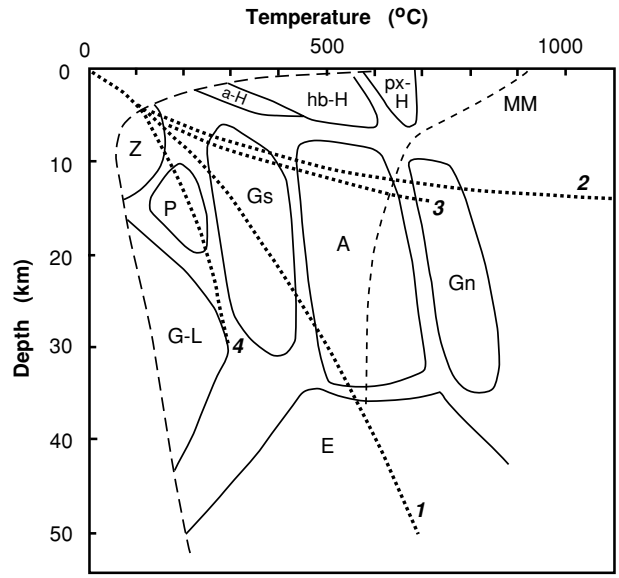


Figure 7.24. Pressure–temperature curves (metamorphic facies series) obtained from exposed rocks in the models after erosion. Thick dotted lines show facies series: curve 1, equilibrium series (Fig. 7.18), high grade; curve 2, basalt intrusion after erosion (Fig. 7.25); curve 3, granite intrusion after erosion (Fig. 7.25); and curve 4, overthrust model (Fig. 7.26), low grade. Facies fields: Z, zeolite; P, prehnite–pumpelleyite; G-L, glaucophane–lawsonite; Gs, greenschist; A, amphibolite; E, eclogite; Gn, granulite; and a-H, hb-H and px-H, albite, hornblende and pyroxene hornfels. Dashed line MM, minimum melting for some metamorphic mineral assemblages. (From Fowler and Nisbet (1982).)

has only a restricted contact zone. Figure 7.23(b) again shows depth–temperature paths followed by individual points exposed on the surface after 20 Ma.

On the other hand, when the intrusion is eroded, and deposition takes place on the country rock, marked effects are seen because deep-seated rocks close to the intrusion are now being eroded. The resulting facies series is one of very low dP/dT (high temperatures at low pressures: facies series 2 for basalt and facies series 3 for granite, as shown in Fig. 7.24). It can be seen from Fig. 7.25 that, although the original intrusions have been almost completely eroded, the metamorphic imprint of intrusion and erosion is widespread and lasting. The small body of granite present today (Fig. 7.25(c)) appears to have had a major metamorphic effect. Figures 7.25(b) and (d) show some depth–temperature paths for points exposed on the surface after 20 Ma.

The overthrust model with erosion

Erosion of the overthrust model provides clues regarding the origin of paired metamorphic belts. For thinner thrust sheets, the results would be similar to those illustrated here but with lower temperatures and pressures.

Initially, only the overthrust unit is exposed at the surface. Since the thrusting event has a general cooling effect on the upper block, the highest temperature any part of the upper block experiences is its initial temperature prior to thrusting. Rocks from the upper block probably retain relict minerals from their previous high-grade environment. The extent of re-equilibration of the mineral is probably dependent on the availability of volatiles rising from the underthrust sheet into

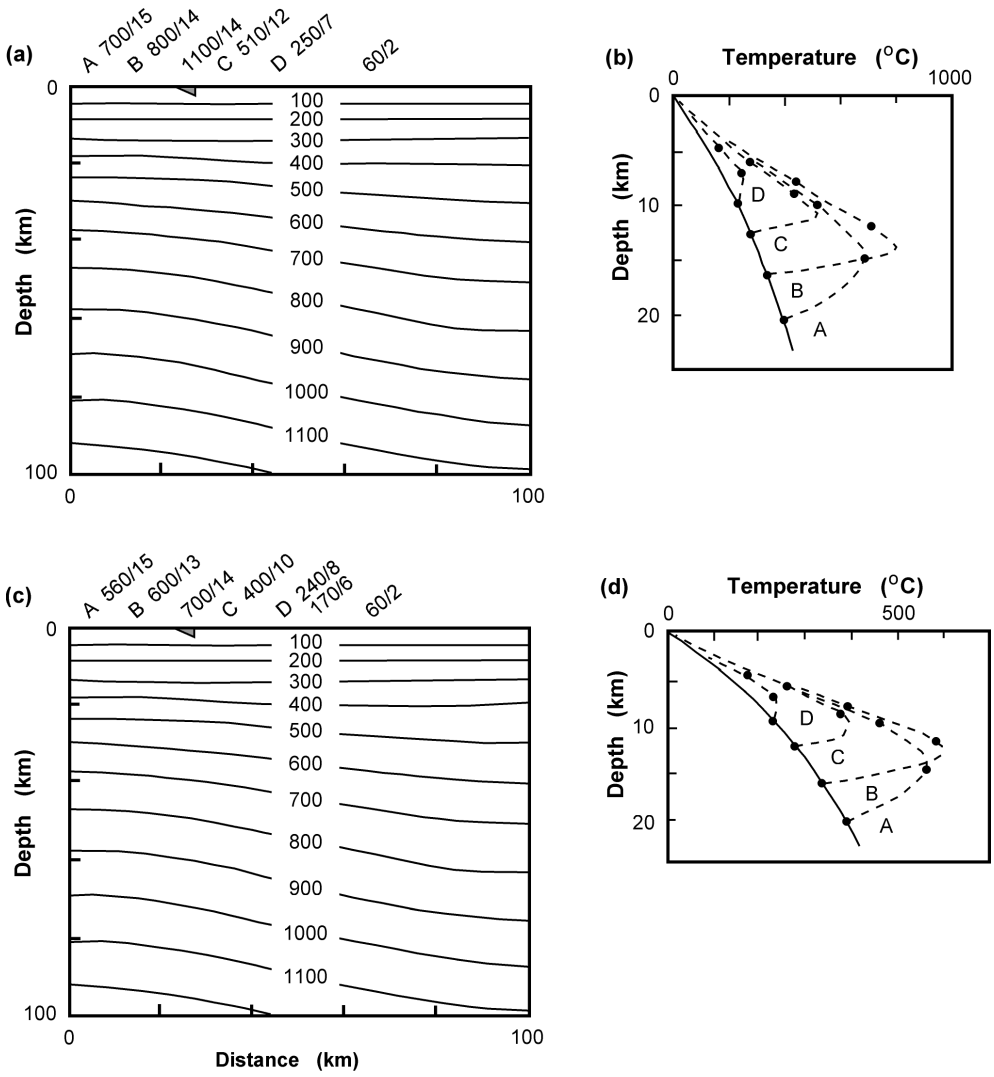


Figure 7.25. (a) The basalt-intrusion model after 20 Ma of erosion of intrusion and country rock. (b) Temperature–depth paths for rocks finally exposed at the surface in model (a). (c) Granite intrusion after 20 Ma of erosion of intrusion and country rock. (d) Temperature–depth paths for rocks finally exposed at the surface in model (c). Notation is as in Fig. 7.23. (From Fowler and Nisbet (1982).)

the hot base of the upper sheet. If local partial melting occurs (e.g., in the upper sheet on the extreme left of the model), prograde effects are seen.

Simultaneously with the metamorphism of the upper block, the underthrust block experiences prograde metamorphism. As erosion takes place, the deep-level rocks are lifted towards the surface. This has a cooling effect, which progressively halts and reverses the rise in temperature of the rocks in the underthrust block.

The maximum temperatures attained in this underthrust block are controlled by the time of initiation and the rate of erosion and also by any shear heating in the upper slab. Eventually the thrust surface is exposed by the erosion, and the following metamorphic events are seen.

1. An early (prior to thrusting) high-grade event in the overthrust block, shown as facies series 1 in Fig. 7.24.
2. A post-thrusting retrograde event in the remaining part of the overthrust block which overprints (1). In real rocks this event is controlled by the introduction of large amounts of volatiles from the underthrust rocks. The overthrust block is rapidly cooled and hydrated, and the degree of overprinting depends on reaction kinetics (exponentially related to temperature) and the availability of volatiles.
3. A high-pressure, low-temperature event in the underthrust rocks, shown as facies series 4 in Fig. 7.24.

An overthrust of the dimensions modelled here produces twin metamorphic belts: one of high grade (facies series 1) and one of low grade (facies series 4). Much smaller overthrusts (e.g., at shallow levels in mountain belts) would be qualitatively similar, but the metamorphic effects might be restricted by kinetic factors or be small or difficult to distinguish.

7.8.5 Dating and metamorphism

The age of a radiometric system such as a mineral generally depends on the way in which the daughter product became sealed into the system as it cooled. For a typical system a blocking temperature exists. This is the temperature below which the system can be thought to be closed (see Section 6.2). Different minerals, dating techniques and rates of cooling all produce different blocking temperatures; thus, we have a powerful tool for working out the thermal history of a mountain belt.

For moderate cooling rates, suggested closure temperatures are 80–110 °C for apatite fission tracks, about 175–225 °C for zircon fission tracks, 280 ± 40 °C for biotite K–Ar and 530 ± 40 °C for hornblende K–Ar (see Table 6.4 and Fig. 6.7). In many cases, the Rb–Sr whole-rock ages probably reflect the original age of the rock. Consider first the basaltic and granitic intrusions illustrated in Fig. 7.19 and 7.20. Underneath the intrusion, the radiometric clocks of minerals that close at about 500 °C would start recording about 1–2 Ma after the start of cooling. Under the basalt intrusion, which has little internal heat generation, minerals that close at about 300 °C would not start recording until 20 Ma. In the case of the granite, a mineral with a blocking temperature of 300 °C would not close at all. The depth–temperature paths in Fig. 7.25 show the great spread in radiometric ages that would be obtained from various mineral ‘clocks’. Furthermore, the blocking temperatures are dependent on the rate of cooling, which is different in each case.

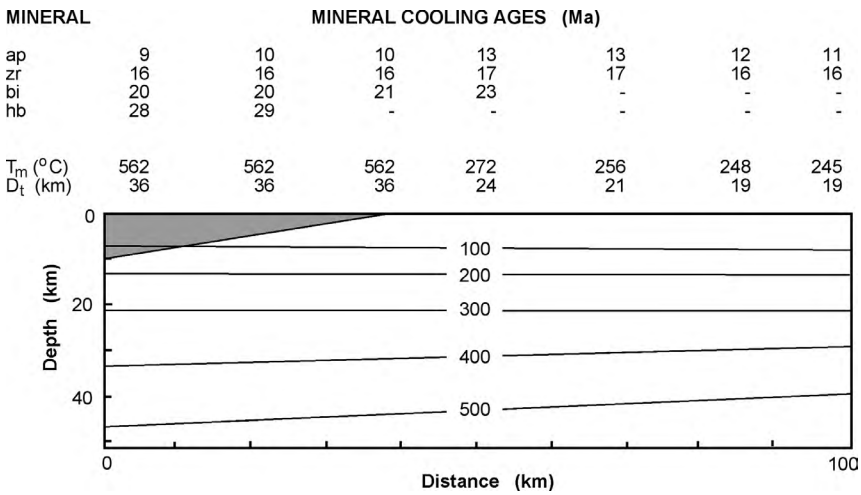


Figure 7.26. The overthrust model after 30 Ma of erosion. Stipple indicates remnants of overthrust block. T_m is the maximum temperature attained by rock finally at the surface; D_t is the depth at which T_m was reached. This figure also shows the times of closure (Ma before the present) of various radiometric systems: ap, apatite fission track; zr, zircon fission track (maximum age based on closure at ~175 °C); bi, biotite K/Ar; and hb, hornblende K/Ar. (From Fowler and Nisbet (1982).)

Figure 7.26 shows the effect of using these various dating techniques across the overthrust model's 30-Ma erosion surface; the dates shown are the dates that would be measured by a geologist working on this surface. It can be seen that these dates are useful for studying the cooling and erosional history of the pile. In many natural examples, there can be profound differences in closure ages of minerals that were initially produced by the same tectonic event. An example of the use of such methods for the Southern Alps of New Zealand is provided in Fig. 7.27. The plate boundary between the Pacific and Indo-Australian plates runs the length of the South Island (Fig. 2.2). Convergence between the plates of over $1\,\mathrm{cm\,yr^{-1}}$ causes comparable uplift in the Southern Alps. Plotting the apparent ages against the closure temperature provides a cooling history for rocks along the plate boundary. This cooling history can be understood in the context of the tectonic setting of the Southern Alps – initially the rocks move horizontally eastwards; only within about 25 km of the Alpine Fault do they start to rise (Fig. 7.27(a)). On the cooling-history graph therefore this appears as almost no cooling until about 2 Ma ago, followed by rapid exhumation.

Implications

To interpret a metamorphic terrain fully, it is not sufficient to know the pressure–temperature conditions undergone by the rocks. A full interpretation of the thermal history of the rocks also involves studying the erosional and radiometric

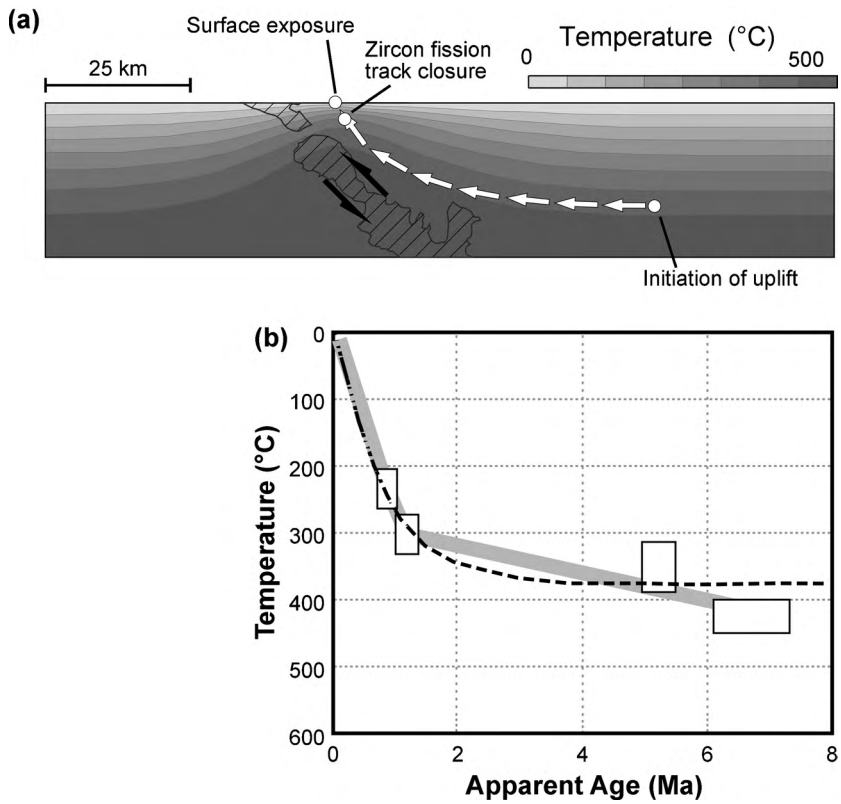


Figure 7.27. The cooling history for the Alpine Fault region, Southern Alps, New Zealand. (a) The modelled exhumation history (white arrows) for a particle moving westwards towards the Alpine Fault. The convergence rate is 1 cm yr^{-1}. The thermal structure shown is that developed after 5 Ma of deformation as appropriate for the Southern Alps. Cross-hatching marks the peak strain and equates with the Alpine Fault. There is no vertical exaggeration. (b) The cooling history for rocks adjacent to the Alpine Fault (white boxes and grey line). The white boxes are (in order of increasing temperature) zircon fission-track, biotite K–Ar, muscovite K–Ar and inferred pre-rift temperatures of the region. The dashed line shows the cooling history of a particle in the numerical model shown in (a). (Reprinted from *Tectonophysics*, **349**, Batt, G. E. and Brandon, M. T., Lateral thinking: 2-D interpretation of thermochronology in convergent orogenic settings, 185–201, Copyright 2002, with permission from Elsevier.)

history of the terrain. These thermal models demonstrate that a general knowledge of the stratigraphy of an area is essential before its metamorphic history can be unravelled.

Convective movement of fluid has not been discussed, but it can be a major factor in heat transport around plutons and during dehydration of overthrust terrains. As a general rule, fluid movement speeds up thermal re-equilibration and reduces to some degree the extent of aureoles around intrusions; at the same time it

promotes metamorphic reactions and has a major impact on the thermodynamics of re-equilibration.

Problems

1. Estimate Table 7.2 for 4000 and 4500 Ma. Discuss the implications of the relative heat-generation values in this table, particularly with respect to the Archaean Earth.
2. Calculate the phase difference between the daily and annual surface temperature variations that would be measured at depths of 2 and 5 m in a sandstone.
3. Taking T_0 as 40 °C for the annual variation in surface temperature, calculate the depth at which the variation is 5 °C. What is the phase difference (in weeks) between the surface and this depth?
4. Calculate an equilibrium geotherm for $0 \leq z \leq d$ from the one-dimensional heat-flow equation, given the following boundary conditions:
   (i) $T = 0$ at $z = 0$ and
   (ii) $T = T_d$ at $z = d$.
   Assume that there is no internal heat generation.
5. Calculate an equilibrium geotherm from the one-dimensional heat-flow equation given the following boundary conditions:
   (i) $\partial T/\partial z = 30\ ^\circ\mathrm{C\,km^{-1}}$ at $z = 0$ km and
   (ii) $T = 700\ ^\circ$C at $z = 35$ km.
   Assume that the internal heat generation is 1 $\mu\mathrm{W\,m^{-3}}$ and the thermal conductivity is $3\ \mathrm{W\,m^{-1}\,^\circ C^{-1}}$.
6. On missions to Venus the surface temperature was measured to be 740 K, and at three sites heat-producing elements were measured (in percentage of total volume) as follows (ppm, parts per million).

	Venera 8	Venera 9	Venera 10
K (%)	0.47 ± 0.08	0.30 ± 0.16	4 ± 1.2
U (ppm)	0.60 ± 0.16	0.46 ± 0.26	2.2 ± 0.2
Th (ppm)	3.65 ± 0.42	0.70 ± 0.34	6.5 ± 2

   The density of the Venusian crust can be taken, from a measurement by Venera 9, to be $2.8 \times 10^3\ \mathrm{kg\,m^{-3}}$. Calculate the heat generation in $\mu\mathrm{W\,m^{-3}}$ at each site. (From Nisbet and Fowler (1982).)
7. Using the one-dimensional equilibrium heat-conduction equation, calculate and plot the Venus geotherms (Aphroditotherms) of Problem 6 down to 50 km depth at each site. Assume that the conductivity is $2.5\ \mathrm{W\,m^{-1}\,^\circ C^{-1}}$ (a typical value for silicates) and that, at a depth of 50 km, the heat flow from the mantle and deep lithosphere of Venus is $21 \times 10^{-3}\ \mathrm{W\,m^{-2}}$. What have you assumed in making this calculation? What do these Aphroditotherms suggest about the internal structure of the planet?

8. Calculate the geotherms for the models shown in Fig. 7.14. Discuss the reason for the difference at depth between these geotherms and the geotherm shown as a solid line in Fig. 7.14.
9. Calculate an equilibrium geotherm for the model Archaean crust shown in Fig. 7.4. Discuss your estimates.
10. To what depth are temperatures in the Earth affected by ice ages? (Use thermal conductivity 2.5 W $m^{-1}\,°C^{-1}$ and specific heat 10^3 J $kg^{-1}\,°C^{-1}$.)
11. Calculate the equilibrium geotherm for a two-layered crust. The upper layer, 10 km thick, has an internal heat generation of 2.5 μW m^{-3}, and the lower layer, 25 km thick, has no internal heat generation. Assume that the heat flow at the base of the crust is 20×10^{-3} W m^{-2} and that the thermal conductivity is 2.5 W $m^{-1}\,°C^{-1}$.
12. Repeat the calculation of Problem 11 when the upper layer has no internal heat generation and the lower layer has internal heat generation of 1 μW m^{-3}. Comment on the effect that the distribution of heat-generating elements has on geotherms.
13. Calculate geotherms for a layered continental crust and comment on the significance of your results for the following cases.
    (a) A 10-km-thick upper layer with heat generation of 2.5 μW m^{-3} overlying a 30-km-thick layer with heat generation of 0.4 μW m^{-3}.
    (b) A 30-km-thick upper layer with heat generation of 0.4 μW m^{-3} overlying a 10-km-thick layer with heat generation of 2.5 μW m^{-3}. For both cases, assume a surface temperature of zero, heat flow from the mantle of 20×10^{-3} W m^{-2} and thermal conductivity of 2.5 W $m^{-1}\,°C^{-1}$.
14. A 1-m-wide dyke with a temperature of 1050 °C is intruded into country rock at a temperature of 50 °C.
    (a) Calculate how long the dyke will take to solidify.
    (b) After two weeks, what will the temperature of the dyke be?
    (Assume a diffusivity of 10^{-5} m^{-2} s^{-1} and a solidus temperature of 800 °C.)
15. Volcanic flood basalts can be several kilometres thick and extend over very large areas (the Karoo basalt in southern Africa is one example). A 2-km-thick basalt is erupted at 1200 °C. If the solidus temperature is 900 °C, estimate the time required for the basalt to solidify. If this basalt is later eroded and the underlying rocks exposed, indicate how far you would expect the metamorphism to extend from the basalt. State all your assumptions in answering this question.
16. (a) Calculate the difference in depth of the seabed at the intersection of a mid-ocean ridge and a transform fault. Assume that the ridge is spreading at 5 cm yr^{-1} and that the ridge axis is offset 200 km by the transform fault.
    (b) Calculate the difference in depth on either side of the same fault 1000 km from the ridge axis and 3000 km from the ridge axis. (See Section 9.5 for information on transform faults.)
17. Calculate the 60-Ma geotherm in the oceanic lithosphere for the simple model of Section 7.5.2. What is the thickness of the 60-Ma-old lithosphere? Use an asthenosphere temperature of 1300 °C and assume a temperature of 1150 °C for the base of the lithosphere

18. Assume that the Earth is solid and that all heat transfer is by conduction. What value of internal heat generation distributed uniformly throughout the Earth is necessary to account for the Earth's mean surface heat flow of 87×10^{-3} W m^{-2}? How does this value compare with the actual estimated values for the crust and mantle?
19. Calculate the rate at which heat is produced in (a) the crust and (b) the mantle. Assume that the crust is 10 km thick and that the volumetric heat-generation rates are 1.5×10^{-6} W m^{-3} in the crust and 1.5×10^{-8} W m^{-3} in the mantle.
20. Calculate the steady-state surface heat flow for a model solid Earth with the following constant thermal properties: k, 4 W m^{-1} °C^{-1}; and A, 2×10^{-8} W m^{-3}.
21. It takes about 4 min to boil a hen's egg of mass 60 g to make it edible for most people. For how long would it be advisable to boil an ostrich egg weighing about 1.4 kg? (From Thompson (1987).)
22. (a) Calculate the conductive characteristic time for the whole Earth.
    (b) Calculate the thickness of the layer that has a characteristic time of 4500 Ma.
    (c) Comment on your answers to (a) and (b).
23. (a) A sphere has radius r and uniform density ρ. What is the gravitational energy released by bringing material from infinitely far away and adding a spherical shell, of density ρ and thickness δr, to the original shell?
    (b) By integrating the expression for the gravitational energy over r from 0 to R, calculate the gravitational energy released in assembling a sphere of density ρ and radius R.
    (c) Use the result of (b) to estimate the gravitational energy released as a result of the accretion of the Earth.
    (d) Assume that all the energy calculated in (c) became heat and estimate the rise in temperature of the primaeval Earth. Comment on your answer.

References and bibliography

Alfe, D., Gillan, J. and Price, G. D. 1999. The melting curve of iron at the pressures of the Earth's core from *ab initio* calculations. *Nature*, **401**, 462–4.

Batt, G. E. and Brandon, M. T. 2002. Lateral thinking: 2-D interpretation of thermochronology in convergent orogenic settings. *Tectonophysics*, **349**, 185–201.

Bloxham, J. and Gubbins, D. 1987. Thermal core–mantle interactions. *Nature*, **325**, 511–13.

Bott, M. H. P. 1982. *The Interior of the Earth: Its Structure, Composition and Evolution*. Amsterdam: Elsevier.

Bukowinski, M. S. T. 1999. Taking the core temperature. *Nature*, **401**, 432–3.

Carlson, R. L. and Johnson, H. P. 1994. On modeling the thermal evolution of the oceanic upper mantle: an assessment of the cooling plate model. *J. Geophys. Res.*, **99**, 3201–14.

Carslaw, H. S. and Jaeger, J. C. 1959. *Conduction of Heat in Solids*, 2nd edn. New York: Oxford University Press.

Célérier, B. 1988. Paleobathymetry and geodynamic models for subsidence. *Palaios*, **3**, 454–63.

Clark, S. P. 1966. Thermal conductivity. In S. P. Clark, ed., *Handbook of Physical Constants*. Vol. 97 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America, pp. 459–82.

Clark, S. P., Peterman, Z. E. and Heir, K. S. 1966. Abundances of uranium, thorium and potassium. In S. P. Clark, ed., *Handbook of Physical Constants*. Vol. 97 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America, pp. 521–41.

da Silva, C. R. S., Wentzcovitch, R. M., Patel, A., Price, G. D. and Karato, S. I. 2000. The composition and geotherm of the lower mantle: constraints from the elasticity of silicate perovskite. *Phys. Earth Planet. Interiors*, **118**, 103–9.

Davis, E. E. and Lister, C. R. B. 1974. Fundamentals of ridge crest topography. *Earth Planet. Sci. Lett.*, **21**, 405–13.

England, P. C. and Richardson, S. W. 1977. The influence of erosion upon the mineral facies of rocks from different tectonic environments. *J. Geol. Soc. Lond.*, **134**, 201–13.

Fowler, C. M. R. and Nisbet, E. G. 1982. The thermal background to metamorphism 11. Simple two-dimensional conductive models. *Geosci. Canada*, **9**, 208–14.

Gubbins, D., Masters, T. G. and Jacobs, J. A. 1979. Thermal evolution of the Earth's core. *Geophys. J. Roy. Astr. Soc.*, **59**, 57–99.

Harrison, T. M. 1987. Comment on 'Kelvin and the age of the Earth'. *J. Geol.*, **95**, 725–7.

Hayes, D. E. 1988. Age–depth relationships and depth anomalies in the southeast Indian Ocean and South Atlantic Ocean. *J. Geophys. Res.*, **93**, 2937–54.

Jeanloz, R. 1988. High pressure experiments and the Earth's deep interior. *Phys. Today*, **41** (1), S44–5.

1990. The nature of the Earth's core. *Ann. Rev. Earth Planet. Sci.*, **18**, 357–86.

Jeanloz R. and Richter, F. M. 1979. Convection, composition and the state of the lower mantle. *J. Geophys. Res.*, **84**, 5497–504.

Jessop, A. M. 1990. *Thermal Geophysics*. Developments in Solid Earth Geophysics 17. Amsterdam: Elsevier.

Jessop, A. M. and Lewis, T. 1978. Heat flow and heat generation in the Superior Province of the Canadian Shield. *Tectonophysics*, **50**, 55–77.

Lachenbruch, A. H. 1970. Crustal temperature and heat production; implications of the linear heat flow relation. *J. Geophys. Res.*, **75**, 3291–300.

McDonough, W. F. and Sun, S. S. 1995. The composition of the Earth. *Chem. Geol.*, **120**, 223–53.

McKenzie, D. P. 1967. Some remarks on heat flow and gravity anomalies. *J. Geophys. Res.*, **72**, 6261–73.

1969. Speculations on the consequences and causes of plate motions. *Geophys. J. Roy. Astr. Soc.*, **18**, 1–32.

McKenzie, D. and Bickle, M. J. 1988. The volume and composition of melt generated by extension of the lithosphere. *J. Petro.*, **29**, 625–79.

McKenzie, D. P. and O'Nions, R. K. 1982. Mantle reservoirs and ocean island basalts. *Nature*, **301**, 229–31.

McLennan, S. M. and Taylor, S. R. 1996. Heat flow and the chemical composition of continental crust. *J. Geol*., **104**, 369–77.

Morgan, P. 1984. The thermal structure and thermal evolution of the continental lithosphere. In H. N. Pollack and V. R. Murthy, eds., *Structure and Evolution of the Continental Lithosphere*. Physics and Chemistry of the Earth 15. Oxford: Pergamon Press, pp. 107–93.

Nisbet, E. G. and Fowler, C. M. R. 1982. The thermal background to metamorphism. I. Simple one-dimensional conductive models. *Geosci. Canada*, **9**, 161–4.

Nyblade, A. A. and Pollack, H. N. 1993. A global analysis of heat-flow from Precambrian terrains – implications for the thermal structure of Archaean and Proterozoic lithosphere. *J. Geophys. Res.*, **98**, 12 207–18.

Oldenberg, D. W. 1975. A physical model for the creation of the lithosphere. *Geophys. J. Roy. Astr. Soc.*, **43**, 425–52.

Parker, R. L. and Oldenberg, D. W. 1973. Thermal models of mid-ocean ridges. *Nature Phys. Sci.*, **242**, 137–9.

Parsons, B. and McKenzie, D. P. 1978. Mantle convection and the thermal structure of the plates. *J. Geophys. Res.*, **83**, 4485–96.

Parsons, B. and Sclater, J. G. 1977. An analysis of the variation of ocean floor bathymetry and heat flow with age. *J. Geophys. Res.*, **32**, 803–27.

Pollack, H. N., Hurter, S. J. and Johnson, J. R. 1993. Heat loss from the Earth's interior – analysis of the global data set. *Rev. Geophys.*, **31**, 267–80.

Richter, F. M. 1986. Kelvin and the age of the Earth. *J. Geol.*, **94**, 395–401.

Richter, F. M. and McKenzie, D. P. 1978. Simple plate models of mantle convection. *J. Geophys.*, **44**, 441–71.

Richter, F. M. and Parsons, B. 1975. On the interaction of two scales of convection in the mantle. *J. Geophys. Res*., **80**, 2529–41.

Roy, R. F., Decker, E. R., Blackwell, D. D. and Birch, F. 1968. Heat flow in the United States. *J. Geophys. Res.*, **73**, 5207–21.

Rudnick, R. L., McDonough, W. F. and O'Connell, R. J. 1998. Thermal structure, thickness and composition of continental lithosphere. *Chem. Geol.*, **145**, 395–411.

Sclater, J. G., Jaupart, C. and Galson, D. 1980. The heat flow through oceanic and continental crust and the heat loss of the Earth. *Rev. Geophys. Space Phys.*, **18**, 269–311.

Sclater, J. G., Parsons, B. and Jaupart, C. 1981. Oceans and continents: similarities and differences in the mechanisms of heat loss. *J. Geophys. Res.*, **86**, 11 535–52.

Stacey, F. D. 2000. Kelvin's age of the Earth paradox revisited. *J. Geophys. Res.*, **105**, 13 155–8.

Stein, C. A. 1995. Heat flow of the Earth. In *Global Earth Physics: A Handbook of Physical Constants*. AGU Reference Shelf 1. Washington: American Geophysical Union, pp. 144–58.

Stein, C. A. and Stein, S. 1992. A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature*, **359**, 123–9.

1994. Constraints on hydrothermal heat flux through the oceanic lithosphere from global heat flow. *J. Geophys. Res.*, **99**, 3081–95.

1996. Thermo-mechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes (overview). In G. E. Bebout, D. W. Scholl, S. H. Kirkby and J. P. Platt, Subduction: Top to Bottom. Geophysical Monograph 96. Washington: American Geophysical Union, pp. 1–17.

Stephenson, D. J. 1981. Models of the Earth's core. *Science*, **214**, 611–19.

Thompson, N., ed., 1987. *Thinking Like a Physicist: Physics Problems for Undergraduates*. Bristol: Adam Hilger.

Thompson, W. 1864. On the secular cooling of the Earth. *Trans. Roy. Soc. Edinburgh*, **XXIII**, 157–69.

Turcotte, D. L. and Schubert, G. 2002. *Geodynamics*, 2nd edn. Cambridge: Cambridge University Press.

Chapter 8
The deep interior of the Earth

8.1 The internal structure of the Earth

8.1.1 Seismic velocities for the whole Earth

The variations of seismic velocities in the Earth are determined from the travel-time curves for seismic waves (Fig. 4.16). There are two approaches to this determination: one is an inverse method and the other is forward. In the inverse problem used in the early determinations, such as that of Jeffreys (see Appendix 3 for details), the velocities are obtained directly from the travel times. In the forward method, a velocity–depth model is assumed and travel times calculated and compared with observations. The model is then adjusted until agreement at the desired level is attained. The surface-wave-dispersion and normal-mode data (Sections 4.1.3 and 4.1.4) are also used in velocity determination. Standard radial reference velocity models PREM (Fig. 8.1) and *iasp91* are listed in Appendices 8 and 9. Other radial-velocity models have been constructed to have the best fit to P- and S-wave arrival times (e.g., *SP6* and *ak135*). The zones in which particular refinements of the structure are still taking place are the low-velocity zone in the upper mantle, the core–mantle boundary region and the transition zone between the inner core and the outer core – everywhere else the gross P- and S-wave velocities are known to within better than $\pm 0.01\ \mathrm{km\,s^{-1}}$.

Crust

The seismic structure of the continental crust is variable, but it has an average thickness of 35 km and an average P-wave velocity of about $6.5\ \mathrm{km\,s^{-1}}$ (see Section 10.1.2). The oceanic crust is thinner, 7–8 km thick, with an average P-wave velocity of more than $6\ \mathrm{km\,s^{-1}}$ (see Section 9.2.1).

Mantle

The discontinuity between the crust and mantle, which is compositional (Section 8.1.4), is called the Mohorovičić discontinuity. The normal P-wave velocity at the top of the mantle is $8.1\ \mathrm{km\,s^{-1}}$.

The uppermost mantle is very heterogeneous, its structure being dependent upon plate processes and history. There does not seem to be a universal

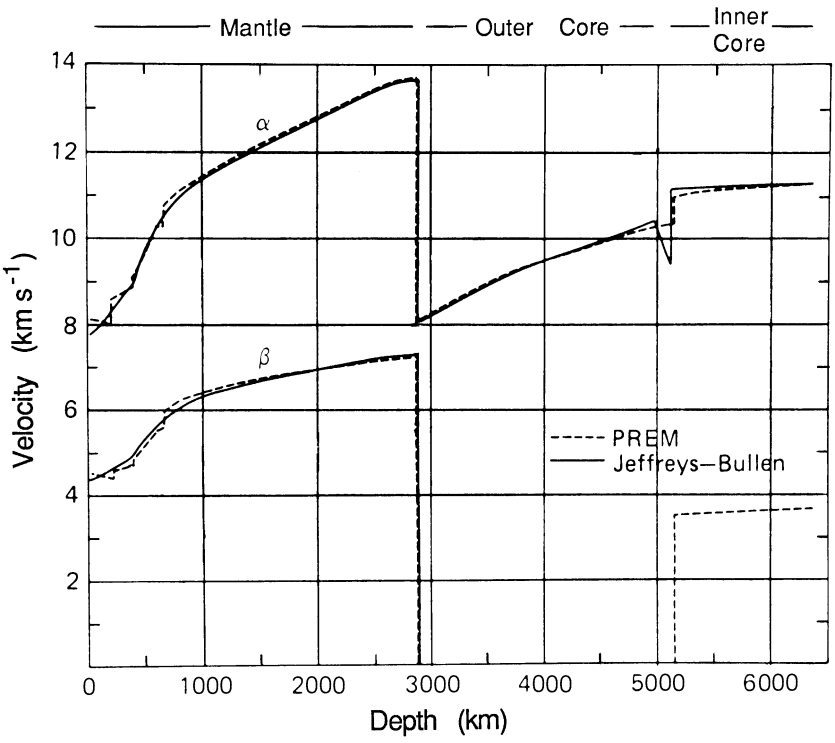


Figure 8.1. Seismic velocity–depth models for the whole Earth. Since the early determination by Jeffreys (1939), which was based on the Herglotz–Wiechert inversion of the Jeffreys–Bullen compilation of travel-time and angular-distance data, there have been many revisions, but the agreement among them is good. The two regions where the models have been most revised and refined are the low-velocity zone in the upper mantle (asthenosphere) and the inner-core–outer-core transition zone. The Preliminary Reference Earth Model (PREM) of Dziewonski and Anderson (1981) was determined by a joint inversion of the free oscillation periods of the Earth, its mass and moment of inertia as well as the travel-time–distance data. (After Bullen and Bolt (1985).)

discontinuity at 220 km, but the region above 220 km is sometimes referred to as the *lid*. Standard velocity models (e.g., Fig. 8.1) vary in representation of the uppermost mantle depending upon the data used and the assumptions made. A *low-velocity zone* for S-waves down to about 220 km is well established by the surface-wave-dispersion data. The low-velocity zone for P-waves is based on a shadow-zone effect for P-waves out to about 15° (Fig. 8.2) and on a matching of waveforms of P-wave arrivals with synthetic seismograms computed for possible velocity structures and shows up in PREM. In contrast, *iasp91* has no low-velocity zone for P- or S-waves. Beneath the low-velocity zone, P- and S-wave velocities increase markedly until about 400 km depth. At depths of 400 and 670 km, there are sharp changes in velocity; both P- and S-wave velocities increase by 5%–7%.

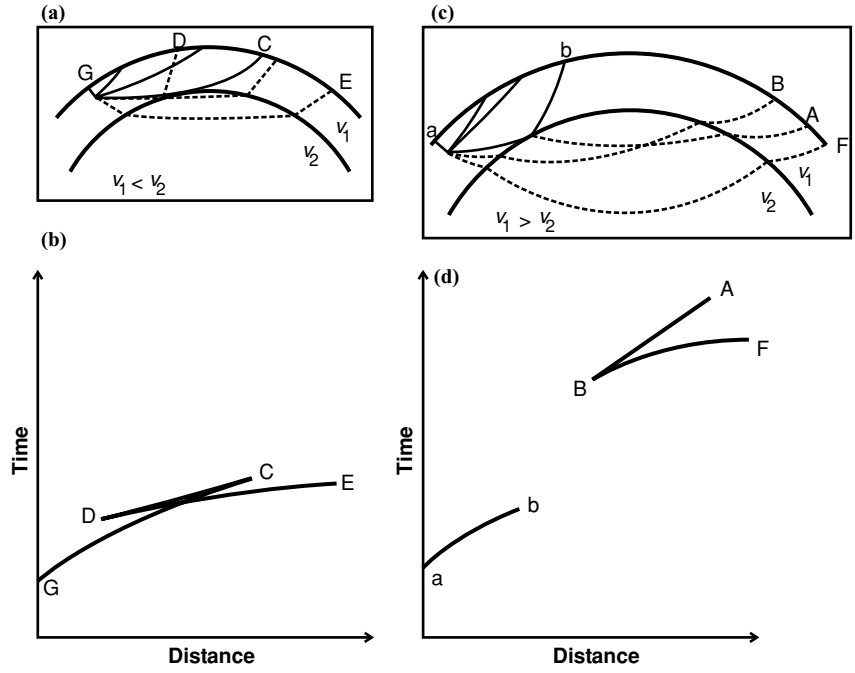


Figure 8.2. The 'shadow zone' resulting from a low-velocity zone. As an example, consider a two-layered sphere for which the seismic velocity increases gradually with depth in each layer. The seismic velocity immediately above the discontinuity in the upper layer is V_1 and that immediately below the discontinuity is V_2. The ray paths for the case $V_2 > V_1$ (the velocity increases at the discontinuity) are shown in (a). If $V_2 < V_1$ (the velocity decreases at the discontinuity, resulting in a low-velocity zone at the top of the second layer), then the ray refracted into the inner layer bends towards the normal (Snell's law), yielding the ray paths shown in (c). The travel-time–distance curves for (a) and (c) are shown in (b) and (d), respectively. When $V_2 > V_1$, arrivals are recorded at all distances, but when $V_2 < V_1$, there is a distance interval over the shadow zone. The angular extent of the shadow zone (b to B) and the corresponding delay in travel time (b to B) are dependent on the depth and extent of the low-velocity zone and on the reduction of velocity in the low-velocity zone. (After Gutenberg (1959).)

These increases have been verified independently by computing synthetic seismograms to match earthquake and nuclear-explosion amplitudes and waveforms. Earthquake activity in subduction zones ceases at about 670 km depth, and this depth is also commonly taken as the boundary between upper and lower mantle. Global maps of topography on this discontinuity between upper and lower mantle reveal variations of up to 30 km. The depressions in this discontinuity seem to be correlated with subduction zones, suggesting that it provides some impediment to the continuation of subduction into the lower mantle. The entire region between 400 and 670 km depth is often called the *mantle transition zone*. The

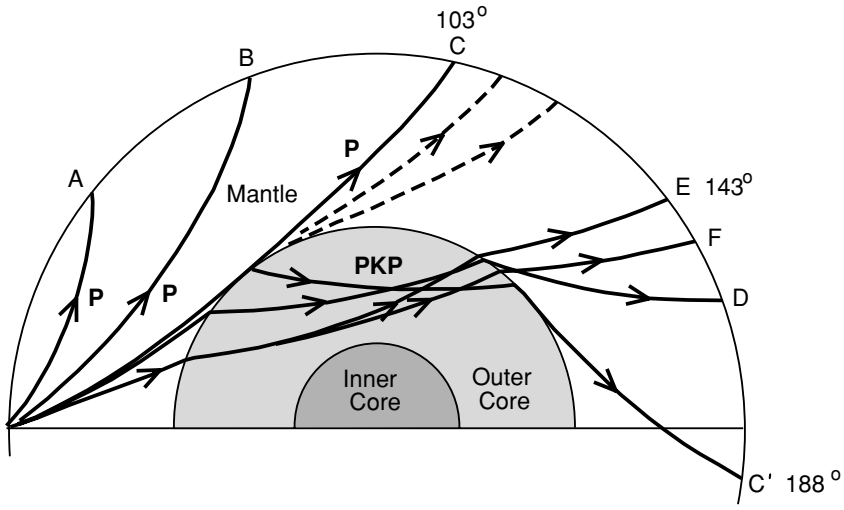


Figure 8.3. Ray paths for PKP, the direct P-wave passing through the mantle and outer core. The mantle P-wave (C) which has grazing incidence on the core has an epicentral distance of 103°. Beyond this distance, there can be no direct mantle P-waves, although PP and PPP can be recorded at greater distances, and a weak diffracted P can be recorded out to about 120°. Because there is a sharp decrease in velocity for P-waves refracted into the core, rays bend towards the normal at the mantle–core boundary and give rise to a *shadow zone* for P-waves. The PKP ray with the shallowest angle of incidence on the outer core (C′) is refracted and finally emerges at an epicentral distance of 188°. With increasing angle of incidence (C′, D, E, F), PKP rays emerge at epicentral distances decreasing to 143° and then increasing to about 155°. Each ray penetrates deeper into the outer core than does its predecessor. Thus, no direct P-waves are recorded at epicentral distances between 103° and 143°; this is the shadow zone. At 143° there is a caustic; the amplitude of PKP is large (this shows clearly in Fig. 4.18). (Based on Gutenberg (1959).)

lower mantle at depths down to 2700 km is referred to as the D′ shell.[1] The lowermost 150–200 km of the mantle (depth ~2700–2900 km) is referred to as the D″ shell. Velocities increase slowly with depth through the lower mantle. The direct P-wave through the mantle can be observed out to 103°. At epicentral distances between 103° and 120°, a weak P-wave is diffracted (Section 4.4.4) at the core–mantle boundary into what is called the *shadow zone* (Fig. 8.3). There is evidence that velocity gradients are much reduced in the D″ shell. This could be due to chemical heterogeneity and interaction between the core and mantle and/or to a thermal boundary layer that would conduct, not convect, heat (refer to Sections 8.2 and 8.3): if the temperature at the base of the mantle is about 3000 K, there is

[1] The names D′ and D″ are the sole survivors of a labelling of the internal layering of the Earth from A to G introduced by Bullen in 1947. A was the crust; B the mantle down to 400 km depth; C the mantle from 400 to 1000 km depth; D the mantle from 1000 km depth to the core; E and F the outer core; and G the inner core.

a temperature contrast between the mantle and core of 1000 ± 500 K (Fig. 7.16). In this sense the D″ layer is similar to the lithosphere, the boundary layer at the top of the mantle. The D″ layer is a region of considerable lateral heterogeneity: lateral variations of up to 4% are suggested by data on diffracted P- and S-waves. Some variations can be explained by temperature anomalies of 200–300 K, others require temperature anomalies coupled with local variation in the silicate/oxide ratio.

Core

At the core–mantle boundary (CMB, also known as the *Gutenberg discontinuity* after its discoverer) the P-wave velocity drops sharply from about 13.7 to about 8.1 km s^{-1}, and the S-wave velocity drops from about 7.3 km s^{-1} to zero. This structure is determined by the strong reflections PcP, ScS and so on. The P-wave velocity increases slowly through the outer core until the boundary of the inner core. This is determined mainly by the rays PKP and SKS. However, since PKP rays do not sample the outermost core (Fig. 8.3), the velocities there are based on SKS rays. The inner core was discovered in 1936 by Inge Lehmann, a Danish seismologist (who died in 1993 aged 104), using seismograms from an earthquake near Buller on the Southern Alpine Fault in New Zealand. She realized that, to explain particular phases (observed at epicentral distances greater than 120° with travel times of 18–20 min), the core must contain a distinct inner region. The phases she identified are then explained as being refractions through the higher-velocity inner core (PKIKP in today's notation), which therefore arrive earlier than does the PKP phase. The depth of the inner-core–outer-core transition can be determined from the travel times of PKiKP (the reflection from this transition), and the velocity increase/velocity gradient occurring controls the amplitude of this reflected arrival. It has been suggested that the boundary between inner and outer core should be termed the *Lehmann discontinuity*. The name has been used for a discontinuity at ~220 km depth beneath North America, but it seems most appropriate in the core.

A zero S-wave velocity for the outer core, which is consistent with its being liquid, is in agreement with studies on the tides, which require a liquid core. This conclusion is supported by all other seismological evidence and, indeed, is essential if the Earth's magnetic field (and its secular variation) is to be accounted for by convection currents in the outer core (Section 8.3.2). There is a transition zone at the outer-core–inner-core boundary. The increase in velocity shown in Fig. 8.1 is a feature both of PREM and of *iasp91*; the low-velocity zone of the early Jeffreys–Bullen model is not required. The P-wave velocity increases from 10.36 to 11.03 km s^{-1} and the S-wave velocity from zero to 3.50 km s^{-1} at the outer-core–inner-core transition.

The P- and S-wave velocities are both almost constant through the inner core. The structure of the inner core is mainly determined by using the ray path PKIKP (Fig. 8.4). The phases with an S-wave leg (J) through the inner core are very hard

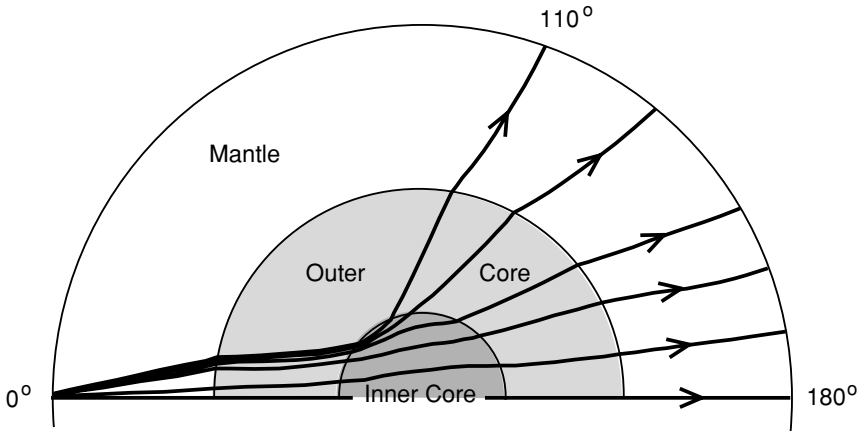


Figure 8.4. Ray paths for PKIKP, the direct P-wave passing through the mantle, outer core and inner core. (Based on Gutenberg (1959).)

to observe because they have a very low amplitude[2] and are easily masked by interferences of other phases. The first clear identification of J phases, pPKJKP and SKJKP, was made for a 1996 earthquake in the Flores Sea.[3] The S-wave velocity of the inner core (3.50–3.67 km s^{-1}) can, however, be determined using normal-mode data.

8.1.2 Density and elastic moduli for the whole Earth

The variation of seismic velocity with depth has been discussed in the previous section. To understand the internal structure of the Earth and its composition further, it is also necessary to know how density and the elastic moduli vary with depth.

We have already seen (Eqs. (4.4) and (4.5)) that the bulk or compressibility modulus K, shear modulus μ and density ρ are related to the P- and S-wave velocities by

$$a = \sqrt{\frac{K + \frac{4}{3}\mu}{\rho}} \tag{8.1}$$

$$\beta = \sqrt{\frac{\mu}{\rho}} \tag{8.2}$$

Even if we know that α and β vary with depth in the Earth, these two equations alone cannot tell us how K, μ and ρ vary with depth because they contain three unknowns. A third equation, which allows us to determine these three unknowns, is the *Adams–Williamson* equation.

[2] There is very poor conversion of P- to S-waves at the outer-core–inner-core boundary: the amplitude of an inner-core S-wave will be some five times smaller than that of a comparable inner-core P-wave. In addition, the attenuation of short-period waves in the inner core is high.

[3] The Flores Sea earthquake occurred on 17 June 1996 (depth 584 km, M_w 7.9). The identification of inner-core shear phases has been achieved for observations of this earthquake (Deuss *et al*. 2000).

Let us assume that the Earth is made up of a series of infinitesimally thin, spherical shells, each with uniform physical properties. The increase in pressure $\mathrm{d}P$ which results during the descent from radius $r + \mathrm{d}r$ to radius r is due only to the weight of the shell thickness $\mathrm{d}r$:

$$\mathrm{d}P = -g(r)\rho(r)\mathrm{d}r \tag{8.3}$$

where $\rho(r)$ is the density of that shell and $g(r)$ the acceleration due to gravity[4] at radius r.

On writing Eq. (8.3) in the form of a differential equation, we have

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -g(r)\rho(r) \tag{8.4}$$

where $\mathrm{d}P/\mathrm{d}r$ is simply the gradient of the hydrostatic pressure. There is a minus sign in Eqs. (8.3) and (8.4) because the pressure P decreases as the radius r increases. The gravitational acceleration at radius r can be written in terms of the gravitational constant G and M_r, the mass of the Earth within radius r:

$$g(r) = \frac{GM_r}{r^2} \tag{8.5}$$

Therefore, Eq. (8.4) becomes

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{GM_r\rho(r)}{r^2} \tag{8.6}$$

To determine the variation of density with radius, it is necessary to determine $\mathrm{d}P/\mathrm{d}r$.

Using Eq. (8.6), we can write

$$\frac{\mathrm{d}\rho}{\mathrm{d}r} = \frac{\mathrm{d}P}{\mathrm{d}r}\frac{\mathrm{d}\rho}{\mathrm{d}P} \tag{8.7}$$

$$= -\frac{GM_r\rho(r)}{r^2}\frac{\mathrm{d}\rho}{\mathrm{d}P} \tag{8.8}$$

The compressibility or bulk modulus for adiabatic compression K (Eq. (A2.31)) is used to obtain $\mathrm{d}\rho/\mathrm{d}P$, the variation of density with pressure, as follows:

$$K = \frac{\text{increase in pressure}}{\text{fractional change in volume}}$$

$$= -\frac{\mathrm{d}P}{\mathrm{d}V/V} \tag{8.9}$$

There is a minus sign in Eq. (8.9) because volume decreases as pressure increases. Since density ρ is the ratio of mass to volume,

$$\rho = \frac{m}{V} \tag{8.10}$$

[4] Outside a spherical shell the gravitational attraction of that shell is the same as if all its mass were concentrated at its centre. Within a spherical shell there is no gravitational attraction from that shell. Together, the preceeding statements mean that, at a radius r within the Earth, the gravitational attraction is the same as if all the mass inside r were concentrated at the centre of the Earth. All of the mass outside radius r makes no contribution to the gravitational attraction and so can be ignored. This is proved in Section 5.2.

we can write

$$\frac{\mathrm{d}\rho}{\mathrm{d}V} = -\frac{m}{V^2} = -\frac{\rho}{V} \tag{8.11}$$

Substituting this into Eq. (8.9) gives

$$K = \rho \frac{\mathrm{d}P}{\mathrm{d}\rho} \tag{8.12}$$

Equation (8.8) can therefore be written

$$\frac{\mathrm{d}\rho}{\mathrm{d}r} = -\frac{GM_r\rho(r)}{r^2}\frac{\rho(r)}{K} \tag{8.13}$$

Combining Eqs. (8.1) and (8.2) gives

$$\frac{K}{\rho} = \alpha^2 - \frac{4}{3}\beta^2 \tag{8.14}$$

so that

$$\frac{\mathrm{d}\rho}{\mathrm{d}r} = -\frac{GM_r\rho(r)}{r^2\left(\alpha^2 - \frac{4}{3}\beta^2\right)} = -\frac{GM_r\rho(r)}{r^2\phi} \tag{8.15}$$

where ϕ, the seismic parameter, is equal to $\alpha^2 - \frac{4}{3}\beta^2$,

Equation (8.15) is the *Adams–Williamson equation* and is used to determine density as a function of radius. To use the equation, it is necessary to start at the Earth's surface and to work inwards, applying the equation successively to shells of uniform composition. It is important to remember that M_r is the mass within radius r:

$$M_r = 4\pi \int_{a=0}^{a=r} \rho(a)a^2 \,\mathrm{d}a \tag{8.16}$$

$$M_r = M_\mathrm{E} - 4\pi \int_{a=r}^{a=R} \rho(a)a^2 \,\mathrm{d}a \tag{8.17}$$

where M_E is the mass of the Earth (known from study of periods of rotation of satellites and from direct measurements of gravity, see Chapter 5) and R is the radius of the Earth. Thus, at each stage of the calculation, working from the Earth's surface inwards, all the terms on the right-hand side of the equation refer to material outside or at the radius r and so have already been determined.

A density structure for the whole Earth obtained by using the Adams–Williamson equation is called a *self-compression model* because the density at each point is assumed to be affected only by compression by the material above it. In practice, it is pointless to assume that the whole Earth has uniform composition since we know this to be untrue, so the density determination is usually begun at the top of the mantle, with a chosen density (the crustal thickness and density vary widely). Then densities can be calculated all the way to the base of the mantle using the Adams–Williamson equation. The core clearly has a different composition from the mantle since the dramatic changes in seismic velocity occurring at the core–mantle boundary could hardly be due to pressure alone. A new starting density is therefore chosen for the top of the core, and the densities

down through the core can be calculated using the Adams–Williamson equation. Because the total mass in the model must equal the mass of the Earth, successive guesses at the density at the top of the core must be made until this constraint is satisfied.

Although such a self-compression density model for the Earth satisfies the seismic-velocity data from which it was derived, it does not satisfy data on the rotation of the Earth. In particular, the Earth's moment of inertia, which is sensitive to the distribution of mass in the Earth, is significantly greater than the moment of inertia for the self-compression model. (To appreciate the importance of mass distribution, test the difference between opening the refrigerator door when all the heavy items in the door compartments are next to the hinge and when they are all next to the handle). There must be more mass in the mantle than the self-compression model allows. To determine the reasons for this discrepancy, it is necessary to re-examine the assumptions made in determining the Adams–Williamson equation. First, it was assumed that the temperature gradient in the Earth is adiabatic (Section 7.7). However, since we know that convection is occurring both in the mantle (Section 8.2) and in the liquid outer core (Section 8.3), the temperature gradients there must be superadiabatic. Equation (8.15) can be modified to include a non-adiabatic temperature gradient:

$$\frac{\mathrm{d}\rho}{\mathrm{d}r} = -\frac{GM_r\rho(r)}{r^2\phi} + \alpha\rho(r)\tau \tag{8.18}$$

where α is the coefficient of thermal expansion and τ the difference between the actual temperature gradient and the adiabatic temperature gradient. This modification means that, in the case of a superadiabatic gradient ($\tau > 0$), the density increases more slowly with depth. Conversely, in the case of a subadiabatic temperature gradient, the density increases more rapidly with depth. This means that corrections for the temperature gradient, which, in practice, are found to be fairly small, act in the opposite direction to that required to explain the missing mantle mass. Another explanation must be found.

The second assumption made in deriving the Adams–Williamson equation was that there were no chemical or phase changes in the Earth (other than differences amongst crust, mantle and core, which have already been included in the model). This assumption provides the answer to the problem of the missing mantle mass. In the mantle transition zone (400–1000 km), there are jumps in seismic velocity that seem to be due to changes of state (phase changes). An example of a change of state is the change from liquid to solid such as occurs when water freezes. This is not the only type of change of state possible; there are also solid–solid phase changes in which the atoms in a solid rearrange and change the crystal structure. Examples of this are the change of carbon from graphite to diamond under increasing pressure and the changes which take place in the transformation from basalt to greenschist to amphibolite to pyroxene granulite (at high temperatures) or to blueschist (at low temperatures) and finally to eclogite

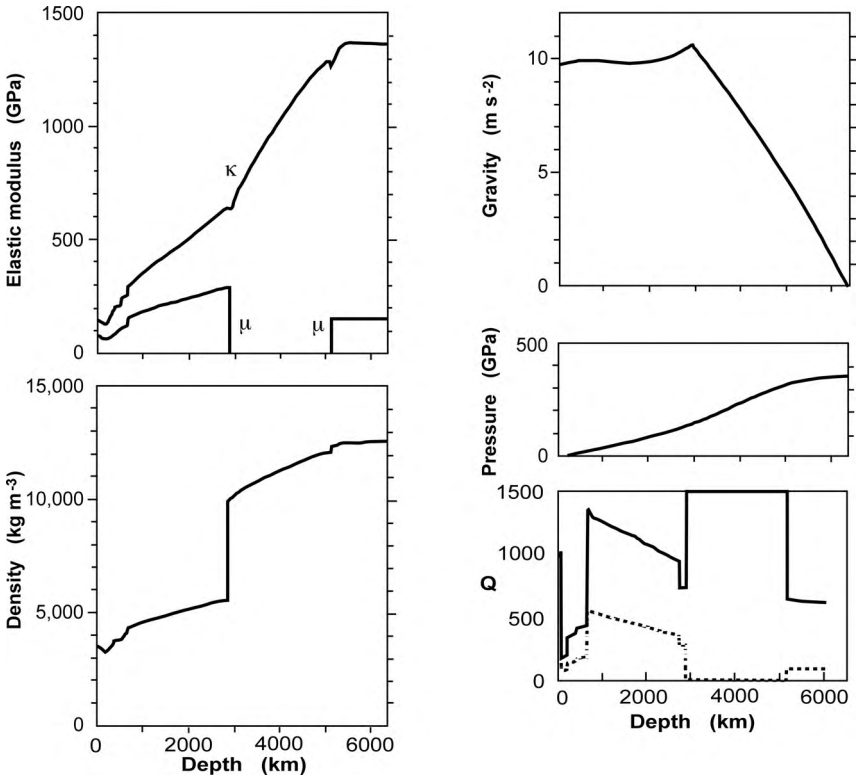


Figure 8.5. The elastic moduli K (bulk), μ (rigidity), density ρ, acceleration due to gravity g, pressure P and quality factors Q_s (dashed line) and Q_p (solid line) in the interior of the Earth. The elastic moduli and pressure are given in GPa. To convert GPa into kilobars (kbar), multiply by ten. Thus, the pressure at the centre of the Earth is 361.7 GPa or 3617 kbar. (After Hart *et al.* (1977), Anderson and Hart (1976) and Montagner and Kennett (1996).)

under increasing pressure and temperature. The phase changes thought to occur in the transition zone are olivine to spinel and pyroxene to garnet, at about 400 km, and spinel to post-spinel forms, at about 700 km. Increases in density of about 10% are associated with these phase changes and are sufficient to account for the self-compression model's low estimate of the mantle mass.

Figure 8.5 shows a density model for the Earth that is based on the Adams–Williamson equation and the additional constraints provided by free oscillations, moment of inertia and total mass. These models are continually being updated and modified, but the densities are unlikely to change substantially from those shown here, though the details of the model, particularly in the transition zone and inner core, may alter. After a density model has been determined, it is straightforward to work backwards using Eqs. (8.1) and (8.2) to determine the elastic moduli (Fig. 8.5).

$$\mu = \rho\beta^2 \tag{8.19}$$

$$K = \rho\alpha^2 - \tfrac{4}{3}\beta^2 \tag{8.20}$$

Table 8.1 provides a comparison of the volume, mass and density of the Earth taken region by region. We know most about the structure of the crust, yet it

Table 8.1 *Volume, mass and density of the Earth*

			Volume		Mass		
	Depth (km)	Radius (km)	(10^{18} m^3)	(%)	10^{21} kg	(%)	Density[a] (10^3 kg m^{-3})
Crust	0–Moho	Moho–6371	10	0.9	28	0.5	2.60–2.90
Upper mantle	Moho–670	5701–Moho	297	27.4	1064	17.8	3.38–3.99
Lower mantle	670–2891	3480–5701	600	55.4	2940	49.2	4.38–5.56
Outer core	2891–5150	1221–3480	169	15.6	1841	30.8	9.90–12.16
Inner core	5150–6371	0–1221	8	0.7	102	1.7	12.76–13.08
Whole Earth	0–6371	6371–0	1083	100	5975	100	

[a] After Dziewonski and Anderson (1981).

constitutes only 0.5% of the total by volume and 0.3% by mass. Uncertainty increases with depth and mass.

8.1.3 Attenuation of seismic waves

In a perfectly elastic medium no elastic energy would be lost during the passage of a seismic wave. However, in practice the Earth is not perfectly elastic, and some energy is dissipated (i.e., turned into heat) as a seismic wave passes. The amount of energy lost as a seismic wave passes through any medium is used to define a parameter Q for that medium. The *quality factor* Q is defined as

$$Q = -\frac{2\pi \times \text{elastic energy stored in the wave}}{\text{energy lost in one cycle or wavelength}} \tag{8.21}$$

Thus, for a perfectly elastic material Q is infinite, whereas for a totally dissipative medium Q is zero. A highly attenuative region in the Earth is often referred to as a low-Q region.

Equation (8.21) can be written in differential form as

$$Q = -\frac{2\pi E}{T\,\mathrm{d}E/\mathrm{d}t}$$
$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{2\pi E}{QT} \tag{8.22}$$

where E is energy, t time and T the period of the seismic wave. Integrating Eq. (8.22) gives

$$E = E_0 \mathrm{e}^{-2\pi t/(QT)} \tag{8.23}$$

where E_0 was the energy of the wave time t ago. Alternatively, since the amplitude of the wave A is proportional to the square root of its energy E (Section 4.2.6),

Eq. (8.23) can be written

$$A = A_0 \mathrm{e}^{-\pi t/(QT)} \tag{8.24}$$

or

$$A = A_0 \mathrm{e}^{-\omega t/(2Q)}$$

where ω is the angular frequency and A_0 the amplitude of the wave time t ago. By performing similar calculations on the spatial form of Eq. (8.21), one obtains

$$A = A_0 \mathrm{e}^{-\pi x/(Q\lambda)} \tag{8.25}$$

where x is the distance travelled by the wave from the point at which it had amplitude A_0, and λ is the wavelength. Thus, after one allows for geometrical spreading, Q can be estimated by taking the ratio of the amplitudes of a body wave of a particular frequency at various distances or times. The quality factor determined by using Eqs. (8.24) and (8.25) is for one particular wave type (P or S) only. Q for P-waves, Q_p, is higher than Q for S-waves, Q_s; in general Q_p is approximately twice Q_s. Figure 8.5 shows the variation of Q_p, and Q_s within the Earth.

8.1.4 The three-dimensional structure of the Earth

Much more detailed velocity models of the mantle can be obtained by using seismic tomography, a technique similar in method to the whole-body scanning method used by medical physicists. The technique requires a network of digital seismic stations. The number of such stations (compared with the WWSSN) meant that the method was previously not capable of resolving structures in the Earth on a horizontal scale of less than about 1500 km and a vertical scale of about 200 km. This has now changed with the establishment of digital seismographic networks. First, the travel times, phase and/or group velocities and/or waveforms are measured for hundreds of earthquakes and recording stations. A best-fitting three-dimensional model of the velocity structure of the mantle is then constructed; the methods are varied and complex and are summarized in Romanowicz (1991) and Ritzwoler and Lively (1995). There are several current tomographic models of the mantle, each determined using differing seismic phases, methods and approximations. There is, however, general agreement amongst them on the broad structure of the mantle (http://mahi.ucsd.edu/Gabi/rem.html). Figures 8.6(a) and (b) (Plates 9 and 10) show perturbations of an S-wave velocity from a one-dimensional standard Earth structure. The model comprises eighteen layers of thickness ~100 km in the upper mantle and 200 km in the lower mantle, each with an equal surface area ($4^\circ \times 4^\circ$ at the equator). The long-wavelength velocity perturbations that can be seen in

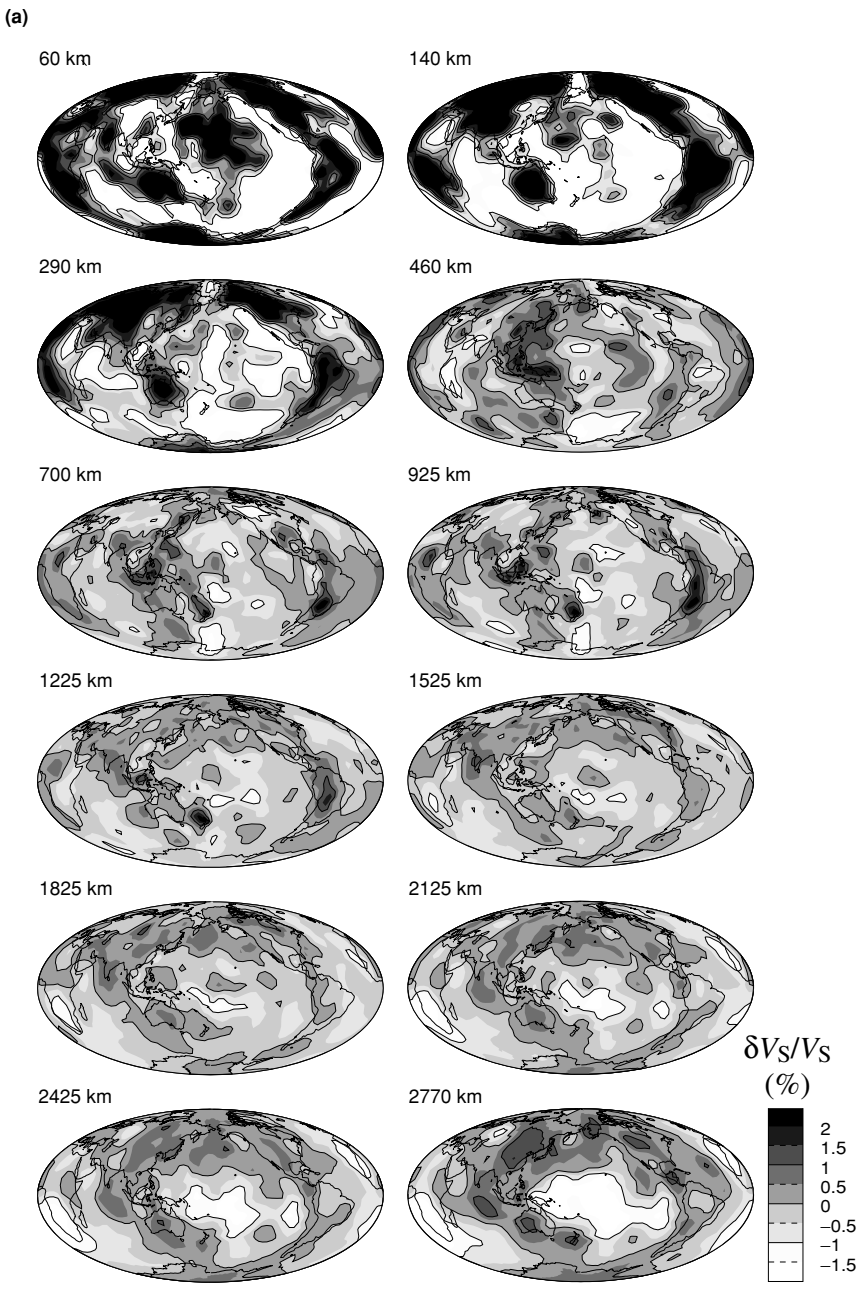


Figure 8.6. (a) Long-wavelength perturbations of S-wave velocity from a standard whole-Earth model at depths of 60, 140, 290, 460, 700, 925, 1225, 1525, 1825, 2125, 2425 and 2770 km. The model was calculated using surface-wave phase-velocity maps, free-oscillation data and long-period body-wave travel times. The gradation in shading indicates increasing perturbation from the standard model. The maximum deviations decrease from $\geq$2% at the top of the upper mantle to $\pm$1% in the lower

this model are of similar magnitude to the velocity jumps in the upper mantle (at 220, 410 and 670 km). It is clear that the old, cold continental shields have high-velocity roots or 'keels' that extend down for several hundred kilometres (Fig. 8.6(a) at 290 km), whereas at these depths the oceanic areas have low velocities indicative of the asthenosphere. The young, hot mid-ocean-ridge systems are associated with very low velocities as plate tectonics would predict (Fig. 8.6(a) at 60 km): the Canadian Shield has the largest positive anomaly and the East Pacific Rise the largest negative anomaly. At 60 km depth the oldest oceanic lithosphere in the northwestern Pacific shows up clearly as having high velocities. However, the continents are not all underlain by high-velocity mantle – although North and South America have high velocities, Asia is characterized by low velocities. By 700 km depth the mantle beneath the subduction zones has, for the most part, high velocities and the oceanic regions have low velocities. Continuing down into the lower mantle, there is a change in the wavelength of the heterogeneities: wavelengths are shorter in the outer part of the lower mantle than in the upper mantle. The old, cold subducting Tethys and Farallon slabs show up as high-velocity zones in the outer parts of the lower mantle (Fig. 8.6(a) at 925–1525 km; see also Fig. 9.60(a)). The basal 800 km or so of the lower mantle is characterized by a merging of these shorter-wavelength anomalies into extensive lateral anomalies. There appear to be two slow regions at the base of the lower mantle, one beneath Africa and the other beneath the Pacific, and two fast linear regions, one beneath India and the other beneath the Americas, which almost appear to be encircling the Pacific. It is tempting to interpret these slow and fast anomalies as being hot upwelling zones (plumes) and cold descending regions (a 'graveyard' for subducting plates). Figure 8.6(c) (Plate 11) shows a comparison between two lower-mantle body-wave models – one using P-wave travel times and the other using S-wave travel times. There is good agreement between them. Figure 8.7 shows the effects of a change in temperature on seismic velocities. Some velocity anomalies suggest that variations in temperature of up to $\pm 250\,^{\circ}\mathrm{C}$ may be present in the mantle. The standard colour scheme for tomographic images has low velocities red and high velocities blue, which is consistent with differences in velocity being caused by variations in temperature. Figure 8.8 (Plate 13) shows an attenuation (Q) model for the upper mantle in terms of variation in the logarithm of $1/Q$. Areas with higher than normal Q (low attenuation) show up as blue on these maps while areas with lower than normal Q (high attenuation) show up as

←

mantle and then increase to $\pm 2\%$ just above the CMB. Colour version Plate 9. (From Masters *et al*. The relative behaviour of shear velocity, bulk sound speed and compressional velocity in the mantle: implications for chemical and thermal structure. Geophysical Monograph 117, 2000. Copyright 2000 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

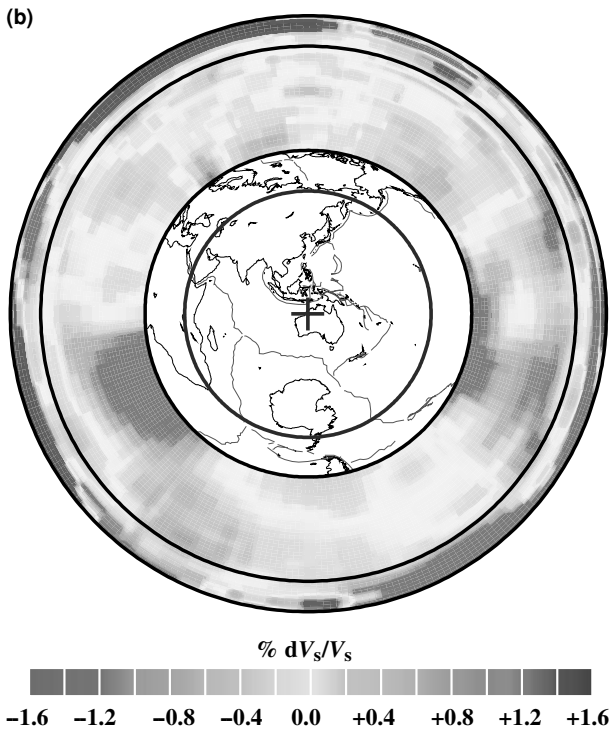


Figure 8.6. (b) The long-wavelength S-wave velocity-perturbation model viewed as a slice through the centre of the Earth along the great circle shown as the blue circle in the central map. Note the two extensive slow (red) anomalies beneath Africa and the central Pacific that start at the CMB, the base of the lower mantle. Black circle, boundary between upper and lower mantle at 670 km. Colour version Plate 10. (T. G. Masters, personal communication 2003.
http://mahi.ucsd.edu/Gabi/rem2.dir/shear-models.htm.)

orange. At shallow depths in the upper mantle the mid-ocean-ridge system has low Q while the continental regions have high Q.

Higher-resolution, global body-wave models of the mantle have until recently been limited by (i) the quality of the short-period data and (ii) uncertainties in earthquake locations. However, there have been significant advances and

Figure 8.6. (c) A comparison of P- and S-wave velocity models for the lower mantle at depths of 800, 1050, 1350, 1800, 2300 and 2750 km. The models are shown as perturbations from a standard whole-Earth model. The P-wave model was calculated using 7.3 million P and 300 000 pP travel times from ~80 000 well-located teleseismic earthquakes, which occurred between 1964 and 1995 (van der Hilst *et al.* 1997). The S-wave model used 8200 S, ScS, Ss, SSS and SSSS travel times (Grand 1994). The number at the side of each image indicates the maximum percentage difference from the standard model for that image. White areas: insufficient data sampling. Colour version Plate 11. (From Grand and van der Hilst, personal communication 2002, after Grand *et al.* (1997).)

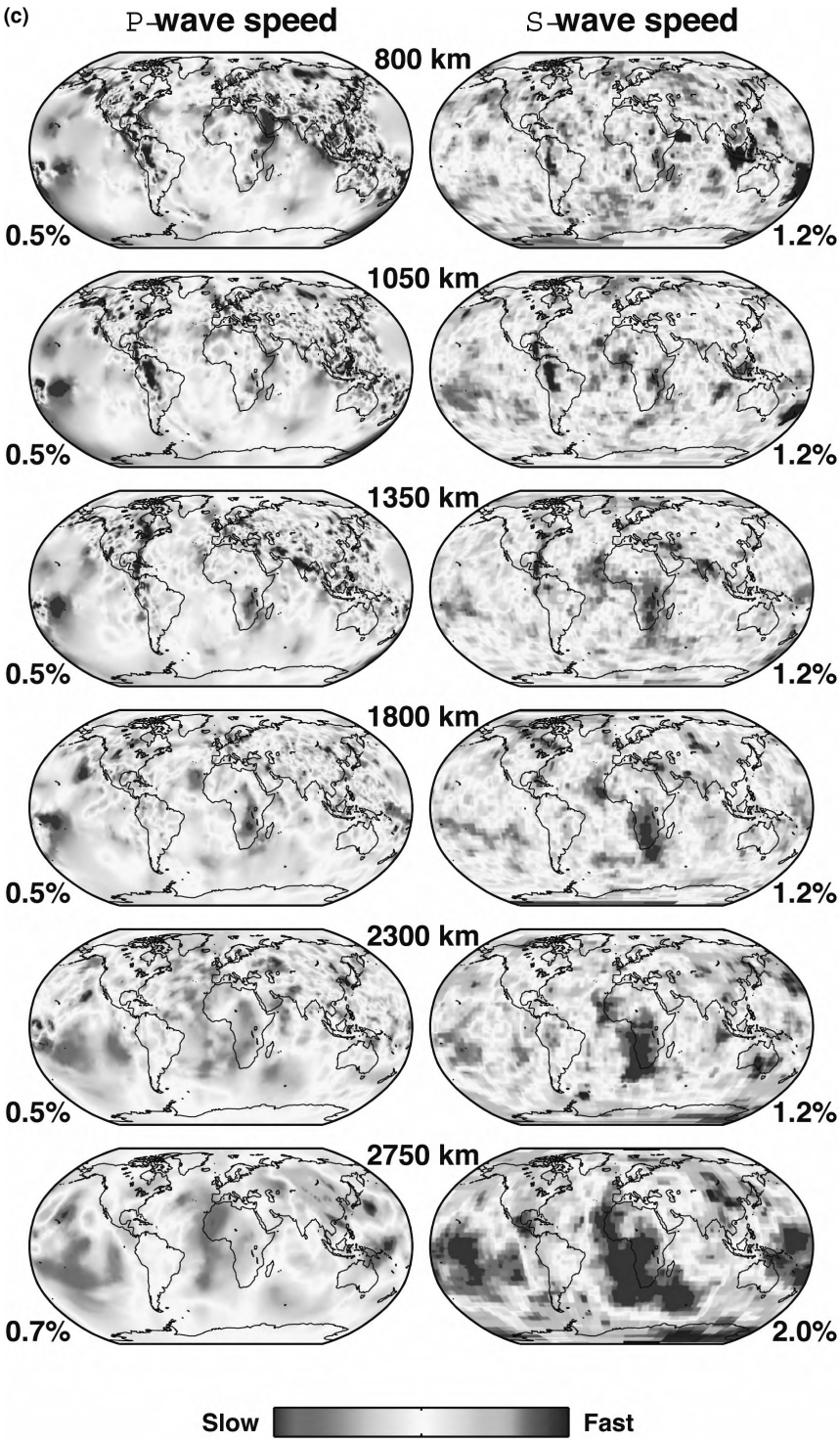


Figure 8.6(c). (*cont.*)

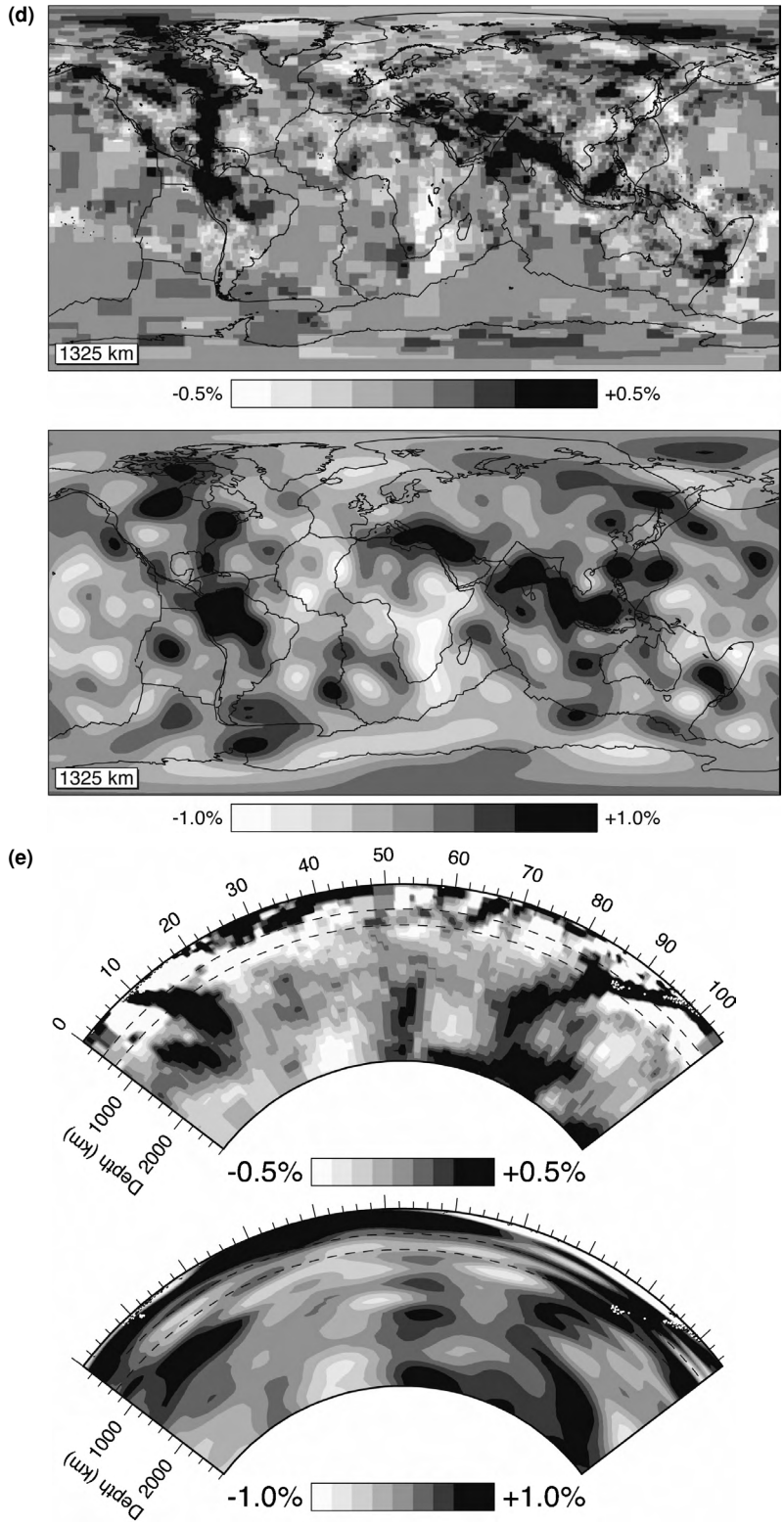


Figure 8.6. (d) A comparison of perturbations from a standard whole-Earth model shown at 1325 km depth. Colour version Plate 12(a). (e) A cross section from the Aegean (left) to Japan (right). Colour version Plate 12(b). Upper panels for (d) and (e), are P-wave mode based on travel times from Bijwaard *et al.* (1998). Lower panels are degree 20 S-wave model S20RTS based on Rayleigh-wave dispersion from Ritsema and van Heijst (2000). Note the different perturbation scales and the different resolutions for the two images. (W. Spakman, personal communication 2003.)

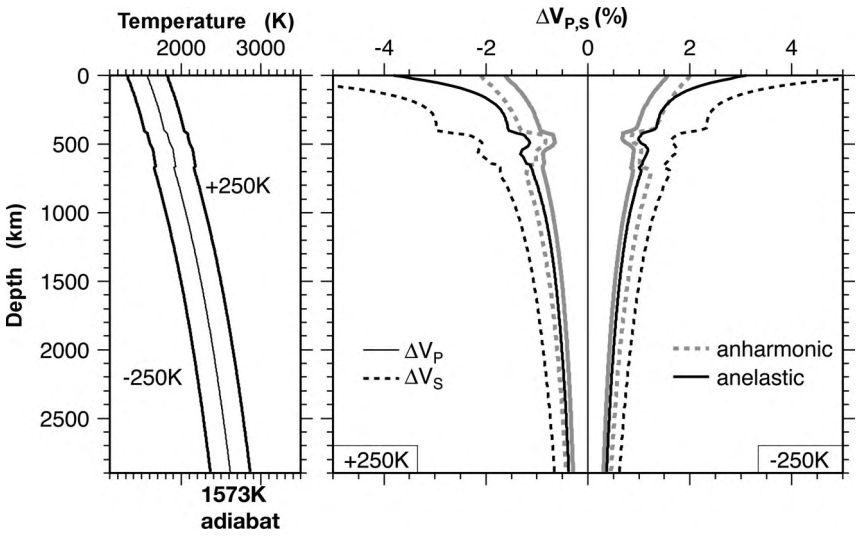


Figure 8.7. The change in seismic velocity that results from a change in mantle temperature. Left-hand panel: mantle adiabats for 250 °C either side of a mantle adiabat with 1300 °C potential temperature. Right-hand panel, percentage changes in P- and S-wave velocities that result from the changes in temperature shown on the left. The grey lines are for anharmonic effects only; the black lines also include anelasticity. The S-wave velocity is more sensitive to temperature than is the P-wave velocity. (Saskia Goes, personal communication 2004.)

high-resolution studies can now be performed. The models from these studies are more detailed than that shown in Fig. 8.6(a); large numbers of constant-velocity blocks and the shorter-wavelength signals allow the detail of major anomalies such as subducting plates to be investigated. Overall the results from higher-resolution studies of the upper mantle are in broad agreement with those from the longer-wavelength studies: young oceanic regions and other tectonically active regions have low velocities down to 250 km, whereas the continental shields have high velocities. Some, but not all, hotspots seem to be underlain by slow regions in the deep mantle. However, there are no plume-shaped anomalies extending directly from the CMB to the upper mantle beneath the hotspots: the resolution of data is not sufficient to detect a plume less than 100 km in diameter. It is clear that some high-velocity zones cross the boundary between upper and lower mantle (Fig. 9.60). Figure 8.6(d) (Plate 12) shows a comparison between two models at 1325 km depth, namely a high-resolution P-wave model and a longer-wavelength S-wave model. The P-wave velocity model has the mantle divided into some 300 000 constant velocity blocks (dimension $2^\circ \times 2^\circ \times 100$–200 km in the lower mantle and as little as $0.6^\circ \times 0.6^\circ \times 35$ km in the upper mantle) and used P, pP and pwP data from 82 000 earthquakes. Both images show a fast

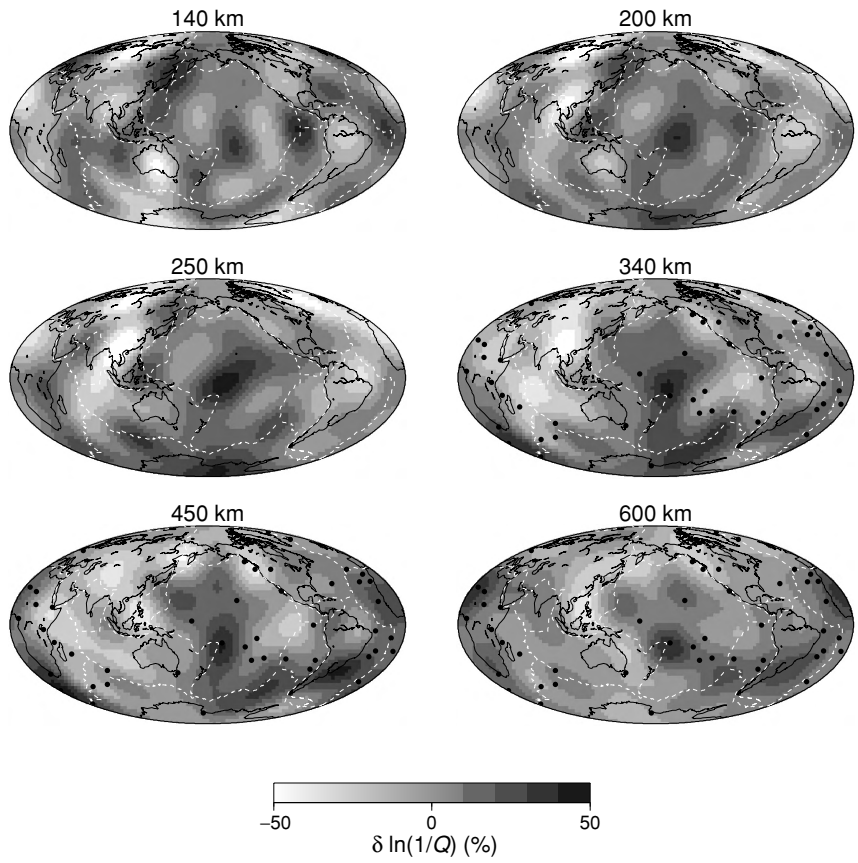


Figure 8.8. Perturbations in the quality factor, Q, of the upper mantle. Colour version Plate 13. (From Gung and Romanowicz (2004).)

anomaly running north–south beneath North America and another extending from the Mediterranean to Southeast Asia. These anomalies are interpreted as the subducted Farallon and Tethys slabs (Chapter 3). Figure 8.6(e) (Plate 12), a cross section from the Aegean through Asia to Japan, suggests that clues to the complex accretionary process that assembled the continent may be present in the underlying mantle.

Figure 8.9 shows the locations of high-density subducted slabs in the mantle for the present day and 56 Ma ago. The high-velocity anomalies around the Pacific correlate with positions of lithosphere that has been subducted into the mantle over the last 100 Ma or so. High velocities through Eurasia appear to mark the location of the Tethys subduction. This indicates that the upper mantle and the lower mantle cannot be two totally separate systems as some geochemical models imply. However, the upper and lower mantle are distinct: the fast region beneath South America does not extend below ∼1400 km. Had the Farallon plate descended through the lower mantle at the same rate as it did through the upper mantle, there would be a high-velocity zone extending all the way to the CMB.

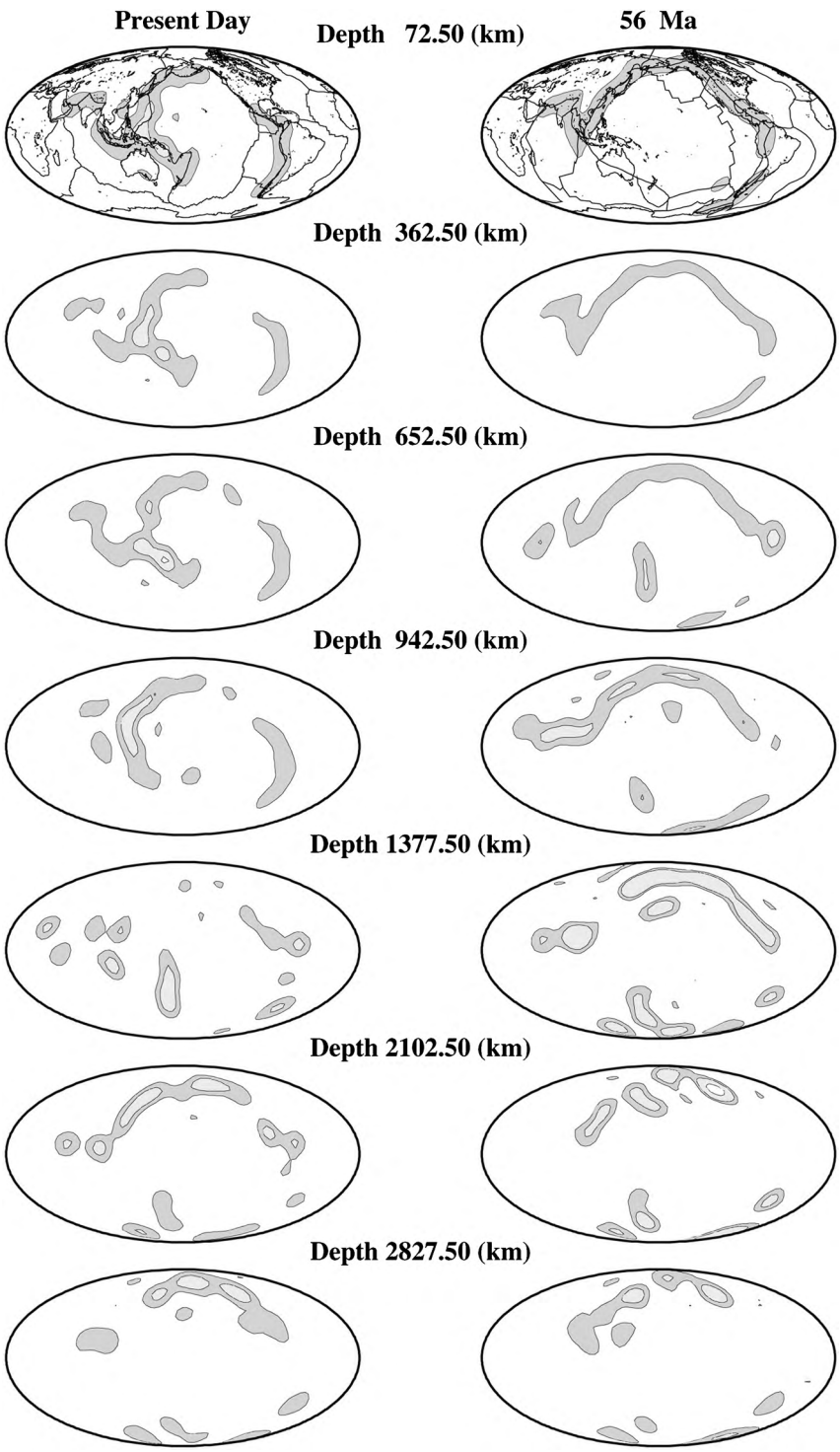


Figure 8.9. Subduction zones are identifiable as areas of high density within the mantle, shown here for the present day and 56 Ma ago. Locations of density anomalies were calculated by mapping subduction zones in the hotspot reference frame and allowing slabs to sink into the mantle, taking into account the major increase in viscosity at the 670-km boundary. (Lithgow-Bertollini and Richards, The dynamics of Cenozoic and Mesozoic plate motions, *Rev. Geophys.*, **36**, 27–78, 1998. Copyright 1998 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

That it does not could indicate that the rate of descent in the lower mantle was much reduced, that descent was in some way hindered or that the plate was deformed (mantle convection is discussed in Section 8.2). The large-scale high- and low-velocity anomalies at the base of the lower mantle visible in Fig. 8.6 are also a feature of these body-wave models. These high-velocity regions may mean that the D″ zone is the ultimate destination of subducted plates.

The fine-scale structure of the D″ zone is the subject of considerable current research. There is a localized stratification at the top of the D″ zone with increases of up to 3% in P- and S-wave velocities. These need not be straightforward increases in velocity but may rather be transition zones up to 50 km thick. Regional reductions in both P- and S-wave velocities of over 10% have been imaged in a thin layer (5–40 km vertical extent) immediately above the CMB. This low-velocity feature, referred to as the ultra-low-velocity zone (ULVZ), is thought to be very heterogeneous. The seismic phases used to image the ULVZ include (1) precursors to the short-period reflections PcP and ScP and (2) the longer-period SKS phase and the later associated phase SPdKS. SPdKS is a phase in which energy is diffracted as a P-wave (code Pd) along the CMB before continuing through the outer core as a P-wave. The ULVZ is not a global feature: it has so far been imaged only beneath the central Pacific, northwestern North America, Iceland and central Africa; it is absent beneath most of Eurasia, North America, South America and the south Atlantic. Major velocity reductions of 10% occurring in such thin localized zones imply major changes in physics and chemistry – partial melting seems possible.

The processes occurring in the D″ boundary layer between mantle and core are matters of much research and conjecture. Seismology gives glimpses of the structures, velocities and anisotropy present there. The details of the interaction of cold, downgoing, subducted plates with the lowermost mantle and the of generation of plumes are not yet well understood. The extent to which the ULVZ and any partial melting may be linked to plume location and the role of a chemical boundary layer above the CMB are also far from clear. Investigating links between hotspot volcanism and past properties of the CMB may seem far-fetched but may provide information on CMB chemistry and processes.

Anisotropy

The velocity of seismic waves through olivine (which is a major constituent of the mantle) is greater for waves travelling parallel to the *a* axis of the olivine crystal than it is for waves travelling perpendicular to the *a* axis. Such dependence of seismic velocity on direction is called velocity *anisotropy* (i.e., the material is not perfectly isotropic). Anisotropy is not the same as *inhomogeneity*, which refers to a localized change in physical parameters within a larger medium. Any flow in the mantle will tend to align the olivine crystals with their *a* axes parallel to the direction of flow. For this reason, measurement of anisotropy in the mantle can indicate whether any flow is vertical or horizontal. Plots of

the difference between vertically (SV) and horizontally (SH) polarized S-wave velocities determined by tomography show flow directions in the upper mantle: horizontal flow beneath the shields and vertical flow beneath mid-ocean ridges and subduction zones. It has been suggested that the longstanding debate on how far into the mantle the continental roots or keels extend (~200–250 km on geochemical, thermal and isostatic evidence, but as much as 400 km from some seismic-velocity models) may be, in part, reconciled when seismic anisotropy is taken into account.

The lower mantle is generally isotropic, but the D″ zone is locally anisotropic for S-waves. Beneath Alaska and the Caribbean the D″ zone is transversely anisotropic, with SH faster than SV. This anisotropy could be due to the presence of a stack of thin horizontal layers in the upper part of D″ or could arise from hexagonal crystals with their symmetry axes aligned vertically. Beneath the central Pacific the anisotropy is very variable but seems to be confined to the lowermost levels of D″. Anisotropy in the D″ zone could be caused by structural laminations (perhaps oriented inclusions of partial melt or subducted oceanic crust) or could result from a change in deformation in the boundary layer relative to the lower mantle. Data are sparse, but there may be some correlation between the form of anisotropy and the presence or absence of the ULVZ.

The seismic velocity of the inner core is anisotropic with an amplitude of 2–4%. It has a cylindrical symmetry about an axis that is approximately aligned (tilted at 8–11°) with the Earth's north–south spin axis. In early 1996 the inner-core symmetry axis was at 79°N, 169°E. The inner-core anisotropy has been determined from measurements of the travel times of body waves: paths parallel to the spin axis are fastest. Additionally, normal modes (Section 4.1.4), which have significant energy in the inner core, undergo some splitting, indicating that the inner core is anisotropic. The anisotropy is, however, not completely uniform – while the Western hemisphere is strongly anisotropic, the outer half of part of the Eastern hemisphere is only weakly anisotropic. The anisotropy is thought to be caused by preferential alignment of the hexagonal close-packed (h.c.p) phase of iron (Section 8.1.5) in the inner core. The reason for the development of the anisotropy is not understood, but it is possible that convective flow in the inner core could preferentially align iron, just as flow in the mantle leads to alignment of olivine. Another possibility is that shear forces due to the axially aligned corkscrew-like magnetic-field lines (Fig. 8.25) may cause a preferential crystal alignment in the inner core.

Repeated measurements of the difference in travel time between P-waves that penetrate the inner core and those on close ray paths that only pass through the outer core have shown that, over three decades, the position of the inner core's fast axis has moved with respect to the crust and mantle. This movement is a rotation: the inner core is rotating faster than the rest of the Earth. Estimates of the rate of rotation are varied, but it is probable that the inner core is rotating relative to

the crust and mantle at several tenths of a degree per year. A complete revolution would therefore take many centuries. This differential rotation of the inner core may affect many aspects of the workings of the planet, including the magnetic field. Conservation of the total angular momentum of the Earth means that any slowing of the rotation of the mantle must be balanced by an increase in the rotation of the atmosphere, oceans and core and vice versa. The atmosphere changes on a short timescale, whereas the core will respond over decades. Changes in the observed length of a day are well explained by atmospheric variation and there is no reason to suppose that this rotation measured over the last thirty years is not a long-term feature of the inner core.

8.1.5 The composition of the Earth

The continental crust varies greatly in the variety of its igneous, sedimentary and metamorphic rocks. However, on average, the continental crust is silica-rich and, very loosely speaking, of granitoid composition (see Section 10.1.3). In contrast, the oceanic crust is basaltic and richer in mafic (Mg, Fe-rich) minerals (see Section 9.1).

Our knowledge of the composition of the deep interior of the Earth is largely constrained by seismic and density models; we have little direct evidence regarding the compositions of the mantle and core. Chemistry is important, but unfortunately we lack *in situ* measurements, since no one has yet drilled that deeply. Our chemical knowledge of the deep interior has to be inferred from the chemistry of the volcanic and intrusive rocks derived from liquids that originated in the mantle, from structurally emplaced fragments of mantle, from nodules brought up during volcanism and from geochemical models of the various seismic and density models discussed in the previous sections.

This lack of direct evidence might suggest that the mantle composition is pretty much unknown, but geochemistry is a sophisticated and powerful branch of Earth science and has developed many techniques in which we use the compositions of rocks available to us on the surface to model the composition at depth. In addition, experiments at high temperatures and pressures have enabled the behaviour of minerals thought to exist in the deep mantle to be studied in some detail. There are many compositional models of the upper mantle, some of which are more popular than others. These models include imaginary rocks such as pyrolite, which is a mixture of basalt and residual mantle material. The main constituent of the mantle is magnesian silicate, mostly in the form of olivine. The core is iron-rich, and the core–mantle boundary is a very major compositional boundary.

Olivine in the mantle lies between two end members: forsterite (Fo), which is Mg_2SiO_4, and fayalite (Fa), which is Fe_2SiO_4. Normal mantle olivine is very forsteritic, probably in the range Fo_{91}–Fo_{94}, where 91 and 94 represent percentages of Mg in $(Mg, Fe)_2SiO_4$. Other trace components of mantle olivine include nickel (Ni) and chromium (Cr). The other major mantle minerals include

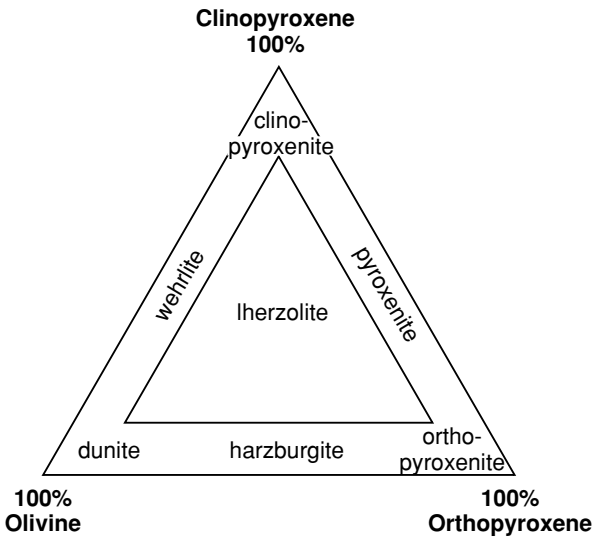


Figure 8.10. Common ultramafic rocks. The three outer corners of the triangle represent 100% compositions: dunite is ultramafic rock with close to 100% olivine and hardly any clinopyroxene or orthopyroxene; wehrlite is made up of approximately equal parts of olivine and clinopyroxene but no orthopyroxene. The interior triangle is the 90% contour. The classification of igneous rocks is discussed in Section 9.1.

orthopyroxene (Opx) and clinopyroxene (Cpx). Orthopyroxene varies between the end members enstatite ($MgSiO_3$) and ferrosilite ($FeSiO_3$), with about the same Mg to Fe ratio (94–91) in the mantle as olivine. Clinopyroxene also contains calcium, as $Ca(Mg, Fe)Si_2O_6$. Figure 8.10 illustrates the relationships amongst the common ultramafic rocks.

Experimental work on olivine has shown that it undergoes phase changes to denser structures at pressures equivalent to depths of 390–410, 520 and 670 km (see also Sections 8.2.3 and 9.6.3). This region of the mantle is called the *transition zone*. A phase change does not involve a change in chemical composition but rather a reorganization of the atoms into a different crystalline structure. With increasing pressure, these phase changes involve closer packing of the atoms into denser structures (Fig. 8.11(a)). At 390–410 km, olivine changes to a β spinel structure, passing through a mixed $\alpha + \beta$ phase region, with a resultant 10% increase in density (pyroxene also changes to a garnet structure at this depth). This change of olivine to spinel structure is accompanied by a release of heat (the reaction is exothermic; Clapeyron slope 2–3 MPa K^{-1}). At depth about 520 km the β spinel phase changes, again through a mixed-phase region, into the γ spinel phase. (This change has much less effect on seismic velocity than do those at 400 and 670 km.) At 670 km the γ spinel structure undergoes another change to minerals with a post-spinel structure: perovskite $(Mg, Fe)SiO_3$ and magnesiowüstite (Mg, Fe)O (Fig. 8.11(b)). These reactions, which also involve increases in density of about 10%, are endothermic (heat is absorbed during the reaction; Clapeyron slope $-(2–3)$ MPa K^{-1}). The silicon atoms in olivine and the spinel structure are both surrounded by four oxygen atoms, but this changes for the perovskite structure. Figure 8.11(a) shows these phase changes alongside the S-wave velocity profiles of the upper mantle. The depths at which the phase

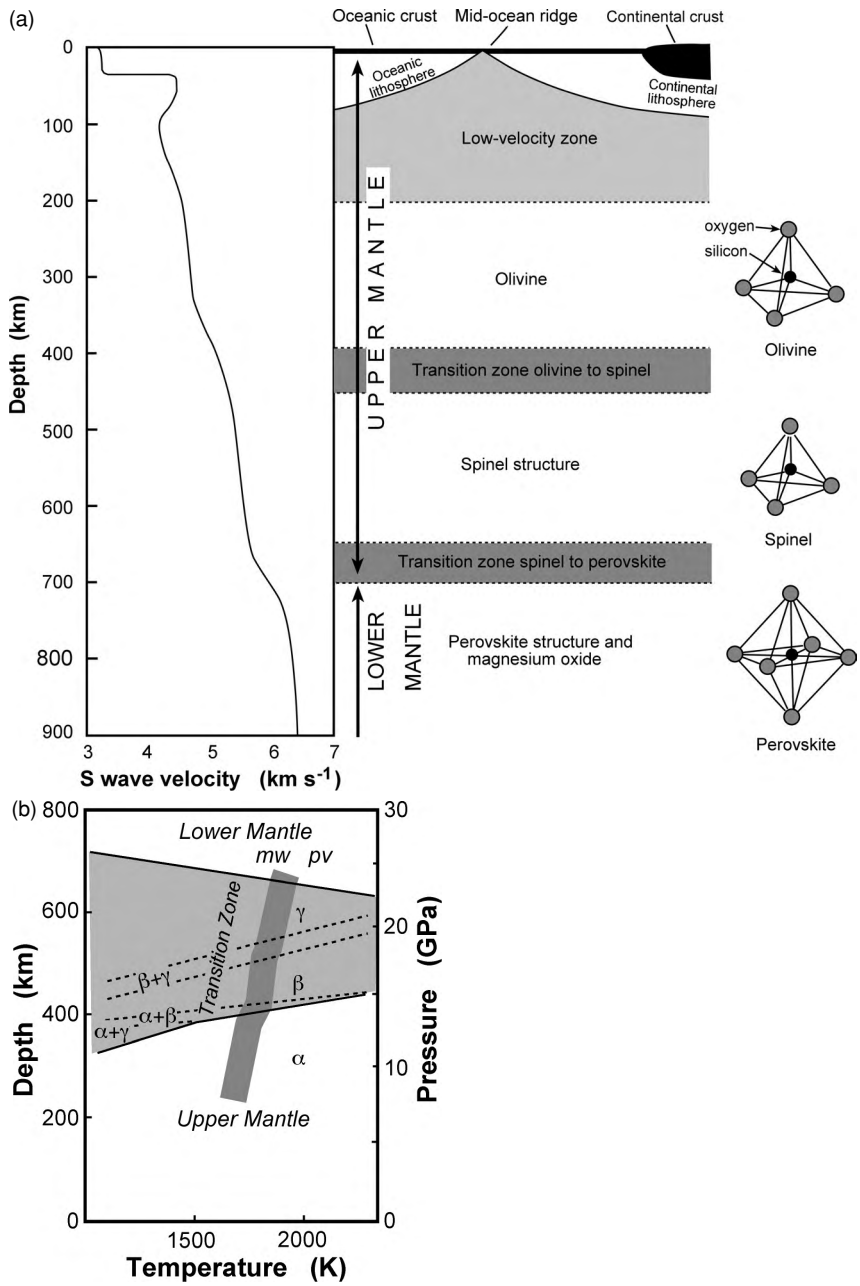


Figure 8.11. (a) The S-wave velocity profile of the upper mantle compared with the phases and transition zones for olivine in the upper mantle. (After McKenzie (1983). Copyright (©) 1983 by Scientific American Inc. All rights reserved.) (b) Phase transformations (shaded) for olivine in the mantle: olivine (α) through wadsleyite (β) and ringwoodite (γ) to perovskite (pv) and magnesiowüstite (mw). The major seismic discontinuities at 410 and 670 km as well as the discontinuity at 520 km correspond to these phase transformations. Dark grey band, upper mantle adiabat with a potential temperature of ~1500–1600 K.

changes take place coincide with those at which the seismic velocity increases more rapidly. The two main phases in the lower mantle, magnesiowüstite and Mg-perovskite, have been shown experimentally to have melting temperatures of over 5000 and 7000 K, respectively. This means that the lower mantle has always been solid (Section 7.7).

The situation for the core is far more difficult: no one has ever had a sample of the core to analyse. The closest approach to sampling the core is to consider the abundance of elements in the Sun and in meteorites. The Earth is believed to have formed from an accretion of meteoritic material. Meteorites are classified into two types: *stony* and *iron* (Sections 6.8 and 6.10). The stony meteorites are similar to the mantle in composition, whereas the iron meteorites may be similar to the core. If so, the core should be rich in iron with a small proportion of nickel ($\sim$5%). The major problem with theories of core composition is that they depend on theories of the origin of the Earth and its chemical and thermal evolution, which are also poorly understood. Solar abundances of iron are slightly higher than that in stony meteorites. If the solar model is taken, the lower mantle may have as much as 15% FeO. The core is very iron-rich and may in bulk be roughly Fe_2O in composition.

More direct evidence for the composition of the core can be inferred from its seismic velocity and density structure. Pressures appropriate for the core can be attained in experiments using shock waves or diamond anvils. Thus, laboratory measurements can be made on test samples at core pressures. When corrections for temperature are made, such laboratory velocity measurements can be compared with seismic models. Figure 8.12(a) shows the square root of the seismic parameter ϕ (Eq. (8.15)) plotted against density for a number of metals. The ranges of values appropriate for the mantle and core, indicated by the seismic-velocity and density models, indicate that, although magnesium and aluminium are possible candidates for a major proportion of the mantle, such low-atomic-number metals are quite inappropriate for the core. All the evidence on the properties of iron at high temperatures and pressures points unequivocably to a core that is predominantly composed of iron.

The outer core is probably an iron alloy: iron with a small percentage, 10% by weight, of lighter elements. Amongst the favoured candidates for the minor alloying element(s) are nickel, oxygen, sulphur, hydrogen, silicon and carbon. Figure 8.12(b) is a plot of density against pressure (obtained from shock-wave experiments) for pure iron (molten and solid) and possible iron compounds compared with the *in situ* values for the core. This plot shows that the outer core cannot be composed of either pure iron or the nickel–iron compound found in meteorites: both of these materials are too dense. Each possible lighter alloying element has its advantages and disadvantages, with sulphur and oxygen the strongest candidates. Cosmochemical evidence suggests that a maximum of 7% sulphur may be present in the core, but, since this amount of sulphur is insufficient to account for the density of the outer core, there must be additional light element(s).

Iron is the presumed constituent of the inner core. The data for the inner core indicate that it may well be virtually pure iron. There are several possible crystalline forms for iron in the inner core: the body-centred cubic (b.c.c.) phase and the face-centred cubic (f.c.c.) phase are probably unstable under inner-core conditions, but the hexagonal close-packed (h.c.p.) phase should be stable. If

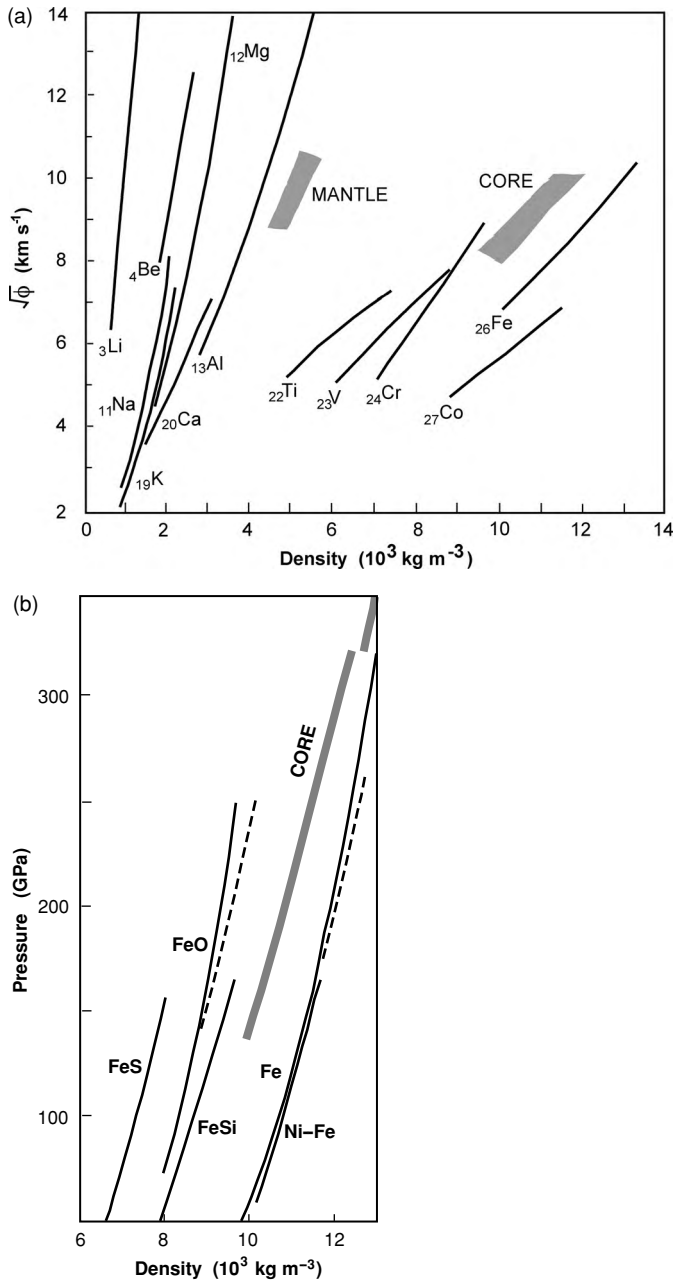


Figure 8.12. (a) The seismic parameter ϕ, $\sqrt{\phi} = \sqrt{K/\rho} = \sqrt{\alpha^2 - \frac{4}{3}\beta^2}$ from Eq. (8.20) plotted against density for metals. These values were obtained from shock-wave experiments. The shaded regions show the ranges of values for the mantle and core given by the seismic models. (After Birch (1968).) (b) Pressure and density as measured in shock-wave experiments for iron and the iron compounds which may be present in the core. The heavy line shows values for the core calculated from a seismic-velocity model. (Based on Jeanloz and Ahrens (1980) and Jeanloz (1983).)

iron in the inner core is in the f.c.c. phase then there seems no need for any light impurity. The h.c.p. phase has a higher density than the inner core so, if iron is in this phase, there must also be an impurity. Clearly whether or not any light impurity is present in the inner core cannot be established until the phase diagram for pure iron at core temperatures and pressures has been well determined (see also Section 8.3.1). Calculations show that, when sulphur or silicon impurities

are present, the b.c.c. phase is more stable than the h.c.p. phase under inner-core conditions, which is opposite to the results for pure iron under inner-core conditions. There is some chemical but not physical evidence in favour of nickel in the inner core: nickel has no effect on density; nickel alloys easily with iron; Fe–Ni phases are observed in meteorites; and nickel in the core would balance its depletion in the mantle compared with cosmic abundances.

At low temperatures, FeO is non-metallic and forms an immiscible liquid with Fe. In the past this led to doubts about the presence of oxygen in the core. High-pressure and -temperature experiments on iron oxide, however, have shown that it becomes metallic at pressures greater than 70 GPa and temperatures greater than 1000 K. This means that oxygen can alloy with iron in the core and suggests that oxygen is very probably a constituent of the core. Since oxygen raises the melting temperature of iron, the presence of oxygen elevates estimates of core temperatures. Experiments at temperatures and pressures appropriate for the core–mantle boundary have shown that liquid iron and iron alloys react vigorously with solid oxides and solid silicates. Thus an iron-rich core would react chemically with the silicate mantle. This may well be the explanation for the seismic complexity of the core–mantle boundary (Section 8.1.4): the mantle and core are not in chemical equilibrium; rather this is the most chemically active part of the Earth. It is probable that the core contains oxygen and the outer core may contain as much oxygen as sulphur. It is, however, not yet experimentally or computationally possible to establish the concentrations of the various lighter alloying elements present in the core.

8.2 Convection in the mantle

8.2.1 Rayleigh–Bénard convection

Convection in liquids occurs when the density distribution deviates from equilibrium. When this occurs, buoyancy forces cause the liquid to flow until it returns to equilibrium. Within the Earth convection occurs in the mantle and the outer core. Density disturbances in the Earth could be due to chemical stratification or to temperature differences. Chemical stratification is the main cause of convection in the outer core, but in the mantle the convection is of thermal origin. The simplest illustration of thermal convection is probably a saucepan of water, or soup heating on the stove.

For a *Newtonian viscous fluid*, stress is proportional to strain rate, with the constant of proportionality being the dynamic viscosity of the fluid.[5]

[5] strain rate $\propto$ stress

$$= \text{dynamic viscosity} \times \text{stress}$$

For materials with a power-law relationship between strain rate and stress

$$\text{strain rate} \propto (\text{stress})^n$$

Thus, when the stress increases by a factor of ten, the stress increases by 10^n. A Newtonian fluid has $n = 1$. Higher values of n are used in modelling some Earth behaviour.

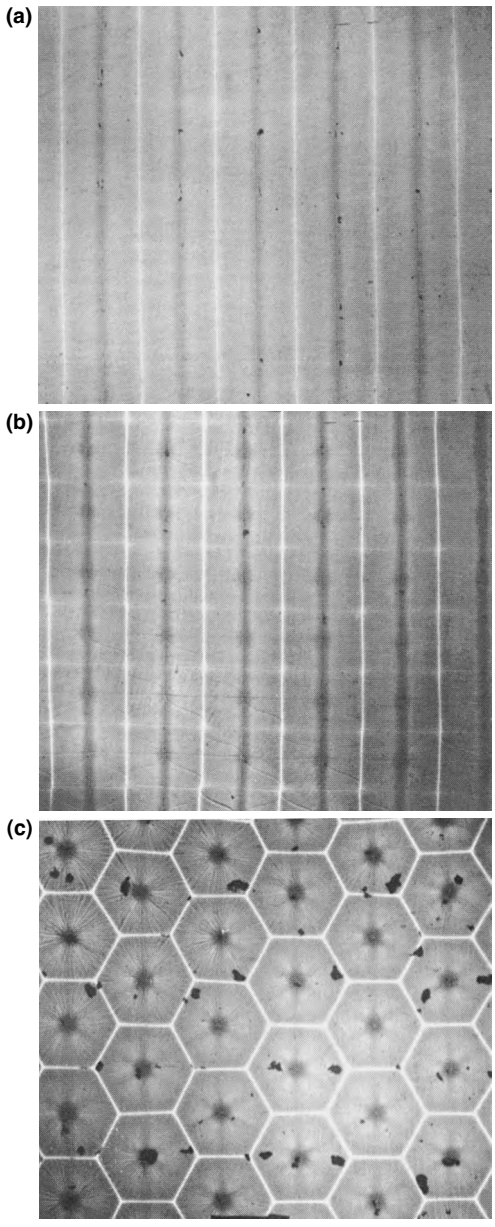


Figure 8.13. Photographs of the planforms of convection in a layer of viscous fluid. (From White (1988).) (a) Horizontal rolls of rising, hotter fluid (dark lines) and sinking, colder fluid (light lines) are stable over a wide range of Rayleigh numbers. (b) The bimodal pattern has a primary set of horizontal rolls with a weaker perpendicular set of rolls. The develops at a higher Rayleigh number than in (a). (c) A hexagonal pattern with a central rising plume and six sinking sheets.

Rayleigh–Bénard convection occurs when a tank of Newtonian viscous fluid is uniformly heated from below and cooled from above. Initially, heat is transported by conduction, and there is no lateral variation. As heat is added from below, the fluid on the bottom of the tank warms and becomes less dense, so a light lower fluid underlies a denser upper fluid. Eventually, the density inversions increase to a magnitude sufficient for a slight lateral variation to occur spontaneously and a convective flow starts. In plan view, the first convective cells are

two-dimensional cylinders that rotate about their horizontal axes. The hot material rises along one side of the cylinder, and the cold material sinks along the other side (Fig. 8.13(a)). As heating proceeds, these two-dimensional cylinders become unstable, and a second set of cylindrical cells develops perpendicular to the first set (Fig. 8.13(b)). This rectangular planform is called *bimodal*. As the heating continues, this bimodal pattern changes into a hexagonal and then a spoke pattern. Figure 8.13(c) shows hexagonal convection cells in plan view, with hot material rising in the centres and cold material descending around the edges. With heating, the fluid convects more and more vigorously, with the upgoing and downgoing limbs of a cell confined increasingly to the centre and edges of the cell, respectively. Finally, with extreme heating, the regular cell pattern breaks up, and hot material rises at random; the flow is then irregular.

8.2.2 Equations governing thermal convection

The derivation and discussion of the full differential equations governing the flow of a heated viscous fluid are beyond the scope of this book, but it is of value to look at the differential equations governing the simplified case of two-dimensional thermal convection in an incompressible Newtonian viscous fluid. The *Boussinesq approximation* to the most general convection equations is often used to simplify numerical calculations. In that approximation the fluid is incompressible (Eq. (8.26)) and the only result of a change in density considered is buoyancy (Eq. (8.30)).

The general equation of conservation of fluid (i.e., there are no sources or sinks of fluid, its volume is constant) is

$$\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z} = 0 \tag{8.26}$$

where $\mathbf{u} = (u_x, u_z)$ is the velocity at which the fluid is flowing.

The two-dimensional heat equation in a moving medium (Eq. (7.19) with no internal heat generation, $A = 0$) is

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial z^2}\right) - u_x\frac{\partial T}{\partial x} - u_z\frac{\partial T}{\partial z} \tag{8.27}$$

The horizontal equation of motion is

$$\frac{\partial P}{\partial x} = \eta\left(\frac{\partial^2 u_x}{\partial x^2} + \frac{\partial^2 u_x}{\partial z^2}\right) \tag{8.28}$$

and the corresponding vertical equation of motion is

$$\frac{\partial P}{\partial z} = \eta\left(\frac{\partial^2 u_z}{\partial x^2} + \frac{\partial^2 u_z}{\partial x^2}\right) - g\rho' \tag{8.29}$$

where P is the pressure generated by the fluid flow, η is the dynamic viscosity, g is the acceleration due to gravity and ρ' is the density disturbance. The convective flow of the fluid is maintained by the buoyancy forces resulting from differences

in density between different parts of the fluid. When the density disturbance is of thermal origin,

$$\rho' = \rho - \rho_0 = -\rho_0 \alpha (T - T_0) \tag{8.30}$$

where ρ_0 is the density at a reference temperature T_0, and α is the volumetric coefficient of thermal expansion.

In order to use Eqs. (8.26)–(8.30) to evaluate the form of convective flow, it is usual to present the equations in a parametric form, which means that the values of density, viscosity, length, time etc. are all scaled to a dimensionless form (e.g., see Hewitt *et al*. 1980). For doing this, several dimensionless numbers that completely describe the flow are routinely used in fluid dynamics. The dimensionless *Rayleigh number Ra* is given by

$$Ra = \frac{\alpha g d^3 \Delta T}{\kappa \upsilon} \tag{8.31}$$

where α is the volume coefficient of thermal expansion, g the acceleration due to gravity, d the thickness of the layer, ΔT the temperature difference in excess of the adiabatic gradient across the layer, κ the thermal diffusivity and υ the kinematic viscosity (kinematic viscosity = dynamic viscosity/density, i.e., $\upsilon = \eta/\rho$). The Rayleigh number measures the ratio of the heat carried by the convecting fluid to that carried by conduction. Flow at a particular Rayleigh number always has the same form regardless of the size of the system. Thus it is straightforward for laboratory experiments to use thin layers of oils or syrups over short times and then to apply the results directly to flow with the same Rayleigh number in the Earth (with viscosity, length and time scaled up appropriately).

To evaluate thermal convection occurring in a layer of thickness d, heated from below, the four differential equations (8.26)–(8.30) have to be solved with appropriate boundary conditions. Usually, these boundary conditions are a combination of the following:

(i) $z = 0$ or $z = d$ is at a constant specified temperature (i.e., they are isotherms), or the heat flux is specified across $z = 0$ or $z =$ d;
(ii) no flow of fluid occurs across $z = 0$ and $z = d$; and
(iii) $z = 0$ or $z = d$ is a solid surface, in which case there is no horizontal flow (no slip) along these boundaries, or $z = 0$ or $z = d$ is a free surface, in which case the shear stress is zero at these boundaries.

Solution of the equations with appropriate boundary conditions indicates that convection does not occur until the dimensionless Rayleigh number, Ra, exceeds some critical value Ra_c. For this layer the Rayleigh number can be written

$$Ra = \frac{\alpha g d^4 (Q + Ad)}{k \kappa \upsilon} \tag{8.32}$$

where Q is the heat flow through the lower boundary, A the internal heat generation and k the thermal conductivity. The critical value of the Rayleigh number further depends on the boundary conditions.

1. For no shear stress on the upper and lower boundaries, the upper boundary held at a constant temperature and all heating from below ($A = 0$), $Ra_c = 27\pi^4/4 = 658$. At this Rayleigh number the horizontal dimension of a cell is $2.8d$.
2. For no slip on the boundaries, the upper boundary held at a constant temperature and all heating from below ($A = 0$), $Ra_c = 1708$. At this Rayleigh number the horizontal dimension of a cell is $2.0d$.
3. For no slip on the boundaries, a constant heat flux across the upper boundary and all heating from within the fluid ($Q = 0$), $Ra_c = 2772$. At this Rayleigh number the horizontal dimension of a cell is $2.4d$.
4. For no shear stress on the boundaries, a constant heat flux across the upper boundary and all heating from within the fluid ($Q = 0$), $Ra_c = 868$. At this Rayleigh number the horizontal dimension of a cell is $3.5d$.

Thus, although the exact value of the critical Rayleigh number Ra_c depends on the shape of the fluid system, the boundary conditions and the details of heating, it is clear in all cases that Ra_c is of the order of 10^3 and that the horizontal cell dimension at this critical Rayleigh number is two-to-three times the thickness of the convecting layer. For convection to be vigorous with little heat transported by conduction, the Rayleigh number must be about 10^5. If the Rayleigh number exceeds 10^6, then convection is likely to become more irregular.

A second dimensionless number[6] used in fluid dynamics, the *Reynolds number* (*Re*), measures the ratio of the inertial to viscous forces,

$$\begin{aligned} Re &= \frac{\rho u d}{\eta} \\ &= \frac{u d}{\upsilon} \end{aligned} \tag{8.33}$$

where u is the velocity of the flow, d the depth of the fluid layer and υ the kinematic viscosity. *Re* indicates whether fluid flow is laminar or turbulent. A flow with $Re \ll 1$ is laminar, since viscous forces dominate. A flow with $Re \gg 1$ is turbulent. *Re* for the mantle is about 10^{-19}–10^{-21}, so the flow is certainly laminar.

A third dimensionless number, the *Nusselt number* (*Nu*), provides a measure of the importance of convection to the heat transport:

$$Nu = \frac{\text{heat transported by the convective flow}}{\text{heat that would be transported by conduction alone}}$$

$$Nu = \frac{Qd}{k\,\Delta T} \tag{8.34}$$

where Q is the heat flow, d the thickness of the layer, k the thermal conductivity and ΔT the difference in temperature between the top and bottom of the layer. The Nusselt number is approximately proportional to the third root of the Rayleigh number:

$$Nu \approx (Ra/Ra_c)^{1/3} \tag{8.35}$$

[6] Fluid dynamics utilizes several dimensionless numbers, all named after prominent physicists.

Table 8.2 *Possible Rayleigh numbers for the mantle*

	Thickness (km)	Rayleigh number, *Ra*
Upper mantle	670	10^6
Whole mantle	2900	6×10^7

Thus, as convection becomes the dominant mechanism of heat transport, the Nusselt number increases (when $Ra \sim Ra_c$, $Nu \sim 1$; for the situation in the mantle with $Ra \sim 10^3 Ra_c$, $Nu \sim 10$).

In regions where upwelling occurs, such as beneath a mid-ocean-ridge axis, heat is carried upwards, or advected, by the rising material. The thermal *Péclet number* (Pe_t), the ratio of convected to conducted heat, is another measure of the relative importance of convective to conductive heat transport,

$$Pe_t = \frac{ul}{\kappa} \tag{8.36}$$

where u is the velocity at which the material is moving, l a length scale and κ the thermal diffusivity. If Pe_t is much larger than unity, advection dominates; if Pe_t is much smaller than unity, conduction dominates. In the mantle Pe_t is about 10^3, showing that the heat is transported mainly by advection.

The dimensionless *Prandtl number*,

$$\begin{aligned} Pr &= \upsilon/\kappa \\ &= Pe_t/Re \end{aligned} \tag{8.37}$$

indicates the relative importance of viscous forces in diffusing momentum of the fluid compared with heat. The Prandtl number is a physical property of the material and is independent of any flow. For the mantle with $\upsilon \sim 10^{18}\ \mathrm{m^2\ s^{-1}}$ and $\kappa \sim 10^{-6}\ \mathrm{m^2\ s^{-1}}$, $Pr \sim 10^{24}$, demonstrating that the viscous response to any perturbation is instantaneous compared with the thermal response.

Table 8.2 gives approximate values of *Ra* calculated for the mantle using Eq. (8.32) with $A = 0$ and assuming the following values: for the upper mantle $\alpha = 2 \times 10^{-5}\ ^\circ\mathrm{C}^{-1}$, $\kappa = 10^{-6}\ \mathrm{m^2\ s^{-1}}$, $Q/k = 1.5 \times 10^{-2}\ ^\circ\mathrm{C\ m^{-1}}$ and $\upsilon = 3 \times 10^{17}\ \mathrm{m^2\ s^{-1}}$; for the whole mantle $\alpha = 1.4 \times 10^{-5}\ ^\circ\mathrm{C}^{-1}$, $\kappa = 2.5 \times 10^{-6}\ \mathrm{m^2\ s^{-1}}$, $\upsilon = 4 \times 10^{17}\ \mathrm{m^2\ s^{-1}}$ and $\mathrm{Q/k} = 7 \times 10^{-3}\ ^\circ\mathrm{C\ m^{-1}}$. Similar values for the Rayleigh number are obtained with $Q = 0$ and the internal heat generation of the mantle about $10^{-11}\ \mathrm{W\ kg^{-1}}$.

It is clear from Table 8.2 that the Rayleigh number is much greater than the critical Rayleigh number ($\sim 10^3$) irrespective of whether flow is considered to occur throughout the whole mantle or to be separate in the upper and lower mantle. Although the exact value of the Rayleigh number depends on the values chosen for the properties of the mantle, it is clear that convection in the mantle is vigorous. The most unrealistic assumption in these calculations is that of a constant-viscosity mantle. It is possible that the viscosity of the mantle is highly

temperature-dependent and may change by as much as an order of magnitude for each change by 100 °C in temperature.

8.2.3 Models of convection in the mantle

Patterns of mantle convection can be investigated in two ways. Numerical models can be simulated on a computer, or physical laboratory models can be made by choosing material of an appropriate viscosity to yield observable flow at appropriate Rayleigh numbers on a measurable timescale. The dimensionless numbers described in Section 8.2.2 are particularly important because they control the fluid behaviour. Therefore, by careful choice of appropriate fluids, it is possible to conduct laboratory experiments at Rayleigh numbers that are appropriate for the mantle – Tate and Lyle's golden syrup, glycerine and silicone oil are frequent choices. The dynamic viscosity of water is 10^{-3} Pa s and that of thick syrup is perhaps 10 Pa s; compare these values with the values of 10^{21} Pa s for the mantle (Section 5.7.2).

Simple two-dimensional numerical models of flow in rectangular boxes at high Rayleigh numbers (10^4–10^6) appropriate for the upper mantle cannot be compared directly with the Earth. The problem with these numerical models is that the exact form of instabilities and secondary flow depends upon the particular boundary conditions used. Figure 8.14(a) shows an example of the temperature and flow lines for a numerical model with heat supplied from below and the temperature fixed on the upper boundary. There is a cold thermal boundary layer at the surface that could represent the lithospheric plates. This cold material, which sinks and descends almost to the base of the box, could represent the descending plate at a convergent plate boundary. A hot thermal boundary layer at the base of the box rises as hot material at the 'ridges'. Therefore, if the flows in the upper and lower mantle are indeed separate, then simple models such as this imply that the horizontal scale of the cells in the upper mantle should be of the order of their depth (the aspect ratio of the cells is about unity). Cells with a large aspect ratio were unstable with these boundary conditions. Thus, this particular model implies that it is not possible for convection in the upper mantle to be directly related to the motions of the plates, with the downgoing cold flow representing the descending plates along the convergent boundary and the upwelling hot flow representing the mid-ocean-ridge system, because such a flow would have an aspect ratio much greater than unity (the horizontal scale of these motions is ~10 000 km). However, changing the boundary conditions results in a dramatic change in the flow. Figure 8.14(b) shows the flow that results when there is a constant heat flow across the upper boundary instead of a constant temperature on the upper boundary. In this case with constant heat flow across both the upper and the lower boundary, large-aspect-ratio convection cells are stable. The small-scale instabilities that develop on both boundaries of this model do not break up the large-scale flow. Figure 8.14(c) shows the results of the same experiment but

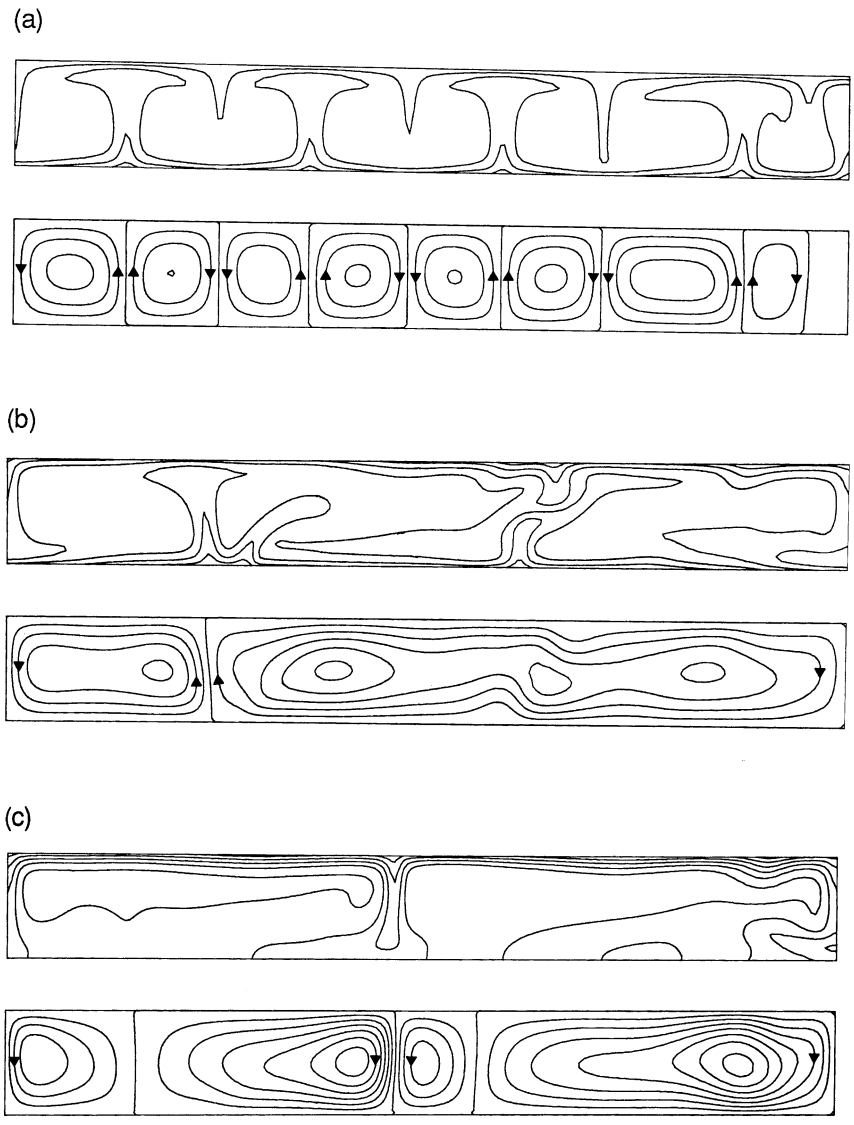


Figure 8.14. Temperature (upper) and fluid flow lines (lower) for computer models of convection in the upper mantle. The Rayleigh number is 2.4×10^5. There is no vertical exaggeration. (a) Temperature constant on the upper boundary, heat flow constant on the lower boundary. (b) Heat flow constant across upper and lower boundaries. (c) All heat supplied from within, heat flow constant on the upper boundary. Notice that the temperature varies rapidly in the boundary layers but is fairly constant in the interior of the cells. The flow of fluid is too fast for conduction of heat to be important: changes in temperature are primarily caused by changes in pressure. This is characteristic of high-Rayleigh-number flow. (After Hewitt *et al.* (1980).)

this time with all the heat supplied from within. Again, large-aspect-ratio cells are stable. The main difference between Fig. 8.14(b) and (c) is that, when all the heat is supplied from within, no sheets of hot material rise from the lower boundary. Thus, depending upon the particular boundary conditions chosen to model the upper mantle, one can draw disparate conclusions concerning the form of upper-mantle convection.

It has been proposed that there could be a two-scale convective flow in the upper mantle. The large-scale flow would represent the plate motions, with the upper boundary layer as the strong, cold mechanical plate. The small-scale flow

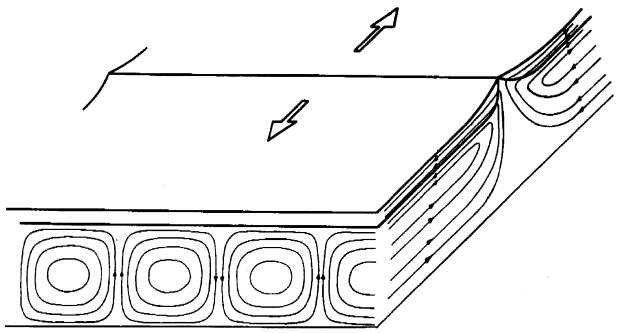


Figure 8.15. Laboratory experiments with a moving rigid upper boundary indicate that flow in the upper mantle could take this type of form. (After Richter and Parsons (1975).)

aligned in the direction of shear would exist beneath the plates; its upper boundary layer, not rigidly attached to the plate, would be the thermal boundary layer. Three-dimensional rectangular laboratory experiments with silicone oil and a moving, rigid upper boundary have indicated that such a two-scale flow can occur. Figure 8.15 shows such a convection system. Another laboratory experiment, which modelled the thermal effect of the subducted lithosphere by cooling one of the side walls, gave rise to a single, stable, large-aspect-ratio convection cell. Again these experiments illustrate that large-aspect-ratio cells can be stable; however, the exact form of instabilities and secondary flow depends on the particular physical characteristics of the experimental model, its geometry and boundary conditions.

Isotopic ratios of oceanic basalts are very uniform and are quite different from those of the bulk Earth, which means that the mantle must be very well mixed. This is confirmed by numerical models. Figure 8.16 shows a computer model of mantle convection in a two-dimensional rectangular box. A square patch of mantle with physical properties identical to those of the rest of the model is marked, and its deformation and distribution throughout the mantle are traced at subsequent times. Within several hundred million years, the convective process is able to mix upper-mantle material thoroughly. This time is short compared with the half-lives of the measured radioactive isotopes, indicating that upper-mantle convection should be well able to account for the general uniformity of isotopic ratios in oceanic basalts. For the upper-mantle model illustrated in Fig. 8.16, any body smaller than 1000 km is reduced to less than 1 cm thick within 825 Ma.

The isotopic ratios of oceanic-island basalts (OIB) require a source for these magmas that is less depleted than the source of mid-ocean-ridge basalts (MORB). Efficient mixing of either the upper mantle or the whole mantle by convection suggests that the source of OIB must be a recent addition to the mantle or an unmixed reservoir. If this were not the case, the source would be mixed into the mantle too well to allow the characteristic isotopic signatures of OIB to have developed. The source of OIB is thus a matter of considerable conjecture. It is possible that they originate from the base of the lower mantle and that this is a

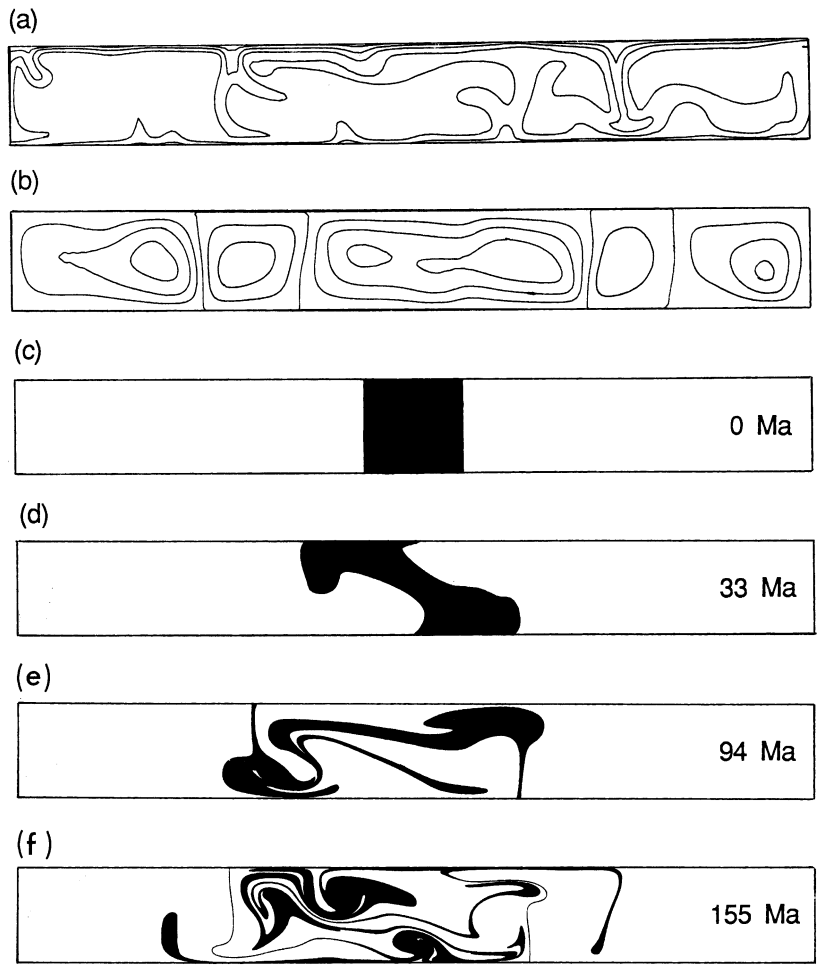


Figure 8.16. A computer model of convection in the upper mantle. Half the heat is supplied from below, and half is supplied internally. The Rayleigh number is 1.4×10^6. The model has several adjacent cells, each with separate circulation, although over time the cell boundaries move and material is exchanged between adjacent cells. (a) Isotherms (temperature contours), (b) fluid flow lines and (c) locations of marked fluid; (a), (b) and (c) are all at the starting time. Deformation of the marked fluid at subsequent times: (d) 33 Ma, (e) 97 Ma and (f) 155 Ma. (From Hoffman and McKenzie (1985).)

'slab graveyard'. Another idea, which is in agreement with the convection and geochemical models, is that the sub-continental lithosphere provides a source for OIB. Isotopic anomalies can easily form in the deep lithosphere beneath the continents. Deep continental material could become denser and delaminate or fall into the upper mantle, this process perhaps being triggered by a continent–continent collision. Such a cold body would descend at least to the base of the upper mantle, where it would warm before rising to the surface as part of the convection system. It would remain a viable magma source for about 100–300 Ma. After 150 Ma a body that was originally 100 km thick would be mixed into 5-km-thick sheets. Another proposal for the origin of OIB is that they are the result of partial melting of material that has risen from a separately convecting primitive lower mantle (Fig. 8.17). Even though this model cannot explain why these basalts do not have the same isotopic composition as the bulk Earth, it is

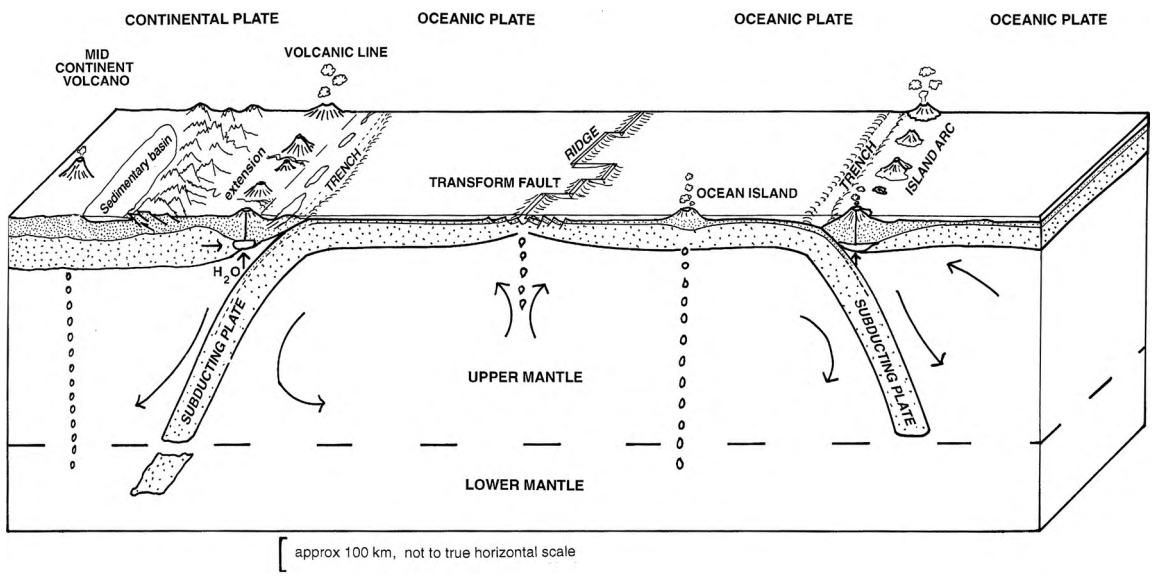


Figure 8.17. A schematic diagram illustrating the formation of new oceanic lithosphere along the mid-ocean ridges and its eventual subduction back into the mantle. Lithosphere is stippled. Crust is indicated by dense stippling. Oceanic-island basalts may be derived from the lower mantle.

attractive in its simplicity. Detection of a rising plume by seismology is difficult: their probable diameter of ~100 km or so is less than the resolution currently attainable (Fig. 8.6). Nevertheless there is some indication of low velocities (probable temperature difference ~300 K) at a depth of 700 km in a zone with diameter 150 km, close to the expected position of the Bowie hotspot in the northeast Pacific. With further careful work it should be possible to select earthquakes and seismic stations to provide information on the structure of the mantle beneath some other hotspots and so to answer some of the questions about their origin.

The possibility that the convective flows in the upper and lower mantle could be separate systems has been proposed on the basis of a number of observations. There is a jump in the seismic P-wave velocity and density at 660–670 km (Section 8.1.1), which is due to an endothermic phase change of mantle olivine from spinel to post-spinel forms (Sections 8.1.5 and 9.6.3). Along the convergent plate boundaries 670 km is observed to be the maximum depth at which earthquakes occur, and it seems that the descending slab may sometimes break off or be deflected at this level. These results are in agreement with geochemical models implying that the upper mantle is depleted in incompatible elements and has been almost separate from the lower mantle throughout the Earth's history. The amount of ^{3}He emitted at the mid-ocean ridges is apparently much less than that produced by radioactive decay in the mantle: a sink in a

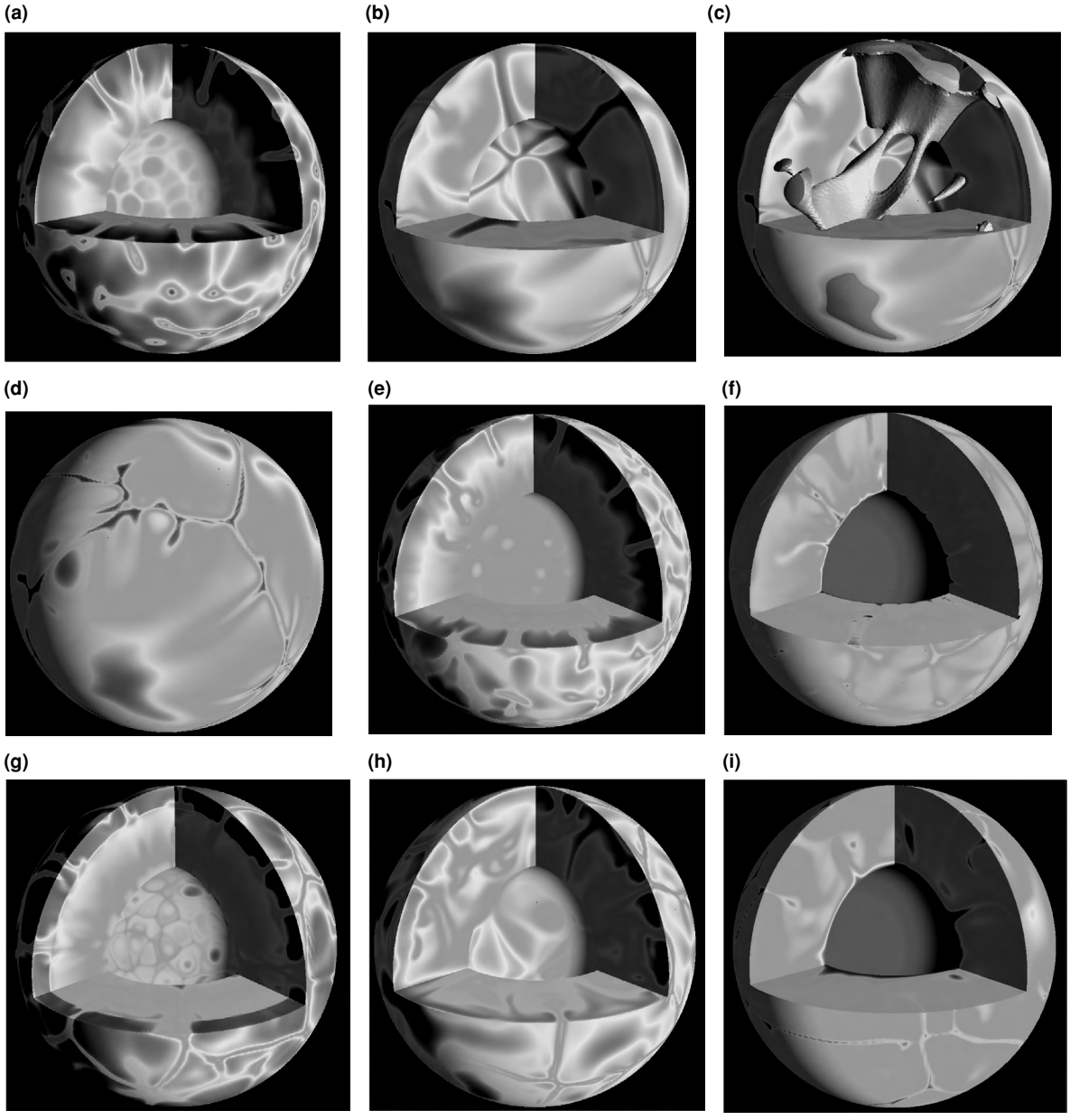


Figure 8.18. Spherical three-dimensional convection models of the mantle (the uppermost 200-km boundary layer is not shown). Superadiabatic temperatures: red, hot; blue, cold. (a) Incompressible mantle, constant viscosity, internal heating only, $Ra = 4 \times 10^7$. (b) As (a) but the viscosity of the lower mantle is thirty times the viscosity of the upper mantle. (c) As (b) showing the isosurface. (d) As (b) showing the planform. (e) Compressible mantle, constant viscosity, $Ra = 10^8$. (f) As (e) but with 38% of heating from the core. (g) As (e) but with an endothermic phase change of $-4\,\mathrm{MPa\,K^{-1}}$ at 670 km depth. (h) As (e) but with the viscosity of the lower mantle thirty times the viscosity of the upper mantle. (i) As (h) and with 38% of heating from the core. Colour version Plate 14. (Bunge, personal communication 2003 after Bunge *et al.* (1996, 1997).)

separate lower mantle would be a possible repository. Likewise a lower-mantle repository could account for the observation that there is much less ^{40}Ar in the atmosphere and continental crust than should have been produced by the decay of primordial ^{40}K. However, the tomographic images reveal that, although the descent of some subducting plates is impeded at 670 km, in general the mantle seems to be one system. The location and extent of any geochemical mantle reservoir has not been established – a self-consistent model for the Earth that reconciles geochemistry, plate tectonics and mantle convection remains a goal and a subject for much research interest.

Advances in computer technology have benefited those making numerical models of mantle convection. Much more realistic models than the simple rectangular two-dimensional models of Figs. 8.14 and 8.16 are now achievable. Figure 8.18 and Plate 14 show a series of three-dimensional spherical convection models of the whole mantle. Figure 8.18(a) is the simplest model, with a constant-viscosity incompressible mantle with all heating being internal. The wavelength of the convection cells is short compared with the size of the Earth and there are numerous downwellings. Figures 8.18(b)–(d) show the change in convection pattern that takes place when the viscosity of the lower mantle is increased to a more realistic value – thirty times that of the upper mantle. With this change the wavelength of the convection cells increases and the flow itself is dominated by sheets that extend right through the mantle. Figure 8.18(e) is another constant-viscosity model, but for a compressible mantle, rather than an incompressible mantle (Fig. 8.18(a)). Again this has numerous downwellings and the flow has a short wavelength. If heating from the core is included in the model (Fig. 8.18(f)), there is a thermal boundary layer at the base of the mantle and the convection pattern is dominated by hot upwellings. Figure 8.18(g) shows the major effects caused by inclusion of an endothermic phase change at 670 km. Both downgoing and upwelling material is inhibited by the phase change, but the wavelength of convection cells is not substantially affected and the overall timescale of the flow is not significantly affected. Figure 8.18(h) shows that, just as for an incompressible mantle (Fig. 8.18(b)), the wavelength increases when the viscosity of the lower mantle is increased to thirty times that of the upper mantle. Figure 8.18(i) shows the effect of including heating from the core on a model with a compressible layered mantle. Now that three-dimensional models such as those shown here can be made, it is possible to investigate separately and together many of the physical parameters which may contribute to the way the Earth's mantle convects. Important factors are the viscosity of the mantle, phase changes in the mantle, sources of heat and the inclusion of the plates as the outer boundary layer.

A three-dimensional numerical model of convection with an exothermic phase change at 400 km and an endothermic phase change at 670 km resulted in a layered convection pattern. The upper and lower shells are effectively separate: downwelling cold sheets in the upper mantle do not penetrate the 670-km horizon

but collect above it. These sheets are typically several thousand kilometres apart, a scale similar to, but less than, the spacing of subduction zones in the Earth. This cold material is gravitationally unstable; when enough has collected at the base of the upper mantle a catastrophic avalanche into the lower mantle ensues. Sudden avalanches of cold material into the lower mantle may take place at several locations at one time and descend as cylinders directly to the CMB. This pattern of downwelling could explain the images of the mantle determined from seismic tomography, which have extensive high-velocity regions at the CMB and a lower mantle that is characterized by long-wavelength anomalies (Figs. 8.6 and 8.8). The hot wide upwelling regions that developed in the upper mantle are not associated with features in the lower mantle. Occasionally, however, narrow hot plumes of material rise from the CMB and can penetrate the 670-km discontinuity to pass into the upper mantle. These are, though, neither stable nor weak enough to be analogous to plume hotspots in the real mantle. With ever-improved computer modelling using realistic Rayleigh numbers (achieved by using lower and better estimates of viscosity) and inclusion of the plates, it is expected that the phase change may prove to be a major factor in the partial separation of the flow regimes in the upper and lower mantle. Better numerical parameterization of the phase changes should also enhance their effects; layering is enhanced by narrower phase transitions and is sensitive to the magnitude and sign of the Clapeyron slope – the exothermic change at depth 400 km acts against stratification while the endothermic change at 670 km causes stratification. The metastability of olivine in the subducting slab (Section 9.6.3) could have a considerable impact, reducing the negative buoyancy of the slab as well as decreasing the heat released at 400 km.

Figure 8.19 illustrates the dramatic effect that the Rayleigh number has on the stratification of convection in a three-dimensional rectangular box. That layering tends to develop at high Rayleigh numbers can be viewed in a simple manner as being due to the thinning of the boundary layer and hence its decreasing ability to penetrate the upper–lower-mantle phase boundary. It is presumed that, during the Archaean, mantle temperatures were higher, the mantle viscosity was lower and plate velocities were high: together these imply that the Rayleigh number was higher. This could have had a major influence on the style of Archaean mantle convection, with a totally stratified system operating until such time that sufficient cooling had occurred for penetration of upper-mantle material into the lower mantle to take place. It is important that factors such as depth-dependent physical properties, temperature-dependent viscosity and plates on the surface be included in realistic spherical-shell convection models in order to determine whether these patterns of convection are similar to what actually occurs in the mantle.

Thus it seems that the geochemists and the geophysicists may perhaps *both* be correct – the mantle *is* stratified, but descending cold lithosphere *can* periodically can descend to the CMB, subducting plates *are* impeded at 670 km

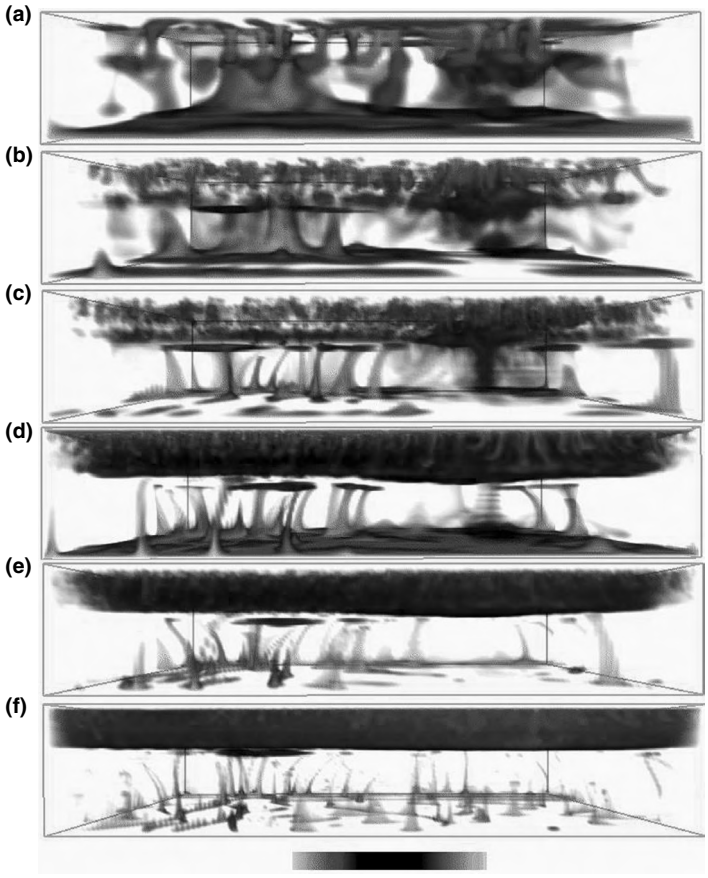


Figure 8.19. Convection in a compressible three-dimensional rectangular model mantle in which viscosity increases with depth and includes phase changes at 400 and 660 km. Rayleigh number: (a) 2×10^6, (b) 1×10^7, (c) 4×10^7, (d) 6×10^7, (e) 1×10^8 and (f) 4×10^8 Note the change in flow pattern that takes place with increasing Rayleigh number: upper and lower mantle become stratified, with episodic avalanches ceasing at higher Rayleigh numbers. Colour scheme: red, hotter; blue, colder. Colour version Plate 15. (Reprinted from *Phys. Earth Planet. Interiors*, **86**, Yuen, D. A. *et al*. Various influences on three-dimensional mantle convection with phase transitions, 185–203, Copyright (1994), with permission from Elsevier.)

and hotspots *may* originate at the CMB. There is clearly much to be learned from further study of convection processes in increasingly realistic mantle models: the dynamics of mantle convection remains the subject of much research activity.

8.2.4 Forces acting on the plates

The cold upper thermal boundary layer which forms in models of thermal convection of the mantle is assumed to represent the lithosphere. The motion of these lithospheric plates relative to each other and the mantle is associated with a number of forces, some of which drive the motion and some of which resist the motion. Figure 8.20 shows the main driving and resistive forces. If the plates are moving at a constant velocity, then there must be a force balance: driving forces = resistive forces.

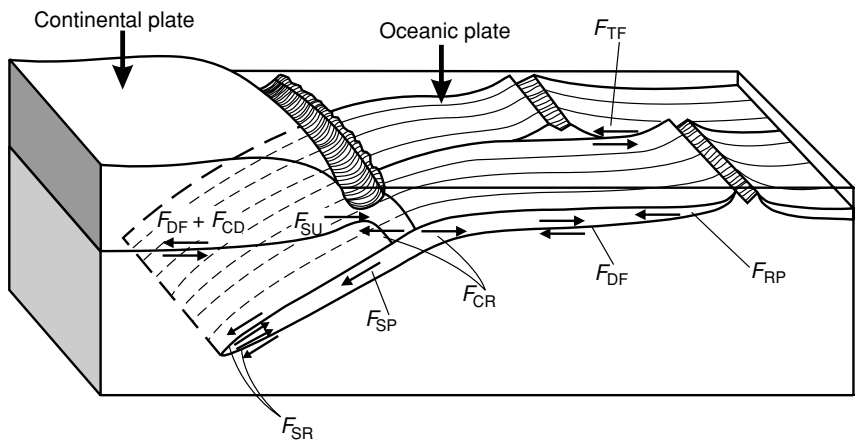


Figure 8.20. Possible forces acting on the lithospheric plates: F_{DF}, mantle-drag; F_{CD}, extra mantle-drag beneath continents; F_{RP}, ridge-push; F_{TF}, transform-fault resistance; F_{SP}, slab-pull; F_{SR}, slab resistance on the descending slab as it penetrates the asthenosphere; F_{CR}, colliding resistance acting on the two plates with equal magnitude and opposite directions; and F_{SU}, a suctional force that may pull the overriding plate towards the trench. (From Forsyth and Uyeda (1975).)

Driving forces

The *ridge-push* force acts at the mid-ocean ridges on the edges of the plates. It is made up of two parts: the pushing by the upwelling mantle material and the tendency of newly formed plate to slide down the sides of the ridge. Of these two, the sliding contribution is approximately an order of magnitude smaller than the upwelling contribution.

An estimate of the total ridge-push per unit length of the ridge axis, F_{RP}, is

$$F_{RP} = ge(\rho_m - \rho_w)\left(\frac{L}{3} + \frac{e}{2}\right) \tag{8.38}$$

where e is the elevation of ridge axis above the cooled plate, ρ_m the density of the mantle at the base of the plate, ρ_w the density of sea water and L the plate thickness (Richter and McKenzie 1978). Equation (8.38) gives F_{RP} as $2 \times 10^{12}\,\mathrm{N\,m^{-1}}$ (N, newton) for the following values: L, 8.5×10^4 m; e, 3×10^3 m; ρ_w, $10^3\,\mathrm{kg\,m^{-3}}$; ρ_m, $3.3 \times 10^3\,\mathrm{kg\,m^{-3}}$; and g, $9.8\,\mathrm{m\,s^{-2}}$.

The other main driving force is the negative buoyancy of the plate being subducted at a convergent plate boundary. This arises because the subducting plate is cooler and therefore more dense than the mantle into which it is descending. This force is frequently known as *slab-pull*. An estimate of the slab-pull force per unit length of subduction zone, $F_{SP}(z)$, acting at depth z and caused by the density contrast between the cool plate and the mantle is given by

$$F_{SP}(z) = \frac{8g\alpha\rho_m T_1 L^2 Re_t}{\pi^4}\left[\exp\left(-\frac{\pi^2 z}{2Re_t}\right) - \exp\left(-\frac{\pi^2 d}{2Re_t}\right)\right] \tag{8.39}$$

where z is the depth beneath the base of the plate, α the coefficient of thermal expansion, T_1 the temperature of the mantle, $d + L$ the thickness of the upper mantle and Re_t the *thermal Reynolds number*, given by

$$Re_t = \frac{\rho_m c_P v L}{2k} \tag{8.40}$$

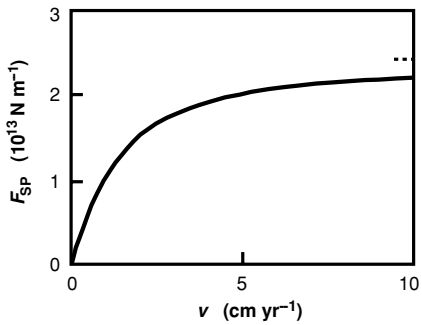


Figure 8.21. The total driving force available from the subducting slab, $F_{SP}(0)$ as a function of subduction velocity v. The horizontal dashed line shows the asymptotic limit as v increases to infinity. (After Richter and McKenzie (1978).)

where c_P is the specific heat and v is the rate at which the slab sinks. The total force available is F_{SP} evaluated at $z = 0$, $F_{SP}(0)$. $F_{SP}(z)$ decreases with depth into the mantle, until, by $z = d$, it is zero, $F_{SP}(d) = 0$. Figure 8.21 shows the dependence of this total force on the consumption velocity v. An additional driving force in the sinking slab would arise if the olivine–spinel phase change within the slab were elevated compared with the mantle (Fig. 9.44). However, if there is a metastable olivine wedge in the slab, then this would be a resistive force rather than a driving force. The magnitude of this force is about half that caused by the difference in temperature between the slab and the mantle. The total slab-pull force is estimated to be $10^{13}\,\mathrm{N\,m^{-1}}$ in magnitude, which is greater than the $10^{12}\,\mathrm{N\,m^{-1}}$ of the ridge-push force. Both slab-pull and ridge-push are caused by the difference in density between hot and cold mantle; hot mantle can rise only because cold mantle sinks.

Resistive forces

Resistive forces occur locally at the ridge axis (occurrence of shallow earthquakes), along the bases of the plates as *mantle-drag* (assuming that the mantle flow is less than the plate velocity; if the reverse is true, then this would be a driving force), along transform faults (earthquakes) and on the descending slab. Estimates of these forces suggest that the resistive force acting on the top of the sinking slab is greater than the shear force acting on its sides. The resistive force acting on the base of the plate is proportional to the area of the plate but is of the same magnitude as the resistive forces acting on the descending slab. These resistive forces cannot easily be estimated analytically and must be calculated numerically from the differential equations for flow in a fluid. The forces are proportional to the product of mantle viscosity η and plate velocity v and are about $10^{13}\,\mathrm{N\,m^{-1}}$ in magnitude (depending on the value of mantle viscosity assumed). For a 6000–10 000-km-long plate, they would total $(80–100)\eta v$.

It is difficult to estimate the resistive forces acting on faults. However, the stress drop for large earthquakes is $\sim 10^6\,\mathrm{N\,m^{-2}}$ in magnitude. Earthquakes at ridge axes are shallow and small, and their contribution to the resistive forces can be ignored in comparison with the fluid-dynamic drag forces. The resistive forces acting

on transform faults are harder to evaluate. Earthquakes on transform faults are usually shallow, even though the plates can be perhaps 80 km in thickness (Section 9.5.3). It is probable that their total resistive contribution is of the same magnitude as the ridge-push driving force, or smaller. Estimates of the resistive force acting on thrusts at the convergent plate boundaries, as indicated by earthquakes, give values of 10^{12} N m^{-1}. Again, this is less than the mantle-drag force.

To summarize, the main driving force is slab-pull, and the main resistive forces occur as drag along the base of the plate and on the descending slab.

Does mantle convection control plate tectonics?

Whether convection in the mantle drags the plates around or whether the forces acting at the edges of the plates drive the plates, which in turn drag the mantle, is a complicated 'chicken or the egg' type of question. From analysis of the driving and resistive forces, it is clear that the pull of the descending slab is a factor in determining the form of mantle flow. If the only locations for ridges were above the rising limbs of convection cells, then in simple schemes (e.g., Fig. 8.14(a)) each plate should have one edge along a ridge and the other along a subduction zone. Clearly this is not the case; for example, the Antarctic and African plates are bounded almost entirely by ridges. Where could the return flow go? In these instances it seems reasonable to assume that the ridges form where the lithosphere is weakest and that mantle material rises from below to fill the gap. Plates with subducting edges move with higher velocities than do those without (see Figs. 2.2 and 2.20), in agreement with the earlier estimate of the importance of the slab-pull driving force. To first order, ridge-push and continental collisional forces control the stress regimes in the plate interiors (Fig. 2.21).

Analysis of the stress within the North American continent permits analysis of the forces which drive and deform the continental part of that plate. The main driving force is the ridge-push from the Mid-Atlantic Ridge. Since the resistive forces amount to only about a quarter of the driving forces, the continent is being compressed against the Pacific plate to the west. The implication of the low values for resistive forces is that the 'root' beneath the North American continent which extends down to ~300 km is moving as one with the underlying asthenosphere.

Thus, in conclusion, though there is still much that is not understood about flow in the mantle and the motion of lithospheric plates, the pull of the descending plate at convergent boundaries due to its decrease in temperature seems to be a major factor both in the thermal modelling of the mantle flow and in the mechanical models of the forces involved.

Did plate tectonics operate during the Archaean?

A force-balancing model can be used to investigate the possibility of plate tectonics operating during the Archaean and to estimate probable plate velocities. The Earth was probably much hotter then than it is now, with temperatures at the top of the asthenosphere of about 1700 °C compared with 1300–1400 °C today.

The ridge-push force would then be about $4 \times 10^{11}\,\mathrm{N\,m^{-1}}$ and the slab-pull $8 \times 10^{12}\,\mathrm{N\,m^{-1}}$ (from Eqs. (8.38) and (8.39)). Equating driving and resistive forces enables estimates of viscosity and plate velocity to be made. Plate tectonics could operate very effectively over an upper mantle with dynamic viscosity 10^{18} Pa s. Velocities could have been high, about $50\,\mathrm{cm\,yr^{-1}}$. High plate velocities may have been necessary during the Archaean in order to maintain a high rate of heat loss through the oceans since, despite the higher temperatures and heat generation prevalent at that time, the thermal gradients determined from Archaean continental metamorphic rocks are relatively low. This topic is discussed further in Section 10.5.

8.3 The core

8.3.1 Temperatures in the core

Attempts to calculate the temperature at the centre of the Earth using conduction models (Section 7.4) fail because heat is primarily convected through much of the Earth. The fine detail of the temperature structure of the mantle depends on its dynamic structure. Figure 7.16(a) shows two possible temperature models, one with the upper and lower mantle convecting separately and the other for the whole mantle convecting with no boundary at 670 km depth. The temperature structure of the core is another important constraint on the temperature structure of the mantle because it controls the amount of heat crossing the core–mantle boundary. Conversely, to calculate the temperatures in the core, it is necessary to start with a temperature for the base of the mantle. Over 20% of the heat lost from the Earth's surface may originate from the core. This means that the core has an important role in mantle convection and plate tectonics. Since the surface area of the core is about one-quarter of the Earth's surface area, the heat flow across the CMB is comparable to that at the Earth's surface (Table 7.3). The other major unknowns are the physical properties, at very high temperatures and pressures, of the iron and iron alloys of which the core is composed (see Section 8.1.5). High-pressure melting experiments for iron alloys show that the presence of sulphur lowers the melting temperature of iron, whereas oxygen seems to raise it. Nickel is presumed to lower the melting temperature. Thus the details of the composition of the core affect its temperature. Nevertheless, despite these difficulties core temperatures can be estimated, albeit subject to large errors.

Diamond-anvil laboratory equipment that allows material to be studied at the very high temperatures and pressures of the core has recently been developed. Experiments involving diamond anvils differ from the shock-wave experiments in that they allow samples under study to be maintained at core temperatures and pressures. Pressures up to 150 GPa are attainable. The pressure at the core–mantle boundary is about 136 GPa (1.36 million times atmospheric pressure), whereas the pressure at the centre of the Earth is about 362 GPa (see Section 8.1.2). In

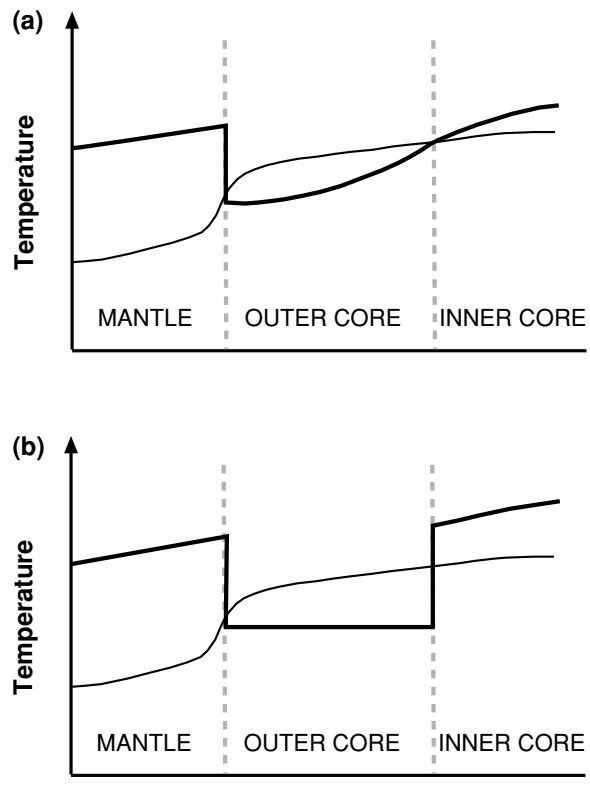


Figure 8.22. Schematic diagrams of possible melting temperatures for the mantle and core and the actual temperature profile. Heavy line, melting curve; lighter line, actual temperature profile. (a) Chemically homogeneous core. As the core cools, the inner core grows. (b) The inner and outer core have different chemical compositions and hence different melting temperatures. An outer core composed of an Fe–S or Fe–O alloy would have a much lower melting temperature than would a pure-iron inner core.

the shock-wave experiments, the samples are subjected to core pressures only instantaneously. High-pressure and -temperature diamond-anvil and shock-wave experiments on iron and iron compounds have produced differing results that are difficult to reconcile. Hence there is considerable uncertainty about temperatures, pressures and the resulting phase diagram for iron. However, present estimates based on these experiments and on *ab initio* theoretical calculations are ~6000 K for the melting temperature of pure iron and ~5600 K for the melting temperature of an iron alloy, both at the outer-core–inner-core interface. The temperature at the centre of the Earth is 6000 ± 500 K. Figure 7.16(b) shows estimates of the probable temperature structure within the Earth. The higher temperatures for the mantle are similar to those shown in Fig. 7.16(a) for a two-layer mantle, while the lower temperatures are for a single-layer mantle.

That the outer core is liquid and the inner core solid is a consequence of the melting curve for iron. The temperature in the outer core is above the melting temperature of iron and so the outer core is molten. The temperature in the inner core is below the melting temperature and so the inner core is solid (Fig. 8.22). If the core is chemically homogeneous and if it is slowly cooling, the inner core will progressively grow with time and the inner-core–outer-core boundary will be at the melting temperature of iron. If the inner core and outer core are of

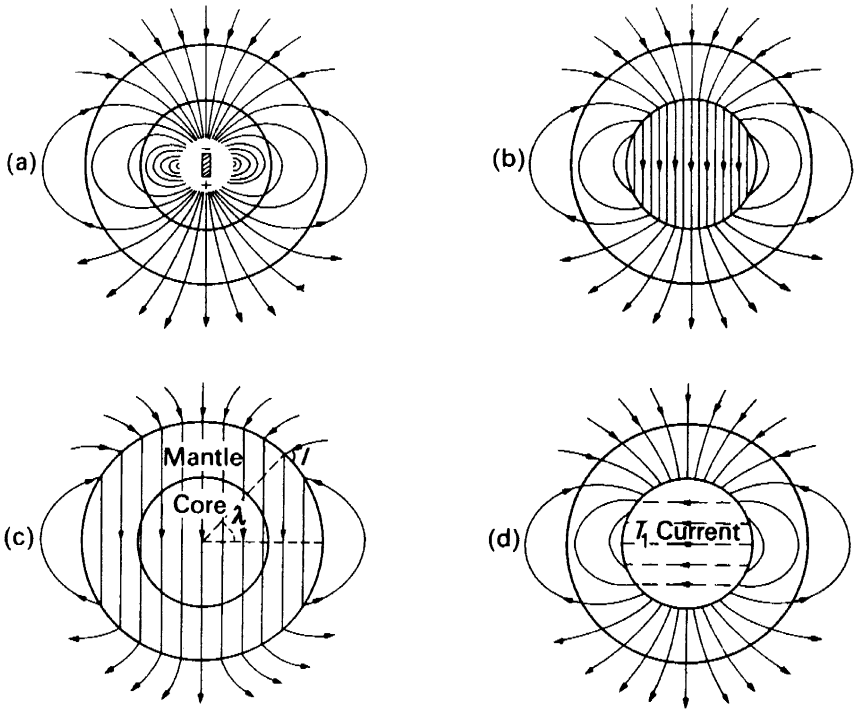


Figure 8.23. Four possible models for producing the Earth's main dipole field: (a) a dipole at the centre of the Earth, (b) a uniformly magnetized core, (c) a uniformly magnetized core and mantle ($\tan I = 2\tan\lambda$) and (d) a current system flowing east–west around the core–mantle boundary. (From Bott (1982).)

different compositions, the depression of the melting temperature in the liquid outer core due to impurities may mean that the temperature at that boundary is below the melting temperature of pure iron. It is essential that the high-pressure and -temperature phase diagram for iron be well determined, since it controls both the geochemistry and the geophysics of the core as well as the evolution of the Earth as a whole.

8.3.2 Convection in the outer core and the Earth's magnetic field

The first suggestion that the Earth's magnetic field is similar to that of a uniformly magnetized sphere came from William Gilbert in 1600 (see Section 3.1.2). Carl Friedrich Gauss (1777–1855) later formally showed that the magnetized material or the electrical currents which produce the field are not external to the Earth but are internal. Figure 8.23 shows four possible models for producing the Earth's main dipole field: (a) a magnetic dipole at the centre of the Earth, (b) a uniformly magnetized core, (c) a uniformly magnetized Earth and (d) an east–west electrical current flowing around the core–mantle boundary. Because the mantle is composed of silicates (see Section 8.1.5), it is not a candidate for the origin of the magnetic field. Permanent magnetization of the mantle or core cannot produce the Earth's magnetic field because temperatures in the deep interior far

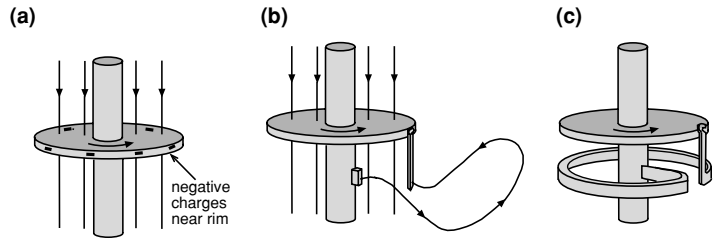


Figure 8.24. The development of a self-exciting dynamo. (a) A metal disc rotating on an axle in a magnetic field. Charge collects on the rim of the disc but cannot go anywhere. (b) A wire joining the rim of the disc to the axle enables current to flow. (c) The wire joining the rim to the axle is modified so that it is a coil looping around the axle. Now the current flowing reinforces the magnetic field, which will induce more current, thus sustaining the magnetic field. This is a self-exciting dynamo. (From Bullard (1972).)

exceed the Curie temperatures for magnetic minerals (see Section 3.1.3). These two facts rule out the model of a uniformly magnetized Earth. The core is predominantly composed of iron and could produce the magnetic field. The Earth's magnetic field is not a constant in time but at present is slowly decreasing in strength and drifting westwards. It undergoes irregular reversals as discussed and used in Chapter 3. This changeability indicates that it is unlikely that the core is uniformly magnetized or that there is a magnetic dipole at the centre of the Earth. This leaves an electrical-current system as the most plausible model for producing the magnetic field. The problem with such an electrical-current system is that it must be constantly maintained. If it were not, it would die out in much less than a million years;[7] yet we know from palaeomagnetic studies that the magnetic field has been in existence for at least 3500 Ma.

The model that best explains the magnetic field and what we know of the core is called the *geomagnetic dynamo* or *geodynamo*. A mechanical model of a *self-exciting dynamo* was developed in the 1940s by W. M. Elsasser and Sir Edward Bullard. Figure 8.24 shows how it works. A simple dynamo is sketched in Fig. 8.24(a): a metal disc on an axle rotating in a magnetic field. The disc is constantly cutting the magnetic field, and so a potential difference (voltage) is generated between the axle and the rim of the disc. However, since there is nowhere for current to flow, the charge can only build up around the rim. In Fig. 8.24(b), a wire is connected between the rim and the axle so that current is able to flow, but, if the external magnetic field is removed, the current stops flowing. In Fig. 8.24(c), the wire connecting the rim to the axle is coiled around the axle; now the current flowing in the coil gives rise to a magnetic field that

[7] The ohmic decay time (R^2/η, where η is the magnetic diffusivity of the core and R its radius) of the core is about 60 000 yr.

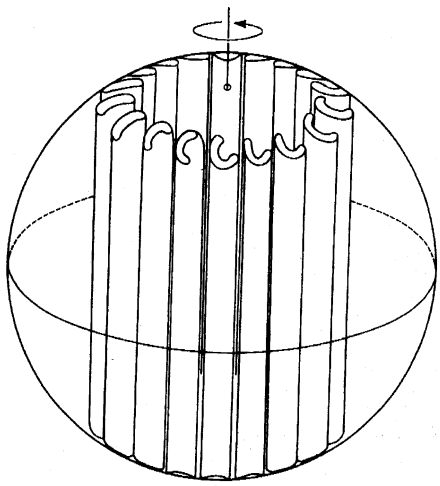


Figure 8.25. Convection currents in a laboratory model of the outer core, a rotating sphere containing a concentric liquid shell and an interior sphere. Thermal convection in the fluid was produced by maintaining a temperature difference between the inner and outer spheres. The convection cells which resulted were slowly spinning rolls; those in the northern and southern hemispheres had opposite polarity. Such a convection system in an electrically conductive outer core would be capable of generating the Earth's dipole field. These convection rolls drift in the same direction as the rotation (arrow). (From Gubbins (1984).)

reinforces the original field. So, when the disc rotates fast enough, the system is self-sustaining, producing its own magnetic field. Unlike a bicycle dynamo, which has a permanent magnet, this dynamo does not need a large constant magnetic field to operate; a slight transient magnetic field can be amplified by the dynamo. All that is necessary is for the disc to be rotating. For this reason this model is often called a self-exciting dynamo. The input of energy to power the dynamo is that required to drive the disc.

An interesting feature of the dynamo shown in Fig. 8.24(c) is that it works either with the current and field as illustrated or with both reversed. This means that, like the Earth's dynamo, such a dynamo is capable of producing a reversed magnetic field. However, unlike the Earth's dynamo, the dynamo in Fig. 8.24(c) cannot reverse itself unless the circuit includes a shunt. It is most unlikely that the Earth's dynamo is like this self-exciting disc dynamo. To start with, because the disc dynamo has a hole in it and is antisymmetrical, it is topologically different from the core. Also, it is hard to imagine that such a simple electrical-current system could operate in the core without short-circuiting itself somewhere. Nevertheless, it has been demonstrated that there are fluid motions in the liquid outer core that can generate a magnetic field that can undergo random reversals.

The whole subject of magnetic fields in fluids is known as *magnetohydrodynamics*. The mathematical equations governing fluid motion in the outer core and generation of a magnetic field are a very complex interrelated set of non-linear partial differential equations. They can, however, be separated (Jacobs 1987) into four groups:

(a) the electromagnetic equations relating the magnetic field to the velocity of the fluid in the outer core;

(b) the hydrodynamic equations, including conservation of mass and momentum and the equation of motion for the fluid in the outer core;
(c) the thermal equations governing the transfer of heat in a flowing fluid or the similar equations governing compositional convection; and
(d) the boundary and initial conditions.

Simultaneous solution of all these equations is exceedingly difficult, in part because the equations are non-linear. However, in special situations solutions can be found for some of the equations. One such simplified approach is to assume a velocity field for the flow in the outer core and then to solve the electromagnetic equations of group (a) to see what type of magnetic field it would generate. Another line of work has been to investigate group (b), possible fluid motions in a fluid outer core sandwiched between a solid mantle and a solid inner core. Figure 8.25 shows the fluid motions observed in a scaled laboratory experiment using a rotating spherical model, with the fluid outer core subjected to a temperature gradient. The convection cells in this model core were cylindrical rolls, with the fluid spiralling in opposite directions in the northern and southern hemispheres. The Coriolis force means that the rolls are aligned with the rotation axis. The dynamics of the flow are significantly affected by the inner core: the rolls are unstable close to the axis and can touch the inner core. The problem with applying flow patterns such as these directly to dynamo models is that any flow pattern is markedly altered by the magnetic field it generates. Figure 8.26 shows schematically the interaction between magnetic field and fluid flow for one dynamo model, the Parker–Levy dynamo. For this particular dynamo model to be self-sustaining, four conditions must be satisfied.

1. The initial dipole field must be aligned along the Earth's spin axis.
2. The fluid outer core must be rotating.
3. There must be upwelling thermal convection currents in the outer core.
4. A spiralling motion of the convection system caused by the Coriolis force is required. The spiralling motions have opposite polarities in the northern and southern hemispheres.

The rotation of the electrically conducting fluid in the outer core will stretch the original dipole magnetic-field lines and wind them into a toroidal field. The interaction of this toroidal magnetic field with the convecting rolls then results in a magnetic field with loops that are aligned with the rotation axis. If the loops have the same sense as the original field, that dipole field can be regenerated; but if the loops have the opposite sense, the original dipole field can be reversed.

Although it can be shown that a convecting outer core can act as a dynamo that undergoes intermittent polarity reversals, exactly why these reversals occur is not clear. They could be due to the random character of the irregular fluid convection and the non-linear coupling of the fluid motion with the magnetic

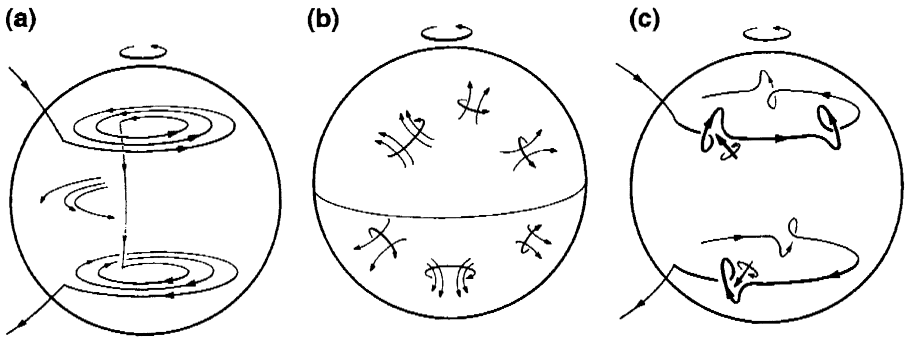


Figure 8.26. The Parker–Levy dynamo. (a) Rotation of an electrically conducting fluid outer core results in the stretching of the magnetic-dipole field lines; they are wound into a toroidal field. Toroidal magnetic fields cannot be detected at the Earth's surface because of the intervening insulating mantle. (b) The Coriolis force acting on the convecting fluid gives rise to spiralling motions (as in Fig. 8.25). The motion has opposite polarity in the northern and southern hemispheres. Such cyclonic motions are analogous to atmospheric cyclones and anticyclones. (c) The toroidal field lines shown in (a) are further deformed into loops by the spiralling motions shown in (b). These loops tend to rotate into longitudinal planes and so effectively regenerate the original dipole field. (From Levy (1976). Reproduced with permission from the *Annual Reviews of Earth and Planetary Sciences,* Vol. 5, © 1976 by Annual Reviews Inc.)

field or to changing boundary conditions, or to the influence of the inner core on the flow. It seems from palaeomagnetic measurements that, during a reversal, the magnitude of the field diminished to about 10% of its normal value, and the path followed by the north magnetic pole was a complex wandering from north to south rather than a simple line of longitude from north to south. The length of time necessary to complete a reversal is short, approximately 5000 yr or less. Since the inner core is a conductor, there is an electromagnetic coupling between the outer core and the inner core. This is likely to be the main coupling between the inner core and the outer core because the major viscosity contrast there (Table 8.3) means that viscous coupling will be weak. There may, however, be some coupling resulting from topography on the boundary if it is not smooth. The electromagnetic coupling between the inner and outer core means that the inner core has a stabilizing effect on the geodynamo: field reversals will take place only when fluctuations exceed a threshold value. Models of the geodynamo have often ignored the fact that the inner core is a conductor.

Three-dimensional dynamic computer simulations of a rotating electrically conducting fluid outer shell surrounding a solid inner sphere have produced a magnetic field that underwent reversal and had a differential rotation of the solid inner sphere. However, such numerical simulations cannot be directly applied to

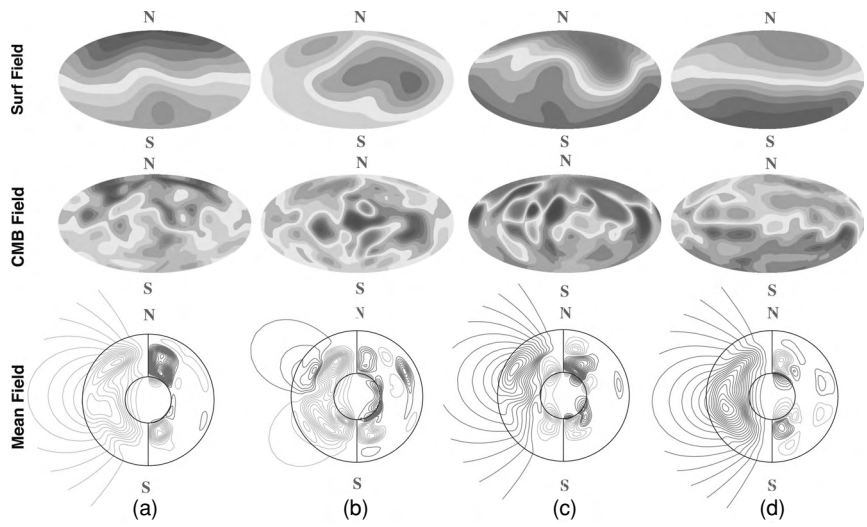


Figure 8.27. A sequence of images of a numerical dynamo model during a reversal of the magnetic field. Columns (a)–(d) show images for every 3000 yr. Top row, map view of the radial field at the earth surface. Middle row, map view of the radial field at the core–mantle boundary. Orange, outward field; blue, inward field. Note that the magnitude of the surface field is dispayed magnified by a factor of ten. Bottom row, longitudinally averaged magnetic field through the core. Outer circle, core–mantle boundary; inner circle, inner core. The right-hand half of each plot shows contours of the toroidal field direction and intensity (red lines, eastwards; blue lines, westwards). The left-hand half of each plot shows magnetic-field lines for the poloidal field (red lines, anticlockwise; blue lines, clockwise). Colour version Plate 16. (Reprinted with permission from *Nature* (Glatzmaier *et al.*, *Nature*, **401**, 885–90) Copyright 1999 Macmillan Magazines Ltd.)

the Earth yet, since realistic values for all parameters[8] are not yet computationally achievable. Figure 8.27 shows stages of a magnetic reversal in progress in a numerical dynamo model. Increasing computer power should mean that, in the next decade, simulations of realistic core models that can aid our understanding of the geodynamo and the dynamics of the inner and outer core may be achievable.

The liquid outer core has a viscosity of about 10^{-3} Pa s (comparable to that of water at room temperature and pressure) while the viscosity of the inner core is about $10^{13\pm3}$ Pa s (Table 8.3). The flow velocity in the outer core is about 10^4 m yr^{-1} and locally even faster flows occur over short distances. These

[8] For the core the Rossby number (*Ro*), the ratio of the inertial forces to the Coriolis forces, is about 10^{-8}. The Ekman number (*E*), the ratio of the viscous forces to the Coriolis forces, is $\ll 10^{-9}$. The Roberts number (*q*), the ratio of thermal diffusivity to magnetic diffusivity, is $\sim 10^{-5}$.

Table 8.3 *Physical properties of the core*

	Outer core	Inner core
Density ($kg\,m^{-3}$)	9900–1216	1276–1308
Volume ($10^{18}\,m^3$)	169	8
Mass (10^{21} kg)	1841	102
Viscosity[a] (Pa s)	$\sim 10^{-3}$	$10^{13\pm3}$
Seismic quality factor[a]		
Q_p (at 1 Hz)	6×10^4	250–600
Q_s (at 10^{-3} Hz)		100–400
Electrical conductivity[a] ($10^5\,S^{-1}\,m^{-1}$)	6 ± 3	6 ± 3
Specific heat[a] ($10^3\,J\,kg^{-1}\,K^{-1}$)	0.5 ± 0.3	0.5 ± 0.3
Coefficient of thermal expansion[a] ($10^{-6}\,K^{-1}$)	8 ± 6	7 ± 4
Thermal diffusivity[a] ($10^{-5}\,m^2\,s^{-1}$)	1.5 ± 1	1.5 ± 1

[a] After Jeanloz (1990).

velocities are orders of magnitude greater than velocities of flow in the mantle. The convection in the outer core is not completely separate from the convection in the mantle; the two convection systems are weakly coupled. The difference in viscosity across the CMB is so great that viscous coupling (via shear forces across the CMB) is not important, but there is some thermal coupling. This means that convection cells in the outer core tend to become aligned with convection cells in the mantle, with upwelling in the outer core beneath hot regions in the mantle and downwelling in the outer core beneath cold regions in the mantle. The heat flux from the core is a major driving force for convection in the mantle.

The changes in density resulting from convection in the outer core are much less than density variations in the mantle and cannot be imaged by seismic methods. Changes in the magnetic field, secular variation, can, however, be used to make estimates of the fluid flow at the surface of the core just below the CMB. The steady toroidal flow is fairly well determined at a maximum of $20\,km\,yr^{-1}$ with two main cells: a strong westward flow along the equator extends from the Indian Ocean to the Americas and then diverges, with the return flows extending north and south to high latitudes (Fig. 8.28). The toroidal flow beneath the Pacific region is very small. This flow is the probable cause of the 'westward drift' of the magnetic field. If the angular momentum of the core is constant, this westward flow at shallow levels in the outer core must be balanced by eastward movements at depth – the dfferential rotation of the inner core seems to satisfy that requirement. In contrast to the toroidal flow, estimates of the poloidal flow and the time dependence of the flow are poor.

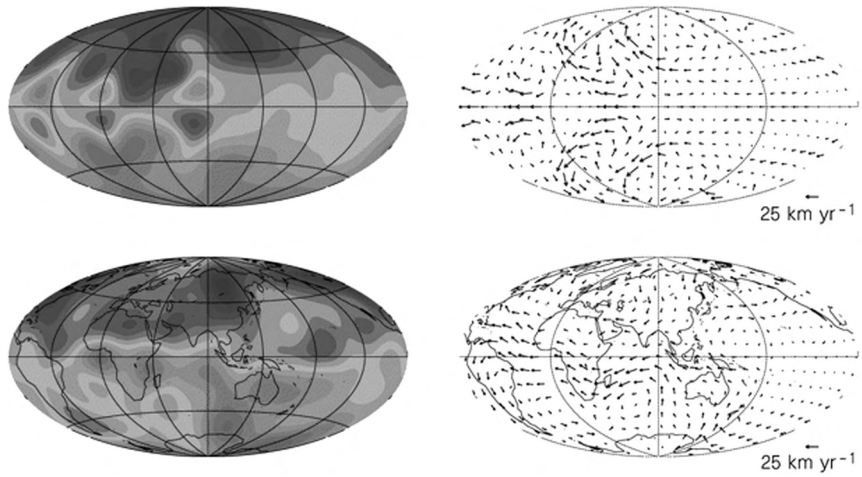


Figure 8.28. Images show the radial component of the magnetic field (left) and the fluid flow at the core–mantle boundary (right). Upper pair, numerical model of the geodynamo with viscous stress-free boundary conditions at the rigid boundaries; lower pair, Earth's field averaged over the years 1840–1990. Colour version Plate 17. Reprinted with permission from *Nature* (Kuang and Bloxham, *Nature*, **389**, 371–4) Copyright 1997 Macmillan Magazines Ltd.)

8.3.3 What drives convection in the outer core and how has the core changed with time?

What drives or powers the dynamo? Estimates of the power needed to drive the dynamo are 10^{11}–10^{12} W, a fraction of the $\sim 4 \times 10^{12}$ W heat flow through the CMB. Any convection in the outer core must involve an inherent density instability, with less dense material lying beneath denser material. Such a density instability could be due to changes in the Earth's rotation or to heat (thermal buoyancy), or it could result from chemical differences in the core. A number of heat sources could contribute to thermally driven convection in the outer core. One possible source is the radioactive isotopes ^{235}U and ^{40}K, which are present in the crust and the mantle, and possibly in the core, but this can provide only part of the energy needed to drive the dynamo. Another possible source of heat is the *primordial heat*: that heat which resulted from the formation of the Earth, which the core is slowly losing. If the inner core is cooling, solidifying and separating from the liquid outer core, then the latent heat of solidification (crystallization) would also provide heat to help power the dynamo. However, it is likely that chemical differences provide most of the energy needed to power the dynamo, with density instabilities arising as the outer core crystallizes dense iron crystals. These crystals, being denser than the liquid iron alloy of the outer core, fall towards the inner core, and the less dense liquid rises. The gravitational energy released in this process is then sufficient to drive the dynamo. Some energy is

ultimately dissipated as heat in the outer core and so contributes to the heat flow into the mantle.

The formation of the core, the evolution of the core and mantle and the magnetic field are thus inextricably linked. We know that the Earth has had a magnetic field since at least the early Archaean, so the inner core must have formed early, with crystallization proceeding at a sufficient rate to power a dynamo. However, the crystallization of the inner core has proceeded sufficiently slowly that the inner core is still only a fraction of the core. The Rayleigh number of the flow in the outer core must be fairly low. Since the mantle contains only low concentrations of Ni and S, the Ni and S in the core must date from the time of core formation. Ni and S can both alloy with Fe at low pressures, so it is probable that Fe alloyed with the available Ni and S early during the process of core formation. This explains their low concentrations in the mantle and means that the Ni level in the Earth matches its cosmic abundance. However, it is thought that some S may have been lost during the accretion of the Earth because even allowing for the maximum amount of S in the core that the seismic data imply still leaves the Earth as a whole depleted in S compared with cosmic abundances. Unlike Ni and S, oxygen cannot have been an early constituent of the core because it does not alloy with Fe at low pressures. This suggests that core composition has gradually changed with time, with the oxygen concentration of the outer core slowly increasing as the result of vigorous chemical reactions at the CMB. Numerical geodynamo models constructed to investigate the magnetic field for a younger Earth with a smaller core still yield a magnetic field that is primarily dipolar. However, the pattern of heat flux across the mantle–core boundary (and hence the pattern of convection in the mantle) may be an important factor in controlling the complexity of the magnetic field.

Much of what has been suggested in these sections about the core is still not certain. In our quest to understand the workings of the core we are hampered by the very high temperatures and pressures that must be attained in experiments, by the thick insulating mantle which prevents complete measurement and understanding of the magnetic field, by not having any sample of core material and by not yet being able to perform realistic numerical simulations. It has not been possible to construct a laboratory model in which the convective flow of a fluid generates a magnetic field. Materials available for laboratory experiments are not sufficiently good conductors for models to be of a reasonable size. Queen Elizabeth I, whose physician was William Gilbert, regarded Canada as *Terra Meta Incognita*. In this century the label should perhaps be transferred to the core.

References and bibliography

Ahrens, T. J. 1982. Constraints on core composition from shock-wave data. *Phil. Trans. Roy. Soc. Lond*. A, **306**, 37–47.

Aki, K. and Richards, P. G. 2001. *Quantitative Seismology – Theory and Methods*, 2nd edn. Columbia Press.

Akaogi, M., Ito, E. and Navrotsky, A. 1989. Olivine–modified spinel–spinel transitions in the system Mg_2SiO_4–Fe_2SiO_4: calorimetric measurements, thermochemical calculation and geophysical application. *J. Geophys. Res.*, **94**, 15 671–85.

Allègre, C. J., Hofmann, A. W. and O'Nions, R. K. 1996. The argon constraints on mantle structure. *Geophys. Res. Lett.*, **23**, 3555–7.

Anderson, O. L. 1999. A thermal balancing act. *Science*, **283**, 1652–3.

1986. Properties of iron at the Earth's core conditions. *Geophys. J. Roy. Astr. Soc.*, **84**, 561–79.

1978. *Q* of the Earth. *J. Geophys. Res.*, **83**, 5869–82.

Anderson, O. L. and Dziewonski, A. M. 1984. Seismic tomography. *Sci. Am.*, **251** (4), 60–8.

Anderson, O. L. and Hart, R. S. 1976. An Earth model based on free oscillations and body waves. *J. Geophys. Res.*, **81**, 1461–75.

Ballentine, C. J. 2002. Tiny tracers tell tall tales, *Science*, **296**, 1247–8.

Båth, M. 1979. *Introduction to Seismology*, 2nd edn. Basel: Birkhäuser-Verlag.

Bijwaard, H. and Spakman, W. 2000. Non-linear global P-wave tomography by iterated linearized inversion. *Geophys. J. Int.*, **141**, 71–82.

Bijwaard, H., Spakman, W. and Engdahl, E. R. 1998. Closing the gap between regional and global travel time tomography. *J. Geophys. Res.*, **103**, 30 055–78.

Birch, F. 1968. On the possibility of large changes in the Earth's volume. *Phys. Earth Planet. Int.*, **1**, 141–7.

Bloxham, J. 2000. Sensitivity of the geomagnetic axial dipole to thermal core–mantle interactions. *Nature*, **405**, 63–5.

Bloxham, J. and Gubbins, D. 1985. The secular variation of Earth's magnetic field. *Nature*, **317**, 777–81.

1987. Thermal core–mantle interactions. *Nature*, **325**, 511–13.

Bloxham, J. and Jackson, A. 1991. Fluid flow near the surface of the Earth's outer core. *Rev. Geophys.*, **29**, 97–120.

Boehler, R. 2000. High-pressure experiments and the phase diagram of lower mantle and core materials. *Rev. Geophys.*, **38**, 221–45.

Bolt, B. A. 1987. 50 years of studies on the inner core. *EOS Trans. Am. Geophys. Un.*, **68**, 73.

Bott, M. H. P. 1982. *The Interior of the Earth: Its Structure, Composition and Evolution*. Amsterdam: Elsevier.

Brett, R. 1976. The current status of speculations on the composition of the core of the Earth. *Rev. Geophys. Space Phys.*, **14**, 375–83.

Buffett, B. A., Huppert, H. E., Lister, J. R. and Woods, A. W. 1992. Analytical model for solidification of the Earth's core. *Nature*, **356**, 329–31.

Bullard, E. C. 1949. Electromagnetic induction in a rotating sphere. *Proc. Roy. Soc. Lond*. A, **199**, 413.

1949. The magnetic field within the Earth. *Proc. Roy. Soc. Lond*. A, **197**, 433–53.

1972. The Earth's magnetic field and its origin. In I. G. Gass, P. J. Smith and R. C. L. Wilson, eds., *Understanding the Earth*, 2nd edn. Horsham: Artemis Press, pp. 71–9.

Bullard, E. C. and Gellman, H. 1954. Homogeneous dynamos and terrestrial magnetism. *Phil. Trans. Roy. Soc. Lond*. A, **247**, 213–78.

Bullen, K. E. and Bolt, B. A. 1985. *An Introduction to the Theory of Seismology*, 4th edn. Cambridge: Cambridge University Press.

Bunge, H.-P. and Grand, S. P. 2000. Mesozoic plate-motion history below the northeast Pacific Ocean from seismic images of the subducted Farallon slab. *Nature*, **405**, 337–40.

Bunge, H.-P., Richards, M. A. and Baumgartner, J. R. 1996. Effect of depth-dependent viscosity on the planform of mantle convection. *Nature*, **379**, 436–8.

1997. A sensitivity study of three-dimensional spherical mantle convection at 10^8 Rayleigh number: effects of depth dependent viscosity, heating mode, and an endothermic phase change. *J. Geophys. Res*., **102**, 11 991–12 007.

Bunge, H.-P., Richards, M. A., Lithgow-Bertelloni, C., Baumgartner, J. R. and Romanowicz, B. A. 1998. Time scales and heterogeneous structure in geodynamic Earth models. *Science*, **280**, 91–5.

Busse, F. H. 1975. A model of the geodynamo. *Geophys. J. Roy. Astr. Soc.*, **42**, 437–59.

1983. Recent developments in the dynamo theory of planetary magnetism. *Ann. Rev. Earth Planet. Sci.*, **11**, 241–68.

Busse, F. H. and Carrigan, C. R. 1976. Laboratory simulation of thermal convection in rotating planets and stars. *Science*, **191**, 31–8.

Carrigan, C. R. and Busse, F. H. 1983. An experimental and theoretical investigation of the onset of convection in rotating spherical shells. *J. Fluid Mech.*, **126**, 287–305.

Carrigan, C. R. and Gubbins, D. 1979. The source of the Earth's magnetic field. *Sci. Am.*, **240** (2), 118–30.

Christensen, U. 1995. Effects of phase transitions on mantle convection. *Ann. Rev. Earth Planet. Sci.*, **23**, 65–87.

Claerbout, J. F. 1985. *Imaging the Earth's Interior*. Oxford: Blackwell Scientific.

Clark, S. P. 1966. *Handbook of Physical Constants*. Vol. 97 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America.

Cox, A. and Hart, R. B. 1986. *Plate Tectonics: How It Works*. Oxford: Blackwell Scientific.

Davies, G. F. 1999. *Dynamic Earth: Plates, Plumes and Mantle Convection*. Cambridge: Cambridge University Press.

Deuss, A., Woodhouse, J. H., Paulssen, H. and Trampert, J. 2000. The observation of inner core shear waves. *Geophys. J. Int.*, **142**, 67–73.

Dobrin, M. B. and Savit, C. H. 1988. *Introduction to Geophysical Prospecting*, 4th edn. New York: McGraw-Hill.

Dreibus, G. and Palme, H. 1996. Cosmochemical constraints on the sulfur content in the Earth's core. *Geochim. Cosmochim. Acta*, **60**, 1125–30.

Dziewonski, A. M. and Anderson, D. L. 1981. Preliminary reference Earth model. *Phys. Earth Planet. Interiors*, **25**, 297–356.

Dziewonski, A. M. and Woodhouse, J. H. 1987. Global images of the Earth's interior. *Science*, **236**, 37–48.

Elsasser, W. M. 1946. Induction effects in terrestrial magnetism. Part I: theory. *Phys. Rev*., **69**, 106–16.

1946. Induction effects in terrestrial magnetism. Part II: the secular variation. *Phys. Rev*., **70**, 202–12.

1947. Induction effects in terrestrial magnetism. Part III: electric modes. *Phys. Rev*., **72**, 821–33.

1950. The Earth's interior and geomagnetism. *Rev. Mod. Phys*., **22**, 1–35.
Forsyth, D. W. and Uyeda, S. 1975. On the relative importance of the driving forces of plate motion. *Geophys. J. Roy. Astr. Soc*., **43**, 163–200.
Garland, G. D. 1979. *Introduction to Geophysics – Mantle, Core and Crust*, 2nd edn. Philadelphia, Pennsylvania: Saunders.
Giardini, D., Xiang-Dong Li and Woodhouse, J. M. 1987. Three-dimensional structure of the Earth from splitting in free-oscillation spectra. *Nature*, **325**, 405–11.
Glatzmeier, G. A. 2002. Geodynamo simulations – how realistic are they? *Ann. Rev. Earth Planet. Sci*., **30**, 237–57.
Glatzmaier, G. A., Coe, R. S., Hongre, L. and Roberts, P. H. 1999. The role of the Earth's mantle in controlling the frequency of geomagnetic reversals. *Nature*, **401**, 885–90.
Glatzmeier, G. A. and Roberts, P. H. 1995. A three-dimensional convective dynamo solution with rotating and finitely conducting inner core and mantle, *Phys. Earth Planet. Int*., **91**, 63–75.
1995. A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature*, **377**, 203–9.
Grand, S. 1994. Mantle shear structure beneath the Americas and surrounding oceans. *J. Geophys. Res*., **99**, 11 591–621.
1997. Global seismic tomography: a snapshot of convection in the Earth. *GSA Today*, **7**, 1–7.
Grand, S., van der Hilst, R. D. and Widiyantoro, S. 1997. Global seismic tomography: a snapshot of convection in the Earth. *Geol. Soc. Am. Today*, **7** (4), 1–7.
Gubbins, D. 1974. Theories of the geomagnetic and solar dynamos. *Rev. Geophys. Space Phys*., **12**, 137–54.
1976. Observational constraints on the generation process of Earth's magnetic field. *Geophys. J. Roy. Astr. Soc*., **46**, 19–39.
1984. The Earth's magnetic field. *Contemp. Phys*., **25**, 269–90.
1987. Mapping the mantle and core. *Nature*, **325**, 392–3.
1989. *Seismology and Plate Tectonics*. Cambridge: Cambridge University Press.
1991. Harold Jeffreys Lecture 1990: Convection in the Earth's core and mantle. *Q. J. Roy. Astr. Soc*., **32**, 69–84.
1997. The Earth's inner core: its discovery and other surprises. *Astr. Geophys. J. Roy. Astr. Soc*., **38** (4), 15–18.
Gubbins, D. and Bloxham, J. 1986. Morphology of the geomagnetic field: implications for the geodynamo and core–mantle coupling. *EOS Trans. Am. Geophys. Un*., **67**, 916.
1987. Morphology of the geomagnetic field and its implications for the geodynamo. *Nature*, **325**, 509–11.
Gubbins, D. and Kelly, P. 1993. Persistent patterns in the geomagnetic field over the last 2.5 myr. *Nature*, **365**, 829–32.
Gubbins, D. and Masters, T. G. 1979. Driving mechanisms for the Earth's dynamo. *Adv. Geophys*., **21**, 1–50.
Gubbins, D., Masters, T. G. and Jacobs, J. A. 1979. Thermal evolution of the Earth's core. *Geophys. J. Roy. Astr. Soc*., **59**, 57–99.
Gubbins, D. and Roberts, P. H. 1987. Magnetohydrodynamics of the Earth's core. In J. A. Jacobs, ed., *Geomagnetism Vol. 2*. London: Academic Press, pp. 1–183.

Gung, Y., Panning, M. and Romanowicz, B. 2003. Global anisotropy and the thickness of continents. *Nature*, **422**, 707–11.

Gung, Y. and Romanowicz, B. 2004. *Q* tomography of the upper mantle using three component seismic waveforms. *Geophys. J. Int*. UPDATE.

Gurnis, M., Buffett, B. A., Knittle, E. and Wysession, M. E., eds., 1998. *The Core–Mantle Boundary*. Washington: American Geophysical Union.

Gutenberg, B. 1959. *Physics of the Earth's Interior*. New York: Academic Press.

Gutenberg, B. and Richter, C. F. 1954. *Seismicity of the Earth and Associated Phenomena*. Princeton, Massachusetts: Princeton University Press.

Hager, B. H., Clayton, R. W., Richards, M. A., Cormer, R. P. and Dziewonski, A. M. 1985. Lower mantle heterogeneity, dynamic topography and the geoid. *Nature*, **313**, 541–5.

Hart, R. S., Anderson, D. L. and Kanamori, H. 1977. The effect of attenuation on gross Earth models. *J. Geophys. Res*., **82**, 1647–54.

Hart, S. R. and Zindler, A. 1986, In search of a bulk-Earth composition. *Chem. Geol*., **57**, 247–67.

Headlin, M. A. H., Shearer, P. M. and van der Hilst, R. D. 1997. Seismic evidence for small-scale heterogeneity throughout the Earth's mantle. *Nature*, **387**, 145–50.

Helmberger, D. W. 1968. The crust–mantle transition in the Bering Sea. *Bull. Seism. Soc. Am*., **58**, 179–214.

Herrin, E. 1968. Introduction to 1968 seismological tables for P phases. *Bull. Seism. Soc. Am*., **58**, 1193–241.

Hewitt, J. M., McKenzie, D. P. and Weiss, N. O. 1975. Dissipative heating in convective fluids. *J. Fluid Mech*., **68**, 721–38.

1980. Large aspect ratio cells in two-dimensional thermal convection. *Earth Planet. Sci. Lett*., **51**, 370–80.

Hoffman, N. R. A. and McKenzie, D. P. 1985. The destruction of geochemical heterogeneities by differential fluid motions during mantle convection. *Geophys. J. Roy. Astr. Soc*., **82**, 163–206.

Hollerbach, R. and Jones, C. A. 1993. Influence of the Earth's outer core on geomagnetic fluctuations and reversals. *Nature*, **365**, 541–3.

Houseman, G. 1983. Large aspect ratio convection cells in the upper mantle. *Geophys. J. Roy. Astr. Soc*., **75**, 309–34.

Jackson, I., ed., 1998. *The Earth's Mantle: Composition, Structure and Evolution*. Cambridge: Cambridge University Press.

Jacobs, J. A. 1984. *Reversals of the Earth's Magnetic Field*. Bristol: Adam Hilger.

1987. *The Earth's Core*, 2nd edn. London: Academic Press.

Jeanloz, R. 1983. The Earth's core. *Sci. Am*., **249** (3), 13–24.

1990. The nature of the core. *Ann. Rev. Earth Planet. Sci*., **18**, 357–86.

1988. High-pressure experiments and the Earth's deep interior. *Phys. Today*, **41** (1), S44–5.

Jeanloz, R. and Ahrens, T. J. 1980. Equations of state of FeO and CaO. *Geophys. J. Roy. Astr. Soc*., **62**, 505.

Jeanloz, R. and Richter, F. M. 1979. Convection, composition and the thermal state of the lower mantle. *J. Geophys. Res*., **84**, 5497–504.

Jeffreys, H. 1939. The times of P, S and SKS and the velocities of P and S. *Mon. Not. Roy. Astr. Soc. Geophys. Suppl*., **4**, 498–533.

1939. The times of PcP and ScS. *Mon. Not. Roy. Astr. Soc. Geophys. Suppl*., **4**, 537–47.
1976. *The Earth*, 6th edn. Cambridge: Cambridge University Press.
Jeffreys, H. and Bullen, K. E. 1988. *Seismological Tables*. Cambridge: British Association Seismological Investigations Committee, Black Bear Press.
Kellogg, L., Hager, B. H. and van der Hilst, R. D. 1999. Compositional stratification in the deep mantle. *Science*, **283**, 1881–4.
Kennett, B. L. N. 1977. Towards a more detailed seismic picture of the oceanic crust and mantle. *Marine Geophys. Res*., **3**, 7–42.
1995. Seismic traveltime tables. In T. J. Ahrens, ed., *Global Earth Physics: a Handbook of Physical Constants, AGU Reference Shelf 1*. Washington: American Geophysical Union, pp. 126–43.
Kennett, B. L. N. and Engdahl, E. 1991. Traveltimes for global earthquake location and phase identification. *Geophys. J. Int*., **105**, 429–65.
Kennett, B. L. N., Engdahl, E. R. and Buland, R. 1995. Constraints on seismic velocities in the Earth from traveltimes. *Geophys. J. Int*., **122**, 108–24.
King, S. D. 2001. Subduction zones: observations and geodynamic models. *Phys. Earth Planet. Interiors*, **127**, 9–24.
Knittle, E. and Jeanloz, R. 1986. High-pressure metallization of FeO and implications for the earth's core. *Geophys. Res. Lett*., **13**, 1541–4.
Knittle, E., Jeanloz, R., Mitchell, A. C. and Nellis, W. J. 1986. Metallization of $Fe_{0.94}O$ at elevated pressures and temperatures observed by shock-wave electrical resistivity measurements. *Solid State Commun.*, **59**, 513–15.
Kuang, W. and Bloxham, J. 1997. An Earth-like numerical dynamo model. *Nature*, **389**, 371–4.
Kumazawa, M. and Anderson, O. L. 1969. Elastic moduli, pressure derivatives and temperature derivatives of single-crystal olivine and single-crystal forsterite. *J. Geophys. Res*., **74**, 5961–72.
Larmor, J. 1919. How could a rotating body such as the sun become a magnet? *Rep. Br. Assoc. Adv. Sci*., 159.
Lay, T. 1994. The fate of descending slabs. *Ann. Rev. Earth Planet Sci*., **22**, 33–61.
Lay, T., Ahrens, T. J., Olson, P., Smyth, J. and Loper, D. 1990. Studies of the Earth's deep interior: goals and trends. *Phys. Today*, **43** (10), 44–52.
Lay, T., Williams, Q. and Garnero, E. J. 1998. The core–mantle boundary layer and deep Earth dynamics. *Nature*, **392**, 461–8.
Levy, E. H. 1972. Kinematic reversal schemes for the geomagnetic dipole. *Astrophys. J*., **171**, 635–42.
1972. On the state of the geomagnetic field and its reversals. *Astrophys. J*., **175**, 573–81.
1976. Generation of planetary magnetic fields. *Ann. Rev. Earth Planet. Sci*., **4**, 159–85.
1979. Dynamo magnetic field generation. *Rev. Geophys. Space Phys*., **17**, 277–81.
Lithgow-Bertelloni, C. and Richards, M. A. 1998. The dynamics of Cenozoic and Mesozoic plate motions. *Rev. Geophys*., **36**, 27–78.
Loper, D. E. 1984. Structure of the core and lower mantle. *Adv. Geophys*., **26**, 1–34.
Loper, D. E. and Lay, T. 1995. The core–mantle boundary region. *J. Geophys. Res*., **100**, 6397–420.

Lowrie, W. 1997. *Fundamentals of Geophysics*. Cambridge: Cambridge University Press.

Masters, G. and Gubbins, D. 2003. On the resolution of density within the Earth. *Phys. Earth Planet. Interiors*, **140**, 159–67.

Masters, G., Laske, G., Bolton, H. and Dziewonski, A. 2000. The relative behaviour of shear velocity, bulk sound speed and compressional velocity in the mantle: implications for chemical and thermal structure. In S. Karato, A. M. Forte, R. C. Liebermann, G. Masters and L. Sixtrude, eds., *Earth's Deep Interior*. American Geophysical Union Monograph 117. Washington: American Geophysical Union, pp. 63–87.

Masters, T. G. and Shearer, P. M. 1990. Summary of seismological constraints on the structure of the Earth's core. *J. Geophys. Res*., **95**, 21 691–5.

1995. Seismic models of the Earth: elastic and anelastic. In T. J. Ahrens, ed., *Global Earth Physics: A Handbook of Physical Constants.* AGU Reference Shelf 1. Washington: American Geophysical Union, pp. 88–103.

Masters, T. G. and Widmer, R. 1995. Free oscillations: frequencies and attenuations. In T. J. Ahrens, ed., *Global Earth Physics: A Handbook of Physical Constants.* AGU Reference Shelf 1. Washington: American Geophysical Union, pp. 104–25.

McDonough, W. F. and Sun, S. S. 1995. The composition of the Earth. *Chem. Geol*., **120**, 223–53.

McKenzie, D. P. 1969. Speculations on the consequences and causes of plate motions. *Geophys. J. Roy. Astr. Soc*., **18**, 1–32.

1983. The Earth's mantle. *Sci. Am*., **249** (3), 66–113.

McKenzie, D. P. and O'Nions, R. K. 1982. Mantle reservoirs and ocean island basalts. *Nature*, **301**, 229–31.

McKenzie, D. P. and Richter, F. 1976. Convection currents in the Earth's mantle. *Sci. Am*., **235** (5), 72–89.

1981. Parametrized thermal convection in a layered region and the thermal history of the Earth. *J. Geophys. Res*., **86**, 11 667–80.

McKenzie, D. P., Roberts, J. M. and Weiss, N. O. 1974. Convection in the Earth's mantle: towards a numerical simulation. *J. Fluid Mech*., **62**, 465–538.

Mégnin, C. and Romanowicz, B. 2000. The shear velocity structure of the mantle from the inversion of body, surface and higher modes waveforms, *Geophys. J. Int*., **143**, 709–28.

Meissner, R. 1986. *The Continental Crust, a Geophysical Approach*. Vol. 34 of International Geophysics Series. Orlando, Florida: Academic Press.

Merrill, R. T., McElhinny, M. W. and McFadden, P. L. 1996. *The Magnetic Field of the Earth: Palaeomagnetism, the Core and the Deep Mantle*. International Geophysics Series 63. San Diego, California: Academic Press.

Moffatt, H. K. 1978. *Magnetic Field Generation in Electrically Conducting Fluids*. Cambridge: Cambridge University Press.

Montagner, J.-P. 1994. Can seismology tell us anything about convecton in the mantle? *Rev. Geophys*, **32**, 115–37.

Montagner, J.-P. and Kennett, B. L. N. 1996. How to reconcile body-wave and normal mode reference Earth models? *Geophys. J. Int*., **125**, 229–48.

Morelli, A. and Dziewonski, A. M. 1987. Topography of the core–mantle boundary and lateral homogeneity of the outer core. *Nature*, **325**, 678.

Müller, G. and Kind, R. 1976. Observed and computed seismogram sections for the whole Earth. *Geophys. J. Roy. Astr. Soc*., **44**, 699–716.

Officer, C. B. 1974. *Introduction to Theoretical Geophysics*. Berlin: Springer-Verlag.
Ohtani, E. and Ringwood, A. E. 1984. Composition of the core, I. Solubility of oxygen in molten iron at high temperatures. *Earth Planet. Sci. Lett*., **71**, 85–93.
Ohtani, E., Ringwood, A. E. and Hibberson, W. 1984. Composition of the core. II. Effect of high pressure on solubility of FeO in molten iron. *Earth Planet. Sci. Lett*., **71**, 94–103.
O'Nions, R. K. and Carter, S. R. 1981. Upper mantle geochemistry. In C. Emiliani, ed., *The Sea*, Vol. 7. New York: John Wiley and Sons, pp. 49–71.
Parker, E. N. 1955. Hydromagnetic dynamo models. *Astrophys. J*., **122**, 293–314.
1969. The occasional reversal of the geomagnetic field. *Astrophys. J*., **158**, 815–27.
1979. *Cosmical Magnetic Fields*. New York: Oxford University Press.
Parsons, B. and McKenzie, D. P. 1978. Mantle convection and the thermal structure of the plates. *J. Geophys. Res*., **83**, 4485–96.
Peltier, W. R., ed., 1989. *Mantle Convection, Plate Tectonics and Global Dynamics*. New York: Gordon and Breach Sciences Publishers.
Phillips, O. M. 1968. *The Heart of the Earth*. San Francisco, California: Freeman, Cooper & Co.
Richards, M. A., Bunge, H.-P. and Lithgow-Bertollini, C. 2000. Mantle convection and plate motion history – toward general circulation models. In. M. A. Richards, R. G. Gordon and R. D. van der Hilst, eds., *The history and Dynamics of Global Plate Motions.* Geophysical Monograph 121. Washington: American Geophysical Union, pp. 289–307.
Richter, F. M. and McKenzie, D. P. 1978. Simple plate models of mantle convection. *J. Geophys*., **44**, 441–71.
Richter, F. M. and Parsons, B. 1975. On the interaction of two scales of convection in the mantle. *J. Geophys. Res*., **80**, 2529–41.
Ringwood, A. E. 1977. Composition of the Earth and implications for origin of the Earth. *Geochem. J*., **11**, 111–35.
1979. Composition and origin of the Earth. In M. W. McElhinny, ed., *The Earth: Its Origin, Structure and Evolution*. New York: Academic Press, pp. 1–58.
Ritsema, J. and van Heijst, H.-J. 2000. Seismic imaging of structural heterogeneity in Earth's mantle: evidence for large scale mantle flow. *Sci. Prog*., **83**, 243–59.
Ritzwoller, M. H. and Lively, E. M. 1995. Three-dimensional seismic models of the Earth's mantle. *Rev. Geophys*., **33**, 1–66.
Rochester, M. G. and Crossley, D. I. 1987. Earth's third ocean: the liquid core. *EOS Trans. Am. Geophys. Un*., **67**, 481.
Romanowicz, B. 1991. Seismic tomography of the Earth's mantle. *Ann. Rev. Earth Planet. Sci*., **19**, 77–99.
1995. A global tomographic model of shear attenuation in the upper mantle. *J. Geophys. Res*., **100**, 12 375–94.
Romanowicz, B. and Y. C. Gung. 2002. Superplumes from the core–mantle boundary to the lithosphere: implications for heat-flux. *Science*, **296**, 513–16.
Runcorn, S. K., Creer, K. M. and Jacobs, J. A., eds., 1982. The Earth's core: its structure, evolution and magnetic field. *Phil. Trans. Roy. Soc. Lond*. A, **306**, 1–289.
Schubert, G., Turcotte, D. L. and Olson, P. 2001. *Mantle Convection in the Earth and Planets*. Cambridge: Cambridge University Press.
Shearer, P. M. and Masters, G. T. 1992. Global mapping of topography on the 660-km discontinuity. *Nature*, **355**, 791–6.

Silver, P. G., Carlson, R. W. and Olson, P. 1988. Deep slabs, geochemical heterogeneity and the large-scale structure of mantle convection: investigation of an enduring paradox. *Ann. Rev. Earth Planet. Sci*., **16**, 477–541.

Sleep, N. H. and Fujita, K. 1997. *Principles of Geophysics*. Malden, Massachusetts: Blackwell Science.

Solheim, L. P. and Peltier, W. R. 1994. Avalanche effects in phase transition modulated thermal convection: a model of Earth's mantle. *J. Geophys. Res*., **99**, 6997–7018.

Song, X. and Richards, P. G. 1996. Seismological evidence for differential rotation of the Earth's inner core. *Nature*, **382**, 221–4.

Stacey, F. D. 1992. *Physics of the Earth*, 3rd edn. Kenmore, Queensland: Brookfield Press.

Stacey, F. D. and Loper, D. E. 1984. Thermal histories of the core and mantle. *Phys. Earth Planet. Interiors*, **36**, 99–115.

Tackley, P. J. 1995. Mantle dynamics: influence of the transition zone. *Rev. Geophys. Suppl*., 275–82.

2000. Mantle convection and plate tectonics: toward an integrated physical and chemical theory. *Science*, **288**, 2002–7.

Tackley, P. J., Stevenson, D. J., Glatzmeier, G. A. and Schubert, G. 1993. Effects of an endothermic phase transition at 670 km depth in a spherical model of convection in the Earth's mantle. *Nature*, **361**, 699–704.

1994. Effects of multiple phase transitions in a three-dimensional model of convection in Earth's mantle. *J. Geophys. Res*., **99**, 15 877–901.

van der Hilst, R. D. and Kárason, H. 1999. Compositional heterogeneity in the bottom 1000 kilometers of the earth's mantle: toward a hybrid convection model. *Science*, **283**, 1885–7.

van der Hilst, R. D., Widiyantoro, S. and Engdahl, E. R. 1997. Evidence for deep mantle ciculation from global tomography. *Nature*, **386**, 578–84.

Vidale, J. E. and Hedlin, M. A. H. 1998. Evidence for partial melt at the core–mantle boundary north of Tonga from the strong scattering of seismic waves. *Nature*, **391**, 682–4.

Vocadlo, L., Alfe, D., Gillan, M. J., Wood, I. G., Brodholt, J. P. and Price, G. D. 2003. Possible thermal and chemical stabilization of body-centred-cubic iron in the Earth's core. *Nature*, **424**, 536–9.

Whaler, K. and Holme, R. 1996. Catching the inner core in a spin. *Nature*, **382**, 205–6.

White, D. B. 1988. The planforms and onset of convection with a temperature dependent viscosity. *J. Fluid Mech*., **191**, 247–86.

White, R. S. and McKenzie, D. 1995. Mantle plumer and flood basalts. *J. Geophys. Res.*, **100**, 17 543–85.

Williams, Q., Jeanloz, R., Bass, J., Svendsen, B. and Ahrens, T. J. 1987. The melting curve of iron to 250 GPa: constraint on the temperature at the Earth's centre. *Science*, **236**, 181–2.

Woodhouse, I. H. and Dziewonski, A. M. 1984. Mapping the upper mantle: three dimensional modelling of Earth structure by inversion of seismic waveforms. *J. Geophys. Res*., **89**, 5953–86.

Wysession, M. E. 1996. Large-scale structure at the core–mantle boundary from core-diffracted waves. *Nature*, **382**, 244–8.

Wysession, M. E., Okal, E. A. and Bina, C. R. 1992. The structure of the core–mantle boundary from diffracted waves. *J. Geophys. Res*., **97**, 8749–64.

Yuen, D. A., Reuteler, D. M., Balachandar, S., Steinbach, V., Malevsky, A. V. and Smedsmo, J. L. 1994. Various influences on three-dimensional mantle convection with phase transitions. *Phys. Earth Planet. Int*., **86**, 185–203.

Zhong, S. and Gurnis, M. 1994. Role of plates and temperature-dependant viscosity in phase change dynamics. *J. Geophys. Res*., **99**, 15 903–17.

Zhou, H.-W. 1996. A high-resolution P wave model for the top 1200 km of the mantle. *J. Geophys. Res*., **101**, 27 791–810.

Chapter 9

The oceanic lithosphere: ridges, transforms, trenches and oceanic islands

9.1 Introduction

Three-fifths of the surface of the solid Earth is oceanic lithosphere, all of which has been formed during the last 160 Ma or so along the mid-ocean ridges. Understanding the structure of the oceanic lithosphere and the mid-ocean ridges is particularly important because it provides a key to understanding the mantle.

9.1.1 Beneath the waves

Bathymetric profiles across the oceans reveal the rugged nature of some of the seabed and something of the scale of its topography (Figs. 9.1 and 9.2). The deepest point on the surface of the Earth was discovered during the voyage of H. M. S. Challenger (1872–6). The bottom of the Challenger Deep in the Mariana Trench (western Pacific Ocean) is 10.92 km below sea level, and Mauna Kea on the island of Hawaii rises to 4.2 km above sea level from an ocean basin more than 5 km deep. Such features dwarf even Mount Everest (8.84 km above sea level). The average global land elevation is 0.84 km, whereas the average depth of the oceans is 3.8 km. Although the seabed is hidden from us by the oceans, the imprint of its shape is revealed by the sea surface and gravity (see Fig. 5.4 and Plates 7 and 8).

The seafloor can be classified into four main divisions: mid-ocean ridges, ocean basins, continental margins and oceanic trenches.

Mid-ocean ridges

The mid-ocean ridges are a chain of undersea mountains with a total length of over 60 000 km. The ridges typically rise to heights of more than 3 km above the ocean basins and are many hundreds of kilometres in width. The Mid-Atlantic Ridge was discovered while the first trans-Atlantic telegraph cable was being laid. The East Pacific Rise was discovered by H. M. S. Challenger in 1875. As was discussed in Chapter 2, the mid-ocean ridges mark the constructive boundaries of the plates. Hot material rises from the asthenosphere along the axes of the mid-ocean ridges and fills the space left by the separating plates; as the material cools, it becomes

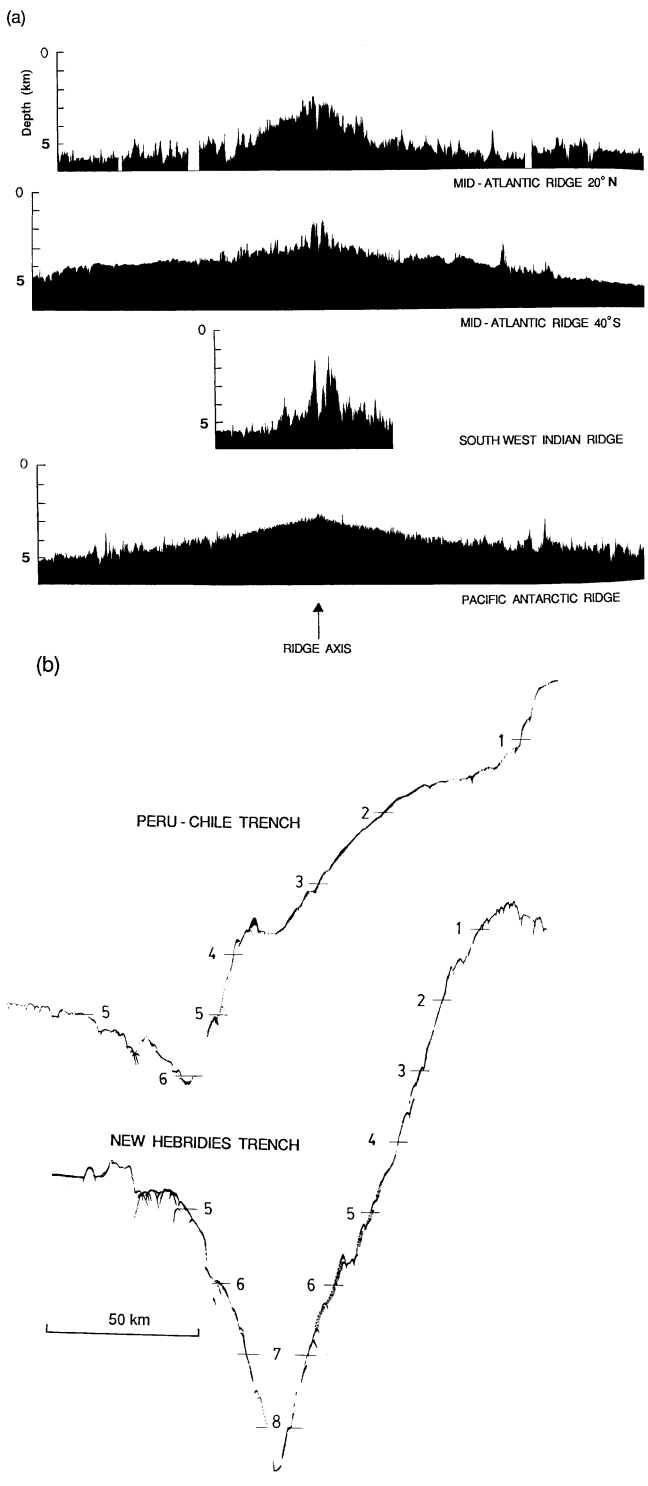


Figure 9.1.
(a) Topographic profiles across the axial regions of the Mid-Atlantic Ridge near 20°N, the Mid-Atlantic Ridge near 40°S, the South West Indian Ridge near 40°E and the Pacific–Antarctic Ridge near 55°S. Water depths are in kilometres. The vertical exaggeration is approximately 60 : 1. (After Heezen (1962).) (b) Topographic echo-sounder cross sections across the Peru–Chile Trench at 12°S, 78°W and the New Hebrides Trench at 12°S, 166°E, the surface expressions of oceanic–continental and oceanic–oceanic plate collisions, respectively. Water depths are in kilometres. The vertical exaggeration is approximately 22 : 1. (From Menard (1964).)

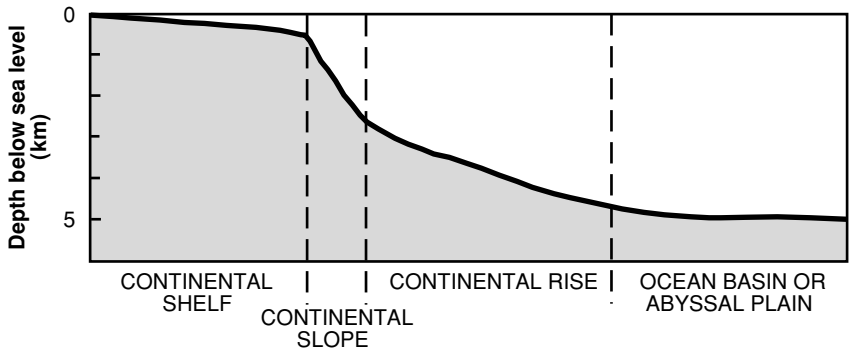


Figure 9.2. The bathymetry of the regions of a passive continental margin such as the margins of the Atlantic Ocean. Slopes are approximate; the vertical exaggeration is 100–200 : 1.

part of the plates. For this reason, mid-ocean ridges are often called spreading centres. The examples in Chapter 2 show that present-day spreading rates of the mid-ocean ridges vary between approximately 0.5 and 10 cm yr^{-1}. Spreading rates are generally quoted as half the plate-separation rate. For example, the North American and Eurasian plates are separating at approximately 2 cm yr^{-1}, so the Mid-Atlantic Ridge is said to be spreading at a rate of 1 cm yr^{-1}. The fastest spreading ridge today is that portion of the East Pacific Rise between the Nazca and Pacific plates; its rate is almost 10 cm yr^{-1}. It is often helpful to consider ridges in fast-, intermediate- and slow-spreading categories (see Table 9.3 later). The East Pacific Rise (half-rate 5–10 cm yr^{-1}) is a fast-spreading ridge, the Juan de Fuca Ridge (half-rate 2.5–3.3 cm yr^{-1}) is an intermediate-spreading ridge, the Mid-Atlantic Ridge (half-rate 1.2–2.5 cm yr^{-1}) is a slow-spreading ridge and the very slow-spreading ridges are the Southwest Indian Ridge and the Arctic ridges (half-rates <1 cm yr^{-1}).

Bathymetric profiles across the Mid-Atlantic Ridge, Southwest Indian Ridge and Pacific–Antarctic Ridge are shown in Fig. 9.1. The relation between the mean bathymetric depth and its age is discussed in detail in Section 7.5 and presented in Table 7.5. Depth increases linearly with the square root of age over young oceanic crust. This holds for all ridges irrespective of their spreading rate. The East Pacific Rise is wider than the more slowly spreading Mid-Atlantic Ridge because the width of any ridge is proportional to its spreading rate. The detailed cross sections of the axial regions of these three ridges reveal more differences among them. These are discussed in Section 9.4.1.

Ocean basins

Sedimentation occurs throughout the oceans. As a result of this and the subsidence of the seabed (due to cooling and contraction of the plates), over time the rugged, faulted topography of the mid-ocean ridges is buried under sediments. The almost flat regions of the seabed, thousands of kilometres in width and some 5–6 km below sea level, are often called ocean basins or abyssal plains. Isolated volcanic islands occur in most of the oceans, often in chains (e.g., the

Emperor Seamount chain and the Hawaiian Islands in the northern Pacific Ocean, as shown in Fig. 2.19). Such island chains apparently form when a plate passes over a 'hotspot', a localized region where magma is rising from deep in the mantle. The depth from which the magma comes can be roughly estimated from the chemistry of the lava. In addition, the islands can be dated and the history of the passage of the plate over the magma source determined. This information was used in the determination of the plate motions in Chapter 2. The aseismic ridges (such as the Walvis Ridge and the Rio Grande Rise in the South Atlantic Ocean, and the Iceland–Faeroe Ridge in the North Atlantic Ocean) are submarine volcanic mountain chains, typically elevated some 2–4 km above the seabed which formed as the plate passed over a hot region in the mantle. They are not tectonically active (hence their name).

Continental margins

The continental margins, as their name suggests, mark the transition between the continent and the ocean floor. At a passive margin (Fig. 9.2), such as occurs off the east coast of the Americas and the matching west coast of Europe and Africa, the first structure seawards of the land is the gently sloping continental shelf, which is merely continent covered by shallow water. At the outer edge of the shelf, the gradient abruptly increases and the seabed deepens rapidly; this is the continental slope. At the base of the slope, the gradient is again much less; this region is the continental rise. The transition from continental crust to oceanic crust occurs beneath the continental slope.

An *active margin*, such as the west coast of North and South America, is so called because of the igneous and tectonic activity occurring at the plate boundary. There the continental shelf is often, but not always, very narrow. Where the plate boundary is a transform fault, the seabed characteristically drops rapidly from the shelf to oceanic depths. Where the plate boundary is a subduction zone, there is usually a trench, typically many kilometres deep.

Oceanic trenches

The oceanic trench marks the surface location of a subduction zone, at which one plate is overriding another; for example, the continental South American plate is overriding the oceanic Nazca plate along the west coast of South America. Here the Peru–Chile Trench extends to depths of 8 km and is considerably less than 200 km in width (see Fig. 9.1). Many of the trenches around the western margin of the Pacific plate occur where an oceanic plate is overriding the oceanic Pacific plate. At such destructive plate boundaries, an island arc runs parallel to the trench, and frequently seafloor spreading occurs on the consuming plate behind the island arc, thus forming a *back-arc*, or *marginal basin*. The trenches of the western Pacific are frequently 10 km or more deep.

Classification of igneous rocks

The classification of igneous rocks is a vexed business that has kept igneous petrologists amused, annoyed and employed for the best part of a century. The discussion that follows is not rigorous but merely attempts to outline the arguments for those who are not igneous petrologists.

The problem is that an igneous rock can be classified in many ways, each being useful in some circumstances. Classification began in the obvious way when geologists simply looked at igneous rocks and the minerals they contain. One obvious difference is between coarse-grained plutonic rocks and fine-grained or partly glassy volcanic rocks. Plutonic rocks are derived from magmas that cooled slowly at depth in the interior of the Earth and therefore grew large crystals. Volcanic lavas that erupted at the Earth's surface cooled quickly and thus the rocks grew small crystals. Indeed, if a magma cools very rapidly, much or all of the rock may be glassy. Of course, to confuse the issue, a magma may carry to the surface some larger crystals that formed at depth (called *phenocrysts*).

In general, the colour of igneous rocks is a good guide to their chemical composition, and it usually reflects the mineralogy of the rock (Table 9.1). Light-coloured rocks, rich in silica and alumina, are *felsic* rocks. These include granite (a plutonic rock) and rhyolite (the volcanic equivalent of granite). The detailed nomenclature of the granitoid rocks is very complex: *granites* and *rhyolites* are rich in quartz and alkali feldspar minerals; granodiorites (plutonic) and their volcanic equivalents, dacites, are rich in plagioclase feldspar minerals and quartz but have less alkali feldspar; tonalites are plutonic rocks that contain quartz and plagioclase feldspar but very little alkali feldspar.

Igneous rocks that contain moderate quantities of dark minerals, such as pyroxene and hornblende, are termed *intermediate*. An example is diorite, a plutonic rock that contains plagioclase feldspar, is moderately rich in calcium, and has some dark minerals such as hornblende, biotite and pyroxene. Its volcanic equivalent is andesite, a rock that is typical of many volcanoes above subduction zones.

Igneous rocks that are rich in dark minerals are called *mafic*, shorthand for rich in magnesium and iron (Fe). These rocks include gabbro, a plutonic rock containing calcic plagioclase, pyroxene and often olivine or hornblende; its volcanic equivalent is basalt.

The *ultramafic* igneous rocks (Fig. 8.10) are almost exclusively made of dark minerals. They include peridotites, which are plutonic rocks rich in olivine, with some pyroxene or hornblende. Dunite is a peridotite that is 90% or more olivine. Harzburgite is mostly olivine, plus some orthopyroxene and perhaps a little clinopyroxene; lherzolite contains olivine, orthopyroxene and clinopyroxene. The mantle of the Earth is ultramafic. Ultramafic rocks have rarely erupted as lavas in the relatively recent geological past, but they were more common in Archaean strata. Ultramafic lavas are called komatiites; fresh komatiites have small crystals of olivine and pyroxene set in a glassy matrix.

Table 9.1 *Classification of igneous rocks*

Colour (% dark minerals)	Type	Example
Less than 40	Felsic	Granite
40–70	Intermediate	Diorite
70–90	Mafic	Gabbro
Over 90	Ultramafic	Peridotite

Silica (% SiO_2)	Type	Example
Over 66	Acid	Rhyolite
52–66	Intermediate	Andesite
45–52	Basic	Basalt
Less than 45	Ultrabasic	Komatiite

Average chemical composition (%)

	Granite	Grano-diorite	Diorite	Gabbro	Peridotite	Andesite	Basalt	Komatiite
SiO_2	70.4	66.9	51.9	48.4	43.5	54.9	50.8	44.9
TiO_2	0.4	0.6	1.5	1.3	0.8	1.0	2.0	0.3
Al_2O_3	14.4	15.7	16.4	16.8	4.0	17.7	14.1	3.1
Fe_2O_3	1.0	1.3	2.7	2.6	2.5	2.4	2.9	2.3
FeO	1.9	2.6	7.0	7.9	9.8	5.6	9.1	11.5
MnO	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2
MgO	0.8	1.6	6.1	8.1	34.0	4.9	6.3	33.0
CaO	2.0	3.6	8.4	11.1	3.5	7.9	10.4	3.8
Na_2O	3.2	3.8	3.4	2.3	0.6	3.7	2.2	0.01
K_2O	5.0	3.1	1.3	0.6	0.3	1.1	0.8	0.01
P_2O_5	0.2	0.2	0.4	0.2	0.1	0.3	0.2	0.03

Note: These are families of rocks and exhibit much variation around the typical; thus, an average is somewhat meaningless. The table presents only the rough chemical content to be expected.
Source: Chemical compositions are from Nockolds *et al.* (1978) and Smith and Erlank (1982).

Another good way to classify rocks is by dividing them according to their chemical content. Rocks rich in SiO_2 such as rhyolite are called acid; intermediate rocks have rather less SiO_2; basic rocks are about 50% SiO_2 or less and include basalt; ultrabasic rocks are very poor in SiO_2.

Chemistry can be of further help. Not content with analysing fine-grained and glassy rocks, petrologists considered what a lava would have turned into if it had cooled slowly, crystallizing completely. Various schemes were devised to estimate the minerals that would have formed if a lava of given composition had crystallized. Of course, these schemes are artificial because so many variables are involved, but eventually standard ways of recalculating the chemical analysis were accepted. The standard, or *normative*, composition can then be used to classify the rock.

Another system of classification, which dates back to the 1930s but has introduced names that are deeply embedded in the literature, is to study a suite of rocks for their CaO and ($Na_2O + K_2O$) contents. Rock suites that are rich in the alkalis are called alkaline; rock suites that are rich in CaO are calcic. Andesite volcanoes typically have calc-alkaline trends.

For geophysicists, the difference between basalt and andesite is especially important. Basalts must have less than 52% SiO_2, whereas andesites have more. Basalts have over 40% by weight of dark minerals, typically pyroxene, whereas andesites have less than 40% of dark minerals and often (though not always) contain hornblende. The plagioclase in basalt is more calcium-rich than that in andesite. In thin section, basalt and andesite can be seen to have different microscopic textures. The rock types between basalt and andesite are called andesitic basalts and basaltic andesite.

Basalts are, in fact, a family of lavas, divided according to their content of normative minerals into tholeiites (which have slightly more silica relative to some other components) and *alkali-olivine basalts* (which are relatively more alkaline), with a transitional group known as *olivine tholeiites*. Most ocean-floor basalts and most but not all basalts in Hawaii are olivine tholeiites. Alkali basalts occur in zones of continental rifting. For more detail, the reader should consult any textbook on igneous petrology.

9.2 The oceanic lithosphere

9.2.1 Oceanic crust

Worldwide gravity surveys indicate that the oceanic regions are in approximate isostatic equilibrium with the continents; that is, the pressure at an arbitrary depth in the mantle is the same beneath continents and oceans. This means that, at arbitrary depth in the mantle, a column of continental crust and underlying mantle and a column of oceanic crust and its underlying mantle have the same mass. This fact enables us to make a simple estimate of the thickness of the oceanic crust (see Section 5.5 for the method). If we assume Airy-type compensation, densities of sea water, crust and mantle of 1.03×10^3, 2.9×10^3 and $3.3 \times 10^3\,\mathrm{kg\,m^{-3}}$, respectively, and an average ocean-basin depth of 5 km, then a typical 35-km-thick continental crust would be in isostatic equilibrium with an oceanic crust 6.6 km in thickness. This rough calculation tells us the important fact that the oceanic crust is approximately one-fifth the thickness of the continental crust.

The details of the seismic structure of the oceanic crust have been determined by using seismic-refraction and -reflection profiling and wide-angle-reflection techniques. In the absence of direct sampling of the crust, its composition must be estimated from measurements of its physical properties (e.g., seismic velocity and density), which vary with lithology. These estimates are frequently ambiguous.

The most direct way to determine the composition of the oceanic crust is to collect rock samples from each of the oceanic plates. Dredging samples from the seabed is not particularly difficult or expensive, but it is often frustrating: imagine

trying to manoeuvre a large bucket, hanging on 5 km of wire, over a scarp slope, which you can see only on the ship's echo sounder, and then attempting to collect rocks from the debris at the scarp base. Dredging is also not particularly representative: a sample might not be typical. A much more precise way of sampling the seabed is to use a submersible, a minisubmarine capable of descending to great depths; there the geologist can sample the exact outcrop and rock type wanted. Submersibles (both manned and remotely operated) have enabled scientists to make detailed studies of small areas of the seabed, particularly the axial zones of the mid-ocean ridges. Such operations are, however, extremely expensive and still sample only the surface. Sampling the rock outcropping at the seabed does not tell us what rocks make up the lower oceanic crust (even if the fault scarps are such that deeper rocks could be exposed at their bases). However, it does enable us to make informed guesses; for instance, since seabed samples include basalts, gabbros, serpentinites and recent sediments, one would not guess at a deep crust made up of granite (it took scientists a while to realize that granite samples dredged from the seabed were not representative of the ocean crust but had been dropped by icebergs, these samples being ice-rafted granite). Another way to obtain samples is to drill. Drilling into the oceanic crust is an expensive and difficult operation compared with drilling on land. Not only is the rock hard and frequently fractured, but also there are many kilometres of sea water between the drilling ship and the top of the drillhole (Fig. 9.3). Drilling into the oceanic crust is an international cooperative venture overseen by the Joint Oceanographic Institutions for Deep Earth Sampling (JOIDES). Drilling, which started in 1968 as the Deep Sea Drilling Project (DSDP), in 1985 entered a new phase as the Ocean Drilling Program (ODP) and then in 2003 became the Integrated Ocean Drilling Program (IODP), has tremendously advanced our knowledge of the geological and geophysical structure of the uppermost crust. While 1.5 km of the lower crust has been drilled on Atlantis Bank (a 5-km-high bank adjacent to the Atlantis II transform fault on the ultra-slowly spreading Southwest Indian Ridge), drilling into the upper mantle remains a project for the future.

Most of the seabed is sedimented. Figure 9.4 shows the sediment thickness over the western North Atlantic. Thicknesses gradually increase from zero at the active ridge axis to 0.5–1 km in the ocean basins. The pelagic sedimentation rate over mid-ocean ridges is very low, millimetres per thousand years (m per Ma). The greatest accumulation of sediments occurs beneath the continental margins; over 10 km of sediment is not uncommon, but thicknesses rarely reach 15 km. These sediment accumulations, a consequence of the rifting apart of the old continent, are discussed in Section 10.3. For detail of shallow sediments on continental margins see Figs. 4.43 and 4.44.

Researchers began widespread oceanic seismic-refraction experiments in the 1950s and by the 1960s had made enough determinations of crustal seismic velocities over the oceans for it to be apparent that, unlike the continents, the oceanic crustal structure varied little. Table 9.2 gives the standard seismic structure of

Table 9.2 *Layering of the oceanic crust beneath the ocean basins*

Layer	Thickness (km)	Range of seismic velocity (km s^{-1})
Sea water	4.5	1.5
Layer 1 (sediment, variable)	About 0.5	About 2
Layer 2 (volcanic layer)	2.11 ± 0.55	2.5–6.6
Layer 3 (oceanic layer)	4.97 ± 0.90	6.6–7.6
Layer 4 (upper mantle)		7.9–8.1
Total igneous crust	7.08 ± 0.78	

Note: This layering is the mean seismic structure of 'normal' oceanic crust.
Source: After White *et al.* (1992).

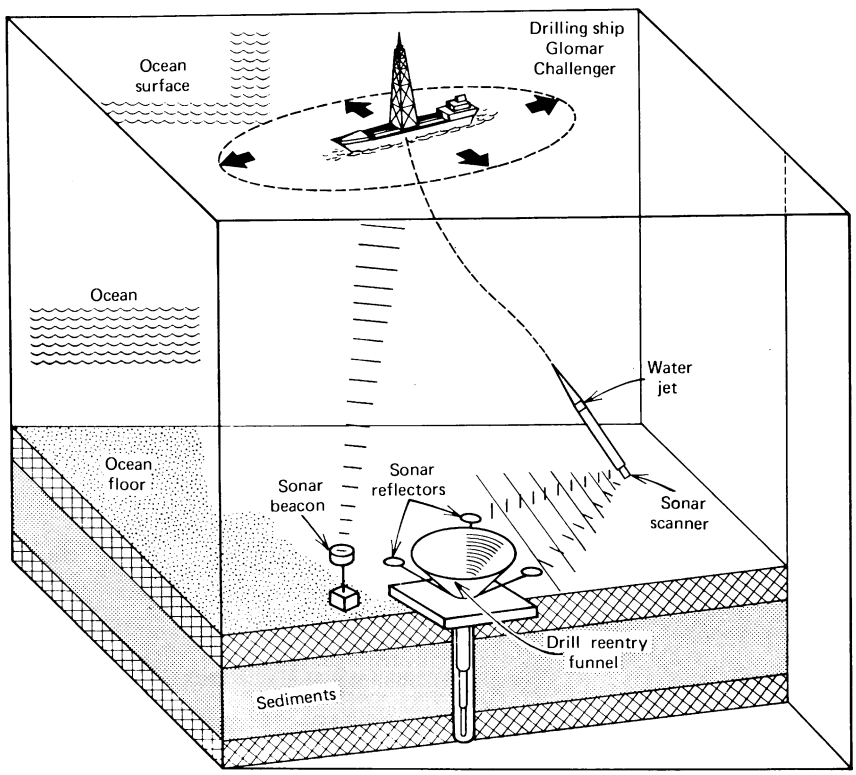


Figure 9.3. Drilling into the oceanic crust. Finding the hole for re-entry, a long way down. Development of deep drilling techniques led to much improved knowledge of the oceanic crust and palaeoclimate. (From Wyllie (1976).)

mature oceanic crust beneath the ocean basins as determined from first-P-wave-arrival travel times. However, subsequent improvements in the techniques for the inversion of seismic wide-angle reflection and refraction data (including computation of synthetic seismograms; see Section 4.3) have provided a much improved resolution of seismic structure of the oceanic crust. The original four layers of

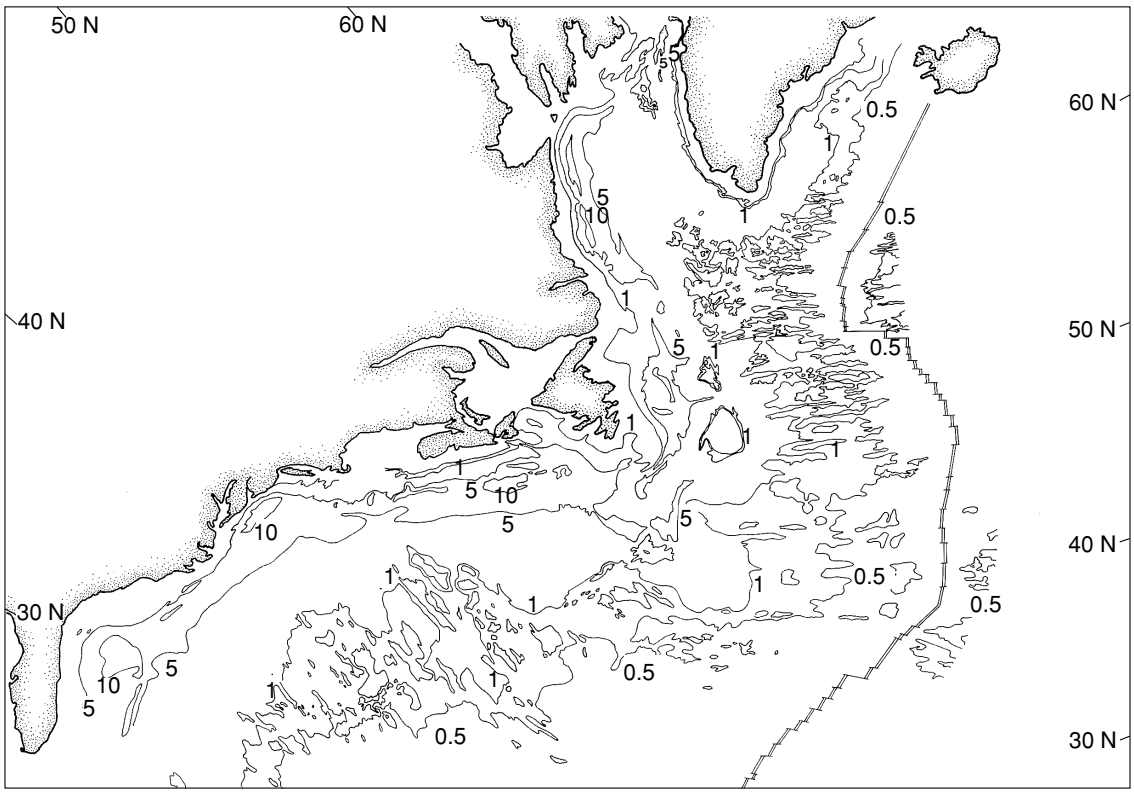


Figure 9.4. Sediment thickness over the western North Atlantic (contours are 0.5, 1, 5 and 10 km). (After Tucholke (1986).) See http://www.ngdc.noaa.gov/mgg/image/sedthick.jpg for a global marine sediment thickness map.

constant velocity[1] have given way to a discontinuous series of velocity gradients in which the velocity increases with depth (Table 9.2). Figure 9.5(a) shows velocity–depth profiles for seismic-refraction lines in the Atlantic and Pacific that were constrained by synthetic seismogram modelling. It is clear that, although there are local differences, the general variation of seismic velocity with depth is fairly constant within each age zone. It should be remembered when looking at velocity–depth profiles such as these that seismic-refraction experiments generally cannot resolve structure much less than one wavelength in thickness. (For layer 2, with velocity of 5 km s^{-1} and a 7-Hz signal, a wavelength is 0.7 km.) Therefore, much of the fine detail and staircase-like appearance of these profiles is a representation of velocity gradients.

[1] This layering is sometimes referred to as 'Raitt–Hill layering' after the scientists who first established that the uniform structure of the oceanic crust could be modelled as four constant-velocity layers (Raitt 1963; Hill 1957).

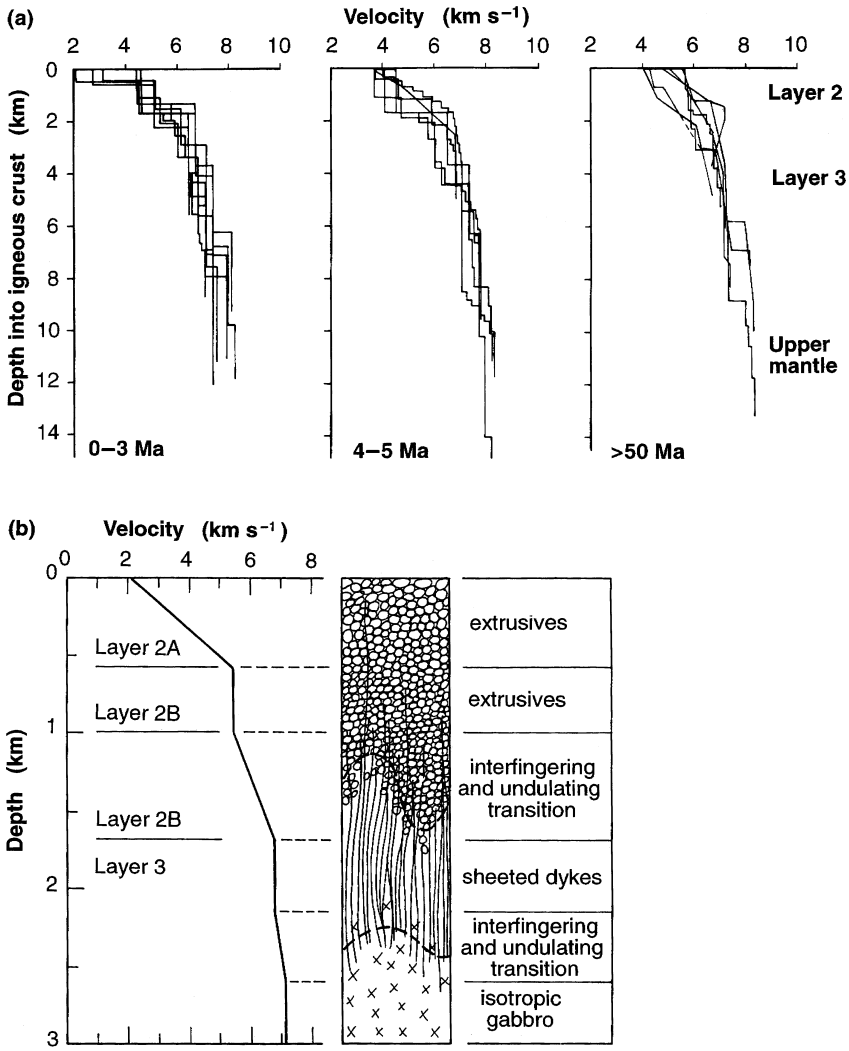


Figure 9.5. (a) Velocity–depth models from the North Atlantic, grouped according to the age of the lithosphere. All these models were determined from seismic-refraction data by calculating synthetic seismograms and adjusting the model until the travel times and the amplitudes agreed well with the data. (From White (1984).) (b) A velocity–depth model obtained from the East Pacific Rise correlated with the DSDP hole 504B drilled into 6.6-Ma oceanic crust formed at the Costa Rica Rift with a half-spreading rate of 3.0–3.4 cm/yr^{-1}. Hole 504B at 2200 m is the deepest hole drilled into oceanic crust. (From Bratt and Purdy (1984).)

Layer 1, the sedimentary layer, thickens with distance from the ridge axis. The P-wave velocity at the top of the sediments is generally close to 1.5 km s^{-1} (the velocity of sound in sea water) but increases with depth as the sediments consolidate. (This layer is not shown in Fig. 9.5.)

At the top of *layer 2*, the *volcanic layer*, there is a sudden increase in P-wave velocity to approximately 4 km s^{-1}. Reflection profiling has shown that the top of layer 2, the *oceanic basement*, is very rough for crust that was formed at a slow-spreading ridge. It was proposed in the 1970s that layer 2 should be subdivided into three layers, 2A, 2B and 2C, with P-wave velocities of 3.5, 4.8–5.5 and 5.8–6.2 km s^{-1}, respectively, and this terminology is often used in the literature. However, layer 2 is probably best described as a region of the oceanic crust in

which seismic velocity increases rapidly with depth (gradients of 1–2 km s^{-1} per kilometre of depth). In very young crust, the P-wave velocity in the top few hundred metres of layer 2 may be less than 2.5 km s^{-1}. Drilling into the top of layer 2 has shown that it is made up of sediments, extrusive basaltic pillow lavas and lava debris in varying degrees of alteration and metamorphism. Deeper drilling has found more consolidated basalts. Towards the base of this basaltic layer are sheeted dykes, which cause the P-wave velocity to increase further. Together, this variability in composition and gradation from sediments to basalt to dykes appears to account for the rapid and variable increase in seismic velocity with depth (Fig. 9.5(b)).

Layer 3, sometimes called the *oceanic layer*, is thicker and much more uniform than layer 2. Typical P-wave velocities are 6.5–7.2 km s^{-1}, with gradients of about 0.1–0.2 km s^{-1} per kilometre depth. Layer 3 is generally presumed to be gabbroic in composition. Some seismic experiments have shown that the basal part of this layer (sometimes termed layer 3B) has a higher P-wave velocity (7.2–7.7 km s^{-1}), indicating perhaps a change to more cumulate-rich gabbros at the base of the crust. A few experimenters have suggested that there is a low-velocity P-wave zone at the base of layer 3, but this is not generally thought to be a universal feature of the oceanic layer. In 1962, Hess proposed that layer 3 is partially serpentinized peridotite formed as a result of hydration of the upper mantle. His proposal was largely discounted in the 1970s, but it is possible that, in regions where the oceanic crust is significantly thinner than normal (beneath fracture zones; crust formed at very-slow-spreading ridges), significant serpentinization of the upper mantle may well occur. Hole 735B penetrated 1.5 km into 11-Ma-old layer 3 on Atlantis Bank (Southwest Indian Ridge). The rocks are strongly heterogeneous olivine gabbros – no cumulates were intersected. This is consistent with a crust constructed from many small intrusions, each crystallizing *in situ*, and there being no steady-state magma chamber.

Crust formed in back-arc basins is very similar to normal oceanic crust. In the Lau back-arc basin layer 2, and hence the crust, is 1–1.5 km thicker than normal. This difference is presumably due to a combination of the physical and chemical influences of the adjacent Tonga–Kermadec volcanc arc.

In the same way as that in which P-waves are used to model the seismic structure of the crust, we can use S-waves to make an S-wave velocity model. Refraction data of sufficient quality for making detailed S-wave models are very rare, so, unfortunately, crustal S-wave velocity models are uncommon. Figure 9.6 shows a detailed model of the P- and S-wave velocity structures of the oceanic crust east of Guadalupe Island off the west coast of Baja California. The seafloor at this location is about 15 Ma old and is part of what is left of the Farallon Plate (see Section 3.3.4). Of particular interest here are the low velocities in the shear-wave model for layer 3. The evidence for this low-velocity zone is very good. Synthetic S-wave seismograms, computed for the best P-wave velocity model assuming that Poisson's ratio s is 0.28 throughout the crust, do not match

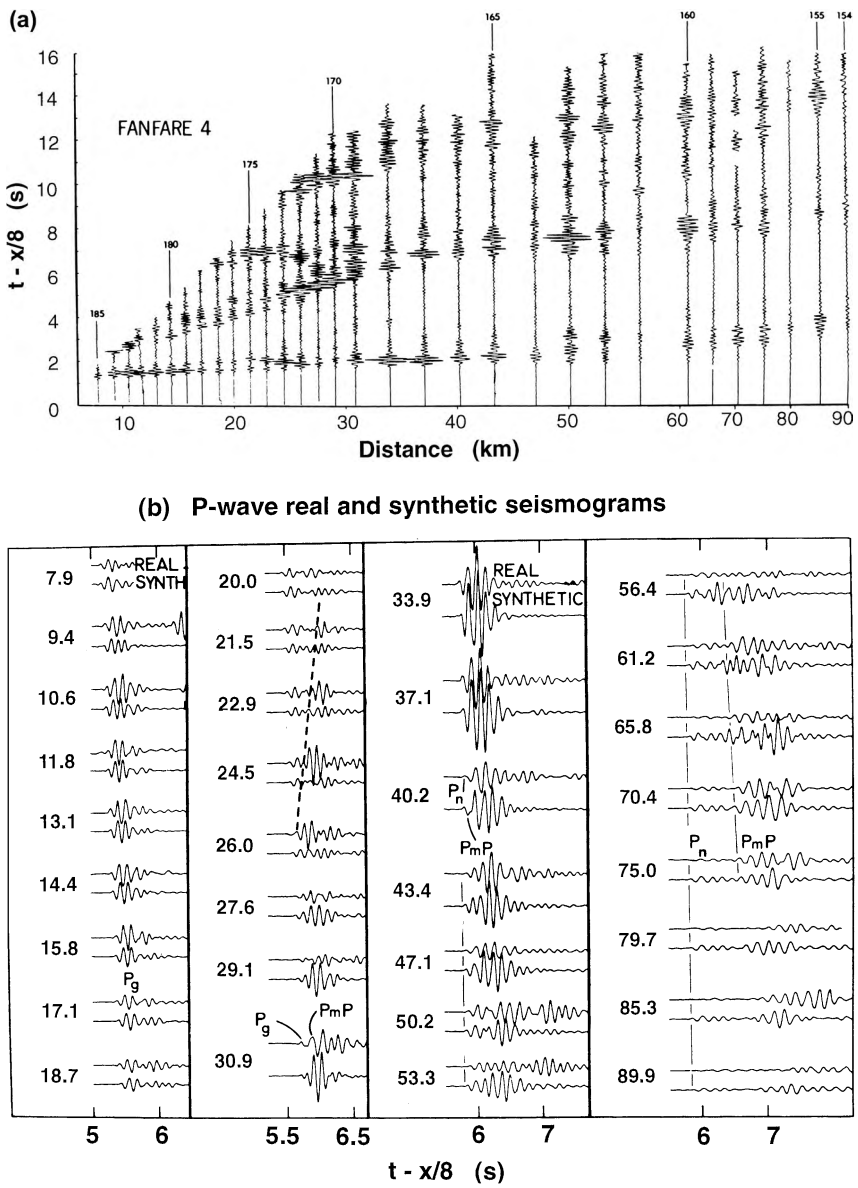
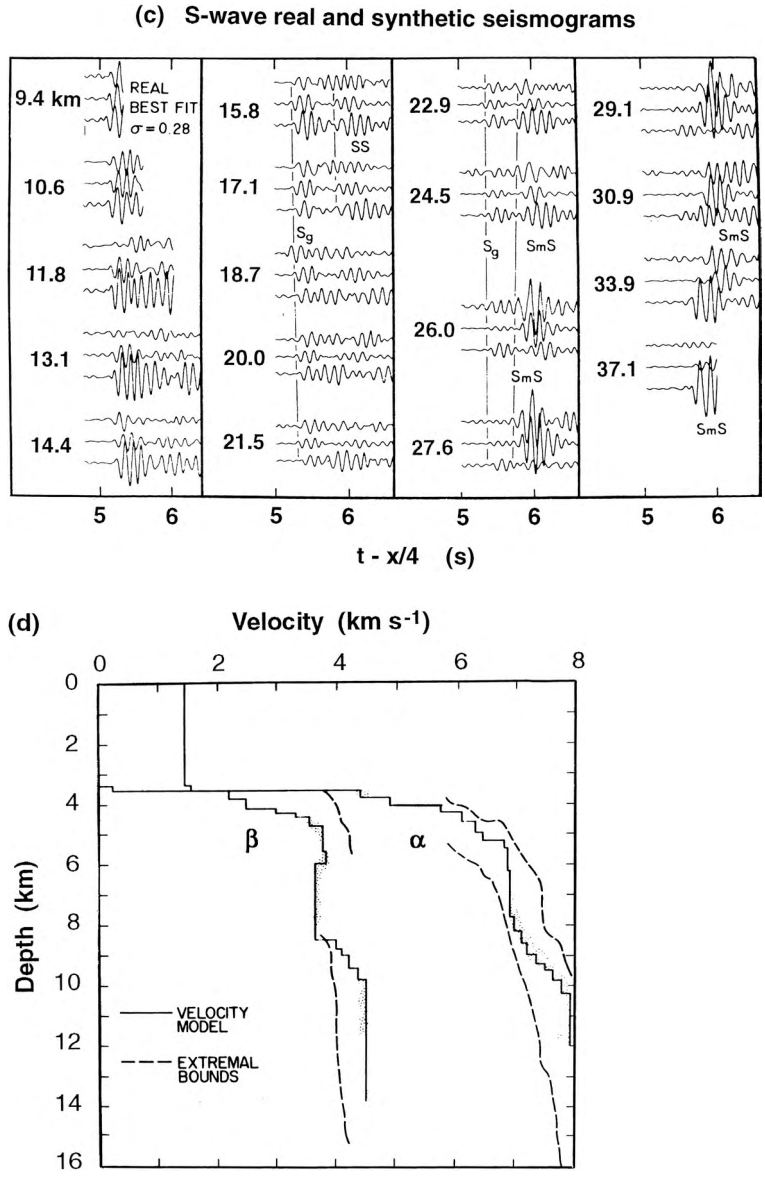


Figure 9.6. Real and synthetic seismograms for a refraction line shot between Guadalupe Island and Baja California on 15-Ma-old oceanic crust. (a) Real data: a reduced record section. (b) P-wave seismogram pairs. Synthetic seismograms computed from the model in (d). (c) Real and synthetic S-wave seismograms. Best-fitting synthetic seismograms were computed from the S-wave model in (d); synthetic seismograms marked $\sigma = 0.28$ (the third in each set) were computed from the P-wave velocity model in (d) with Poisson's ratio $\sigma = 0.28$. (d) P-wave and S-wave velocity models. All velocity–depth models satisfying the travel-time data lie within the extremal bounds as determined from the travel-time inversion. (From Spudich and Orcutt (1980).)

the shear reflection from the mantle (S_mS): The critical distance for S_mS is too large, and the arrival time is too early. The critical distance could be decreased by elevating the S-wave Moho, but this would cause S_mS to arrive earlier still. A solution can be obtained, however, by introducing a low-velocity S-wave zone into the lower crust; this reduces the S_mS critical distance and delays the arrival times without affecting the P-wave structure at all (see Fig. 9.6). It is suggested that these well-defined low S-wave velocities could indicate the relative abundance

Figure 9.6. (*cont.*)



of hornblende in comparison with augite and olivine in the gabbros which make up layer 3. This interpretation is not the only solution: the low velocities could also be produced by fluid-filled pores.

In a major two-ship multichannel seismic experiment, the North Atlantic Transect, some 4000 km of reflection data were collected between the east coast of the U.S.A. at 32°N and the Mid-Atlantic Ridge at 23°N. Along much of the

transect a sharp, apparently continuous reflection from the Moho was observed at about 2.5 s beneath the top of the oceanic basement. Eleven expanding-spread (wide-angle) profiles were also recorded along this transect; Fig. 9.7(a) shows the record section for one of these profiles. Modelling of these travel times and amplitudes yielded a detailed structure of the oceanic crust at this location, which is some 7.5 km in thickness. Study of the amplitudes of those seismograms shows that there may be a low-velocity zone in the lower crust, but it is not necessarily needed to satisfy the data. A reduced record section from a seismic refraction line shot over 140-Ma-old oceanic crust in the western central Atlantic Ocean is shown in Fig. 9.7(b). The main features of these typical seismograms are the large-amplitude mantle reflections P_mP, the weak mantle refraction P_n and the clear shear waves.

Thus, several lithologies have been proposed for layer 3, and we will not be able to determine definitively the composition of layer 3 and its subdivision until a hole (better yet, several holes) is drilled completely through the oceanic crust.

Ophiolites (a sequence of rocks characterized by basal ultramafics overlain by extensive thicknesses of basaltic and gabbroic material: gabbro, dykes, lavas and deep-sea sediments) are often regarded as examples of oceanic crust. However, because they are now tectonically emplaced on land, they are atypical and might not represent normal oceanic crust. Ophiolites are probably samples of young oceanic crust produced in back-arc basins, or fore-arc basins, associated with subduction zones (Section 9.6). It is more likely that such crust, rather than typical oceanic crust from a section of old cold lithosphere, could become emplaced on land. Figure 9.8 shows the principal ophiolite belts of the world. One well-studied ophiolite is the Semail ophiolite in the Sultanate of Oman, which is believed to have been formed some 95 Ma ago at a spreading centre in the Tethyan Sea and later emplaced on the Arabian Plate as Tethys closed when Africa and Arabia moved northwards and collided with Asia (see Section 3.3.3). The results of the extensive geological and seismic-velocity studies of this classical ophiolite are summarized in Fig. 9.9. It is apparent that the seismic-velocity structure is similar to that of the oceanic crust and upper mantle, with gradients throughout layer 2, a rapid increase in velocity at the layer-2–layer-3 boundary and a relatively high-velocity layer 3B at the base of the crust. The transition from crust to mantle is sharp. Similar results have been determined for the Blow Me Down Massif of the Bay of Islands ophiolite in Newfoundland, Canada. However, it is unwise to make a direct comparison between the seismic velocities measured on ophiolite rock samples at ultrasonic frequencies in a laboratory and the seismic velocities determined from reflection and refraction experiments using very-low-frequency signals; the frequency scale and the length scale of the velocity determinations differ.

The oceanic crust has more or less the same thickness and velocity structure throughout the world's oceans, on all plates. This observation can be explained

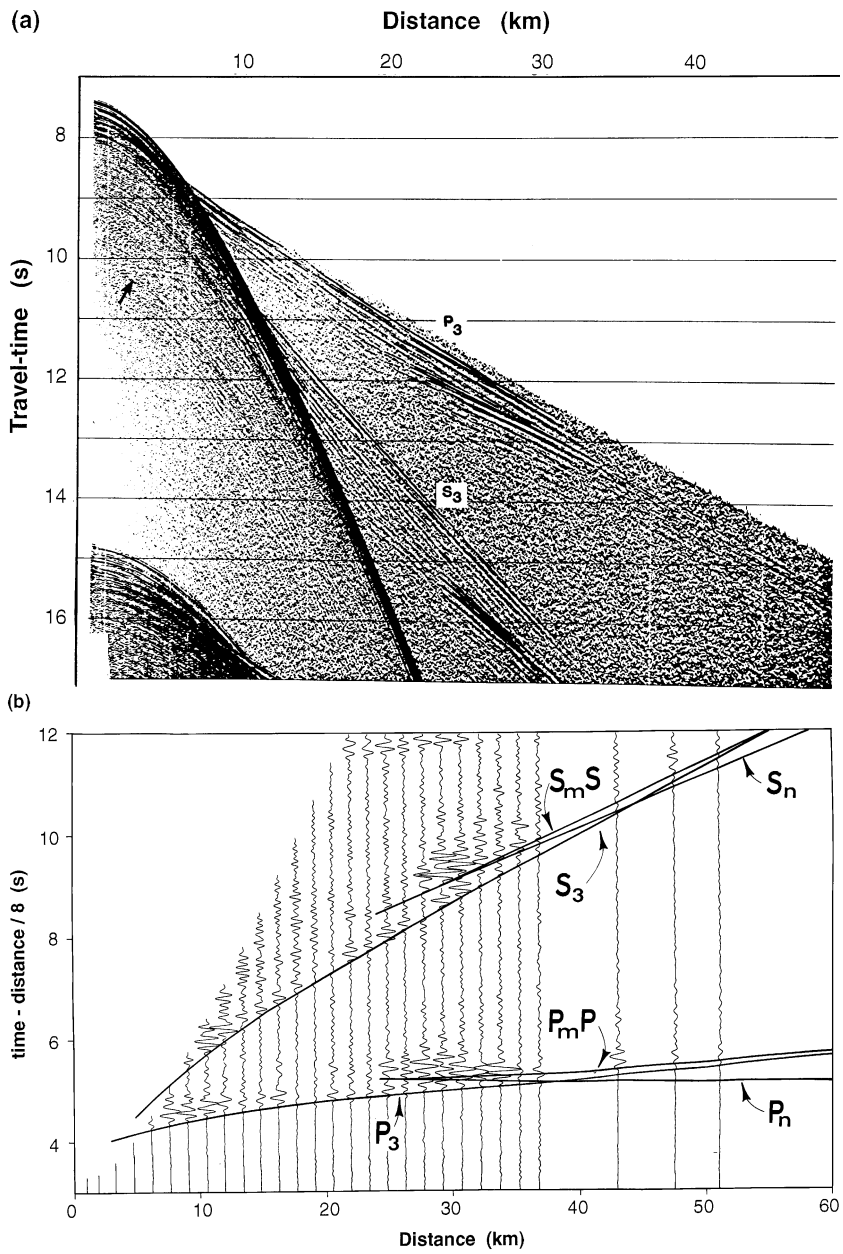


Figure 9.7. (a) The record section of an expanding-spread (wide-angle) profile shot over 118-Ma-old oceanic crust southwest of Bermuda on the North American Plate. The records have been band-pass filtered at 6–30 Hz and plotted with a gain (amplification) that increases linearly with distance and as the square of the travel time. The P-wave arrival, P_3, is a lower-crustal refraction. The S_3 arrival is the lower-crustal S-wave. The strong secondary P-wave arrival seen between 20 and 35 km is a refraction from the upper mantle. Arrow; near-normal-incidence reflection from the Moho. The large-amplitude hyperbola, which is the first arrival at

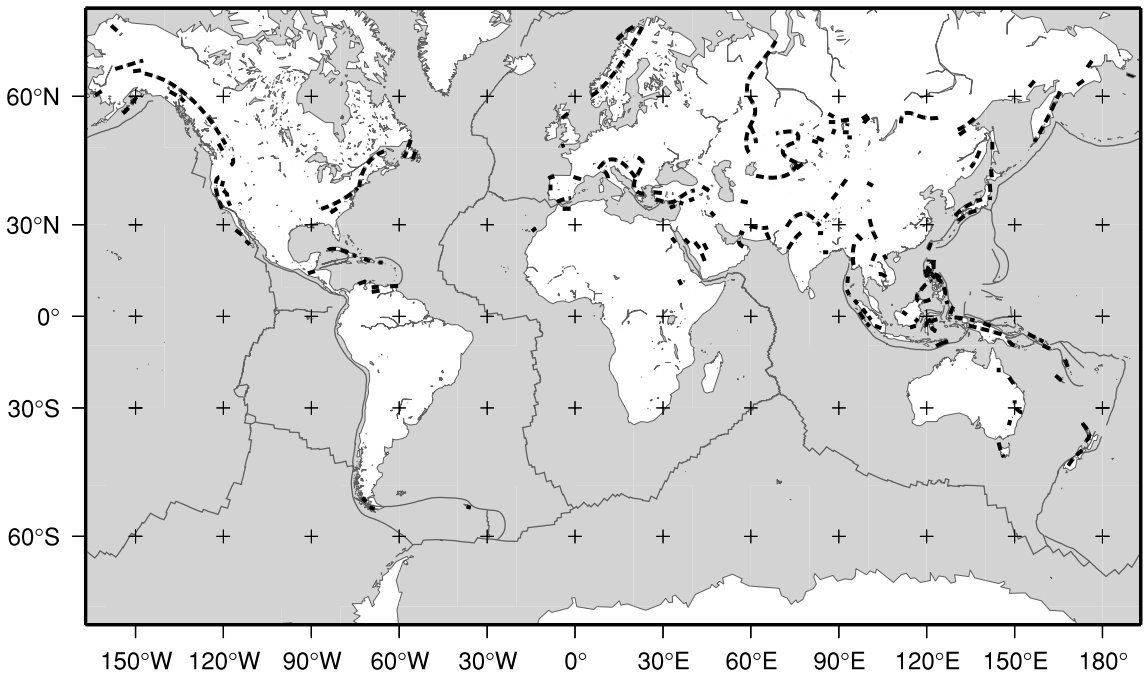


Figure 9.8. Principal ophiolite belts of the world. (Based on Coleman (1977).)

by the fact that the depth and temperature of melting in fast- and slow-spreading ridges are similar, which means that the extent of melting is similar (Section 9.3.2). It has been suggested that the oceanic crust gradually thickens slightly with age, but there are insufficient results from seismic experiments to confirm this. The Pacific oceanic crust is similar to, but rather more uniform than, the Atlantic crust, due to its having a faster and more constant magma-supply rate. The very or ultra-slow-spreading mid-ocean ridges (e.g., the Arctic Gakkel Ridge, half-rates 0.3–0.75 cm yr^{-1}; and the Southwest Indian Ridge, half-rate 0.57–0.85 cm yr^{-1}) produce much thinner crust than normal as a result of there being less melting in the mantle. To understand the structure and composition of the oceanic crust, it is necessary to determine the structure of the mid-ocean ridges and the processes occurring there where the crust is formed (Sections 9.3 and 9.4).

←

distances less than about 6 km, is the reflection from the seabed. (From NAT Study Group 1985.) (b) The reduced record section of a refraction profile shot over 140 Ma oceanic crust in the western central Atlantic Ocean on the North American Plate. The seismograms have been low-pass filtered at 15 Hz, corrected for varying sizes of explosive charge and plottend with a gain that increases linearly with distance. P_3 and S_3 are the crustal refractions, P_mP and S_mS the reflections from the crust–mantle boundary and P_n and S_n the upper-mantle refractions. The S phases are converted shear waves (converted from P to S and then back from S to P at the sediment–crust interface beneath the shot and receiver, respectively). (From Purdy (1983).)

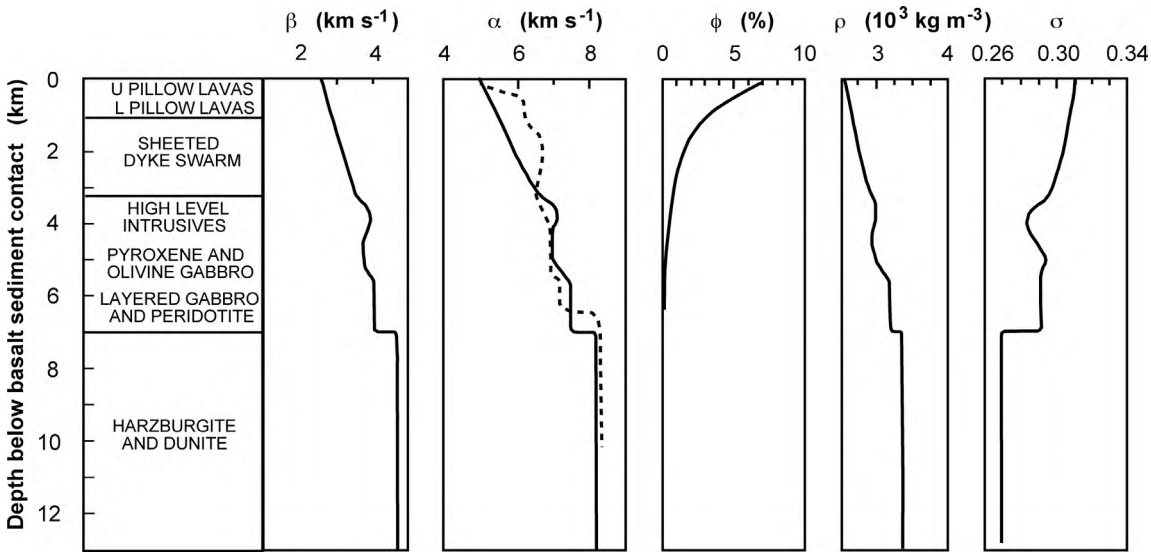


Figure 9.9. Shear-wave velocity β, compressional-wave velocity α, effective porosity ϕ, density ρ and Poisson's ratio σ, versus depth and rock type as measured in the laboratory, on samples taken from the Semail ophiolite in Oman. The dashed line shows the compressional-wave velocity measured on samples from the Blow Me Down Massif, an ophiolite in Newfoundland, Canada. (After Christensen and Smewing (1981).)

9.2.2 Oceanic upper mantle

The seismic P-wave velocity of the uppermost oceanic mantle averages 8.1 km s^{-1}. The oceanic crust–mantle boundary may be a sharp transition but can have a finite, variable thickness and is often a thin gradient zone. It marks the transition from the basaltic–gabbroic crust to the peridotitic mantle. As with all seismic and geological boundaries, there is some uncertainty about the exact correlation between the seismic boundary and the petrological boundary – hence expressions such as the 'seismic Moho' and the 'petrological Moho', which need not be exactly coincident. Correctly speaking, the Moho is a seismic boundary, and presumably it is the transition downwards into ultramafic rock. Petrologists, however, distinguish between zones of cumulate ultramafic rock (such as dunite, which is made of olivine; see Fig. 8.10), which have precipitated from a melt, and the underlying deformed residual upper mantle known as tectonite ultramafic. To a petrologist, this boundary is the significant dividing line because it separates the top of the residual mantle from the overlying material which is ultimately derived from the partial melt at the ridge axis. This boundary is termed the petrological Moho.

There is a gradual increase of seismic P-wave velocity with depth in the oceanic upper mantle, amounting to about 0.01 km s^{-1} per kilometre of depth. In addition, there is evidence of anisotropy: the P-wave velocity of the upper mantle

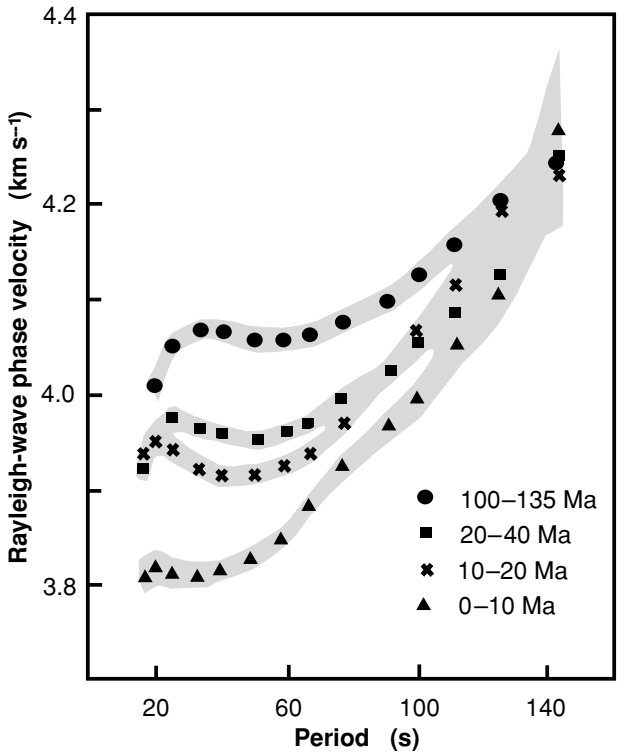


Figure 9.10. Rayleigh-wave phase-velocity-dispersion curves (see Section 4.1.3) for oceanic lithosphere of various ages. Shading shows one standard deviation in the measurements. Notice that, for periods less than 100 s, the phase velocity increases with the age of the plate. For periods greater than 100 s, the velocities are within error of each other. (After Forsyth (1977).)

is dependent on the azimuth of the ray path. Velocities measured perpendicular to the mid-ocean-ridge axis are greater than velocities measured parallel to the ridge axis. This velocity anisotropy is believed to be caused by the preferential aligning of olivine crystals in the mantle parallel to the direction of flow (see Section 8.1.4). Mantle S-wave velocities are reduced beneath ridges when compared with the ocean basins, down to about 150 km depth.

Surface-wave-dispersion curves (Section 4.1.3) for oceanic paths show that phase velocities of Rayleigh waves are significantly reduced for young oceanic lithosphere and asthenosphere (Fig. 9.10). Interpretation of these dispersion curves suggests that the lithospheric thickness increases from some 30 km at 5 Ma to 100 km at 100 Ma, whereas lithospheric S-wave velocities increase from 4.3 to 4.6 km s^{-1} with age. Similarly, asthenosphere S-wave velocities increase from 4.1 to 4.3 km s^{-1} with age.

9.3 The deep structure of mid-ocean ridges

9.3.1 Geophysical evidence

The free-air gravity anomaly across the mid-ocean ridges is not zero, indicating that the ridges are not in total isostatic equilibrium (see Section 5.5 for the definition and use of these terms). Partial compensation is attained by virtue

of the presence of low-density material in the upper mantle beneath the ridge. Because gravity models are non-unique, they need to be constrained. Figure 9.11 shows four density models for the Mid-Atlantic Ridge. In the first model, the low-density zone is narrow and extends 200 km deep into the mantle; in the other three models, compensation is achieved by a broad and fairly shallow low-density zone. All four models adequately satisfy the gravity data. For reasons explained later, the first model (Fig. 9.11(c)) is thought to be the best representation of the density structure beneath the ridge.

The most direct evidence of the deep structure of the mid-ocean ridges has come from earthquake seismology. Results of early studies of the Mid-Atlantic Ridge north of 50°N, using data on earthquakes recorded in Iceland and Greenland, indicated that the upper mantle there had a very low P-wave velocity and that this low-velocity zone extended to perhaps 250 km and was a few hundred kilometres wide. The detailed three-dimensional S-wave-velocity structure of the upper mantle obtained by the inversion of long-period seismic recordings shows that, beneath the mid-ocean ridges in the depth interval 25–250 km, the S-wave velocity is reduced by 2%–8%. Similarly, Q_β values are considerably reduced in the depth interval 0–150 km, being ~300 at 0–30 km and ~70 at 30–150 km depths, compared with values of 600 at 0–80 km and 80 at 80–150 km depths in PREM. Seismic tomography (Section 8.1) confirms the depth extent of the low-velocity regions beneath the mid-ocean ridges as being about 250 km (Figs. 8.6–8.9). Seismologists working with teleseismic data from earthquakes have reported a 'gap in the lithosphere' beneath the mid-ocean-ridge system, across that Sn (a seismic shear wave that propagates in the uppermost mantle) is very poorly propagated. Sn propagates well across stable regions such as continental shields and deep ocean basins but only very inefficiently when its path crosses the mid-ocean-ridge system or the concave side of an island arc. In addition, it is generally noted that magnitudes of ridge–crest earthquakes are often lower than those of earthquakes with comparable surface-wave magnitude but which are located away from the mid-ocean-ridge system. Frequently, surface waves are observed from mid-ocean-ridge earthquakes for which no body waves are detected. All of these observations can be explained by the presence of an absorptive zone in the upper mantle beneath the mid-ocean ridges. Such a zone, which is also limited in extent, can most readily and reasonably be explained as being due to partial melting, occurring because there the upper mantle material is raised above its solidus. Results of seismic-refraction experiments along and parallel to the crest of the Mid-Atlantic Ridge also suggest that, in some places, there is a zone of strong P- and S-wave absorption extending some 25 km on either side of the ridge axis and indicate that the top of this zone lies at least 7 km beneath the seabed.

Results of detailed studies of the source mechanism of large, ridge-axis earthquakes on the Mid-Atlantic, South West Indian and American–Antarctic ridges and the Gorda Rise and Galapagos spreading centres indicate that all the foci

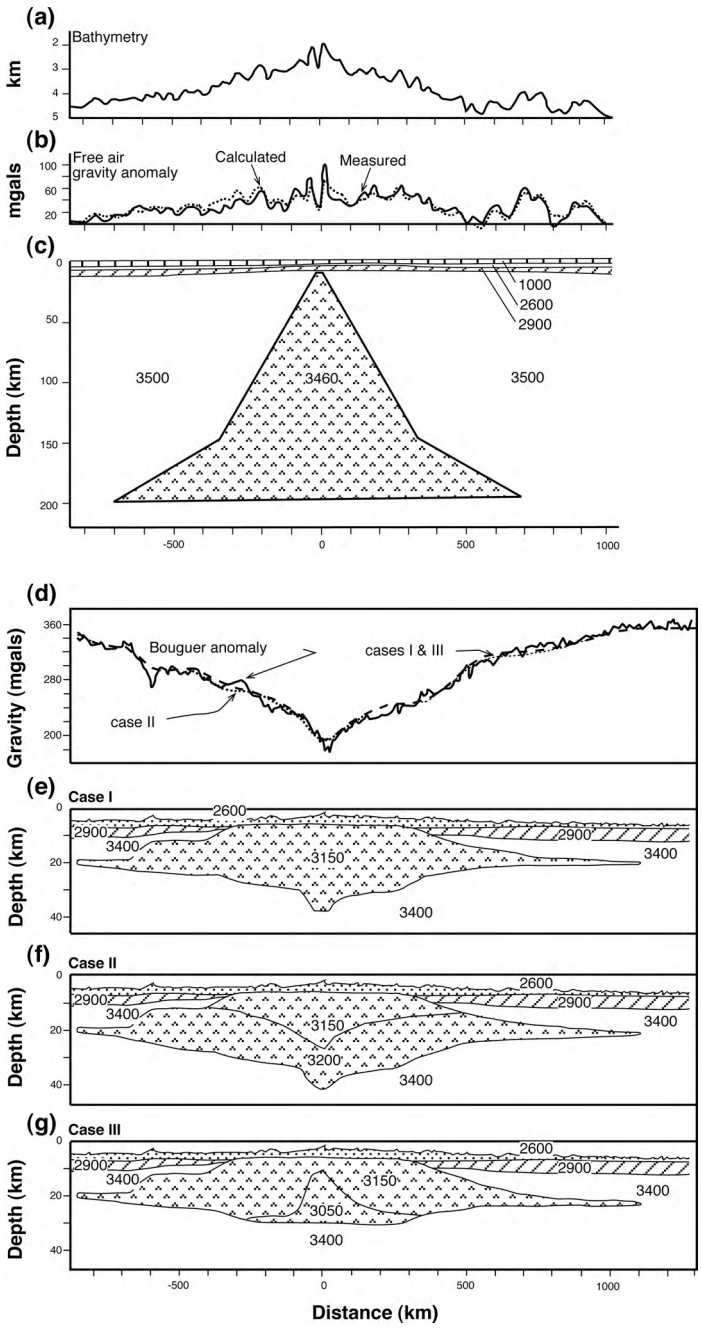


Figure 9.11. Gravity models for the Mid-Atlantic Ridge: (a) bathymetry; (b) free-air gravity anomaly and (c) density model for the Mid-Atlantic Ridge at 45°N; (d) Bouguer gravity anomaly and (e), (f) and (g) density models that also all satisfy the anomaly shown in (d). These four density models – (c), (e), (f) and (g) – illustrate the non-uniqueness of models based on gravity data. A low-density zone lies beneath the ridge, but its dimensions need to be constrained by other methods also. The oceanic crust is assumed to be continuous across the ridge axis in model (c); but in models (e), (f) and (g) there is a zone some 800 km wide and centred on the ridge axis from which normal oceanic crust and uppermost mantle are absent. The density model in (c) is in better agreement with everything that is known about mid-ocean-ridge structure than are the models in (e), (f) and (g). Densities are given in $kg\,m^{-3}$. (After Talwani *et al.* (1965) and Keen and Tramontini (1970).)

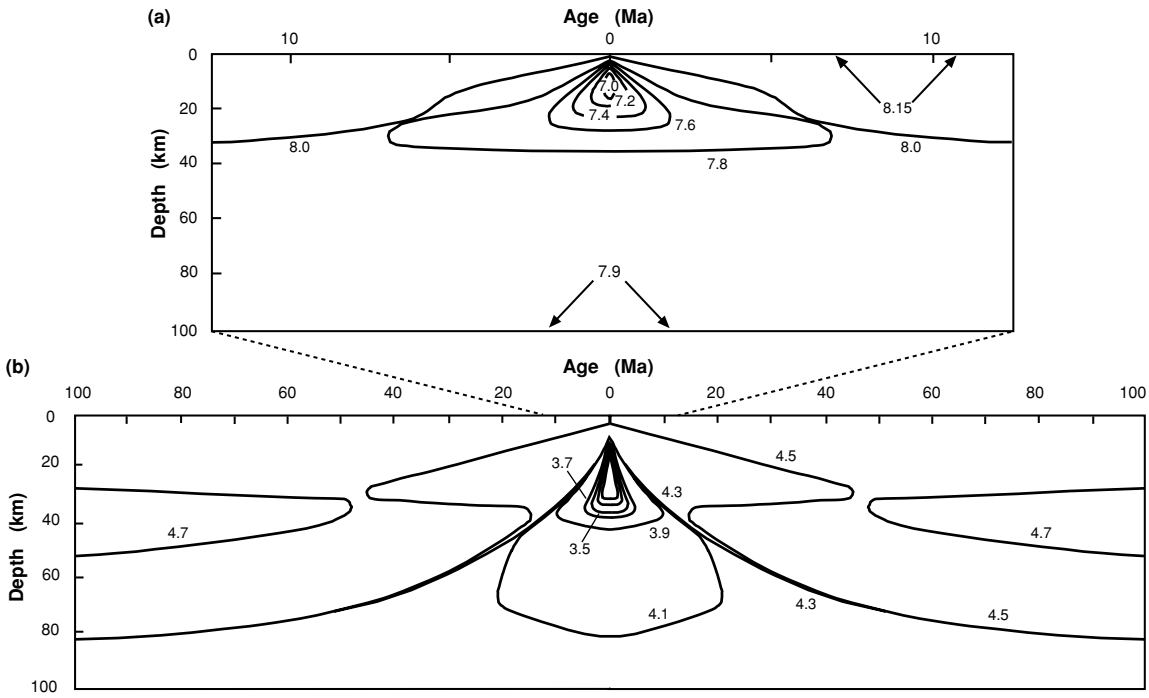


Figure 9.12. Seismic-velocity models of mid-ocean ridges calculated from a thermal model similar to those in Section 7.5.2 and theoretical estimates of the dependence of seismic velocity on temperature and melting of mantle material (assumed to be peridotite with 0.1% water). (a) Contours of constant P-wave velocity in the oceanic lithosphere and asthensphere. The 5-Ma-wide, 30-km-deep low-velocity region centred on the ridge axis is modelled as a zone of extensive melting. (After Solomon and Julian (1974).) (b) Contours of constant S-wave velocity in the oceanic lithosphere and asthenosphere. The hydrous (wet) solidus and anhydrous (dry) solidus for peridotite are approximately delineated by the 4.3- and 3.7-km s^{-1} contours, respectively. The 4.3-km s^{-1} contour therefore represents the boundary between the lithosphere and asthenosphere if the base of the lithosphere is defined as the depth at which partial melting first starts. Note the different horizontal scales for (a) and (b). (From Duschenes and Solomon (1977).)

were extremely shallow (1–6 km) and were located beneath the median valley. The mechanisms were nearly pure normal faulting on planes dipping at 45° with strike parallel to the local trend of the ridge axis. In addition, the focal depths decreased with increasing spreading rate, which is consistent with the theory that the maximum epicentral depth is representative of the depth at which the lithosphere ceases to deform in a brittle manner and ductile deformation takes over.

Figure 9.12 shows detailed P- and S-wave-velocity models of a mid-ocean ridge. These models were derived from a thermal model similar to the models discussed in Section 7.5.2 by estimating the dependence of seismic velocity on temperature and melting for a wet (0.1% water) peridotitic mantle. The large

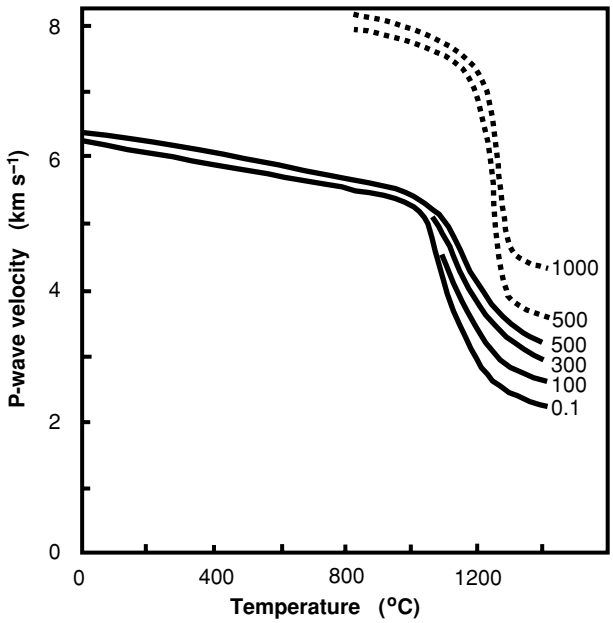


Figure 9.13. Effects of temperature and pressure on P-wave-velocities. Ultrasonic P-wave-velocity measurements were made on dry basaltic (lower four curves) and peridotitic (upper two curves) samples in a laboratory. The numbers beside each curve denote the pressure (MPa) at which the measurements were made. Note the sudden decrease in velocity at the solidus temperature and the steady decrease over the solidus–liquidus interval (approximately 1000–1200 °C for basaltic samples and 1200–1300 °C for peridotitic samples). These measurements were made on dry samples. The presence of water would reduce the solidus and liquidus temperatures but would not greatly change the shape of curves: they would merely be shifted to the left and slightly downwards. (After Khitarov *et al.* (1983).)

velocity gradients in the low-velocity zone (1 km s^{-1} over distances of 15–25 km) are the result of partial melting of mantle material. The S-wave-velocities predicted by this model are in broad agreement with all the surface-wave-dispersion results. Were the mantle drier than was assumed for these models, the melting temperatures and hence the velocities would be somewhat higher.

All of the seismic-velocity models are dependent on the elastic properties of partially molten systems. The effects of partial melting on seismic velocity and attenuation are significant, as illustrated by an NaCl–ice system: for experiments on a sample with 1% concentration of NaCl, reductions in the P- and S-wave velocities of 9.5% and 13.5% were measured at the onset of melting, and attenuation increased by 37% and 48%, respectively. These measurements were made for a 3.3% melt. Partial melting can have a large effect on seismic velocity and attenuation even at very low melt percentages. Results of experiments on basaltic and peridotitic samples indicate that a large decrease in velocity occurs over the liquidus–solidus temperature range (Fig. 9.13). Basalt P-wave velocities

decreased from about 5.5 km s^{-1} to about 3.5 km s^{-1}, and peridotite P-wave velocities decreased from 7.5 to 5.5 km s^{-1}. The pre-melting dependence of P-wave velocity on temperature for these rock samples was an almost linear decrease from 0 °C to the onset of melting.

None of these geophysical investigations can distinguish between *in situ* partial melting and an aggregation of melt that has separated from a region of lesser melt content. However, geochemistry broadly implies that melting beneath the ridge axis occurs in the top 50 km of a region more than 100 km wide. Partial melting of 18% of the mantle in this zone would be sufficient to produce the oceanic crust and to maintain the overall melt fraction in the zone below 2%.

9.3.2 Melting under mid-ocean ridges

Ridges are mostly passive structures. The interior of the mantle is everywhere so hot that, if mantle rocks were brought up to the surface without temperature loss, the decompression from mantle pressures to atmospheric pressure would cause them to melt. Fertile mantle (that is, mantle that has not already been depleted of some of its chemical components by a partial-melting event) will always melt when it is brought up adiabatically to 40 km or less below the surface. Upwelling must occur under any rift, to fill the vacated space: if the extension on rifting is infinite, the melting produces a mid-ocean ridge.

The global variation in major-element composition of the oceanic crust can be accounted for by 8%–20% partial melting of the mantle at pressures of 0.5–1.6 GPa (depths of 15–50 km). There is little variation to first order, hence the term MORB[2] (mid-ocean-ridge basalt). This is because the most important control on melting is the temperature of the mantle and that is fairly uniform except in the vicinity of mantle plumes. However, there are some slight variations. The proportions of some incompatible elements (e.g., Na) present in mid-ocean-ridge basalts correlates with the spreading rate of the ridge: basalts from very slowly spreading ridges have greater concentrations of incompatible elements. This is a consequence of the fact that a small degree of partial melting means that the incompatible-element concentration is high, whereas, with a greater degree of partial melting, the incompatible elements are diluted and so their concentration is lower. There is also a correlation between the concentration of rare-earth elements (REE) in MORB and crustal thickness (and hence spreading rate and mantle temperature). Some correlation of major-element chemistry with the depth of the ridge axis (the depth of the seabed at the intrusion zone) is also observed,

[2] MORB from the mid-ocean ridges in the Indian Ocean differs isotopically from Atlantic and Pacific MORB. Indian MORB has high $^{87}Sr/^{86}Sr$ and low $^{206}Pb/^{204}Pb$ ratios. The eastern boundary of this isotopic difference approximately coincides with the Australian–Antarctic Discordance (AAD) south of Australia. The western boundary is near the Andrew Geddes Bain Fracture Zone at 31°E. These differences probably reflect mantle convection patterns and reservoirs beneath the Indian Ocean.

which indicates that the axial depth is dependent on the thermal structure of the underlying mantle. Where the mantle is hottest (and the ridge axis shallowest), the rising mantle material intersects the solidus earlier, at greater depth (Fig. 7.17), and so melts more. Maximum differences in temperature of some 200 °C in the subsolidus mantle may be necessary to account for this observed correlation. Results of seismic and electromagnetic studies (MELT) of the East Pacific Rise suggest that small amounts of melt (<2%) are present over a very wide region within the mantle. The melt then converges on the narrow active ridge axis. Variations in chemistry of erupted lavas are attributed to the details of the path taken by the melt from source to surface. The fine details of the melting, rising, interaction, solidification and eruption processes are still matters of current research.

In order to calculate how much partial melting occurs when mantle at a temperature T upwells and so is decompressed, it is necessary first to express the solidus and liquidus temperatures for the mantle as functions of pressure (e.g., Fig. 7.17). For a garnet-peridotite mantle the solidus temperature T_s (in °C) and the pressure P (in GPa) are related by the following equation:

$$P = \frac{T_s - 1100}{136} + 4.968 \times 10^{-4} e^{1.2 \times 10^{-2}(T_s - 1100)} \tag{9.1}$$

The liquidus temperature T_l (in °C) also varies with pressure P (in GPa):

$$T_l = 1736.2 + 4.343P + 180 \tan^{-1}(P/2.2169)$$

The degree of melting, as a fraction by weight of the rock, x, is then given by

$$x - 0.5 = T'(T'^2 - 0.25)(0.4256 + 2.988T') \tag{9.2}$$

where T', a dimensionless temperature, is defined as

$$T' = \frac{T - (T_s + T_l)/2}{T_l - T_s} \tag{9.3}$$

Surprisingly, there is no clear evidence for variation of the degree of melting $x(T')$ with pressure.

The thickness of the oceanic crust is almost constant on all the plates; 7 ± 1 km (Table 9.2), provided that seismic measurements are made away from fracture zones where the crust is on average some 3 km thinner (Section 9.5.2), and plumes where the crust is thicker (e.g., Iceland). In order to generate 7-km-thick oceanic crust, 7 km of melt is needed from the upwelling mantle. Equations (1)–(3) show that, to produce this amount of melt, the potential temperature (Eq. (7.95)) of the source region needs to be about 1280 °C (McKenzie and Bickle 1988). Rising mantle will then cross the solidus at 1300 °C at a depth of about 45 km, and melt will reach the surface at 1200 °C (Fig. 7.17). These values yield an average melting depth of 15 km and a melt fraction of 10%–15%. The magma would be about 10% MgO, and the melt fraction would not exceed 24%. Estimates of melting based on matching crustal thickness are broadly similar: a normal potential temperature of 1300 ± 20°C, with melting starting at 50 km depth and

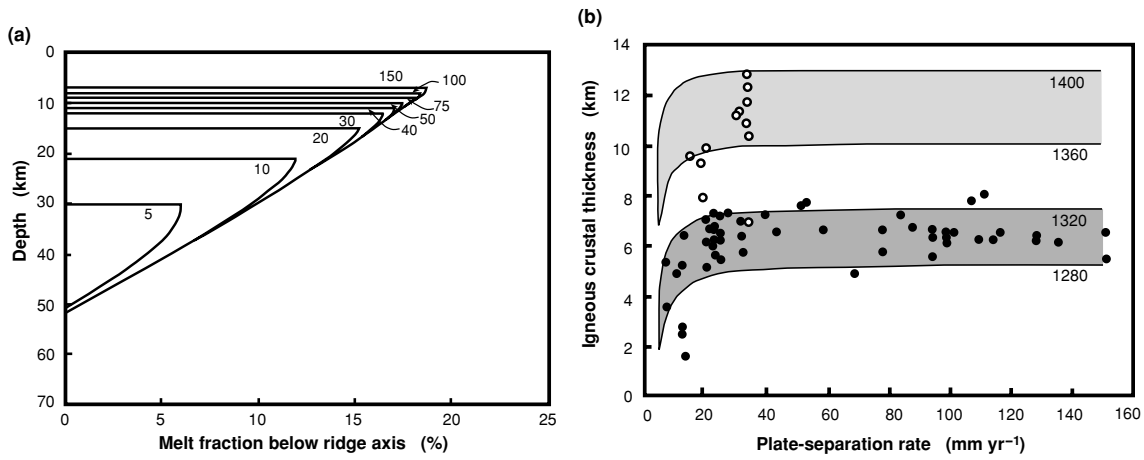


Figure 9.14. (a) The predicted melt fraction in the mantle beneath a mid-ocean ridge calculated for a mantle potential temperature of 1300 °C. The curves are labelled with the half-spreading rate (mm yr^{-1}). For very slow spreading rates, below 1 cm yr^{-1}, there is a marked drop in the percentage of the melt fraction in the mantle and a restriction on the depth interval over which melting takes place as the depth of the upper limit on melting increases. (b) The predicted crustal thickness calculated for potential temperatures of 1280–1320 °C appropriate for normal oceanic mantle and 1360–1400 °C for mantle in the vicinity of a hotspot. Crustal thicknesses determined by seismic experiments are shown as dots. (See White *et al.* (2001) for other models.) (Reprinted from *Earth Planet. Sci. Lett.*, **121**, Bown, J. W. and White, R. S., Variation with spreading rate of crustal thickness and geochemistry, 435–49, copyright 1994, with permission from Elsevier.)

the maximum melt fraction of 15%–18% localized between 8 and 15 km depth (Fig. 9.14).

Oceanic crust produced at ridges with very slow spreading rates (half-rate <1 cm yr^{-1}, full rate <2 cm yr^{-1}) is very much thinner than normal (4–5 km only on the Southwest Indian Ridge, which is spreading at 0.8 cm yr^{-1}, and less in the Arctic). These are the spreading rates below which the loss of heat by conduction from the ascending mantle becomes crucial. The consequences of very slow spreading are that the depth to the top of the melting zone increases and the melt fraction is reduced. Together these factors mean that, at very slow spreading rates, the total production of melt is much reduced and so the crust is thinner (Fig. 9.14). The 1800-km-long Gakkel Ridge (Figs. 2.2 and 9.18) which crosses the Arctic Ocean beneath the ice from northwest of Svalbard to the Lapatev Sea is the slowest-spreading mid-ocean ridge on the Earth (the half-spreading rate decreases from 0.75 cm yr^{-1} in the west to $\sim$0.3 cm yr^{-1} in the east). It may spread so slowly at its eastern end (where it is very close to the rotation pole) that that there is effectively no basaltic crust, very enriched magmas erupt and peridotite may be exposed over large areas.

The crust in the vicinity of Iceland, which is situated over a plume, is much thicker than normal. The crustal thickness beneath central Iceland averages some 35 km, decreasing to ∼20 km on the Reykjanes Peninsula in southwest Iceland and 10–13 km beneath the Reykjanes Ridge. In order to generate this amount of melt, the potential temperature of the mantle must be considerably higher than normal, up to about 1480 °C beneath Iceland and 1360–1400 °C beneath the Reykjanes Ridge.

9.4 The shallow structure of mid-ocean ridges

9.4.1 Topography

The axial topography of mid-ocean ridges varies widely (Figs. 9.1 and 9.15). The topography and the formation of the oceanic crust are controlled by the magmatic supply, tectonic strain and cooling by hydrothermal circulation but the spreading rate is the most important overall control on the topography. The crest of the slow-spreading Mid-Atlantic Ridge is very rugged, with faulted blocks and a narrow axial *median valley* (20–30 km wide and 1–2 km deep overall). The very-slow-spreading Southwest Indian Ridge is also very rugged, in places the median valley is >2 km deep, with the bathymetric depth at the axis over 4 km (c.f. Eq. (7.72)). The slowest-spreading ridge, the Gakkel Ridge, is very rugged, with large scarps along a 2-km-deep continuous median valley and a bathymetric depth at the axis of ∼5 km. In contrast, the axial regions of the faster-spreading Pacific–Antarctic Ridge and the East Pacific Rise are much smoother and generally lack the median valley, having instead an axial high some 1–2 km wide. Globally the changeover from median valley to no valley is at half-spreading rates of 3–4 cm/yr^{-1} (Fig. 9.16). However, the presence of a median valley is not controlled purely by the spreading rate, with slow-spreading ridges having median valleys and fast-spreading ridges lacking them: there are exceptions. For example, much of the slow-spreading Reykjanes Ridge does not have a median valley, and in some locations the fast-spreading East Pacific Rise has a shallow axial valley. These differences in axial topography can be modelled by the ductile (i.e., non-brittle) extension of a viscous or plastic lithosphere, which thickens with distance from the ridge axis, as well as the effect of accumulation of magma beneath the plate. For slow-spreading ridges, the depth of the rift valley primarily depends on the strength of the model lithosphere and the rate at which it thickens. On fast-spreading ridges the lithosphere is weaker and cannot support significant stress. This means that the seabed acts as a free surface and the topography is due mainly to the buoyancy of the magma. The Reykjanes Ridge, even though it is a slow-spreading ridge, does not have a median valley because the overall effect of the hotter underlying mantle close to the Iceland plume means that the lithosphere is hotter and hence weaker than normal. On the Reykjanes Ridge the change from a 700-m-deep median valley in the south to an axial high happens

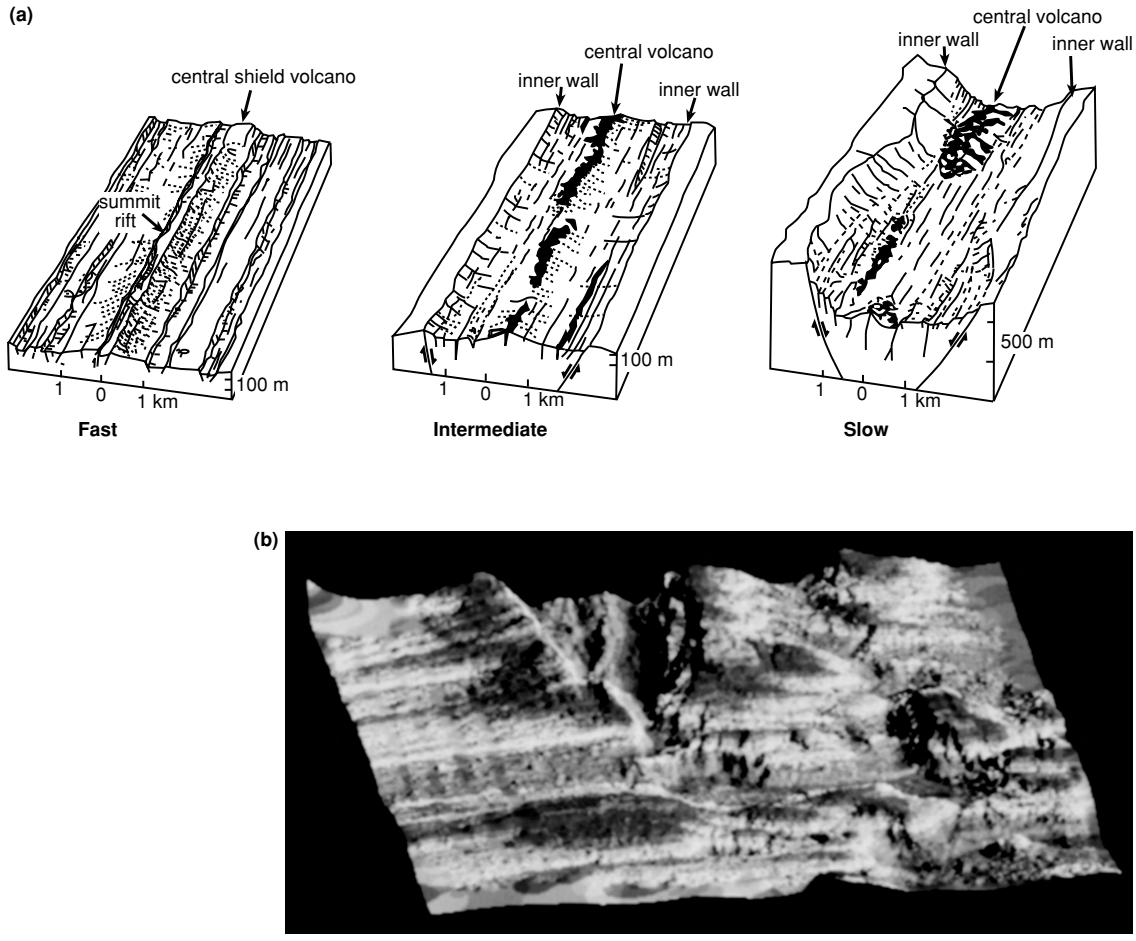


Figure 9.15. (a) Schematic diagrams to show the major differences in the axial morphology amongst fast-, intermediate- and slow-spreading mid-ocean ridges. As the spreading rate decreases, the height of the normal-faulted inner wall increases and the central volcano ceases to be a continuous feature. The vertical is exaggeration is 2 : 1. (After Macdonald (1982)). Reprinted, with permission, from the *Annual Review of Earth and Planetary Sciences*, volume 10 © 1982 by Annual Reviews www.annualreviews.org.) (b) A three-dimensional bathymetric image of the Rodriguez or Indian Ocean triple junction (Fig. 2.2) looking towards the west. This is the RRR triple junction (Fig. 2.16) where the Antarctic, Australian and African plates meet. The very-slow-spreading Southwest Indian Ridge (top of image) has a deep rift valley and uplifted rift flanks. The two intermediate-spreading ridges, the Southeast Indian Ridge (left of image) and the Central Indian Ridge (right of image), have almost the same north-south strike, and are broad and regular. The image covers approximately 90 km $\times$ 90 km. Colour version Plate 18. (N. J. Mitchell 1989 data collected by M. Munschy and R. Schlich, IPG Strasbourg, France.)

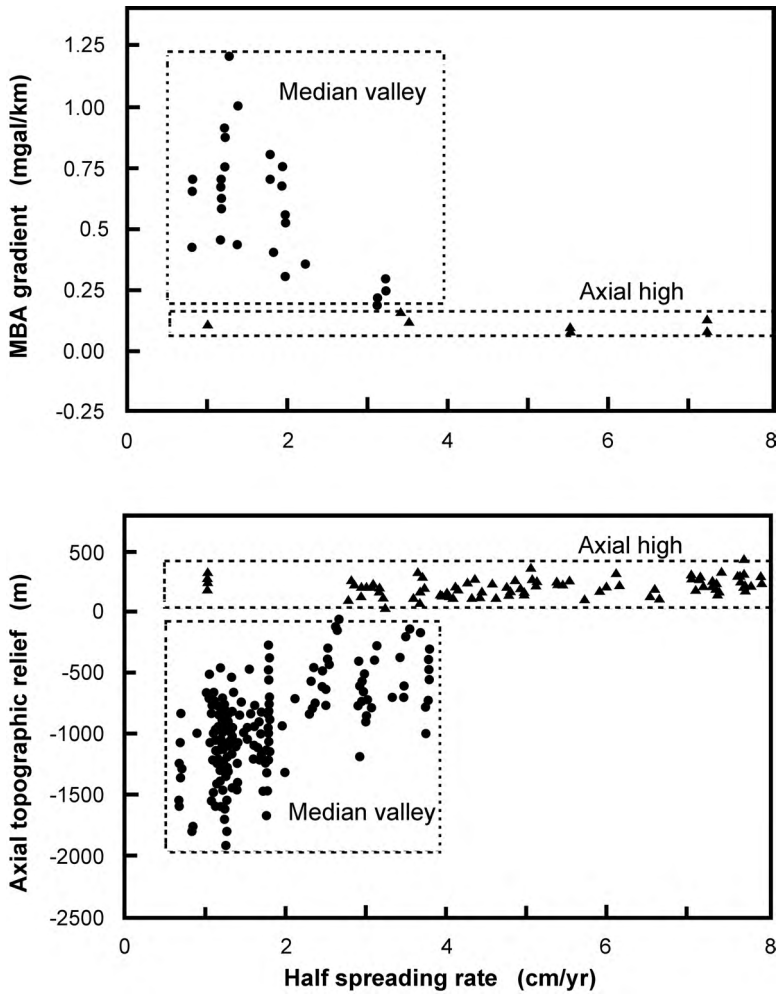


Figure 9.16. The gradient of the mantle Bouguer anomaly (MBA) and the axial topographic relief plotted against the half-spreading rate. The major difference between ridges with a median valley and those without suggests that an axial high is indicative of efficient transport of melt along the ridge axis. (After Wang and Cochran (1995).)

abruptly at 58° 40′N. On the East Pacific Rise there is an axial high where a magma chamber is present but where no magma chamber is evident (at 23°N), there is instead an 18-km-wide, 1-km-deep axial valley.

Side-scan sonar and swath bathymetry yield detailed images of the fabric of the seabed (Figs. 9.15(b), 9.19 and 9.31 and Plates 18, 20 and 21). On slow-spreading ridges the median valley generally has a 5–12-km-wide inner valley floor within which the volcanic activity takes place. Within each segment the width and depth of the median valley is thermally/magmatically controlled, with the valley being shallower and narrower at segment centres where the magma supply is greatest. The magma supply may be a direct feature of the upwelling process or it may be that melt migrates laterally along segments from the centre. Often small individual volcanic eruptions build a linear volcanic ridge, an axial volcanic ridge (AVR).

Table 9.3 *Volcanic characteristics of mid-ocean ridges*

	Full spreading rate (cm yr^{-1})	Width of neovolcanic zone (km)	Recurrence interval for eruptions (year)	Volume of individual eruptions (10^6 m^3)
Fast	8–16	0.1–0.2	5	1–5
Intermediate	4–8	0.2–2.0	500	5–50
Slow	1–4	2–12	5000	50–1000

Source: Perfit and Chadwick (1998).

An AVR can be up to a few hundred metres high and tens of kilometres long. On either side of the inner valley floor normal faults raise the crust up to the crestal mountains. Individual faults can be hundreds of metres high. Outward-facing fault scarps are rare. Results of studies on the Mid-Atlantic Ridge have shown that most of the plate separation is achieved by magmatic accretion (or some diffuse deformation within the crust) since only ~10%–20% is accommodated by brittle faulting on visible faults. The zone of active deformation extends to 15 km from the ridge axis but almost all the volcanic activity takes place within 5 km of the axis (Table 9.3). It is possible that the geometry of bounding faults may constrain the location of the magmatism with the median valley.

Figure 9.17 illustrates the extent of variations in topography, tectonics and geology found on the Mid-Atlantic Ridge. The area which has undergone the most recent volcanic constructional event is shown at the top of Fig. 9.17, and the area at the bottom has the most developed extensional tectonic features. The other areas are arranged in order between these two extremes. It seems that this variability, which can occur over short distances along the axis, is the result of a cyclic tectono-magmatic process (with a period of perhaps 10 000–50 000 years) with more or less continuous extension (moving apart of the two plates) but only periodic or intermittent magmatism. The magmatism in neighbouring ridge segments, or cells (Section 9.5.2), need not be synchronous. The last major volcanic event would have occurred within the last few hundred years for the area at the top of Fig. 9.17, whereas those areas at the bottom are in the last extensional phases of their cycle. The time-scale and length scale of tectono-magmatic evolution are large for slow-spreading ridges but much less for fast-spreading ridges.

As is evident from Fig. 2.1, earthquakes of large magnitude are not common along the mid-ocean-ridge system. Most large earthquakes occur along the transform faults. Earthquake activity associated with small-scale faulting and the movement of magma is generally of magnitude too low for detection at teleseismic distances. Local seismic networks do, however, record earthquake swarms. Unusually there was a major period of earthquake activity that was recorded at

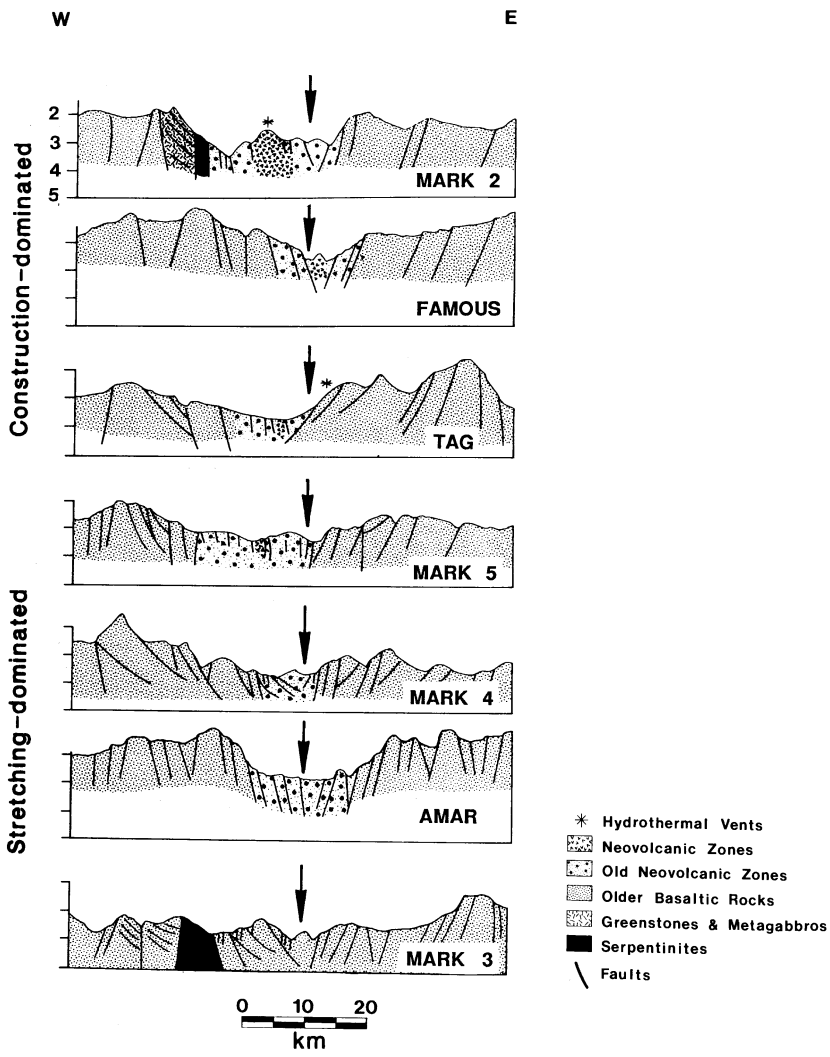


Figure 9.17. Schematic geological cross sections of the axis of the Mid-Atlantic Ridge. The MARK cross sections are at intervals of about 15 km along the axis at 23°N, with MARK 2 being the most northerly and MARK 5 the most southerly. FAMOUS is at 37°N, AMAR at 36°N and TAG at 26°N. Arrows indicate the ridge axis. The vertical exaggeration is 3 : 1. Depths beneath the sea surface are in kilometres. (From Karson *et al.* (1987).)

teleseismic distances from the Gakkel Ridge in the Arctic Ocean in 1999. Over 200 earthquakes were recorded over a 7-month period, a sudden and huge increase on the seismic activity recorded over previous decades (Fig. 9.18, Plate 19). Initially earthquakes had normal-faulting mechanisms, as is usual for an extensional tectonic environment. Later the activity changed to focal mechanisms of a non-double-couple type typical of caldera-collapse volcanic activity. It is possible that the earthquakes were caused by a major dyke propagating along the ridge axis through the cold, brittle crust. Subsequent side-scan sonar images confirmed the presence of fresh lava flows. This is quite unlike seismic activity recorded from slow-spreading ridges, for which earthquakes occurring in swarms are generally too small to be detected teleseismically, are far fewer in number and have a shorter

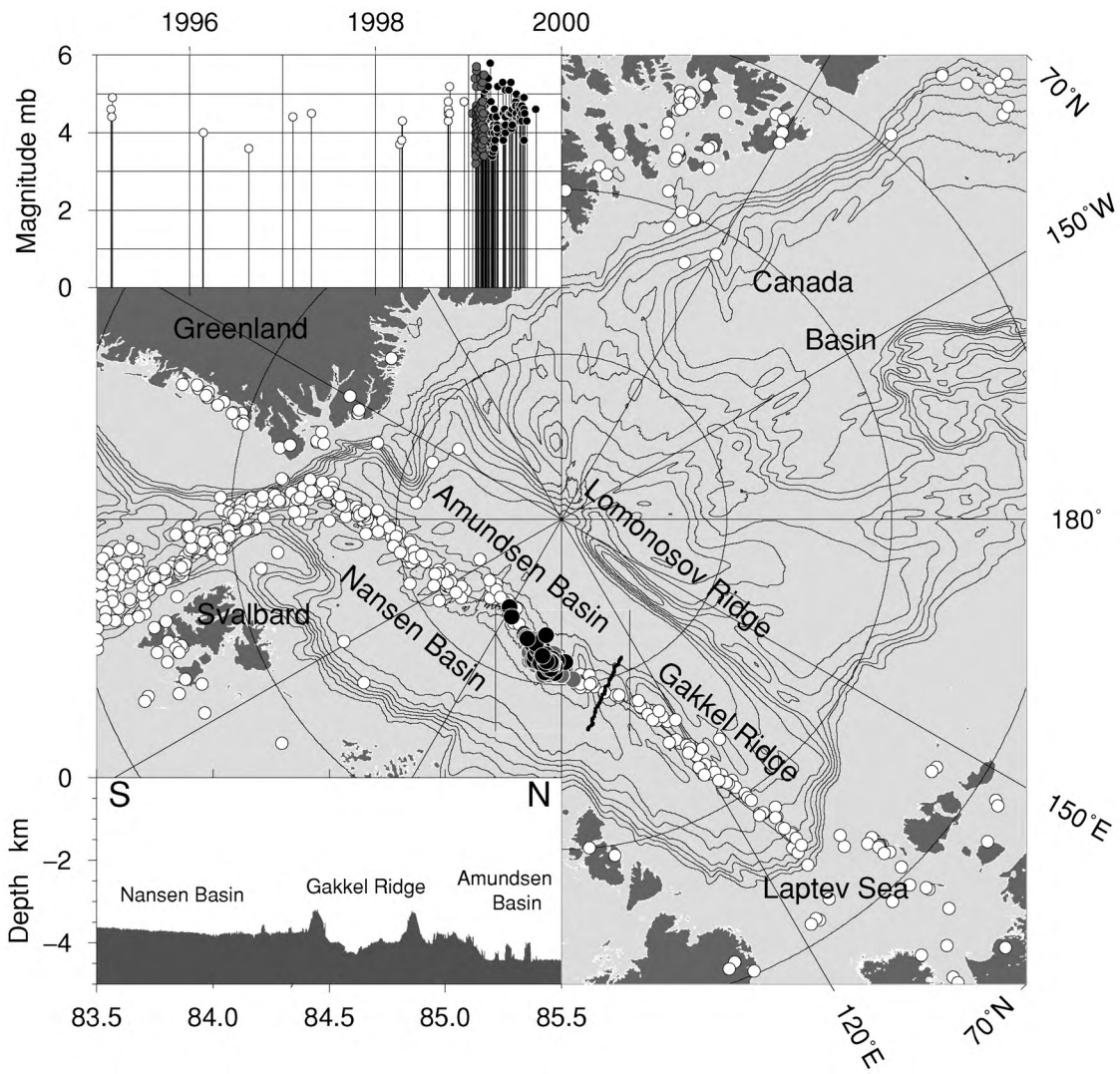


Figure 9.18. The Gakkel Ridge, the slowest-spreading mid-ocean ridge, runs across the Arctic. Grey shading shows bathymetry. Dots show locations of earthquakes in 1999. The upper inset shows their magnitude and the occurrence of earthquakes on the Gakkel Ridge from 1995 to 2000. Colour version Plate 19. (Müller and Jokat, 2000. Seismic evidence for volcanic activity at the Eastern Gakkel Ridge. *EOS Trans. Am. Geophys.*, **81**, 265, 9).

duration. The evidence therefore suggests that that activity on the Gakkel Ridge was due to intense magmatic activity and may indicate that, on ultra-slow ridges, melt generation and volcanic activity is very strongly episodic.

Faulting plays an important role in the accretion of new material along the axis of any mid-ocean ridge. There must be an overall balance between the accretion

of new oceanic crust along the axis of the ridge and slip on the normal faults. The heterogeneous structure of slow-spreading ridges is a result of the spatial and temporal variability of tectonic and volcanic processes (Fig. 9.17). Within individual ridge segments (Section 9.5.3) the fault distribution is such that there are many small faults towards the segment centre whereas the segment ends are characterized by just a few large faults. Fault size and spacing both increase from the segment centre to its end, although the tectonic strain is constant. Understanding of the distribution and size of faults along mid-ocean-ridge segments is greatly improving with the increased availability of high-resolution technology in the form of multibeam and side-scan sonar, coupled with ground-truthing from submersibles.

Multibeam and side-scan sonar imaging have led to spectacular discoveries of what seem to be major detachment surfaces on the Mid-Atlantic Ridge (near the Kane and Atlantis transform faults) and on the Central Indian Ridge (50 km north of the Rodriguez triple junction) (Fig. 9.19 and Plate 20). The surfaces are at inner corners (RT) of ridge–transform intersections (see Section 9.5.2). The surfaces are corrugated and striated platforms that dip gently towards the ridge axis for tens of kilometres parallel to the ridge axis. The geometry of these surfaces is such that, with over 10–15 km of displacement, the lower crust and upper mantle should be exposed. Initial sampling by dredging has yielded gabbros and serpentinite from the corrugated surface, while it is overlain by upper-crustal basaltic rocks on its median valley side. The corrugations which are parallel to the spreading direction have a wavelength of $\sim$1.0 km and an amplitide of tens of metres. The fine-scale (wavelength 50–100 m) striations which also parallel the spreading direction seem to be alternate exposures of hard rock and sediment-filled depressions. These corrugations and striations seem to be directly related to the formation of the detachment fault. Although erosion has removed the surface corrugations, Atlantis Bank, a platform adjacent to the Atlantis II fracture zone on the Southwest Indian Ridge, seems to be another example. The 1500-m deep ODP hold 735B has drilled through 1300 m of lower-crustal gabbro and lherzolite. The mechanisms for formation of these surfaces are similar to those suggested for the formation of continental-core complexes: slip on low-angle detachment faults and subsequent tilting of upper-crustal fault blocks. Submersible diving and detailed shallow drilling, linked with magnetic and geochemical studies, should provide information on the detailed operation of processes at slow-spreading ridges and may also have implications for our understanding of continental geology.

9.4.2 Crustal magma chambers

As has been shown, the region of partial melting in the upper mantle is probably relatively broad. This is in sharp contrast to the very narrow zone of intrusion at the ridge axis: magnetic data (see Section 3.2) suggest that the standard deviation of the distribution of dyke injections must be only a few kilometres or less, though

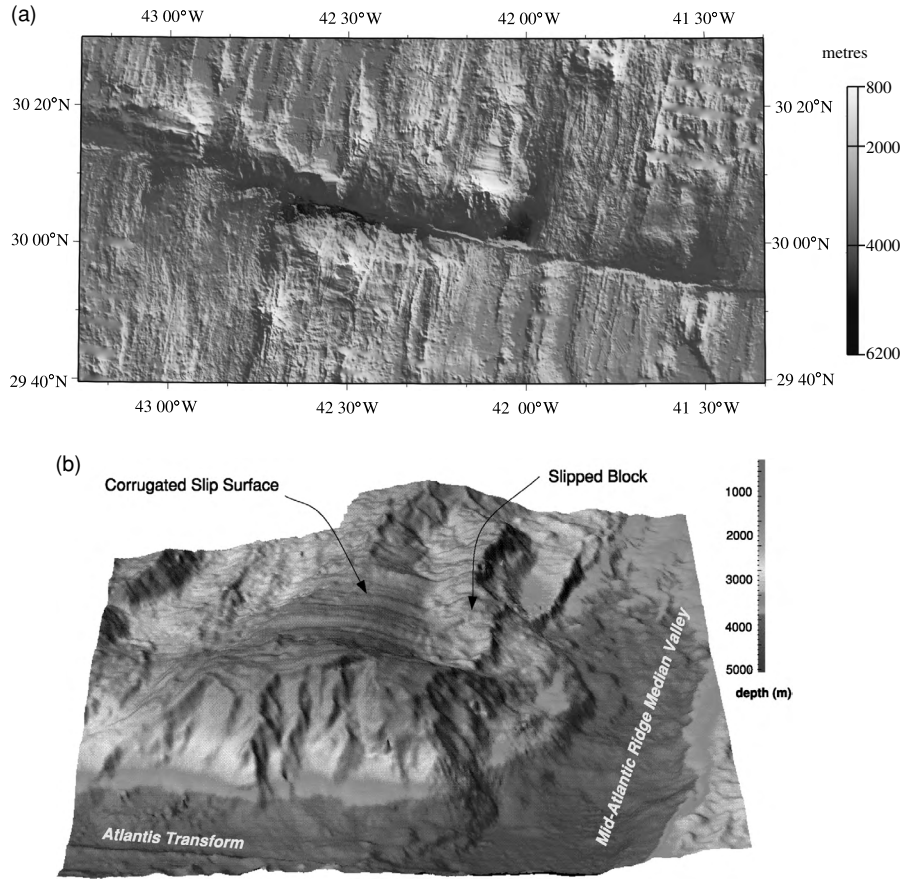


Figure 9.19. (a) Bathymetry of a 90 km × 180 km area on the Mid-Atlantic Ridge, showing the Atlantis transform fault. The deepest parts of the transform are at, and between, the two ridge–transform intersections (RTI) at 30°N, 42°W and 30° 05′N, 42° 40′W. The fracture zone continues as a clear linear feature trending 103° both on the African plate and on the North American plate. The seabed away from the ridge axis is characterized by linear abyssal hills. These volcanic hills are usually a few hundred metres high, parallel the ridge and extend up to the fracture zone. Three topographic highs exhibit the corrugations suggestive of detachment surfaces: at 30° 10′N, 42° 05′W, the active inside corner north of the transform fault; at 30° 15′N, 43° 00′W, an 8-Ma-old fossil inside corner north of the transform fault; and at 29° 55′N, 42° 30′W, a fossil inside corner south of the transform fault. Colour version Plate 20(a). (From Blackman *et al.*, Origin of extensional core complexes: evidence from the Mid-Atlantic Ridge at Atlantis Fracture Zone, *J. Geophys. Res.*, **103**, 21 315–33, 1998. Copyright 1998 American Geophysical Union. Reprinted by permission of American Geophysical Union.) (b) A three-dimensional shaded relief image of the active inside corner shown in (a). The image is viewed from the south and illuminated from the northwest. The area is 30 km east to west and 35 km north to south. The corrugated surface dips at ∼10° towards the median valley. The slipped block is volcanic. Colour version Plate 20(b). (Reprinted with permission from *Nature* (Cann *et al.*, *Nature*, 385, 329–32). Copyright 1997 Macmillan Magazines Ltd.)

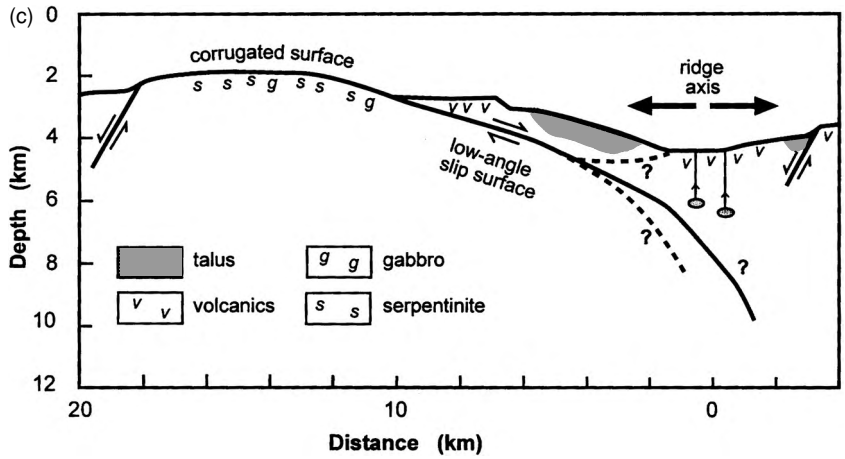


Figure 9.19. (*cont.*) (c) A cartoon cross section of an active inside corner close to a transform fault on a slow-spreading ridge. Compare this with (a) and (b). The geometry of the detachment surface at depth is unknown. Three possible subsurface positions for the active fault are indicated by '?': (1) it could steepen towards the median valley, effectively detaching much of the lower crust and upper mantle from the upper basaltic crust; (2) it could have a more or less constant low-angle dip (although this would mean that it crossed the ridge axis); and (3) it could be a major failure surface outcropping at the seabed on the median valley floor at the foot of a slipped upper-crustal block. (Reprinted with permission from *Nature* (Cann *et al. Nature*, 385, 329–32). Copyright 1997 Macmillan Magazines Ltd.).

the actual extrusion zone of lavas onto the sea floor is probably only a few hundred metres wide.

The comparison of ophiolites with the oceanic crust initially resulted in a number of models (of varying degrees of complexity) of magma chambers in the crust at shallow depth beneath the ridge axis. These models were nicknamed 'infinite onions', infinite because of their continuity in time and along the axis of the mid-ocean-ridge system, terminated only by transform faults, and onion because of their shape and because they peel off layers of oceanic crust at their sides. These magma chambers were thought to produce the basaltic lavas which are erupted onto the seafloor, as well as the underlying dykes and cumulate layers which are present in ophiolites and are thought to make up much of seismic layers 2 and 3. Several thermal models of crustal magma chambers have been made in order to determine limits on their size and dependence on spreading rate, crystal settling and thermal parameters. Clearly, an infinite-onion magma chamber, extending through much of the crust and being several kilometres in width, should be easily detectable seismically. Many detailed seismic-reflection, seismic-refraction, gravity and deep-tow surveys have been carried out in order to delineate magma chambers on the Mid-Atlantic Ridge, Reykjanes Ridge, East

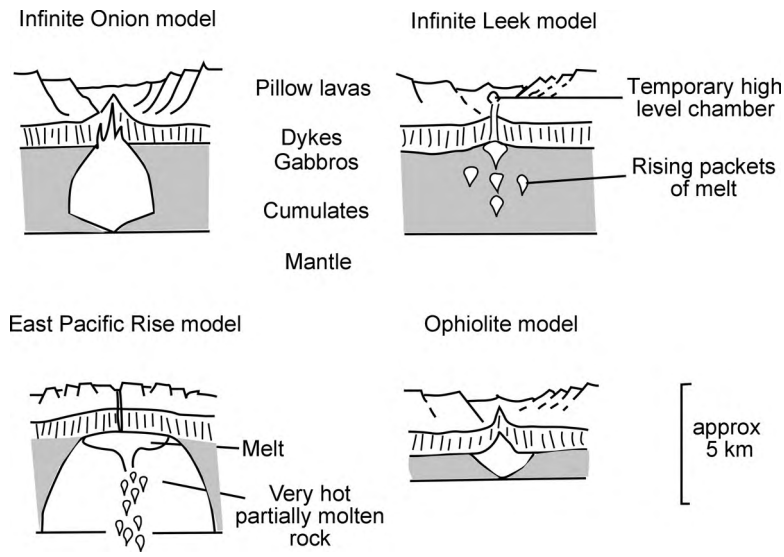


Figure 9.20. Four models of mid-ocean-ridge magma chambers show the development of ideas about the formation of oceanic crust along the mid-ocean ridges (left-hand pair, fast spreading; right-hand pair, slow spreading). There is no vertical exaggeration. The early models, namely the 'infinite-onion' model (Cann 1974; Bryan and Moore 1977) and the 'ophiolite' model (Pallister and Hopson 1981) had large steady-state magma chambers in the lower crust, to produce the lavas and dykes (layer 2) and the gabbros and cumulates (layer 3) of the oceanic crust. Seismic experiments should have been able to detect such chambers. The 'infinite-leek' model (Nisbet and Fowler 1978) did not require a steady-state magma chamber; rather, magma was envisaged as rising intermittently in packets from the mantle, nucleating a crack in the basaltic crust and rising to the surface. Each crack is likely to be in the same location as the previous crack. A small, shallow magma chamber could be a temporary feature of this model. The 'East Pacific Rise' model (Macdonald 1986; 1989) had a mushroom-shaped central magma chamber overlying a zone of very hot rock, which is at most a few per cent partially molten. The infinite-leek and East Pacific Rise models approximate fast- and slow-spreading ridges. The infinite-onion and infinite-leek models could be regarded as extreme end members of a continuum of models whose character depends on the spreading rate and degree of inflation.

Pacific Rise and Juan de Fuca Ridge. Figure 9.20 shows some of the variations of magma-chamber models and their possible settings.

9.4.3 Thermal models

It is easy to see how a large infinite-onion magma chamber could produce the layered lava, dyke and gabbro cumulate crust observed in ophiolites and presumed to comprise the oceanic crust. Thermal models of the formation of the oceanic crust indicate, however, that such an extensive steady-state magma chamber cannot be present beneath slow-spreading ridges such as the Mid-Atlantic Ridge.

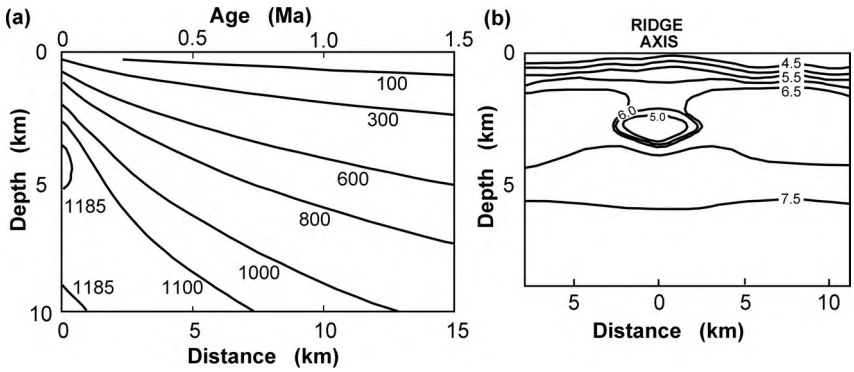


Figure 9.21. (a) Temperature in oceanic crust calculated for a mid-ocean ridge with a half-spreading rate of 1 cm yr^{-1}. The edge of the magma chamber is indicated by the 1185-°C isotherm (the solidus). The chamber is confined to depths appropriate for layer 3 and is only about 1 km wide. (From Sleep (1975).) (b) A seismic P-wave velocity model for the East Pacific Rise at 9°N (half-spreading rate 6.1 cm yr^{-1}) based on the first thermal model discussed in Section 9.4.3. (From Hale *et al.* (1982).)

To calculate thermal models of mid-ocean ridges, we need to know something about the eruption and solidification temperatures of likely crustal and upper-mantle materials. Estimates of 1203–1246 °C have been made for the eruption temperature of basalts collected from the FAMOUS area of the Mid-Atlantic Ridge. In melting experiments on a Mid-Atlantic Ridge basaltic glass, clinopyroxene disappeared at 1166 °C, plagioclase at 1215 °C and olivine at 1245 °C. When melting a Mid-Atlantic Ridge olivine tholeiite (at 0.15 GPa), it was found experimentally that clinopyroxene disappeared at 1160 °C, plagioclase at 1180 °C and olivine at 1210 °C.

Thermal models of mid-ocean ridges are computed numerically by using the heat-conduction equation, Eq. (7.59). The exact size and shape of any magma chamber depends on the assumptions made about the latent heat of solidification and the temperatures for the boundary conditions, as well as the spreading rate. Figure 9.21(a) shows one steady-state model made for a ridge spreading with a half-rate of 1 cm yr^{-1}. In this model, the temperature at the base of the lithosphere was assumed to be 1290 °C, the edge of the magma chamber was defined by the 1185-°C isotherm, and all the latent heat was released at the axis. The maximum half-width of this magma chamber is about 0.5 km. In another model, the temperature at the base of the lithosphere and along the intrusion axis was assumed to be 1125 °C, the edge of the magma chamber was defined by the 1000 °C isotherm and the latent heat was distributed over the liquidus–solidus interval. The maximum half-width of this chamber is about 4 km (or 3 km if the latent heat is released entirely at the solidus). Some of the apparent difference between these two models is due to the different assumptions regarding temperature, and some is due to the different treatments of latent heat and upper boundary conditions.

The temperature measurements discussed previously suggest that the temperature assumptions of the second model are too low. When this second model is recalculated using more appropriate liquidus temperatures of 1230–1250 °C and solidus temperatures of 1160–1185 °C, the maximum half-width of the magma chamber is about 2 km (or less if the latent heat is released entirely at the solidus). Latent heat of about 4.2×10^5 J kg^{-1} is a reasonable estimate for basalt.

The effect of accumulation of crystals settling to the bottom of a magma chamber can also be included in the thermal models. The width of any chamber is reduced by crystal settling: a 2 : 1 ratio of solidified material forming the chamber roof to that falling to the bottom reduces the maximum width of the magma chamber to about one-half its previous value, whereas a 1 : 1 ratio results in a reduction of width by one-third. Crystal settling also has a significant effect on the shape of the magma chamber, changing it from triangular (apex at the top) to a diamond shape.

Thermal models calculated for medium- and fast-spreading ridges are very similar to those for slow-spreading ridges, the only significant difference being that the magma chambers are much wider, as one would expect. Thermal considerations therefore imply that a steady-state crustal magma chamber can exist on fast-spreading ridges, though, for half-spreading rates less than 1 cm yr^{-1}, it is questionable whether any steady-state crustal magma chamber can exist. For the Mid-Atlantic Ridge, the maximum half-width is probably 0.5–1.0 km, and any magma chamber would then be confined to depths equivalent to the lower part of layer 3 (i.e., 4–6 km below the seafloor). For the East Pacific Rise, a half-width of 3–5 km would be generally appropriate. Note that, if much hydrothermal cooling of the crust occurs, then these conductive thermal models are not applicable because cooling would be very much faster. This means that the size of any magma chambers would be much reduced.

Figure 9.21(b) shows a possible velocity model for a fast-spreading ridge, the East Pacific Rise at 9°N (half-spreading rate 6.1 cm yr^{-1}). This model was derived using the first thermal model discussed, P-wave velocities measured on the Semail ophiolite samples, a rate of change of seismic velocity with temperature $\mathrm{d}a/\mathrm{d}T$ of -8×10^{-4} km s^{-1} °C^{-1} and an arbitrary 1.5 km s^{-1} reduction in seismic velocity in the magma chamber to model the partial melt (compare with Fig. 9.13). The maximum width of this magma chamber is about 5 km, and its top is some 2 km beneath the seabed.

9.4.4 Hydrothermal circulation in young oceanic crust

Hydrothermal circulation through the ocean floor is one of the most important geochemical and geophysical processes on Earth. It plays a major role in controlling the chemistry of sea water, in the operation of subduction zones, in the growth of continents and in managing the Earth's heat budget. The loss of heat by hydothermal circulation is estimated to be one-third of the oceanic heat flux or

one-quarter of the global heat flux (Table 7.3). This cooling affects the plate thickness (Section 7.5) because the top few kilometres of plate are 'instantly' cooled. In effect, the cooling reduces the mean age of the oceanic crust before subduction. As the oceanic plate moves away from the ridge and cools, the hydrothermal systems slowly lose their vigour. Thus hydrothermal circulation is the cooling radiator of the front of the Earth's engine, like the splendid radiator on a Rolls-Royce, but it is much more than just a radiator.

Hot water is less dense than cold water. Any water that enters the porous, rubbly flank of a volcano heats up and rises, to be replaced by cold water from rain or from the sea. At mid-ocean ridges this process is intense, and water penetrates very deeply into the new oceanic crust. The system is so vigorous that huge, well-organized circulations become established, with cold water entering, warming as it penetrates to depths of perhaps several kilometres and then rising. The rising water may pass through well-defined channels to debouch back into the ocean at temperatures that are nearly critical, often approaching 400 °C at the ambient pressures of the ocean floor. These hydrothermal fluids are generally very acid, with a pH as low as 2. Enormous volumes of water pass through these systems, so great that in total a volume equal to the entire volume of the oceans passes through the ridge hydrothermal systems in about ten million years. The spacing between hydrothermal vent fields along the mid-ocean ridges is dependent upon the spreading rate (and hence the magmatic budget): it is predicted to be a linear function of the spreading rate. Data from the East Pacific Rise fit this prediction with frequent vent fields, but spacing on the slow-spreading Atlantic ridges is much more variable: 30–130 km on the Mid-Atlantic Ridge (11–38°N), but there is only one vent field along 600 km of the Reykjanes Ridge. Hydrothermal vent fields on the very-slow-spreading Southwest Indian Ridge are likely to be rare.

Off the ridge axis, hydrothermal circulation continues to take place as the young lithosphere cools. Here the circulation is probably confined to the uppermost permeable crust. Water temperatures and circulation rates are low for off-axis systems, but this circulation continues until the volcanic surface is sealed by sediments and the fractures become clogged. This average *sealing age* for the oceanic lithosphere is 65 ± 10 Ma. Passive off-axis hydrothermal circulation is not dramatic but may overall account for over 80% of the heat loss by hydrothermal circulation. Overall, the volume of water circulating through 1–65-Ma-old oceanic crust may be 10^2–10^3 times that circulating through young (<1 Ma) oceanic crust.

There are immediate chemical consequences of hydrothermal circulation. As sea water passes through rock, it exchanges cations such as sodium and calcium, dissolving some, moving or precipitating others. This changes the chemistry and the mineralogy of the rock, altering and metamorphosing it, depending on the depth and extent of hydration. Slightly away from the axis, the new basalt is already a much rearranged metamorphic rock, typically with its fine-grained minerals and with any glassy groundmass partly or substantially changed. Water

has been added; and, in many places, CO_2 has also been added, typically as calcium carbonate precipitated in the pores and fissures through which the water flowed. In the nineteenth century, geologists attempted to calculate the age of the Earth by measuring the amount of salt in the sea and dividing it by the amount of salt annually brought down by rivers from the land. We now know that that early calculation was invalid, because hydrothermal circulation is one of the principal controls on sodium in sea water.

The characteristic feature of a hydrothermal vent field on a mid-ocean ridge is the vent chimneys. The rising hydrothermal fluid is suddenly cooled as it re-enters the ocean (which is at about 4 °C, just above freezing) and minerals immediately precipitate from the fluid. Chimneys build up as the minerals precipitate. When the fluids venting from the chimneys have temperatures over ~330 °C, clouds of dark minerals are precipitated, giving rise to the name *black smokers*. *White smokers* are similar chimneys formed by somewhat cooler (200–330 °C) milky-white venting water. Warm (5–60°C) clear, shimmering water is discharged as a diffuse flow over surrounding areas several hundreds of metres square. Chimneys can grow very fast (over 2 m in five days has been documented on the Juan de Fuca Ridge), but they are fragile. Hydrothermal vent fields are dependent upon the vagaries of the underlying magmatic heat source, so fluid-circulation paths can change rapidly. However, long-lived hydrothermal systems have been discovered: U-series geochronology has shown that the TAG field at 26°N on the Mid-Atlantic Ridge has been episodically active for over 50 000 years.

Exotic communities of life have been discovered around these hydrothermal vents: giant tubeworms up to 3 cm in diameter and 3 m long, giant clams, crabs and unusual bacteria. These communities of unusual organisms were discovered by geophysicists in 1977, much to the surprise of biologists. The lives of these organisms do not depend directly on sunlight (as other life does) but on chemical energy from the hydrothermal system. They are not wholly disconnected from the rest of the biosphere, however, since they depend on the oxidation state of the environment, which is set by photosynthetic life. The bacteria live by exploiting the contrast between the relatively oxidized sea water (a product of surface photosynthesis) and the reduced fluids in the volcanic plumbing system – this process is known as *chemosynthesis*. The bacteria grow in abundance, covering surfaces of rock and fauna and are the first life to colonize any new vent. Many of the organisms have a symbiotic relationship with the bacteria, some even do without a digestive system – the bacteria live within them and meet all their energy requirements. Microbial life deep within the oceanic crust could prove to be an important aspect of hydrothermal circulation. Heat-tolerant *hyperthermophile* bacteria can withstand temperatures up to 113 °C and have been discovered in material from 3 km depth.

Interestingly, the gross features of the biogeographic distribution of the deep-sea hydrothermal vent fauna are controlled by plate tectonics. Of roughly 500 species and about 250 genera catalogued, over 90% of the species and half of

the genera are endemic, that is, found only at hydrothermal vents. Not only do the mid-ocean ridges act as the highway along which fauna can migrate/disperse between the hydrothermal vents (even though individual vents and vent fields are transient features), but also the timings of the past links between mid-ocean ridges (e.g., Galapagos, Gulf of California, Juan de Fuca Ridge and Mid-Atlantic Ridge) all seem to be partially reflected in the present-day distribution of fauna.

At subduction zones, the influence of the hydrothermal circulation is again felt and not only in the hydrothermal vent fields on spreading centres in back-arc basins. On a totally dry planet, the subducting crust (if subduction took place at all) would be without water. On Earth, the descending plate includes much water and CO_2 (even if virtually all the sediments lying on top of the oceanic crust are scraped off; see Section 10.2.2) because the slab is so heavily hydrated and carbonated. The water and CO_2 in the descending plate are eventually driven off and rise into the overlying mantle wedge. There the water plays a critical role in promoting melting, since wet rocks melt at much lower temperatures than do dry rocks. This melting (Section 10.2.1) is the process that leads eventually to the magmas that have formed the continents. Finally, a very small amount of water and CO_2 can descend with the plate into the deeper interior and can recirculate through the upper mantle over billions of years, eventually to rise again with the mantle under mid-ocean ridges. Even the very small amount of water that exists in ridge lavas has important geochemical and tectonic effects, such as promoting the melting process.

9.4.5 Seismic structure

The possibility of the existence of extensive, crustal infinite-onion magma chambers (Fig. 9.20) has led to a number of seismic experiments over the axial regions of the mid-ocean ridges in an attempt to delineate these chambers. Such magma chambers should be characterized by low seismic velocities and high attenuation.

Early models of the seismic structure of the oceanic crust (Figs. 9.11(e)–9.11(g)) had a very wide zone centred on the ridge axis in which normal layer-3 and upper-mantle velocities were not measured. However, subsequent experiments have revealed that the crust has the 'normal' oceanic structure except over a very narrow axial zone (about 20 km wide in some cases). Within this zone, layer 3 often appears to be absent or has a reduced velocity, and normal upper-mantle velocities are frequently not measured. Instead, the highest velocity is generally 7.1–7.6 km s^{-1}.

Slow-spreading ridges

Seismic experiments shot on the Mid-Atlantic and Reykjanes Ridges have yielded upper-mantle velocities at the ridge axis of 7.1–7.6 km s^{-1}. The final, well-constrained seismic-velocity models for one of these experiments, for which synthetic seismograms were used as part of the modelling procedure, are shown

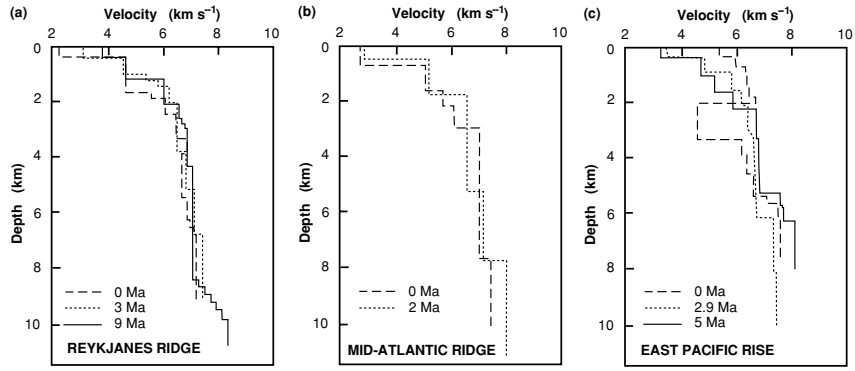


Figure 9.22. Seismic P-wave velocity structures for the axial regions of mid-ocean ridges: (a) Reykjanes Ridge at 60°N, (b) Mid-Atlantic Ridge at 37°N and (c) East Pacific Rise at 9°N. All these velocity structures were determined using synthetic seismograms. (After Bunch and Kennett (1980), Fowler (1976) and Orcutt *et al.* (1976).)

in Fig. 9.22. However, an experiment on the Mid-Atlantic Ridge at 23°N, just south of the Kane Fracture Zone, revealed near-normal oceanic crustal and upper-mantle velocities beneath a portion of the median valley. These differences are thought to reflect the time since the beginning of the last volcanic cycle. At 23°N, the crust is older and cooler and so is closer to the velocity of normal oceanic crust. None of these experiments gave results indicating the presence of a crustal low-velocity zone.

A seismic-reflection survey shot across the axis of the Mid-Atlantic Ridge at 37°N detected a small, shallow low-velocity zone in the crust. This appears to be a zone of extensive cracking at the injection centre associated with the extrusion of lavas at the seabed. Seismic experiments on the Juan de Fuca Ridge produced no evidence of crustal magma chambers.

Perhaps the most conclusive test of the presence of a partially molten zone can be provided by shear waves, which should be greatly attenuated when they cross such a zone. Arrays of three-component seismographs, with which it is possible positively to identify shear waves, have been used to monitor the microearthquake activity on the Mid-Atlantic Ridge (Fig. 9.23). There is no evidence for any special crustal attenuation of the shear waves since the microearthquakes produced large-amplitude shear waves. The crustal shear waves have travel times consistent with their having crossed the axial zone at depths down to about 5 km beneath the seabed. At these depths they would have had a wavelength of up to 1 km.

All this seismic and thermal evidence strongly suggests that no large, continuous crustal magma chamber (as in Fig. 9.20) is normally present beneath the axes of slow-spreading ridges, though the presence of small pockets of melt, up to perhaps 2 km in diameter, cannot be excluded.

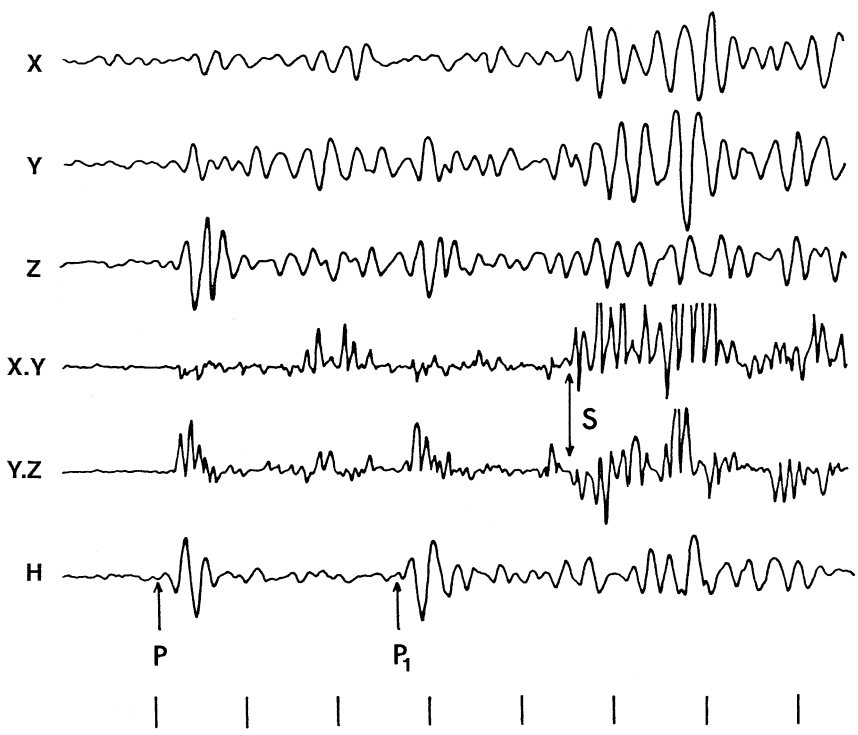


Figure 9.23. A record from an ocean-bottom seismometer (OBS) of an explosive charge 33 km away and on the opposite side of the axis of the Mid-Atlantic Ridge. X and Y (horizontal) and Z (vertical) denote seismometer components; H, hydrophone; time marks every second; X.Y (product of X and Y) and Y.Z (product of Y and Z) are displayed with a smaller scaling factor than X, Y and Z. P is the first-arriving P-wave; it has crossed the ridge axis in the lower part of layer 3; P_1 is the sea-surface (multiple) reflection of P; S is the first-arriving shear wave, which has also crossed the ridge axis in the lower crust. The arrival time of the shear wave can be picked most accurately from the products X.Y and Y.Z. (From Fowler (1976).)

However, results of detailed seismic and electromagnetic experiments at 57°N on the Reykjanes Ridge have delineated a significant crustal magma body: a lens of very-low-velocity material is present 2–3 km below the axial volcanic ridge. This lens, which is about 3 km across and 200 m thick, has a P-wave velocity of 3.3 km s^{-1}. Either side of the ridge axis the P-wave velocity at this depth is 6.5 km s^{-1}, which corresponds to the top of layer 3. Immediately above and below the lens the layer-2 and -3 velocities are 1.0–1.5 km s^{-1} lower than off-axis velocities at comparable depths. By 10 km from the ridge axis the crust is a normal 7 km thick and is underlain by a normal Moho. Results of electrical-resistivity studies delineate a small body with a very low resistivity (0.2 Ω m) coincident with the seismic low-velocity zone. Such resistivity values are appropriate for a basaltic melt. The location for these experiments was carefully chosen to be on a magmatically active section of the Reykjanes Ridge. It seems therefore that, during the most magmatic time of their tectono-magmatic cycle (Section 9.4.1), short-lived, transient melt bodies can be expected to underlie the axes of slow-spreading ridge segments.

Seismic wide-angle and normal-incidence experiments have imaged a melt body along at least 50 km of the axis of the Valu Fa Ridge in the Lau back-arc basin (half-spreading rate 3 cm yr^{-1}). There the magma chamber is at most 4 km wide and extends from 2.9 km below the seabed, through layer 3 to within

2 km of the Moho. The uppermost 250 m of the chamber has a very low P-wave velocity of 3.0 km s^{-1} over a 1.25-km-wide zone, implying the presence of interconnected melt. Beneath this thin and narrow lens, the P-wave velocity anomaly is -0.2 km s^{-1}, which is consistent with a 1% melt fraction. The main difference between this body and that imaged beneath the Reykjanes Ridge is in its lateral extent – this is a much narrower feature with smaller velocity anomalies.

Fast-spreading ridges

In contrast to the evidence from the slow-spreading ridges, seismic-refraction and -reflection experiments shot on the East Pacific Rise have given good images of crustal magma chambers on this fast-spreading ridge. At 9°N, where the half-spreading rate is 6 cm yr^{-1}, a zone with a velocity of about 5 km s^{-1} has been imaged at the ridge axis by seismic-refraction experiments using ocean-bottom seismometers (Fig. 9.22(c)). This low-velocity zone was not present in the models for 2.9- and 5.0-Ma-old crust. Along the ridge axis and at age 2.9 Ma, the mantle arrivals indicate low P-wave velocities of 7.6–7.7 km s^{-1}. However, by age 5 Ma, the mantle P-wave velocity is 8 km s^{-1}, and a more typical oceanic crustal layering appears to have developed. This model was determined from an analysis of the travel-time and distance data, as well as from waveform and amplitude studies using synthetic seismograms. Results of subsequent seismic experiments have confirmed the presence of an axial low-velocity anomaly (maximum -2 km s^{-1}) ~4 km in width, located between 1 and 3 km beneath the seabed.

Results of another refraction experiment on the East Pacific Rise (at 11° 20′N) using three-component ocean-bottom seismographs indicated that crustal P- and S-waves both crossed the ridge crest with high amplitudes and no apparent time delay. If magma is present within the crust at this location, these results indicate that it must be confined to narrow dykes or isolated bodies less than about 1 km in vertical extent.

Evidence for the existence of a narrow magma chamber beneath the East Pacific Rise axis at 21°N was provided by the differing attenuation of P- and S-waves (Fig. 9.24). The P-waves were well transmitted to a seismometer located on the ridge axis, but the S-waves were severely attenuated. In contrast, both P- and S-waves were well transmitted to a seismometer located 10 km from the ridge axis. This suggests that a narrow magma chamber is present beneath the ridge axis at this location. A schematic diagram of the crest of the East Pacific Rise at 21°N is shown in Fig. 9.25.

The ultra-fast East Pacific Rise (half-rate 7.5–8.1 cm yr^{-1}) south of the Garrett transform fault (13° 30′S) has also been the subject of seismic-reflection and -refraction experiments. A high-amplitude, reverse-polarity reflector 0.5 s below the seabed reflector is continuous over tens of kilometres along the ridge axis. This reflector, which is only about 1 km wide, coincides with the top of a crustal low-velocity zone, and appears to be from a narrow melt sill. The reflector is very similar in width and thickness to those imaged on other ridges but is shallower

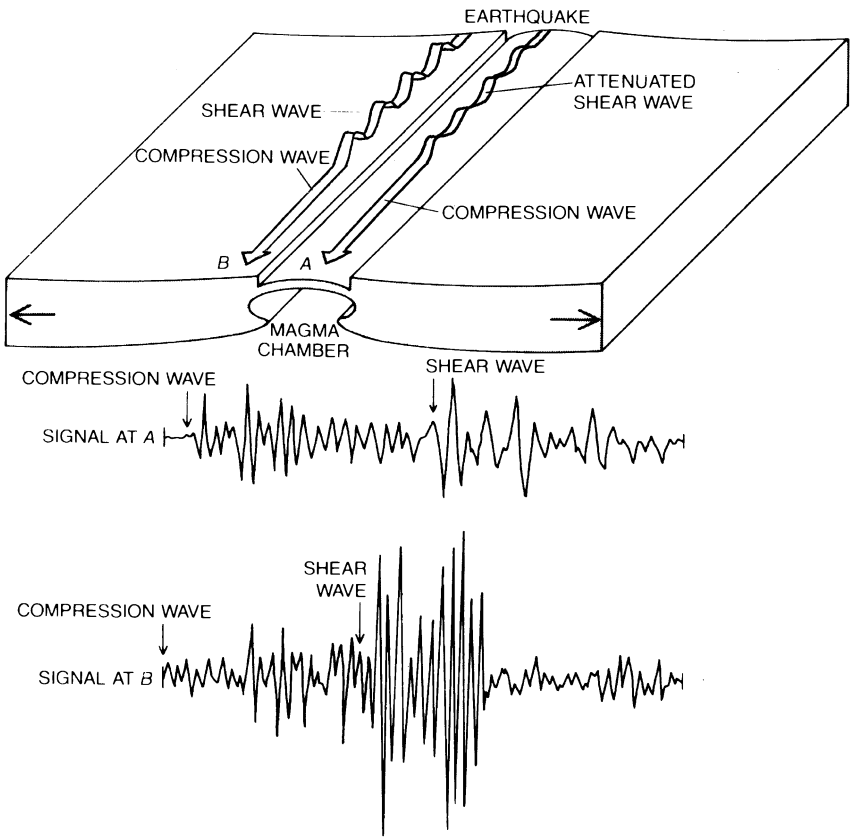


Figure 9.24. Compressional waves from a distant earthquake are well transmitted to an ocean-bottom seismometer A, situated on the East Pacific Rise axis, but the shear waves are attenuated. Both compressional and shear waves are well transmitted to a seismometer B, 10 km from the ridge axis. These observations suggest that a narrow magma chamber is present beneath the ridge axis. (From The crest of the East Pacific Rise, Macdonald and Luyendyk. Copyright © 1981 by Scientific American, Inc. All rights reserved.)

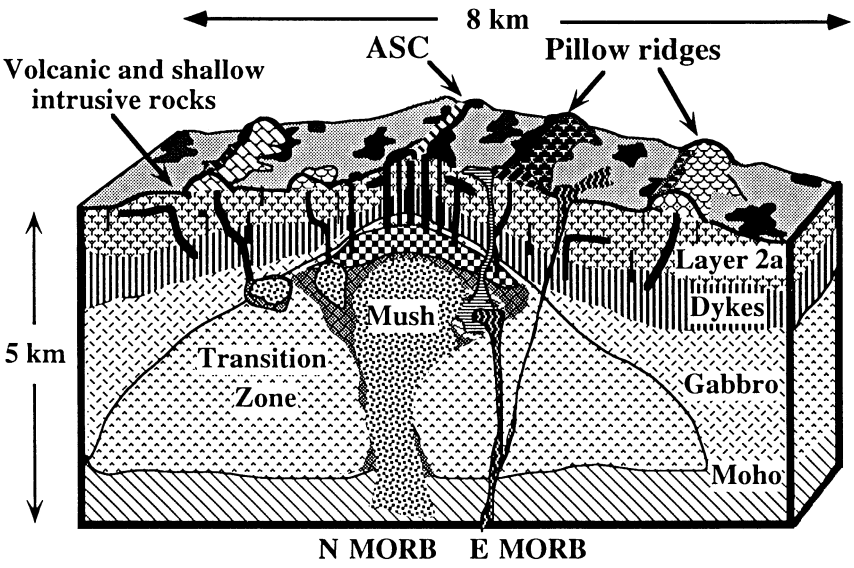


Figure 9.25. A schematic diagram of the northern East Pacific Rise. The checker- board pattern beneath the axial zone represents melt. Note the difference between on-axis eruptions of basalt and off-axis volcanism of pillow lavas from fissures and faults. The off-axis volcanism tends to be more enriched (E MORB) than the on-axis volcanism (N MORB) and is fed by separate conduits. (From Perfit *et al.* (1994).)

Example: seismic velocity and width of a low-velocity zone

The velocity of the lower crust in a model of the East Pacific Rise at 12°N is 7 km s^{-1}. Calculate the seismic P-wave velocity of a lower-crustal magma chamber if ray paths crossing the axis are delayed by 0.1 s and the width of the magma chamber is assumed to be (a) 1 km, (b) 3 km and (c) 6 km.

If w is the width of the chamber, the travel time in normal lower-crustal material is $w/7.0$ s. Ray paths crossing the axis are 0.1 s slower than this, so their travel time is $0.1 + w/7.0$ s. If the seismic velocity in the magma chamber is a km s^{-1}, then

$$\frac{w}{\alpha} = 0.1 + \frac{w}{7.0}$$

Rearranging this equation gives the seismic velocity in the chamber as

$$\alpha = \frac{7w}{w + 0.7}$$

Therefore, (a) for $w = 1$ km, $a = 4.12$ km s^{-1}; (b) for $w = 3$ km, $a = 5.68$ km s^{-1}; and (c) for $w = 6$ km, $a = 6.27$ km s^{-1}.

(9°N, 1.3–1.6 km; Valu Fa Ridge, ~3 km; Reykjanes Ridge, 2–3 km). It seems that the depth and accumulation of melt within the crust is controlled by a balance between the magmatic heat input and the loss of heat by hydrothermal cooling: the melt lens is likely to be shallow on fast-spreading ridges and deep on slow-spreading ridges. At 9°N and 13°S the Moho reflection extends laterally to within 2–3 km of the ridge axis, but determining exactly what happens beneath these low-velocity axial bodies is very difficult.

Several thousands of kilometres of common-depth-point seismic-reflection data have been obtained from the axial region of the East Pacific Rise in the search for the elusive crustal magma chamber. Data from a reflection line that crossed the ridge axis are shown in Fig. 9.26(a). A large-amplitude reflector about 3 km wide and centred on the ridge axis can be seen at about 0.6 s beneath the seabed. This reflector, which is continuous beneath the ridge axis for many tens of kilometres (Fig. 9.26(b)), has negative polarity, indicating that it arises at an interface with a strong negative velocity contrast. All these data are consistent with this arrival being a reflection from the lid of a magma chamber. The magma-chamber reflection was observed on about 60% of the 3500 km of reflection profile of this particular survey. The depth to the top of the magma chamber varies between about 1.2 and 2.4 km beneath the seabed and can be correlated with the depth of the ridge axis. Where the ridge axis is shallowest, the magma-chamber lid is shallowest; where it is deepest, the magma chamber is discontinuous or nonexistent, which shows the effect of magma supply on the axial topography. The presence of a magma chamber is also indicated by three expanding-spread reflection profiles, which were shot parallel to the East Pacific Rise at 13°N (Fig. 9.27). The profile shot along the ridge axis (ESP 9) shows a pronounced

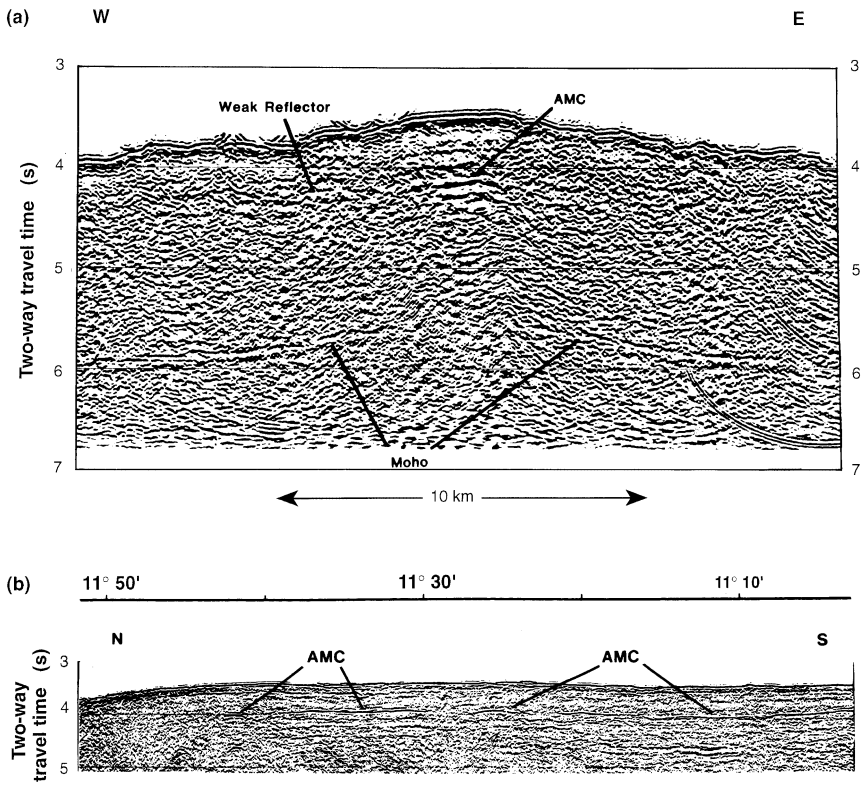


Figure 9.26. (a) A migrated CDP reflection record section across the axis of the East Pacific Rise at 9° 30′N. The first reflector at 3.5–3.9 s is the seabed reflection (water depth 2.6–2.9 km). The large-amplitude event beneath the ridge axis is interpreted as the reflection from the lid of an axial magma chamber (AMC). Its extension as a very weak reflector out onto the flanks of the ridge could perhaps represent the top of a frozen chamber or the change from extrusives to dykes and gabbros. Reflections arriving at the correct time for Moho reflections extend to within 2–3 km of the ridge axis. (b) A CDP reflection record section along 100 km of the axis of the East Pacific Rise (11–12°N). The AMC can be traced as a continuous event for tens of kilometres. The AMC reflection was not recorded north of 11° 50′N. See also Fig. 9.35(b). (From Detrick *et al.* (1987).)

shadow zone for crustal arrivals beyond 11 km range, indicating the presence of a crustal low-velocity zone. The profile shot 1.5 km off the ridge axis (ESP 11) shows some evidence for a thin low-velocity zone (note the discontinuity in first arrivals between 8 and 10 km). The profile 3.6 km away from the axis (ESP 13) shows no evidence for a low-velocity zone and has a typical oceanic velocity structure. Thus, at 13°N the low-velocity zone has a total width less than 5 km. The event at 6 s two-way time (Fig. 9.26(a)) is interpreted as being a reflection from the Moho. This reflection can be traced to within 2–3 km of the ridge axis, which again limits the maximum width of the magma chamber to about

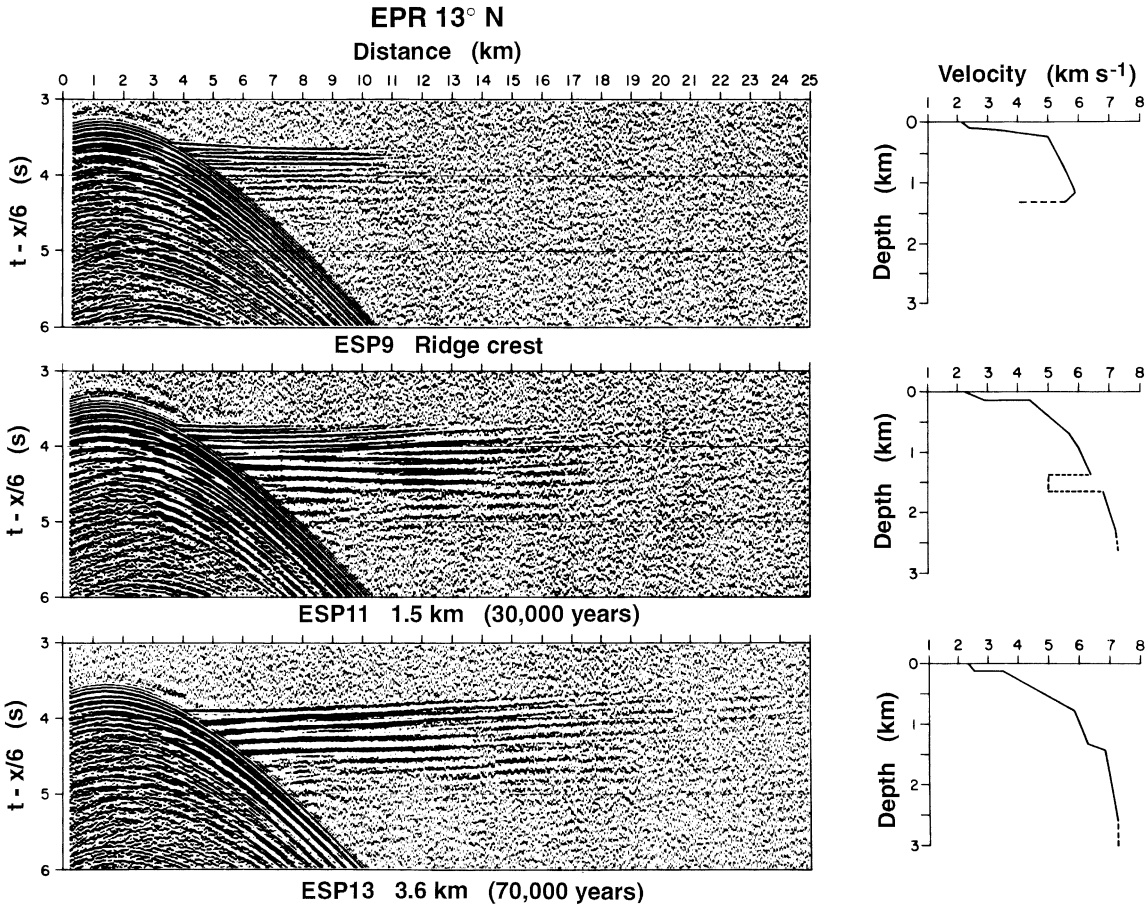


Figure 9.27. Reduced record sections for three expanding spread profiles (ESP) shot along and parallel to the East Pacific Rise axis at 13°N. ESP9 was shot along the ridge axis, ESP11 was shot 1.5 km away on 0.03-Ma crust and ESP13 was shot 3.6 km away on 0.07-Ma crust. To the right of each record section is the velocity–depth model determined from it by travel-time analysis. Note the low-velocity zone at the ridge axis, which produces a shadow zone on ESP9. The slight discontinuity in first arrivals at 8–10 km on ESP11 may indicate that a thin low-velocity zone is still present there. Beneath ESP13, however, there is typical oceanic crust, and clearly no low-velocity zone is present. (From Detrick *et al.* (1987).)

5 km. There is no evidence as to whether the magma chamber is literally that – a chamber filled with liquid magma – or whether it is a chamber filled with partially solidified crystal mush. There were no reflections from beneath the chamber to enable estimates of interval velocities or attenuation to be made. A schematic diagram of the magma chamber on the East Pacific Rise is shown in Fig. 9.28. This model magma chamber is very much smaller than many that have been proposed as results of studies of ophiolites. Figure 9.29 indicates where melt and crystal mush are located beneath a fast-spreading ridge.

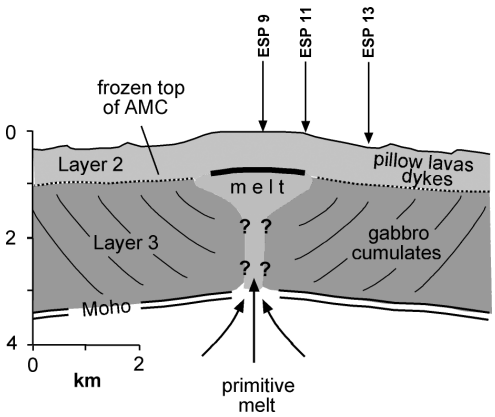


Figure 9.28. A schematic diagram of the magma chamber beneath the East Pacific Rise, based on the results of the reflection profiling illustrated in Figs. 9.26 and 9.27. (From Detrick *et al.* (1987).)

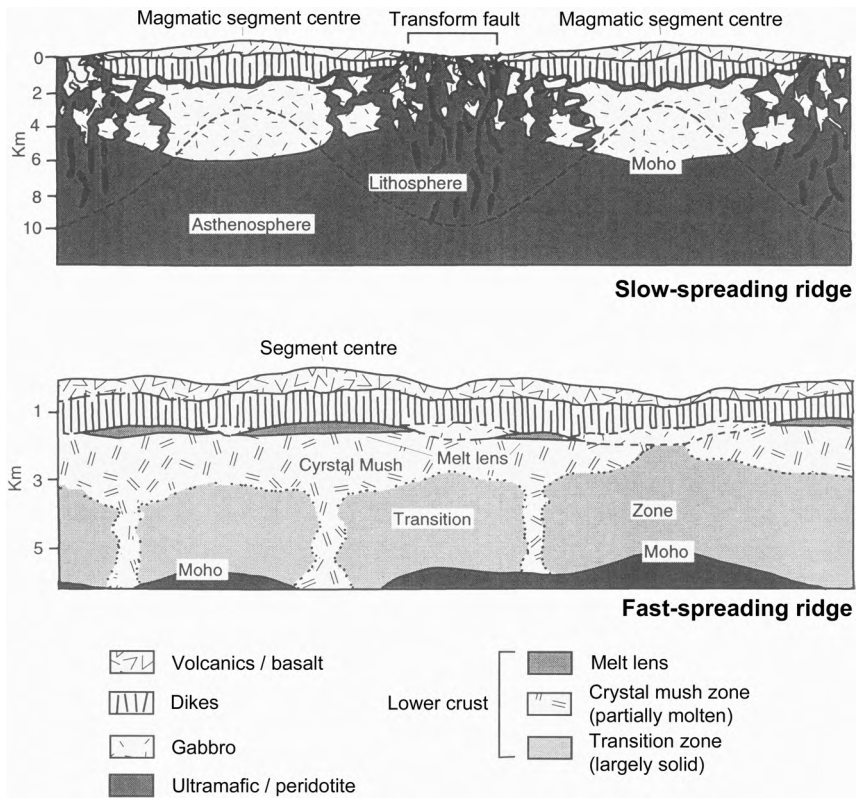


Figure 9.29. Sections along the axis of a slow-spreading ridge and along the axis of a fast-spreading ridge illustrating the differences in crustal structure at segment centres and ends. (From Dilek *et al.*, Structure of modern oceanic crust and ophiolites and implications for faulting and magmatism at oceanic spreading centres, Geophysical Monograph 106, 219–65, 1998. Copyright 1998 American Geophysical Union. Reproduced by permission of American Geophysical Union, after Cannat *et al.* (1995); and Sinton and Detrick, Mid-ocean ridge magma chambers, *J. Geophys. Res.*, **80**, 4037–42, 1992. Copyright 1992 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

The preceding discussion of some of the seismic experiments carried out over the East Pacific Rise shows clearly that, just as is indicated for the Mid-Atlantic Ridge, the East Pacific Rise is not such a simple one-dimensional feature as was first thought. Although a magma chamber seems to be present in some locations, it apparently varies in size and extent along the axis and is not a continuous feature. In contrast, the Moho discontinuity appears to be continuous under all oceanic crust except for a very narrow zone some few kilometres in width along the ridge axis.

9.5 Transform faults

9.5.1 Geometry

Transform faults are plate boundaries at which material is neither created nor destroyed. The relative motion between the two plates is parallel to the strike of the fault. By referring to Sections 2.3 and 2.4, you will appreciate that, since lines of constant velocity are small circles about the pole of rotation between two plates, transform faults are arcs of small circles. From this fact, the relative motions of the plates can be determined, as discussed in Chapter 2. If the relative motion between the two plates is not exactly parallel to the fault, there is a small component of convergence or divergence. A transform with a small component of divergence is called a *leaky transform fault*. An example of a leaky fault is the plate boundary between Eurasia and Africa, from the Azores Junction on the Mid-Atlantic Ridge eastwards towards Gibraltar (Fig. 9.30). Earthquake fault-plane solutions for this boundary near the Azores illustrate the small amount of extension occurring there. Transform faults that have a significant amount of compression in addition to strike–slip are referred to as transpressional. An example of such a fault is the plate boundary through the South Island of New Zealand.

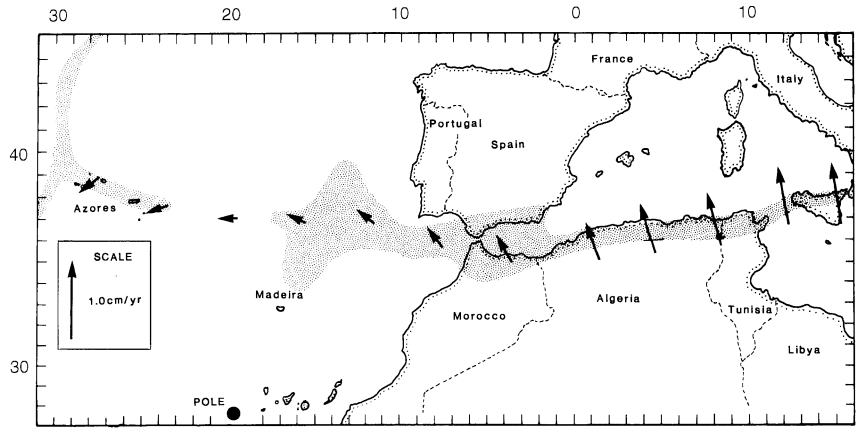


Figure 9.30. Slip vectors for the motion of Africa relative to Eurasia along their plate boundary. The length of each arrow indicates the magnitude of the velocity. The solid circle denotes the pole position. Shading indicates the regions of most intense seismicity. (From Anderson and Jackson (1987).)

Table 9.4 *Details of some transform faults*

Name	Location[a]	Offset (km)	Approximate age contrast (Ma)	Approximate topographic relief (km)	Approximate truncated lithosphere thickness[b] (km)	Present slip rate (cm yr^{-1})
Charlie Gibbs	MAR 52°N	340	20	2.5	50	2.5
Kurchatov	MAR 40°N	22	2	1.5	15	2.5
A)	FAMOUS area:	20	2	1.5	15	2
B)	MAR 37°N	22	2	1	15	2.2
Oceanographer	MAR 35°N	130	12	4	40	2
Atlantis	MAR 30°N	70	6	3	27	2.5
Kane	MAR 24°N	150	11	3	35	2.8
Vema	MAR 11°N	320	20	3	50	3
Romanche	MAR equator	935	60	4	85	3
Blanco	JdFR 44°N	350	10		35	6
Mendocino	JdFR 40°N	1150	27	2	55	2–4
Tamayo	EPR 23°N	80	3	1	20	6
Clipperton	EPR 10°N	85	1.5	1	14	11
Gofar	EPR 5°S	180	2	2	15	15
Wilkes	EPR 9°S	150	2	1	15	15
Andrew Bain	SWIR 30°E	500	–	4	–	1.6
Atlantis II	SWIR 57°E	200	20	3	50	1.6
Udintsev	PAR 58°S	300	10	1	35	6
Eltanin system	PAR 55°S	900	30–34		60	6

[a] MAR, Mid-Atlantic Ridge; JdFR, Juan de Fuca Ridge; EPR, East Pacific Rise; SWIR, South West Indian Ridge; PAR, Pacific–Antarctic Ridge.

[b] From Eq. (7.65).

Table 9.4 lists some of the main transform faults on the mid-ocean-ridge system. Some transform faults, such as the San Andreas Fault in California, occur on land, but they usually occur between ridge segments at sea. Many Indian and Pacific transform faults are not listed because data are not yet available for them; the details of much of the ocean bed remain unknown. Owing to financial and logistical considerations, the most easily accessible parts of the ocean bed are the most studied and best understood. The term transform fault is correctly applied only to the active region of a fault (i.e., the part between the offset ridge segments). Extending beyond the active region is a zone of fracturing called a *fracture zone*, or an 'inactive fossil transform fault'; examples are the Kurchatov Fracture Zone and the Clipperton Fracture Zone (Fig. 9.31, Plate 21). There is no horizontal slip motion on the fracture zones, but slight vertical adjustments do occur.

Transform faults range from very long to very short. The equatorial transform faults in the Atlantic are many hundreds of kilometres long. However, the

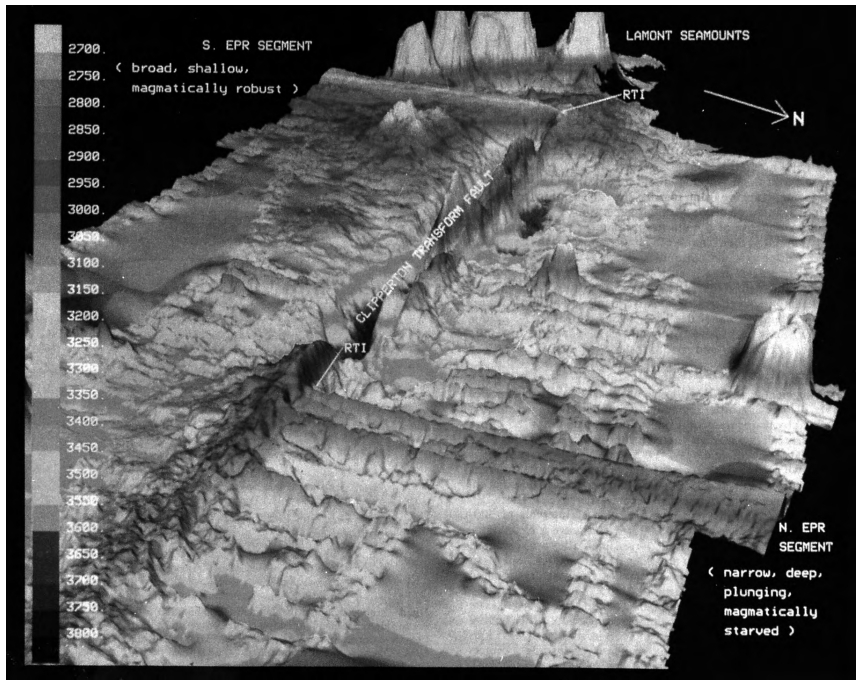


Figure 9.31. Shaded bathymetry of the Clipperton transform fault (Table 9.4) on the East Pacific Rise at 10°N. The image is 100 km (north–south) × 175 km (east–west). Depths are metres below sea level. The transform fault, a straight narrow cleft that cuts through the rugged transform valley, has had this configuration for 9 Ma. The northern ridge segment is starved of magma and deepens steadily for 70 km towards the transform. In contrast, the southern ridge segment is elevated and is underlain by a magma chamber right up to the transform. Colour version Plate 21. (Reprinted with permission from *Nature* (Macdonald *et al.* 1998, *Nature*, **335**, 217–25) Copyright 1988 Macmillan Magazines Ltd.)

more detailed the bathymetric, magnetic and reflection surveys become, the more numerous small discontinuities are seen to be. In the FAMOUS area at 37°N on the Mid-Atlantic Ridge, the ridge segments are 20–60 km in length and are offset by small discontinuities about 20 km in length. Even on this small scale, the regular geometrical pattern is maintained.

Figure 9.32 illustrates a transform fault between two ridge segments. In this example, the offset is 200 km. The fault juxtaposes materials that differ in age by 20 Ma. Note that, although the spreading rate is 1 cm yr^{-1}, the slip rate along the fault (which is twice the spreading rate) is 2 cm yr^{-1}. This slip motion occurs on the fault only between the two ridge segments. This is demonstrated by the location of earthquake epicentres: outside the active transform faults, seismic activity is negligible (Fig. 9.33).

Even though transform faults are active only between the two offset ridge segments, they remain major features on bathymetric, gravity (Fig. 5.4) and magnetic-anomaly maps outside the active zone because of the constant age contrast across the fault. The majority of Atlantic fracture zones are bathymetric and gravity features from America to Africa and start at the continental margin. This observation was puzzling to Earth scientists before the development of plate tectonics, and the problem of *transcurrent faults* (as they were then called) was the subject of much debate. Before seafloor spreading was understood, the fault shown in Fig. 9.32 would have been understood as a right lateral offset of the ridge

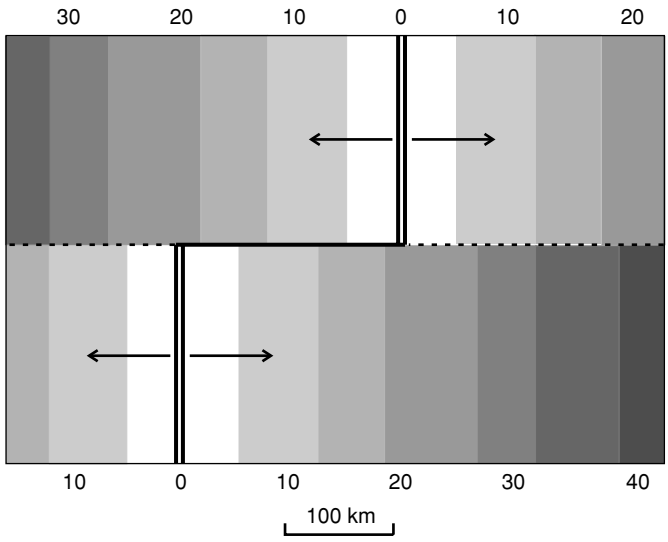


Figure 9.32. An idealized map of a transform fault offsetting two ridge segments. Numbers are ages in Ma. Age provinces are progressively shaded. Note that the motion on the fault occurs only between the ridge offsets and is sinistral (to the left looking across the fault) not dextral (to the right) as it would have been described in the pre-plate-tectonic, pre-magnetic-stripe literature.

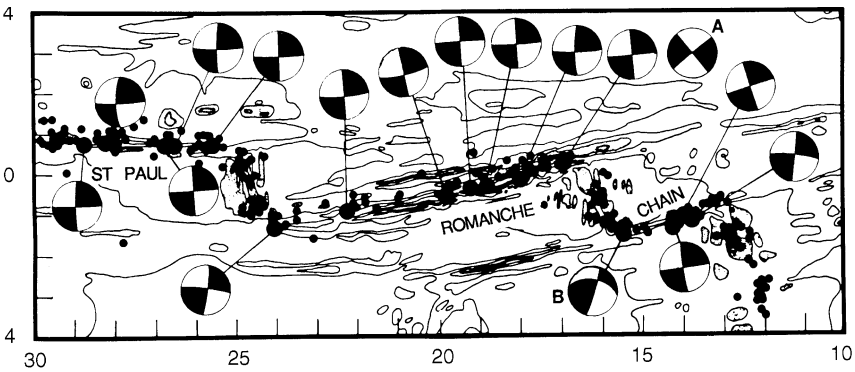


Figure 9.33. A bathymetric map of the St Paul, Romanche and Chain transform faults in the equatorial Atlantic (Table 9.4). Dots indicate earthquake epicentres. The large dots are those for which the fault-plane solutions are shown. Notice how closely the epicentres follow the ridge axis and transform faults. Because the foci were all shallow (less than 5 km), and the depth range of faulting is about 0–10 km, the foci were above (shallower than) the 600 °C isotherm. (From Engeln *et al.* (1986).)

axis. Had this been the case, the entire fault would be active, not just the portion between the two offset ridge axes, and the slip motion on the fault would be in the opposite direction to that shown in Fig. 9.32. This gave rise to the question 'What happens to the fault on the continent?' or 'Where do transcurrent faults end?' Plate tectonics neatly solved the problem.

There are several reasons why there are transform faults offsetting ridge axes. First, the major faults are determined by the geometry of the initial break between the two continents; this break occurs along zones of weakness, which are probably partly associated with old geological structures. The break need not be a straight line and may be quite irregular. Transform faults tend to develop in locations where the relative motion is approximately parallel to the break and are necessary in order to accommodate this geometry with the pole of rotation. Thus, the preferred spreading-geometry configuration consists of straight ridge segments offset by orthogonal transform faults. That segmentation of slow-spreading ridges

is generally long-lived and may be inherited from the earliest stages of continental rifting is supported by three-dimensional modelling of melting and upwelling in a mantle with non-uniform rheology: model segments persist for tens of millions of years. Second, smaller offset discontinuities probably form to accommodate the change in relative plate motion with distance from the rotation pole and the magmatic supply regime. In this way, ridge segments with constant spreading rates are created.

The orthogonal transform-fault/ridge-segment geometry, while being *normal*, is not universal. The orthogonal geometry develops provided that the resistance to slip at a transform (shearing yield stress) is less than the resistance to spreading (tensional yield stress) at the ridge axis. At slow-spreading ridges where the lithosphere is strong and thick close to the ridge axis (Eq. (7.65)), it may be difficult for transform faults to develop, or for the ridge to adjust to any change in the pole position. Oblique spreading is more likely to occur on slow-spreading ridges than on fast-spreading ridges. Both the Reykjanes Ridge and the Mohns Ridge (Fig. 2.2) are spreading obliquely, in the case of the Mohns Ridge at an angle of 30–40° to the ridge. The spreading direction along the Mohns Ridge changed 27 Ma ago with a major reorganization in the North Atlantic. Since then the Ridge has kept its old trend in the form of an axial valley. Within the valley, however, there is a series of *en échelon* horsts and grabens (perpendicular to the current spreading direction) linked by orthogonal transfer zones. The thicker crust on the Reykjanes Ridge (Section 9.3.2) apparently affects the relative yield stresses, giving rise to oblique spreading without transform faults. There are no transform faults between 57°N and Iceland (a distance of over 900 km) while to the south, where the crust is thinner, the normal orthogonal fault/ridge geometry prevails.

9.5.2 Topography and crustal structure

Transform faults are major bathymetric features, visible on magnetic-anomaly maps and marked by earthquake epicentres. Figure 9.32 and Problem 9 suggest that transform faults might ideally be marked by a single vertical fault; but, instead, active transforms are anomalous linear valleys, usually less than 15 km wide, bounded by inward-facing scarps and depressed by from 1 to 5 km. In the idealized single-fault model, the oceanic crust is exposed along the fault wall and easily accessible to geologists for sampling by dredging and drilling. However, in practice, instead of a single fault scarp, there are usually many, each typically with a throw of only a few hundreds of metres or less.

Some of the jargon used in the literature to describe the various provinces along transform faults is illustrated in Fig. 9.34, a schematic diagram of the bathymetry and faulting associated with a slow-slipping transform. The valley along the line of a transform fault is the transform valley. The transform domain or transform tectonized zone (TTZ) is the region which has been affected by deformation

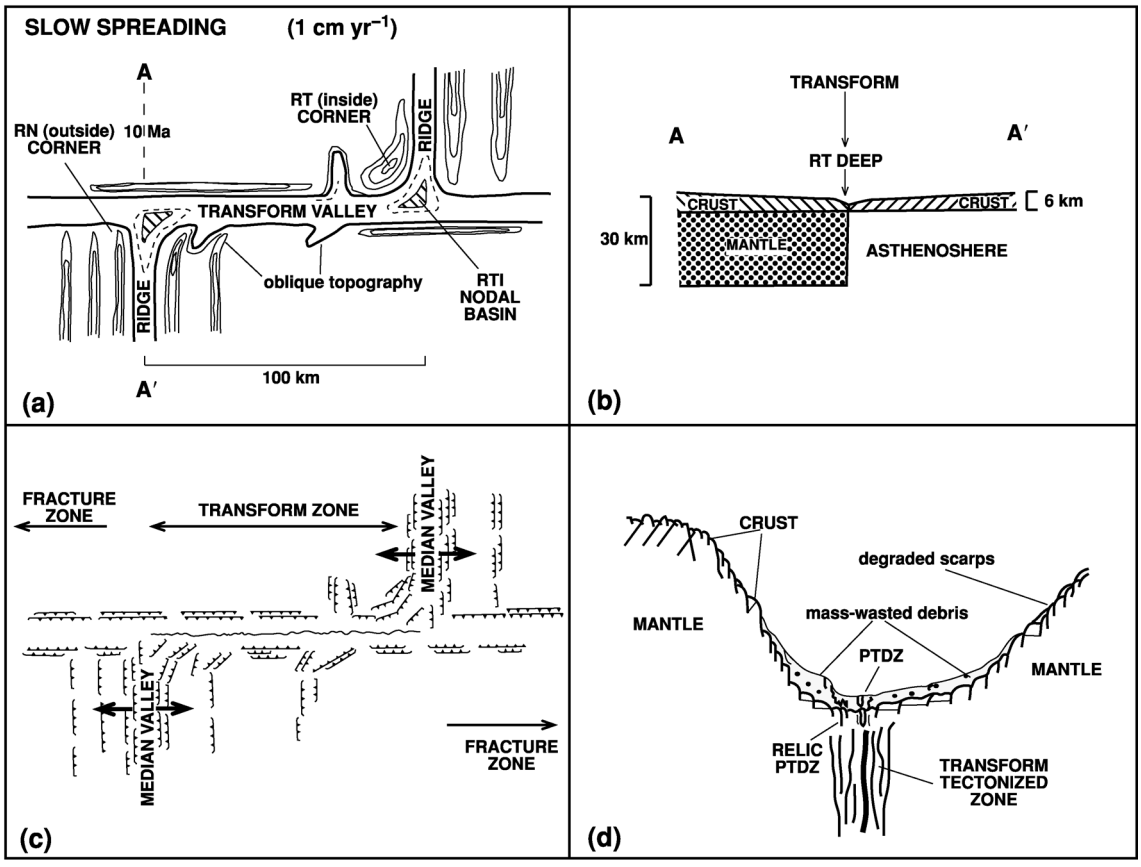


Figure 9.34. A slow-slipping transform fault with an offset of 100 km (10 Ma): (a) topography, (b) schematic vertical cross section, (c) structural grain and (d) geological cross section. A fast-slipping transform fault with an offset of 100 km (0.8 Ma): (e) topography, (f) schematic cross section and (g) structural grain. (After Fox and Gallo (1984) and Goud and Karson (1986).)

associated with the ridge offset (i.e., the shear zone). The zone of interconnecting faults that are active at present (the active fault zone) is called the principal transform displacement zone (PTDZ). At the ridge–transform-fault intersection (RTI), there is a deep, closed-contour depression called a nodal basin. The junctions of the median-valley walls with the transform-valley walls are called the ridge–transform (RT) and ridge–non-transform (RN) corners or alternatively the inner corner (IC) and outer corner (OC). On slow-spreading ridges the inside corner generally has elevated topography while the outside corner is deeper. This can be interpreted as the IC being the elevated footwall of a large detachment fault (Section 9.4.1) in a paired half-graben model for the ridge axis.

On fast-spreading ridges, transform faults need not occur as distinct faults at regular intervals; instead, they are less frequent and often extend over wide shear

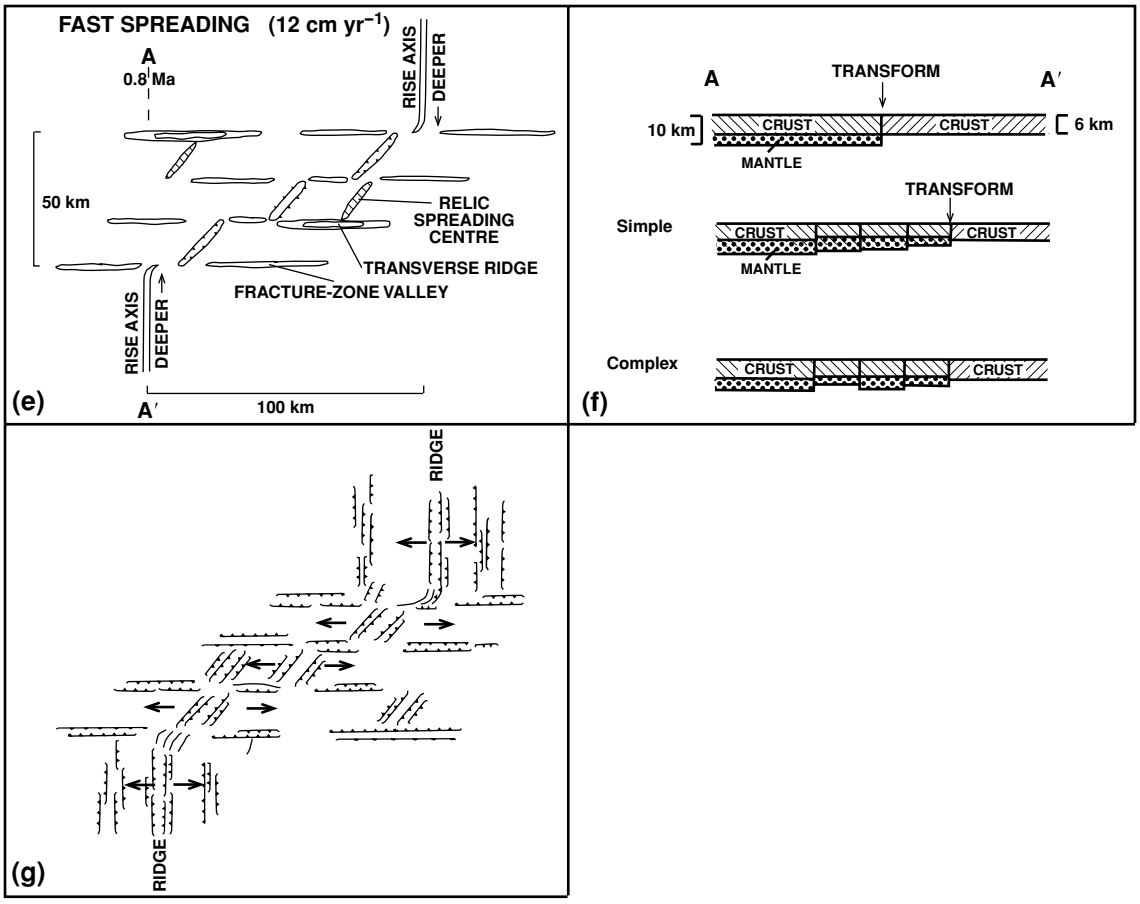


Figure 9.34. (*cont.*)

zones, perhaps up to 50 km wide, and have a series of ridge–fault segments. It should not be surprising that fast-slipping transforms operate somewhat differently from slow-slipping transforms. There is a difference of up to a factor of ten in their slip rates, which means that their thermal structures are very different. Figure 9.34 shows the features which might reasonably be expected to result at a large-offset, fast-slipping transform fault. It is clear that there is a wide shear zone made up of a series of ridge–fracture segments. The East Pacific Rise between the Nazca and Pacific plates, the fastest-spreading ridge segment in the world, is offset by a number of long transform faults. Results from various bathymetric and tectonic studies of these faults show that they can be best described as 20–150-km-wide shear zones of complex topography.

Work on the East Pacific Rise has resulted in more new terminology. Although the terms *transform fault* and *fracture zone* are used, respectively, for the active and inactive portions of faults, the terms *deviations in axial linearity* (DEVAL)

Table 9.5 *Segmentation of mid-ocean ridges*

	First order	Second order	Third order	Fourth order
Fast-spreading ridges	*Transform fault*	*Overlapping spreading centre*	*Overlapping spreading centre*	*DEVAL*
Segment length	600 ± 300 km	140 ± 90 km	50 ± 30 km	14 ± 8 km
Offset	>30 km	2–30 km	0.5–2 km	<1 km
Segment lifetime	>5 Ma	0.5–5 Ma	0.01–0.1 Ma	0.0001–0.01 Ma
Depth anomaly	300–600 m	100–300 m	30–100 m	0–50 m
Geochemical anomaly	Yes	Yes	Usually	Sometimes
Off-axis anomaly	Fracture zone	V-shaped discordant zone	None	None
Slow-spreading ridges	*Transform fault*	*Oblique shear zone, depressions*	*Inter-volcano gap*	*Intra-volcano gap*
Segment length	400 ± 200 km	50 ± 30 km	15 ± 10 km	7 ± 5 km
Offset	>30 km	2–30 km	0.5–2 km	<1 km
Segment lifetime	>5 Ma	0.5–30 Ma	?	?
Depth anomaly	500–2000 m	300–1000 m	50–300 m	0–100 m
Geochemical anomaly	Yes	Yes	Usually	?
Off-axis anomaly	Fracture zone	Broad irregular valley	Faint or None	None

Source: After Macdonald (1998).

and *overlapping spreading centre* (OSC) are also used. DEVAL refers to very small offsets or deviations, which are not regular transform faults but sometimes mark a petrological segmentation of the ridge; OSC refers to a phenomenon that occurs on fast-spreading ridges when the transform fault does not mark the abrupt termination of both ridge segments, but, instead, the segments overlap slightly (Figs. 9.35(a) and 9.35(b)).

Discontinuities along the mid-ocean ridges divide the ridge up into a number of discrete segments: a transform fault is a rigid plate boundary that offsets the ridge axis normally by over 30 km and partitions the ridge into distinct tectonic units on a length scale of hundreds of kilometres. However, the segmentation of the mid-ocean ridges also occurs on shorter scales described as second, third and fourth order. The main features of mid-ocean-ridge segmentation are summarized and the terminology put into context in Table 9.5 and they are illustrated schematically in Fig. 9.35(c). Transform faults represent the first-order scale of mid-ocean-ridge discontinuity: they persist for many millions of years and are easily identified by their large depth anomalies and by magnetic-anomaly patterns. Second-order discontinuities partition the ridge on a length scale of ∼50 km or more, with shorter offsets of the ridge axis. The orthogonal linearity of the ridge/transform system

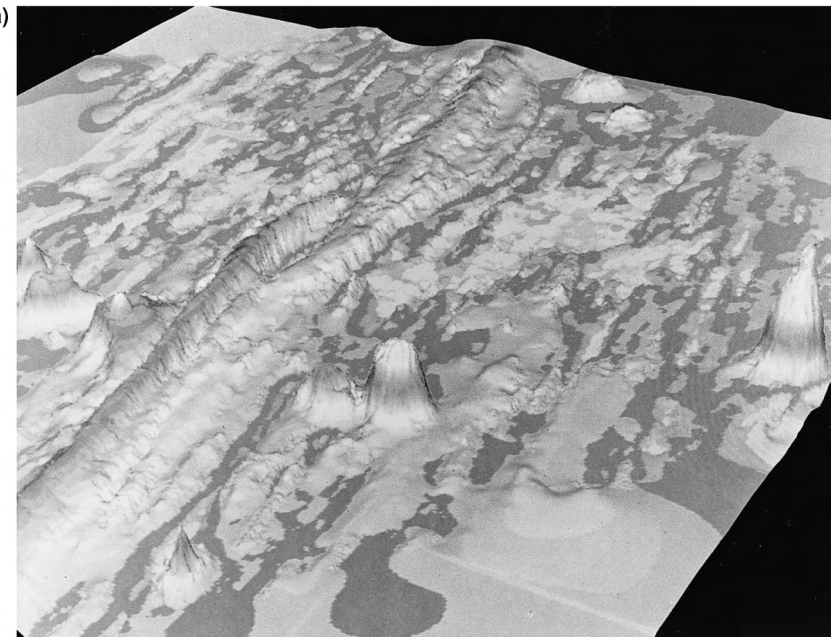
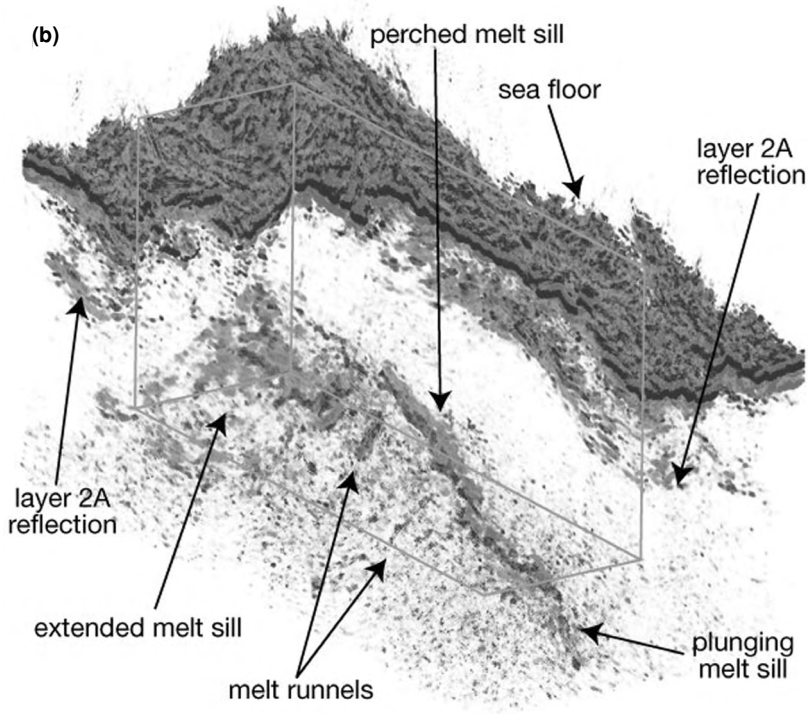


Figure 9.35. (a) A sonar image of the overlapping spreading centres on the East Pacific Rise at 9°N. There is no transform fault between the two tips of the ridge segments which extend past each other. The deeper region between the two spreading centres is termed the overlap basin. Such overlapping spreading centres are temporary phenomena caused by misalignment of tensional cracks in neighbouring ridge segments. As the magma rises episodically, the tips of the ridge segments move back, forth and laterally. (Reprinted with permission from *Nature* (Macdonald *et al*. 1988, *Nature*, **335**, 217–25). Copyright 1988, Macmillan Magazines Ltd.)

(b) A three-dimensional seismic reflection survey across the eastern limb of the overlapping spreading centre on the East Pacific Rise at 9° 03′N, with north to the top and south to the bottom of the image. A rectangular box (orange frame) has been cut away to reveal the reflectivity at depth. The top surface (shown red) is the seafloor. Reflections from layer 2A are blue. Reflections from the magma chamber are orange and blue–green. The image shows approximately 8 km × 14 km. The two-way time from seafloor to base of image is ~1.5 s. Bodies of melt underlie the eastern limb of the ridge and apparently pond beneath the overlap basin. The melt bodies may be only 100 m thick. Imaging highly heterogeneous structures is extremely difficult and the fine detail of melt transport beneath the OCS is complex, but there must be a robust vertical magma supply from the underlying mantle. Colour version Plate 22. Reprinted with permission from *Nature* (Kent, G. M. *et al*. 2000, *Nature*, **406**, 614–18) Copyright (2000) Macmillan Magazines Ltd.)

(c) The hierarchy of the segmentation of mid-ocean ridges. Ridge segments from first to fourth order are S1, S2, S3 and S4. Discontinuities from first to fourth order are D1, D2, D3 and D4. (Reprinted with permission from Macdonald, K. C., Sheirer, D. S. and Carbotte, S., Mid-ocean ridges: discontinuities, segments and giant cracks, *Science*, **253**, 986–94. Copyright 1991 AAAS.)



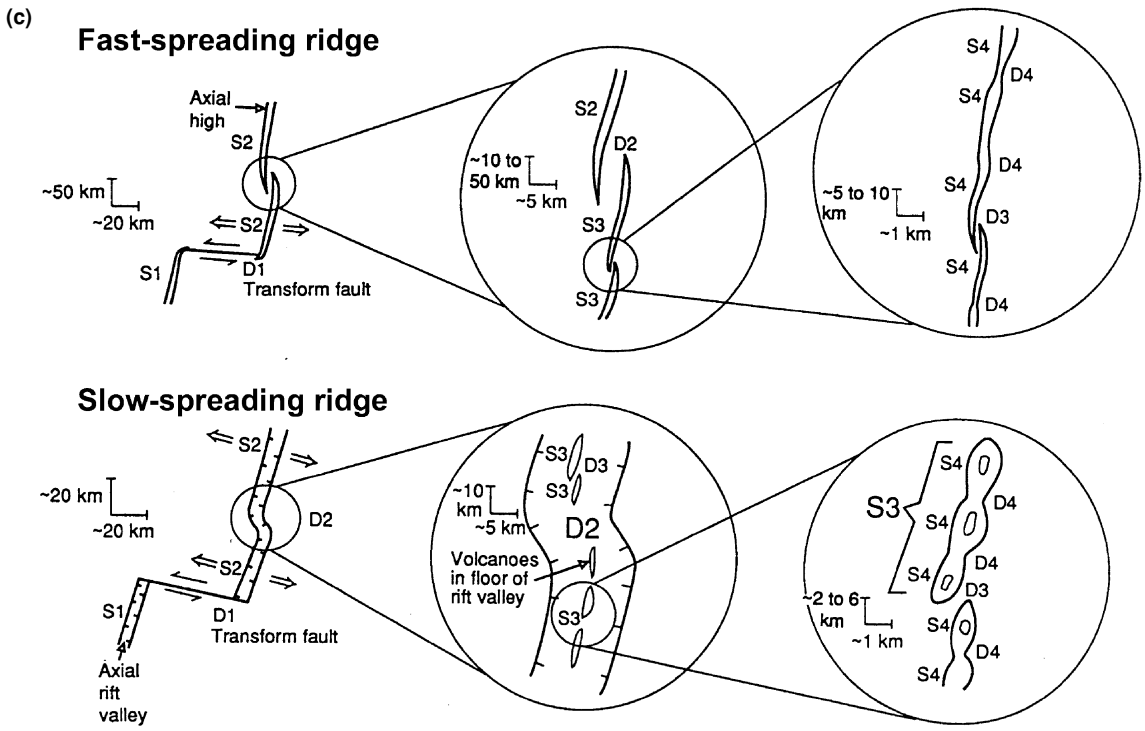


Figure 9.35. (*cont.*)

need not be maintained at second-order discontinuities. On intermediate- and fast-spreading ridges, second-order discontinuities include overlapping spreading centres, whereas at slow spreading ridges they are represented by oblique offsets of the ridge axis and deep depressions in the seabed. Third-order discontinuities at intermediate- and fast-spreading ridges are also associated with overlapping spreading centres, but with much reduced offset and depth anomalies. Fourth-order discontinuities have very small offsets in the summit graben or deviations in the axis; they rarely have a depth anomaly. On slow-spreading ridges these fine-scale third- and fourth-order discontinuities are short-lived features, marked by gaps between and within individual volcanoes.

While classification of ridge segments and discontinuities into four orders is very helpful, it should be appreciated that this is not a physical constraint but a description – reality is complex and time-dependent.

Detailed studies of the Oceanographer Transform Fault on the Mid-Atlantic Ridge have shown that its active region is centred on the deepest part of the valley and is confined to a zone a few hundred metres to a few kilometres wide. Figure 9.36(a), a seismic-reflection profile across the Vema transform fault at 6°N on the Mid-Atlantic Ridge, shows a narrow region of faulting, which marks the boundary between the African and South American plates. The deformation in the sediments is confined to a narrow zone less than a kilometre across.

Several detailed seismic-refraction experiments have been shot along and across large- and small-offset Atlantic transform faults, the results of which indicate that their crustal structures are anomalous. Instead of normal 6–8-km-thick oceanic crust, the fracture-zone crust is considerably thinner, generally 3–5 km thick but in places only 2–3 km thick. The thin crust is confined to a region less than 10 km wide. In addition, lower than normal compressional crustal velocities are characteristic, and layer-3 velocities are frequently not observed. Figures 9.37(a)–9.37(e) show seismic models crossing transform faults at slow- and very-slow-spreading ridges along the strike of the median valley. It is clear that, over some tens of kilometres, the crustal thickness decreases towards the transform fault, in addition to the very thin crust occurring in the transform fault itself. The Kane and Oceanographer transform faults have been studied in some detail by submersibles and ODP drilling has taken place on Atlantis Bank adjacent to the Atlantis II fracture zone. It is clearly documented that rocks usually associated with the lower crust and upper mantle (i.e., gabbroic and ultramafic rocks) outcrop on escarpments at transform faults. This is hard to explain if the crust is of normal thickness and again suggests that transform-fault crust is thinner than normal.

A detailed wide-angle seismic-reflection survey has been made over the Blake Spur fracture zone (a small western Atlantic fracture zone with about 12 km offset on 140-Ma-old lithosphere). Normal oceanic crust is present up to 10 km from the fracture zone, but the seismic-velocity structure in a zone 10–20 km wide centred on the fracture zone is anomalous (Fig. 9.37(f)). Beneath the sediments there is

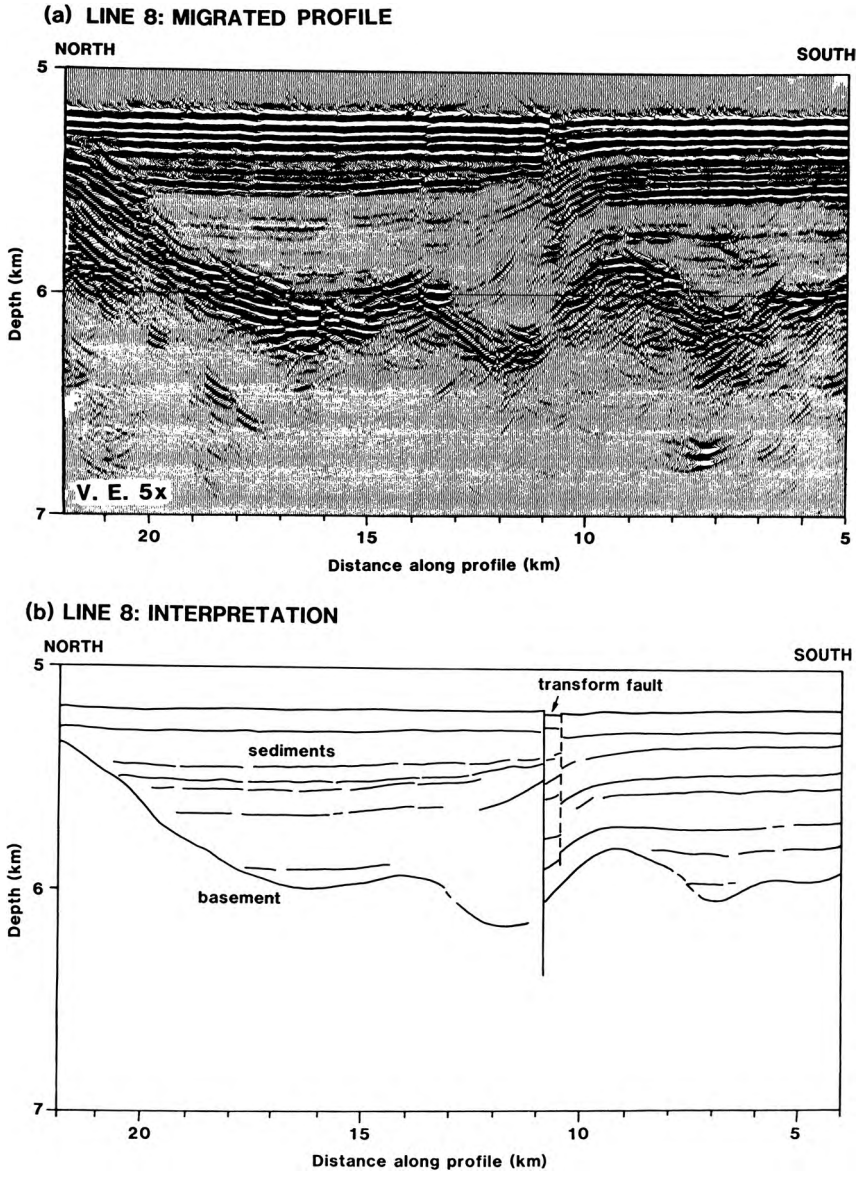


Figure 9.36. (a) A migrated seismic-reflection profile across the Vema Transform Fault on the Mid-Atlantic Ridge at 6°N, approximately midway between the two ridge segments. (b) Interpretation of the reflection profile shown in (a). Note that the sediments are undisturbed except in a narrow zone about 0.5 km wide. On the seabed, the position of the active fault is marked by a narrow groove 30 m deep and about 500 m wide. (From Bowen and White (1986).)

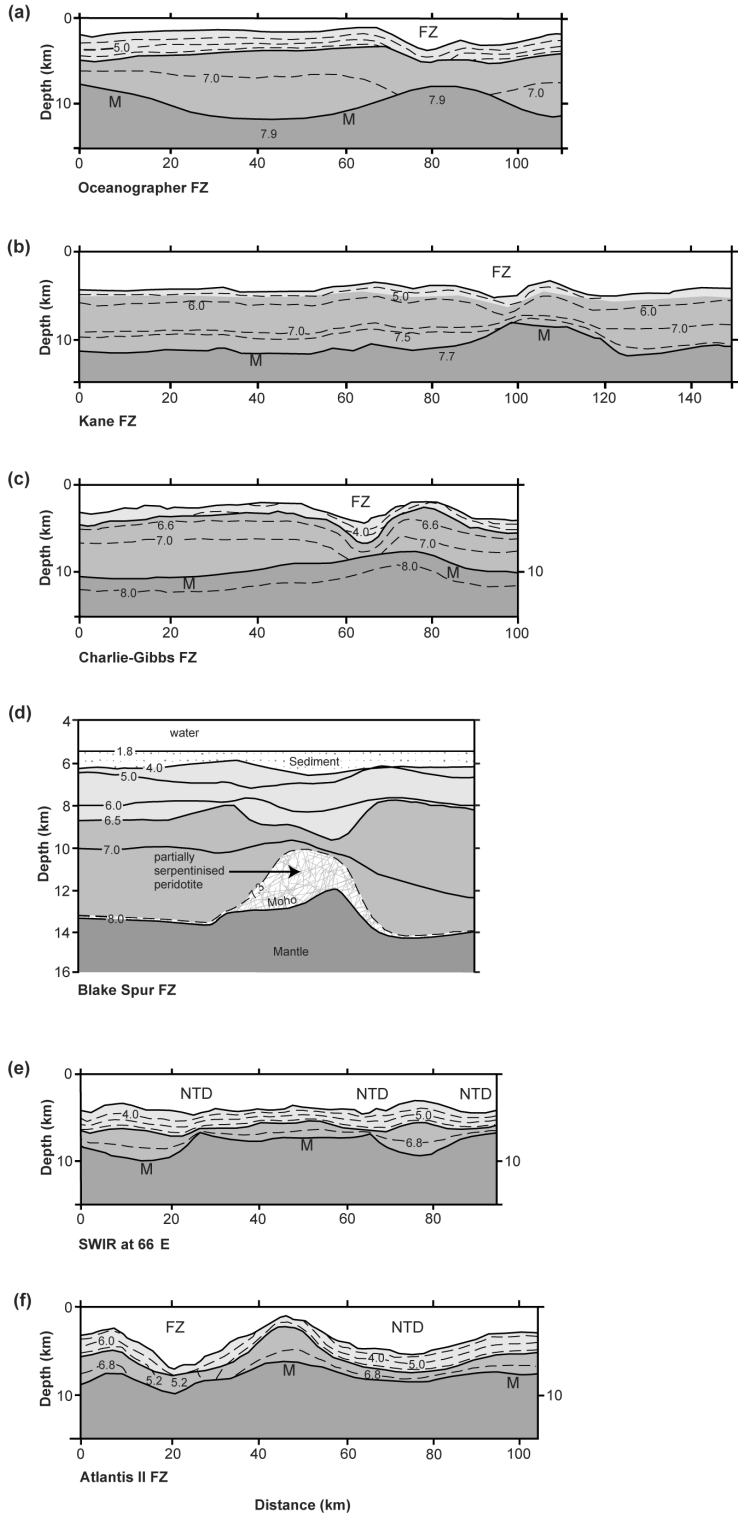


Figure 9.37. Seismic P-wave velocity cross sections of slow-slipping transform faults: (a) Oceanographer, Mid-Atlantic Ridge at 35°N; (b) Kane Mid-Atlantic Ridge at 24°N; (c) Charlie Gibbs Mid-Atlantic Ridge at 52°N; (d) Blake Spur fracture zone in the western Atlantic, showing the thin crust beneath this 12-km-offset feature on 140-Ma-old lithosphere; (e) thee non-transform discontinuities on the Southwest Indian Ridge at 66°E; and (f) Atlantis II, Southwest Indian Ridge. All sections are normal to the transform fault (i.e., parallel to the ridge) and cross the fault between the two offset ridge segments. Note the low velocities and thin crust in the fault and the gradual thinning of the crust towards the fault. (After Sinha and Louden (1983), Abrams *et al*. (1988), Whitmarsh and Calvert (1986), White *et al*. (1984), Minshull *et al*. (1991) and Muller *et al*. (1999; 2000).)

a 2–4-km-thick zone with high velocity gradients and velocities appropriate for layer 2 and the upper part of layer 3. This is underlain by a 2–3-km-thick layer with velocity of 7.2–7.6 km s^{-1} and then by the Moho. In trying to decide the origin of this 7.2–7.6 km s^{-1} layer it is important here to keep in mind that the Moho here is strictly a seismic boundary. This seismic Moho and the petrological Moho need not be locally coincident beneath fracture zones. Water penetrating through the fracture zone could have serpentinized the uppermost mantle, resulting in a reduction of the seismic velocity of peridotites. This would mean that the seismic Moho would occur beneath any serpentinized peridotite, whereas the petrological crust–mantle boundary would be drawn above it. If this is the case, hydrothermal circulation must have penetrated the cracked and therefore more permeable lithosphere along the fracture zone down to depths of at least 6 km. Alternatively, the 7.2–7.6 km s^{-1} material may perhaps represent interlayered gabbro and dunite, a much thickened transition zone between crust and mantle. Whichever explanation is correct, the fracture-zone crust is thinner than the normal adjacent crust, with most of the thinning taking place in layer 3, and the sudden change in crustal structure beneath the fracture zone strongly implies that there is a major reduction in magma supply at the ends of ridge segments. A further example of the importance of realizing that the seismic Moho cannot uniquely be assigned to a specific geological contact comes from the Mid-Atlantic Ridge at 35°N. In a 90-km-long segment south of the Oceanographer transform fault, the seismic Moho corresponds to a gabbro-plus-dunite/ultramafic contact; to a gabbro/ultramafic contact; and to a serpentinized alteration front beneath the thin crust of a non-transform offset within the segment and beneath the inside corner at the southern end of the segment.

Results of a three-dimensional tomographic study across the Clipperton transform fault (Table 9.4) on the East Pacific Rise do not show any substantial reduction in crustal thickness. The P-wave velocities are significantly reduced, by $\sim$1 km s^{-1}, over a 10-km-wide zone centred on the fault and extending throughout the crust. The low velocities are consistent with an increase in porosity caused by brittle deformation of the crust. Sea water can therefore penetrate throughout the crust along the fault. The constant crustal thickness is consistent with a two-dimensional plate-driven mantle flow regime, rather than a three-dimensional segment-centred flow, beneath the ridge – some models suggest that the flow should be two-dimensional for spreading rates in excess of 4 cm yr^{-1} (Fig. 9.38). As yet no seismic models of the fastest-slipping transform faults have been made. However, beneath the Tamayo and Orozco transform faults on the northern East Pacific Rise, which both have slip rates of approximately 6 cm yr^{-1}, the crust is 1–3 km thinner than normal oceanic crust.

Gravity anomalies are usually associated with transform faults; the local topography is uncompensated. The interpretation of the excess mass is a matter of current debate (gravity models are not unique), but it could be partially due to anomalous material at a shallow depth. Until the details of the tectonics

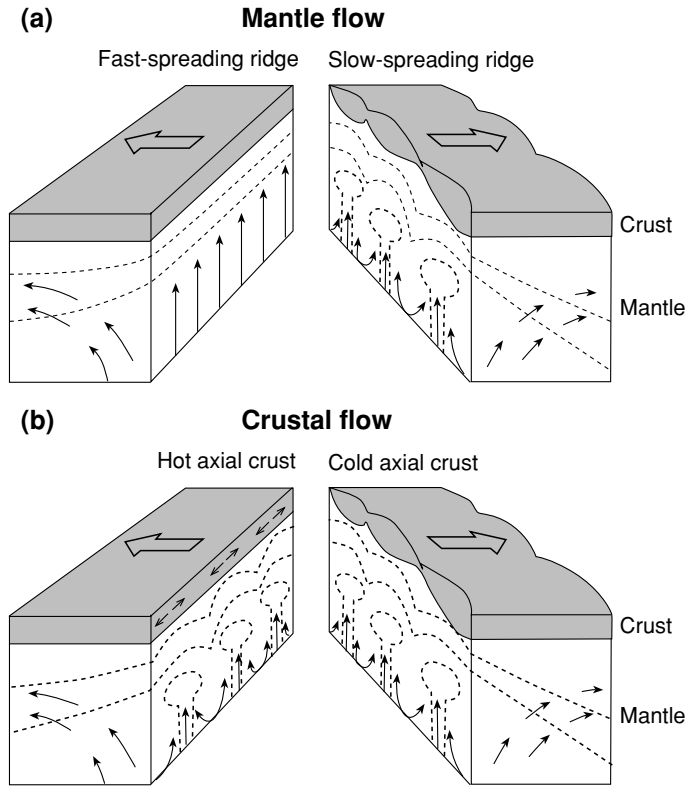


Figure 9.38. Two possible mechanisms for the crustal thickness along fast-spreading ridges to vary little, whereas along slow-spreading ridges it decreases towards segment ends. (a) Mantle flow is two-dimensional beneath fast-spreading ridges and three-dimensional beneath slow-spreading ridges. (From Lin and Phipps Morgan, The spreading rate dependence of three-dimensional mid-ocean ridge gravity structure, *Geophys. Res. Lett.*, **19**, 13–16, 1992. Copyright 1992 American Geophysical Union. Reproduced by permission of American Geophysical Union.) (b) Mantle flow is three-dimensional both beneath fast-spreading ridges and beneath slow-spreading ridges, with axial flow in the crust at fast-spreading ridges. Reprinted with permission from *Nature* (Bell and Buck, *Nature*, **357**, 583–6) Copyright 1992, Macmillan Magazines Ltd.)

of transform fault are better understood, this issue probably will not be fully resolved.

The age contrast across transform faults varies greatly (Table 9.4). This means that the thickness of the lithosphere changes across the fault (see Section 7.5.2). A 37 Ma age contrast across the St Paul transform fault in the North Atlantic implies an increase in lithospheric thickness from perhaps 7 km at the ridge axis to over 60 km, almost an order of magnitude. It is this presence of young, hot lithosphere next to old, cold, thick lithosphere which probably controls much of the topography, magmatism and morphology of the transform faults. The nodal basins are considered to be the result of a thermally induced loss of viscous head in the rising asthenosphere beneath the ridge axis, where the ridge is truncated by the older, cold thick lithosphere.

Figure 9.34 illustrates the two extremes in transform-fault structures: large-offset slow slipping and large-offset fast slipping. It is important to realize that it is not the length of the offset but the age difference across the transform which controls the magnitude of the topography and the type of transform fault formed. In the slow-slipping example, the faulting takes place across a narrow zone because the thick, cold lithosphere places severe mechanical constraints on

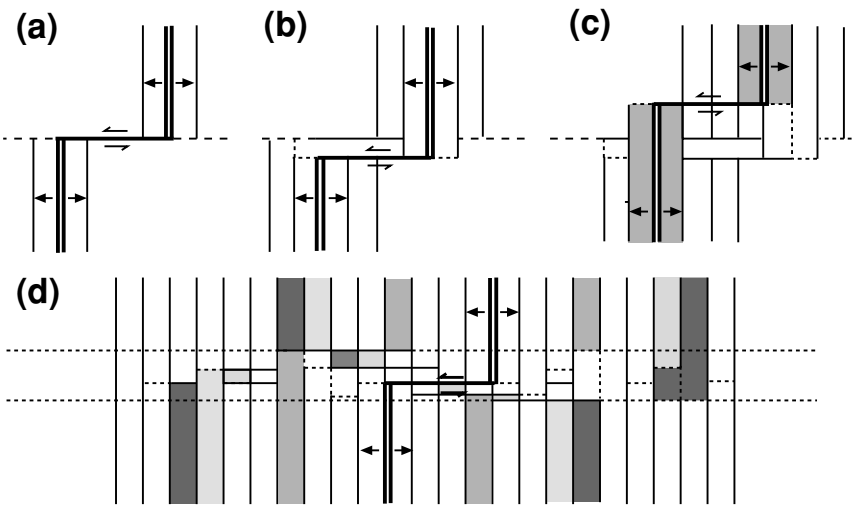


Figure 9.39. A diagram illustrating the evolution to topography and the lithospheric age distribution for a right-lateral transform fault: (a) simple model; (b), (c) and (d) progressive evolution when the ridge shown in (a) propagates and recesses. Dashed lines mark the extent of the fracture zone. At any one time the active transform fault would link the termination of the ridge segments. By time (d), it is clear that, near the active transform, the age of the lithosphere and the faulting are very complicated and that many old small lithospheric blocks are retained there. (From Goud and Karson (1986).)

any fault. In the fast-slipping example, the lithosphere is much thinner, so it is much easier for a wide shear-zone system to develop. Detailed studies of the Mid-Atlantic Ridge have been made in the FAMOUS area, which includes a number of short-offset ($\sim$20 km) transform faults. Two of these transform faults, A and B, which are at opposite ends of a 50-km ridge segment, illustrate the inadvisability of generalization. They are similar in terms of their broad structure (Table 9.4), but results of detailed submersible studies of morphology, faulting, location of scarps and exposures of fresh volcanic material show that transform fault A looks very like the large-offset slow-slipping model of Fig. 9.34 with faulting confined to a narrow zone, whereas in transform fault B, shear deformation is occurring across a wide ($>$6 km) zone. Interpretations of the bathymetry and magnetics support these observations and suggest that, at transform fault B, the ridge may well have undergone propagation and recession in the past (instability in the location of the PTDZ), giving rise to the more complicated structure observed there. Figure 9.39 shows schematically such an unstable PTDZ.

The thinner than normal crust in the fracture zones and the gradual thinning of the crust over some tens of kilometres towards the fracture zones, as illustrated for the Atlantic in Fig. 9.37, are thought to be the result of generalization of two effects. The old, cold lithosphere juxtaposed against the ridge axis has a significant thermal effect on the magma supply and tectonics. Thermal modelling

(detailed in Chapter 7) shows that, for a 10-Ma offset fracture zone, temperatures anomalous by $\pm 100\,^{\circ}C$ would be observed within about 10 km of the transform fault, whereas on the ridge axis close to the intersection with the transform fault the temperature would be about $300\,^{\circ}C$ too low (Fig. 9.40). From such a temperature difference, it is fairly simple to calculate the resulting contribution to the change in depth of the median valley as the transform fault is approached. These anomalous temperatures provide an explanation for thin crust in large-offset transform faults, such as the Kane, Oceanographer and Charlie Gibbs faults, but cannot explain thin crust in small-offset transform faults, the gradual thinning of the crust adjacent to the transform fault and the deepening of the median valley towards the nodal basin. These effects are better explained by a magma 'plumbing' system, in which each ridge segment is fed by a single subcrustal, centrally located magma-injection zone. In this way, the crust near large- and small-offset transform faults would be thinner because these faults are at the far ends of the magma-supply systems. Additionally, the regular 40–80 km spacing between transform faults may indicate the horizontal distance over which magma from a single central supply point can feed a slow-spreading ridge segment. The more infrequent transform faults on fast-spreading ridges could in this way be an effect of a more plentiful supply of magma.

9.5.3 Seismic activity at transform faults

The active part of a transform fault is that portion between the two ridge segments. Figure 9.33 shows how closely the epicentres follow the axis of the mid-ocean ridge and transform faults. These epicentres are determined by observations of the earthquakes made by the WWSSN system, which, being on land, are necessarily far from many of the mid-ocean-ridge epicentres. This means that only fairly large-magnitude earthquakes (body-wave magnitude greater than about 5 in the case of the Mid-Atlantic Ridge) are detected and located and that location is subject to some error (to determine the location, one has to assume a velocity structure, which might not be particularly close to the exact structure at the epicentre). Nevertheless, it is clear that the earthquake activity is confined to a very narrow zone centred on the median valley at the ridge axis and to the transform valley on the transform faults. To detect smaller-magnitude earthquakes on the mid-ocean ridges, it is necessary to deploy seismometers directly on the seabed in an array close to and within the median and transform valleys. These instruments, called ocean-bottom seismographs (OBS), must withstand pressures of at least 5 km of water and record for many days; they are necessarily expensive. Launching and recovery and the determination of their exact positions on the seabed are difficult and important procedures. To use these instruments to locate earthquakes, at least three OBS must be recording simultaneously because four unknowns – origin time and epicentre location (x, y, z) – must be determined (see Section 4.2.1). The level of seismic activity along transform faults is

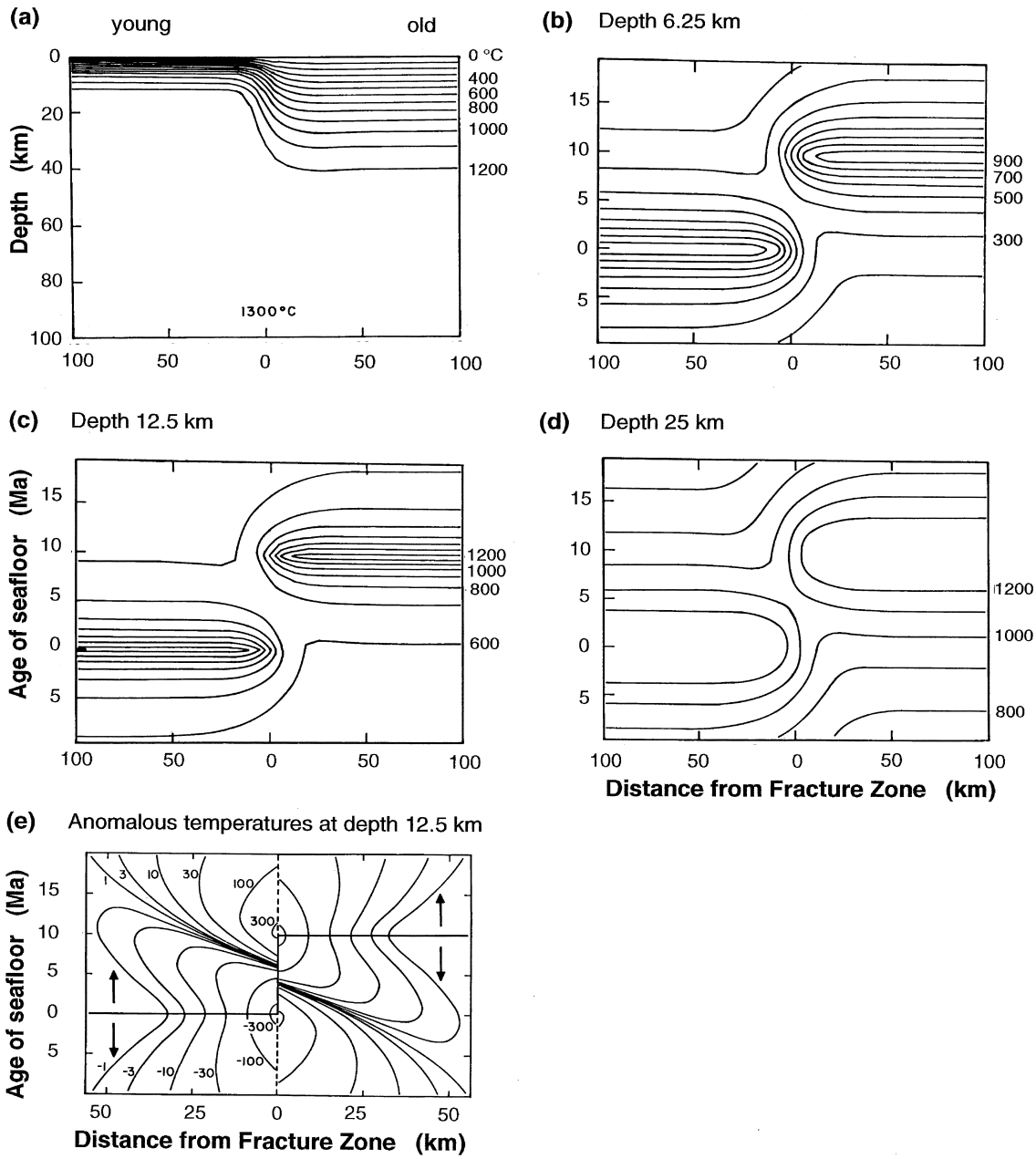


Figure 9.40. Temperatures (°C) near a transform fault with an age offset of 10 Ma and a slip rate of 3 cm yr^{-1}. The half-spreading rate is 1.5 cm yr^{-1}. (a) A vertical cross section through the ridge axis perpendicular to the fault; (b), (c) and (d) horizontal cross sections at depths of 6.25, 12.5 and 25 km, respectively. (e) Anomalous temperatures derived from (c) at 12.5 km depth. Note the different horizontal scale in (e). (From Forsyth and Wilson (1984).)

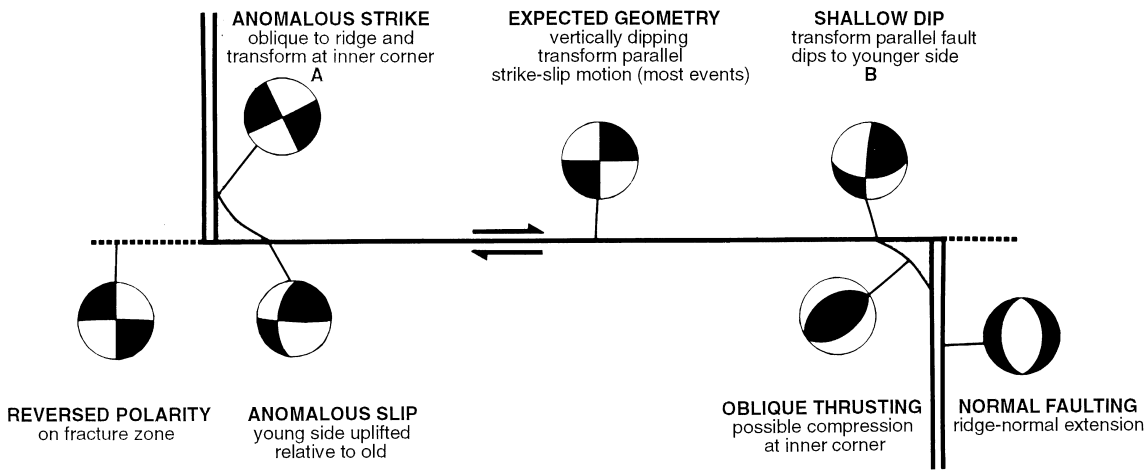


Figure 9.41. A schematic transform fault showing the types of earthquake which would occur at each location. Earthquakes A and B on the Romanche and Chain transform fault (Fig. 9.33) have the anomalous geometries illustrated here. (From Engeln *et al.* (1986).)

quite variable; often 10–50 microearthquakes per day are recorded over regions 40–50 km long, though this has been observed to vary by up to two orders of magnitude during a few months on adjacent transform faults. Earthquake swarms are common. Figure 9.41 shows the type of earthquake which could occur along a transform fault and at its intersections with the mid-ocean ridge. Notice in particular that anomalous fault-plane solutions could occur at the inner corner between the ridge and the transform fault. Figure 9.33 shows how fault-plane solutions along the equatorial Atlantic transform faults conform to this model.

Depths determined for microearthquakes and for teleseismically recorded events are generally shallower than 7–9 km. This confirms that the faulting is occurring above the 400–600 °C isotherms (half-space and cooling-plate models, Section 7.5.2, are indistinguishable at these young ages). In 1925, a magnitude-7.5 earthquake occurred on the eastern half of the Vema Transform; that was the largest Atlantic transform earthquake recorded to date. The largest recorded strike-slip earthquake, a magnitude 8.0, occurred in 1942 on the Andrew Bain transform fault.

9.6 Subduction zones

9.6.1 Introduction

A *subduction zone*, a convergent plate boundary, is the zone where old, cold lithospheric plate descends (is subducted) into the Earth's mantle. The surface expression of a subduction zone is the deep *oceanic trench* on the oceanic plate

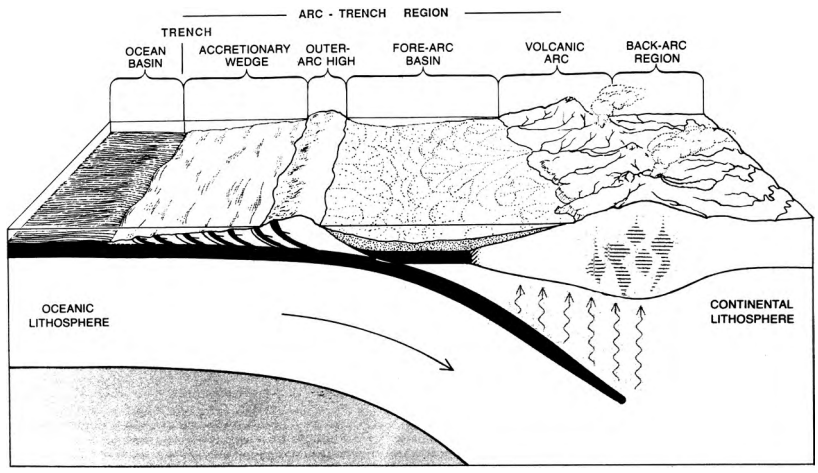


Figure 9.42. An active convergent plate boundary is characterized by a sequence of geological features. Extension in the back-arc region forms a back-arc, or marginal, basin. (From The continental crust, Burchfiel. Copyright © 1983 by Scientific American, Inc. All rights reserved.)

and the line of volcanoes on the overriding plate (Fig. 9.42). The volcanoes either form an island arc, such as those around the western Pacific, Indonesia and Caribbean, or are continental, as in the Andes. The horizontal distance between the trench and the volcanic arc is hundreds of kilometres – some 270 km for the Japan subduction zone.

Back-arc basins, otherwise known as *marginal basins*, occur behind the island arcs of the Western Pacific and Caribbean as well as the Scotia Arc. The crust in these basins is similar to oceanic crust, and magnetic lineations have been deciphered over some of them (e.g., the West Philippine Basin and the South Fiji Basin). The ophiolites which today are found obducted (thrust up) on land may be crust formed in back-arc basins.

The deep oceanic trench and the outer rise are a consequence of the bending of the lithospheric plate. Trenches often exceed 8 km in depth, the deepest being the Mariana Trench (see Section 9.1.1). The deformation of a thin elastic plate under an external load has been discussed in Section 5.7.1. This theoretical deformation has been used to explain the bending of the lithosphere at oceanic trenches and to yield estimates of its elastic parameters (see Figs. 5.14 and 5.15). For a subduction zone adjacent to a continent, the trench can be almost nonexistent because a high sedimentation rate may have filled it with sediments. In such cases, deformed sediments form a substantial *accretionary wedge* and *outer-arc high*. The *volcanic arc* itself is behind this sedimentary fore-arc basin. In the case of the Makran subduction zone (part of the boundary between the Arabian and Eurasian plates), which is an example of a shallow-dipping subduction zone with an extensive accretionary wedge resulting from a 7 km thickness of undisturbed sediment on the oceanic plate, the distance from the first deformation of sediment in the accretionary wedge (the *frontal fold*; see Fig. 10.9) to the volcanic arc is more than 400 km.

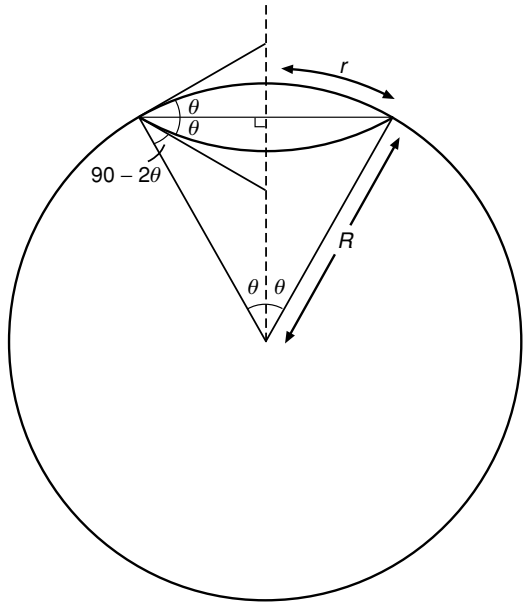


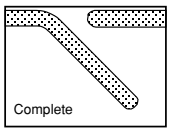
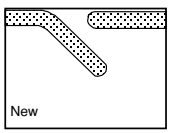
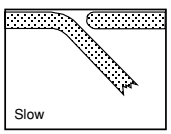
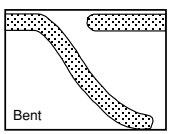
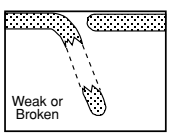
Figure 9.43. The geometry of an indented spherical cap on an inextensible spherical shell (e.g., a ping-pong ball). Radius of shell, R; radius of spherical cap, r; and angle subtended by spherical cap at centre of sphere, 2θ.

Oceanic trenches and island arcs are characteristically concave towards the overriding plate. This shape can be explained by imagining the lithosphere to be an inextensible spherical shell (like a ping-pong ball). If a portion of the shell is bent inwards, its edge is circular, and the indented portion has the same curvature as the original sphere (Fig. 9.43). The radius r of the indented circle is therefore $R\theta$, where R is the radius of the sphere and 2θ is the angle subtended by the circle at the centre of the sphere. The angle of dip of the indented circle, measured at the surface, is equal to 2θ. This simple model can be applied to the Earth and tested against observed dip angles of trenches and subduction zones and their radii of curvature. An arc radius of about 2500 km gives a dip angle of 45°, which is in reasonable agreement with values for many subduction zones but certainly not for all (Table 9.6 and Fig. 2.2). Thus, the 'ping-pong-ball' model provides a partial explanation for the concave shape of many oceanic trenches and island arcs even though it is an oversimplification of the problem.

9.6.2 Thermal structure

Measurements of heat flow along a cross section perpendicular to a subduction zone follow a standard pattern. The heat flow is low in the region between the trench and the volcanic arc. At the volcanic arc, there is a sudden jump in heat flow from about 40×10^{-3} to $(75–100) \times 10^{-3}\,\mathrm{W\,m^{-2}}$ within a very short horizontal distance. Heat flow then remains high across the volcanic arc and into the marginal basin. The sudden increase in heat flow at the volcanic arc is due

Table 9.6 *Details of major subduction zones*

Subduction zone	Plates	Length of zone (km)	Approx. subduction rate (cm yr^{-1})	Approx. dip angle (degrees)	Thermal parameter (1000 km)	Geometry
Kurile–Kamchatka–Honshu	Pacific under Eurasia	2800	6–13	40	5–7.5	
Tonga–Kermadec–New Zealand	Pacific under India	3000	8	60	12–17	
Middle America	Cocos under N. America	1900	9	70	<1	
Aleutians	Pacific under N. America	3800	6	50	2.5	
Sundra–Java–Sumatra–Burma	India under Eurasia	5700	7	70	4	
Cascadia	Juan de Fuca under N. America	1400	2	15	<1	
South Sandwich–Scotia	S. America under Scotia	650	2			
Caribbean	S. America under Caribbean	1350	0.5	50		
Aegean	Africa under Eurasia	1550	3			
Solomon–New Hebrides	India under Pacific	2750	10	70	9–12	
Izu–Bonin–Marianas	Pacific under Philippine	4450	10	60	7–12	
Iran	Arabian under Eurasia	2250	5	5		
Himalayan	India under Eurasia	2400	5			
Ryukyu–Philippine	Philippine under Eurasia	4750	7	45	2–6	
Peru–Chile	Nazca under S. America	6700	9	30	1.5–5	

Note: The subducted plate is oceanic except for the Iran and Himalayan subduction zones, for which all or part of the subducted plate is continental. Thermal parameter = age of subducting plate × descent rate.

Sources: After Toksöz (1975), Furlong and Chapman (1982), Stein and Stein (1996) and Kirby *et al*. (1996).

to shallow heat sources, namely, magma beneath the volcanoes. (Examples from the Japan and Cascadia subduction zones are shown in Figs. 9.58(e) and 9.61(a)).

Subduction zones have been modelled by many researchers, and, although there are differences between their starting models and the resulting isotherms and heat flow, the results are broadly the same for all and could probably already have been predicted by a reader of Chapter 7. Figure 7.32 shows the thermal re-equilibration of a thrust fault. A subduction zone is merely a type of thrust that keeps moving as new cold plate is continually subducted. Figure 9.44 shows two conductive thermal models of the temperature structure in and around a subducting oceanic plate. The 95-km-thick plate is heated by the surrounding mantle by conduction. Since no allowance is made for shear, latent or radiogenic heating of the slab, the temperatures shown are low estimates. Note that these model subducting slabs retain low temperatures compared with adjacent mantle, even at depths of more than 700 km, which result in high seismic velocities and densities for the slabs. A slab subducting at 45° at 5 cm yr^{-1} would take only 20 Ma to reach 700 km depth; a slab subducting at 10 cm yr^{-1} would reach that depth in 10 Ma. The *thermal parameter*, equal to the product of the vertical rate of descent of the plate into the mantle and the age of the subducting lithosphere, is a useful parameter. The temperature at any depth varies smoothly with the thermal parameter: the thermal parameter is effectively proportional to the maximum depth of the 'V' in the temperature contours (isotherms). A low thermal parameter, representing slow subduction of a young (hot) plate (Fig. 9.44(a)), means that a given isotherm would be at a shallow depth (the plate heats up quickly). A high value for the thermal parameter, representing fast subduction of old (cold) plate (Fig. 9.44(b)), means that isotherms are deeper (the plate is much slower to heat up).

Many factors affect the fine details of subduction-zone temperature structure. Estimates of the magnitude of frictional heating shear stresses have ranged up to 100 MPa, though values in the range 10–40 MPa are reasonable on the basis of values of the heat flow in the trench–volcanic-arc region. Of central importance to the shallow thermal structure is the role of water, which transports heat along the thrust fault and through the overriding plate. In addition, the melting behaviour of the subducting plate and the overriding mantle depends very strongly on the amount of water present. Heat is advected by rising magma beneath the volcanic arc and the back-arc basin. A small-scale convection system tends to operate beneath the back-arc basin, giving rise to the ocean-type crust and magnetic anomalies there (which are discussed further in Section 10.2.1).

The main changes which occur in the subducting plate are the shallow reaction of the oceanic crust to eclogite and the changes deeper in the mantle of olivine to a spinel structure and then to post-spinel structures (see Section 8.1.5). These changes result in increases in the density of the subducting slab. Figure 9.45 shows equilibrium pressure–temperature curves for the olivine–spinel and spinel–post-spinel structures. The change of olivine to spinel is exothermic (heat-releasing)

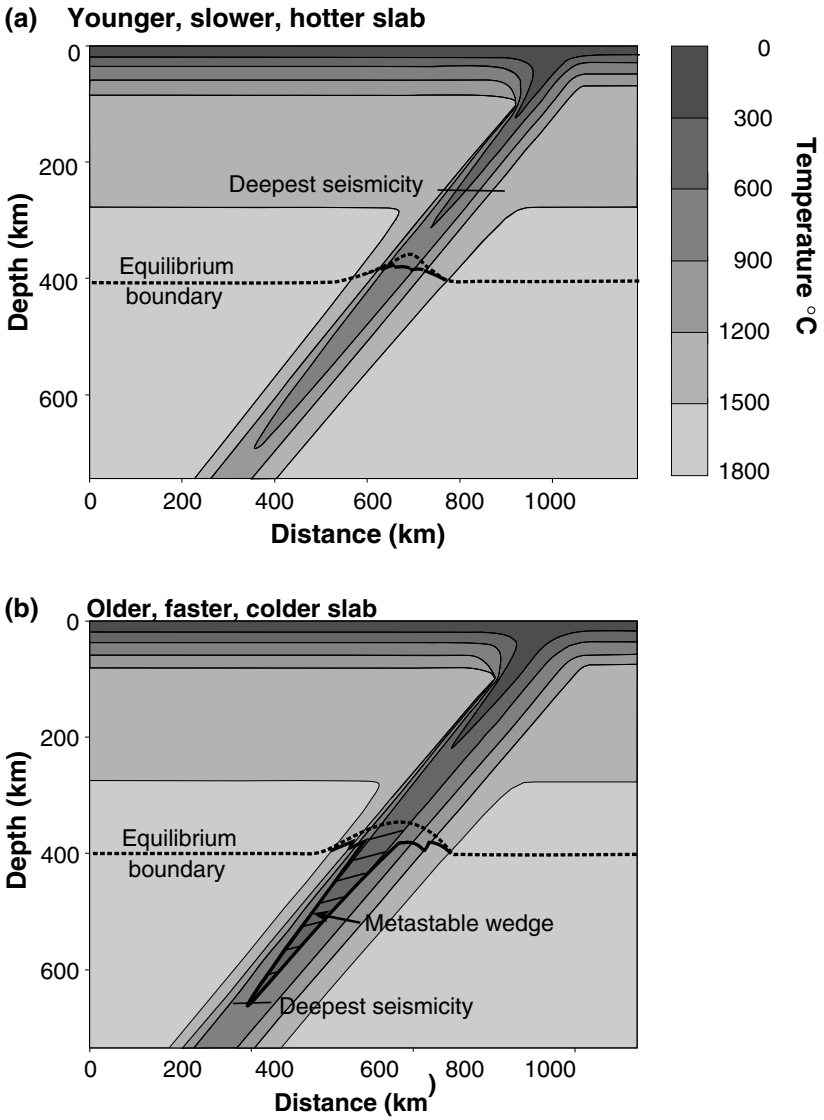


Figure 9.44. Thermal structure and predicted region of metastability of olivine for lithospheric plates with (a) a low thermal parameter of 2500 km similar to the younger and slowly subducting Aleutian slab and (b) a high thermal parameter of 17 000 km similar to the older and rapidly subducting Tonga slab. Within the subducting plate, the equilibrium position of the olivine–spinel phase change (dashed line) is elevated relative to the mantle. When the kinetics of the reaction are included in the calculation, the position of the olivine–spinel phase change (solid line) is depressed below its equilibrium position (dashed line). The thermal parameter is the product of the plate age and its vertical rate of descent into the mantle. The metastable wedge is present only in subduction zones that are sufficiently cold that the olivine–spinel change takes place more slowly than the descent of the plate. (After Stein and Stein, Thermo-mechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes (overview), 1996. Copyright 1996 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

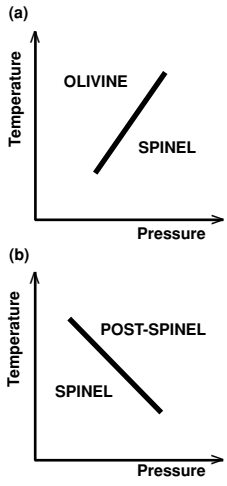


Figure 9.45. Equilibrium pressure–temperature curves (Clapeyron curves) for the olivine (α and β) to spinel (γ) and the 670-km spinel to perovskite and magnesiowüstite (Pv + Mw) phase changes. Estimates of the slopes are 2.0–3.0 MPa K^{-1} for the exothermic olivine–spinel change and -3 ± 1 MPa K^{-1} for the endothermic spinel–perovskite and spinel–magnesiowüstite change.

and at lower temperatures occurs at a lower pressure (shallower depth). Thus, in the subducting plate this phase change is elevated (i.e., occurs closer to the surface). For the models in Fig. 9.44 the equilibrium phase change is elevated by over 50 km in the subducting plate. The increase in density associated with this phase change is about 300 kg m^{-3}. The lower temperature of the subducting plate means that the spinel–post-spinel phase change occurs there at a higher pressure (greater depth) than it does in the mantle. This phase change is endothermic (heat-absorbing). Thus, the downward force on the subducting plate has three components: that due to the thermal contraction of the plate, that due to the elevation of the olivine–spinel phase boundary and that due to the depression of the spinel–post-spinel boundary. The third contribution is of opposite sign to the first two, acting to impede the descent of the plate, but is smaller in magnitude. Thermal contraction provides the greatest contribution to the overall driving force. (See Section 8.2.4 for discussion of the driving forces of plate tectonics.)

9.6.3 Seismic activity at subduction zones

Subduction zones are delineated by earthquake activity (see Figs. 2.1 and 2.2) extending from the surface down to depths of, in some cases, almost 700 km. The volcanic and seismic *ring of fire* around the margins of the Pacific Ocean is due to the subduction of oceanic plates (Fig. 2.2). It has been estimated that over 80% of the total energy at present being released worldwide by earthquakes comes from earthquakes located in this ring. The remaining 15% of the total energy is released by earthquakes in the broader seismic belt which extends eastwards from the Mediterranean and across Asia and includes the Alps, Turkey, Iran and the Himalayas.

Table 9.6 shows the distribution of earthquake foci with depth for the major subduction zones. These deeply dipping seismic zones are sometimes termed Wadati–Benioff zones after the Japanese discoverer of deep earthquakes, Kiyoo Wadati, and his American successor, Hugo Benioff. Fig. 9.46 shows the shallow geometry of the major subduction zones, and Fig. 9.47 shows the shape of the Tonga–Kermadec subduction zone as defined by earthquake foci. This subduction zone, which extends to over 600 km in depth, is S-shaped in plan view. Figure 9.48 shows the focal distribution of microearthquakes along a cross section perpendicular to the trench axis beneath northeastern Japan. A large amount of very shallow seismicity is associated with volcanism and shallow deformation and thrusting. A cluster of low-frequency events (black circles) in the lower crust/uppermost mantle beneath the volcanic arc seems to be caused by deep magmatic activity of mantle diapirs. The deeper foci, however, clearly delineate the descending lithospheric plate. These foci apparently define two almost parallel planes (which are also observed beneath the Aleutians): the upper plane is defined by earthquakes with reverse faulting or down-dip compressional stresses and the lower plane is defined by earthquakes with down-dip extensional stresses. The

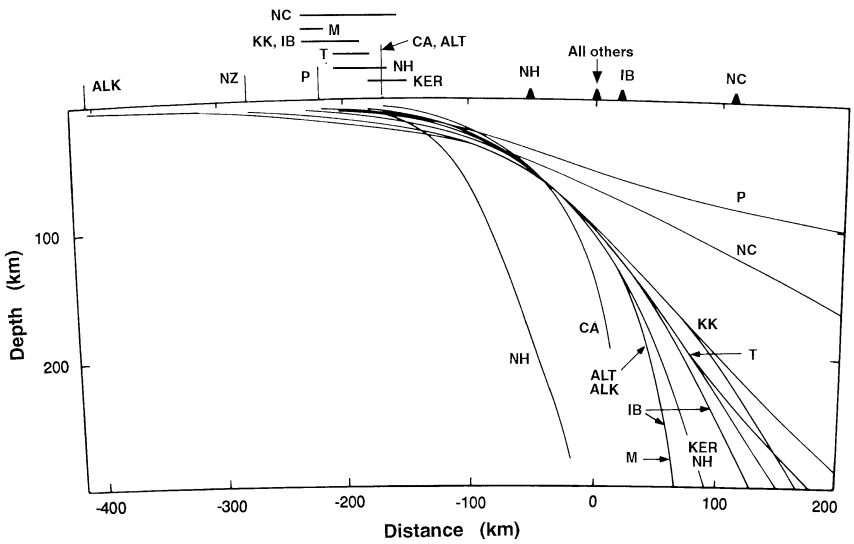


Figure 9.46. The shallow geometry of subduction zones as defined by earthquake foci. Abbreviations of subduction-zone names: NH, New Hebrides; CA, Central America; ALT, Aleutian; ALK, Alaska; M, Mariana; IB, Izu–Bonin; KER, Kermadec; NZ, New Zealand; T, Tonga; KK, Kurile–Kamchatka; NC, North Chile; CC, Central Chile; SC, South Chile; P, Peru. Solid triangles marks the volcanic lines; vertical or horizontal lines mark the location or extent of the oceanic trench. Some sections have, for clarity, been offset from the others. (After Isacks and Barazangi (1977).)

focal mechanisms of these intermediate-depth earthquakes being due to brittle rupture shows that melting within the slab is not taking place at these depths. The upper and lower boundaries of an 80-km-thick model subducting plate are shown superimposed on the foci. This subducting plate has P- and S-wave velocities which are 6% higher than those in the surrounding mantle. The tomographic inversion of the seismic travel times of P- and S-waves yields a detailed image of the perturbations both in P- and in S-wave velocity beneath Japan. Low-velocity zones in the crust are present beneath active volcanoes – many of those in the upper crust may be caused by the presence of water (on the basis of low values of v_p/v_s), whereas those in the lower crust apparently result from the presence of partial melt (on the basis of higher values of v_p/v_s). Within the mantle wedge the low-velocity regions are consistent with the presence of partial melts. Thus the process imaged here can be summarized as follows: water lost by the subducting plate at depths of 140–150 km moves into the mantle wedge where it causes partial melting; these partial melts move upwards and oceanwards, apparently following the flow-lines of the induced convection in the mantle wedge, i.e., along the fabric of the shear structures; melts rise and intrude the lower crust beneath active volcanoes; water expelled from cooling lower-crustal melt bodies rises into the upper crust; and some rising melts are erupted along the volcanic arc. (See Section 10.2.1 for further detailed discussion.)

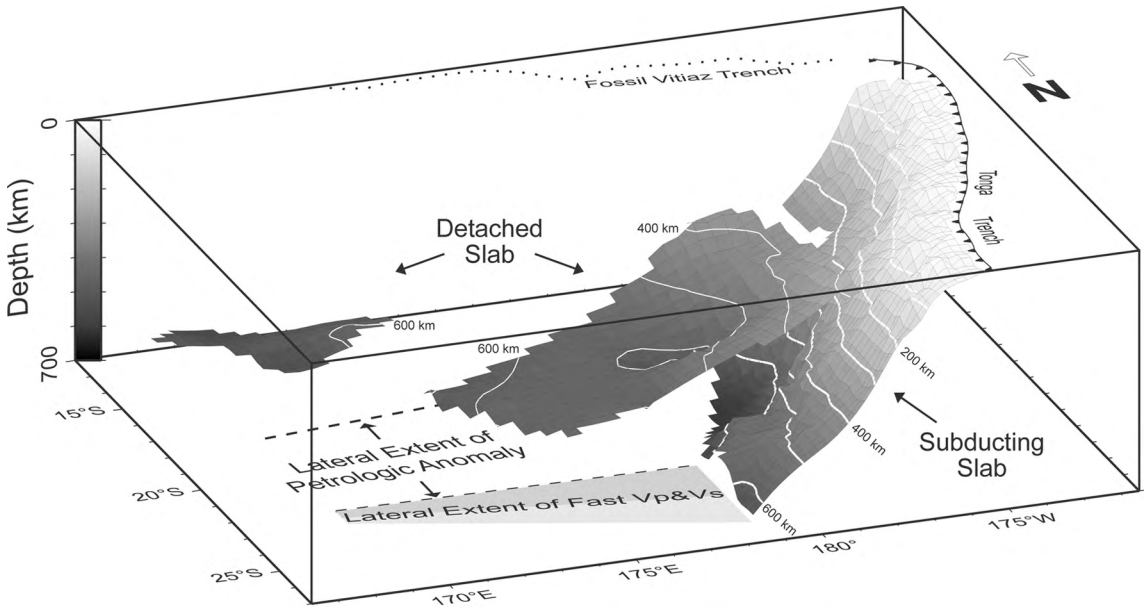


Figure 9.47. The seismicity of the Tonga–Kermadec subduction zone. Two separate surfaces are present south of 15°S, representing a detached remnant of slab and the actively subducting Pacific plate. There is also a separate cluster of earthquakes in the northwest. Contour lines are every 100 km in depth. (From Brudzinski and Chen, A petrologic anomaly accompanying outboard earthquakes beneath Fiji–Tonga: corresponding evidence from broadband P and S waveforms, *J. Geophys. Res.,* **108**, 2299, doi: 10.1029/2002JB002012, 2003. Copyright 2003 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

Results of a detailed study of the propagation of the short-wave seismic phases P_n and S_n in the region extending from the Tonga Trench to the Fijian islands are shown in Fig. 9.49. There is a zone of extremely high attenuation ($Q_\alpha < 100$) in the upper 100 km above the subducted plate beneath the spreading Lau back-arc basin, which is consistent with the low-velocity zones and partial melting. The 200–300-km-deep mantle wedge in the back-arc region has $Q_\alpha < 200$ – it is possible that these low Q values are associated with dehydration reactions occurring deep within the subducting slab. To the east of the subduction zone, at these depths Q_α is considerably higher. The subducted plate itself, down to 670 km depth, is characterized by low attenuation, with $Q_\alpha > 900$. A similar highly attenuative zone is present in the mantle wedge above the subducting Pacific plate beneath Japan. The low-Q regions are roughly coincident with the low-velocity zones shown in Fig. 9.48. The high attenuation and low velocities above these subducting plates are consequences of the dehydration of the subducting plate, the rise of water into the mantle wedge, the partial melting and then the accumulation of melt at high levels beneath and within the crust. When examined in detail, the low-velocity/low-Q zones are only rarely connected to the subducting plate. This is consistent with volatiles from the subducting plates

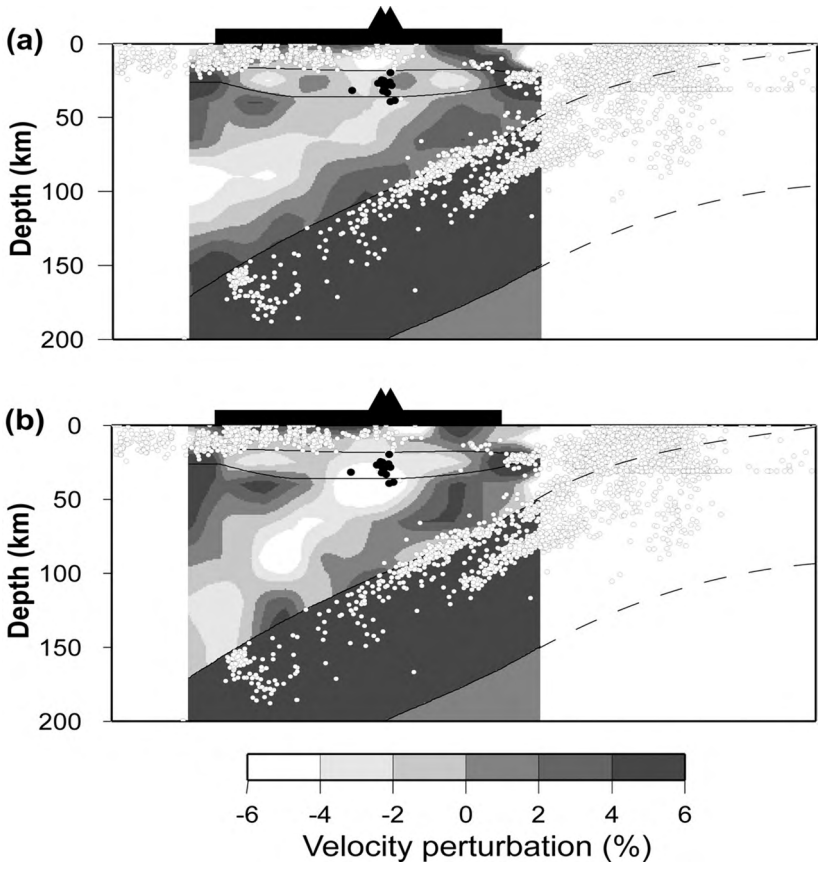


Figure 9.48. Cross sections perpendicular to the Japan Trench in northeastern Japan (at 37–38°N) where the Pacific plate is subducting beneath the North American/Okhotsk plate showing (a) P-wave and (b) S-wave velocity perturbations determined by tomographic methods. Circles, focal-depth distribution of earthquakes recorded by local seismic networks; black circles, low-frequency microearthquakes. The thick horizontal bar indicates the position of mainland Japan. The solid triangle marks the volcanic front. The 80-km-thick subducting slab has P- and S-wave velocities 6% higher than those of the surrounding mantle. A general assumption for the double seismic zone is that it is due to the unbending of the subducting plate. Alternative mechanisms include the transformation of lenses of subcrustal basaltic heterogeneities into eclogite and the dehydration of serpentine. (After Nakajima *et al.*, Three-dimensional structure of V_p, V_s and V_p/V_s beneath northeastern Japan: implications for are magnatism and fluids. *J. Geophys. Res.,* **106**, 21 843–57, 2001. Copyright 2001 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

rising and eventually causing partial melting within the overlying mantle wedge (see Section 10.2.1 for more discussion).

Seismic activity at subduction zones is not constant along the length of the zone: many quiet regions, or seismic gaps, have been observed. Such gaps are receiving particular attention since it is likely that a gap will be the location of

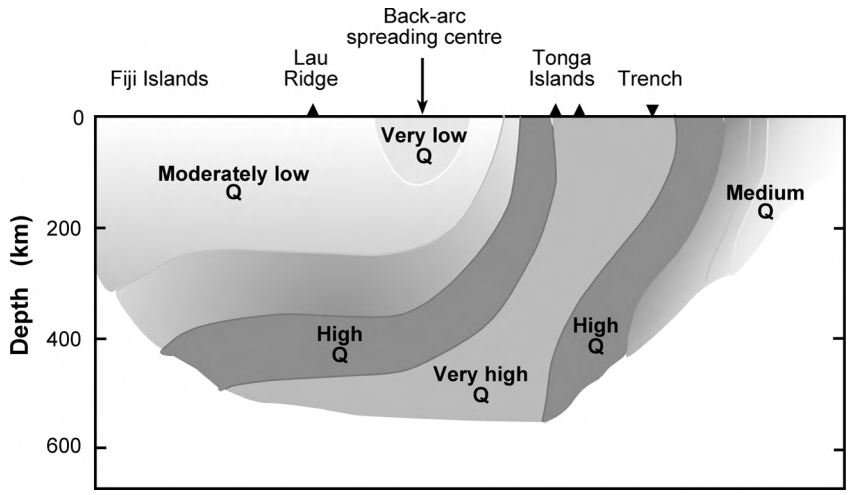


Figure 9.49. A cross section perpendicular to the Tonga Trench, showing the location of the subducting plate relative to the zones of high and low seismic attenuation (*Q*). (Based on Roth *et al.* (1999).)

a future large earthquake. Variations in the level of seismicity along the length of a subduction zone may also be controlled by factors such as geological or bathymetric structures in either of the two plates. Figure 9.50 shows the seismicity along the Tonga–Kermadec Arc. The Tonga–Kermadec subduction zone has very high seismicity; approximately one earthquake with m_b greater than 4.9 occurred every 2 km along the arc during the eighteen years of recording shown. There are clear variations in the level of the seismicity on a length scale ranging from tens to hundreds of kilometres.

One particular 250-km-long seismic gap in the Aleutian Arc is the Shumagin gap (Figs. 9.51 and 9.52). Analysis of background seismicity along the Aleutian Arc shows that activity there is low, although very large earthquakes have occurred along this arc in the past. The amount of stress built up along the Shumagin gap since the last great rupture there would be sufficient to result in an $M_w = 8.3$–8.4 earthquake within the next few decades, unless aseismic slip is occurring (i.e., the plates are moving smoothly past each other). An $M_S = 6.9$ earthquake in 1993 ruptured a small part of the gap but did not significantly relieve the build-up of strain. The available strain, tiltmeter and GPS measurements are unable to determine strain accumulation: aseismic slip could but might not be occurring. Unfortunately for those who live in the vicinity, only time may provide the answer to the question of the origin of the Shumagin gap.

Two other subduction zones where large earthquakes have long repeat times are the Cascadia subduction zone where the Juan de Fuca plate is descending beneath North America and the Rivera subduction zone where the tiny Rivera plate is being thrust beneath the Jalisco region of Mexico. Small earthquakes occur frequently along the 1400-km-long Cascadia subduction zone but there has been no large earthquake on the subduction thrust fault since record keeping began. Coastal geological evidence indicates that very large earthquakes with repeat times of

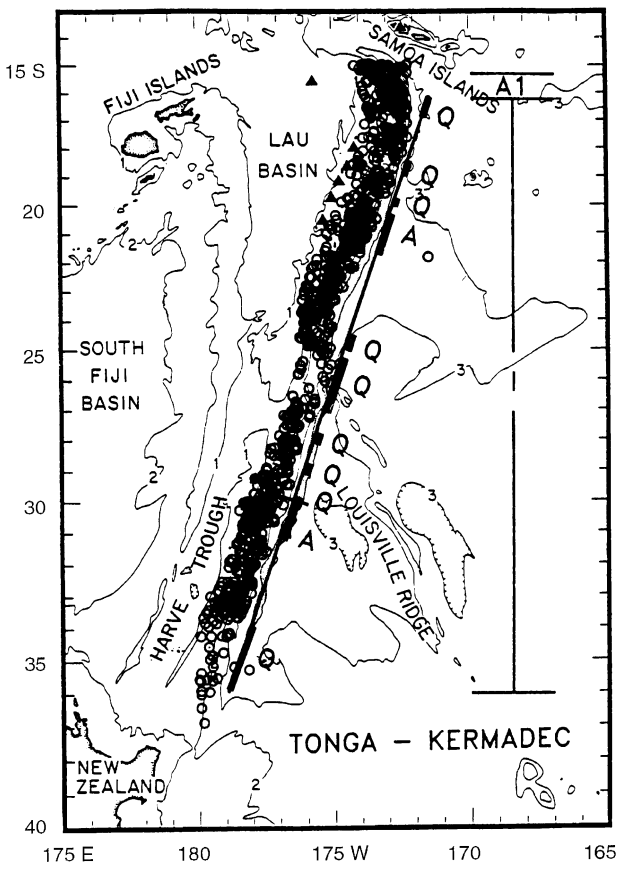


Figure 9.50. A map of the Tonga–Kermadec Arc showing the location of all the shallow, main shock epicentres (aftershocks are not shown) with m_b greater than 4.9 which occurred between 1963 and 1981. Regions with high and low seismicity are marked A (active) and Q (quiet), respectively. The very-active zone marked A1 is associated with the change in strike of the plate boundary from northerly to westerly. Bathymetric contours are in kilometres. (From Haberman *et al.* (1986).)

about 500 yr have occurred along this subduction zone. A damaging tsunami that affected Japan on 27 and 28 January 1700 is thought to have been caused by a great Cascadia earthquake (probable magnitude >9), which would have occurred at 9pm on 26 January. The present pattern of deformation along the subduction zone indicates that most accumulated strain is stored elastically and so will ultimately be released in earthquakes. While some of the strain has resulted in permanent deformation of the North America plate, there is no evidence that any aseismic slip is currently occurring: the subduction zone is locked. The amount of accumulated strain may be sufficient to generate a great Cascadia earthquake. The seismic moment accumulated there over 500 yr would be released by one magnitude-9 earthquake, or about thirty magnitude-8 earthquakes (Section 4.2.3). Potential implications for the coastal cities of British Columbia, Washington and Oregon are severe. Evidence suggests that the 1700 earthquake may have ruptured almost the whole length of the subduction zone. Far to the south the Rivera plate is being subducted beneath North America at 2.0 cm yr^{-1}, slower than the 5.2 cm yr^{-1} of the adjacent Cocos plate, which complicates the subduction processes. The

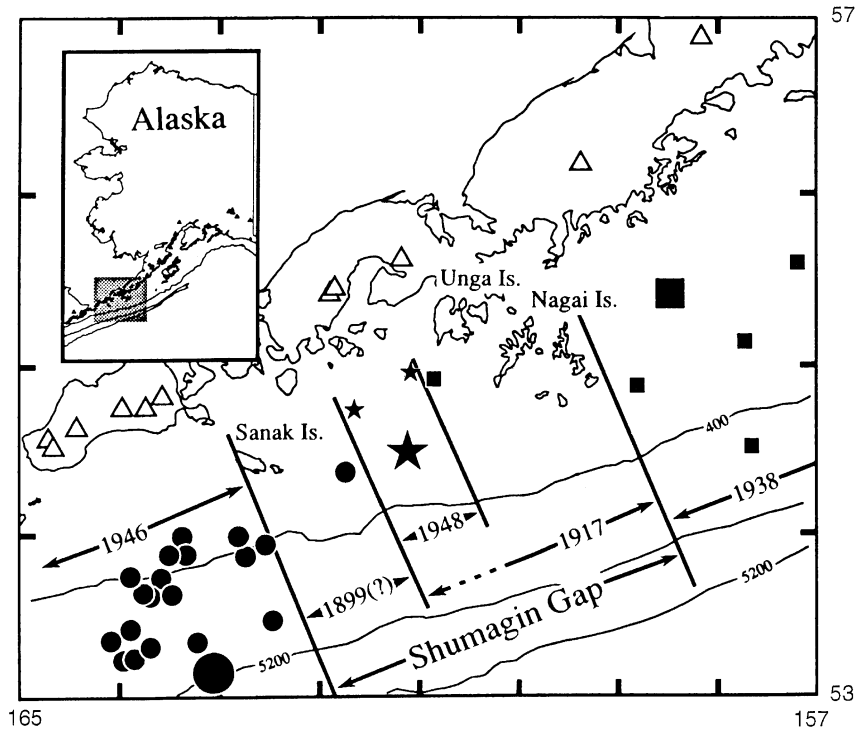


Figure 9.51. Part of the Aleutian Arc, showing the location of the Shumagin gap in relation to the previous large earthquakes in the region. Main shocks and aftershock sequences are shown as well as estimates of the extent of rupture for each event: ■ 1938, $M_w = 8.3$; ● 1946, $M_w = 7.3$ and ★ 1948, $M_S = 7.5$. The 1993, $M_S = 6.9$, earthquake ruptured a small part of the Shumagin gap. Probable rupture extents of the 1917, $M_S = 7.9$, and the 1899, $M_S = 7.2$, events are also shown. The great 1964, $M_w = 9.2$, earthquake was to the east at 154°W, 56.5°N. △, locations of volcanoes. Bathymetric contours are in metres. (After Boyd *et al.* (1988).)

largest earthquake ($M \sim 8.2$) ever to affect Mexico took place on the Rivera subduction zone in 1932. Until a shallow $M_w = 8.0$ earthquake occurred in 1995, there was concern that this subduction zone was locked. Faulting from this earthquake, which originated near the Cocos–Rivera plate boundary, was mainly along the Rivera–North America plate boundary.

Figure 9.53 shows the slow build-up of deformation that takes place when a subduction zone is locked and the sudden motions that occur when a great earthquake occurs. In the 1964 Alaska earthquake ($M_W = 9.2$) there was over 2 m of subsidence behind and over 6 m of uplift in front of the epicentre. That the Cascadia subduction zone is currently locked is illustrated in Fig. 9.54; the measured uplift is consistent with a numerical thermal model having part of the subduction zone locked. The uplift measured in Japan is also consistent with such a model.

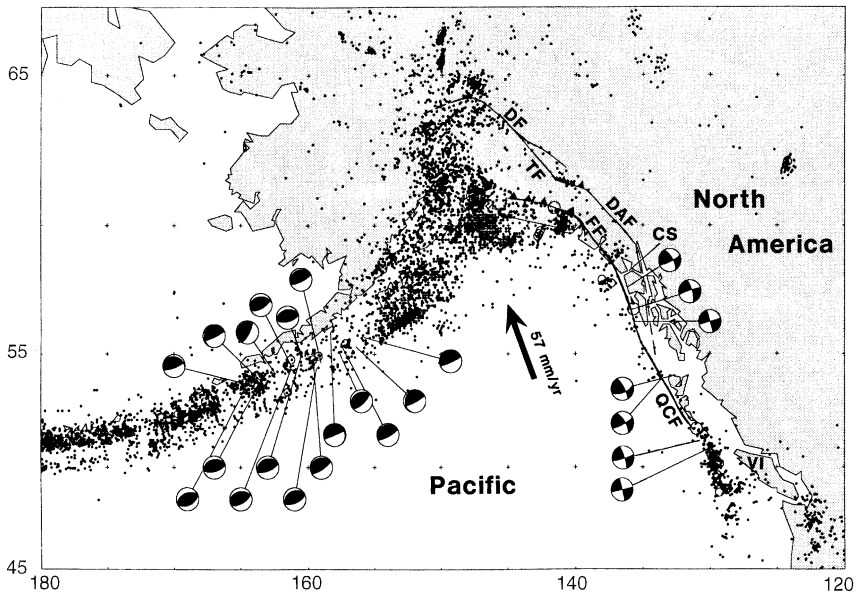


Figure 9.52. Focal mechanism for selected earthquakes along the Queen Charlotte Islands transform fault (QCF) and along the Alaskan end of the Aleutian subduction zone. VI, Vancouver Island; DF, TF, FF, DAF and CS, major faults. Blacks dots, epicentres of earthquakes shallower than 50 km. (Courtesy of C. Demets and R. G. Gordon.)

A detailed study of the focal mechanisms of large subduction-zone earthquakes indicates that they are depth-dependent. Earthquakes occurring within the outer rise of the subducting oceanic plate are predominantly shallow and extensional with the tension axis perpendicular to the trench. Shallow and medium-depth earthquakes usually have focal mechanisms with down-dip extension, whereas deeper earthquakes have down-dip compression. Some subduction zones have no seismic activity at intermediate depths; this could be due to the slab being broken or result from the thermal structure of the slab. The shallow earthquakes occurring in the overriding plate are predominantly thrust-faulting, whereas shallow earthquakes in the subducting plate are generally normal-faulting.

Deep earthquakes amount to about 20% of all earthquakes and their mechanism has long been enigmatic. There is no major change in apparent focal mechanism for intermediate and deep earthquakes and yet estimates of the pressure and temperatures indicate that, at these levels in the mantle, the strain should be released by creep or flow, not fracture. Temperature is an important parameter in controlling seismic activity at subduction zones. With increasing temperature, movement on faults will change from stick–slip (earthquakes) to stable sliding to ductile. Although rupture can be initiated only in the 'stick–slip' region, it can then propagate below that, but earthquakes below about 300 km are far too deep for frictional 'stick–slip' to be their cause. Although the bimodal form of the earthquake depth distribution suggests that there is a different cause for earthquakes below about 300 km, the pattern of seismic energy radiated from deep earthquakes looks very similar to that from shallower earthquakes. The observation that intermediate and deep earthquakes have down-dip compressive focal mechanisms has

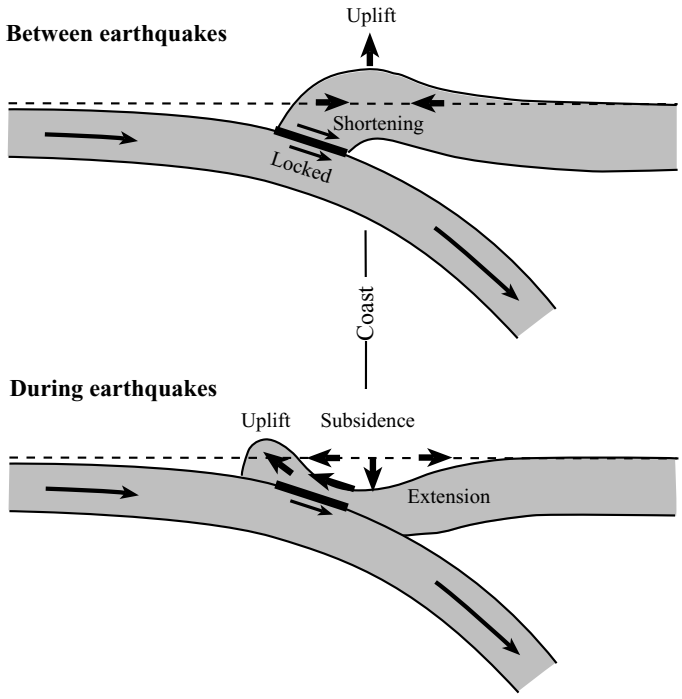


Figure 9.53. Cross sections of a subduction zone, illustrating the deformation that occurs between major earthquakes (*interseismic* deformation) and during an earthquake (*coseismic* deformation). Between earthquakes the seaward edge of the plate is dragged downwards while just inland the flexural bulge means that there is uplift. During the earthquake, the locked zone is suddenly released, causing the seaward edge to rebound and the uplifted bulge to subside. (From Hyndman and Wang, Tectonic constraints on the zone of major thrust earthquake failure: the Cascadian subduction zone, *J. Geophys. Res.*, **98**, 2039–60, 1993. Copyright 1993 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

traditionally been used as evidence in support of the hypothesis that there is a physical barrier at 650–700 km depth and an upper-mantle convection system that is separate from the lower mantle (Section 8.2.3). A mechanism that may account for intermediate earthquakes is the sudden release of strain on pre-existing faults lubricated by the dehydration of hydrous minerals, but any applicability of this mechanism to deep earthquakes below 300 km remains unproven. Another mechanism is the transformational faulting of metastable olivine in a wedge-shaped region in the interior of the subducting plate (Fig. 9.44). These mechanisms imply that earthquakes should occur within a narrow part of the subducted plate.

High-pressure laboratory experiments on olivine[3] indicate that deep earthquakes may be associated with the change of olivine to spinel in the subducting plate. Figure 9.44 shows that the equilibrium position of this phase change in the subducting plate is elevated above its position in the mantle. However, the kinetics of the reaction suggests that a wedge of olivine several tens of kilometres in thickness can persist in the centre of the slab well below the equilibrium depth, even down to 670 km. This situation arises because the kinetics of the olivine–spinel change mean that, in a fast subducting cold slab, the rate at which the change takes place is slower than the actual subduction rate. (The reaction rate is slow because it is temperature-dependent and decreases as temperature decreases.) This mechanism, which may be the origin of deep earthquakes, is termed ‘anticrack’ or

[3] Peridotite, the upper-mantle rock type, is mostly olivine, with about 10% pyroxene.

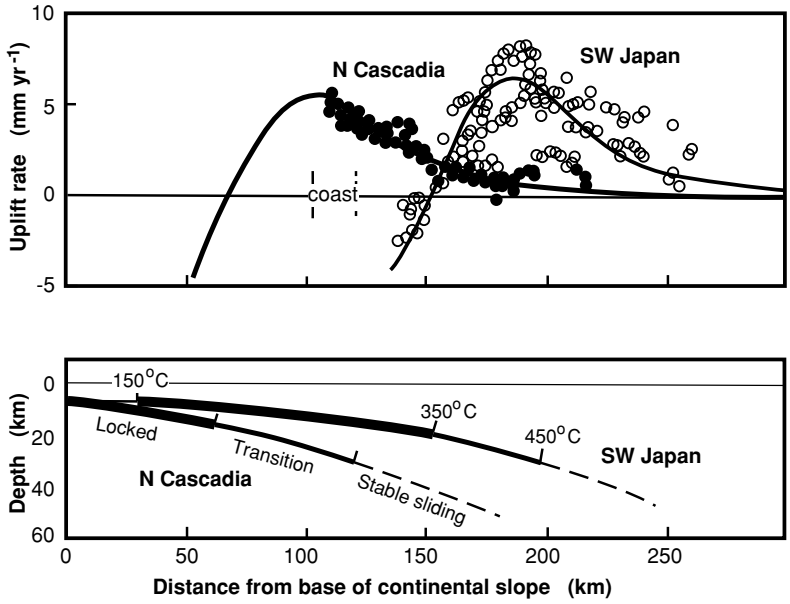


Figure 9.54. (a) The measured present-day uplift of coastal regions over the northern-Cascadia (black circles) and southwestern-Japan subduction zones (white circles). Solid lines show uplift predicted by a numerical thermal model of locked subduction zones. Position of coastline: Cascadia, solid line; Japan, dashed line. (b) Cross sections of the temperature models show the extent of the locked portion of the subduction zones. The behaviour of deformation on the subduction zone is temperature-controlled: locked zones (150–350 °C), thick solid line; transition (350–450 °C), thin solid line; and stable sliding (>450 °C), dashed line. (From Hyndman and Wang, The rupture zone of Cascadia great earthquakes from current deformation and the thermal regime, *J. Geophys. Res.*, **100**, 22 133–54, 1995. Copyright 1995 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

'transformational' faulting. When under stress the metastable olivine in the central wedge transforms to spinel in crack-like inclusions that form perpendicular to the direction of maximum stress. With sufficient cracks in a small region of the wedge the microspinel crystals become superplastic and a catastrophic slip occurs – an earthquake. Theoretically, transformational faulting should occur for reactions for which both the latent heat and the change in volume are negative, in other words, for exothermic rections with an associated increase in density. The olivine–spinel change satisfies these critera, as does the other major mantle mineral change, clinoenstatite[4]–ilmenite. With the maximum compressive stress oriented parallel to the slab, predicted earthquake mechanisms are consistent with those observed. The focal mechanism for these deep earthquakes, slip on a fault, is thus very similar to mechanisms for shallow earthquakes, although the mechanical cause is completely different. Earthquakes are not associated with spinel–perovskite and spinel–magnesiowüstite or ilmenite–perovskite reactions at 670 km because these are endothermic, absorbing, rather than releasing, heat. Any olivine or clinoenstatite still remaining in the metastable wedge at 670 km depth will also decompose to perovskite without producing earthquakes because the latent heat for these reactions is positive. Since there are no further phase changes in the lower mantle, there is no opportunity for any deeper earthquakes to occur. Thus the cessation of seismicity at the base of the upper mantle seems to be a direct consequence of mantle mineralogy.

[4] Clinoenstatite is a pyroxene mineral.

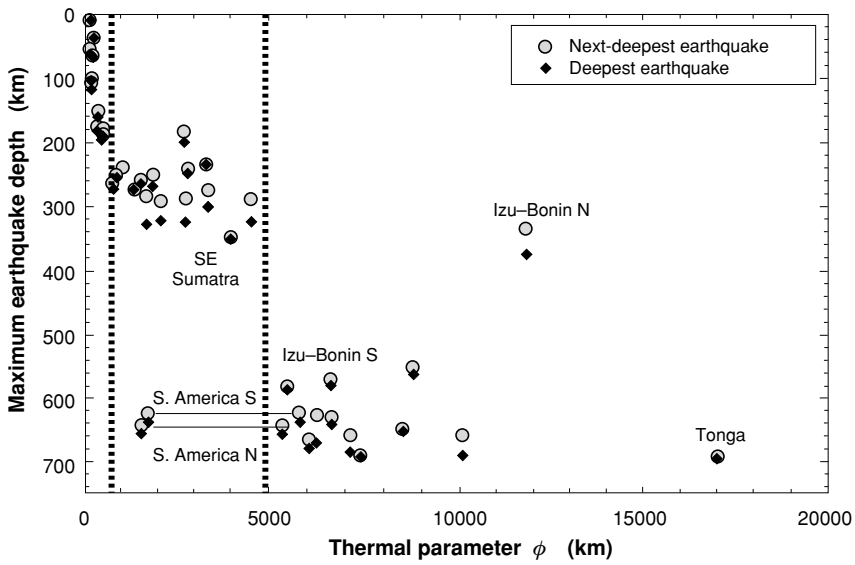


Figure 9.55. Thermal parameter plotted against maximum depth for subduction- zone earthquakes. If the occurrence of deep earthquakes were directly controlled by temperature, their depth should vary steadily with the thermal parameter. The subduction zones fall into two groups, those with $\phi > 5000$ having deep earthquakes and those with $\phi < 5000$ not having deep earthquakes. This is consistent with an abrupt phase change causing the earthquakes. (From Kirby *et al.*, Metastable mantle phase transformations and deep earthquakes in subducting lithosphere, *Rev. Geophys.*, **34**, 261–306, 1996. Copyright 1996 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

Figure 9.44 shows the temperature structure and the predicted regions of instability for two subduction zones, one with a low thermal parameter for which a metastable olivine wedge would not develop and the other with a high thermal parameter in which a metastable olivine wedge would be expected to develop. Table 9.6 and Fig. 9.55 confirm that subduction zones with a thermal parameter less than 5000 km do not have deep seismicity whereas those with a thermal parameter greater than 5000 km do. As the kinetics of the transformation of metastable olivine to spinel is crucially dependent upon temperature, these transformational earthquake focii should occur along the hotter outer edges of the wedge. While very deep earthquakes in the Tonga and Izu–Bonin subduction zones (the two subduction zones with the largest thermal parameters) seem to occur along double seismic zones, other seismic evidence for a metastable olivine wedge is hard to obtain. Since olivine is less dense and has a lower seismic velocity than spinel at the same temperature, a metastable olivine wedge would have a low seismic velocity and so could be detected, although the imaging of the deep structure of a subducting slab at the resolution of both velocity and depth required is a particularly difficult task. The coincidence of low v_p and v_s velocities, deep earthquakes and subducting lithosphere collecting just above the lower mantle behind the Tonga arc (Fig. 9.47) provides strong support for their origin being metastable olivine.

Some deep earthquakes take place in regions not associated with present-day subduction or in deep remnants of slabs that are apparently detached from the subducting plate. These earthquakes have been very difficult to explain. Transformational faulting is, however, able to account for this isolated seismicity. Remnants of subducted slabs containing regions of metastable peridotite will continue to transform: a physical connection to the surface is unnecessary. Thus

the $M_w = 7.9$ earthquake that occurred 626 km beneath Spain in 1954 and is associated with a seismic velocity anomaly may be the remnant of a slab detached following Africa–Eurasia subduction. Likewise the zone of seismicity at 600–670 km beneath the North Fiji Basin, which is separate from the seismicity of the Tonga and Vanuatu subduction zones, may be due to transformational faulting within a slab that was subducted at the fossil Vitayz trench, but then subsequently detached and foundered at the base of the upper mantle.

Deep earthquakes do not have many aftershocks compared with shallow earthquakes. Temperature is an important control on aftershocks following deep earthquakes but not on aftershocks following shallow earthquakes. Subduction zones with high thermal parameters (cold) have far more deep aftershocks than do subduction zones with low thermal parameters (hot). The *b*-value (Section 4.2.5) for deep aftershocks is also related to the thermal parameter: high-thermal-parameter slabs have high *b*-values and low-thermal-parameter slabs have low *b*-values. This is consistent with a transformational-faulting origin for deep earthquakes.

Two large deep earthquakes that occured in 1994 promoted much new research on deep earthquakes. The $M_w = 7.6$ Tonga event at 564 km depth was unprecedented in that 82 aftershocks with magnitudes between 3.6 and 6.0 were recorded during the following six weeks. The main event and many of the aftershocks occurred on a near-vertical plane consistent with one of the nodal planes from the main event. The rupture zone was 50 km $\times$ 65 km in extent and extended beyond the expected metastable olivine wedge and out of the known seismic zone (Fig. 9.56). It has been proposed that ductile faulting, triggered by tranformational faulting in the cool slab immediately outside the cold metastable wedge, caused the two outlying aftershocks. The $M_w = 8.3$ earthquake on the Nazca subduction zone beneath Bolivia was rather different in that it occurred in a region with no previous recorded seismicity and had only three aftershocks with $m_b \geq 4.5$. The geometry of the Nazca slab is therefore not known, but, depending upon its shape and extent, this earthquake with a 30–50-km sub-horizontal fault plane may have ruptured beyond any metastable wedge. Alternatively, this earthquake might not have been the result of transformational faulting but may instead have resulted from ductile faulting or plastic instabilities in a warmer spinel slab. This might account for the very high stress drop ($\sim$110 MPa), which was over an order of magnitude greater than that for the Tonga event and for normal shallower events (Fig. 4.12). Thus present knowledge of the kinetics of mantle reactions, the behaviour of minerals at very high pressures and the details of the assumptions made in thermal calculations may need to be re-examined in order to establish the cause of all deep earthquakes – transformational faulting might not be the only process taking place. Deep earthquakes remain something of a puzzle.

9.6.4 Gravity across subduction zones

Very large gravity anomalies occur over subduction zones. The anomalies across the Aleutian Trench, the Japan Arc and the Andes are shown in Figs. 9.57–9.59.

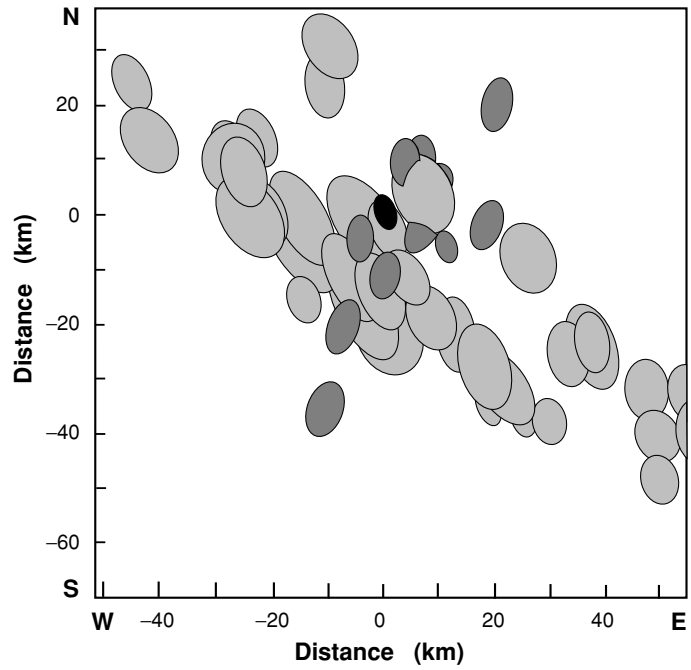


Figure 9.56. A map view of all well-located earthquakes (1980–1987) between 525 and 615 km depth (light grey), the 9 June 1994 deep Tonga earthquake (black) and its best-located aftershocks (dark grey) shown as 95%-confidence ellipsoids. The subducting Pacific plate is vertical here at 17–19°S (Fig. 9.47). The two linear bands of grey ellipsoids suggest that there may be a double seismic zone, which would be consistent with transformational faulting along the edges of a 30–40-km-wide metastable olivine wedge. The aftershocks from the 1994 earthquake clearly cut entirely across the seismic part of the slab into the surrounding aseismic region. The main-event (black) rupture started in the cold core of the slab and terminated some 15–20 km outside the seismic zone, close to the outlying aftershock, where the temperature is ~200 °C higher. This warmer region adjacent to the cold seismic core of the slab may have a different faulting regime for deep earthquakes: rupture can propagate here from the metastable core but aftershocks are rarely initiated in this region. Colour version, Plate 23 (Reprinted with permission from *Nature* (Wiens *et al.*, *Nature*, **372**, 540–3) Copyright 1994 Macmillan Magazines Ltd.)

The general feature of gravity profiles over convergent plate boundaries is a parallel low–high pair of anomalies of total amplitude between 100 and 500 mgal and separated by about 100–150 km. The *low* is situated over the trench; the *high* is near to and on the ocean side of the volcanic arc.

Density models that can account for these gravity anomalies include the dipping lithospheric plate and thick crust on the overriding plate. Details that also have to be included in the modelling are the transformation of the basaltic oceanic crust to eclogite with an increase in density of about 400 kg m^{-3} by about 30 km depth.

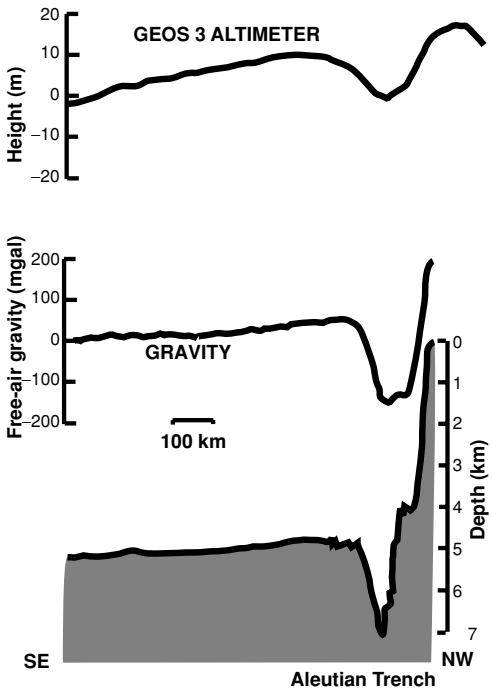


Figure 9.57. The geoid height anomaly as measured by the Geos 3 satellite altimeter, free-air gravity anomaly and bathymetry along a profile perpendicular to the Aleutian Trench. (From Chapman and Talwani (1979).)

9.6.5 Seismic structure of subduction zones

Earthquake data and seismic-refraction and -reflection profiling are all used to determine the seismic-velocity structure around subduction zones. The large-scale and deep structures are determined from earthquake data. The subducting plate, being a cold, rigid, high-density slab, is a high-velocity zone with P- and S-wave velocities about 5%–10% higher than those in normal mantle material at the same depth. The asthenosphere above the subducting plate, which is associated with convection and back-arc spreading in the marginal basin, is a region with low seismic velocities. Evidence from seismic modelling of deep earthquakes has revealed that the subducting plate may penetrate into the lower mantle as an anomalous high-velocity body, reaching depths of at least 1000 km. Figure 9.60 shows the deviations in velocity beneath three Pacific subduction zones: it is clear that there is no standard structure; some plates extend into the lower mantle, others do not. The central Izu–Bonin arc appears to be deflected at 670 km depth and extends horizontally beneath the Philippine Sea as far as the Ryukuyu subduction zone. However, to the south beneath the Mariana subduction zone the high-velocity Pacific slab extends to about 1200 km depth.

The crustal seismic structure across Japan and the Japan Sea is shown in Fig. 9.58(c). Japan has a 30-km-thick continental type of crust whereas the crust in the marginal basin is only 8–9 km thick with velocities near those of oceanic crust. Normal upper-mantle velocities are found beneath the oceanic plate, the

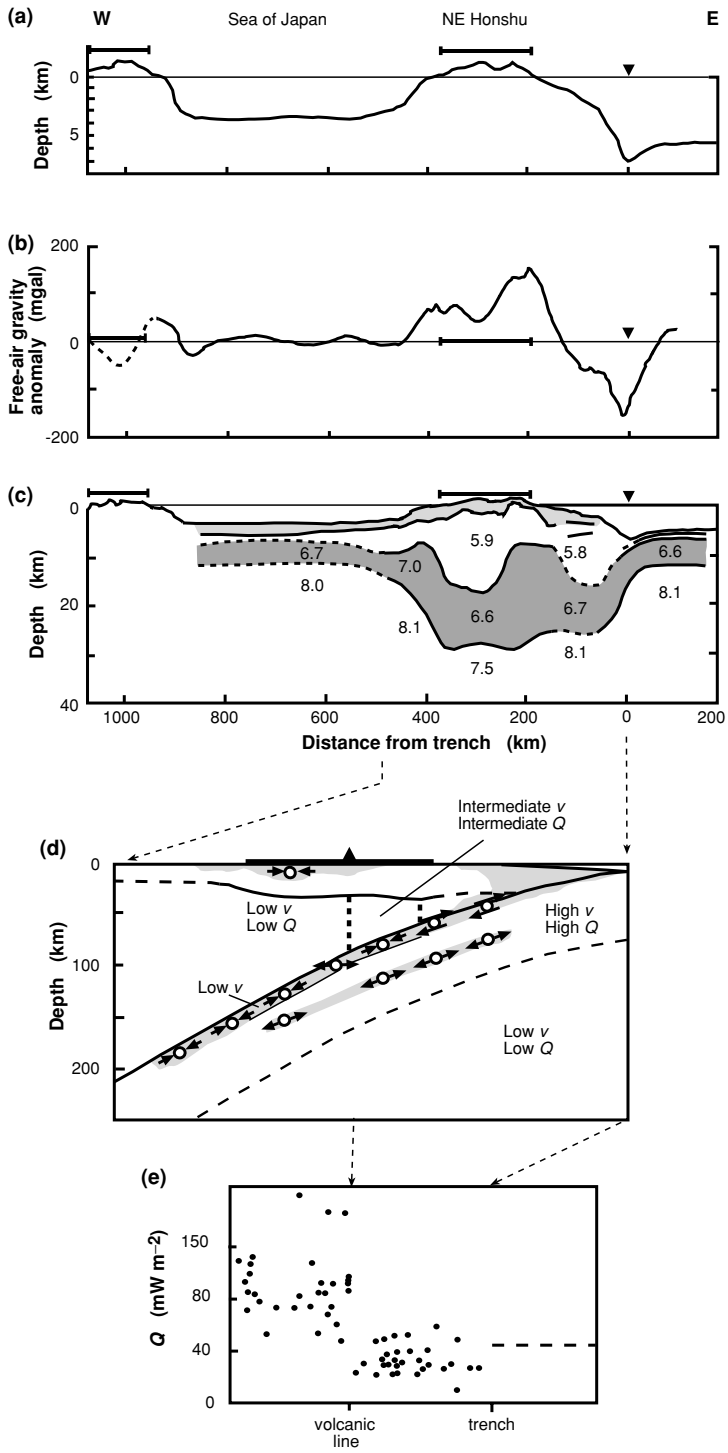


Figure 9.58. Sections across the Japan Trench and Arc. Solid horizontal bars denote land; ▼, trench; ▲, volcanic front. (a) Topography, vertical exaggeration 25 : 1. (b) Free-air gravity anomaly. (c) Crustal seismic P-wave velocity (km s^{-1}) structure, vertical exaggeration 10 : 1. (d) Summary of the seismic structure, true scale. Shaded areas, seismically active regions. Typical earthquake focal mechanisms are shown. The low-*Q*, low-velocity regions are the asthenosphere beneath the 30-km-thick overriding plate and beneath the 80-km-thick subducting Pacific plate. The focal-depth distribution of earthquakes beneath Japan is shown in Fig. 9.48. The thin low-velocity layer at the top of the subducting plate is considered to be the subducted oceanic crust. (e) Heat-flow measurements. The dashed line is the theorectical heat flow for 120-Ma oceanic lithosphere, 120 Ma being the age of the Pacific plate beneath Japan. (After Yoshii (1979), Matsuzawa *et al*. (1986) and van den Beukel and Wortel (1986).)

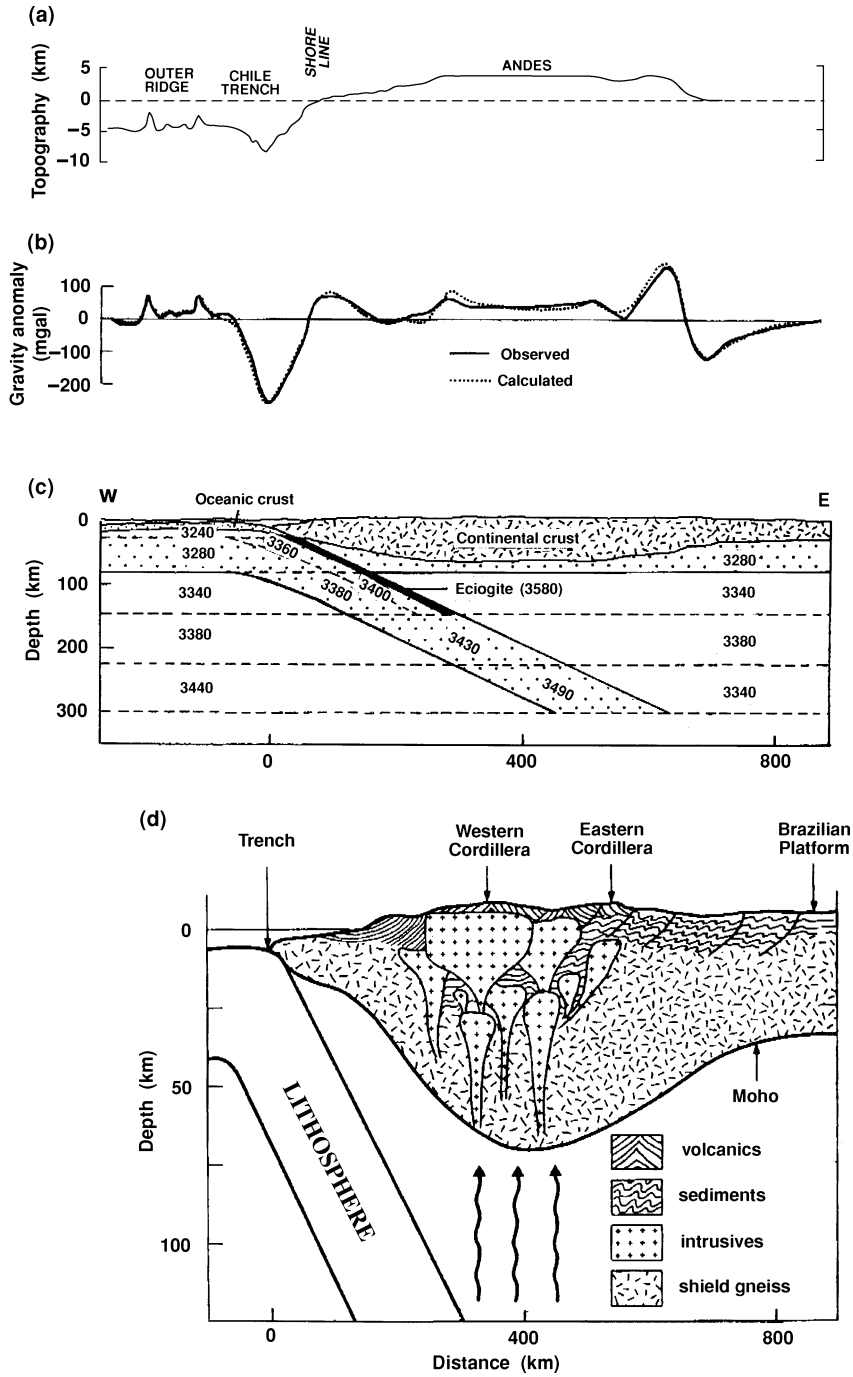


Figure 9.59. Sections across the Chile Trench and the Andes at 23°S: (a) topography, vertical exaggeration 10 : 1; (b) free-air gravity anomaly; and (c) density model, true scale (densities are in 10^3 kg m^{-3}). (After Grow and Bowin (1975).) (d) A schematic geological cross section through the central Andes. Arrows indicate rising magma. Most of the new material being added to the crust at this destructive plate boundary is in the form of huge diorite intrusions. The andesite volcanics provide only a small proportion of the total volume. The vertical is exaggeration is 5 : 1. (From Brown and Hennessy (1978).)

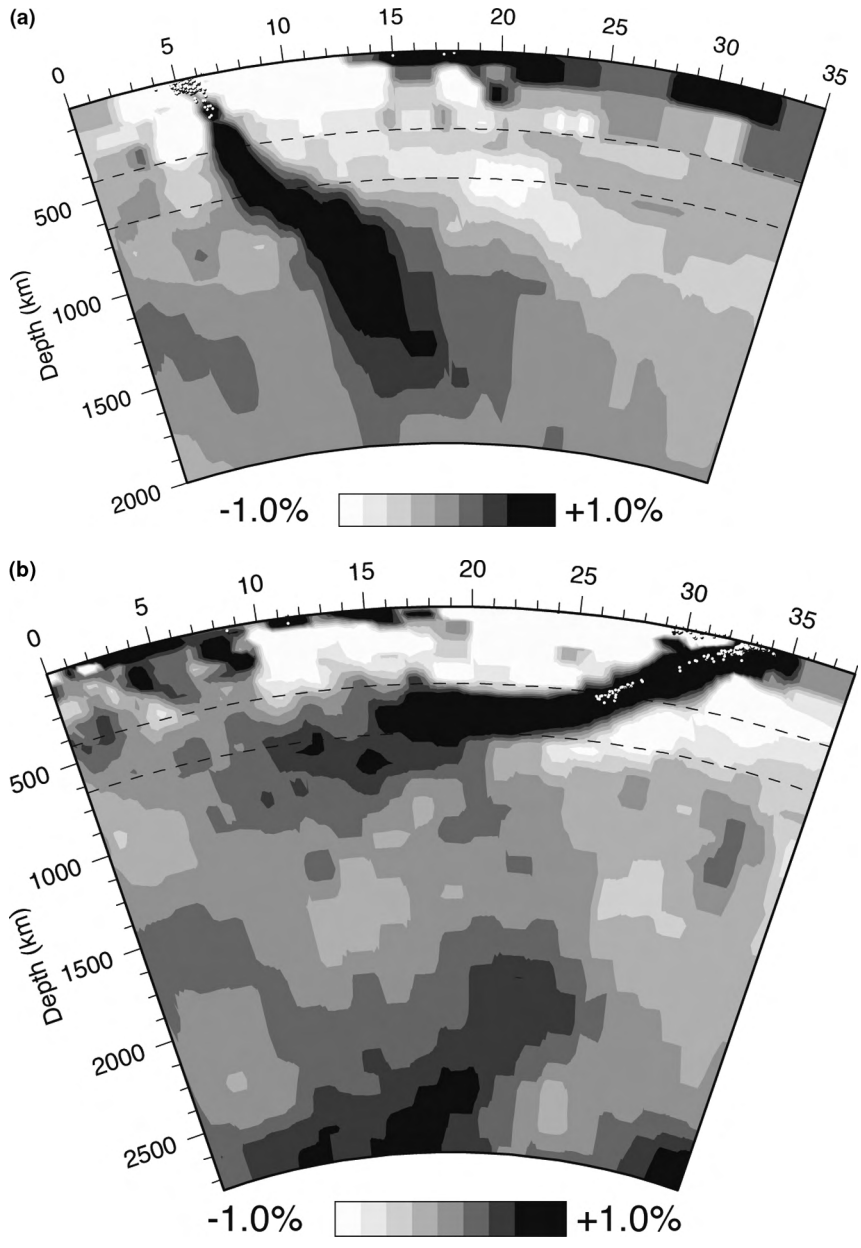


Figure 9.60. Deviations of seismic velocity from a standard model across subduction zones of the Pacific as determined from P-wave travel times: (a) Farallon, (b) Japan and (c) Tonga. (d) A map view of the velocity anomalies at a depth of 145 km beneath the Tonga arc and the Lau back-arc region. Note the high velocities of the cold subducting plate. The plate boundary and coastlines are shown as solid lines, earthquakes as white circles. Colour versions Plate 25. (W. Spakman, personal communication 2003.)

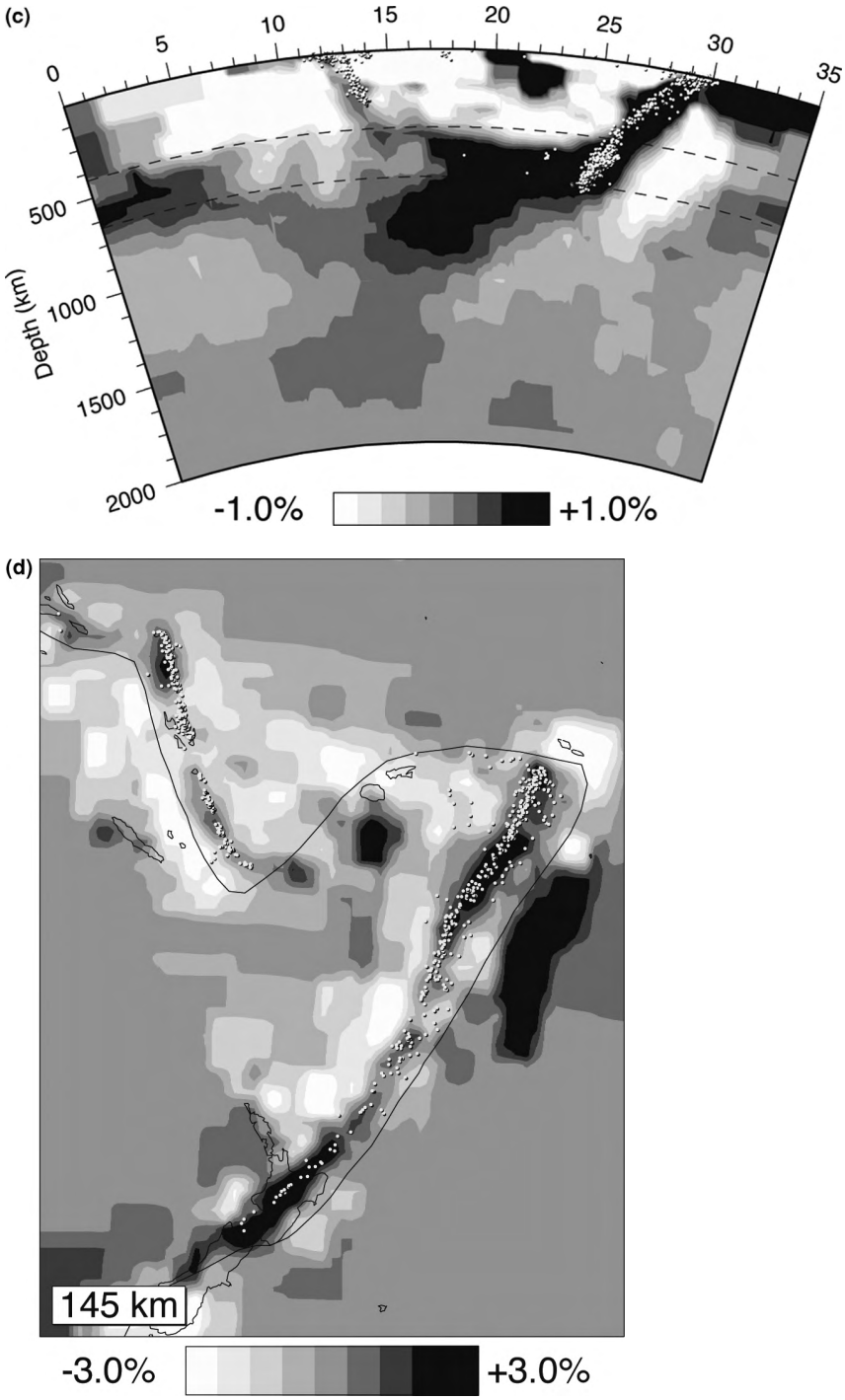


Figure 9.60. (*cont.*)

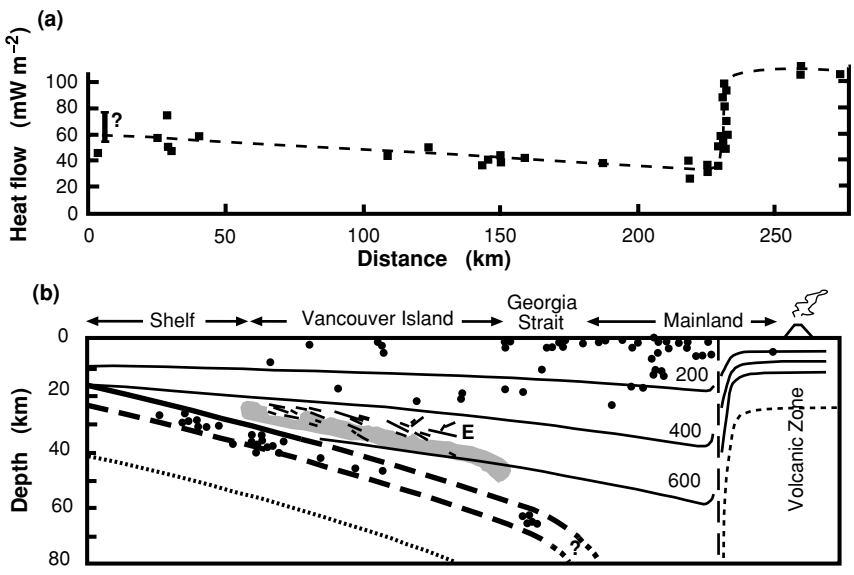


Figure 9.61. Structure across the Cascadia subduction zone of western Canada where the Juan de Fuca plate is descending beneath the North American plate. Land is shown by solid horizontal bar. (a) Surface heat-flow measurements. (b) Estimated isotherms, earthquake foci (dots), strongest E reflectors (short lines) and zone of high electrical conductivity assumed to be associated with water (shaded zone). (From Lewis *et al.* (1988) and Hyndman (1988).)

trench and the marginal basin, but not beneath Honshu. There the highest velocity measured was 7.5 km s^{-1}. This low velocity is characteristic of the asthenosphere.

The density model for the Chile Trench and Andes was constrained by refraction data, which indicate that the crust beneath the Andes is some 60 km thick, the upper 30 km having a seismic P-wave velocity of 6.0 km s^{-1} and the lower 30 km a velocity of 6.8 km s^{-1}. This thick crust appears to have grown from underneath by the addition of andesitic material from the subduction zone rather than by compression and deformation of sediments and pre-existing crustal material.

The convergent plate boundary off the west coast of North America, where the North American plate is overriding the young Juan de Fuca plate (Fig. 2.2) at about 2 cm yr^{-1}, is an example of a subduction zone with no easily distinguishable bathymetric trench. The dip of the subducting plate here is very shallow, about 15°. The subducting oceanic plate is overlain by a complex of accreted terranes, which are exposed on Vancouver Island and the mainland. This is, therefore, not a simple subduction zone but one where subduction has assembled a complex assortment of materials and pushed or welded them onto the North American continent. Figure 9.60(a) shows that there is a major high-seismic-velocity anomaly extending to at least 1500 km depth, namely the subducted Farallon plate, of which now the Juan de Fuca plate is all that remains at the surface (Sections 3.3.4 and 8.1.4). Figures 9.61 and 9.62(b) show the thermal structure and the seismic P-wave velocity structure across this subduction zone. The subducting oceanic plate and the 35-km-thick continental crust are clearly visible in the seismic model. One unusual feature of the velocity model is the wedges of somewhat higher-velocity material and bands of low-velocity material immediately above the Juan de Fuca plate. One interpretation is that the low-velocity material is

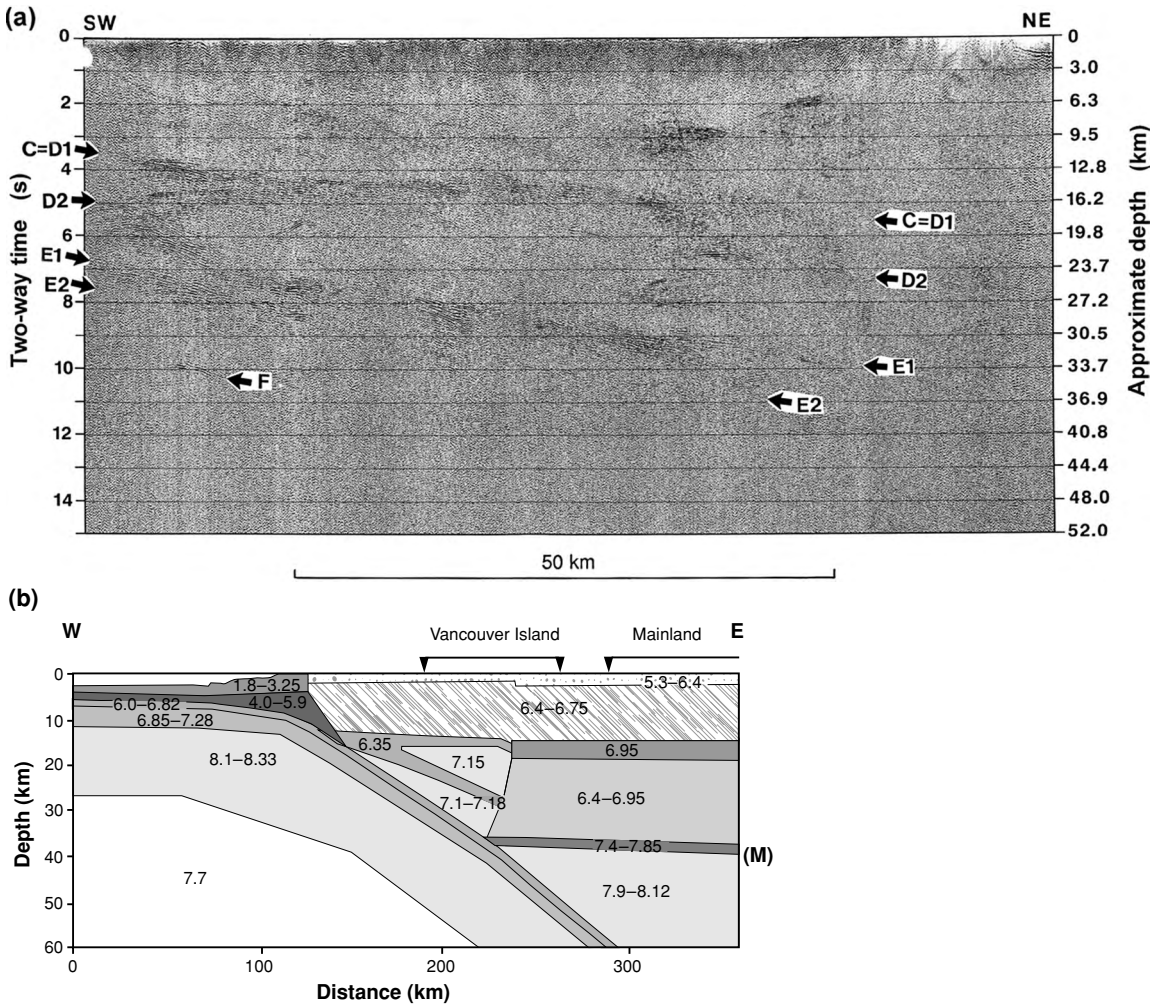


Figure 9.62. The Cascadia subduction zone. (a) A migrated seismic-reflection section across Vancouver Island, Canada. Reflections C and D are associated with the base of a major accreted terrane; reflections E may be associated with water in the crust; reflection F is from the top of the subducting Juan de Fuca plate. (Courtesy of R. M. Clowes.) (b) The P-wave velocity model determined from refraction and reflection data. M, Moho. (From Drew and Clowes (1989).) (c) Perturbation in S-wave velocity on a ~250-km-long section across the subduction zone. This section is ~500 km south of (a) and (b). Solid triangle, volcanic arc. Colour version Plate 26. (After Bostock *et al*. (2002) and Rondenay *et al*. (2001).) Details from reflection lines across the subduction zone. (d) The thin reflection zone where the subduction thrust is locked. (e) Further to the east, where aseismic slip is occurring, the thrust shows up as a thick band of reflections (see Fig. 9.54) (From Nedimovic *et al*. (2003).)

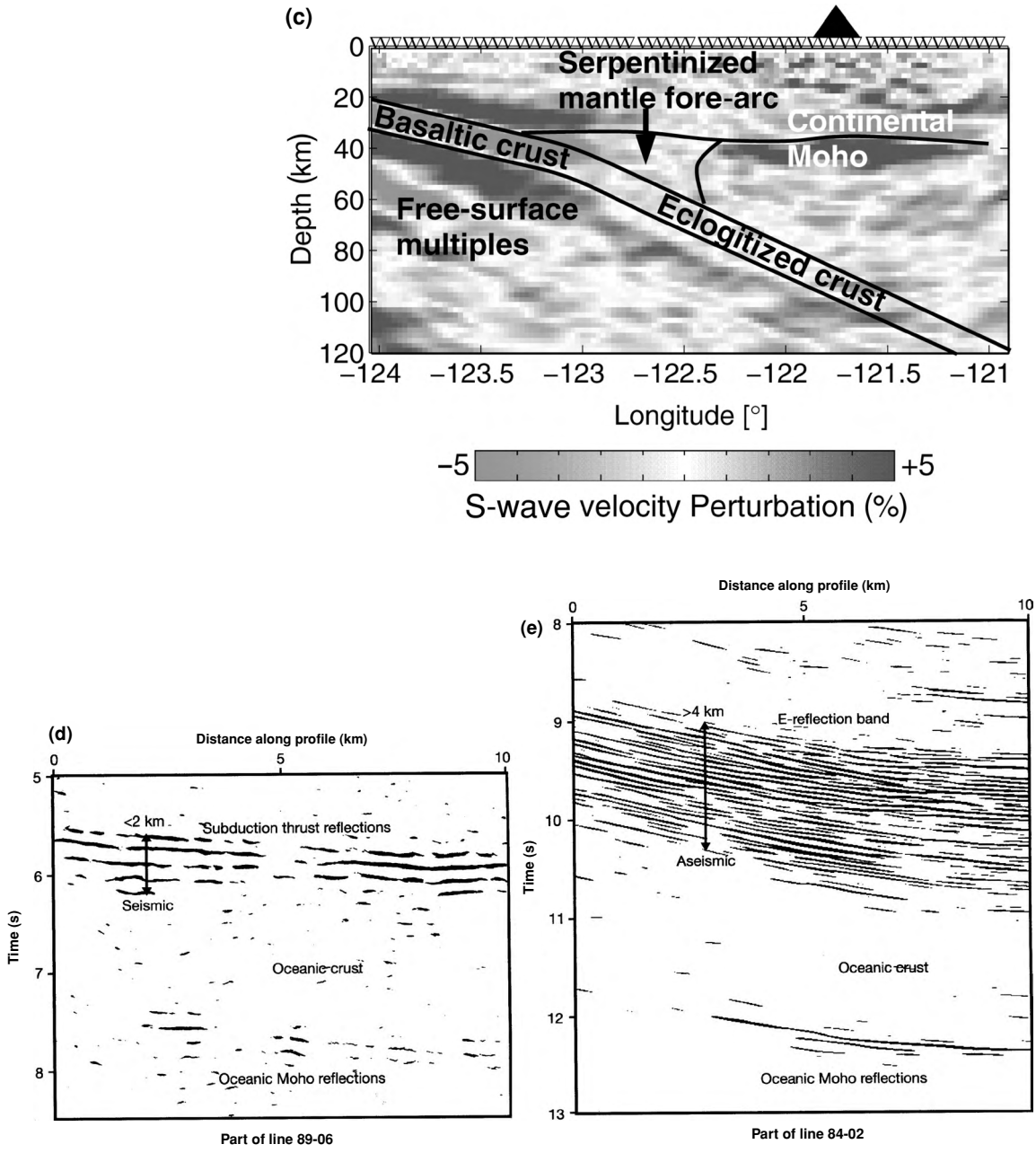


Figure 9.62. (*cont.*)

associated with water lost from the subducting plate. Another interpretation is that the wedge structure is tectonically underplated oceanic lithosphere. Such underplating could have been a continuous process, scraping off slivers of oceanic crust, or it could have occurred rapidly if the subduction zone jumped westwards. Nevertheless, it should be stressed that, however convincing any schematic model

appears, it is only as good as the data on which it is based. Other interpretations of the seismic data may be possible, so this wedge might in reality not be exactly as shown in Fig. 9.62.

The fine seismic structure of this convergent margin has been imaged by several deep reflection lines. (For details of shallow-seismic-reflection data on this margin see Figs. 4.43 and 4.44.) Figure 9.62(a) shows data from a line shot across Vancouver Island coincident with the refraction line shown in Fig. 9.62(b). Two very clear laminated reflections, here marked C and E, were seen on all the Vancouver Island reflection lines. The C reflection, which dips at 5–8°, is believed to be from the decollement zone (detachment surface) at the base of one of the accreted terranes (called Wrangellia). The E reflector, which dips at 9–13°, may represent a zone of porous sediments and volcanic rocks or may mark the location of trapped water in the crust. Figure 9.62(c) shows the perturbation in S-wave velocity on a ∼250-km-long section across this subduction zone. The continental Moho can be clearly seen east of −122.3° as a boundary between low-velocity continental crust and high-velocity mantle. However, to the west there is no clear continental Moho – the very low S-wave velocities in that part of the mantle wedge are consistent with the mantle being highly hydrated and serpentinized peridotite.

9.6.6 Chemistry of subduction-zone magmas

The igneous rocks above subduction zones include granites, basalts and andesites as well as some ultramafic rocks (see Section 9.1.1 and Table 9.1). The igneous rocks of the young Pacific island arcs such as the Tonga and Mariana arcs are primarily basalt and andesite. However, the older island arcs such as the Japan Arc are characterized by andesite volcanoes as well as diorite intrusions. Figure 9.59(d) shows a schematic geological cross section through the Andes. Although the andesite volcanoes provide the surface evidence of the active subduction zone beneath, the considerable thickening of the crust beneath the Andes is presumed to reflect the presence of large igneous intrusions.

The subducting plate produces partial melting in a number of ways. The basalts erupted above subduction zones result from partial melting of the mantle above the subducting plate. The loss even of small quantities of water from the subducting plate into the overriding mantle is sufficient to lower the melting temperature considerably (see Fig. 10.6). However, the magma which produces the andesite volcanics and the diorite intrusions forms either from partial melting of the subducted oceanic crust and sediments or, mostly, from melting in the overriding mantle wedge. The melt then collects beneath the overriding crust, where it fractionates. The subducted oceanic mantle does not undergo partial melting because it is already depleted-mantle material. Thus, the ‘volcanic line’ marks the depth at which material in the subducted plate or overlying mantle first reaches a high enough temperature for partial melting to occur. At shallow levels, partial melting is likely to produce basaltic magma; at greater depth, the degree of partial melting

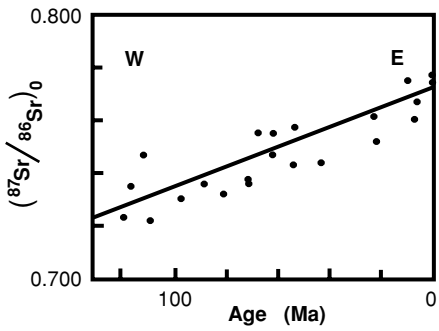


Figure 9.63. Initial $^{87}Sr/^{86}Sr$ ratio versus age for volcanic and plutonic rocks from the central Andes, 26°S–29°S. (Data from McNutt *et al.* (1975).)

decreases because much of the water has been lost from the subducting plate. This means that the magmas produced are likely, and are observed, to be more alkaline (more andesitic). It is also possible that they will be altered on their ascent through the greater thickness of mantle and crust. Ultramafic rocks found above subduction zones are presumably tectonically emplaced pieces of the residue remaining after partial melting of the overriding mantle. This subject is discussed more fully in Section 10.2.1.

The chemical compositions of lavas in island arcs are spatially zoned with respect to the subduction zone. The strontium isotope ratio exhibits some correlation with the depth to the subduction zone. In the Indonesian Arc, this ratio appears to increase slightly with depth to the subduction zone. In the central Andes, the ratio increases with distance from the trench; but this is an age effect as well as a depth effect since the youngest rocks are inland and the oldest on the coast (Fig. 9.63). In this case, the isotopic data indicate that the magma source moved progressively eastwards with time. The smallest, and oldest, initial ratio of 0.7022 is in good agreement with the 0.702–0.704 which would be expected for the oceanic crustal basalts and mantle that presumably melted to form the sampled rocks. As time progressed and the magma source moved eastwards, the rising magma would have to migrate through increasing thicknesses of crust; thus, the likelihood of contamination of the strontium isotope ratio is high. Subducted sea water, which would have a ratio of about 0.707, could also contaminate the magma. Thus, the greatest measured value of 0.7077 is still entirely consistent with a mantle or oceanic-crust origin for these rocks.

Back-arc spreading centres are necessarily influenced by the subduction zone. Results from the Tonga arc and the spreading centres in the Lau Basin immediately behind it show that lavas from the central Lau spreading centre were generated by decompression melting in the garnet stability field and are indistinguishable from MORB. In contrast, lavas from the Valu Far spreading centre to the south, which is within 50 km of the arc volcanoes, are similar to the arc lavas. Results of geochemical studies indicate that, like the arc lavas, they were derived from the mantle wedge, but with lower levels of fluid from the subducted slab.

9.7 Oceanic islands

Oceanic island chains represent anomalies in the oceanic lithosphere, being locations away from the plate boundaries where considerable volcanic and microearthquake activity is taking place. However, because of this they have been very useful in advancing an understanding of the physical properties of the plates and the underlying mantle.

Many of the details of seamount chains and oceanic islands have been discussed elsewhere in this book: dating of seamount chains and the hotspot reference frame (Chapter 2), gravity and flexure of the oceanic lithosphere due to the loading of the Hawaiian Ridge (Section 5.7) and possible origins for oceanic-island basalts (Chapter 8). In this section some details of the seismicity beneath Hawaii and the seismic structure of the crust and upper mantle are presented.

Hawaii is the best-studied active oceanic island. Work by the staff at the Hawaii Volcano Observatory, located on the rim of the Kilauea crater, and others has built up a detailed picture of the processes taking place beneath the island. Figure 9.64 shows the P-wave velocity and density structure of the island of Hawaii on a profile crossing the Mauna Loa volcano, which, together with the neighbouring volcano Kilauea and seamount Loihi, marks the present location of the Hawaiian hotspot. The density values for the layers defined by the P-wave velocity were calculated by assuming a relation between velocity and density similar to that shown in Fig. 4.2. Some small adjustments to the seismic layering were necessary in order to fit the gravity data well. The seismic-velocity and density models, taken together, show that the crust beneath Mauna Loa thickens to some 18 km and that most of this material is of high density and velocity. The oceanic crust on which the volcano has formed is bent downwards by the load, in accordance with the flexural models. A schematic geological interpretation of these models is shown in Fig. 9.64(d). The high-velocity, high-density intrusive core of the volcano is interpreted as a sequence of densely packed dykes similar to the sheeted-dyke complex proposed for the upper portions of oceanic layer 3 (see Fig. 9.5(b)).

Figure 9.65 shows the seismic activity occurring down to depths of 60 km along two profiles across the island of Hawaii. In Fig. 9.65(a) the epicentres are for all earthquakes from 1970 to 1983; in Fig. 9.65(b) the epicentres are for all long-period earthquakes from 1972 to 1984. Both sets of data show an extensive shallow (less than about 13 km depth) zone of activity. On the basis of seismic and density models, these shallow earthquakes, which often occurred in swarms, were all in the crust and are judged to be of volcanic origin. The deeper earthquakes are larger in magnitude and appear to be tectonic. The Kilauea magma conduit can be traced down to about 30 km. Deeper than 30 km the events merge into a broad zone. The magma-transport systems for Mauna Loa, Kilauea and the seamount Loihi appear to be connected to this deep zone.

The chemistry and dynamics of the Hawaiian magma supply have been studied in great detail. The top of the magma-source zone beneath Hawaii is at about

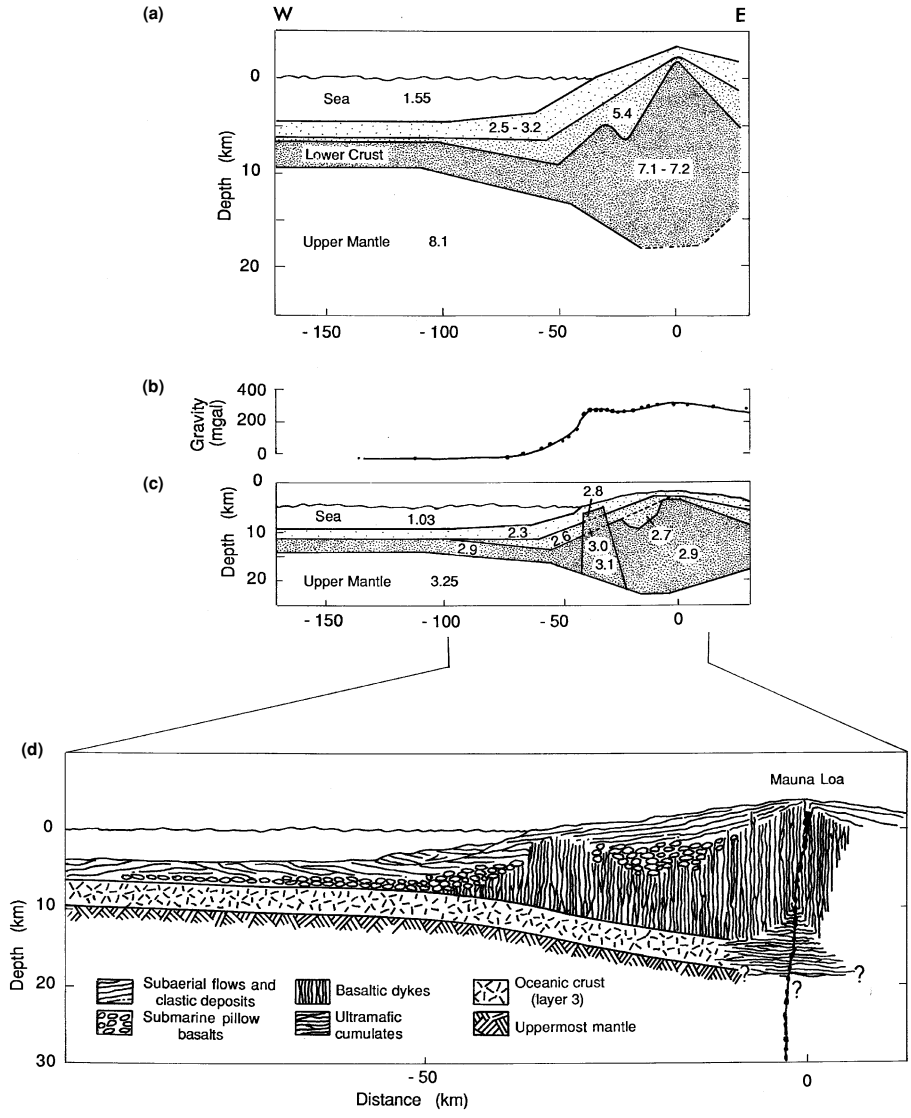


Figure 9.64. WSW–ENE profiles across the west coast of the island of Hawaii and the summit of Mauna Loa volcano. (a) The P-wave velocity structure determined from seismic refraction experiments. The vertical exaggeration is 4 : 1. (b) Bouguer gravity anomaly onshore, free-air gravity anomaly offshore. Dots, measured anomalies; solid line, a matching anomaly computed from the density model in (c). (c) The density structure based on the seismic-velocity structure shown in (a). The vertical exaggeration is 2 : 1. (d) A schematic geological structure based on the velocity and density models. The central magma conduit and the summit magma chamber are shown solid black. True scale. (From Hill and Zucca (1987).)

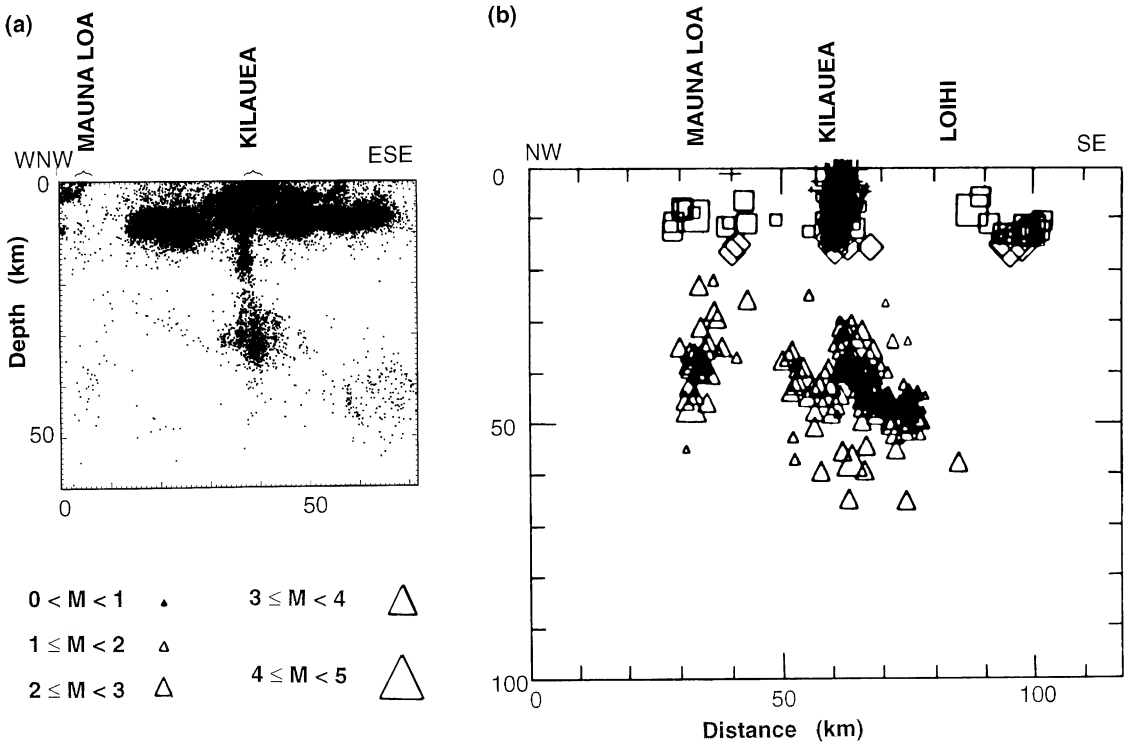


Figure 9.65. The distribution of earthquakes recorded by the Hawaiian Volcano Observatory along two cross sections through the island of Hawaii, approximately perpendicular to those shown in Fig. 9.64. (a) All earthquakes, 1970–1983. True scale. (b) Long-period earthquakes, 1972–1984. The magnitudes of the events are indicated by the size of the symbol: events shallower than □, 13 km; ◇, 13–20 km; and △, 20 km. True scale. (From Klein *et al.* (1987) and Koyanagi *et al.* (1987).)

60 km depth. From there it rises through the plumbing system to the active vents. Figure 9.66 shows a detailed view of the magma conduit beneath Kilauea. This model was defined by well-located, magma-related earthquake foci; such definition is possible because fractures in the rocks around the magma chambers and conduits open as the pressure of the magma increases. Thus, the maximum lateral extent of the conduit at any depth corresponds to the region of hydraulically induced seismicity at that depth. The conduit system enables the primary conduit to supply magma to any of the structure vents or fissures. There are magma-supply pathways from the summit magma reservoir beneath Kiluaea caldera both to the eastern and to the southwestern rift zone.

Oceanic islands are inherently prone to failure. They are great upward-built piles of rubble, set on oceanic sediments and down-warped crust. During the building of an island, some detachment may occur at the contact between the

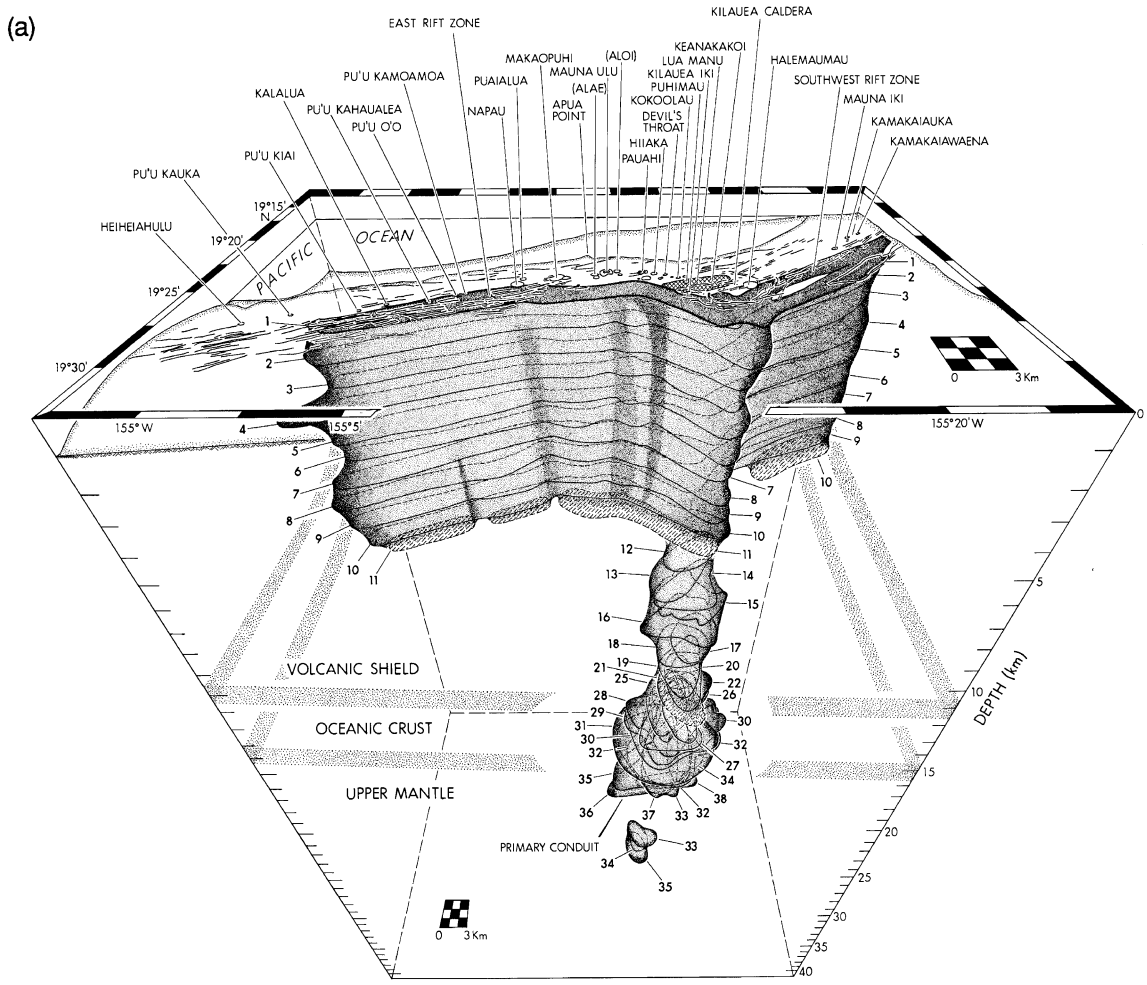


Figure 9.66. (a) A view southward of the magma plumbing system beneath Kilauea volcano. The transition zones from the volcanic shield to oceanic crust and from oceanic crust to the upper mantle are stippled. Square 1-km-scale grids are located at the surface and the base of the model (depth 37 km). Individual conduit cross sections are labelled with their depths beneath Kiluaea's caldera floor. (b) A view eastward of the internal structure of the primary conduit, which is displayed as a series of segments to show its internal structure. The inner transport core of the conduit is shaded light grey. Where parts of two sections overlap, the core is shaded medium grey; where three overlap, dark grey. Solid arrows show the retreat of the zone through which magma ascended from the 1969–1975 interval to the 1975–1982 interval when ascent of magma was restricted to the core of the conduit. (From Ryan (1988).)

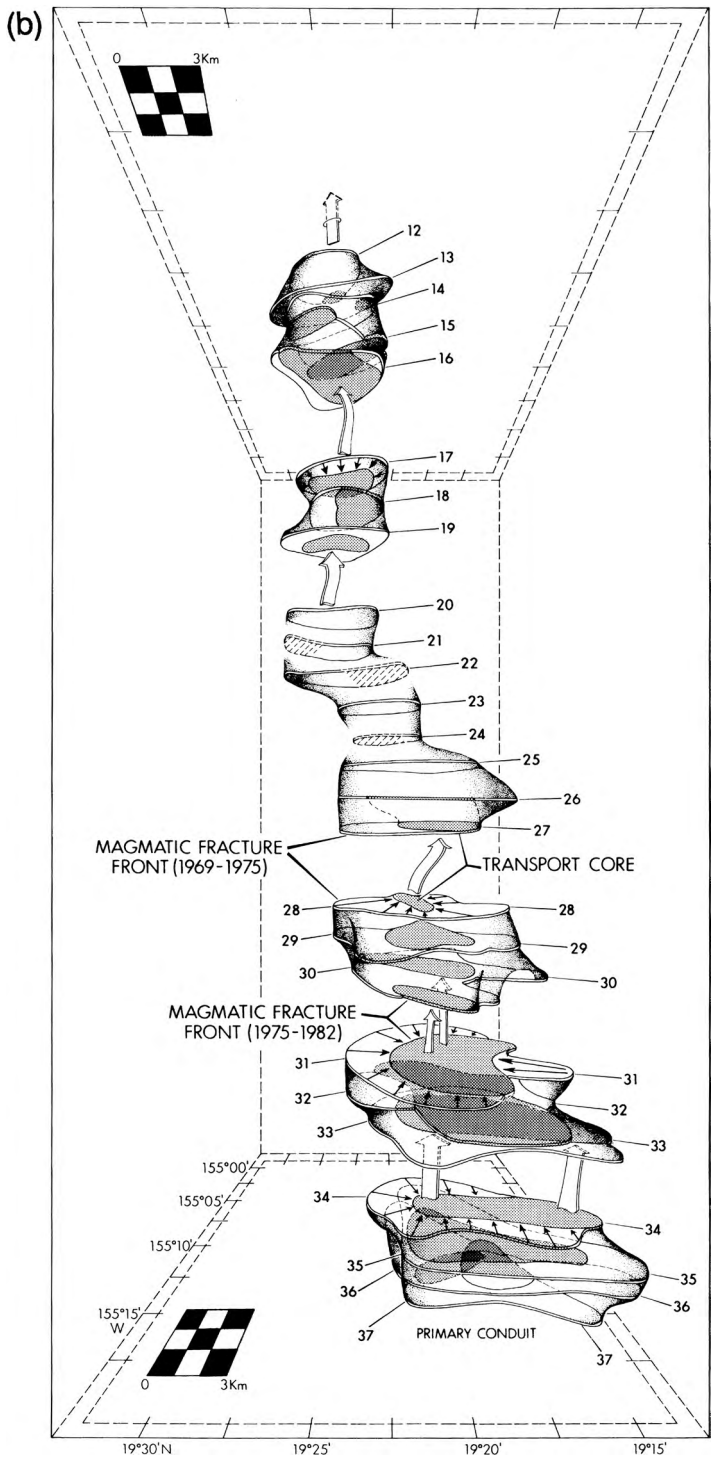


Figure 9.66. (*cont.*)

ocean sediment and the basalt lavas; as the volcano grows, the foot of the massive pile thrusts outwards. These detachments tend to slope upwards away from the centre of volcanism. The upper parts of the pile have bedding planes that slope downwards away from the centre. Eventually, failure occurs in the upper part and massive landslips move out into the surrounding ocean. The landslip headwall becomes a high cliff, up to 1000 m high. Often failure is catastrophic, spawning tsunami, with the debris transported out as deep-water turbidites. Examples of this can be seen in many volcanic islands. Some of the older Hawaiian islands have huge debris fields offshore with giant blocks of slumped rock. In the Atlantic there is concern about potential failure in the Canary Isles, which may be a major present-day natural hazard.

Problems

1. (a) Assuming isostatic equilibrium and Airy-type compensation, calculate the thickness of the oceanic crust if the continents averaged 50 km thick.
   (b) What is the minimum possible thickness of the continental crust? (Use densities of sea water, crust and mantle of 1.03×10^3, 2.9×10^3 and 3.3×10^3 kg m^{-3}; the ocean-basin depth is 5 km.)
2. Determine the oceanic crustal structure for the wide-angle reflection–refraction data shown in Fig. 9.7(a).
   (a) Use the normal-incidence two-way travel times to estimate the depth of the seabed. (Use 1.5 km s^{-1} for the velocity of sound in sea water.)
   (b) Use the slope–intercept method to estimate the following: (i) an upper crustal P-wave velocity (use first arrivals at less than 10 km distance), (ii) the lower-crustal P-wave velocity (use P3), (iii) the upper-mantle velocity, (iv) an upper-crustal S-wave velocity (use S-wave arrivals at 10–15 km distance), (v) the lower-crustal S-wave velocity (use S3), (vi) the upper-mantle S-wave velocity and (vii) the thickness of the oceanic crust at this location.
   (c) Compare the crustal thickness obtained in (vii) with the value obtained using normal-incidence two-way times.
   (d) Calculate the ratios of P-wave velocity to S-wave velocity. Do these values fall within the expected range?
3. Derive and plot the relationship between continental crustal thickness and ocean depth. Assume isostatic equilibrium and Airy-type compensation.
4. If the magma chamber on the East Pacific Rise at 12°N described in the example in Section 9.4 were filled with molten basaltic magma with seismic P-wave velocity 3 km s^{-1}, how wide would it be? At the other extreme, if it contained only 10% partial melt, what would its width be? Comment on the likelihood of detection of such extreme magma chambers by seismic-reflection and -refraction experiments.
5. What would be the likelihood of delineating the magma chambers of Problem 4 by seismic methods if the dominant frequency of your signal were (a) 5 Hz, (b) 15 Hz and (c) 50 Hz?

6. What are the criteria for deciding whether the Earth's crust at any location is oceanic or continental in origin?
7. The structure of the oceanic crust is very much the same irrespective of whether it was created at a slow- or a fast-spreading ridge. Discuss why. Should we also expect to find such worldwide lack of variation in the continental crust?
8. Calculate the lithospheric thicknesses at the ridge–transform-fault intersection shown in Fig. 9.32. Use the lithospheric thickness–age model $z(t) = 11t^{1/2}$ with z in kilometres and t in Ma.
9. (a) Draw a bathymetric profile along the north–south 10-Ma isochron of Fig. 9.32.
   (b) Draw a similar profile along a north–south line crossing the continuation of the fracture zone 500 km to the east.
10. Draw fault-plane solutions for earthquakes occurring within the active zone of the transform fault shown in Fig. 9.32. Discuss the relative frequency of earthquakes occurring between the two ridge segments and outside this zone. Draw a fault-plane solution for an earthquake occurring on the fault at the 50-Ma isochron on (a) the west and (b) the east side of the fault.
11. Calculate the topographic relief across a transform fault if the age offset is (a) 10 Ma, (b) 20 Ma and (c) 50 Ma. Plot the relief against distance (age) from the ridge axis for these three faults. What can you deduce about earthquakes on the inactive portions of these faults?
12. The boundary between the African and Eurasian plates between the Azores Triple Junction and Gibraltar runs approximately east–west. What is the nature of the boundary? Draw fault-plane solutions for earthquakes occurring along it.
13. What would happen to the plates shown in Fig. 9.32 if the rotation pole were suddenly to move so that the spreading rate remained unchanged in magnitude but altered 20° in azimuth? Illustrate your answer with diagrams.
14. It is possible to make a first approximation to the temperatures along a transform fault by assuming that the temperature at any point along the transform is the average of the temperatures on either side. Using the cooling half-space model (Section 7.5.2), plot 400, 600 and 800-°C isotherms beneath a 20-Ma-offset transform fault. What assumptions have you made? If the base of the plate is assumed to be the 1000-°C isotherm, how does this model agree with the depth to which faulting occurs on transform faults?

References and bibliography

Abers, G. A. 1992. Relationship between shallow- and intermediate-depth seismicity in the eastern Aleutian subduction zone. *Geophys. Res. Lett.*, **19**, 2019–22.

Abrams, L. J., Detrick, R. S. and Fox, P. J. 1988. Morphology and crustal structure of the Kane fracture zone transverse ridge. *J. Geophys. Res.*, **93**, 3195–210.

Anderson, H. and Jackson, J. A. 1987. Active tectonics of the Adriatic region. *Geophys. J. Roy. Astr. Soc.*, **91**, 937–84.

Anderson, R. N., DeLong, S. E. and Schwarz, W. M. 1978. Thermal model for subduction with dehydration in the downgoing slab. *J. Geol.*, **86**, 731–9.

1980. Dehydration, asthenospheric convection and seismicity in subduction zones. *J. Geol.*, **88**, 445–51.

Ansell, J. H. and Gubbins, D. 1986. Anomalous high-frequency wave propagation from the Tonga–Kermadec seismic zone to New Zealand. *Geophys. J. Roy. Astr. Soc.*, **85**, 93–106.

Baker, E. T., Chen, Y. J. and Phipps Morgan, J. 1996. The relationship between near-axis hydrothermal cooling and the spreading rate of mid-ocean ridges. *Earth Planet. Sci. Lett.*, **142**, 137–45.

Barazangi, M. and Isacks, B. L. 1971. Lateral variations of seismic-wave attenuation in the upper mantle above the inclined earthquake zone of the Tonga Island arc: deep anomaly in the upper mantle. *J. Geophys. Res.*, **76**, 8493–516.

Batiza, R. 1996. Magmatic segmentation of mid-ocean ridges: a review. *Geol. Soc. Spec. Pub.*, **118**, 103–30.

Bell, R. E. and Buck, W. R. 1992. Crustal control of ridge segmentation inferred from observations of the Reykjanes Ridge. *Nature*, **357**, 583–6.

Benioff, H. 1949. Seismic evidence for the fault origin of oceanic deeps. *Geol. Soc. Am. Bull.*, **60**, 1837–56.

Bergman, E. A. and Solomon, S. C. 1984. Source mechanisms of earthquakes near mid-ocean ridges from body waveform inversion, implications for the early evolution of oceanic lithosphere. *J. Geophys. Res.*, **89**, 11 415–41.

1990. Earthquake swarms on the Mid-Atlantic Ridge: products of magmatism or extensional tectonics? *J. Geophys. Res.*, **95**, 4943–65.

Bevis, M. and Isacks, B. L. 1984. Hypocentral trend surface analysis: probing the geometry of Benioff zones. *J. Geophys. Res.*, **89**, 6153–70.

Bird, P. 1978. Stress and temperature in subduction shear zones. *Geophys. J. Roy. Astr. Soc.*, **55**, 411–34.

Blackman, D. K., Cann, J. R., Janssen, B. and Smith, D. K. 1998. Origin of extensional core complexes: evidence from the Mid-Atlantic Ridge at Atlantis Fracture Zone. *J. Geophys. Res.*, **103**, 21 315–33.

Bostock, M. G., Hyndman, R. D., Rondenay, S. and Peacock, S. M. 2002. An inverted continental Moho and serpentinization of the forearc mantle. *Nature*, **417**, 536–8.

Bowen, A. N. and White, R. S. 1986. Deep-tow seismic profiles from the Vema transform and ridge transform intersection. *J. Geol. Soc. Lond.*, **143**, 807–17.

Bown J. W. and White, R. S. 1994. Variation with spreading rate of crustal thickness and geochemistry. *Earth Planet Sci. Lett.*, **121**, 435–49.

Boyd, T. M., Taber, J. J., Lerner-Lam, A. L. and Beavan, J. 1988. Seismic rupture and arc segmentation within the Shumagin Islands Seismic Gap, Alaska. *Geophys. Res. Lett.*, **5**, 201–4.

Bratt, S. R. and Purdy, G. M. 1984. Structure and variability of oceanic crust on the flanks of the East Pacific Rise between 11° and 13°N. *J. Geophys. Res.*, **89**, 6111–25.

Bratt, S. R. and Solomon, S. C. 1984. Compressional and shear wave structure of the East Pacific Rise at 11° 20′N, constraints from three-component ocean bottom seismometer data. *J. Geophys. Res.*, **89**, 6095–110.

Brown, G. C. and Hennessy, J. 1978. The initiation and thermal diversity of granite magmatism. *Phil. Trans. Roy. Soc. Lond.* A, **288**, 631–43.

Brownlow, A. H. 1979. *Geochemistry*. Englewood Cliffs, New Jersey: Prentice-Hall.

Brudzinski, M. R. and Chen, W.-P. 2003. A petrologic anomaly accompanying outboard earthquakes beneath Fiji–Tonga: corresponding evidence from broadband P and S waveforms, *J. Geophys. Res.*, **108**, 2299, doi: 10.1029/2002JB002012.

Bryan, W. B. and Moore, I. G. 1977. Compositional variation of young basalts in the mid-Atlantic Ridge rift valley near lat. 36° 49′N. *Geol. Soc. Am. Bull.*, **88**, 556–70.

Bunch, A. W. H. and Kennett, B. L. N. 1980. The crustal structure of the Reykjanes Ridge at 59° 30′N. *Geophys. J. Roy. Astr. Soc.*, **61**, 141–66.

Burchfiel, B. C. 1983. The continental crust. *Sci. Am.*, **249** (3), 57–65.

Canales, J. P., Detrick, R. S., Lin, J. and Collins, J. A. 2000. Crustal and upper mantle structure beneath the rift mountains and across a non-transform offset at the Mid-Atlantic Ridge (35°N). *J. Geophys. Res.*, **105**, 2699–719.

Cann, J. R. 1974. A model for oceanic crustal structure developed. *Geophys. J. Roy. Astr. Soc.*, **39**, 169–87.

Cann, J. R., Blackman, D. K., Smith, D. K., McAllister, E., Janssen, B., Mello, S., Avgerinos, E., Pascoe, A. R. and Escartin, J. 1997. Corrugated slip surfaces formed at ridge–transform intersections on the Mid-Atlantic Ridge. *Nature*, **385**, 329–32.

Cannat, M. 1996. How thick is the oceanic crust at slow spreading oceanic ridges? *J. Geophys. Res.*, **101**, 2847–57.

Cannat, M. *et al*. 1995. Thin crust, ultramafic exposures, and rugged faulting at the Mid-Atlantic Ridge (22°–24°N). *Geology*, **23**, 49–52.

Carlson, R. L. and Raskin, G. S. 1984. Density of ocean crust. *Nature*, **311**, 555–8.

Cary, P. W. and Chapman, C. H. 1988. Waveform inversion of expanding spread profile S from the North Atlantic transect. *J. Geophys. Res.*, **93**, 13 575–88.

Chapman, M. and Talwani, M. 1979. Comparison of gravimetric geoids with Geos 3 altimeter. *J. Geophys. Res.*, **84**, 3803–16.

Chen, Y. and Phipps Morgan, J. 1990. Rift valley/no rift valley transition at mid-ocean ridges. *J. Geophys. Res.*, **95**, 17 571–81.

Chiang, C. S. and Detrick, R. S. 1985. The structure of the lower oceanic crust from synthetic seismogram modelling of near-vertical and wide-angle reflections and refractions near DSDP Site 417 in the western North Atlantic. *EOS Trans. Am. Geophys. Un.*, **66**, 956.

Christensen, N. I. and Smewing, J. D. 1981. Geology and seismic structure of the northern section of the Oman Ophiolite. *J. Geophys. Res.*, **86**, 2545–55.

Clague, D. A. and Dalrymple, G. B. 1987. The Hawaiian–Emperor volcanic chain. In R. W. Decker, T. L. Wright and P. H. Stauffer, eds., *Volcanism in Hawaii*. U.S. Geological Survey Professional Paper 1350. Denver, Colorado: U.S. Geological Survey, pp. 5–54.

Clague, J. J. 1997. Evidence for large earthquakes at the Cascadia subduction zone. *Rev. Geophys.*, **35**, 439–60.

Clift, P. D. and Vannuchi, P. 2004. Controls on tectonic accretion versus erosion in subduction zones: implications for the origin and recycling of the continental crust. *Rev. Geophys.*, **42**, RG 2001, doi: 10.1029/2003RG000127.

Coleman, R. G. 1977. *Ophiolites, Ancient Oceanic Lithosphere?* Berlin: Springer-Verlag.

Coulbourn, W. T. 1981. Tectonics of the Nazca plate and the continental margin of western South America, 18°S to 23°S. In L. D. Kulm, J. Dymond, E. J. Dasch and D. M. Hussong, eds., *Nazca Plate Crustal Formation and Andean Convergence*. Vol. 154 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America, pp. 587–618.

Cox, A. 1969. Geomagnetic reversals. *Science*, **163**, 237–45.

Creager, K. C. and Jordan, T. H. 1986. Slab penetration into the lower mantle beneath the Mariana and other island arcs of the northwest Pacific. *J. Geophys. Res*., **91**, 3573–89.

Dässler, R. and Yuen, D. A. 1996. The metastable olivine wedge in fast subducting slabs: constraints from thermo-kinetic coupling. *Earth Planet. Sci. Lett.*, **137**, 109–18.

Dauteuil, O. and Brun, J.-P. 1993. Oblique rifting in a slow-spreading ridge. *Nature*, **361**, 145–8.

Davies, J., Sykes, L., House, L. and Jacob, K. 1981. Shumagin seismic gap, Alaska peninsula: history of great earthquakes, tectonic setting and evidence for high seismic potential. *J. Geophys. Res.*, **86**, 3821–55.

Decker, R. W. 1987. Dynamics of Hawaiian volcanoes: an overview. In R. W. Decker, T. L. Wright and P. H. Stauffer, eds., *Volcanism in Hawaii*, U.S. Geological Survey Professional Paper 1350. Denver, Colorado: U.S. Geological Survey, pp. 997–1018.

DeMets, C., Gordon, R. G., Argus, D. F. and Stein, S. 1990. Current plate motions. *Geophysical J. Int*., **101**, 425–78.

Detrick, R. S., Buhl, P., Vera, E., Mutter, J., Orcutt, J., Madsen, J. and Brocher, T. 1987. Multichannel seismic imaging of a crustal magma chamber along the East Pacific Rise. *Nature*, **326**, 35–41.

Detrick, R. S., Harding, A. J., Kent, G. M., Orcutt, J. A., Mutter, J. C. and Buhl, P. 1993. Seismic structure of the southern East Pacific Rise. *Science*, **259**, 499–503.

Dick, H. J. B. *et al*. 2000. A long in-situ section of the lower ocean crust: results of ODP leg 176 drilling at the Southwest Indian Ridge. *Earth Planet. Sci. Lett.*, **179**, 31–51.

Dilek, Y. 2003. Ophiolite concept and its evolution. In Y. Dilek and S. Newcomb, eds., *Ophiolite Concept and the Evolution of Geological Thought*. Geological Society of America Special Paper 373. Boulder, Colorado: Geological Society of America, pp. 1–16.

Dilek, Y., Moores, E. M. and Furnes, H. 1998. Structure of modern oceanic crust and ophiolites and implications for faulting and magmatism at oceanic spreading centres. In W. R. Buck, P. T. Delaney, J. A. Karson and Y. Lagabrielle, eds., *Faulting and Magmatism at Mid-Ocean Ridges*. Geophysical Monograph 106. Washington: American Geophysical Union, pp. 219–65.

Drew, J. J. and Clowes, R. M. 1989. A re-interpretation of the seismic structure across the active subduction zone of western Canada. In *Studies of Laterally Heterogeneous Structures Using Seismic Refraction and Reflection Data*. Proceedings of the 1987 Commission on Controlled Source Seismology Workshop, Geological Survey of Canada Paper 89-13. Ottawa: Geological Survey of Canada.

Driscoll, M. and Parsons, B. 1988. Cooling of the oceanic lithosphere – evidence from geoid anomalies across the Udintsev and Eltanin fracture zones. *Earth Planet. Sci. Lett.*, **88**, 289–307.

Duschenes, J. D. and Solomon, S. C. 1977. Shear wave travel time residuals from oceanic earthquakes and the evolution of the oceanic lithosphere. *J. Geophys. Res.*, **82**, 1985–2000.

Eastabrook, C. H., Jacob, K. H. and Sykes, L. R. 1994. Body wave and surface wave analysis of large and great earthquakes along the Eastern Aleutian Arc, 1923–1993: implications for future events. *J. Geophys. Res.*, **99**, 11 643–62.

Edmond, J. H. and Von Damm, K. 1983. Hot springs on the ocean floor. *Sci. Am.*, **248** (4), 78–93.

Eicher, D. L., McAlester, A. L. and Rottman, M. L. 1984. *The History of the Earth's Crust*. Englewood Cliffs, New Jersey: Prentice-Hall.

Elliott, T., Plank, T., Zindler, A., White, W. and Bourden, B. 1997. Element transport from slab to volcanic front at the Mariana arc. *J. Geophys. Res.*, **102**, 14 991–15 019.

EMSLAB Group. 1988. The EMSLAB electromagnetic sounding experiment. *EOS Trans. Am. Geophys. Un.*, **69**, 89.

Engeln, J. F., Wiens, D. A. and Stein, S. 1986. Mechanisms and depths of Mid-Atlantic Ridge transform faults. *J. Geophys. Res.*, **91**, 548–78.

Ernst, W. G. 1976. *Petrologic Phase Equilibria*. San Francisco, California: Freeman.

Escartín, J. and Lin, J. 1995. Ridge offsets, normal faulting and gravity anomalies of slow speading ridges. *J. Geophys. Res.*, **100**, 6163–77.

Fisher, A. T. 1998. Permeability within basaltic oceanic crust. *Rev. Geophys.*, **36**, 143–82.

Fischer, K. M., Jordan, T. H. and Creager, K. C. 1988. Seismic constraints on the morphology of deep slabs. *J. Geophys. Res.*, **93**, 4773–83.

Fisher, R. L. and Goodwillie, A. M. 1997. The physiography of the Southwest Indian Ridge. *Marine Geophys. Res*., **19**, 451–5.

Flück, P., Hyndman, R. D. and Wang, K. 1997. Three-dimensional dislocation model for great earthquakes of the Cascadia subduction zone. *J. Geophys. Res.*, **102**, 20 539–50.

Forsyth, D. W. 1977. The evolution of the upper mantle beneath mid-ocean ridges. *Tectonophysics*, **38**, 89–118.

Forsyth, D. W. and Wilson, B. 1984. Three-dimensional temperature structure of a ridge–transform–ridge system. *Earth Planet. Sci. Lett.*, **70**, 355–62.

Forsyth, D. W. *et al*. 1998. Imaging the deep seismic structure beneath a mid-ocean rise: the MELT experiment. *Science*, **280**, 1215–17.

Fowler, C. M. R. 1976. Crustal structure of the Mid-Atlantic Ridge crest at 37°N. *Geophys. J. Roy. Astr. Soc.*, **47**, 459–91.

1978. The Mid-Atlantic Ridge: structure at 45°N. *Geophys. J. Roy. Astr. Soc.*, **54**, 167–82.

Fowler, S. R., White, R. S. and Louden, K. E. 1985. Sediment dewatering in the Makran accretionary prism. *Earth Planet. Sci. Lett.*, **75**, 427–38.

Fox, P. J. and Gallo, D. G. 1984. A tectonic model for ridge–transform–ridge plate boundaries: Implications for the structure of oceanic lithosphere. *Tectonophysics*, **104**, 205–42.

1989. Transforms of the Eastern Central Pacific. In E. L. Winterer, D. M. Hussong and R. W. Decker, eds., *The Geology of North America*, *Vol. N*, *The Eastern Pacific and Hawaii*. Boulder, Colorado: Geological Society of America, pp. 111–24.

Francis, T. J. G., Porter, I. T. and Lilwall, R. C. 1978. Microearthquakes near the eastern end of St. Paul's fracture zone. *Geophys. J. Roy. Astr. Soc.*, **53**, 201–17.

Frank, F. C. 1968. Curvature of island arcs. *Nature*, **220**, 363.

Fuis, G. S. 1998. West margin of North America – a synthesis of recent seismic transects. *Tectonophysics*, **288**, 265–92.

Furlong, K. P. and Chapman, D. S. 1982. Thermal modelling of the geometry of the tectonics of the overriding plate. *J. Geophys. Res.*, **87**, 1786–802.

German, C. R., Parson, L. M. and HEAT Scientific Team. 1996. Hydrothermal exploration near the Azores triple junction: tectonic control of venting at slow-spreading ridges? *Earth Planet. Sci. Lett.*, **138**, 93–104.

Goto, K., Suzuki, Z. and Hamaguchi, H. 1987. Stress distribution due to olivine–spinel phase transition in descending plate and deep focus earthquakes. *J. Geophys. Res.*, **92**, 13 811–20.

Goud, H. R. and Karson, J. A. 1986. Tectonics of short-offset, slow slipping transform zones in the FAMOUS area, Mid-Atlantic Ridge. *Mar. Geophys. Res.*, **7**, 489–514.

Green, A. G., Clowes, R. M., Yorath, C. J., Spencer, C., Kanasewich, E. R., Brandon, M. T. and Sutherland-Brown, A. 1986. Seismic reflection imaging of the subducting Juan de Fuca Plate. *Nature*, **319**, 210–13.

Green, A. G., Milkereit, B., Mayrand, L., Spencer, C., Kurtz, R. and Clowes, R. M. 1987. Lithoprobe seismic reflection profiling across Vancouver Island: results from reprocessing. *Geophys. J. Roy. Astr. Soc.*, **89**, 85–90.

Green II, H. W. 1994. Solving the paradox of deep earthquakes. *Sci. Am.*, **271** (3), 50–7.

Green II, H.W. and Houston, H. 1995. The mechanics of deep earthquakes. *Ann. Rev. Earth Planet. Sci.*, **23**, 169–213.

Grimison, N. L. and Chen, W.-P. 1988. Source mechanisms of four recent earthquakes along the Azores–Gibraltar boundary. *Geophys. J. Roy. Astr. Soc.*, **92**, 391–402.

Grow, J. A. and Bowin, C. O. 1975. Evidence for high-density crust and mantle beneath the Chile Trench due to the descending lithosphere. *J. Geophys. Res.*, **80**, 1449–58.

Haberman, R. E., McCann, W. R. and Perin, B. 1986. Spatial seismicity variations along convergent plate boundaries. *Geophys. J. Roy. Astr. Soc.*, **85**, 43–68.

Hale, L. D., Morton, C. and Sleep, N. H. 1982. Reinterpretation of seismic reflection data over the East Pacific Rise. *J. Geophys. Res*, **87**, 7707–17.

Hasegawa, A., Umino, N. and Takagi, A. 1978. Double-planed deep seismic zone and upper-mantle structure in northeastern Japan Arc. *Geophys. J. Roy. Astr. Soc.*, **54**, 281–96.

Hay, W. W., Sloan II, J. L. and Wold, C. N. 1988. Mass/age distribution and composition of sediments on the ocean floor and the global rate of sediment subduction. *J. Geophys. Res.*, **93**, 14 933–40.

Heezen, B. C. 1962. The deep-sea-floor. In S. K. Runcorn, ed., *Continental Drift*. New York: Academic, pp. 235–68.

Heirtzler, J. R., Dickson, G. O., Herron, E. M., Pitman III, W. C. and LePichon, X. 1968. Marine magnetic anomalies, geomagnetic field reversals and motions of the ocean floor and continents. *J. Geophys. Res.*, **73**, 2119–36.

Hekinian, R. and Walker, D. 1987. Diversity and spatial zonation of volcanic rocks from the East Pacific Rise near 21°N. *Contrib. Mineral. Petrol.*, **96**, 265–80.

Hill, D. P. and Zucca, J. J. 1987. Geophysical constraints on the structure of Kilauea and Mauna Loa volcanoes and some implications for seismomagmatic processes. In R. W. Decker, T. L. Wright and P. H. Stauffer, eds., *Volcanism in Hawaii*, U.S. Geological Survey Professional Paper 1350. Denver, Colorado: U.S. Geological Survey, pp. 903–17.

Hill, M. N. 1957. Recent geophysical exploration of the ocean floor. *Phys. Chem. Earth*, **2**, 129–63.

Hobbs, R. W., Tong, C. H. and Pye, J. W. 2003. *Modelling and Processing of 3-D Seismic Data Collected over the Overlapping Spreading Centre on the East Pacific Rise at 9°03′N.* Special Publication of the Geological Society of London 212. London: Geological Society of London, pp. 251–9.

Honda, S. 1985. Thermal structure beneath Tohoku, northeast Japan: a case study for understanding the detailed thermal structure of the subduction zone. *Tectonophysics*, **112**, 69–102.

Huang, P. Y. and Solomon, S. C. 1987. Centroid depths and mechanisms of midocean ridge earthquakes in the Indian Ocean, Gulf of Aden and Red Sea. *J. Geophys. Res.*, **92**, 1361–82.

1988. Centroid depths of mid-ocean ridge earthquakes: dependence on spreading rate. *J. Geophys. Res.*, **93**, 13 445–7.

Huang, P. Y., Solomon, S. C., Bergman, E. A. and Nabelek, J. L. 1986. Focal depths and mechanisms of Mid-Atlantic Ridge earthquakes from body waveform inversion. *J. Geophys. Res.*, **91**, 579–98.

Humphris, S. E., Zierenberg, R. A., Mullineaux, L. S. and Thompson, R. E., eds., 1995. *Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Constraints*. Geophysical Monograph 91. Washington: American Geophysical Union.

Hyndman, R. D. 1988. Dipping seismic reflectors, electrically conductive zones and trapped water in the crust over a subducting plate. *J. Geophys. Res.*, **93**, 13 391–405.

1995. Giant earthquakes of the Pacific Northwest. *Sci. Am.*, **273** (6), 68–75.

Hyndman, R. D. and Wang, K. 1993. Tectonic constraints on the zone of major thrust earthquake failure: the Cascadia subduction zone. *J. Geophys. Res.*, **98**, 2039–60.

1995. The rupture zone of Cascadia great earthquakes from current deformation and the thermal regime. *J. Geophys. Res.*, **100**, 22 133–54.

Isacks, B. and Molnar, P. 1969. Mantle earthquake mechanisms and the sinking of the lithosphere. *Nature*, **223**, 1121–4.

Isacks, B., Oliver, J. and Sykes, L. 1968. Seismology and the new global tectonics. *J. Geophys. Res.*, **73**, 5855–99.

Isacks, B. L. and Barazangi, M. 1977. Geometry of Benioff zones: lateral segmentation and downwards bending of the subducted lithosphere. In M. Talwani and W. C. Pitman III, eds., *Island Arcs, Deep-sea Trenches and Back-arc Basins*. American Geophysical Union Maurice Ewing Series 1. Washington: American Geophysical Union, pp. 99–114.

Jackson, J. A. and McKenzie, D. P. 1984. Active tectonics of the Alpine–Himalayas belt between western Turkey and Pakistan. *Geophys. J. Roy. Astr. Soc.*, **77**, 185–266.

Jones, J. E. W. 1999. *Marine Geophysics*. Chichester: John Wiley.

Karson, J. A. and Dick, H. J. B. 1983. Tectonics of ridge–transform intersections at the Kane Fracture Zone. *Mar. Geophys. Res.*, **6**, 51–98.

Karson, J. A. *et al*. 1987. Along axis variability in seafloor spreading in the MARK area. *Nature*, **328**, 681–5.

Keen, C. E. and Tramontini, C. 1970. A seismic refraction survey on the Mid-Atlantic Ridge. *Geophys. J. Roy. Astr. Soc.*, **20**, 473–91.

Kent, G. M. *et al*. 2000. Evidence from three-dimensional seismic reflectivity images for enhanced melt supply beneath mid-ocean-ridge discontinuities. *Nature*, **406**, 614–18.

Khitarov, N. I., Lebedev, E. B., Dorfman, A. M. and Bagdasarov, N. S. 1983. Study of process of melting of the Kirgurich basalt by the wave method. *Geochimica*, **9**, 1239–46.

Kirby, S. H., Durham, W. B. and Stern, L. 1991. Mantle phase changes and deep-earthquake faulting in subducting lithosphere. *Science*, **252**, 216–25.

Kirby, S. H., Stein, S., Okal, E. A. and Rubie, D. C. 1996. Metastable mantle phase transformations and deep earthquakes in subducting oceanic lithosphere. *Rev. Geophys.*, **34**, 261–306.

Klein, E. M. 2003. Spread thin in the Arctic. *Nature*, **423**, 932–3.

Klein, E. M., Langmuir, C. H., Zindler, A., Staudigel, H. and Hamelin, B. 1988. Isotope evidence of a mantle convection boundary at the Australian–Antarctic Discordance. *Nature*, **333**, 623–9.

Klein, F. W., Koyanagi, R. Y., Nakata, J. S. and Tanigawa, W. R. 1987. The seismicity of Kilauea's magma system. In R. W. Decker, T. L. Wright and P. H. Stauffer, eds., *Volcanism in Hawaii*. U.S. Geological Survey Professional Paper 1350. Denver, Colorado: U.S. Geological Survey, pp. 1019–186.

Koyanagi, R. Y., Chouet, B. and Aki, K. 1987. Origin of volcanic tremor in Hawaii, part I; data from the Hawaiian Volcano Observatory, 1969–1985. In R. W. Decker, T. L. Wright and P. H. Stauffer, eds., *Volcanism in Hawaii*, U.S. Geological Survey Professional Paper 1350. Denver, Colorado: U.S. Geological Survey, pp. 1221–58.

Kusznir, N. J. and Bott, M. H. P. 1976. A thermal study of the formation of oceanic crust. *Geophys. J. Roy. Astr. Soc.*, **47**, 83–95.

Lambert, I. B. and Wyllie, P. J. 1972. Melting of gabbro (quartz eclogite) with excess water to 35 kilobars, with geological applications. *J. Geol.*, **80**, 693–708.

Langmuir, C. H. 1987. A magma chamber observed? *Nature*, **326**, 15–16.

Langmuir, C. H., Klein, E. M. and Plank, T. 1992. Petrological systematics of mid-ocean ridge basalts: constraints on generation beneath mid-ocean ridges. In J. Phipps Morgan and J. M. Sinton, eds., Mantle Flow and Melt Generation Beneath Mid-ocean Ridges. Geophysical Monograph 71. Washington: Geophysical Union, pp. 183–280.

Larson, R. L. and Hide, T. W. C. 1975. A revised timescale of magnetic reversals for the Early Cretaceous and Late Jurassic. *J. Geophys. Res.*, **80**, 2586–94.

Laughton, A. S., Matthews, D. H. and Fisher, R. L. 1970. The structure of the Indian Ocean. In A. E. Maxwell, ed., *The Sea*, Vol. 4, Part II. New York: Wiley-Interscience, pp. 543–86.

Lay, T. 1994. Seismological constraints on the velocity structure and fate of subducting lithospheric slabs: 25 years of progress. In D. Dmowska and D. Salzman, eds., *Seismological Structure of Slabs*. Advances in Geophysics 36. San Diego, California: Academic Press, pp. 1–185.

Lewis, B. T. R. and Garmany, J. D. 1982. Constraints on the structure of the East Pacific Rise from seismic refraction data. *J. Geophys. Res.*, **87**, 8417–25.

Lewis, T. J., Bentkowski, W. H., Davies, E. E., Hyndman, R. D., Souther, J. G. and Wright, J. A. 1988. Subduction of the Juan de Fuca Plate: thermal consequences. *J. Geophys. Res.*, **93**, 15 207–25.

Lewis, T. J., Jessop, A. M. and Judge, A. S. 1985. Heat flux measurements in southwestern British Columbia: the thermal consequences of plate tectonics. *Canadian J. Earth Sci.*, **22**, 1262–73.

Lilwall, R. C., Francis, T. J. G. and Porter, I. T. 1977. Ocean-bottom seismograph observations on the Mid-Atlantic Ridge near 45°N. *Geophys. J. Roy. Astr. Soc.*, **51**, 357–69.

Lin, J. and Phipps Morgan, J. 1992. The spreading rate dependence of three-dimensional mid-ocean ridge gravity structure. *Geophys. Res. Lett.*, **19**, 13–16.

Lin, J., Purdy, G. M., Schouten, H., Sempère, J.-C. and Zervas, C. 1990. Evidence from gravity-data for focused magmatic accretion along the Mid-Atlantic Ridge. *Nature*, **344**, 627–32.

Lindwall, D. 1988. A two-dimensional seismic investigation of crustal structure under the Hawaiian Islands near Oahu and Kauai. *J. Geophys. Res.*, **93**, 12 107–22.

Louden, K. E., White, R. S., Potts, C. G. and Forsyth, D. W. 1986. Structure and seismotectonics of the Vema Fracture Zone, Atlantic Ocean. *J. Geol. Soc. Lond.*, **143**, 795–805.

Macdonald, K. C. 1982. Fine scale tectonic and hydrothermal processes within the plate boundary zone. *Ann. Rev. Earth Planet. Sci.*, **10**, 155–90.

1986. The crest of the Mid-Atlantic Ridge: models for crustal generation processes and tectonics. In P. R. Vogt and B. E. Tucholke, eds., *The Geology of North America, Vol. M, The Western North Atlantic Region*. Boulder, Colorado: Geological Society of America, pp. 51–68.

1989. Anatony of a magma reservoir. *Nature*, **339**, 178–9.

1998. Linkages between faulting, volcanism, hydrothermal activity and segmentation on fast spreading centers. In R. W. Buck, ed., *Faulting and Magmatism at Mid-ocean Ridges*. Geophysical Monograph 106. Washington: American Geophysical Union, pp. 27–58.

Macdonald, K. C., Castillo, D. A., Miller, S. P., Fox, P. J., Kastens, K. A. and Bonatti, E. 1986. Deep-tow studies of the Vema fracture zone: 1. Tectonics of a major slow slipping transform fault and its intersection with the Mid-Atlantic Ridge. *J. Geophys. Res.*, **91**, 3334–54.

Macdonald, K. C. and Fox, P. J. 1988. The axial summit graben and cross-sectional shape of the East Pacific Rise as indicators of axial magma chambers and recent volcanic eruptions. *Earth Planet. Sci. Lett.*, **88**, 119–31.

Macdonald, K. C., Fox, P. J., Perram, L. J., Eisen, M. F., Haymon, R. M., Miller, S. P., Carbotte, S. M., Cormier, M.-H. and Shor, A. N. 1988. A new view of the mid-ocean ridge from the behaviour of ridge-axis discontinuities. *Nature*, **335**, 217–25.

Macdonald, K. C., Scheirer, D. S. and Carbotte, S. M. 1991. Mid-ocean ridges: discontinuities, segments and giant cracks. *Science*, **253**, 986–94.

Macdonald, K. C., Sempère, J. C., Fox, P. J. and Tyce, R. 1987. Tectonic evolution of ridge axis discontinuities by the meeting, linking, or self-decapitation of neighbouring ridge segments. *Geology*, **15**, 993–7.

Mahlberg Kay, S. and Kay, R. W. 1985. Role of crystal cumulates and the oceanic crust in the formation of the lower crust of the Aleutian arc. *Geology*, **13**, 461–4.

Mahoney, J., LeRoux, A. P., Peng, Z., Fisher, R. L. and Natland, J. H. 1992. Southwestern limits of Indian Ocean ridge mantle and the origin of low $^{206}Pb/^{204}Pb$ mid-ocean ridge basalt: isotope systematics of the central Southwest Indian Ridge (17°–50°E). *J. Geophys. Res.*, **97**, 19 771–90.

Marone, F., van der Lee, S. and Gardini, D. 2004. Three-dimensional upper-mantle S-velocity model for the Eurasia–Africa plate boundary region, *Geophys. J. Int.*, **158**, 109–30.

Matsuzawa, T., Umino, N., Hasegawa, A. and Takagi, A. 1986. Upper mantle velocity structure estimated from PS-converted wave beneath the northeastern Japan Arc. *Geophys. J. Roy. Astr. Soc.*, **86**, 767–87.

Maxwell, A. E., Von Herzen, R. P., Hsü, K. J., Andrews, J. E., Saito, T., Percival Jr, S. F., Millow, E. D. and Boyce, R. E. 1970. Deep sea drilling in the South Atlantic. *Science*, **168**, 1047–59.

McCarthy, J., Mutter, J. C., Morton, J. L., Sleep, N. H. and Thompson, C. A. 1988. Relic magma chamber structures preserved within the Mesozoic North Atlantic crust? *Geol. Soc. Am. Bull.*, **100**, 1423–36.

McGuire, J. J., Wiens, D. A. and Shore, P. J. 1997. The March 9 (M_w 7.6), Deep Tonga earthquake: rupture outside the seismically active slab. *J. Geophys. Res.*, **102**, 15 163–82.

McKenzie, D. P. 1969. Speculations on the consequences and causes of plate motions. *Geophys. J. Roy. Astr. Soc.*, **18**, 1–32.

McKenzie, D. P. and Bickle, M. J. 1988. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.*, **29**, 625–79.

McKenzie, D. P. and Sclater, J. G. 1971. The evolution of the Indian Ocean since the Late Cretaceous. *Geophys. J. Roy. Astr. Soc.*, **25**, 437–528.

McNutt, M. K. 1998. Superswells. *Rev. Geophys.*, **36**, 211–44.

McNutt, R. H., Crocket, J. H., Clark, A. H., Caelles, J. C., Farrar, E., Haynes, S. J. and Zentilli, M. 1975. Initial $^{87}Sr/^{86}Sr$ ratios of plutonic and volcanic rocks of the central Andes between latitudes 26° and 29° South. *Earth Planet. Sci. Lett.*, **27**, 305–13.

Meade, C. and Jeanloz, R. 1991. Deep-focus earthquakes and recycling of water into the earth's mantle. *Science*, **252**, 68–72.

Menard, H. W. 1964. *Marine Geology of the Pacific*. New York: McGraw-Hill.

Mendel, V., Sauter, D., Patriat, Ph. and Munschy, M. 2000. Relationship of Central Indian Ridge segmentation with the evolution of the Rodrigues Triple Junction for the past 8 Myr. *J. Geophys. Res.*, **105**, 16 563–75.

Minshull, T. A., White, R. S., Mutter, J. C., Buhl, P., Detrick, R. S., Williams, C. A. and Morris, E. J. 1991. Crustal structure at the Blake Spur fracture zone from expanding spread profiles. *J. Geophys. Res.*, **96**, 9955–84.

Mitchell, N. C. 1989. Investigation of the structure and evolution of the Indian Ocean Triple Junction using GLORIA and other geophysical techniques, D. Phil. Thesis, Oxford University.

Mitchell, N. C. and Parson, L. M. 1993. The tectonic evolution of the Indian Ocean Triple Junction, anomaly 6 to present. *J. Geophys. Res.*, **98**, 1793–812.

Monger, J. W. H., Clowes, R. M., Price, R. A., Simony, P. S., Riddihough, R. P. and Woodsworth, G. J. 1985. *Continent–Ocean Transect B2: Juan de Fuca Plate to Alberta Plains*. Centennial Continent/Ocean Transect No. 7. Boulder, Colorado: Geological Society of America.

Morgan, J. P. and Forsyth, D. W. 1988. Three-dimensional flow and temperature perturbations due to a transform offset: effects on oceanic crustal and upper mantle structure. *J. Geophys. Res.*, **93**, 2955–66.

Morton, J. L., Sleep, N. H., Normark, W. R. and Tompkins, D. H. 1987. Structure of the southern Juan de Fuca Ridge from seismic reflection records. *J. Geophys. Res.*, **92**, 11 315–26.

Müller, C. and Jokat, W. 2000 Seismic evidence for volcanic activity at the Eastern Gakkel Ridge. *EOS Trans. Am. Geophys. Un.*, **81** (24), 265 and 269.

Muller, M. R., Minshull, T. A. and White, R. S. 1999. Segmentation and melt supply at the Southwest Indian Ridge. *Geology*, **27**, 867–70.

2000. Crustal structure of the Southwest Indian Ridge at the Atlantis II Fracture Zone. *J. Geophys. Res.*, **105**, 25 809–28.

Munschy, M. and Schlich, R. 1989. The Rodriguez Triple Junction (Indian Ocean): structure and evolution for the past one million years. *Mar. Geophys. Res.*, **11**, 1–14.

Nakajima, J., Matsuzawa, T., Hasegawa, A. and Zhao, D. 2001. Three-dimensional structure of V_p, V_s, and V_p/V_s beneath northeastern Japan: implications for arc magmatism and fluids, *J. Geophys. Res.*, **106**, 21 843–57.

NAT Study Group 1985. North Atlantic Transect: a wide-aperture, two-ship multichannel seismic investigation of the oceanic crust. *J. Geophys. Res.*, **90**, 10 321–41.

Navin, D. A., Peirce, C. and Sinha, M. C. 1998. The RAMESSES experiment – II. Evidence for accumulated melt beneath a slow spreading ridge from wide-angle refraction and multichannel reflection seismic profiles. *Geophys. J. Int.*, **135**, 746–72.

Nedimovic, M. R., Hyndman, R. D., Ramachandran, K. and Spence, G. D. 2003. Reflection signature of seismic and aseismic slip on the northern Cascadia subduction interface. *Nature*, **424**, 416–20.

Ness, G., Levi, S. and Couch, R. 1980. Marine magnetic anomaly timescales for the Cenozoic and Late Cretaceous: a precis, critique and synthesis. *Rev. Geophys. Space Phys.*, **18**, 753–70.

Nisbet, E. G. and Fowler, C. M. R. 1978. The Mid-Atlantic Ridge at 37 and 45°N: some geophysical and petrological constraints. *Geophys. J. Roy. Astr. Soc.*, **54**, 631–60.

Nockolds, S. R., Knox, R. W. O'B. and Chinner, G. A. 1978. *Petrology for Students*. Cambridge: Cambridge University Press.

Orcutt, J. A., Kennett, B. L. N. and Dorman, L. M. 1976. Structure of the East Pacific Rise from an ocean bottom seismometer survey. *Geophys. J. Roy. Astr. Soc.*, **45**, 305–20.

Pallister, J. S. and Hopson, C. A. 1981. Semail ophiolite plutonic suite: field relations, phase variation, cryptic variation and layering and a model of a spreading ridge magma chamber. *J. Geophys. Res.*, **86**, 2593–644.

Patriat, Ph. and Segoufin, J. 1988. Reconstruction of the central Indian Ocean. *Tectonophysics*, **155**, 211–34.

Perfit, M. R. and Chadwick Jr, W. W. 1998. Magmatism at mid-ocean ridges: constraints from volcanological and geochemical investigations. In W. R. Buck, P. T. Delaney, J. A. Karson and Y. Lagabrielle, eds., *Faulting and Magmatism at Mid-Ocean Ridges*. Geophysical Monograph 106. Washington: American Geophysical Union, pp. 59–115.

Perfit, M. R., Fornari, D. J., Smith, M. C., Langmuir, C. H., Bender, J. F. and Haymon, R. M. 1994. Small-scale spatial and temporal variations in mid-ocean ridge crest magmatic processes. *Geology*, **22**, 375–9.

Phipps Morgan, J. E., Parmentier, M. and Lin, J. 1987. Mechanisms for the origin of mid-ocean ridge axial topography: implications for the thermal and mechanical structure of accreting plate boundaries. *J. Geophys. Res.*, **92**, 12 823–36.

Pitman III, W. C. and Heirtzler, J. R. 1966. Magnetic anomalies over the Pacific–Antarctic Ridge. *Science*, **154**, 1164–71.

Pratson, L. F. and Haxby, W. F. 1997. Panoramas of the seafloor. *Sci. Am.* (June), 67–71.

Prothero, W. A., Reid, I., Reichle, M. S. and Brune, J. N. 1976. Ocean-bottom seismic measurements on the East Pacific Rise and Rivera Fracture Zone. *Nature*, **262**, 121–4.

Purdy, G. M. 1983. The seismic structure of 140 Myr old crust in the western central Atlantic Ocean. *Geophys. J. Roy. Astr. Soc.*, **72**, 115–37.

Purdy, G. M. and Detrick, R. S. 1986. The crustal structure of the Mid-Atlantic Ridge at 23°N from seismic refraction studies. *J. Geophys. Res.*, **91**, 3739–62.

Purdy, G. M. and Ewing, J. 1986. Seismic structure of the ocean crust. In P. R. Vogt and B. E. Tucholke, eds., *The Geology of North America, Vol. M, The Western North Atlantic Region*. Boulder, Colorado: Geological Society of America, pp. 313–30.

Raitt, R. W. 1963. The crustal rocks. In M. N. Hill, ed., *The sea*, Vol. 3. New York: Interscience, pp. 85–102.

Reid, I., Orcutt, J. A. and Prothero, W. A. 1977. Seismic evidence for a narrow zone of partial melting underlying the East Pacific Rise at 21°N. *Geol. Soc. Am. Bull.*, **88**, 678–82.

Richter, F. M. 1979. Focal mechanisms and seismic energy release of deep and intermediate earthquakes in the Tonga–Kermadec region and their bearing on the depth extent of mantle now. *J. Geophys. Res.*, **84**, 6783–95.

Rihm, R. and Langseth, M. 1997. Arctic ridges: results and planning. In *InterRidge Global Studies Working Group Workshop Report*. Paris: InterRidge.

Rise Project Group 1980. East Pacific Rise: hot springs and geophysical experiments. *Science*, **207**, 1421–33.

Rona, P. A. 1988. Hydrothermal mineralisation at oceanic ridges. *Canadian Mineralogist*, **26**, 431–65.

Rondenay, S., Bostock, M. and Shragge, J. 2001. Multiparameter two-dimensional inversion of scattered teleseismic body waves, 3. Application to the Cascadia 1993 data set. *J. Geophys. Res.*, **106**, 30 795–807.

Roth, E. G., Wiens, D. A., Dorman, I. M., Hildebrand, J. and Webb, S. C. 1999. Seismic attenuation of the Tonga–Fiji region using phase pair methods. *J. Geophys. Res.*, **104**, 4795–809.

Rowlett, H. and Forsyth, D. W. 1984. Recent faulting and microearthquakes at the intersection of the Vema fracture zone and the Mid-Atlantic Ridge. *J. Geophys. Res.*, **89**, 6079–94.

Ryan, M. P. 1988. A close look at parting plates. *Nature*, **332**, 779–80.

1988. The mechanics and three-dimensional internal structure of active magmatic systems: Kiluaea Volcano, Hawaii. *J. Geophys. Res.*, **93**, 4213–48.

Ryan, M. P., Koyanagi, R. Y. and Fiske, R. S. 1981. Modelling the three-dimensional structure of macroscopic magma transport systems: application to Kilauea Volcano, Hawaii. *J. Geophys. Res.*, **86**, 7111–29.

Sacks, I. S. and Okada, H. 1974. A comparison of the anelasticity structure beneath western South America and its tectonic significance. *Phys. Earth Planet. Interiors*, **9**, 211–19.

Schubert, G., Yuen, D. A. and Turcotte, D. L. 1975. Role of phase transitions in a dynamic mantle. *Geophys. J. Roy. Astr. Soc.*, **42**, 705–35.

Schulz, A. and Elderfield, H. 1997. Controls on the physics and chemistry of seafloor hydrothermal circulation. *Phil. Trans. Roy. Soc. Lond.* A, **355**, 387–425.

Schweller, W. J., Kuln, L. D. and Prince, R. A. 1981. Tectonics, structure and sedimentary framework of the Peru–Chile Trench. In L. D. Kuln, J. Dymond, E. J. Dasch and D. M. Hussong, eds., *Nazca Plate, Crustal Formation and Andean Convergence*. Vol. 154 of Memoirs of the Geological Society of America. Boulder, Colorado: Geological Society of America, pp. 323–50.

Searle, R. C. 1983. Multiple, closely spaced transform faults in fast-slipping fracture zones. *Geology*, **11**, 607–10.

Searle, R. C., Keeton, J. A., Owens, R. B., White, R. S., Mecklenburgh R., Parsons, B. and Lee, S. M. 1998. The Reykjanes Ridge: structure and tectonics of a hot-spot influenced,

slow-spreading ridge, from multibeam bathymetry, gravity and magnetic investigations. *Earth Planet. Sci. Lett.*, **160**, 463–78.

Shaw, P. 1992. Ridge segmentation, faulting and crustal thickness in the Atlantic Ocean. *Nature*, **358**, 490–3.

Shen, Y. and Forsyth, D. W. 1995. Geochemical constraints on initial and final depths of melting beneath mid-ocean ridges. *J. Geophys. Res.*, **100**, 2211–37.

Sinha, M. C. and Louden, K. E. 1983. The Oceanographer fracture zone – I. Crustal structure from seismic refraction studies. *Geophys. J. Roy. Astr. Soc.*, **75**, 713–36.

Sinton, J. M. and Detrick, R. S. 1992. Mid-ocean ridge magma chambers. *J. Geophys. Res.*, **97**, 197–216.

Sleep, N. H. 1975. Formation of oceanic crust: some thermal constraints. *J. Geophys. Res.*, **80**, 4037–42.

1983. Hydrothermal convection at ridge axes. In P. A. Rona, K. Bostrom, L. Laubier and K. L. Smith, eds., *Hydrothermal Processes at Seafloor Spreading Centres*. New York: Plenum, pp. 71–82.

Smith, D. K. and Cann, J. R. 1993, Building the crust at the Mid-Atlantic Rodge. *Nature*, **365**, 707–14.

Smith, H. S. and Erlank, A. J. 1982. Geochemistry and petrogenesis of komatiites from the Barberton greenstone belt, South Africa. In N. T. Arndt and E. G. Nisbet, eds., *Komatiites*. London: George Allen and Unwin, pp. 347–97.

Solomon, S. C. and Julian, B. R. 1974. Seismic constraints on oceanic-ridge mantle structure: anomalous fault plane solutions from first motions. *Geophys. J. Roy. Astr. Soc.*, **38**, 265–85.

Solomon, S. C. and Toomey, D. R. 1992. The structure of mid-ocean ridges. *Ann. Rev. Earth Planet. Sci.*, **20**, 329–64.

Spence, G. D., Clowes, R. M. and Ellis, R. M. 1985. Seismic structure across the active subduction zone of western Canada. *J. Geophys. Res.*, **90**, 6754–72.

Spiegelman, M. and Reynolds, J. R. 1999. Combined dynamic and geochemical evidence for convergent melt flow beneath the East Pacific Rise. *Nature*, **402**, 282–5.

Spudich, P. and Orcutt, J. 1980. Petrology and porosity of an oceanic crustal site: results from waveform modelling of seismic refraction data. *J. Geophys. Res.*, **85**, 1409–33.

Stauder, W. and Maulchin, L. 1976. Fault motion in the larger earthquakes of Kuie–Kamchatka Arc and of the Kurile, Hokkaido Corner. *J. Geophys. Res.*, **81**, 297–308.

Stein, S. and Stein, C. A. 1996. Thermo-mechanical evolution of oceanic lithosphere: implications for the subduction process and deep earthquakes. In G. E. Bebout, D. W. Scholl, S. H. Kirby and J. P. Platt, eds., *Subduction Top to Bottom*. Geophysical Monograph 96. Washington: American Geophysical Union, pp. 1–17.

Stoffa, P. L., Buhl, P., Henon, T. J., Kan, T. K. and Ludwig, W. J. 1980. Mantle reflections beneath the crustal zone of the East Pacific Rise from multichannel seismic data. *Mar. Geol.*, **35**, 83–97.

Suyehiro, K. and Sacks, I. S. 1983. An anomalous low velocity region above the deep earthquakes in the Japan subduction zone. *J. Geophys. Res.*, **88**, 10 429–38.

Swift, S. A., Lizarralde, D., Stephen, R. A. and Hoskins, H. 1998. Velocity structure in upper ocean crust at Hole 504B from vertical seismic profiles. *J. Geophys. Res.*, **103**, 15 361–76.

Talwani, M., LePichon, X. and Ewing, M. 1965. Crustal structure of the mid-ocean ridges 2: computed model from gravity and seismic refraction data. *J. Geophys. Res.*, **70**, 341–52.

Toksöz, M. N. 1975. The subduction of the lithosphere. In J. T. Wilson, ed., *Continents Adrift and Continents Aground*. New York: Scientific American Inc., pp. 113–22.

Toomey, D. R., Solomon, S. C. and Purdy, G. M. 1994. Tomographic imaging of the shallow crustal structure of the East Pacific Rise at 9° 30′N. *J. Geophys. Res.*, **99**, 24 135–57.

Toomey, D. R., Solomon, S. C., Purdy, G. M. and Murray, M. H. 1985. Microearthquakes beneath the median valley of the Mid-Atlantic Ridge near 23°N: hypocentres and focal mechanisms. *J. Geophys. Res.*, **90**, 5443–58.

Tsumura, N., Matsumoto, S., Horiuchi, S. and Hasegawa, A. 2000. Three-dimensional attenuation structure beneath the northeastern Japan arc estimated from the spectra of small earthquakes. *Tectonophysics*, **319**, 241–60.

Tucholke, B. E. 1986. Structure of the basement and distribution of sediments in the western North Atlantic. In P. R. Vogt and B. E. Tucholke, eds., *The Geology of North America*, Vol. M. Boulder, Colorado: Geological Society of America, pp. 331–40.

Tunnicliffe, V. and Fowler, C. M. R. 1996. Influence of seafloor spreading on the global hydrothermal vent fauna. *Nature*, **379**, 531–3.

Tunnicliffe, V., Fowler, C. M. R. and McArthur, A. G. 1996. Plate tectonic history and hot vent biogeography. In C. J. MacLeod, P. A. Tyler and C. L. Walker, eds., *Tectonic, Magmatic*, *Hydrothermal and Biological Segmentation of Mid-ocean Ridges*. Geological Society of London. Special Publication 118. London: Geological Society of London, pp. 225–38.

Tunnicliffe, V., McArthur, A. G. and McHugh, D. 1998. A biogeographical perspective of the deep-sea hydrothermal vent fauna. *Adv. Mar. Biol.*, **34**, 353–442.

Turcotte, D. L. and Schubert, G. 1973. Frictional heating of the descending lithosphere. *J. Geophys. Res.*, **78**, 5876–86.

Turner, I. M., Peirce, C. and Sinha, M. C. 1999. Seismic imaging of the axial region of the Valu Fa Ridge, Lau Basin – the accretionary processes of an intermediate back-arc spreading ridge. *Geophys. J. Int.*, **138**, 495–519.

Turner, S. P., Hawkesworth, C. J., Rogers, N., Bartlett, J., Worthington, T., Hergt, J., Pearce, J. A. and Smith, I. 1997. ^{238}U–^{230}Th disequilibria, magma petrogenesis and flux rates beneath the depleted Tonga–Kermadec island arc. *Geochim. Cosmochim. Acta*, **61**, 4855–84.

Van Avendonk, H. J. A., Harding, A. J. and Orcutt, J. A. 2001. Contrast in crustal structure across the Clipperton transform fault from travel time tomography. *J. Geophys. Res.*, **106**, 10 961–81.

van der Hilst, R., Engdahl, E. R. and Spakman, W. 1993. Tomographic inversion of P and pP data for aspherical mantle structure below the northwest Pacific region. *Geophys. J. Int.*, **115**, 264–302.

Vassiliou, M. S. 1984. The state of stress in subducting slabs as revealed by earthquakes analysed by moment tensor inversion. *Earth Planet. Sci. Lett.*, **69**, 195–202.

Vassiliou, M. S., Hager, B. H. and Raefsky, A. 1984. The distribution of earthquakes with depth and stress in subducting slabs. *J. Geodyn.*, **1**, 11–28.

Vine, F. J. 1966. Spreading of the ocean floor: new evidence. *Science*, **154**, 1405–15.

Vine, F. J. and Matthews, D. H. 1963. Magnetic anomalies over oceanic ridges. *Nature*, **199**, 947–9.

Wadati, K. 1928. Shallow and deep earthquakes. *Geophys. Mag.*, **1**, 162–202.
1935. On the activity of deep-focus earthquakes in the Japan Islands and neighbourhoods. *Geophys. Mag.*, **8**, 305–25.
Wang, K., Wells, R., Mazzotti, S., Hyndman, R. D. and Sagiya, T. 2003. A revised dislocation model of interseismic deformation on the Cascadia subduction zone. *J. Geophys. Res.*, **108** (B1), 2009, doi: 10.1029/200/JB001227.
Wang, X. and Cochran, J. R. 1995. Along-axis gravity gradients at mid-ocean ridges: implications for mantle flow and axial morphology. *Geology*, **23**, 29–32.
Watanabe, T., Langseth, M. G. and Anderson, R. N. 1977. Heat flow in back-arc basins of the Western Pacific. In M. Talwani and W. C. Pitman III, eds., *Island Arcs, Deep-sea Trenches and Back-arc Basins*. American Geophysical Union Maurice Ewing Series 1. Washington: American Geophysical Union, pp. 137–62.
Watts, A. B., ten Brink, U. S., Buhl, P. and Brocher, T. M. 1985. A multi-channel seismic study of lithospheric flexure across the Hawaii–Emperor seamount chain. *Nature*, **315**, 105–11.
Watts, A. B., Weissel, J. K. and Larson, R. L. 1977. Sea-floor spreading in marginal basins of the western Pacific. *Tectonophysics*, **37**, 167–81.
Watts, A. B. and Zhong, S. 2000. Observations of flexure and the rheology of oceanic lithosphere. *Geophys. J. Int.*, **142**, 855–75.
Wernicke, B. 1992. Cenozoic extensional tectonics of the U.S. Cordillera. In B. C. Burchfiel *et al*., eds., *The Geology of North America*, *vol. G3*, *The Cordilleran Orogen: Coterminous U.S.* Boulder, Colorado: Geological Society of America, pp. 553–81.
1995. Low-angle normal faults and seismicity: a review. *J. Geophys. Res.*, **100**, 20 159–74.
White, R. S. 1984. Atlantic oceanic crust: seismic structure of a slow spreading ridge. In I. G. Gass, S. J. Lippard and A. W. Shelton, eds., *Ophiolites and Oceanic Lithosphere*. London: Geological Society of London, pp. 34–44.
1993. Melt production rates in mantle plumes. *Phil. Trans. Roy. Soc. Lond.* A, **342**, 137–53.
White, R. S., Detrick, R. S., Mutler, J. C., Buhl, P., Minshull, T. A. and Morris, E. 1990. New seismic images of oceanic crustal structure. *Geology*, **18**, 462–5.
White, R. S., Detrick, R. S., Sinha, M. C. and Cormier, M. H. 1984. Anomalous seismic crustal structure of oceanic fracture zones. *Geophys. J. Roy. Astr. Soc.*, **79**, 779–98.
White, R. S. and Louden, K. E. 1982. The Makran continental margin: structure of a thickly sedimented convergent plate boundary. *Am. Assoc. Petrol. Geol. Mem.*, **34**, 499–518.
White, R. S., McKenzie, D. P. and O'Nions, R. K. 1992. Oceanic crustal thickness from seismic measurements and from rare earth element inversions. *J. Geophys. Res.*, **97**, 19 683–715.
White, R. S., Minshull, T. A., Bickle, M. J. and Robinson, C. J. 2001. Melt generation at very slow-spreading oceanic ridges: constraints from geochemical and geophysical data. *J. Petrol*., **42**, 1171–96.
Whitmarsh, R. B. and Calvert, A. J. 1986. Crustal structure of Atlantic fracture zones – I. The Charlie-Gibbs Fracture Zone. *J. Roy. Astr. Soc.*, **85**, 107–38.
Wiens, D. A. and Gilbert, H. J. 1996. Effect of slab temperature on deep earthquake aftershock productivity and magnitude-frequency relations, *Nature*, **384**, 153–6.
Wiens, D. A., McGuire, J. J., Shore, P. J., Bevis, M. G., Draunidalo, K., Prasad, G. and Helu, S. P. 1994. A deep earthquake aftershock sequence and implications for the rupture mechanism of deep earthquakes, *Nature*, **372**, 540–3.

Wiens, D. A. and Stein, S. 1983. Age dependence of oceanic intraplate seismicity and implications for lithospheric evolution. *J. Geophys. Res.*, **88**, 6455–68.

Wyllie, P. J. 1976. *The Way the Earth Works*. New York: Wiley.

1981. Experimental petrology of subduction, andesites and basalts. *Trans. Geol. Soc. S. Afr.*, **84**, 281–91.

Wyss, M., Hasegawa, A. and Nakajima, J. 2001. Source and path of magma for volcanoes in the subduction zones of northeastern Japan. *Geophys. Res. Lett.*, **28**, 1819–22.

Yoshii, T. 1972. Features of the upper mantle around Japan as inferred from gravity anomalies. *J. Phys. Earth*, **20**, 23–34.

1977. Crust and upper-mantle structure beneath northeastern Japan. *Kagaku*, **47**, 170–6 (in Japanese).

1979. A detailed cross-section of the deep seismic zone beneath north-eastern Honshu, Japan. *Tectonophysics*, **55**, 349–60.

You, C.-F. and Bickle, M. J. 1998. Evolution of an active sea-floor massive sulphide deposit. *Nature*, **394**, 668–71.

Zhao, D., Xu, Y., Wiens, D. A., Hildebrand, J. and Webb, S. 1997. Depth extent of the Lau back-arc spreading centre and its relation to subduction processes. *Science*, **278**, 254–7.

Zobin, V. M. 1997. The rupture history of the M_w 8.0 Jalisco, Mexico, earthquake of 1995 October 9. *Geophys. J. Int.*, **130**, 220–8.

Chapter 10
The continental lithosphere

10.1 Introduction

10.1.1 Complex continents

We have seen something of the general simplicity of the Earth's internal structure and the detailed complexity of the motions of tectonic plates and convective systems. The clues to this simplicity and complexity come from the oceans, the study of whose structures has led to an understanding of the plates, of the mantle beneath and, to some extent, of the core, via its magnetic properties.

Although complex details must be sorted out and theories may change slightly, we can now be reasonably confident that the oceans are understood in their broad structure. In contrast, the continents are not understood at all well. Yet we need to understand the continents because in their geological record lies most of the history of the Earth and its tectonic plates, from the time that continental material first formed over 4400 Ma ago (see Section 6.10). The oldest oceanic crust is only about 160 Ma old, so the oceanic regions can yield no earlier information.

In the broadest terms, the continents are built around ancient crystalline crust, flanked by younger material representing many events of mountain building, collision, rifting and plate convergence and subsidence. Figure 3.30 shows the recent motions of the plates, illustrating how continents have collided and been torn asunder.

A major problem in the geological and geophysical study of continents is that we can observe only what is exposed at or near the surface. To extend that knowledge to tens of kilometres deep, let alone to hundreds of kilometres, demands conjecture that cannot be tested directly. Oil and mineral exploration companies have developed sophisticated techniques for surveying the upper few kilometres of the crust in search of deposits and have significantly advanced our knowledge of sedimentation, oil maturation and ore genesis. The proof of the pudding is in the eating – oil companies are accountable, and they finally have to drill to verify their interpretations and conclusions. If they are wrong too often, they become bankrupt. Their methods must be good since the results are tested.

In our study of the Earth we are at present unable to sample directly the deep interior and so are at the disadvantage of being unable to test our models

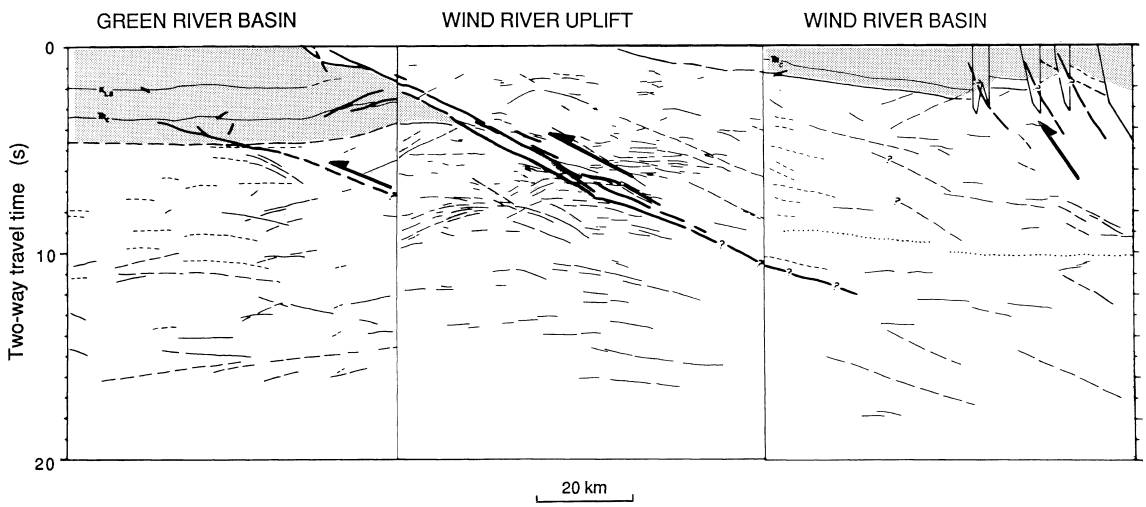


Figure 10.1. A deep reflection line shot by COCORP across the Wind River Mountains in Wyoming, U.S.A. Heavy lines, thrusts and faults. Shading, sediments. Short dashed lines, possible multiples. Dotted line at 10 s, an enigmatic low-frequency event. Event at ~15 s beneath Green River Basin, a possible Moho reflector. Dashed line at 4.0–4.5 s, the reflection from the base of the Green River Basin sediments. This is cut off by the Wind River Thrust. (After Brewer *et al.* (1980).)

straightforwardly; however, we do have the questionable advantage that no one can prove us wrong. Academics need not go bankrupt through their misconceptions. The deepest hole drilled into the crust is in the Kola peninsula of northwestern Russia. Drilling there began in the early 1970s and has penetrated, so far, to some 13 km. The second deepest hole, 9.1 km, has been drilled in Germany, the KTB project. Technically, drilling such holes is an exceedingly difficult enterprise. One of the problems with deep seismic-reflection profiling (see Section 4.5.5) is that differing interpretations of the various reflectors can sometimes be made. Nevertheless, deep seismic profiling is very successful and has given an immediate solution to some geological puzzles; for instance, it enabled COCORP to trace the Wind River Thrust as a 30–35° dipping reflector in Wyoming (U.S.A.), from its surface exposure to some 25 km deep (Fig. 10.1), and to learn without doubt that compressional rather than vertical forces were the cause of the uplifted basement blocks.

Geochemistry is probably less hindered than structural geology by our inability to obtain deep, fresh samples – although, of course, the exact composition of the lower crust and the nature of the Moho and upper mantle are matters of current debate, which fresh samples could resolve.

In this chapter some of the major geophysical and geological features of the continents are described and discussed in terms of their relation to the internal processes of the Earth.

10.1.2 Geophysical characteristics of continents

The continental crust averages 38 km thick. We have already used this value in calculations of isostasy in Section 5.5 and in the calculation of the thickness of the oceanic crust (Section 9.2.1). Although a 'normal' or 'standard' oceanic crustal structure can be defined, it is much more difficult to give a standard continental crustal structure. The variability of the structure of the continental crust is, like all the other properties of the continents, a direct result of the diverse processes involved in their formation and the long time over which they have formed. Figure 10.2(a) shows a global crustal-thickness map. The large variations in crustal thickness are very clear. The thickest crust is found beneath the Tibetan plateau, the Andes and Finland. This global map has a $5^\circ \times 5^\circ$ cell size, which means that features narrower than 200 km, such as mountain belts, are not accurately depicted. Generally, the crust is thick beneath young mountain ranges, moderately thick beneath the ancient shield regions and thin beneath young basins and rifts such as the North Sea and Rhine Graben in Europe, the East African Rift and the Basin and Range Province in the U.S.A. as well as beneath the continental margins, passive margins and active fore-arcs. The global average thickness of continental crust is 38 km, but the thickness typically ranges between 30 and 45 km.

The seismic-velocity structure of the crust is determined from long seismic-refraction lines. The advent of deep reflection lines has delineated the fine structure of the crust very well, but such data usually cannot yield accurate velocity estimates (see the discussion of stacking velocities in Section 4.5.3). The offset between the source and the further receiver must be increased considerably to obtain better deep velocity information from reflection profiling.

Teleseismic earthquake recordings can be used to confirm gross crustal and upper mantle interfaces through the use of P to S mode conversions. The technique is referred to as the 'receiver function' method since interfaces are identified for each 'receiver' or seismograph location. When there are several receivers in a study area it is possible to establish a gross crustal, or lithosphere, thickness map.

The direct wave which travels in the crystalline, continental basement, beneath surface soil and sedimentary cover, termed Pg, normally travels with a velocity of about 5.9–6.2 km s^{-1}. The velocity of the upper 10 km of the crust is usually in the range 6.0–6.3 km s^{-1}; beneath that, in the middle crust, the velocity generally exceeds 6.5 km s^{-1}. At some locations there is another, lower crustal layer with velocity greater than 7 km s^{-1}. For many years, a major discontinuity at the base of the upper crust (the *Conrad discontinuity*) was thought to be a universal feature of continental crust and to be underlain by a basaltic layer having a velocity of 6.5 km s^{-1}. This is no longer thought to be the case. Continental crustal structure is complex and, while some regions have a well-developed discontinuity at the base of the upper crust, not all do. Low-velocity zones at various locations at all depths in the continental crust have been described and represent the complexity of the history of the crust. The continental crust does not have a standard structure: Fig. 10.2(b) shows the general variation. In the middle crust velocities

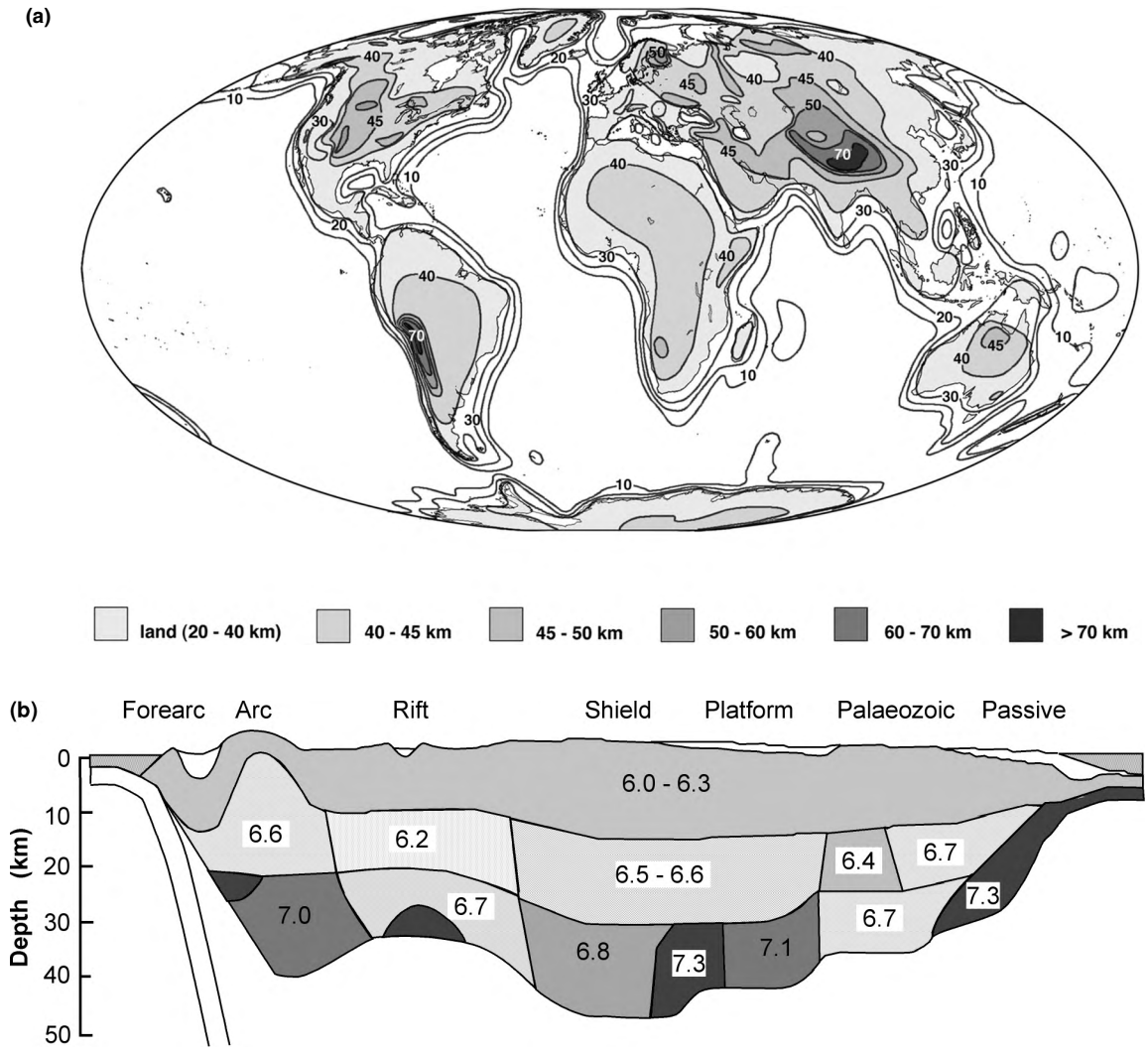


Figure 10.2. (a) A global crustal-thickness map (Mercator projection). The map is contoured on the basis of a $5^\circ \times 5^\circ$ cell size, which means that narrow features are not resolved. The cell size at the equator is 550 km × 550 km. Colour version Plate 27. (From Mooney *et al.*, CRUST 5.1: a global crustal model at $5^\circ \times 5^\circ$, *J. Geophys. Res.*, **103**, 727–47, 1998. Copyright 1998 American Geophysical Union. Reprinted by permission of American Geophysical Union.) (b) A cross section across an idealized continent, showing the average seismic P-wave velocities of the crust in the various tectonic regions. (After Holbrook *et al.* (1992).)

vary between 6.0 and 7.1 km s^{-1}, with 6.4–6.8 km s^{-1} being the 'normal' range. In the lower crust velocities are found to vary between 6.4 and 7.5 km s^{-1}, with most measurements being within the ranges 6.6–6.8 and 7.0–7.2 km s^{-1}. This variability of seismic velocity in the continental crust reflects the bulk composition of the crust as well as its thermal state and metamorphic history: the

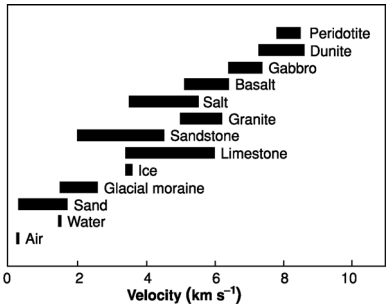


Figure 10.3. Ranges of laboratory measurements of the P-wave velocity in various rock types. (From data in Press (1966).)

low velocities in the middle crust beneath rift zones are probably due to elevated temperatures. Velocities increase with composition on going from felsic to mafic to ultramafic; elevated temperatures cause a reduction in velocity and high-grade metamorphism increases velocities. Thus velocities of 6.6–6.8 km s^{-1} in the lower crust may be typical of regions in which arc magmatism has been the dominant mechanism of continental growth, whereas velocities of 7.0–7.2 km s^{-1} in the lower crust may be typical of regions in which mafic/ultramafic magmatic underplating and/or high-grade metamorphism has dominated.

Figure 10.3 shows the ranges of laboratory measurements of the P-wave velocity for various rock types. For example, not every basalt has a velocity of 6.0 km s^{-1}, but velocities in the range 5.1–6.4 km s^{-1} are reasonable for basalts. Laboratory measurements show that the P-wave velocity increases with pressure. However, this does not necessarily mean that the P-wave velocity for a given rock unit will increase with depth in the crust. The increase of temperature with depth can either counteract or enhance the effect of increasing pressure, depending on the physical properties (e.g., pores and fissures) of the particular rock unit.

The Moho is in some places observed to be a velocity gradient, in other places it is a sharp boundary and in still others it is a thin laminated zone. The thickness of this transition from crust to mantle can be estimated from the wavelength of the seismic signals. Two kilometres is probably a maximum estimate of the thickness of the transition. Geologically, however, the Moho represents the boundary between the lower-crustal granulites and the ultrabasic upper mantle, which is predominantly olivine and pyroxene.

The gross structure of the continental crust as determined from surface waves was discussed in Section 4.1.3 and illustrated in Fig. 4.6; we found the continental lithosphere to be thicker than the oceanic lithosphere. The thermal structure of the continents was discussed in Section 7.6, and we concluded that the oceanic geotherms for lithosphere older than 70 Ma can equally well be applied to the continental lithosphere (see Figs. 7.14 and 7.15).

10.1.3 The composition of the continental crust

The continental crust has been formed from mantle material over the lifetime of the Earth by a series of melting, crystallization, metamorphic, erosional,

Table 10.1 *Estimated compositions of the bulk continental crust and of oceanic crust*

Compound	Continental (%)	Oceanic (%)
SiO_2	57.3	49.5
TiO_2	0.9	1.5
Al_2O_3	15.9	16.0
FeO	9.1	10.5
MgO	5.3	7.7
CaO	7.4	11.3
Na_2O	3.1	2.8
K_2O	1.1	0.15

Source: Taylor and McLennan (1985).

depositional, subduction and endless reworking events. We saw in Chapter 6 how radiometric isotope methods unravel some of the complexity for us by dating samples, indicating whether they came from a crustal or mantle source, whether they were contaminated and where contamination may have occurred. The other tools used to decipher the history of rocks are, of course, all the methods of geology and geophysics.

The continental crust, despite its complexity and variation, has a fairly standard 'average' composition (Table 10.1). This composition is more silica-rich than that of oceanic basalts. In general terms, the composition of the continental upper crust is similar to that of granodiorite, and the lower crust is probably granulite. However, this is a gross oversimplification. The crust is far from being homogeneous and still retains the marks of its origins. Thus, sedimentary material buried during a thrusting event can be found deep in the crust, and oceanic-type rocks or even ultramafic rocks have been thrust up to the surface during mountain building.

Knowledge of the variation of the strength of the crust is based primarily on laboratory measurements of rock samples. Figure 10.4 shows strength envelopes[1] for continental and oceanic crust under compression. In both cases, the stress at which failure occurs increases linearly with depth. The sharp reductions that occur at ~15 and 35 km depth for continental lithosphere and ~35 km depth for oceanic lithosphere are due to the fact that the rocks deform by solid-state creep at these depths and temperatures (rocks that are brittle at low temperatures become ductile at higher temperatures). Earthquakes will tend to nucleate around the brittle–ductile transition. In the continental lithosphere, the upper crust is strong while the lower crust is weak and will deform viscously or viscoelastically. However, at the Moho, with the change in composition, there is a region in which the strength is increased. This maximum depth is dependent upon the strain

[1] See Section 6.6 of Watts (2001) for a full discussion of yield strength.

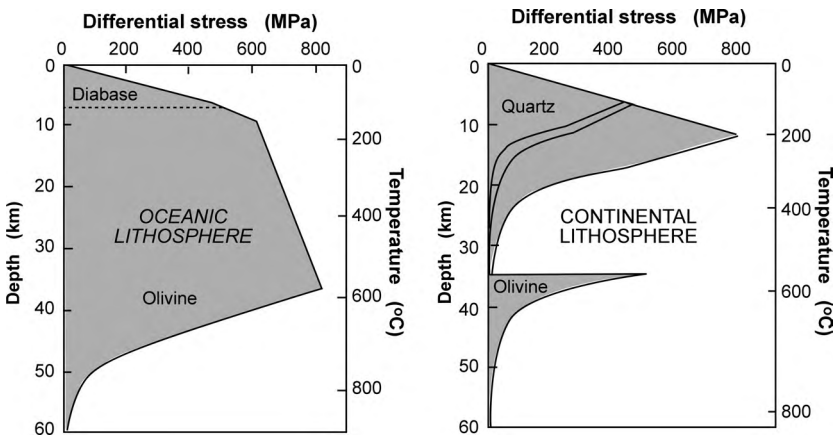


Figure 10.4. Yield-strength envelopes for the oceanic and continental lithosphere in compression at strain rate $10^{-15}\,s^{-1}$. Since rocks are considerably stronger in compression than they are in tension, the yield-strength envelopes for oceanic and continental lithosphere in tension have a broadly similar shape to those shown here but with the differential stress reduced by a factor of two to three. (From Kohlstedt *et al.*, Strength of the lithosphere: constraints imposed by laboratory measurements, *J. Geophys. Res.*, **100**, 17 587–602, 1995. Copyright 1995 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

rate (higher strain rates increase the depth). In contrast to the continental lithosphere, the yield-strength envelope for the oceanic crust and uppermost mantle is simple – strength increases linearly with depth down to ~35 km and below this the lithosphere deforms by solid-state creep. (The yield-strength envelope shown here is for 60-Ma-old oceanic lithosphere; young oceanic lithosphere is much weaker – the depth of the maximum in strength is approximately proportional to the square root of the plate age.) The maximum stress that can be transmitted by the lithosphere can be estimated by calculating the area under the yield-strength curves. For the oceanic lithosphere this is about $(2–3) \times 10^{13}\,N\,m^{-1}$ when it is under compression, but about $8 \times 10^{12}\,N\,m^{-1}$ when it is under tension. For the continental lithosphere the values are $\sim(0.5–2) \times 10^{13}\,N\,m^{-1}$ and $\sim(1–3) \times 10^{12}\,N\,m^{-1}$, respectively, depending upon assumptions of composition and age. These values are considerably in excess of the values for the plate-driving and resistive forces (Section 8.2.4), confirming that the plates are indeed strong enough to transmit such forces without fracturing.

Table 10.2 gives some idea of the worldwide extent of continental crust of various ages. Only 30% of current basement rocks are younger than 450 Ma; the remaining 70% are older. Continental growth rates are discussed in Section 10.2.4.

It is immediately apparent from a map of the ages of the continents that the oldest material tends to concentrate towards the centre of a continent with younger material around it. These old continent interiors are termed *cratons* (Greek *cratos*, meaning strength, power or dominion). On the North American continent these cratons are the stable, flat interior regions (Fig. 10.5). To the east of the Archaean

Table 10.2 *The area of continental basement*

Age (Ma)	Area (10^6 km^2)	Percentage of total area
0–450	38.2	29.5
450–900	41.1	31.8
900–1350	14.6	11.3
1350–1800	8.7	6.7
1800–2250	19.4	15.0
2250–2700	6.2	4.8
2700–3150	1.1	0.9
Total	129.3	

Source: Hurley and Rand (1969).

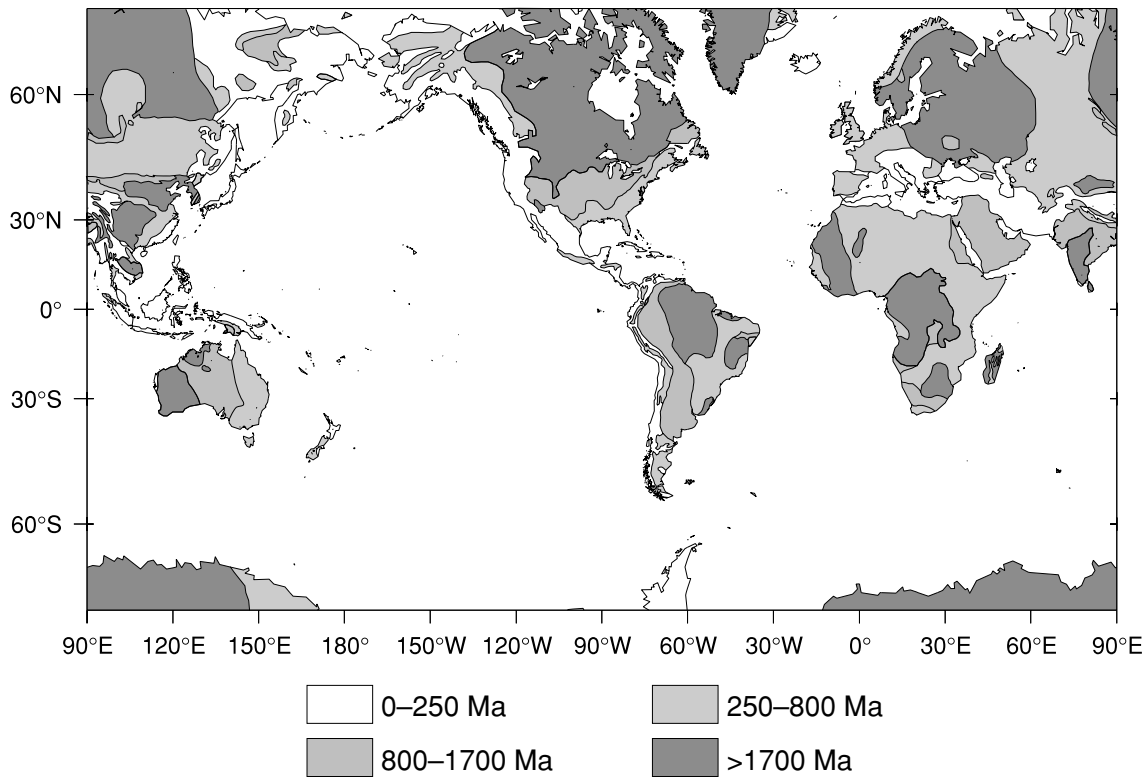


Figure 10.5. The age of the continents. (After Sclater *et al.* (1981).)

cratons are the Grenville and Appalachian rocks, which accreted much later as a result of continental collision. In the west, the Cordilleran rocks (western mountains) comprise a series of *accreted terranes*, which have been added, or accreted, to the continent during the last 200 Ma. These terranes, which comprise material of continental, oceanic and island-arc origin, have been added to the

North American plate as a result of plate tectonics and subduction. Terranes that are suspected of having originated far from the present location and of being transported and then accreted are descriptively referred to as *suspect terranes*! Such terranes were first identified in the early 1970s in the eastern Mediterranean in Greece and Turkey, when it was realized that the region is composed of small continental fragments with very different histories. The present-day subduction and volcanism along the western edge of the North American plate is thus continuing a history that has been occurring there episodically for several hundred million years (see also Section 3.3.4).

This chapter does not proceed in chronological order from the beginning of the continents to the present but instead works from the present back into the past, or, in the case of North America, from the edges to the centre. Our starting point is to continue the discussion of subduction zones from Chapter 9.

10.2 The growth of continents

10.2.1 Volcanism at subduction zones

The geophysical setting of subduction zones has been discussed in Section 9.6. The initial dip of the subducting plate is shallow, typically about 20° for the first 100 km, as seen horizontally from the trench, or on average about 25–30° in the region from the surface to the point at which the slab is 100 km deep. Volcanic arcs are characteristically located more than 150 km inland from their trench. This distance is variable, but it is clear from Fig. 9.46 that, with the exception of the New Hebrides, which is a very steeply dipping subduction zone in a very complex region, the volcanic arcs are located above places where the top of the subducting plate reaches a depth of about 100–125 km. The crust under volcanic arcs is usually fairly thick, in the range 25–50 km. The volcanic arcs are regions of high heat flow and high gravity anomalies. Despite the broad similarities, the settings of arc volcanism vary tremendously, and the styles of volcanism and the chemistry of the lavas vary in sympathy. The settings range from extensional to compressional and from oceanic to continental. In each case the product is different.

The descending slab: dehydration of the crust

The subducting plate or descending slab (both terms are used) is cooler than the mantle. Thus, as it descends it is heated and undergoes a series of chemical reactions as the pressure and temperature increase. The oceanic crust is heavily faulted and cracked and water has usually circulated through it in hydrothermal systems that became active soon after the crust formed at the ridge. The crust of the descending slab is therefore strongly hydrated (up to several per cent H_2O). All the chemical reactions which take place in the descending oceanic crust are *dehydration reactions*; that is, they involve a loss of water, usually in an endothermic process with a reduction in volume of the residue.

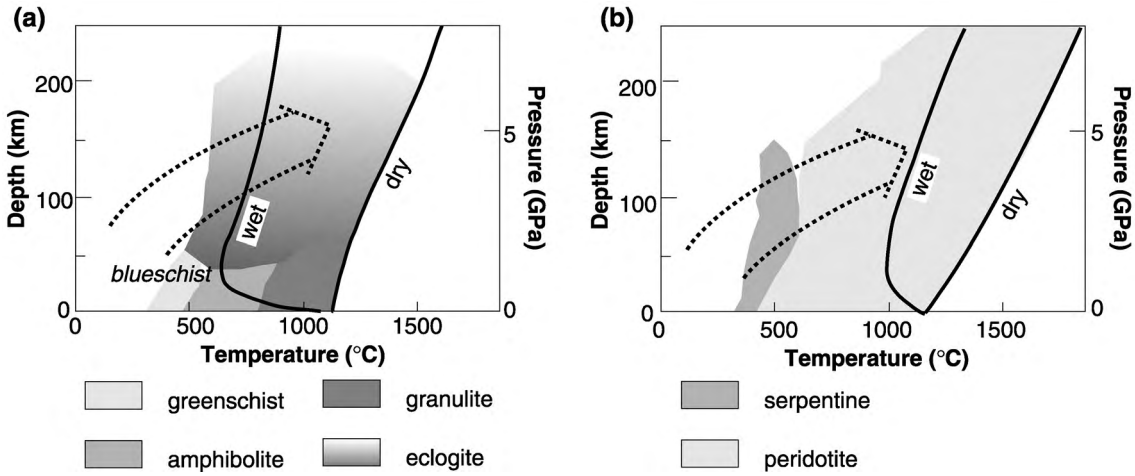


Figure 10.6. (a) A schematic representation of equilibrium metamorphic facies and melting of basaltic oceanic crust. Note that all the boundaries are gradational, not sharp, and that many of the reactions involved have not been particularly well defined. (b) A schematic representation of the breakdown of serpentinite and melting of peridotite. The two heavy lines are the solidus in the presence of excess water (wet melting) and the solidus for the dry rock (dry melting). The large arrow indicates the probable range of the temperature–depth profile of the subducting oceanic crust. (After Wyllie (1981) and Turner (1981).)

Prior to subduction, much of the oceanic crust is altered to low-grade brownstone, or at greater depth is in the greenschist metamorphic facies (Fig. 10.6(a)). This alteration was produced by the hydration and metamorphism of basalt in the near-ridge hydrothermal processes. During the initial stage of subduction, at shallow depths and low temperatures, the oceanic basalt passes through the pressure–temperature fields of the prehnite–pumpellyite and blueschist facies. At this stage, extensive dehydration and decarbonation take place as the basalt is metamorphosed beginning with expulsion of unbound water and followed by significant metamorphic dehydration that commences at a depth of about 10–15 km. During the descent of the slab, the pressure increases and the slab slowly heats up. The heat which warms the slab is transferred from the overlying wedge of mantle and is also produced by friction. The basalt then undergoes further dehydration reactions, transforming from blueschist to eclogite. During this compression, the water released, being light, moves upwards. Any entrained ocean-floor sediment also undergoes progressive dehydration and decarbonation. These processes are illustrated in Fig. 10.6(a).

The temperatures and pressures at which oceanic basalt, in the presence of excess water, produces significant amounts of melt are markedly different from the temperatures and pressures at which dry basalt produces copious melts (Fig. 10.6(a)). At 10 km depth, wet basalt begins to produce significant melt at about 850 °C and dry basalt at 1200 °C. For increasing amounts of water, the point at which copious melting begins lies at positions intermediate between the

dry and wet extremes: voluminous melting can start anywhere between the two curves, depending on the amount of water present. However, except in limited regions where heating is especially rapid, melting of the subducted oceanic crust is unlikely to occur during the early stages of descent.

The subducted oceanic crust carries wet oceanic sediment. At or near the trench, much of this sediment is scraped off and becomes part of the accretionary wedge (see Section 10.2.2). However, a small part of the sediment may be subducted. This subducted sediment melts at comparatively low temperatures (although these temperatures might not be attained on the surface of a subducting slab until depths >150 km) and provides components (such as CO_2, K and Rb) to the stream of volatiles rising upwards. ^{238}U–^{230}Th disequilibrium evidence shows that, beneath the Mariana arc, at least 350 000 yr elapses between melting of sediment and its incorporation into mantle melt.

The lower, plutonic (gabbro) portion of the subducted crust and the uppermost subducted mantle (peridotite) may also have been partially hydrated by sub-seafloor hydrothermal processes. Figure 10.6(b) shows the controls on dehydration of hydrated ultramafic rock in the lowest crust and topmost mantle as water is driven off any serpentine in the rock. The subducted oceanic mantle, which is depleted and refractory, does not usually melt. Being cooler than normal upper mantle at these depths, the subducted mantle simply heats up slowly towards the temperature of the surrounding mantle.

The descending slab: heating

The descending slab heats up by conduction from the overlying hotter mantle, but other factors that combine to speed up the heating process are also operating.

Some contribution to the heating of the slab may come from friction on its upper surface. This frictional or shear-stress heating is, however, not well quantified. Estimates of shear stress (0–100 MPa) are used in thermal modelling. Results of detailed studies of the heat flow measured in the fore-arc region, where there is no thermal contribution from the volcanic arc, indicate that the shear stress is low and probably lies between 10 and 30 MPa. Figure 10.7 shows a series of conductive models of the thermal structure of subduction zones dipping at 26.6°, the average dip of subduction zones down to 100 km depth (see Fig. 9.59) and with shear-stress heating increasing from zero at the surface and then decreasing to zero when the top of the slab reaches 100 km depth. The temperatures of the subducted oceanic crust are not high enough for it to melt in this region; the wet solidus for basalt (shown in Fig. 10.6(a)) is not attained in the upper part of the slab. Thermal models suggest that normally subducted oceanic crust reaches temperatures no greater than 500–700 °C at depths of ~125 km (Fig. 10.6(a)). The subducting oceanic crust is progressively subjected to temperatures and pressures appropriate to blueschist to eclogite facies. Greatly increased values of the shear stress and/or much higher initial temperature gradients for the subducting and overriding plates would be necessary for melting of the crust itself to take place in this region. This is a good check on the validity of models and the

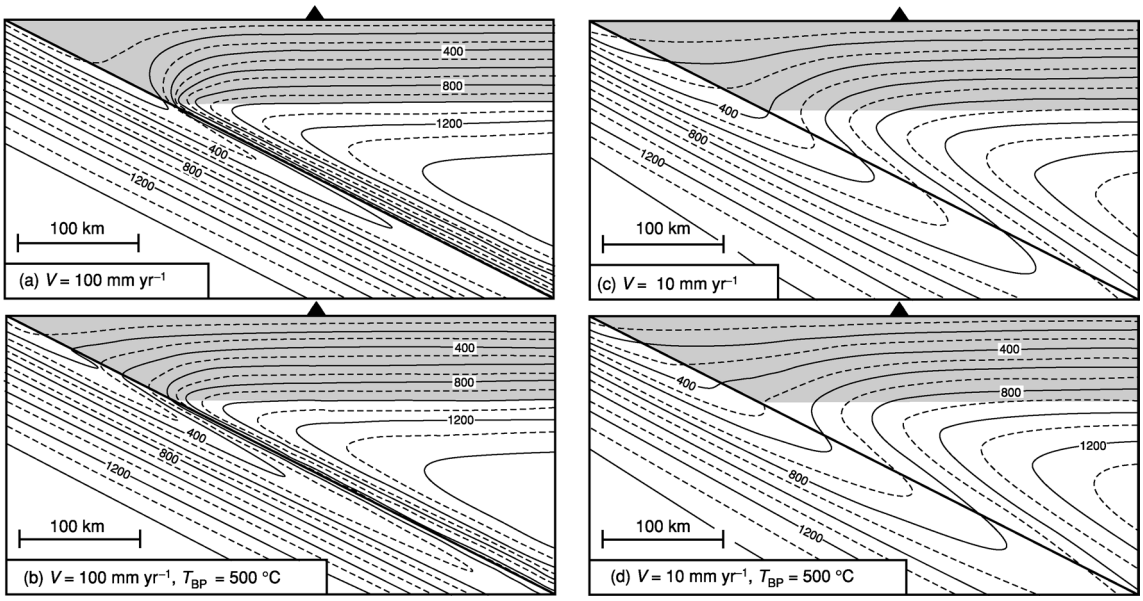


Figure 10.7. The steady-state thermal structure of the upper part of a subduction zone dipping at 26.6°. The age of the downgoing lithosphere at the start of subduction is 50 Ma. Solid triangles mark the location of the volcanic line. The shaded region is the 65-km-thick rigid overriding lithosphere (the base of this mechanical lithosphere is taken as being at 1000 °C). Isotherms are labelled in °C. In the mantle beneath the lithosphere the adiabatic gradient is 0.3 °C km^{-1}. Subduction with no shear heating: (a) at 10 cm yr^{-1} and (c) at 1 cm yr^{-1}. Subduction with shear heating that increases linearly with depth as 5% of lithostatic pressure down to the brittle–ductile transition (taken to be at 500 °C) below which the shear stress then decreases exponentially: (b) at 10 cm yr^{-1} and (d) at 1 cm yr^{-1}. (From Peacock, Thermal and petrologic structure of subduction zones, Geophysical Monograph 96, 113–33, 1996. Copyright 1998 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

physical conditions chosen: any models that imply that melting should take place between the trench and the volcanic line must fail, for melting should first take place only beneath the volcanic line. The oceanic crust probably starts transforming to eclogite by the time it has subducted to a depth of about 50 km, but not all of the reactions need be complete until it is much deeper than that because transformation takes time (though it is hastened by the presence of abundant fluid).

Many dehydration reactions of the oceanic crust are endothermic (the reactions require heat), and lack of heat may constrain the transformation. This means that temperatures in the real slab are lower than those obtained from computer models that do not include such heat requirements. Some estimates of the heat needed are 5.8×10^4 J kg^{-1} for the mineral reactions involved in the greenschist–amphibolite change and 2.5×10^5 J kg^{-1} for the serpentinite–peridotite reactions. The water released in these reactions rises into the overlying mantle, a process that further

slows the heating of the slab and cools the overlying wedge. A total value of about 2.5×10^5 J kg^{-1} for all the dehydration reactions is probably not unrealistic. None of these reactions takes place instantaneously (this point is discussed further in Section 10.3.310.3.5). At low temperatures and pressures, the reactions proceed more slowly than they do at higher temperatures and pressure; furthermore, under very dry conditions, at great depth, transformation is also slow.

The released water moves out of the oceanic crust into the overlying upper mantle. Dehydration probably begins with the onset of subduction but, at shallow depths, has little effect on the overlying wedge of the overriding plate except to stream fluid through it and metamorphose it. Dehydration continues as the plate is subducted: a plate being subducted at an angle of 20° at 8 cm yr^{-1} for 1 Ma descends 27 km, with attendant heating, compression and metamorphic dehydration. Figure10.6(b) indicates that dehydration of serpentine starts when the top of the slab reaches a depth of roughly 70 km. At such depths, but shallower than 100 km, water released from the slab is probably fixed as amphibole in the overlying mantle wedge, or streams upwards if the wedge is fully hydrated; but, at a depth of 100 km, the melting of wet overlying mantle begins. In the descending slab, the wet solidus for basalt probably can be reached only at a depth of about 100–150 km, so melting of the subducted oceanic crust normally cannot take place much shallower than this (but see the end of this section regarding some special circumstances). The precise depth depends on such factors as the age (and hence temperature and degree of hydration) of the slab and the angle and rate of subduction. In reality, probably only the melting of the overlying wedge occurs at 100 km; melting of the slab may not take place until it is much deeper, because by this stage the slab must be highly dehydrated.

Since the released water is unlikely to be able to leave the subducting plate and enter the overlying mantle wedge by porous flow, there must be some other mechanism. A high pore pressure in this non-percolating water would act as a lubricant, reducing friction on the subduction zone. This would facilitate slip and allow intermediate-depth earthquakes to occur. A large earthquake could connect sufficient water along the fault plane to initiate a hydrofracture. This hydrofracture would then transport the water into the overlying mantle wedge, where it would initiate partial melting. If there is down-dip tension in the downgoing slab, the hydrofractures will propagate upwards perpendicular to the slab (perpendicular to the least compressive stress). Beneath Japan intermediate-depth earthquakes occur right at the top of the subducting Pacific plate and are due to brittle ruptures. Figure 9.48 shows clearly that the low-velocity zones within the mantle wedge beneath Japan are inclined within the wedge and that the subducting plate and the volcanic arc are not directly connected vertically by low-velocity wet/molten/hot material. The velocity structures imply strongly that the paths taken by the rising volatiles and melts are along the shear zones in the mantle wedge, i.e., along the flow-lines, which, in the lower part of the wedge, are parallel to the top of the subducting plate. Both volatiles and melt take the easiest route to the surface, which is not necessarily the shortest route.

Isotopic work has indicated that the subducting slab provides only a small component of erupted lavas. Data from Chile have shown that the overriding mantle provides a significant proportion of all elements: only volatile species such as H_2O and the *large-ion lithophile* elements such as Rb, K, Ba, Th and Sr were supplied in any great quantity by the subducting slab. These elements are transported upwards to the overlying mantle in the volatiles ascending from the subducted slab to become incorporated into the rising melt.

The overriding mantle wedge

It is assumed in many thermal models of subduction zones that heat is transferred by conduction alone, which is an extreme simplification; more realistic models incorporate a viscous, convecting mantle. The inflow of overriding mantle ensures a supply of fertile asthenosphere, with a potential temperature (Eq. (7.94a)) of about 1280 °C, to the region above the descending slab.

Two thermal models are shown (Figs. 9.44 and 10.7). Temperatures close to the wet solidus (Fig. 10.6(b)) are reached in the vicinity of the descending slab at depths in excess of about 100 km. This is important because it means that the mantle temperatures there are in the range within which addition of water results in copious partial melting. Most melt is generated in the upper mantle beneath the volcanic arc (the mantle wedge) as a result of the addition of water and other volatiles from the subducting slab. As discussed earlier, water is lost by the slab at all depths down to about 100 km, initially due to compaction (the closing of the pores) and then to the dehydration reaction. However, it is only at depths at which the overlying mantle temperature is more than 1000 °C that partial melting can take place in the mantle wedge; this condition seems to be satisfied at about 100 km depth in most subduction zones.

The effects of subduction on the overriding mantle wedge can be summarized as follows: (1) an influx from the descending slab of upward-moving volatiles; (2) some melt rising from deeper parts of the slab; and (3) the driving of convective flow in the wedge, the flow-lines of which show movement of mantle material from the distant part of the wedge (on the right in Fig. 10.8) and pulled downwards (counterclockwise in Fig. 10.8) by the slab.

The upward-moving volatiles from the descending slab consist of H_2O and CO_2, probably accompanied by a substantial flux of mobile elements such as Rb, K, Ba, Th and Sr. Melting of any subducted sediment may be important. The contribution of melt (as opposed to volatiles, etc.) from the slab itself is probably relatively small. The deeper parts of the slab may leak some melt upwards, leaving residual quartz-eclogite behind. Compositionally, any melt rising from the slab is probably hydrous, siliceous magma, which is roughly similar to calc-alkaline magmas erupted in island arcs but probably with a higher CaO/(FeO + MgO) ratio.

The addition of streams of volatiles, plus perhaps some melt at deeper levels, causes partial melting in the warm peridotite overlying the subducted slab

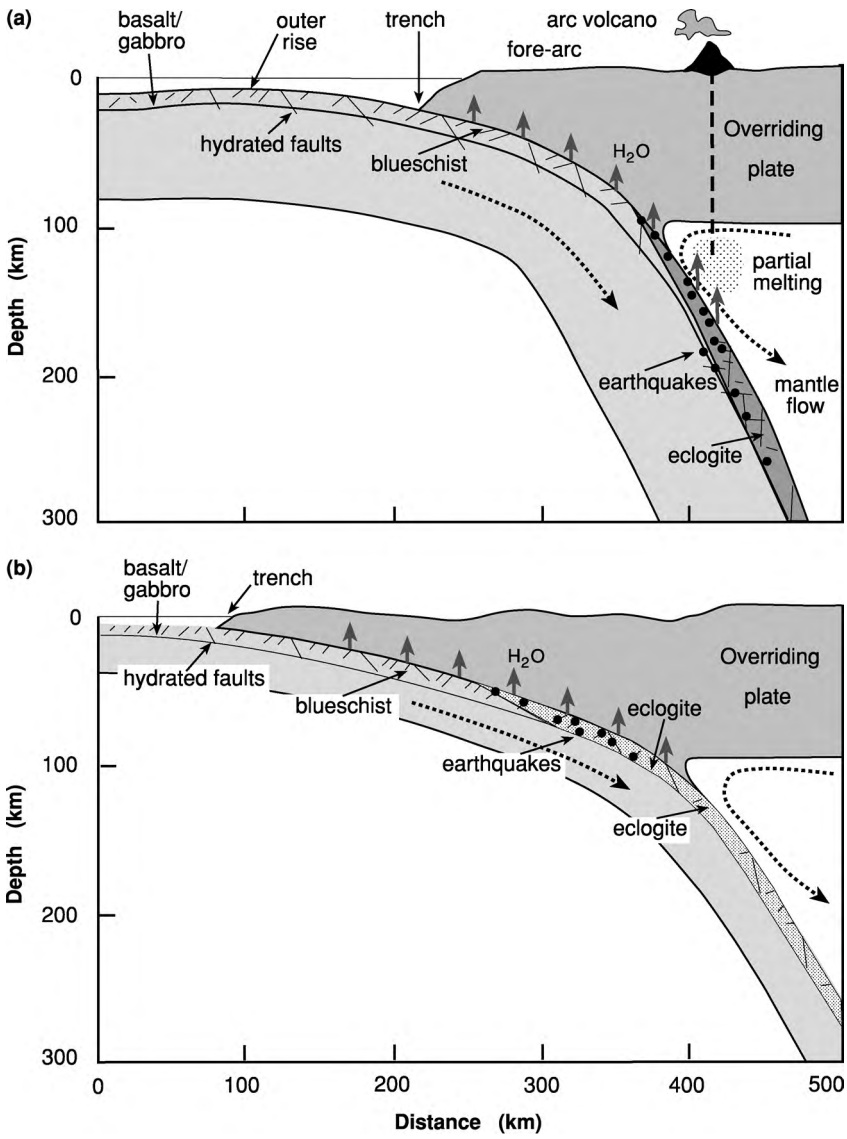


Figure 10.8. Cross sections of subduction zones, showing the metamorphism of the downgoing slab, associated earthquakes, loss of volatiles and arc magmatism. (a) Thermally mature subducting plate – high thermal parameter. Intermediate-depth earthquakes result from the dehydration and transformation to eclogite of the oceanic crust and reactivation of faults. (b) Young and/or slowly subducting plate – low thermal parameter. Since the dehydration, formation of eclogite and reactivation often take place at shallow depths, the amount of partial melting in the mantle wedge and hence arc magmatism may be reduced. (Based on Kirby *et al.* (1996) and Peacock (1996).)

(Fig. 10.9). There is still uncertainty about the exact location of melt generation, but it is probable that much of the melting occurs in the warm overlying wedge immediately above the locus where the slab reaches 100–120 km depth. The ^{238}U–^{230}Th disequilibrium means that, for the Mariana arc, the time delay between dehydration of the slab and eruption of the lavas is less than 30 000 yr. Such a short time interval implies that the fluids are the primary cause of mantle melting and that the melt migration is rapid. In contrast, the timescale for transport of the signature of subducted melted sediment to the melt source region is

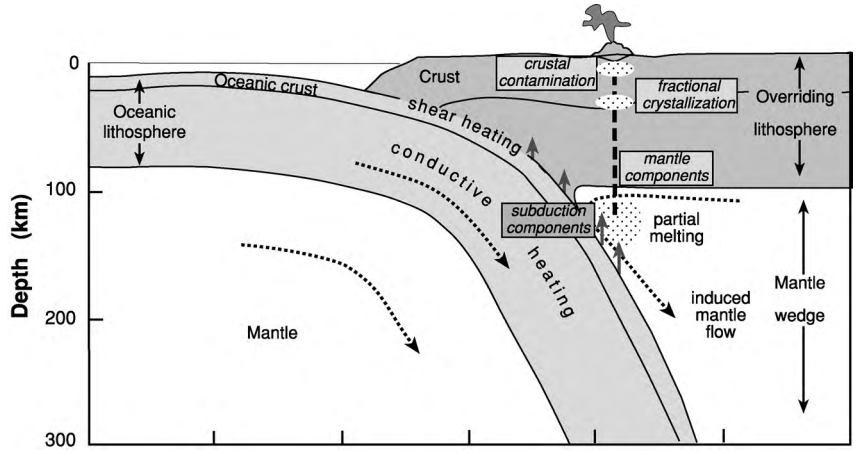


Figure 10.9. A cross section of a subduction zone, showing the main factors and processes which control its thermal structure and the volcanic arc. The age and rate of subduction of the oceanic plate, as well as the dip and amount of shear heating, are important factors controlling the thermal structure. The loss of water and volatiles from the downgoing oceanic plate causes melting in the mantle wedge. The melt is contaminated and fractionated during ascent through the overriding lithosphere. (Based on Pearce (1983).)

over 350 000 yr. Shallower melting in the wedge, at depths of 60–90 km, would produce liquids that may be parental to basaltic andesite. Andesite, which is the commonest magma erupted from arc volcanoes, cannot be directly produced by melting of the mantle wedge except by melting of very wet (about 15% H_2O) mantle peridotite at depths of about 40 km, or under other very restricted circumstances. Most probably, the fluids from the subducted slab promote melting of the mantle wedge, producing basic melts that rise because they are lighter than the surrounding residual peridotite and eventually reach the base of the continental crust.

The whole subduction melt system can be thought of as a conveyor belt emitting volumes of melt that rise upwards: warm mantle convects into place above the subduction zone; fluid enters from the slab below and initiates melting; melt escapes upwards; the residual mantle is carried away by convection; and new warm material takes its place.

The base of the continental crust

Partial melts generated in the mantle wedge rise either as diapirs of melt and crystals, which increasingly melt and become more basic as they rise and decompress, or as rapidly moving segregated melt. The rising melt probably follows a nearly adiabatic pressure–temperature path in a mature subduction system, cooling by up to $1\,^{\circ}C\,km^{-1}$ (this is the adiabatic gradient in a magmatic liquid, in contrast to $0.5\,^{\circ}C\,km^{-1}$ in solid mantle). Much depends on the release of latent heat. At

the base of the continental crust, lighter liquids can pass straight upwards to the surface, but most liquids are probably trapped by their density, the continental crust being less dense than the magma.

Rising magma carries heat with it. Hot magma collecting at the base of the crust loses heat to the overlying continent and coals and fractionates, precipitating minerals such as clinopyroxene together with garnet, olivine or orthopyroxene (depending on the depth of the melt and on its temperature, composition and percentage of water). After fractionation, the lighter liquids rise to the surface, most probably as basalt and basaltic andesite magmas. However, the transfer of heat into the base of the continental crust also produces melting in the crust. High-temperature partial melting of deep continental crust at temperatures of about 1100 °C produces tonalite liquids, which are silica-rich melts that can rise to the surface and erupt as andesites. At somewhat lower temperatures (about 1000 °C), in the presence of more water, the product of melting is granodiorite. Many of the large granitic intrusions of the continents above subduction zones are granodiorite. In the aftermath of large-scale continental collisions (such as in the Himalayas), there is often overthrusting of continental crust with partial melting of the underlying slab. After partial melting, the residual material left behind in the deep continental crust is granulite, which is depleted of all its low-temperature-melting fractions. This depletion includes the removal of the heat-producing elements (U, Th and K), which are carried upwards with the rising granitoid liquids. Because these heat-producing elements are carried upwards, the continents are self-stabilizing: heat production is concentrated at the top, not the bottom, of a continent. This process, which has moved heat production to shallow levels, has had the effect of reducing the continental temperature gradient, making it more difficult to melt the crust (see Chapter 7, Problem 13).

It is chemically unlikely that granitic liquids are produced directly from partial melting of mantle peridotite or subducted oceanic crust; otherwise we should find granitoids in the oceanic lithosphere. The main geographic location of granitoids is above subduction zones and in continental collision zones, modern or ancient, which implies that granitoid generation is strongly linked to the processes of plate tectonics. In the modern Earth, it is probable that most granitoids are generated in the presence of water. In the Archaean, when the mantle may have been hotter, tonalites appear to have been more common, having been generated at 1300 °C from dry crust. Today, under cooler and wetter conditions, melts are granodioritic.

The upper continental crust above a subduction zone is characterized by large granodioritic (granitoid) intrusions. Above these are andesite volcanoes, which erupt melt that originated in the mantle above the subduction zone, but which has fractionated on ascent and perhaps been contaminated by material derived from the continental crust. However, the broad similarity of volcanism in island arcs, where in some cases little continental material is present, to volcanism on continents above subduction zones indicates that magma must be produced

within the mantle, irrespective of the type of overriding plate. The complication introduced by the continental crust is that it allows the possibility of further chemical complexity and variation in the erupted and intruded melts.

There are some particular conditions under which the subducting hydrous basaltic slab can melt – generally, if it is very young (less than 5 Ma) and hot, it can melt at depths of 60–80 km. The melts that result are termed adakites, being dacitic in composition. However, adakites are not confined to volcanism over very young subducting plates – any process that causes the temperature of the subducting plate to be higher than normal will mean that melting of that plate and formation of adakites may be possible. Particular instances where this may occur include a very slow subduction rate or a torn subducting plate. This may explain the presence of adakites in the Aleutian/Kamchatka corner region (Fig. 2.2). In addition, they are also found over older subducted lithosphere, apparently when the subduction zone has temporarily had a very shallow dip – perhaps when some buoyant crust, such as an oceanic plateau, enters the subduction zone – temporary 'flat subduction'.

10.2.2 Sediments at subduction zones

A schematic cross section of a subduction zone was shown in Fig. 9.42, which illustrates the characteristic geological features. In reality, a subduction zone does not necessarily have all these features. The accretionary wedge is the region of folding, and then of faulting and thrusting of the sediments on the subducting oceanic plate. Then comes the outer-arc high. The fore-arc basin is an active sedimentary basin, being filled largely by material (detritus) eroded from the adjacent arc. It may be underlain by oceanic crust marking the position of the old passive continental margin before subduction began. The volcanic arc and the back-arc region were discussed in Sections 10.2.1 and 9.6.

One subduction zone with a well-developed accretionary wedge is the Makran subduction zone in the Gulf of Oman off Iran and Pakistan. There the oceanic part of the Arabian Plate is being subducted beneath the Eurasian Plate. This 900-km-long subduction zone is unusual in a number of ways:

1. the dip of the subducting Arabian Plate is very low;
2. there is no clear expression of an oceanic trench;
3. the background seismicity is very low; and
4. there is a prominent accretionary wedge, much of which is at present exposed on land in Iran and Pakistan.

Figure 10.10 shows a seismic-reflection profile across part of the offshore portion of this thick accretionary wedge. The undeformed abyssal plain sediments on the Arabian plate are 6–7 km thick. The deformation of these sediments as they are pushed against the accretionary wedge is clearly visible. A gentle frontal fold

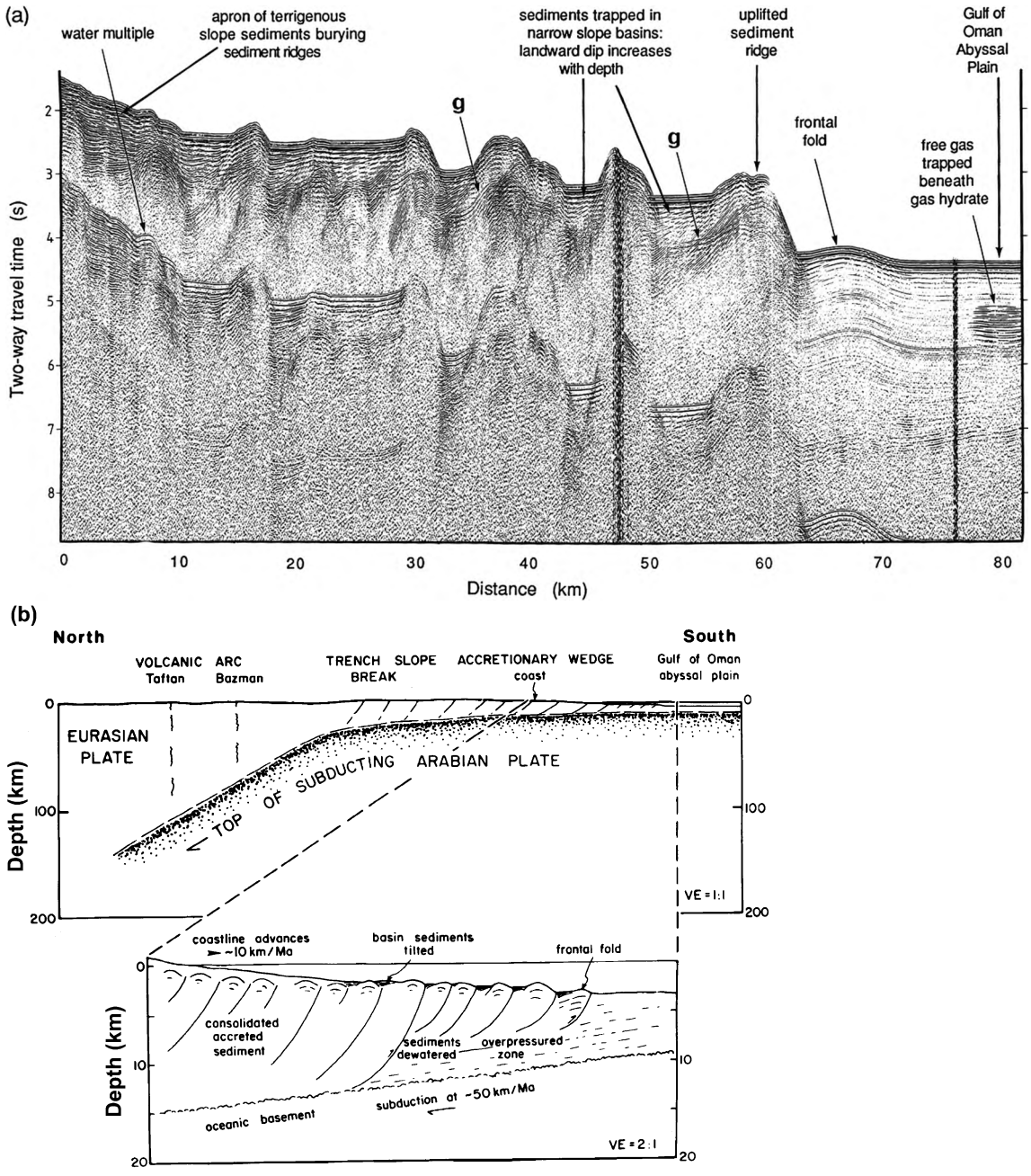


Figure 10.10. (a) A seismic-reflection profile across the accretionary prism of the Makran subduction zone. The bottom-simulating reflector of the base of gas hydrates is marked g. (b) Cross sections through the Makran subduction zone. (From White (1984).)

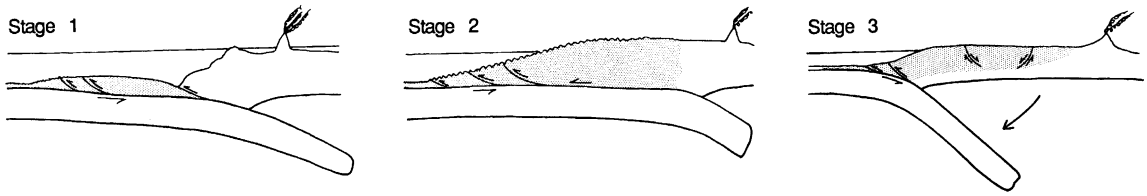
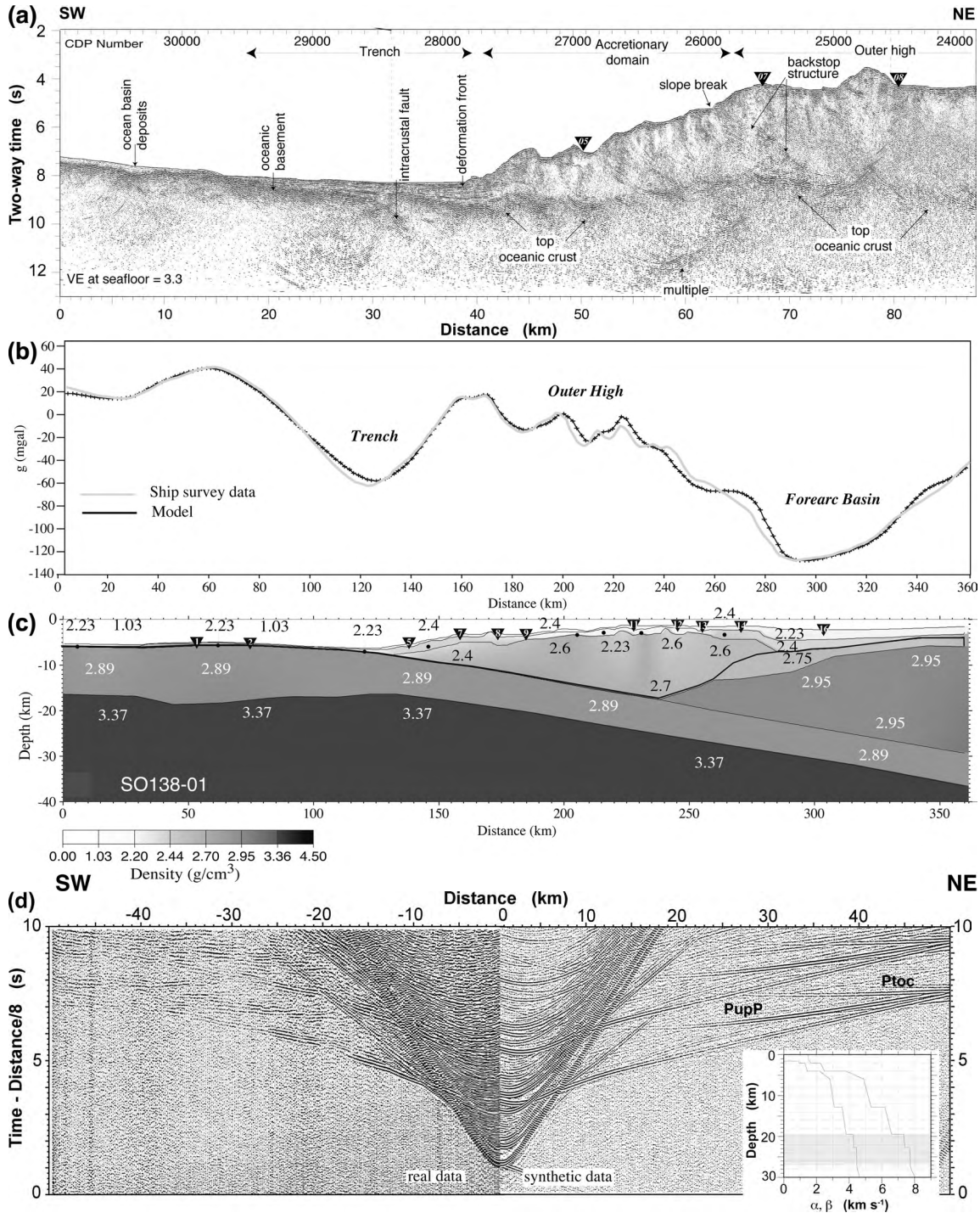


Figure 10.11. Three stages in the development of a subduction zone when there is a great thickness of sediment on the oceanic plate. Stage 1: a trench is visible. Stage 2: the large amount of sediment scraped off the subducting oceanic plate has choked the trench. The accreted sediments all belong to the overriding plate, and thrusting takes place beneath them. The Makran subduction zone is now at this stage. Stage 3: the subducted plate sinks into the asthenosphere, resulting in the gradual extinction of the volcanoes and, in the accretionary wedge, extension and (eventually) new volcanoes. (From Jackson and McKenzie (1984).)

develops first, followed by a major thrust fault, which raises the fold some 1200 m above the abyssal plain. Reflectors can be traced from the abyssal plain into the frontal fold but not beyond because deformation and faulting are too extreme further into the wedge. Detailed wide-angle seismic-velocity measurements show that dewatering of the sediments, and hence compaction, occurs in the frontal fold. The sediments are then sufficiently strong to support the major thrust fault. The continuous process of forming this accretionary wedge results in the southward advance of the coastline by 1 cm yr^{-1}. By this process, a considerable volume of material is being added to the Eurasian plate every year. Figure 10.11 illustrates the possible stages in the development of a thick accretionary wedge. Sometime in the future, the situation shown in Fig. 10.11 may be appropriate for the Makran subduction zone: the subducted Arabian plate may fall into the asthenosphere, resulting in extension and a new volcanic region in the present accretionary wedge.

Figure 10.12(a) shows the detail of the style of deformation in the accretionary wedge in a more 'normal' subduction zone, the Sunda Arc, where the Indian plate is being obliquely subducted beneath Sumatra at ~7 cm yr. The plate dips at about 3° close to the deformation front, but the dip increases with depth. The rugged top of the oceanic crust can be clearly seen beneath the accretionary wedge. The seismic-velocity and density models determined from detailed

Figure 10.12. (a) A seismic-reflection line across the Sumatra Trench. The active part of the accretionary wedge extends from the detachment front to the slope break that marks the backstop structure. Intense faulting within the accretionary prism makes imaging the detail of the structure very difficult. (b) The gravity anomaly. (c) The best-fitting density model. (d) Recorded (left) and synthetic (right) wide-angle seismic data from a strike line located at 230 km on the cross section (b). On the velocity–depth structure, the depth extent of the subducted plate is shaded gray. (From Kopp *et al.* (2001).)


(a)
SW
NE
CDP Number
Trench
Accretionary domain
Outer high
backstop structure
slope break
ocean basin deposits
oceanic basement
intracrustal fault
deformation front
top oceanic crust
multiple
VE at seafloor = 3.3
Two-way time (s)
Distance (km)
(b)
Outer High
Trench
Forearc Basin
Ship survey data
Model
g (mgal)
Distance (km)
(c)
SO138-01
Depth (km)
Distance (km)
Density (g/cm3)
(d)
SW
NE
Distance (km)
Time - Distance/8 (s)
Ptoc
PupP
real data
synthetic data
Depth (km)
α, β (km s-1)

wide-angle-reflection/refraction and gravity data across the plate boundary clearly delineate the two plates (Figs. 10.12(c) and (d)). The density and seismic-velocity models indicate that the crust beneath the fore-arc basin is continental, not oceanic. Figure 10.12(d) shows the wide-angle seismic data from a strike line recorded by one of the ocean-bottom seismometers along the profile and synthetic seismic data that best fit it. The P-waves that have travelled in the overriding plate are labelled $P_{up}P$; those that travelled in the top of the subducting plate are labelled P_{toc}.

10.2.3 Continent–continent collisions

Because continental lithosphere is not dense enough to be subducted as a whole into the mantle, the collision of two continents results in a complex process of thrusting and deformation, involving a reduction and finally a cessation of relative motion. Other plates reorganize to take up the motion elsewhere. Two classic examples of young mountain ranges formed from such continental collisions are the Himalayas, which are the result of the collision of the Indian plate with Eurasia, and the Alps, which are a result of the northward motion of the African plate towards Eurasia.

The Himalayas

Body- and surface-wave studies of earthquake data indicate that the crust beneath the Himalayas and the Tibetan Plateau is over 70 km thick. This is in contrast to the 40 km-thick crust of the Indian shield. India has a typical shield S-wave velocity structure with a thick, high-velocity lithosphere overlying the asthenosphere. However, the lithospheric structure beneath Tibet is complex, and indicates that the Indian plate is underthrusting not all, but only part of, the Tibetan Plateau and that Tibet is not a typical shield region (Fig. 10.13).

The major tectonic blocks of the Himalayas and the Tibetan Plateau and their sutures are shown in Fig. 10.14(a). Figure 10.14(b) shows one attempt to explain the overall evolution of the region. This evolution has been much more complex than just a simple collision of India with Eurasia. The reconstruction starts at 140 Ma with the Kunlun and Qiangtang blocks already sutured to Eurasia and the Lhasa block moving northwards as an oceanic plate is subducted. By 100 Ma, the Lhasa block was attached to Eurasia and may have undergone internal thrusting and intrusion while subduction moved to its southern margin. Shortening by up to 60% seems to have taken place in the northern Lhasa block. Continuing subduction beneath the southern margin of the Lhasa block meant that the Indian continent moved northwards (at about $10\,\mathrm{cm\,yr^{-1}}$; see Section 3.3.3) until, by $\sim$40 Ma, the continental collision occurred, at which time the rate of convergence of India and Eurasia suddenly dropped to $5\,\mathrm{cm\,yr^{-1}}$. Initially thrusting took place along the Yarlung–Zangbo (or Indus–Tsangpo) suture. (Zangbo, Zangpo, Tsangbo and Tsangopo are all used to transliterate the same Chinese word.) Thus

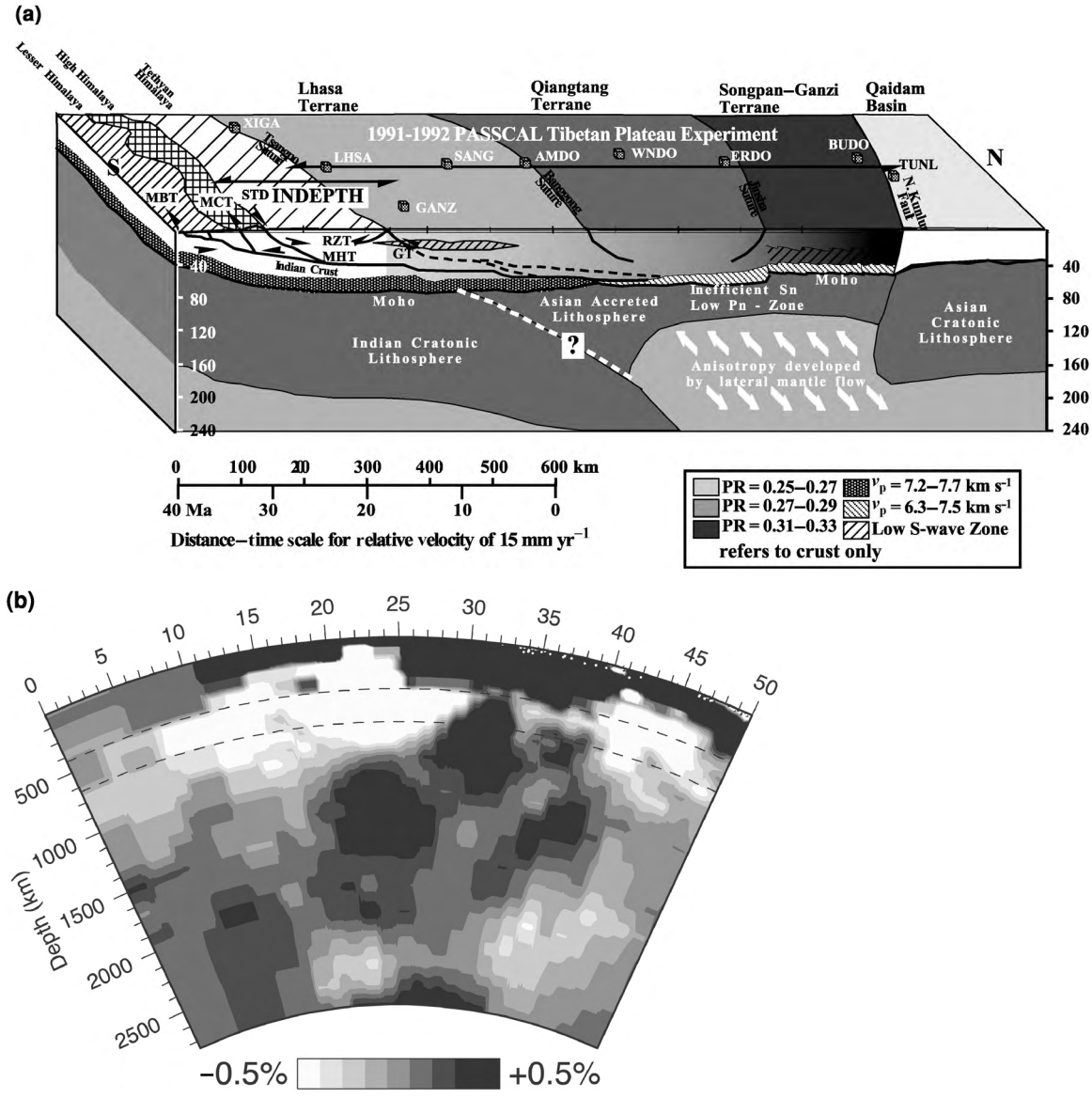


Figure 10.13. (a) A south–north section from India across the Himalayas and Tibet at 92–93°E summarizing the results of teleseismic and controlled-source studies and showing the geometry of the converging plates. PR, Poisson's ratio. There are two possible interpretations of these data: (1) that India underthrusts as far as the Banggong–Nujiang suture and (2) that India underthrusts only as far as the Zangbo suture before the plate dips steeply into the mantle. In case (2) the Lhasa terrane must have a very thick crust with a rigid lower crust and mantle. Colour version Plate 27. (Zandt, personal communication 2001). (b) The deviation of seismic P-wave velocity from a standard model of the mantle along a line across India and the Himalayas to Tibet. Note the thick high-velocity regions associated with the India–Eurasia collision through the upper mantle. Deep high-velocity anomalies may be subducted oceanic plate now in the lower mantle. Colour version Plate 25(d). (Spakman, personal communication 2003.)

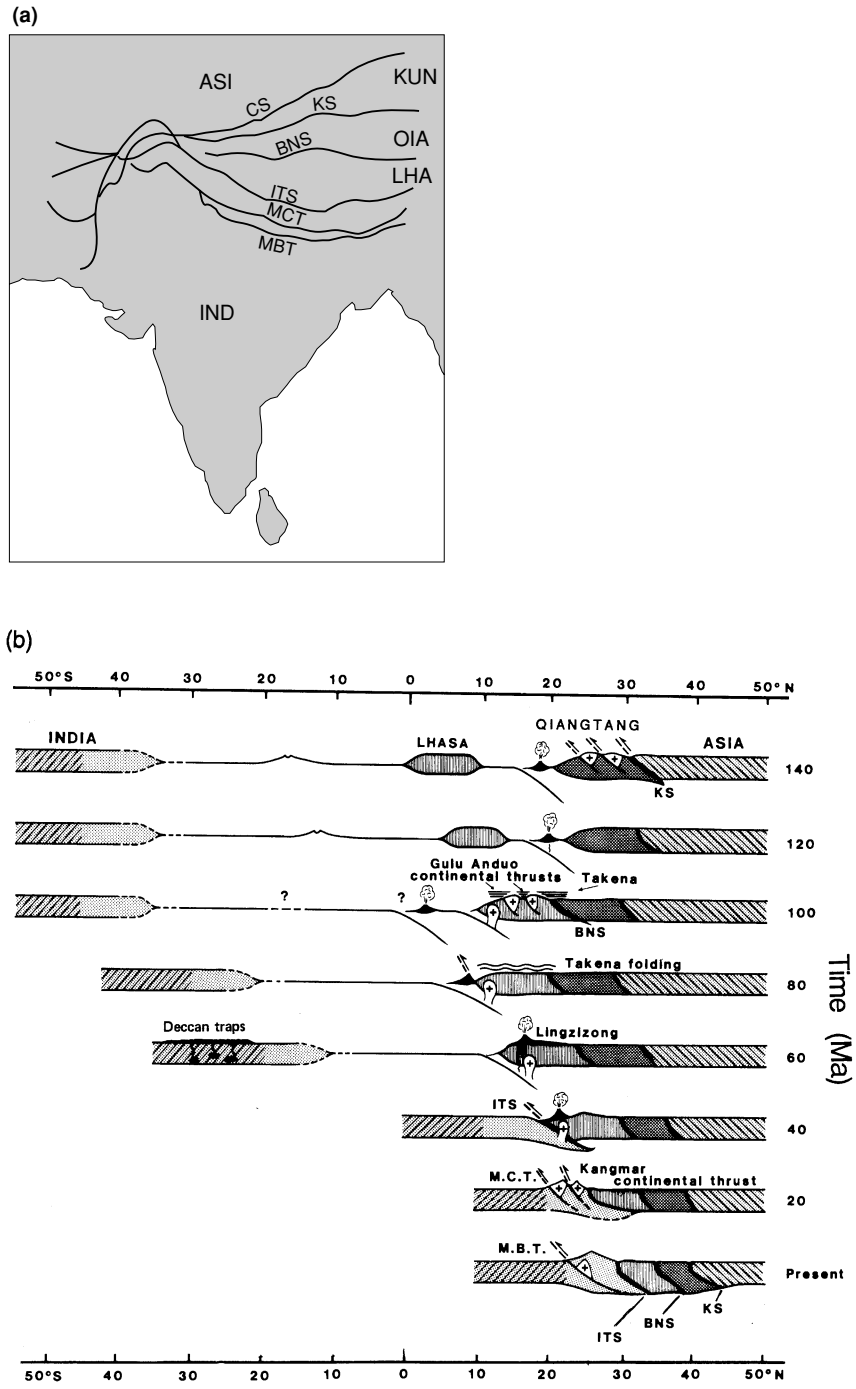


Figure 10.14. (a) The accretion of the terranes that make up the Himalayas. CS, Chilien suture; KS, Kokoxili suture; BNS, Banggong–Nujiang suture; ITS, Indus–Tsangpo (or Yarlung–Zangpo) suture; MCT, main central thrust; MBT, main boundary thrust (or fault). The tectonic blocks are ASI, Asian plate; KUN, Kunlun block; QIA, Qiangtang block; LHA, Lhasa block; and IND, Indian plate. (b) A reconstruction of Tibet and the Himalayas at 20-Ma intervals from 140 Ma to the present. The Qiangtang block is assumed to have sutured to Asia at about 200 Ma. 140–120 Ma: the small ocean basin between Asia and the Lhasa block closes. 100 Ma: the Lhasa block is sutured to Asia along the BNS. 80–60 Ma: subduction takes place beneath the Lhasa continental margin (including the possible subduction of a volcanic arc). 40 Ma: subduction ceases. Continental obduction or shortening occurs as the Indian and Asian plates collide. 20 Ma: MCT is the main thrust. Present: MBT is the main thrust. Note that the deep geometry of the bounding sutures between the terranes is purely schematic – details are shown in other figures in this section. (From Allègre *et al.* (1984).)

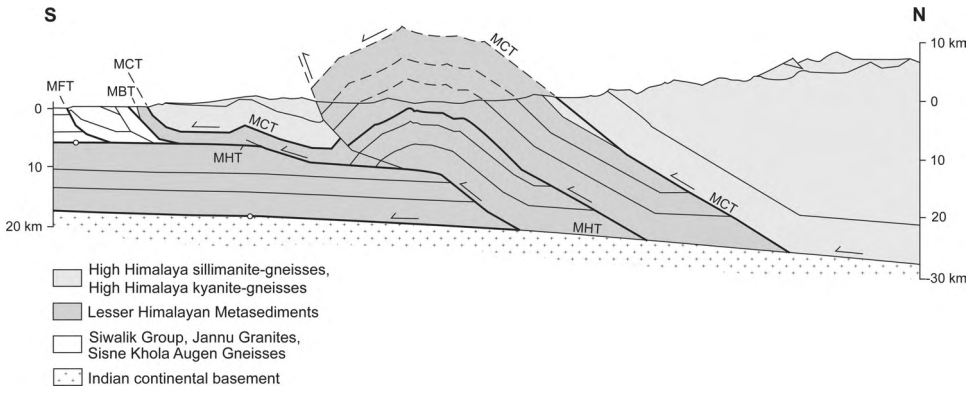


Figure 10.15. A balanced, restored north–south section across Nepal at ~88°E shows the probable sequence of thrusting events that gave rise to the Himalayas as observed today. Shortening of the crust since 16–25 Ma has occurred in three stages as the location of active thrusting moved progressively to the south: (1) along the MCT, (2) along the MBT and MHT and (3) mainly along the MBT and MHT. The total amount of shortening amounts to 200–250 km, with 40–70 km having taken place during stages 2 and 3. Thrusting, uplift and erosion have caused the double exposure of the MCT, the underlying Lesser Himalaya and the overlying lower-crustal rocks of the Higher Himalaya. (Based on Schelling and Arita (1991) and Ratschbacher *et al.* (1994).)

all the present Himalayan rocks were once part of India (Tethys), not Eurasia. The oldest continental sediments from the Himalayan foreland basin provide a constraint upon the timing of the start of erosion from the mountain chain and hence its uplift. Ar–Ar dating on the earliest continental sediments has yielded ages of 36–40 Ma.

The continental collision did not result in a simple overriding of India by Eurasia along the main Himalayan thrust (MHT); instead, as underthrusting proceeded, the active thrust was repeatedly blocked and thrusting migrated southwards, each time leaving a thick slice of Indian crust attached to the Eurasian plate (Fig. 10.15). Isostatic uplift and erosion then further modified the structures. The initial continental collision occurred along the Yarlung–Zango suture and left crust from the Indian continental margin and Tibetan fore-arc basin stacked beneath the Lhasa block. Since that initial collision about 500 km of shortening has taken place south of the Yarlung–Zangbo suture. South of the suture the Himalayan chain consists of three distinct tectonic units, bounded by thrust faults. These units are, from north to south, the Higher Himalaya, the Lesser Himalaya and the sub-Himalaya. After at least 100 km, and perhaps as much as 300 km, of underthrusting of India and its margin beneath Tibet along the Yarlung Zangbo suture had taken place, the main central thrust (MCT) developed to the south.

The Higher Himalaya crystalline rocks (a gneiss/leucogranite tectonic package) were then thrust over the Lesser Himalaya meta-sediments along the MCT. After 150–200 km of thrusting, the main boundary thrust (MBT) developed to the south (as a splay thrust off the main Himalayan thrust (MHT)), and the MCT became inactive. The Lesser Himalaya meta-sediments were then thrust over the sub-Himalaya molasse sediments along the MHT and the MBT. The sub-Himalaya is also underlain by the MHT, since a third stage of thrusting means that the young sediments of the Ganga (Ganges) foreland basin are underthrust beneath the sub-Himalaya. Estimates of the total amount of shortening in the Lesser and sub-Himalaya are ∼40–70 km. Since the initiation of the MCT between 16 and 25 Ma ago, there has therefore been in total some 200–250 km of shortening, with the average rate of shortening being 0.7–1.5 cm yr^{-1}.

The present situation has the main thrust plane between India and Eurasia being the MFT/MHT, while the underthrust Indian continental margin does not extend beneath the entire Tibetan Plateau. The MFT marks the southern extent of the deformation and the thin-skinned tectonics that are taking place in the Himalayan foothills. The main detachment between India and Eurasia is the MHT, which extends as a shallowly dipping plane beneath the Lesser Himalaya. Beneath the Lesser Himalaya the MHT steepens, reaching a depth of ∼25–30 km beneath the Higher Himalaya. The earlier thrusting, folding, uplift and erosion has exposed the MCT, the high-grade lower-crustal rocks in the overlying Higher Himalaya and the underlying medium-grade Lesser Himalaya. The thrust relationships are illustrated in Fig. 10.15. The main reason for uncertainty in the geology and in understanding the formation and structure of the Himalayan region is the extreme size and ruggedness of the terrain, which makes access and working there very difficult.

The Himalayas are seismically active: magnitude-8 earthquakes are not uncommon (there have been eight since 1816). Figure 10.16 shows fault-plane solutions for some earthquakes in the Himalayas. All the fault-plane solutions for the earthquakes immediately north of the MBT exhibit thrust faulting. The nodal plane, which is assumed to be the thrust plane, dips at a shallow angle of ∼15°. The focal depth for these earthquakes is 10–20 km, which, since the epicentres are some 100 km north of the MBT, is consistent with the earthquakes being located at the top of the Indian plate as it is subducted beneath Eurasia (see Figs. 10.13 and 10.19). North of this band of thrust-faulting earthquakes, the style of deformation changes: there is normal faulting and east–west extension at shallow depths over Tibet (Fig. 10.16). The earthquake that occurred at 78°E in the Indian plate well to the south of the Himalayas exhibits normal faulting and therefore is presumably indicative of the extension taking place in the top of the Indian plate as it bends prior to subducting beneath the Himalayas.

The present-day rate of convergence in the Himalayas is estimated to be ∼2 cm yr^{-1}, which is much less than half the estimated convergence between India and Eurasia (∼6 cm yr^{-1}). An estimate of the rate of shortening based on

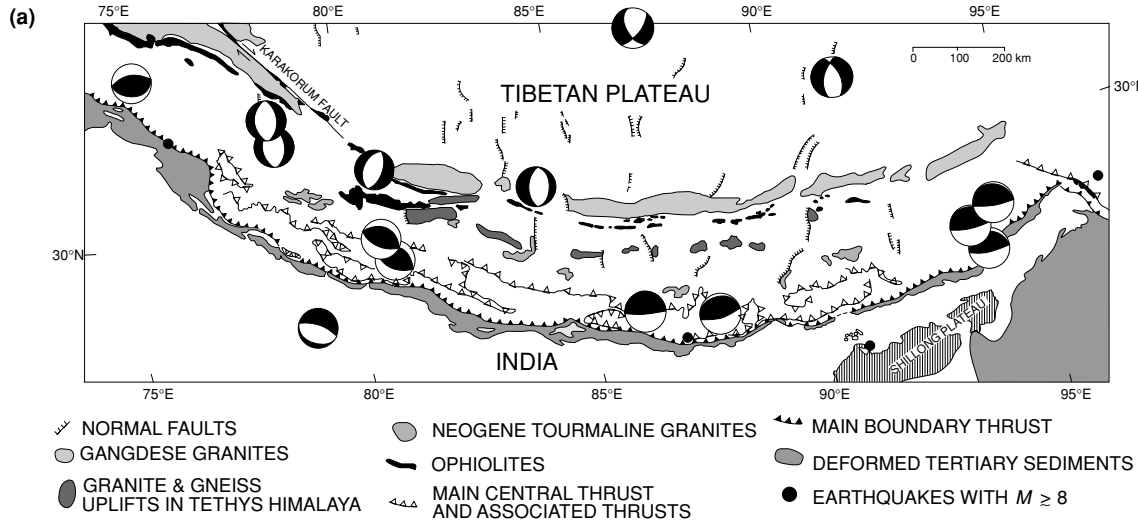


Figure 10.16. (a) A tectonic map and earthquake fault-plane solutions for the Himalayas. The southernmost thurst, the Main Boundary Thrust (MBT), is the present location of the plate boundary between India and Eurasia. The string of ophiolites delineates the Indus–Tsangpo suture (ITS), the original collision zone between India and Eurasia. The rocks between the ITS and the Main Central Thrust (MCT) were all part of the Indian (Tethyan) plate: first are the Tethyan passive-margin sediments; south of these are the crystalline rocks of the Higher Himalaya which were Tethyan crust. The rocks south of the MCT are very-thick, low-grade, clastic Precambrian/Palaeozoic sediments that make up the Lesser Himalaya. The rocks south of the MBT are Tertiary sediments of the sub-Himalaya foreland basin. The black dots show the epicentres of the $M \geq 8$ earthquakes which occurred in 1905, 1934, 1897 and 1950 (from west to east), for which fault-plane solutions are not available. (From Molnar and Chen (1983).)

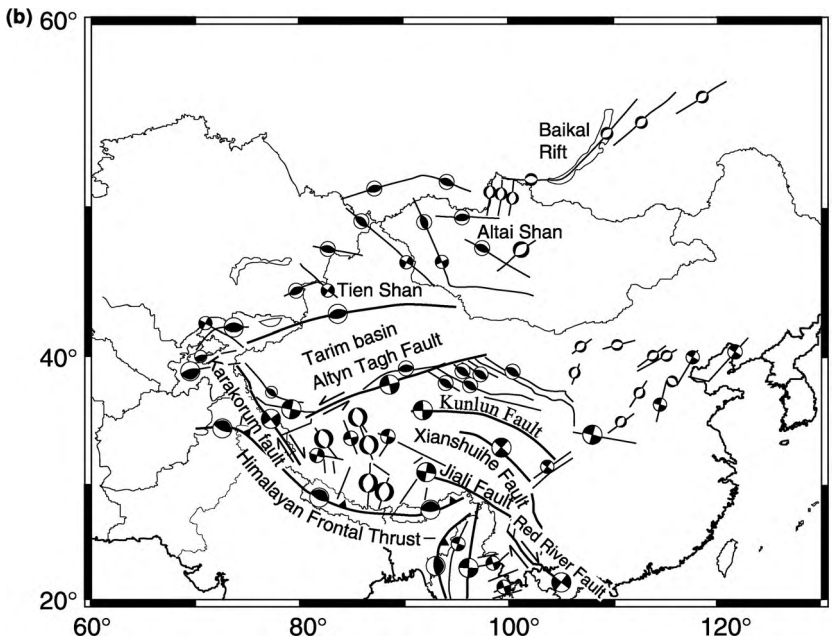
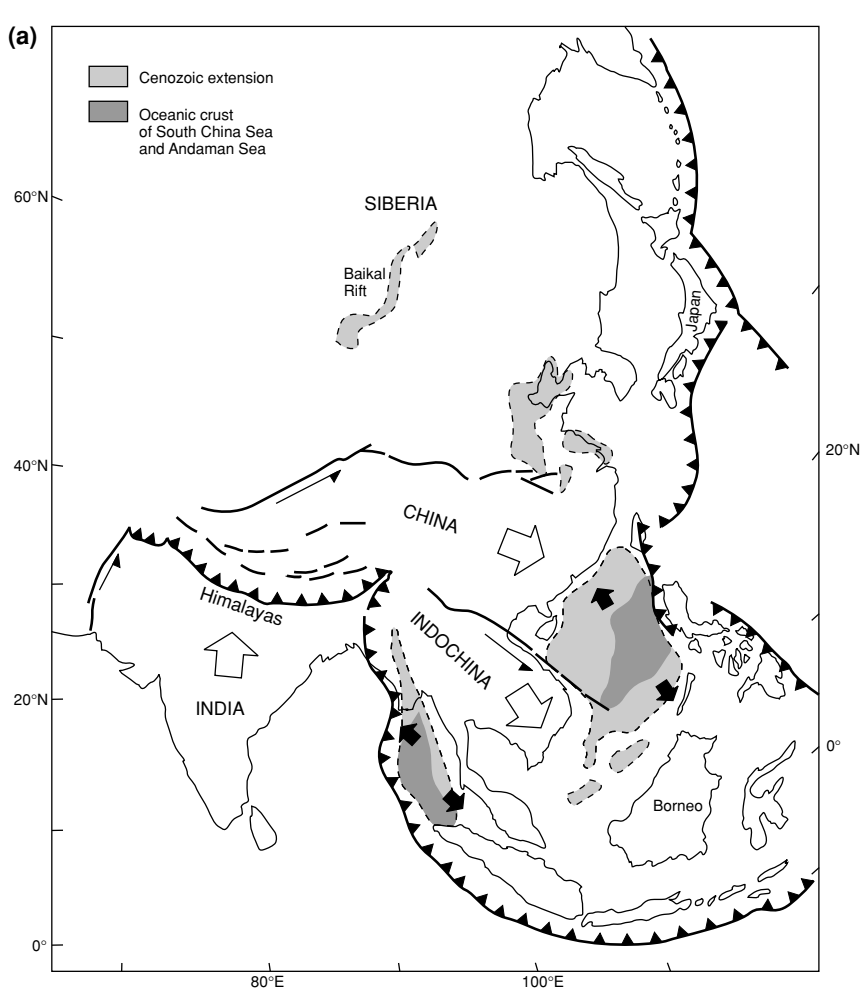
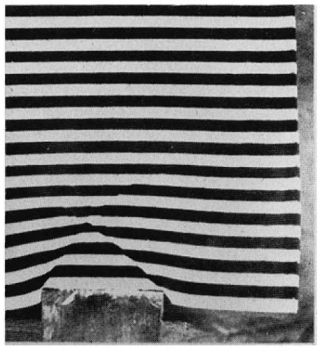
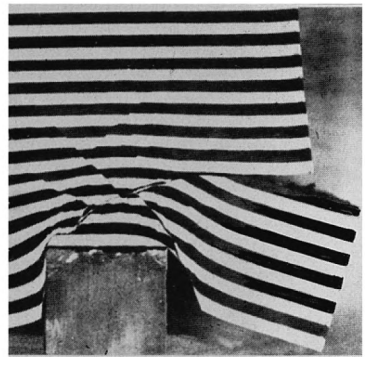


Figure 10.16. (b) Main faults and earthquake focal mechanisms for eastern Asia. Note the normal faulting north of the Himalayas in Tibet and compressive faulting in the Tien Shan north of the Tarim Basin. (After England and Molnar (1997).)







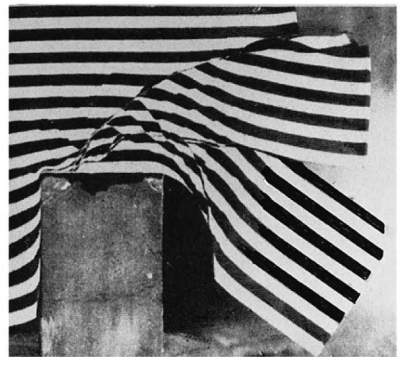


Figure 10.17. (a) The scheme of subduction zones, thrusts, large faults and Cenozoic extension in eastern Asia. Heavy lines, major plate boundaries of faults. White arrows, motion of Indian and the two major extruded blocks (China and Indochina) with respect to the Siberian block. Black arrows, direction of extension. In (b)–(d) are shown plan views of three successive stages in the indentation by a rigid indenter (India) into a striped block of plasticine (Asia). The plasticine was confined on only the left-hand side, leaving the right-hand side representing China and Indochina to deform freely. The resulting extrusion (d) of two large blocks to the right and the faulting and rifting ahead of the indenter are similar to the large-scale deformation, shown in (a). (After Tapponier *et al.* (1982).)

the seismic moment for large earthquakes over a century is $1.8\,\mathrm{cm\,yr^{-1}}$, while GPS measurements indicate that $1.8\,\mathrm{cm\,yr^{-1}}$ of contraction was occurring in Nepal between 1991 and 1995. The GPS data further imply that, south of the Higher Himalaya, the MHT was locked, while to the north a mid-crustal fault was moving at $0.2\,\mathrm{cm\,yr^{-1}}$. The remaining Indo-Asian convergence is taking place over a very large area north and east of the Himalayas, which is consistent with the extensive tectonic, seismic and local volcanic activity in these regions. The Tibetan plateau has anomalously high elevation and the presence of active normal faults indicates that it is extending. This may be visualized as similar to a blof of viscous fluid spreading out and thinning. Convergence across the Tian Shan is due to the clockwise rotation of the apparently rigid Tarim Basin. Figure 10.17(b) shows results from a plasticine model of Southeast Asia and the deformation that resulted when a rigid block (India) was pushed northwards into it. The large-scale internal deformation and eastward squeezing of regions appropriate for Tibet and China show up clearly.

There have been several major seismic experiments to determine the details of crustal and uppermost-mantle structure across the Himalayas and into Tibet. These have been international experiments conducted by American, Chinese and French institutions. Detailed seismic-reflection data from the Higher Himalaya and southern Tibet are shown in Fig. 10.18 and the crustal and lithospheric structure are shown in Fig. 10.13. The main feature of the crustal structure across the Himalayas is the major increase in depth of the Moho from 35–40 km beneath India, to $\sim$70 km beneath the Himalayas, a further increase to about 70–80 km beneath the southern Lhasa block, a decrease to 60–70 km beneath the northern part of the Lhasa block and a further decrease to less than 60 km beneath the Qaidam Basin. The reflections from the MHT beneath the Higher Himalaya show up clearly, as do the reflections from the Moho in the subducting Indian plate. There is a low-velocity zone in the crust to the north of the Zangbo suture. The Tibetan Plateau is a huge region at an elevation of >4.5 km, which suggests that the underlying crust and the compensation mechanism should be uniform for the entire plateau. However, this does not seem to be the case. Across the plateau, the crust has an average P-wave velocity of 6.1–6.3 $\mathrm{km\,s^{-1}}$, lower than is normal for continental crust ($6.45 \pm 0.21\,\mathrm{km\,s^{-1}}$), and a low-velocity zone may be present at mid-crustal levels. However, the crustal thickness decreases by 10–20 km on going from south to north; Poisson's ratio for the crust is much higher than normal in the north; there are zones of low S-wave velocity in the lower crust; seismic velocities for the upper mantle are low; and the upper mantle in the north is anisotropic and does not transmit S-waves well. Overall these results indicate that there are extensive regions of partial melting in the Tibetan crust and mantle and regions of lateral flow in the Tibetan uppermost mantle. The presence of widespread melt within the Tibetan crust is also supported by the facts that seismicity is generally shallower than 10 km, that there is a mid-crustal low-resistivity zone and that the Tibetan plateau is associated with a

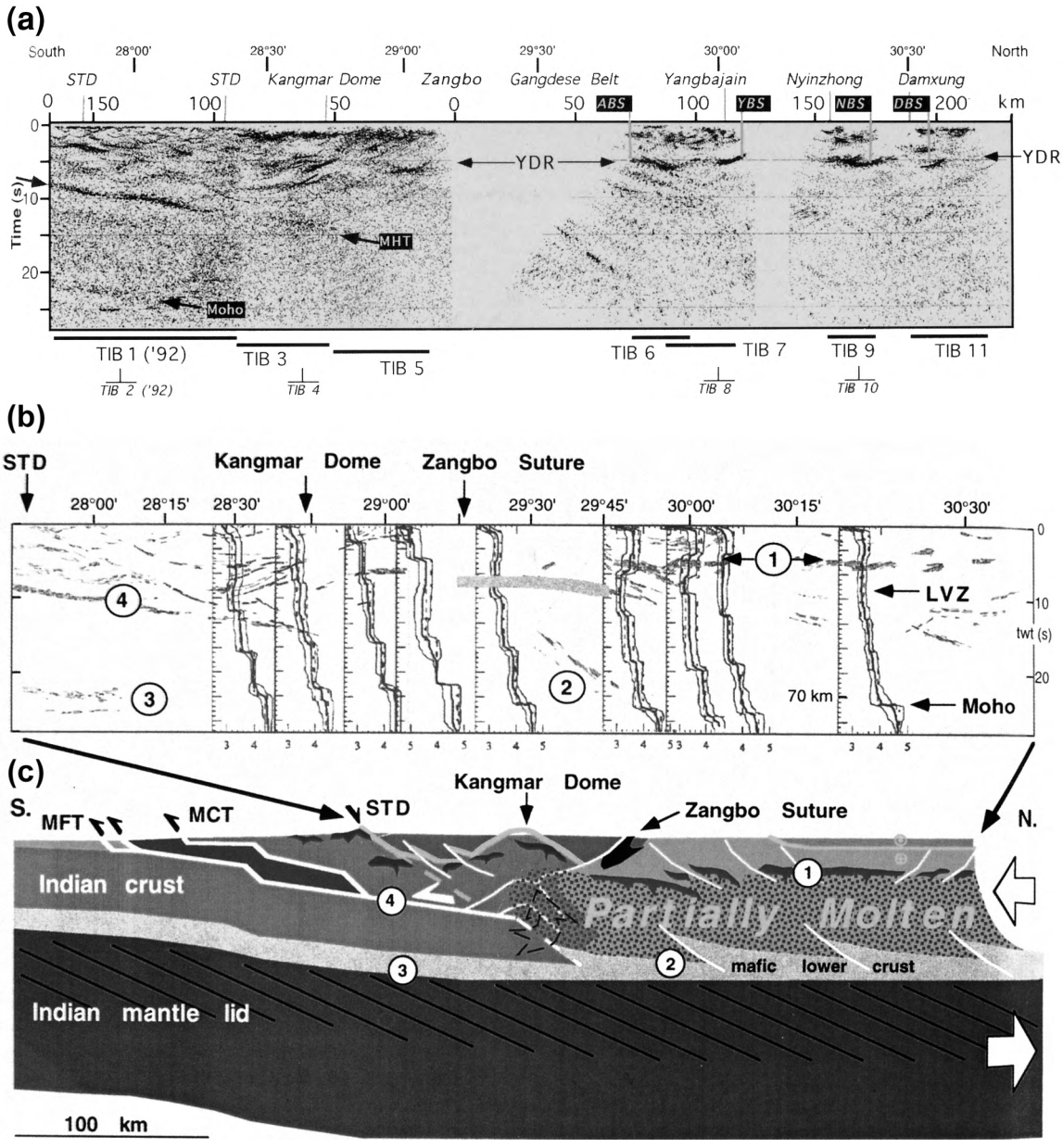


Figure 10.18. Sections across southern Tibet (along the Yadong–Gulu rift at 89–91°E). (a) Migrated reflection profiles (individual sections labelled TIB1–TIB11). Horizontal distances are centred on the Yarlung–Zangbo suture. MHT, main Himalayan thrust; Moho, the top of the subducting Indian mantle; a series of sub-horizontal reflections at ~9 s two-way time in the blank region to the north of the Zangbo suture are thought to be from an ophiolite nappe accretionary wedge; YDR (Yamdrok–Damxung) a series of sub-horizontal reflections extending to the northern end of the reflection profiles and including several high-amplitude, negative-polarity 'bright spots' indicative of solid–fluid contacts (Section 4.4.3). (b) A composite of seismic information: S-wave velocity models derived from waveform modelling of broad-band data (solid line with error bars); stipple, wide-angle reflection beneath and north of the Zangbo suture; background,

pronounced magnetic-anomaly low. Modelling of the magnetic data indicates that the magnetic susceptibility of the Tibetan crust is low. A reasonable interpretation is that the Curie-temperature isotherm (Section 3.1.3) is at a depth of ∼15 km beneath the plateau and that an intracrustal granitic melt is present below about 15 km depth over much of the plateau. Such a mid-crustal layer, if widespread, is consistent with an effective decoupling of the upper crust from the underlying lower crust and mantle. An estimate of the average viscosity of the Tibetan lithosphere is 10^{22} Pa s, only 10–100 times greater than that of the upper mantle.

There are different reasons for the presence of partial melt beneath southern and northern Tibet.

1. The presence of partial melt in the crust in the vicinity of the Zangbo suture results from the effective doubling of the crust: the increase in crustal thickness causes temperatures to rise (see Section 7.3 regarding calculation of geotherms). After a few tens of million years temperatures would have been sufficiently high for partial melting to take place in wet crust (i.e., temperatures exceed the wet solidus) and for granites to form. Hence, in time, a doubling of crust causes the development of a partially molten mid-crustal layer. Figure 10.18(c) suggests that temperatures in the partially molten mid-crust exceed ∼600 °C.
2. The widespread normal faulting and basaltic volcanism in northern Tibet started only ∼8–12 Ma ago, late in the accretionary tectonic history of the region (Fig. 10.14(b)). The distinctive potassium-rich composition of the volcanism indicates that the source was melted lithosphere rather than asthenosphere. A relatively sudden onset to such volcanism, combined with extension, can be explained were part of the lower lithosphere beneath Tibet suddenly removed. This would cause a rapid increase in temperature as the lower lithosphere was replaced by hotter asthenosphere. Another consequence of such a convective removal of the lower lithosphere would be an additional uplift (in excess of 1 km) of Tibet. An additional uplift would have enhanced Tibet's role as a major regulator on the climate of the Indian region and could account for the changes in the monsoon that occurred during the Miocene.

The entire Himalayan mountain chain is a region of large negative Bouguer gravity anomalies. Figure 10.19(a) shows the Bouguer gravity anomaly along profiles perpendicular to the Himalayas at 84–86°E. Also shown in Fig. 10.19(a) is the anomaly that has been calculated by assuming that the surface topography is locally isostatically compensated by crustal thickening (i.e., Airy's hypothesis – see Section 5.5.2). These calculated anomalies are different from those actually

←

reflection sections from (a). (c) A schematic interpretation of the India/Eurasia collision zone based on (a), (b) and structural information. MFT, main frontal thrust; MCT, main central trust; numbers 1–4 indicate structures giving rise to similarly labelled features in (b). Colour version Plate 28. (After Brown *et al.* (1996). Reprinted with permission from Nelson, K. D. *et al.*, Partially molten middle crust beneath southern Tibet: synthesis of project INDEPTH results, *Science*, **274**, 1684–7. Copyright 1996 AAAS.)

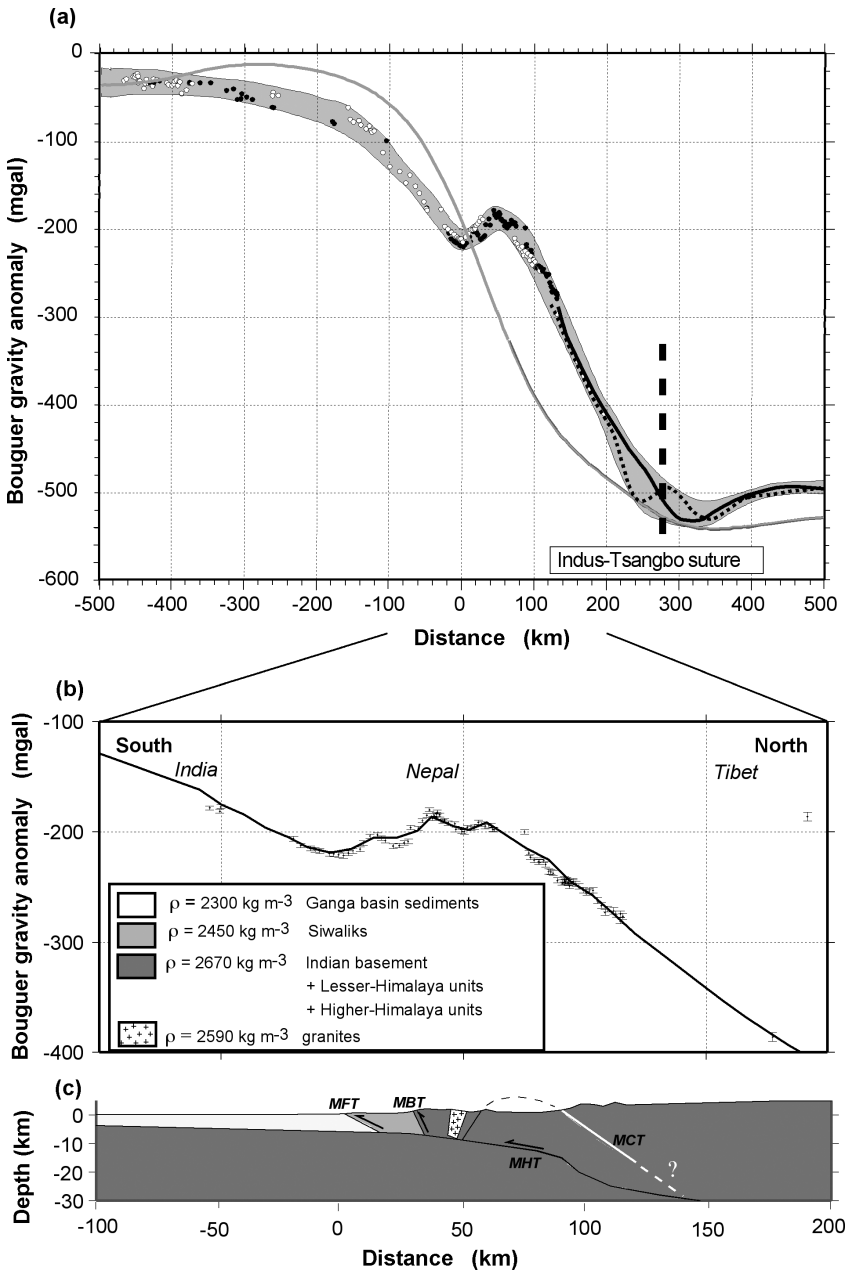


Figure 10.19. Profiles perpendicular to the Himalayas at ~84–86°E. Distances are in kilometres from the main frontal thrust fault (MFT). (a) The Bouguer gravity anomaly. Circles, data; grey shading shows the two-dimensional variation of data (black and white circles and solid and dashed black lines are from two profiles ~70 km apart); the grey line is the expected anomaly assuming local isostatic equilibrium. (b) Details of Bouguer gravity anomalies (with error bars) in the region of the MFT and MCT. Solid line, anomalies calculated from the density model. (c) The density model. (From Cattin *et al.* (2001).)

observed, being too small over the sedimentary Ganga (Ganges) foreland basin to the south of the mountains and too large over the mountains and beneath Tibet. This means that the Ganga Basin is over-compensated (there is a mass deficiency relative to the isostatic model) and that the Himalayas are under-compensated (there is a mass excess relative to the isostatic model) by as much as 100 mgal. A

model in which the Indian plate underthrusts the mountains, is flexed downwards by and supports the load of the mountains, while being heated, can account for these differences. The Ganga Basin, which forms in front of the Himalayas because the Indian plate is flexed downwards there, is filled with sediment eroded from the mountains. The complex jog in the gravity anomaly between 0 and 100 km is due to the lower-density foreland basin and sub-Himalaya sediments, which are underthrust beneath the higher-density Lesser Himalaya (Figs. 10.19(b) and (c)). The model shown in Fig. 10.13 has the Indian plate underthrust as an intact unit beneath the Himalayas and continuing beneath southern Tibet. Estimates of the effective elastic thickness of the Indian lithosphere decrease from south to north, reaching as little as ∼30 km beneath southern Tibet (see Section 5.7 for discussion of flexure and elastic thickness). This is consistent with the overall seismic and thermal picture of the crust and uppermost mantle beneath the Himalayas and Tibet. The first-order agreement between the Bouguer gravity anomaly and the isostatic anomaly to the north of the Indus–Tsangpo suture confirms that the crustal thickness beneath Tibet must be fairly uniform. However, the upwarping of the Bouguer anomaly there (∼30–40 mgal less than the isostatic anomaly) indicates that the region is under-compensated. This could be accounted for by the presence of additional mass at some depth (perhaps formation of eclogite in the Indian lower continental crust) or additional bending moments acting on the Indian plate.

The Alps

Although our understanding of the Himalayas is still incomplete and details will change as further work is undertaken, the Alps have been paid much more attention and are far more accessible. The Alps formed when the Adriatic promontory on the African plate collided with the southern margins of the Eurasian plate. The Alps were not the only mountains formed as a result of the convergence of Africa and Eurasia. Figure 10.20(a) shows the extensive Alpine fold system of the Mediterranean region. The complex present-day tectonics of the Mediterranean involves a number of microplates. The main rigid regions are Africa, Eurasia, Arabia, the Adriatic Sea, central Turkey and central Iran (Figs. 10.20(b) and (c)). Palaeomagnetic data indicate that the Adriatic block, which was a northern promontory of the African plate, has been separate from the African plate since the Cretaceous (estimates are ∼80–130 Ma) and is now rotating separately. The clockwise extrusion of the Anatolian block appears to be a consequence of the northward motions of the African and Arabian plates. The African plate is being subducted beneath Crete along the Hellenic arc and the back-arc Aegean region is undergoing intense localized deformation. The Gulf of Corinth is opening at about 1 cm yr^{-1}, which is about a quarter of the relative motion on the Hellenic arc. The Hellenic arc itself is retreating southwards. GPS and earthquake data show that the Anatolian plate is rotating anticlockwise and moving westwards at ∼2–3 cm yr^{-1} (Fig. 10.20(d)). GPS measurements yield an estimate for slip on the 1600-km-long right-lateral North Anatolian fault of 2–3 cm yr^{-1}, whereas

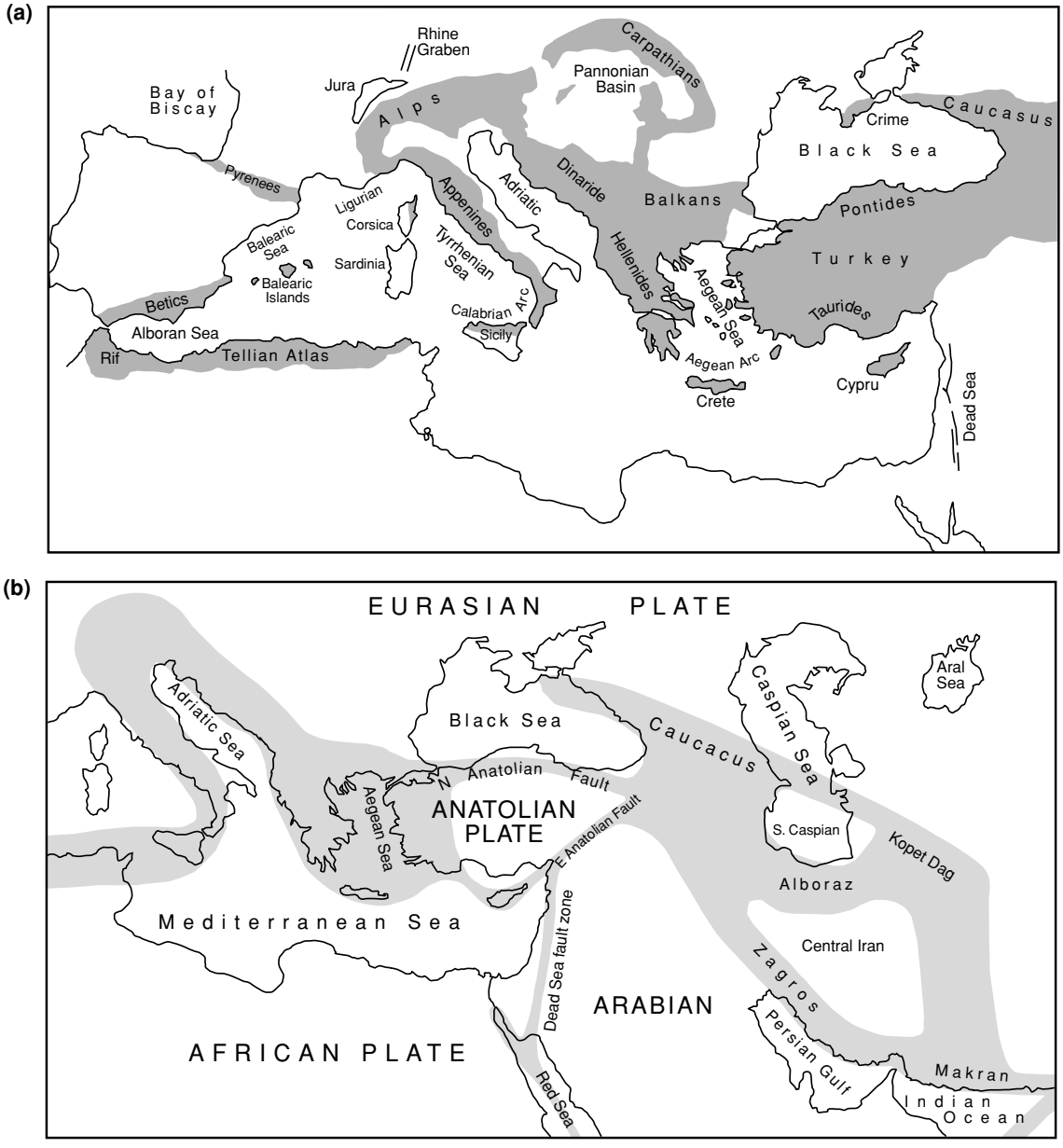


Figure 10.20. (a) The Alpine system of Europe (shaded) forme as a result of the convergence of Africa and Eurasia. (After Smith and Woodcock (1982).) (b) The principal seismic belts in the Mediterranean and Middle East. (After Jackson and McKenzie (1988).) (c) Fault-plane solutions for major earthquakes in the eastern Mediterranean region (1908–1999, including Izmit and Düzce 1999 events). (From Kahle *et al.* (2000).) The seismicity along the North Anatolian Fault suggests a westward migration pattern of large earthquakes. If this is correct, the next large earthquake might be under the Sea of Marmara and so could threaten Istanbul.

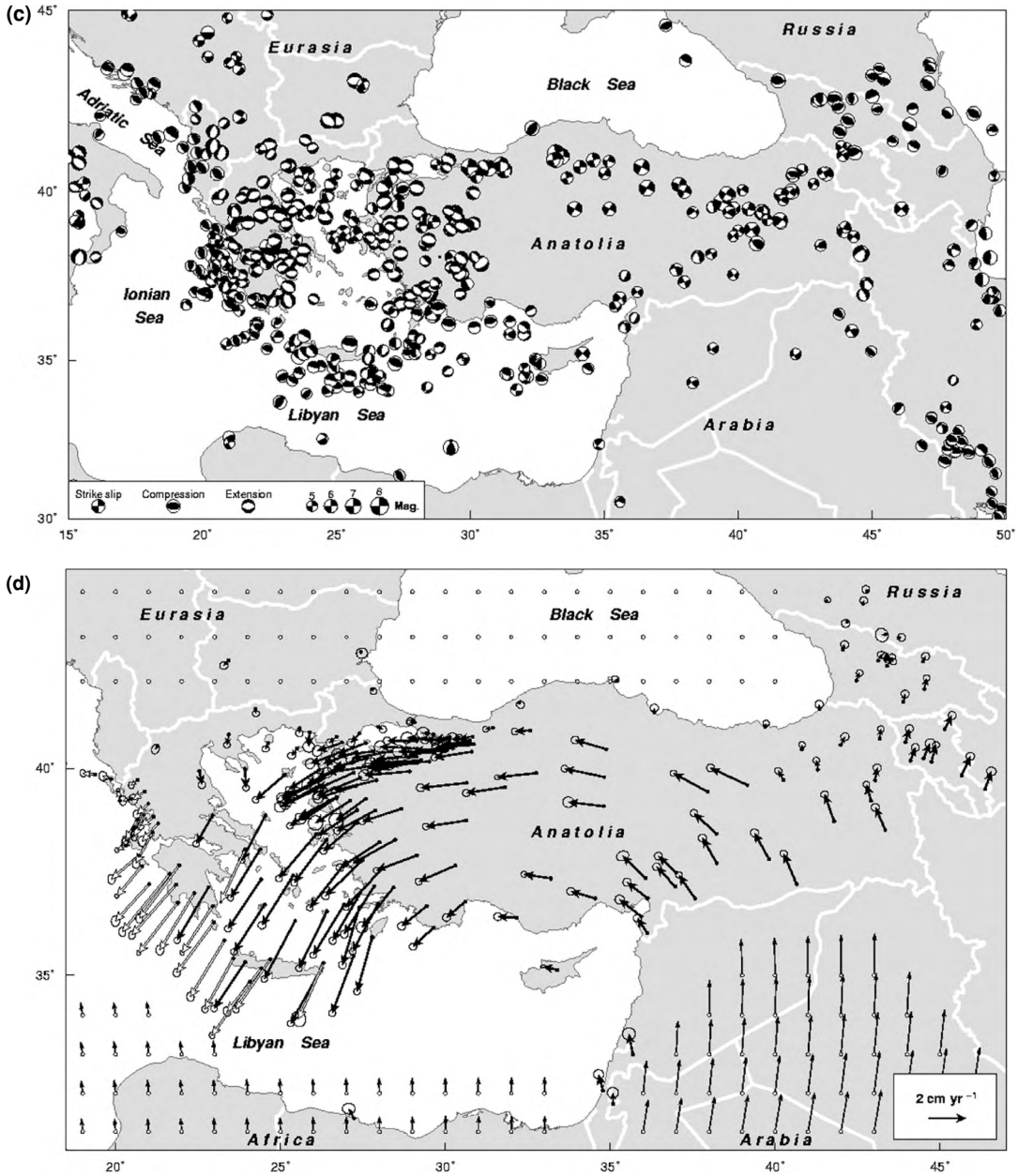


Figure 10.20. (*cont.*) (d) Velocity fields relative to Eurasia in the eastern Mediterranean region. Solid and open arrows, two separate sets of GPS results. Thin arrows on African and Arabian plates, velocities calculated using NUVEL-1A (Table 2.1). (From Kahle *et al.*, GPS-derived strain field rate within the boundary zones of the Eurasian, African and Arabian plates, *J. Geophys. Res.*, **105**, 23 353–70, 2000. Copyright 2000 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

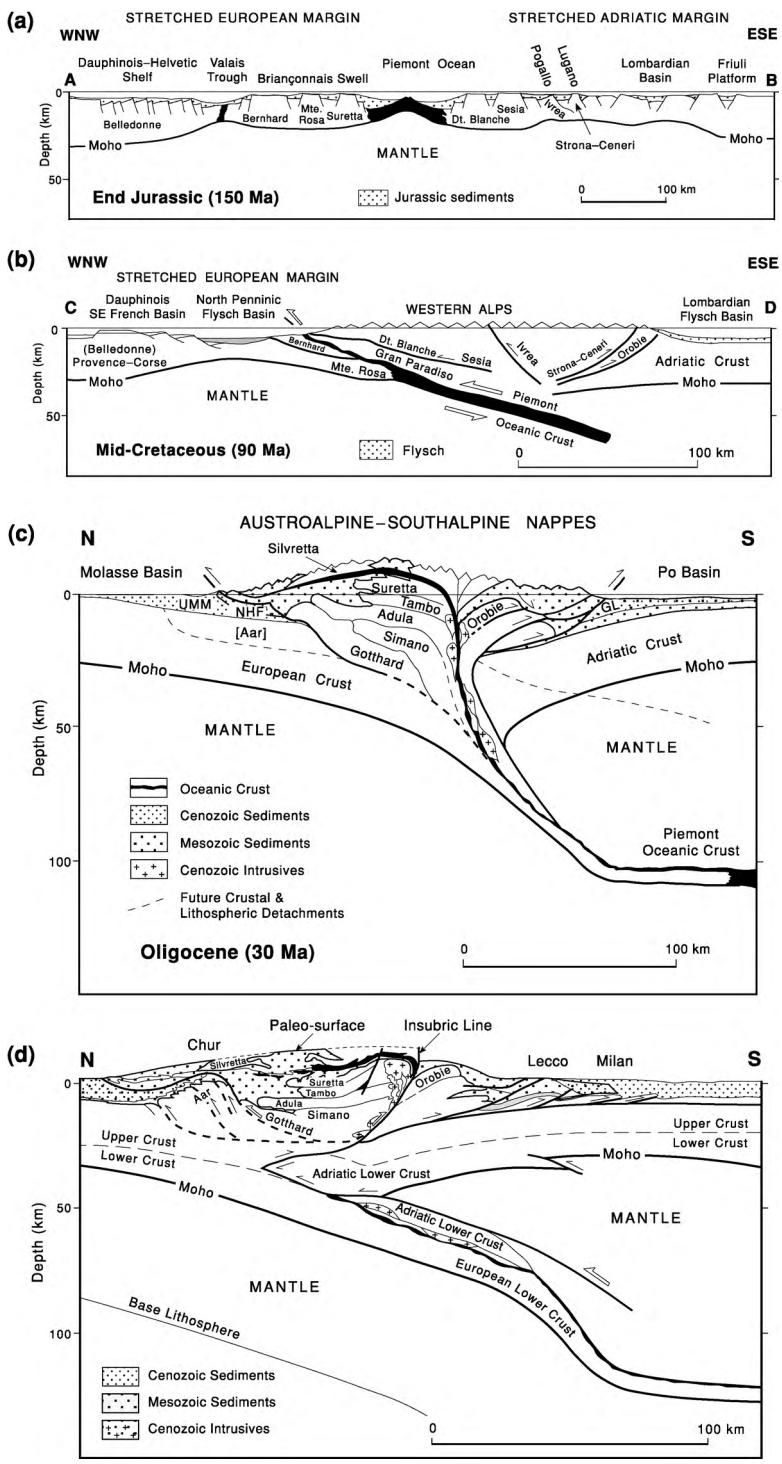


Figure 10.21. North–South sections along the European Geotraverse (EGT) at ~9°E across the Alps, showing the tectonic evolution of the mountain chain: (a) end of the Jurassic (150 Ma), (b) Mid-Cretaceous (90 Ma), (c) Oligocene (30 Ma) and (d) present-day (from Pfiffner (1992)).

motion on the left-lateral East Anatolian fault is $\sim$1.5 cm yr^{-1}. The compression occurring to the east in the Caucasus is also a consequence of the collision between Arabia and Eurasia. The strain there is much greater than can be accounted for by recorded earthquakes, indicating that considerable aseismic deformation is occurring.

An interpretation of the stages of the evolution of the Alps is shown in Fig. 10.21. During the Mesozoic extensional rift systems were operating in the initial crust and mantle of Pangea as Gondwana and Laurasia separated and the Tethys and central Atlantic Oceans formed (Fig. 3.30). By the end of the Jurassic major extension had taken place: the so-called Neo-Tethys or Piemont Ocean had formed and separated the rifted and thinned European and Adriatic continental margins (Fig. 10.21(a)). Estimates of the width of this ocean are 100–500 km. Some of this oceanic material may now be represented in the ophiolite sequences that occur along the length of the Alpine chain. The initial formation of the Alps resulted from northeast–southwest convergence between Europe and Africa. By the mid-Cretaceous (Fig. 10.21(b)) the Piemont oceanic lithosphere had been subducted and the European continental margin was being thrust beneath the Adriatic margin. The main episode of continental collision took place in the Tertiary with north–south convergence. This resulted in the major deformation, uplift and subsequent erosion which formed the Alps as we observe them today. Figure 10.21(c) shows the situation during the Oligocene (30 Ma) with the thinned European continental margin being delaminated – the European upper crust was peeled off and thrust northwards. At the same time, on the Adriatic margin the upper crustal layers were also being removed; they were thrust southwards. During the late Eocene a foreland basin developed to the north of the Alps. The exact method of loading and deformation and the peeling off and stacking of slices of crust from the colliding plates remain matters of research.

Figure 10.22 shows a series of sections across the Swiss Alps. Negative Bouguer gravity anomalies characterize the Alps, which is consistent with major crustal thickening. The details of the geology of the Alpine chain are very complex, but overall the geology is straightforward. The northern ranges are molasse (sediments) from the foreland sedimentary basin. The rocks then progressively age southwards, until finally the highly metamorphosed crystalline core is reached: these crystalline rocks were at deep levels in the crust until thrusting and erosion brought them to the surface. (See Section 7.8.4 on metamorphic belts.) The southern Alps, south of the Insubric Line, originated on the continental shelf of the Adriatic promontory of the African continent. It is estimated that some 100 km of shortening has occurred (by folding and thrusting) across the Alps during the last 40 Ma. In the western Alps, there is a large positive gravity anomaly caused by the *Ivrea body*, a slice of lower-crustal–upper-mantle material that was obducted from the southern (Adriatic) plate and thrust to a shallow level, in some places outcropping at the surface.

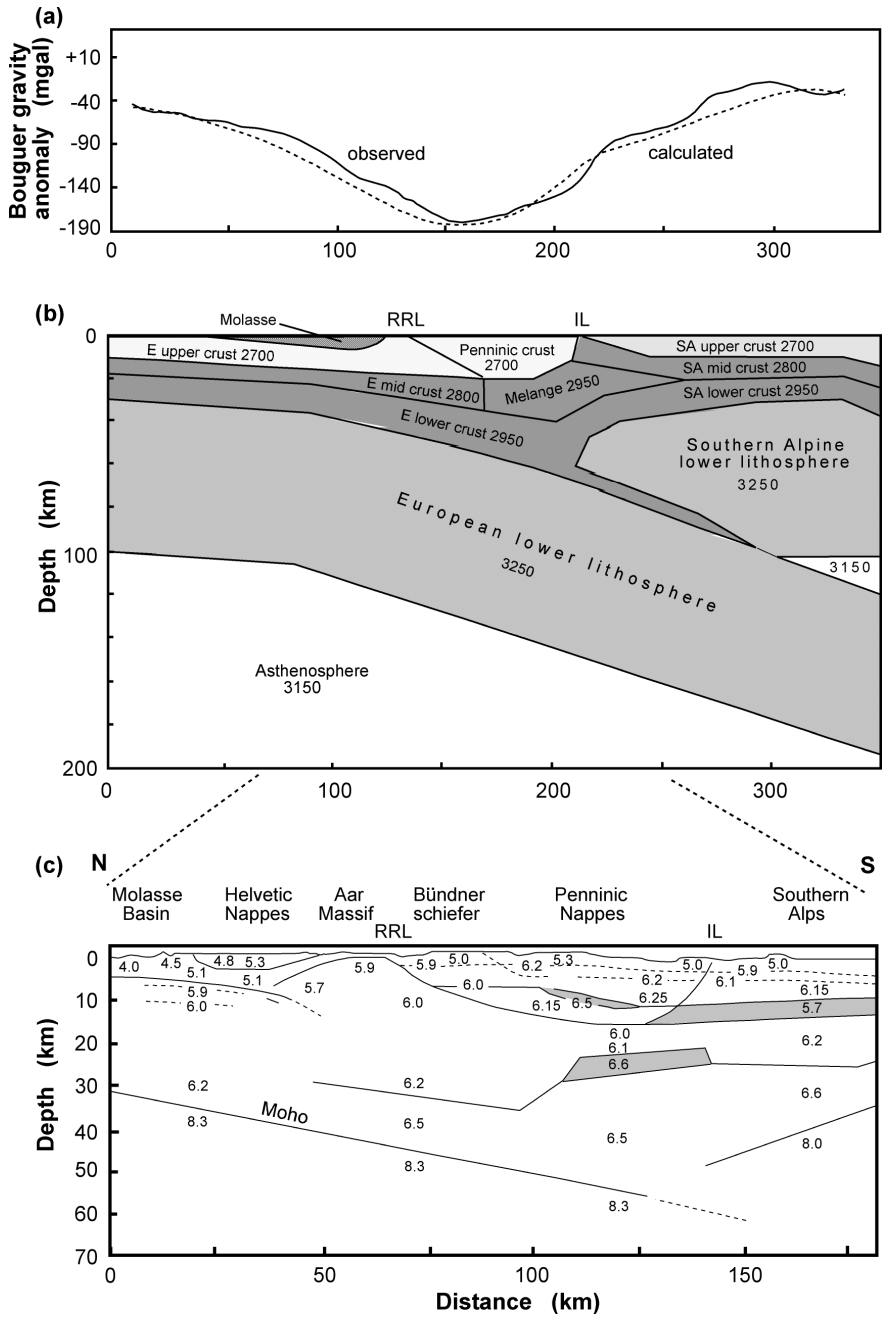


Figure 10.22. Cross sections across the Swiss Alps. (a) The Bouger gravity anomaly. (b) The density structure used to model the gravity data in (a). (c) The P-wave velocity structure determined from seismic-refraction profiles. Low- and high-velocity zones are shaded. Major faults: RRL, Rhine–Rhône Line; and IL, Insubric Line. (d) A reflection profile across the Alps. (From Pfiffner (1992) and Holliger and Kissling (1992).)

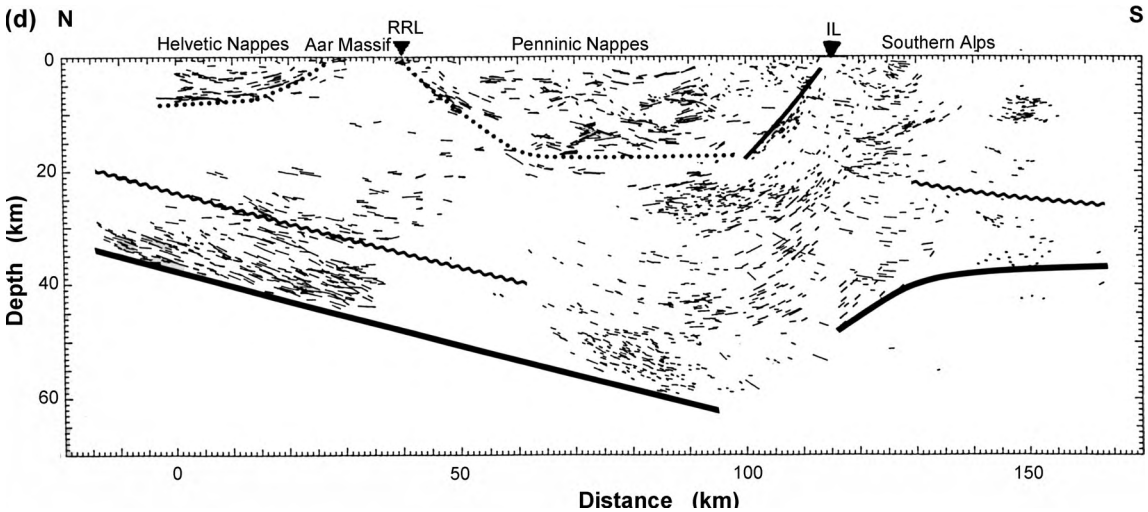


Figure 10.22. (*cont.*)

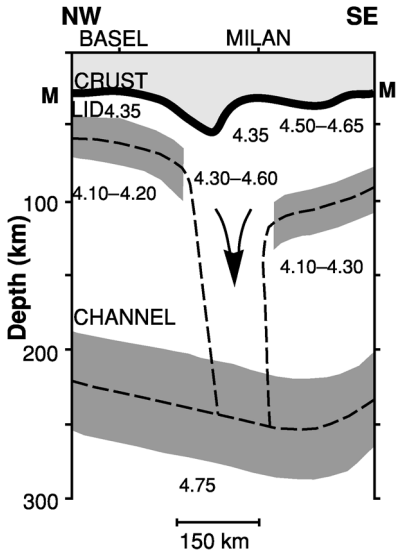


Figure 10.23. The shear-wave velocity structure of the upper mantle beneath the western Alps obtained from the simultaneous inversion of surface-wave-dispersion data. S-wave velocities are shown in km s^{-1}. M, Moho discontinuity; 'Lid', lower lithosphere; 'Channel', asthenosphere; dark shading, uncertainty in boundaries. High velocity beneath the Alps indicates that material may have been subducted to about 200 km depth. (From Panza and Mueller (1979).)

Figure 10.23 shows the shear-wave velocity structure of the upper mantle on a cross section through Switzerland and northern Italy obtained from inversion of surface-wave data. The lithosphere on the northern side of the Alps is somewhat thicker than that beneath Italy. The high shear-wave velocities, extending to ~200 km depth beneath the Alps, are typical of the lithosphere, rather than the asthenosphere, and suggest that lithosphere has been subducted to about that depth.

Many seismic-refraction lines have been shot in the Alps. The results have been used in the preparation of the Moho-depth map in Fig. 10.2: The crust thickens from 25 km beneath the Rhine graben to over 55 km in the Central Alps and then thins again on the southern side of the Alps. Figure 10.24(a) shows a seismic-refraction record section from the Jura region, north of the Alps. There the crust is 27 km thick and has a complex structure with two distinct low-velocity zones, one in the upper crust and the second immediately above the Moho. These low-velocity zones have been detected because of the offset in the travel times between the wave which travelled in the overlying high-velocity material and the wave reflected from the base of the low-velocity zone (see Section 4.4.3). One prominent feature of this record section is the large amplitude of the P_n phase (the Moho headwave), which indicates that there is a strong positive velocity gradient in the upper mantle. Figure 10.24(b), data from the southern Swiss/Italian Alps, shows evidence of a low-velocity zone in the upper crust but not in the lower crust. The large amplitude of the wide-angle reflection from the Moho, P_mP, indicates that there is a large velocity contrast at the Moho beneath the southern Alps.

The crust reaches a maximum thickness of 56 km beneath the Alps and is underlain by the southward-dipping European mantle. Just to the north of the Insubric Line there is a very sudden step offset in the Moho. To the south the Moho is much shallower and dips northwards. Beneath the Southern Alps the Moho is at 33 km depth. The bases of the European and the Adriatic crusts are imaged on the reflection line shown in Fig. 10.22(d). The depth extent of the European Moho and details of the European mantle lithosphere as a continuous interface are not clear. The Insubric Line itself is well imaged on the reflection profile and can be traced to a depth of 17 km.

The crust beneath the Alps is complex, both vertically and laterally along the length of the Alps, but it can be broadly described as upper crust and lower crust. Lower-crustal P-wave velocities are greater than 6.5 km s^{-1}, whereas upper-crustal P-wave velocities are less than 6.2 km s^{-1} (Fig. 10.22(c)). The upper crust is characterized by having a complex geometry and thin layers, but in part this is due to the greater resolution possible in the upper crust (Section 4.4.4). There is a pronounced low-velocity, 5.7 km s^{-1}, zone in the upper crust beneath the Southern Alps – this may be part of the southward-oriented thrust sheets (Fig. 10.21(d)). Beneath the Penninic nappes a thin high-velocity, 6.5 km s^{-1}, layer at 10 km depth can be matched to normal-incidence reflections from an interface within the Penninic nappes. The second high-velocity layer at ~20 km depth beneath the Penninic nappes is also clearly identifiable on the normal-incidence-reflection profile. This reflection horizon may be the top of the Adriatic lower crust.

The Alps are not particularly active seismically. Earthquakes do occur but not frequently, and, although sometimes damaging, they are usually of lower magnitude than Himalayan events. To the north of the Alps and beneath the Southern Alps, earthquakes occur throughout the crust. Beneath the central Alps, however, earthquake activity is restricted to the upper 15–20 km of the crust. This

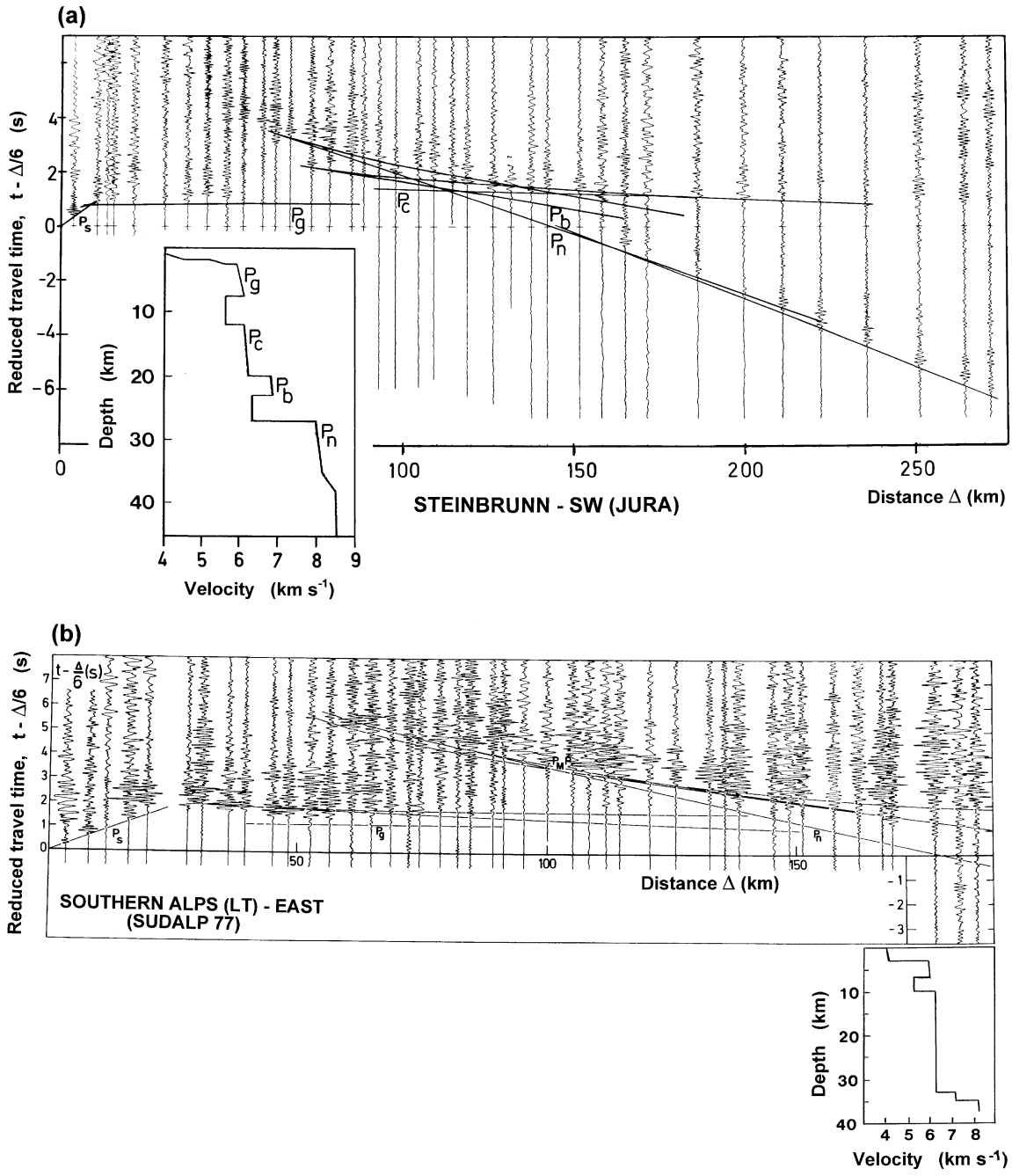


Figure 10.24. (a) A record section, reduced to 6 km s^{-1}, for a refraction line shot in the Jura in the northern part of the cross section shown in Fig. 10.22(b). The time offset between the crustal phases P_g and P_c indicates the presence of a low-velocity zone. Likewise, the low-velocity zone at the base of the crust is indicated by the time offset between phase P_b and the Moho reflection. (b) A record section, reduced to 6 km s^{-1}, for a refraction line shot in the southern Alps perpendicular to the cross section shown in Fig. 10.22(b). The postcritical Moho reflection P_mP is very strong on these records. (From Mueller *et al.* (1980).)

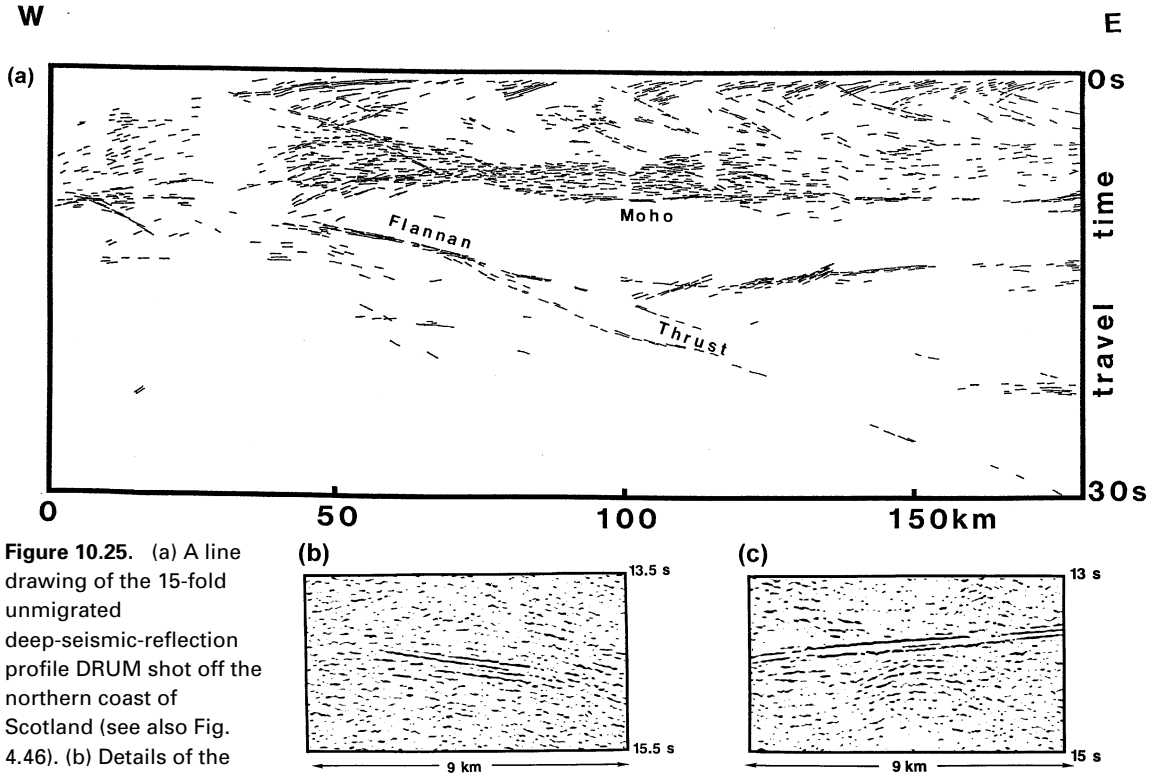


Figure 10.25. (a) A line drawing of the 15-fold unmigrated deep-seismic-reflection profile DRUM shot off the northern coast of Scotland (see also Fig. 4.46). (b) Details of the reflections from the Flannan Thrust. (c) Details of the reflections from the sub-horizontal deep-mantle reflector. (From McGeary and Warner (1985) and Warner and McGeary (1987).)

supports the implication that there is a major detachment surface at this level. The maximum uplift in the Alps is 0.15 cm yr^{-1}, almost an order of magnitude less than Himalayan values.

Ancient continental collisions

Interpreting ancient continental collision zones is a complex geological problem – tectonics and erosion mean that only parts of the jigsaw remain for study.

The *Caledonian orogeny* occurred some 400 Ma ago when the ancient Iapetus Ocean between North America and Europe closed during the formation of the supercontinent Pangea (Fig. 3.30). The remnants of this collision are now in Scotland and eastern North America. The Flannan Thrust off the northern coast of Scotland has been spectacularly imaged by deep-seismic-reflection profiling (Fig. 10.25). This thrust originates in the lower crust, cuts (and may offset) the Moho and extends to a depth of 75–85 km. The upper crust is characterized by rotated half-grabens filled with sediment, which formed during a later period of Mesozoic extension. The lower crust is highly reflective, with the Moho clearly visible as a bright reflector at its base. The dipping crustal reflector which

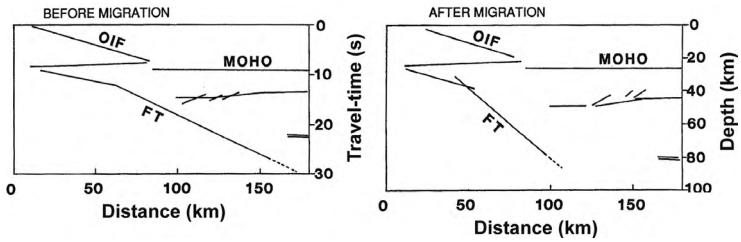


Figure 10.26. The effect of depth migration on the deep reflections of Fig. 10.25(a). The dipping reflectors steepen and migrate up-dip. OIF, Outer Isles Fault; and FT, Flannan Thrust. (From McGeary and Warner (1985).)

is visible between 30 and 80 km is the Outer Isles Fault (see also Fig. 4.57). Two sets of clear, strong reflections originate within the mantle: the first, a 100-km-long sub-horizontal reflector at two-way time 13–15 s, and the second, the dipping Flannan Thrust reflector which extends from 7 s down to at least 27 and possibly 30 s two-way time (the recording time of the survey was 30 s). Figure 10.26 shows the effect of depth migration (see Section 4.5.4) on these deep reflectors: the dipping reflectors steepen and migrate up-dip. The most plausible explanation for the Flannan Thrust reflector is that it is a fossil subduction zone dating from the Caledonian orogeny when the Iapetus Ocean closed. If so, it is a 400-Ma-old thrust, though it could have been reactivated by the later Mesozoic extension in the region. Nevertheless, such strong reflections, which clearly originate from within the lower part of the lithosphere, show that the lower lithosphere can be structurally complex and can support localized strains over a long time.

The western part of North America has been a continent–ocean boundary for some 700–800 Ma (Fig. 3.30). Evidence from the series of accretionary complexes and magmatic arcs that make up the over-600-km-wide northern Cordilleran orogen show that it has been a convergent boundary since the Devonian. Since the early Jurassic the North American plate has moved over a series of oceanic plates and has accreted subduction-zone terranes and intra-oceanic arcs. Magmatic arcs have then been emplaced on and in these accreted terranes. Over this time the west coast of North America may have consumed a region wider than the present-day Pacific Ocean. Extensive programmes of deep-seismic-reflection profiling and associated geological and geophysical work have imaged details of the structure of the crust and uppermost mantle along this margin (see http://www.litho.ucalgary.ca/atlas/index1.html for the Canadian Lithoprobe data).

The core of northern and western North America is made up of a series of Archaean cratons that were assembled during the Proterozoic (Fig. 10.27). The *Trans-Hudson orogen*, now about 500 km wide and exposed in Canada, is what remains of a major Himalayan-scale continental collision. It is part of the 1.75–1.85-Ga collision zone that extends northwards from the central U.S.A. through Canada and then eastwards across Hudson Bay to southern Greenland and on into northern Scandinavia and Russia. The orogen consists of four distinct

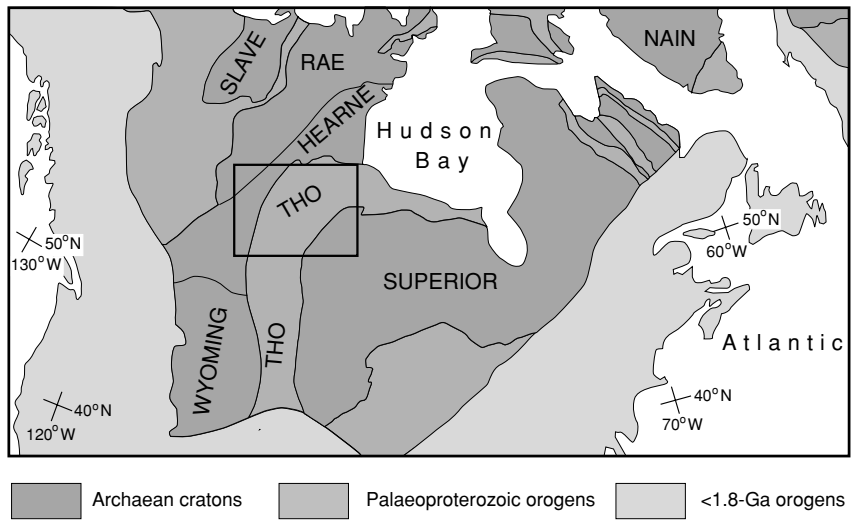


Figure 10.27. Locations of the Archaean provinces of North America and the Trans-Hudson orogen (THO). The orogen extends across Hudson Bay, Ungava and across Greenland to Scandinavia. Greenland and western Scandinavia are shown in the positions that they would have had prior to the opening of the Atlantic Ocean. The box indicates the location of the 'Lithoprobe' work shown in Fig. 10.28. (After Lucas *et al.* (1993) and Hoffman (1989).)

zones, which include subducted continental margins, obducted oceanic crust and accreted juvenile arc terrains, sandwiched between Archaean cratons: the Superior craton to the south and east and the Rae/Hearne craton to the north and west. The four parallel zones, in order from the east, are (1) a foreland belt, (2) a wide internal zone, (3) remnants of a magmatic arc and (4) a fold belt.

1. The narrow foreland belt (Thompson), the boundary zone along the eastern edge of the orogen, consists of heavily reworked rocks of the Superior craton as well as sediments associated with rifting and igneous rocks of unknown origin. At the end of the collision the Superior craton was ultimately thrust westwards over the previously deformed central part of the orogen.
2. The wide internal zone (Reindeer) comprises oceanic and arc volcanic and plutonic rocks together with sediments, all thought to have been formed along a subduction zone. This zone is structurally underlain by an Archaean micro-continent – small structural windows expose Archaean rocks.
3. The remnants of the magmatic arc (Wathaman) are exposed as a 1.85-Ga batholith, which is up to 100 km wide and extends some 800 km along the length of the orogen. It represents the eroded interior of a major calc-alkaline intrusion similar to those in the Andes. The presence of this magmatic arc along the northwest margin of the orogen confirms that oceanic lithosphere was subducted beneath the Hearne craton.
4. Finally, the fold belt (Cree Lake) comprises deformed and metamorphosed continental margin sediments as well as 1.8-Ga intrusions.

The North American Central Plains (NACP) conductivity anomaly, the largest electrical conductor delineated by electromagnetic induction studies, is collinear with the Trans-Hudson orogen and is thought to be caused by rocks within the

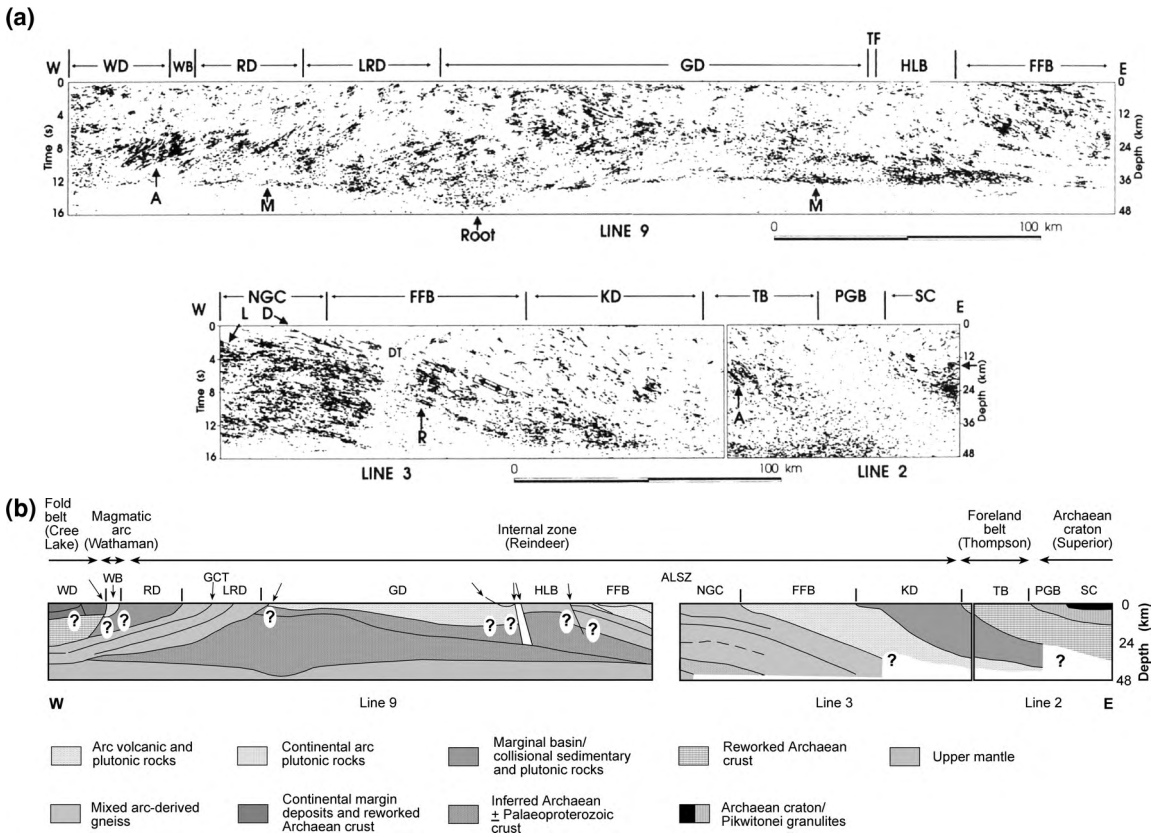


Figure 10.28. (a) 'Lithoprobe' seismic-reflection data right across the Trans-Hudson orogen on an east–west profile located in the south of the box in Fig. 10.27. Data have been migrated, coherency-filtered and displayed using an automated routine that displays the main reflections. The section is to true scale assuming a constant migration velocity of 6 km s^{-1}. Note that Line 9 and Line 3 are offset by ~30 km north–south. 'M', reflection Moho; and 'Root', crustal root. (b) A geological interpretation of (a). Main geological designations: WD, Wollaston domain; WB, Wathaman–Chipewyan batholith; RD Rottenstone domain; LRD La Ronge belt; GD, Glennie domain; HLB, Hanson Lake block; FFB, Flin Flon belt; NGC, Namew gneiss complex KD, Kisseynew domain; TB, Thompson belt; PGB Pikwitonei granulite domain; and SC Superior craton. Arrows show major shear zones, faults and thrusts. Reprinted with permission from *Nature* (Lucas *et al.*, *Nature*, **363**, 339–42. Copyright 1993. Macmillan Magazines Ltd.)

western part of the Reindeer zone. Results from a series of seismic experiments across the orogen in Canada (Lithoprobe) and in the U.S.A. (COCORP) have delineated the deep-crustal and uppermost-mantle structures beneath this orogen (Fig. 10.28). Tectonic activity took place over a region ~1000 km in width. The crust is thicker than normal beneath the orogen (40–54 km), with the thickest crust

being beneath the Archaean micro-continent. The crust is highly reflective and the Moho is very well defined beneath the orogen. The westward-dipping crustal reflections in the west of the orogen are consistent with the geological evidence that subduction took place beneath the Hearne craton and indicate that the arc-derived rocks extend beneath the continental margin and Archaean rocks. There is no apparent connection at depth between the central Archaean micro-continent and the Hearne and Superior cratons to either side – this is consistent with the micro-continent being exotic. The easterly dipping reflections in the east of the orogen seem to be related to the late/post-collisional oblique deformation as the Superior craton was thrust up and over and along the previously deformed orogen and do not imply that there was easterly directed subduction. In addition there is evidence for mantle structures – a high-P-wave-velocity ($8.4\,km\,s^{-1}$) zone in the mantle. This zone is at depths <100 km and dips westwards. Such high P-wave velocities suggest that the material could be a remnant eclogitic slab.

Ultra-high-pressure belts

Ultra-high-pressure (UHP) metamorphic belts are unusual metamorphic complexes that contain relics of low-temperature, high-pressure assemblages (including coesite and diamond). These very rare belts are associated with collisional mountain belts and contain <10-km-thick UHP terranes of partly exhumed and metamorphosed continental crust, tectonically juxtaposed against other, very different continental rocks. UHP belts are found in ancient continental collision zones across Eurasia, which range in age from 300 Ma to ~2.3 Ga (e.g., the Dora Maira massif, western Alps; Western Gneiss Region, Norway; Kokchetav Complex, Kazakhstan; Maksyutov Complex, Urals; Qinling–Dabie–Sulu belt, China). The wide spread of ages implies that the processes that resulted in the burial and subsequent exhumation of these rocks were not unusual. However, their rarity indicates that some aspects of the process must have been very uncommon. The temperature of recrystallization of these UHP rocks was 700–900 °C at pressures of 2.8–4.0 GPa (approximately 90–125 km). Exhumation rates were fast – rates are estimated at 2–12 mm per year for ascent from >100 km depth to upper/mid-crustal levels. The process responsible both for the descent of the UHP rocks to >100 km and for their ascent is apparently subduction. The problems associated with an understanding of the occurrence of UHP rocks do not concern their subduction, but rather concern (1) the way in which they have been returned fairly rapidly to shallow levels and (2) the preservation of the relic UHP metamorphism. It is probable that UHP rocks arise when old, cold continental crust is part of a largely oceanic plate that is being subducted. At depth, thin sheets of continental crust then become detached from the rest of the subducting plate as large amounts of continental crust enter the subduction zone (Fig. 10.29). This may occur when convergence ceases and the continents suture. Then, being less dense than the rest of the lithosphere, these sheets, or tectonic slices, can rise

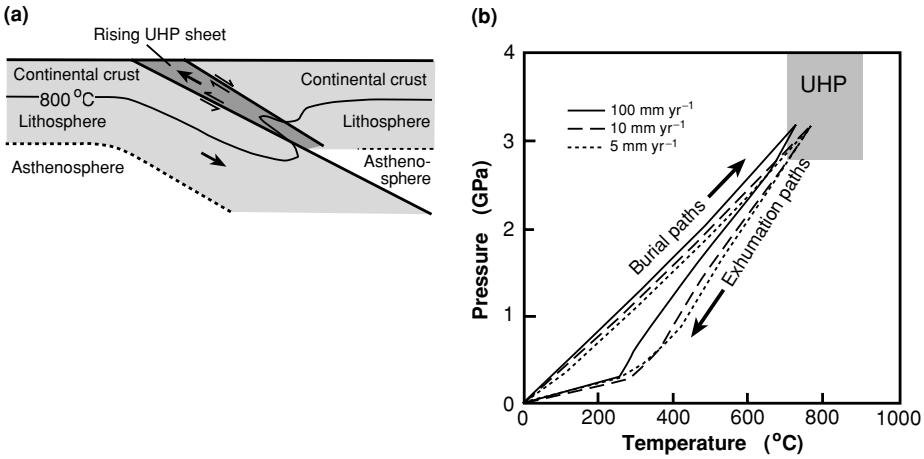


Figure 10.29. (a) A schematic cross section showing a rising UHP sheet during the final stages of continent–continent collision. The thin sheet becomes detached from the subducting plate and rises back up the subduction zone. (b) Burial and exhumation paths for a rock that was subducted to 100 km depth and then immediately exhumed along the same fault (now reactivated as a normal fault). *P–T* paths are shown for three convergence and exhumation rates, 100, 10 and 5 mm yr^{-1}. The shaded region indicates the ultra-high-pressure (UHP) region. (After Ernst and Peacock, A thermotectonic model for preservation of ultra high-pressure phases in metamorphosed continental crust, Geophysical Monograph 96, 171–8, 1996. Copyright 1996 American Geophysical Union. Modified by permission of American Geophysical Union.)

if their upward buoyancy forces exceed the downward forces (shear at the base of the sheet and resistance against the overriding plate). The differences in density can be considerable – the density of basalt eclogite is ~3700 kg m^{-3} and that of granitic gneiss is ~3000 kg m^{-3}. Section 7.8 shows that, when large-scale exhumation returns a thick layer of rock to shallow levels, the temperature of the rising rocks will be adiabatically controlled – it will remain high (Fig. 7.23). In such a situation any relic UHP metamorphism is unlikely to be retained since the minerals will re-equilibrate. However, in the situation in which sheets of UHP material rise by moving back up the subduction zone, the sheets can be cooled by conduction both from their upper and from their lower surfaces, to the overlying plate and to the underlying subducting plate (Fig. 10.29). If the detached sheet is thin, cooling could be efficient enough that complete re-equilibration upon ascent would not take place and for some relic UHP metamorphic mineralization to be retained. Reaction rates are generally reduced in the absence of aqueous fluids, which suggests that ascending UHP rocks may also have to have been dry. Thus, UHP belts are very narrow linear tectonic features that can form only under very rare specific conditions during the final stages of collision between two continents.

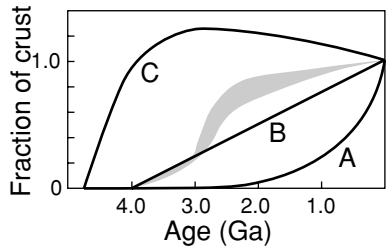


Figure 10.30. Growth curves for the continental crust. Model A has an increasing growth rate with time. In model B the growth rate is linear. Model C has an initially rapid growth rate. The generally accepted range of models is stippled but some influential geologists prefer models closer to C than to B.

10.2.4 Continental growth rates

The growth of continents and the depletion of mantle (see Chapter 6 for mantle-growth curves, etc.) are very much linked; the variation of initial isotope ratios with time illustrates this fact. So, in trying to estimate the growth rate of continental crust, we are also trying to estimate the depletion rate of the mantle. There are many models of crustal growth rates, which are summarized in Fig. 10.30. At one extreme is model A, in which the rate of formation of continental crust increases with time; in this model, loss of sediments into the mantle at subduction zones is of minor importance. Model B has a linear rate of growth of continental crust with time; this model also assumes that all sediments are returned to the continents at subduction zones. Model C represents a rapid growth of crust early in the history of the Earth, with the volume of crust staying more or less constant since then. It is possible to have a constant crustal volume if the sediments which are eroded from the continents and returned to the mantle at subduction zones have the same volume as the material which melts and is added to the continents above subduction zones. When the volume of volcanic material is less than that of the sediments, then the crustal volume decreases with time (as in model C).

Table 10.2 shows that some 70% of the present surficial area of the continents formed during the period which ended 450 Ma ago. Growth model A was suggested by the map of the world from the areas of crust exposed today and their ages. But crust is covered, reworked or destroyed. Most past crust is lost or hidden. Thus, the increasing rate of crustal growth depicted in model A is rather unlikely.

Most popular models for crustal growth lie between model B and model C, with many isotope geochemists favouring models that are closer to B than to C. The stippled region in Fig. 10.30 shows a range of models that assume a growth rate that increased from about 3.0 to 2.5 Ga to account for the large volume of Archaean volcanic rocks formed during that period (see Section 10.5). Such models provide a reasonable estimate of the growth of continental crust and are also consistent with chemical and thermal models. However, some geologists working on early crust strongly favour models in which early continental crust was abundant, more like C than B.

10.3 Sedimentary basins and continental margins

10.3.1 Introduction

Sedimentary basins are of great economic importance because fossil fuels were formed there. Knowledge of the origin of a basin can provide information about the location and the type of thermal maturation of hydrocarbon deposits. A number of tectonic mechanisms can cause the formation of sedimentary basins. These are discussed in the following sections.

10.3.2 The loading effect of sediments

For a sedimentary basin to form, some factor must cause the basement to subside. However, let us initially suppose that a subaqueous depression of depth d_w in the basement exists (Fig. 10.31). Because the Earth has rain and winds, erosion of surrounding terrain automatically fills this depression with sediments. Assuming that it is completely filled with sediments and that isostatic equilibrium is maintained, we can use the methods of Section 5.5.2 to calculate the total thickness of these sediments. This thickness d_s is calculated by equating the mass of the water in the depression plus the mass of displaced mantle to the mass of the sediments:

$$d_w\rho_w + (d_s - d_w)\rho_m = d_s\rho_s$$

which, on rearranging terms, gives

$$d_s = d_w\left(\frac{\rho_m - \rho_w}{\rho_m - \rho_s}\right) \tag{10.1}$$

where ρ_m is the density of the mantle, ρ_w the density of water and ρ_s the density of the sediments. If we take ρ_m as $3.3 \times 10^3\,\mathrm{kg\,m^{-3}}$, ρ_s as $2.5 \times 10^3\,\mathrm{kg\,m^{-3}}$ and ρ_w as $1.0 \times 10^3\,\mathrm{kg\,m^{-3}}$, the presence of the sediments amplifies the original subaqueous depression by a factor d_s/d_w of 2.9. Thus, a basin filled with 5 km of sediments represents a subaqueous depression of the crust of about 1.5–2.0 km, the exact value depending on the density of the infilling sediments. To make comparisons between one sedimentary basin and another, we must first calculate the loading effect of the sediments so that this amplification can be removed. This process is called *backstripping*. The depth at which the basement would have been if there had been no infilling sediments, d_w, is given by Eq. (10.1) as

$$d_w = d_s\left(\frac{\rho_m - \rho_s}{\rho_m - \rho_w}\right) \tag{10.2}$$

However, suppose that the sediments were deposited in water of depth h_w instead of at sea level (Fig. 10.31). In this case, Eq. (10.2) should be written as

$$d_w = h_w + d_s\left(\frac{\rho_m - \rho_s}{\rho_m - \rho_w}\right) \tag{10.3}$$

Sea level has not remained constant throughout geological history but has risen and fallen many times. Figure 10.32 illustrates one model of the variation of

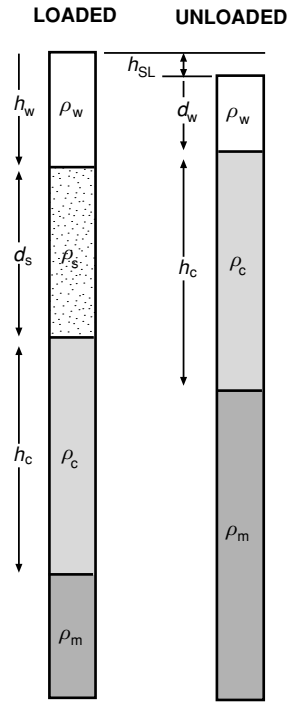


Figure 10.31. The loading effect of sediments. ρ_w is the density of water; ρ_s, the density of sediment; ρ_c, the density of crust; and ρ_m, the density of mantle. h_w is the depth of water in which sediments were deposited; d_s, the thickness of sediments; h_c, the thickness of crust; d, the backstripped depth of basement; and h_{SL}, the change in sea level from the time at which sediments were deposited to the present day.

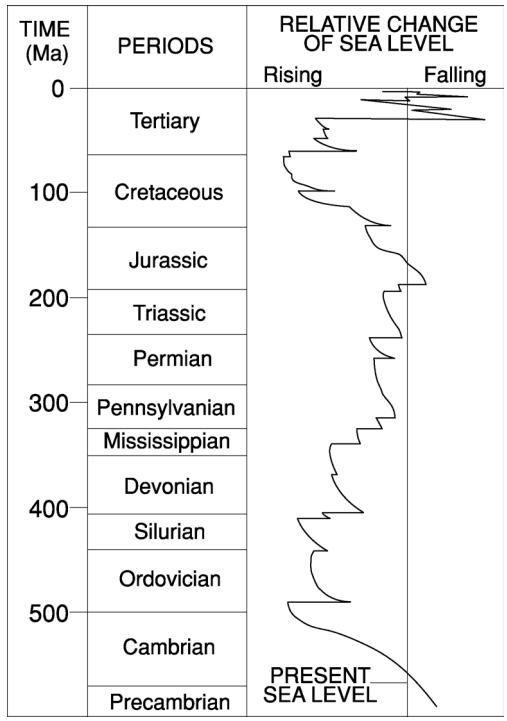


Figure 10.32. Changes in sea level during the Phanerozoic. (After Vail and Mitchum (1978), reprinted by permission.)

Phanerozoic sea level. To take account of this variation when we backstrip subsidence data, we must further modify Eq. (10.3) to

$$d_{\mathrm{w}} = h_{\mathrm{w}} + d_{\mathrm{s}}\left(\frac{\rho_{\mathrm{m}} - \rho_{\mathrm{s}}}{\rho_{\mathrm{m}} - \rho_{\mathrm{w}}}\right) - h_{\mathrm{SL}}\left(\frac{\rho_{\mathrm{m}}}{\rho_{\mathrm{m}} - \rho_{\mathrm{w}}}\right) \tag{10.4}$$

where h_{SL} is the change in sea level as shown in Fig. 10.32. In any detailed study of a region, another factor that should be taken into account is the postdepositional compaction of sediments. As more sediments are deposited, pore water is expelled and the thickness of each sedimentary layer reduced. In Eqs. (10.1)–(10.4), d_{s} represents the thickness of sediments when they were laid down and ρ_{s} their saturated density, not the thickness and density measured *in situ* today. *Decompaction* is the process of calculating the original sediment thickness.

Equations (10.1)–(10.4) assume Airy-type local compensation. However, if compensation is achieved not locally but regionally, the loading effect of the sediments should be calculated by considering the flexural loading of a thin elastic or viscoelastic plate. This is much more complex mathematically. In such models, because the load of the sediments is spread over a broad area, the net result is that the amplification factor is somewhat reduced.

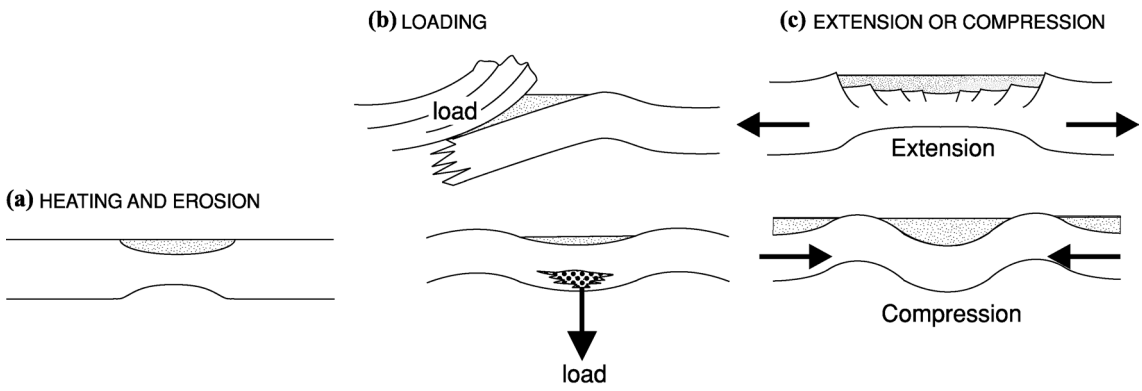


Figure 10.33. Some mechanisms that can give rise to sedimentary basins. (a) Heating of the lithosphere, which results in expansion and uplift. If erosion takes place, a basin forms when the lithosphere cools and contracts. (b) Loading of the lithosphere from above (thrust piles) or below (intrusions, phase changes). (c) Extension, compression and faulting of the lithosphere.

10.3.3 Mechanisms for producing subsidence

A variety of geological processes can lead to the formation of thick sedimentary deposits on continental crust. They can, however, be divided into three main classes: (1) those formed by thermal events, (2) those formed as the result of flexure of the lithosphere by an imposed load and (3) those formed as the result of extension, compression or faulting of the basement (Fig. 10.33). Basins of the first class subside as a result of a heating event that causes uplift and erosion. Then, as cooling and contraction begin, sedimentation occurs. Examples are the Illinois and Michigan basins in the U.S.A. An example of the second class, a basin formed as a result of flexure of the lithosphere under an imposed load, is the Canadian Alberta basin. This foreland basin formed as a result of the loading provided by the Rocky Mountain fold thrust belt. Examples of the third class include the thick deposits laid down on the continental margins as the result of continental rifting (e.g., the Atlantic continental margins), basins formed after stretching of the continental lithosphere (e.g., the Tertiary North Sea basin), basins formed in major fault grabens (e.g., the Basin and Range of the western U.S.A.) and the fault-controlled basins that apparently formed as a result of compression of the lithosphere (e.g., central Australian basins).

These classes of basins are by no means mutually exclusive or comprehensive. For example, the subsidence of the North Sea basin is explained by extension and thinning of the lithosphere by a factor of about 1.25, which gave rise to an initial rapid subsidence. This was followed by cooling, which resulted in a subsequent, slower, exponential thermal subsidence.

10.3.4 Basins of thermal origin

A thermal event probably provides the simplest illustration of the formation of a sedimentary basin. If a region of crust and lithosphere is heated, there is expansion, the density changes and then uplift takes place. Erosion follows, reducing the crustal thickness. Subsequent cooling and contraction of the lithosphere forms a basin. (If no erosion takes place, the surface simply contracts back to its pre-heating level, and no basin forms; nothing changes.) Isostatic calculations based on reasonable values of the increase in temperature during heating give a maximum value for the initial uplift of about 2 km. This would result in a maximum erosion of 4 km, allowing for the isostatic amplification as material is removed (assuming that the erosional time constant is of similar magnitude to the thermal time constant), and therefore would result in a maximum of perhaps 3 km of sediment. This is significantly less than the thickness of sediments on the continental margins, where 10 km of sediment is often present (see Fig. 9.4), and somewhat less than the thickness of sediment in many of the intra-continental sedimentary basins. To deposit 10 km of sediment as a result of heating the lithosphere, an increase in temperature of more than 1000 °C is required.

The subsidence due to thermal contraction is easily modelled mathematically. It is similar to the plate model for formation of the oceanic lithosphere (see Section 7.5.2). To first order, the depth to a particular bed (of age t) in a hole drilled into sediments can be written as

$$d = d_0(e^{t/t_0} - 1) \tag{10.5}$$

where d_0 is a slowly varying function of the position of the well in the basin. The thermal time constant of the lithosphere, t_0, is given by

$$t_0 = \frac{L^2}{\pi^2 \kappa} \tag{10.6}$$

where L is the thickness of the lithosphere and κ is the thermal diffusivity. The thermal time constant for a 125-km-thick lithosphere with thermal diffusivity of 10^{-6} m^2 s^{-1} is 50 Ma. The Michigan and Illinois basins have been modelled on the basis of thermal contraction of the lithosphere (Fig. 10.34). The subsidence of the Illinois basin between 530 and 360 Ma can be explained by one thermal event or by thermal relaxation following an initial extensional event with $\beta = 1.1$–1.2. However, a second thermal event is required in order to account for the rapid Carboniferous subsidence (~360 Ma onwards). The neighbouring Michigan basin has a somewhat different subsidence history. The Cambrian-to-Middle Ordovician (530–470 Ma) subsidence appears to be separate from the Middle Ordovician-to-Carboniferous (460–360 Ma) subsidence, which approximates a thermal-subsidence curve. Some evidence exists for igneous activity at the time of the proposed Cambrian (530 Ma) heating event but not for the others. Invoking a heating event for which there is at present no direct external evidence

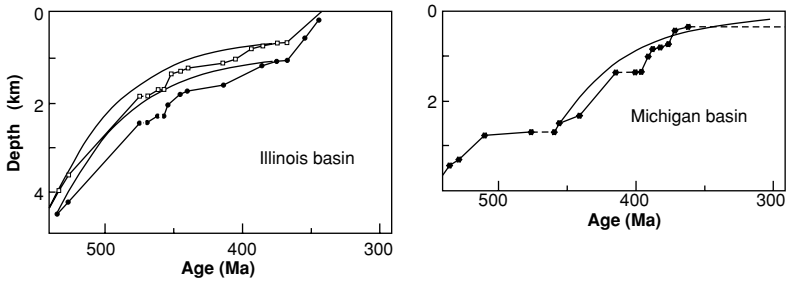


Figure 10.34. Depth to particular sedimentary horizons versus age of the sediment. Data are from two wells drilled in the Illinois basin and one in the Michigan basin. Solid curves are 50-Ma exponentials. If the subsidence were due to thermal contraction of the lithosphere, the depths should plot approximately on these exponential curves. Horizontal dashed lines indicate major unconformities in the sedimentary sequence. The Cambrian-Carboniferous (530–360 Ma) subsidence of the Illinois basin and the Ordovician-Carboniferous (460–360 Ma) subsidence of the Michigan basin could be due to thermal contraction of the lithosphere. (After Sleep *et al.* (1980). Reproduced with permission from the *Annual Reviews of Earth and Planetary Sciences,* vol. 8, © 1980 by Annual Reviews Inc.)

other than the observed subsidence (i.e., no volcanic activity or known uplift) is not altogether satisfactory, and further work on the basement geology and subsidence must be done before these basins can be proven to be of thermal origin. Other explanations are possible.

10.3.5 Flexural basins

Flexure of the lithosphere by a load at the surface produces a foreland basin, of which the Canadian Alberta Basin and the Appalachian Basin are excellent examples. If a load or thrust sheet is placed on the lithosphere, the lithosphere subsides. To study problems of this type, it is not enough to assume local isostatic equilibrium; rather, the lithosphere must be modelled as a thin, elastic or viscoelastic plate, and regional deformation must be calculated as discussed in Section 5.7. The width of a flexural basin is a function of the lithospheric thickness. For a given load, a thick lithosphere supports a wider basin than a thin lithosphere, as one would intuitively expect. Since isostatic balance requires the mass of mantle material displaced from beneath the lithosphere to equal the mass of the load, a wide basin on a thick lithosphere is shallower than a narrow basin on a thin lithosphere (Fig. 10.35(a)).

In the case of a viscoelastic lithosphere, the initial flexure is the same as for an elastic lithosphere, but the lithosphere relaxes with time. A viscoelastic material behaves elastically on a short timescale and viscously on a long timescale. The viscoelastic relaxation time τ, the time taken for any stress to relax to $1/\mathrm{e}$ of its original value, is defined by $\tau = 2\eta/E$, where E is Young's modulus and η

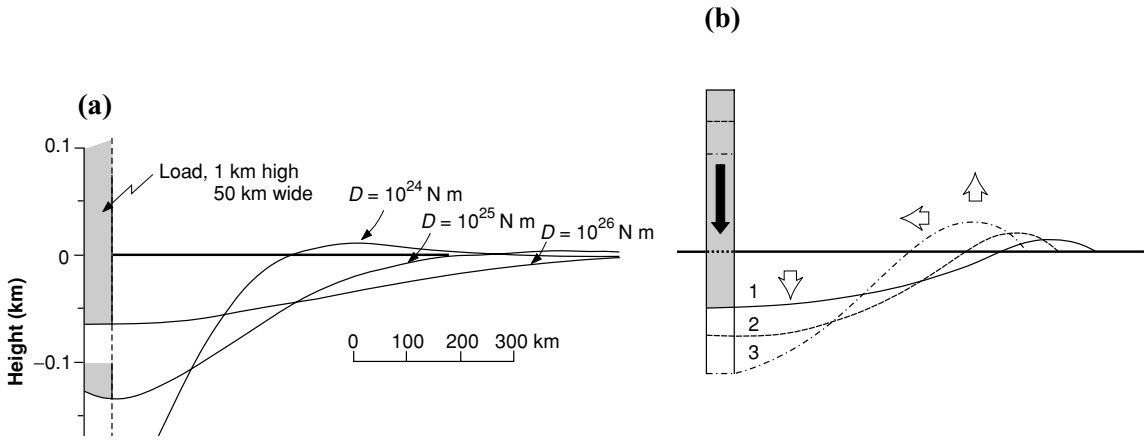


Figure 10.35. (a) Cross sections of sedimentary basins formed on an elastic lithosphere due to a 1-km-high, 50-km-wide two-dimensional load. Load density and sediment density are 2.4×10^3 kg m^{-3}. The three curves show the difference in basin cross sections for a lithosphere with flexural rigidity, D, of 10^{24}, 10^{25} and 10^{26} N m. (From Beaumont (1981).) (b) Deformation of the surface of a viscoelastic lithosphere by a surface load. The initial deformation is the same as for an elastic lithosphere (a). However, with time, the viscoelastic lithosphere allows the stress to relax; so the deformation evolves to curves 2 and 3. The final stage, if the load is left in place, would be local isostatic equilibrium. (From Quinlan and Beaumont (1984).)

the dynamic viscosity. This means that the basin becomes deeper and narrower with time, evolving towards local isostatic equilibrium (Fig. 10.35(b)). The rate at which this relaxation occurs is characterized by the viscoelastic relaxation time τ.

Figure 10.36 shows how loading can produce a foreland basin. Figure 10.37 shows a cross section of the best-fitting flexural model for the Alberta basin. This Canadian basin was formed on the North American plate east of the Rocky Mountains as a consequence of thrusting and emplacement from the west of the Rocky Mountains. Thrusting in the Rocky Mountains began in the late Jurassic (approximately 140 Ma) and lasted about 100 Ma. From 35 Ma to the present, massive erosion (and hence isostatic uplift) has taken place across the region represented by the model. Almost 10 km of material has been eroded from the centre of the thrust pile and about 3 km from its western edge. Most of the Paskapoo formation (dense stipple) has been eroded and many of the older units outcrop away from the Rocky Mountains. The model shown in Figure 10.37 was calculated for a thin viscoelastic plate with flexural rigidity 10^{25} N m and a viscous relaxation time of 27.5 Ma overlying the asthenosphere.

The Appalachian basin has been modelled similarly, but, in this case, the best fit is obtained with a temperature-dependent viscoelastic model that has an approximately elastic thickness for the lithosphere of 80–90 km. Sections across

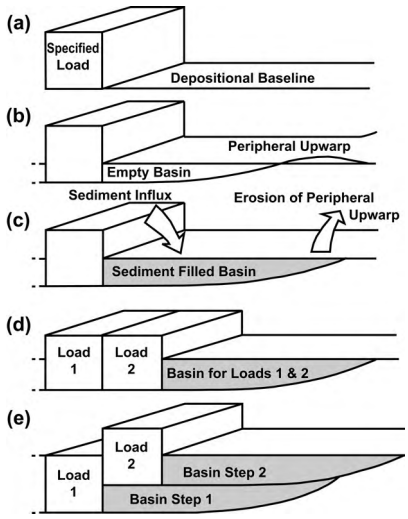


Figure 10.36. How a load on the surface of the lithosphere forms a depression, which, when filled with sediment, becomes a foreland basin. (a) Load is emplaced. (d) Depression forms. (c) Sediment fills the depression, and erosion eliminates the forebulge. (d) Two loads are emplaced simultaneously. (e) Two loads are emplaced sequentially, resulting in a different basin from (d). (After Beaumont (1981).)

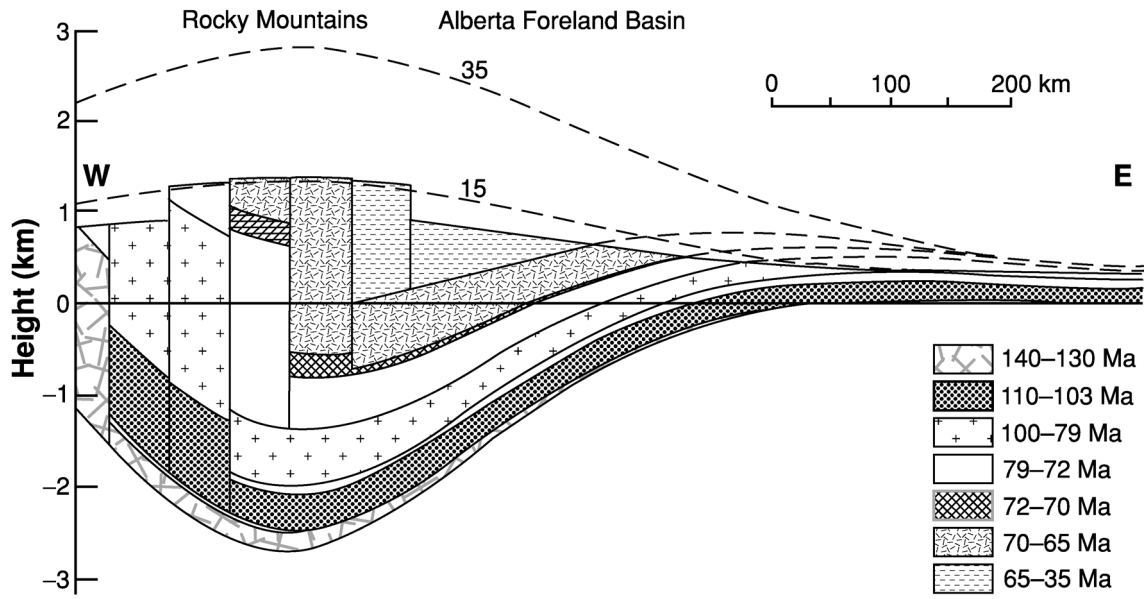


Figure 10.37. A cross section of a loading model for the Alberta Foreland basin. The vertical column loads emplaced during each time interval are shaded like the sediments which filled the resulting depression. Dashed lines denote sediment surfaces that have since been eroded. Since loading stopped 35 Ma ago, massive erosion of the load (thrust pile) and the youngest sediments has taken place. Thus, 35 Ma ago the Paskapoo formation was some 1.5–2.0 km thicker than it is today, and almost 10 km of material has been removed from the centre of the thrust pile. The flexural rigidity of the lithosphere for this model was 10^{25} N m, and the viscoelastic relaxation time was 27.5 Ma. (After Beaumont (1981).)

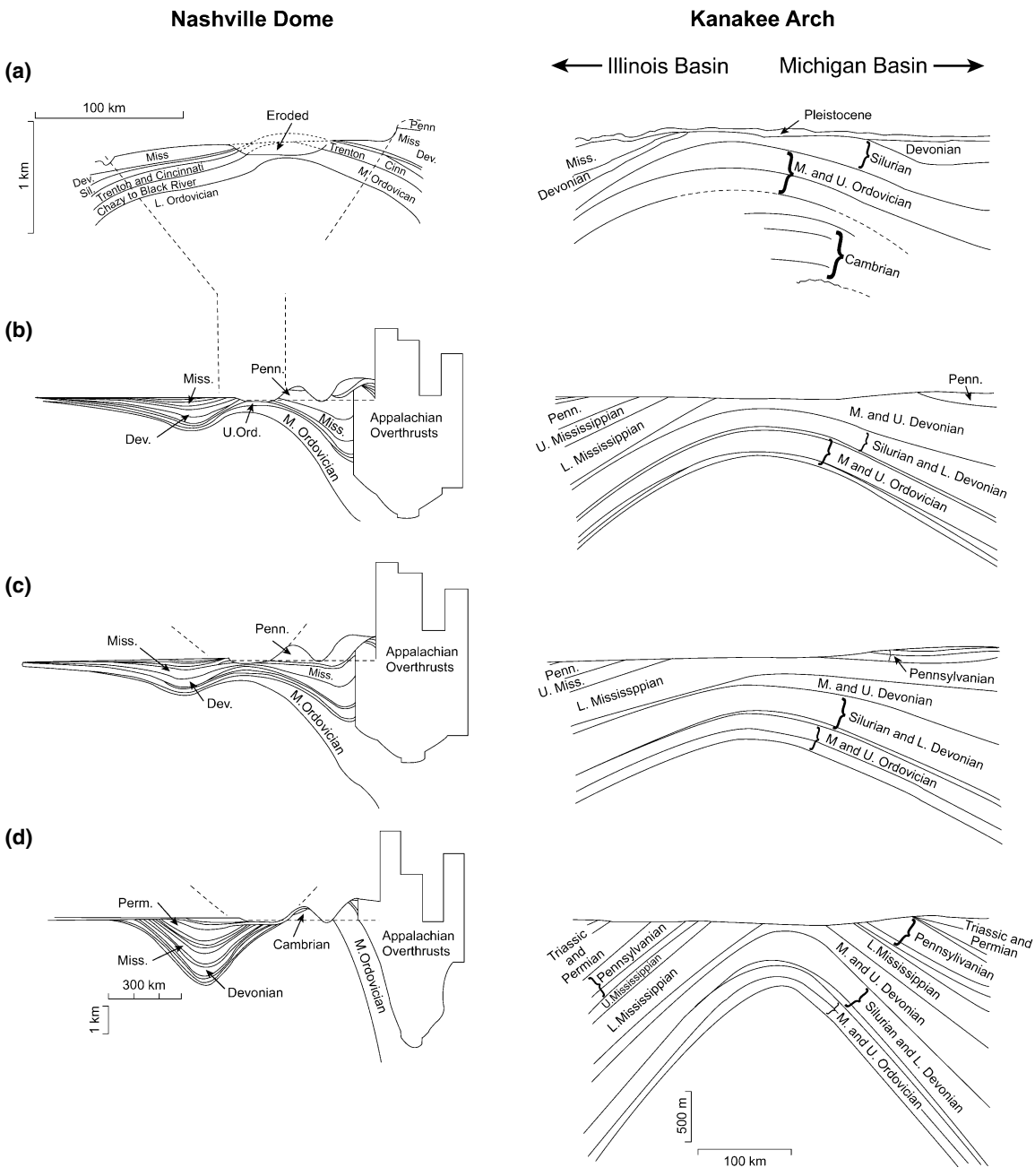


Figure 10.38. Sections across the Nashville Dome between the Appalachians and the Illinois basin, and across the Kanakee Arch between the Illinois and Michigan basins. All models have the same loads. (a) Observed depths to sedimentary horizons. (b) Depths calculated for the best-fitting model which has a temperature-dependent viscoelastic lithosphere. (c) Depths calculated for a model *elastic* lithosphere ($D = 10^{25}$ N m). (d) Depths calculated for a model viscoelastic lithosphere ($D = 10^{25}$ N m, $\tau = 27.5$ Ma). (After Quinlan and Beaumont (1984).)

arches between the Appalachian and Illinois basins and between the Illinois and Michigan basins are shown in Fig. 10.38. Thrusting and emplacement of loads on the Appalachian mountain system took place from about 470 until 200 Ma. The modellers used the observed subsidence of the neighbouring Michigan and Illinois basins, as shown in the previous section, and investigated the effect that their presence has on the intervening arches as a result of the Appalachian loading. Figures 10.38(c) and (d) show the cross sections calculated for the same model loads for an elastic lithosphere and a viscoelastic lithosphere. A model with an elastic lithosphere gives too little uplift over the arches whereas a model with a viscoelastic lithosphere gives far too much relaxation and thus too much uplift and erosion. The best-fitting model has a temperature-dependent viscoelastic lithosphere. This study of the Appalachian basin has shown that the subsidence–time plots for the adjacent Michigan and Illinois basins (see Fig. 10.34) should ideally be corrected to remove the loading effects of the Appalachians.

Foreland basins such as the Appalachian basin, the Alberta basin and the northern Alpine molasse basins are generally linear two-dimensional features, unlike many intercratonic basins, which are roughly circular. It is clear that, since subsidence–time plots for foreland basins are controlled by factors external to the basin and lithosphere, they might not be useful except insofar as they can indicate the possible timing and magnitude of the imposed loads.

The lithosphere can be loaded from below as well as from above. An example of an intercratonic basin that may have formed as a result of loading in or beneath the lithosphere is the Williston basin, which straddles the U.S.–Canadian border just east of the Alberta basin (Fig. 10.39). Prior to the initiation of subsidence during the Cambrian, there had been no tectonic activity in the region since a probable continental collision at about 1800 Ma in the Hudsonian event and some late-Proterozoic events further west. The basin appears to have subsided continuously at a slow, fairly constant rate for most of the Phanerozoic (over 400 Ma). A remarkable feature of the subsidence is that the centre of depression of the basin remained almost in the same place throughout that time (Fig. 10.40). One simple model for the subsidence of the basin involves a steadily increasing load hung under the centre of the basin. This raises the obvious question ‘What is the load?’ One possibility is that, by some means, a region deep in the lithosphere was heated and then slowly cooled and contracted, becoming denser. The problem with this model is that this intrusive body must account for the subsidence without leaving surface volcanism – perhaps a very large, cool intrusion. An alternative explanation is that some part of the lithosphere is undergoing phase changes or metamorphic reaction. Geologically, the most probable reactions are those which involve the growth of high-pressure assemblages in the deep crust or upper mantle. As an example, the complete transformation from gabbro to eclogite at the base of the crust would increase the density from about 3.0×10^3 to 3.4×10^3 kg m^{-3}. This change in density, about thirty times greater in magnitude than that caused by the cooling of gabbro through 150 °C, means that the region undergoing the

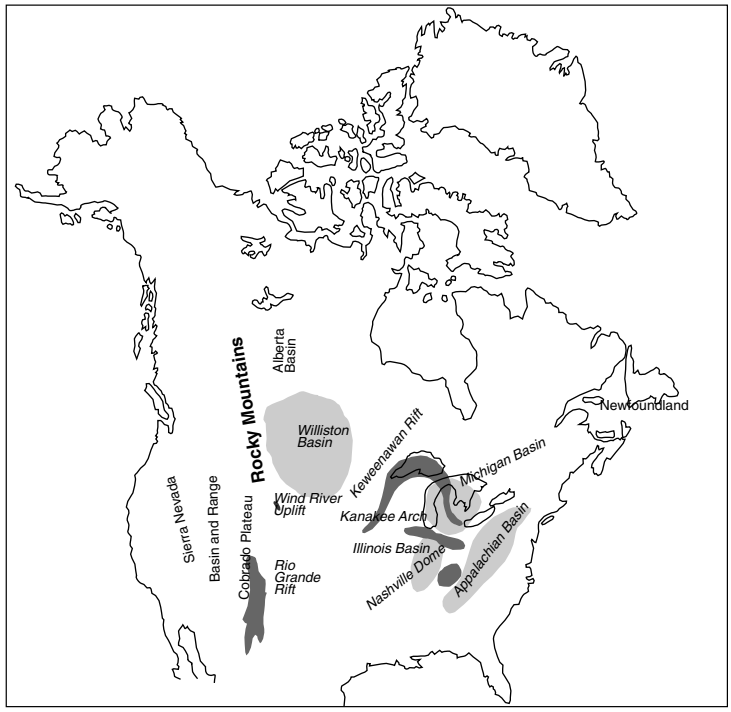


Figure 10.39. Locations of some geographic, geological and geophysical features in North America.

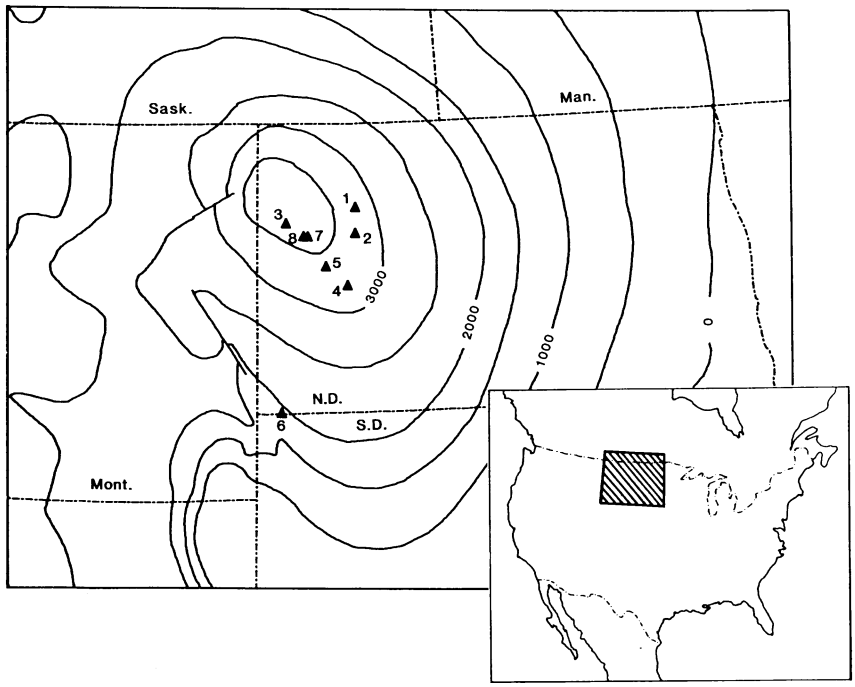


Figure 10.40. Depth to the Precambrian basement beneath the Williston Basin. The contour interval is 500 m. Triangles denote locations of the centre of the basin determined from the best-fitting flexural model of the plate during the following time intervals: 1, 450–352 Ma; 2, 352–328 Ma; 3, 328–324 Ma; 4, 324–239 Ma; 5, 328–239 Ma; 6, 239–168 Ma; 7, 168–0 Ma; and 8, 114–0 Ma. (From Ahern and Ditmars (1985).)

phase changes can be thirty times smaller than needed in an intrusion model and still give rise to the same load. Many other similar metamorphic changes can occur, each with particular rates of change. Such metamorphic changes occur slowly, at rates controlled by the extent to which volatiles are present and by the temperature–time history of the rock. The rate of nucleation of a new phase can be expressed as

$$\text{rate} = K\mathrm{e}^{-\Delta G^*/(RT)}\mathrm{e}^{-H_\mathrm{a}/(RT)}\mathrm{e}^{-P\Delta V/(RT)} \tag{10.7}$$

where ΔG^*, the free energy of activation, is proportional to

$$\frac{1}{(T_{\text{equilibrium}} - T)^2}$$

H_a is the enthalpy of activation, P the pressure, ΔV the change in volume, R the gas constant, T the temperature and K a constant. At temperatures close to equilibrium, ΔG^* is large, so the transformation rate is slow. At intermediate temperatures or at temperatures much greater than the equilibrium temperature, the transformation rate is faster. At low temperatures, however, T is small, so (Eq. (10.7)) the transformation rate is slow. The reaction rates are not known with any accuracy; but, for example, a gabbro layer at depth 50 km – such as would be produced by underplating in a major flood-basalt event (comparable to the Karoo event in southern Africa) or by emplacing a slab of mafic material under the Williston area during the last stages of the Hudsonian plate collision – would cool and transform to eclogite over a time of the order of 10^9 or fewer years. Initially, massive uplift and erosion would occur, but then cooling, contraction and transformation to eclogite would take over and progressively load the lithosphere. This would produce nearly steady subsidence, as may have occurred in the Williston basin. Either of the models discussed could allow for a mid-Proterozoic event to produce Phanerozoic subsidence. Such models are supported by independent seismological evidence: a COCORP deep-reflection profile over the basin shows that the lowermost crust is characterized by relatively high-amplitude reflections. Detailed seismic-refraction surveys have shown that, beneath the basin, where the crust is some 45 km thick, there is a high-velocity lower-crustal layer and some indication of a high P-wave velocity of 8.4 km s^{-1} for the upper mantle. This high-velocity reflective material can most simply be explained as a layer of eclogite. These are simply models, and other models have also been proposed. The origin of the basin is as yet unknown. The inversion of thermal and subsidence data is extremely non-unique.

10.3.6 Extensional basins

Sediments deposited on the passive continental margins record the sedimentary history of the rifting apart of old continents. Passive margins are economically important because the thick sedimentary basins that develop along their length are

major hydrocarbon reservoirs. Figure 9.4 shows that sediment thickness increases rapidly with distance from the coast locally, beneath the continental shelf, reaching a thickness of more than 10 km. This is considerably greater than the thickness of sediment on old oceanic crust. The subsidence rates in these *syn-rift basins*, as recorded in the sedimentary record, decrease with time since rifting. The crust beneath these basins is extensively cut by normal faults, which have been imaged clearly by reflection profiling. These faults are often observed to be *listric faults*, faults for which the angle of dip decreases with depth.

To explain these continental-margin sedimentary basins, it is necessary to consider the effects of stretching the lithosphere and so allowing hotter asthenospheric material to rise. The first effect of such stretching is isostatic: lithospheric material is replaced by asthenospheric material, and, depending on the thicknesses and densities, subsidence or uplift ensures that isostatic equilibrium is reattained. The second effect is thermal: the stretched lithosphere and asthenosphere are not in thermal equilibrium. Cooling, and hence slow contraction and further subsidence, takes place until a new thermal equilibrium is eventually reached. Thus, the sedimentary basins develop as a direct consequence of the stretching, cooling and subsidence of the lithosphere. Stretching models can also be used to explain continental-margin sedimentary deposits formed during continental rifting and some extensional basins that formed when continental lithosphere was stretched, but did not split.

Results of detailed seismic studies combined with ocean drilling have shown that there are two types of passive continental margins – *volcanic* and *non-volcanic* (Fig. 10.41). Volcanic rifted margins are those associated with a significant amount of mantle melting, with volcanism occurring either immediately prior to, or during, the process of continental break-up. In contrast, non-volcanic margins form with no associated melting of the mantle. Adiabatic decompression

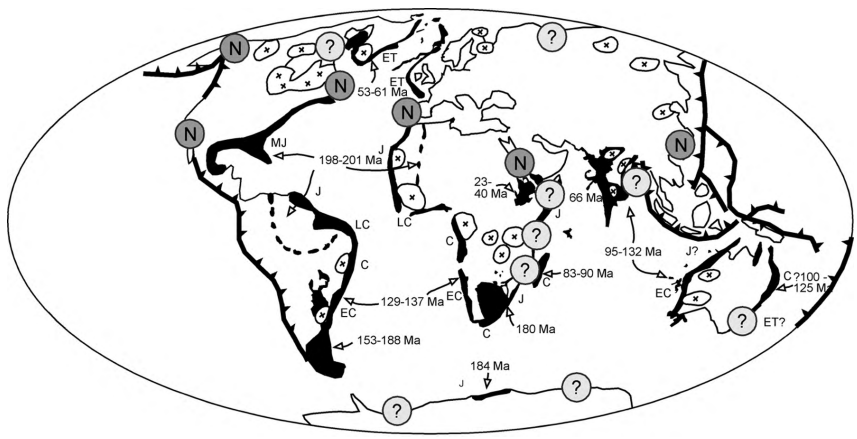


Figure 10.41. Global distribution of rifted margins younger than 200 Ma; volcanic (solid black), non-volcanic (N) and unknown (?) – 90% of continental margins are volcanic. The approximate age of the oldest oceanic crust is shown as well as the age of the continental flood basalts (J, Jurassic; C, Cretaceous; T, Tertiary – prefixes E, M and L denote Early, Middle and Late). ⊗, Archaean cratons. (From Menzies *et al.* (2003).)

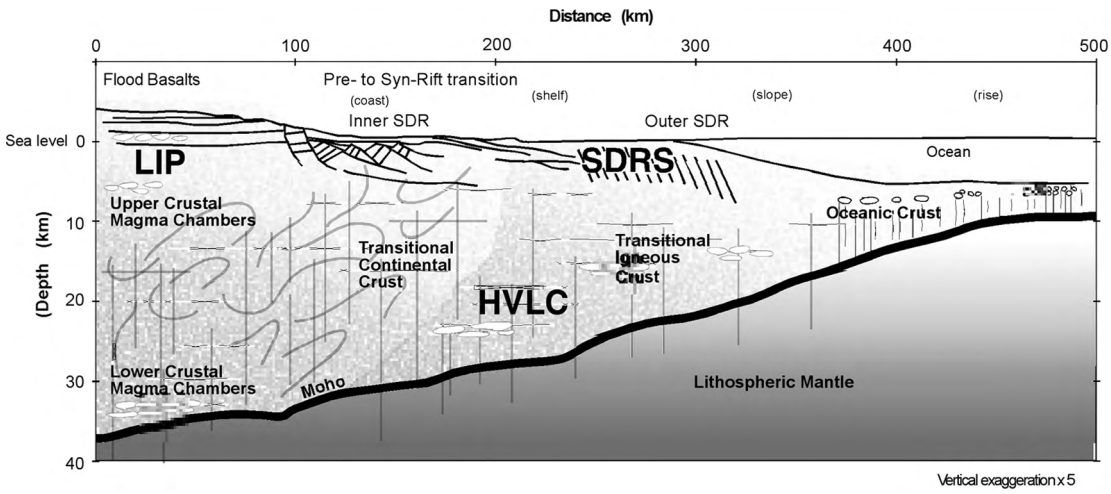


Figure 10.42. A schematic cross section of a volcanic rifted margin, showing the flood volcanism (large igneous province, LIP), seaward-dipping reflectors (SDR) and the high-velocity lower crust (HLVC). (From Menzies *et al.* (2003).)

melting can result in melt production at rifted margins but, for large amounts of melt to be generated, the potential temperature of the underlying mantle needs to be higher than normal (Section 7.7).

Many volcanic continental margins are associated with continental flood basalts, or 'large igneous provinces' (LIP). It has been estimated that more than 2×10^6 km^3 of basalt was extruded onto the rifting margins in the northern North Atlantic over a period of about 2 Ma and that perhaps twice as much igneous rock was added to the lower crust (Fig. 10.42). The extruded volcanic rocks show up on reflection profiles as a set of seaward-dipping reflectors. A thick lower crust with an abnormally high P-wave velocity (7.1–7.6 km s^{-1}) is interpreted as the igneous rocks which were added to the lower crust. The volcanism that took place along the northern North Atlantic margins seems to have been caused by elevated temperatures in the asthenosphere. Were the temperature of the asthenosphere 100–150 °C higher than normal, rapid rifting would generate large quantities of melt (McKenzie and Bickle 1988). Proximity to the hot mantle currently located beneath Iceland was probably the reason for the elevated temperatures beneath the northern North Atlantic at the time of rifting. However, numerical modelling has shown that rifting can induce unstable small-scale mantle convection beneath the continental margin (Boultier and Keen 1999). If the mantle becomes dehydrated when melt is removed, such localized convection can result in some increase in production of magma at the margin, even with normal mantle temperatures, and matches some observations along the North American margin. In addition, since the convection may continue beneath the margin for tens of millions of years after

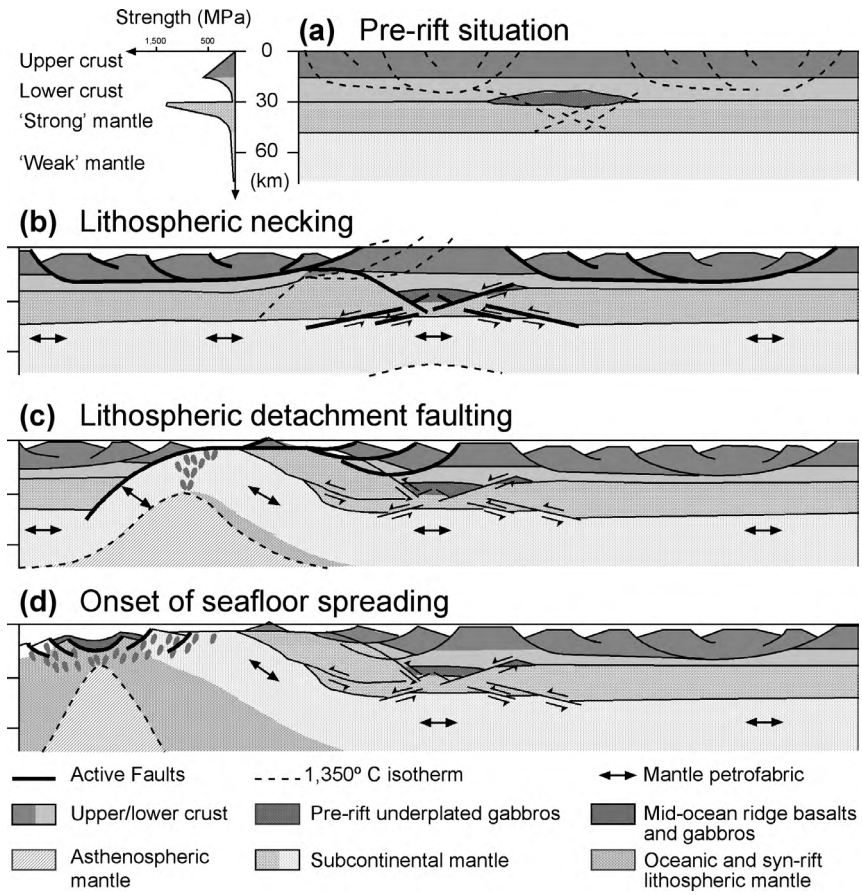


Figure 10.43. A schematic model of the development of a non-volcanic rifted margin, based on the Iberian margin and the Tethyan margin as exposed in the Alps. (a) The initial lithosphere has four layers – strong upper crust, weak lower crust, strong upper mantle and weak underlying mantle. (b) Lithosphere 'necks' beneath an underplated gabbro; ductile lower crust controls rift basins. (c) Rising asthenosphere has a major influence on the style and location of rifting; extensive thinning of continental crust; subcontinental mantle exposed. (d) Mid-ocean-ridge basalts intruded into and onto exposed subcontinental mantle, seafloor spreading begins; margin has a very wide transition zone between oceanic crust and continental crust. The conjugate margin need not be symmetrical. Shown to true scale. Reprinted with permission from *Nature* (Whitmarsh *et al.*, *Nature*, **413**, 150–4. Copyright 2001 Macmillan Magazines Ltd.)

the start of seafloor spreading, coupling between the flow at the margin and that at the ridge axis may be possible.

In contrast, further to the south in the Atlantic off the Iberian penninsula (the conjugate margin to Newfoundland formed at ~130 Ma) is a non-volcanic margin. It lacks any associated volcanic province and seaward-dipping reflectors

and formed with little melting in the mantle. Such non-volcanic margins have very thin stretched continental crust, or even lack any crust, in the continent–ocean transition zone. Along the Iberian margin, the ocean–continent transition is a region of exposed mantle (drilled at basement highs) that has been serpentinized (Fig. 10.43). The width of this transition zone varies considerably along the margin, at its greatest being over 150 km. The stretching/extension along this margin must have been so slow that the unroofed continental mantle was cooled by conduction rather than undergoing any decompression melting. Since exposed mantle peridotite would become serpentinized through reaction with sea water, its seismic velocity and density would be appropriate for crust rather than mantle.

Whether the rifting process itself focusses mantle thermal/plume activity or whether the thermal activity initiates the rifting is much debated – chicken-or-egg-wise. Rifting of continents is a slow process – the time interval between the start of flood volcanism and the formation of oceanic crust can be tens of millions of years. Following flood volcanism and with the onset of extension, the magmatism changes from silicic to felsic and in style and volume. Thick extrusive layers (subaerial and submarine volcanic and sedimentary rocks) form at high levels and are imaged as the sequences of 'seaward-dipping reflectors'. Additionally, the thick zones of igneous material are underplated to the lower crust in the transitional region between the stretched continental crust and the normal oceanic crust. The thickness of these high-velocity zones can be well in excess of the thickness of normal oceanic crust, reaching 10–15 km (Fig. 10.42).

Figure 10.44 outlines the situations that occur when the lithosphere is uniformly stretched by a factor of β. The initial subaqueous subsidence S_i can be determined by assuming the *before* and *after* columns to be in isostatic equilibrium (see Section 5.5.2). In this case, we can write

$$h_c\rho_c + (h_l - h_c)\rho_l = \frac{h_c}{\beta}\rho_c + \frac{h_l - h_c}{\beta}\rho_l + \left(h_l - \frac{h_l}{\beta} - S_i\right)\rho_a + S_i\rho_w \qquad (10.8)$$

where ρ_c is the average density of the crust, ρ_l the average density of the subcrustal lithosphere, ρ_a the density of the asthenosphere, ρ_w the density of water, h_c the thickness of the crust, h_l the thickness of the lithosphere and β the stretching factor. Rearranging Eq. (10.8) gives an expression for the initial subsidence:

$$S_i = \frac{h_l(\rho_a - \rho_l) + h_c(\rho_l - \rho_c)}{\rho_a - \rho_w}\left(1 - \frac{1}{\beta}\right) \qquad (10.9)$$

If the temperature gradient in the lithosphere is assumed to be linear and the temperature of the asthenosphere T_a is assumed to be constant, the densities ρ_c,

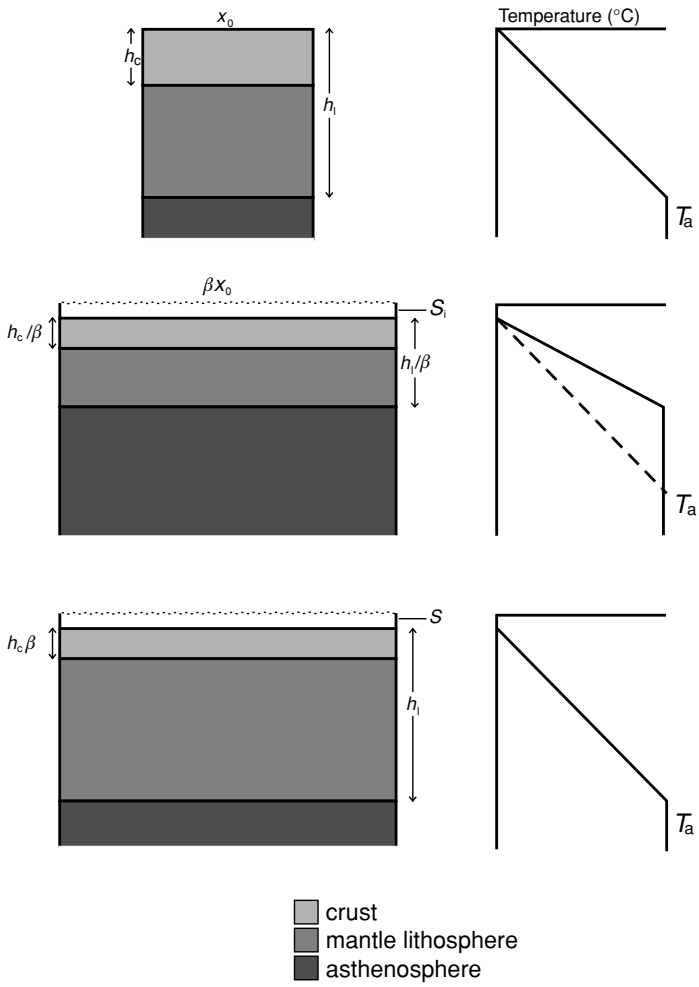


Figure 10.44. At time $t = 0$ the continental lithosphere has initial length x_0 and thickness h_1 (top). It is instantaneously stretched by a factor β (middle). By isostatic compensation, hot asthenospheric material rises to replace the thinned lithosphere. The temperature of the lithosphere is assumed to be unaffected by this stretching, but the stretched lithosphere is out of thermal equilibrium (dashed line). The stretched lithosphere slowly cools and thickens until it finally reattains its original thickness (bottom). An initial subaqueous subsidence S_i occurs as a result of the replacement of light crust by denser mantle material. Further subsidence occurs as the lithosphere slowly cools. The final subsidence is S. (After McKenzie (1978).)

ρ_l and ρ_a are given by

$$\rho_c = \rho_{c_0}\left(1 - \frac{\alpha T_a h_c}{2h_l}\right)$$

$$\rho_l = \rho_{m_0}\left(1 - \frac{\alpha T_a}{2} - \frac{\alpha T_a h_c}{2h_l}\right)$$

and

$$\rho_a = \rho_{m_0}(1 - \alpha T_a)$$

where α is the coefficient of thermal expansion and ρ_{c_0} and ρ_{m_0} are the densities of crust and mantle, respectively, at 0 °C.

After this initial subsidence, the lithosphere gradually cools (Fig. 10.44), which results in a period of slower thermal subsidence. This can be modelled mathematically by using the one-dimensional heat-flow equation (Eq. (7.13))

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c_P}\frac{\partial^2 T}{\partial z^2}$$

with a constant temperature in the asthenosphere (T_a) and temperature increasing linearly with depth z through the lithosphere,

$$\left.\begin{array}{ll} T = \dfrac{T_\mathrm{a}\beta z}{h_\mathrm{l}} & \text{for} \quad 0 \le z \le \dfrac{h_\mathrm{l}}{\beta} \\ T = T_\mathrm{a} & \text{for} \quad \dfrac{h_\mathrm{l}}{\beta} \le z \le h_\mathrm{l} \end{array}\right\} \text{at } t = 0$$

Problems such as this are best solved by Fourier expansion; for details of these methods readers are referred to Carslaw and Jaeger (1959). To a first approximation, the thermal subsidence is an exponential with a time constant equal to the time constant of the oceanic lithosphere (e.g., Eqs. (10.5) and (10.6)):

$$S_t = E_0 r(1 - \mathrm{e}^{-t/\tau}) \tag{10.10}$$

where

$$E_0 = \frac{4h_\mathrm{l}\rho_{\mathrm{m}_0}\alpha T_\mathrm{a}}{\pi^2(\rho_{\mathrm{m}_0} - \rho_\mathrm{w})}$$

is a constant that depends on initial values and

$$r = \frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\right)$$

depends on the stretching factor

$$\tau = \frac{h_\mathrm{l}^2}{\pi^2\kappa}$$

is called the relaxation time.

Equation (10.10) can be used to provide a value of the stretching factor β from the variation of thermal subsidence S_t with time t. A plot of S_t against $1 - \mathrm{e}^{-t/\tau}$ is a straight line through the origin with slope $E_0 r$. So, by making reasonable assumptions for h_l, ρ_{m_0}, α and T_a, β can be calculated.

The *total amount of subaqueous subsidence S* occurring after an infinite time can be most simply expressed by assuming Airy-type isostasy:

$$S = \frac{h_\mathrm{l}(\rho_\mathrm{l}' - \rho_\mathrm{l}) + h_\mathrm{c}\left(\rho_\mathrm{l} - \dfrac{\rho_\mathrm{l}'}{\beta} + \dfrac{\rho_\mathrm{c}'}{\beta} - \rho_\mathrm{c}\right)}{\rho_\mathrm{a} - \rho_\mathrm{w}} \tag{10.11a}$$

which can be rearranged to give

$$S = h_\mathrm{c}\frac{(\rho_{\mathrm{m}_0} - \rho_{\mathrm{c}_0})\left(1 - \dfrac{\alpha T_\mathrm{a} h_\mathrm{c}}{2h_\mathrm{l}}\right)}{\rho_\mathrm{a} - \rho_\mathrm{w}}\left(1 - \frac{1}{\beta}\right) \tag{10.11b}$$

where the average densities of the lithosphere and crust at infinite time are

$$\rho_{\mathrm{l}}' = \rho_{\mathrm{m}_0}\left(1 - \frac{\alpha T_{\mathrm{a}}}{2} - \frac{\alpha T_{\mathrm{a}} h_{\mathrm{c}}}{2\beta h_{\mathrm{l}}}\right)$$

and

$$\rho_{\mathrm{c}}' = \rho_{\mathrm{c}_0}\left(1 - \frac{\alpha T_{\mathrm{a}} h_{\mathrm{c}}}{2\beta h_{\mathrm{l}}}\right)$$

The *total amount of thermal subsidence,* S_t, is then the difference between the total subsidence S (Eq. (10.11)) and the initial subsidence S_{i} (Eq. (10.9)):

$$S_t = S - S_{\mathrm{i}} \tag{10.12}$$

As the temperature gradient changes through the extension and cooling process, the heat flow must also change with time. Before stretching the heat flow was $Q_0 = -kT_{\mathrm{a}}/h_{\mathrm{l}}$. Immediately after stretching the temperature gradient is $\beta T_{\mathrm{a}}/h_{\mathrm{l}}$ and the heat flow is βQ_0. As the lithosphere cools the heat flow decreases back to the equilibrium value of Q_0.

Obviously, models such as this, though complicated, are grossly oversimplified approximations to the real Earth, neglecting lateral conduction of heat and the time over which extension takes place. However, it can be shown that, provided that stretching takes place over less than 20 Ma, an instantaneous-stretching model such as this is adequate. Figure 10.45(a) shows the linear relationships between initial and thermal subsidence and relative thinning $(1 - 1/\beta)$ for particular values of the parameters appropriate for the Bay of Biscay. The usual depth of the mid-ocean ridges is 2.5 km below sea level; thus, this is the isostatic equilibrium level of the asthenosphere or, in other words, the surface of the asthenosphere's geoid. If crust is submerged to depths of more than 2.5 km, extensive volcanism and rupture of the lithosphere occur. Thus, the boundary between oceanic and continental crust should occur at the locus of points offshore at a depth of 2.5 km. For the Bay of Biscay, an initial subsidence of 2.5 km is attained for $1 - 1/\beta = 0.68$, namely for $\beta = 3.24$; this means that asthenospheric material could break through to the surface for β greater than 3.24. However, since the time of breakthrough of asthenospheric material (when the oceanic crust was first formed), additional thermal subsidence will have occurred. It is necessary to know the approximate age of the continental margin in order to calculate this thermal subsidence. Once this is known, the present depth at which the boundary between stretched continental crust and oceanic crust occurs can be calculated. This is useful for making accurate continental reconstructions (see Chapter 3), since it is desirable to calculate the original (pre-stretching) width of the continental margin rather than using an arbitrary present-day bathymetric contour.

Faulting is the surface expression of the initial subsidence caused by extension. Evidence from multichannel seismic-reflection profiles and from fault-plane solutions for earthquakes indicates that, although the normal faults outcropping at the surface have steep dips, the dips decrease with depth. Such faults are called listric, or rotational, faults. Estimates of β made from supposed normal

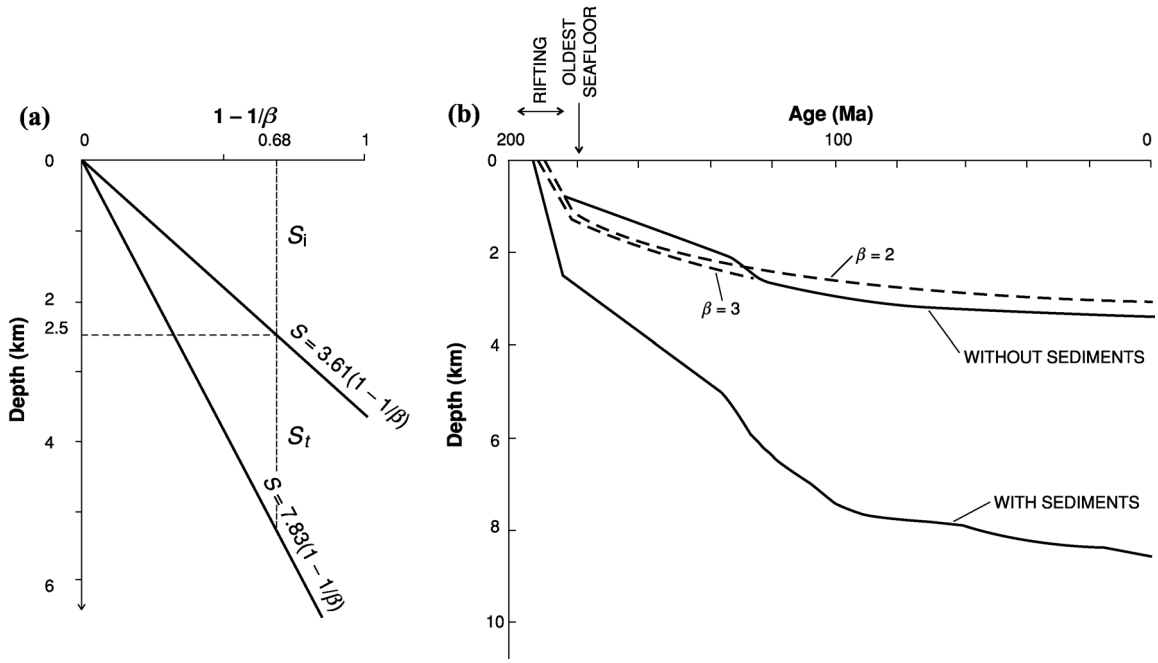


Figure 10.45. (a) A plot of relative thinning (1–1/β) against subsidence for the simple stretching model of the lithosphere. The upper line represents the initial (immediately post-stretching) subsidence S_i as given by Eq. (10.9). The lower line represents the total subsidence, $S = S_i + S_t$ at infinite time as given by Eq. (10.11), where S_t is the thermal subsidence. Values of parameters are h_1, 125 km; h_c, 30 km; ρ_{m_0}, 3350 kg m^{-3}; ρ_{c_0}, 2780 kg m^{-3}; ρ_w, 1030 kg m^{-3}; α, 3.28 $\times$ 10^{-5} °C^{-1}; T_a, 1333 °C. (After Le Pichon and Sibuet (1981).) (b) Observed and computed subsidence curves for a well on the Nova Scotia continental margin off eastern Canada. The rifting period indicates the time during which extension was taking place and thus shows the initial subsidence S_i. Subsidence occurring since that time is due to thermal re-equilibration. (After Keen and Cordsen (1981).)

faulting of the basement surface are, therefore, likely to be lower than the actual value. Measurement of the crustal thickness by seismic-refraction experiments provides another method of estimating β, provided that an adjacent unextended crustal thickness is known. However, in locations with a significant thickness of material underplated to the lower crust, estimates of the stretching factor β obtained from changes in crustal thickness will be too low.

Figure 10.45(b) shows observed and computed subsidence curves for a well on the continental margin off eastern Canada. Figure 10.46 shows a cross section of the passive continental margin of the eastern U.S.A. The transition between stretched and unstretched crust, termed the *hinge zone*, is narrow, being only 50 km across. Figure 10.47 shows observed and computed subsidence curves for five wells situated across the central graben of the North Sea (on a line between Scotland and Norway) and the variation in crustal thickness across the graben.

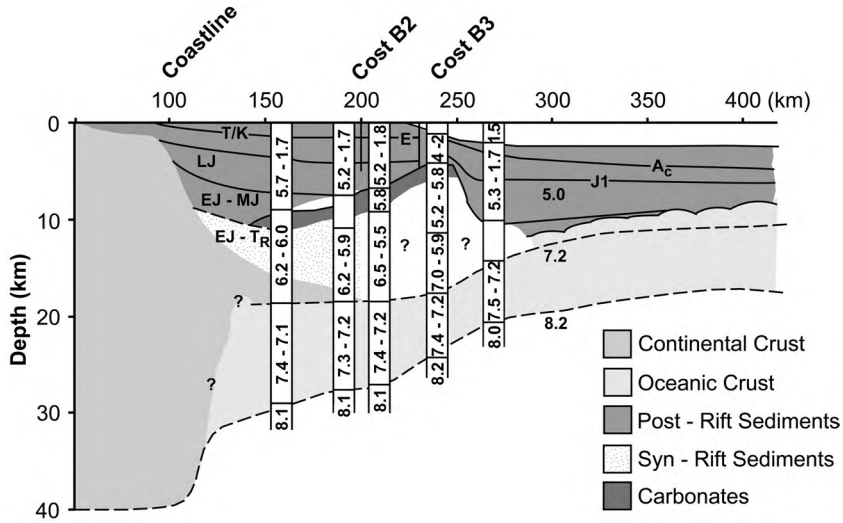


Figure 10.46. A cross section of the passive continental margin off the east coast (New Jersey) of the U.S.A. The numbers are the seismic P-wave velocities in km s^{-1} as determined from expanding-spread (wide-angle) profiles. Vertical solid lines labelled Cost give the locations of two deep commercial wells. The deep crustal layer having P-wave velocity 7.1–7.5 km s^{-1} is believed to be igneous material intruded at the time of rifting; thus, it is very similar to the oceanic layer 3 into which it merges. (From LASE Study Group (1986).)

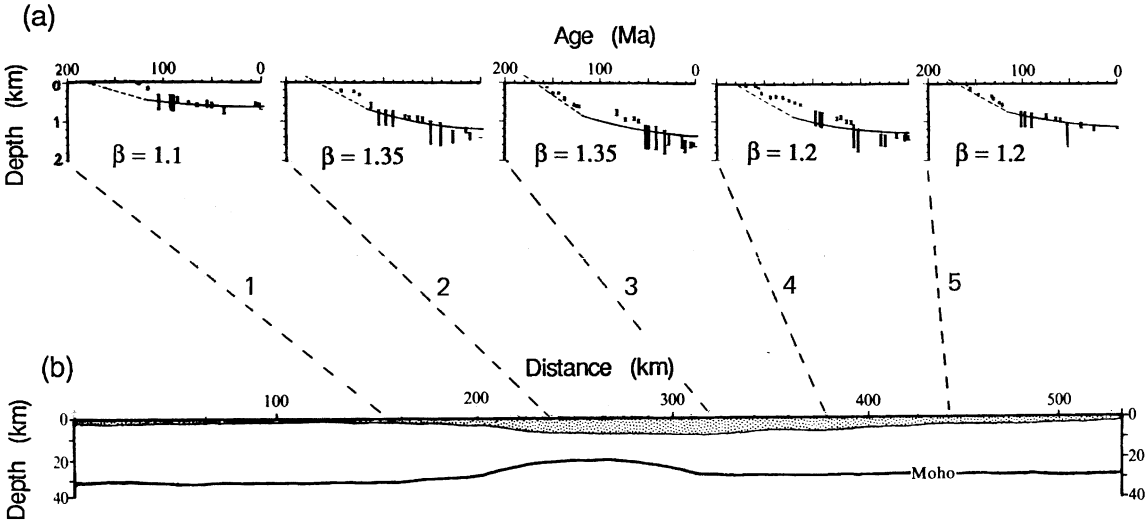


Figure 10.47. (a) Subsidence of the North Sea basin when filled with water only (corrections have been made for sediment compaction, loading and palaeowater depth) for five wells that straddle the central graben of the basin. Bars denote range of basement level; dashed line, calculated fault-controlled subsidence for a model with steady extension occurring over 60 Ma; solid line, calculated subsequent thermal subsidence. The model had an initial crustal thickness of 31 km for wells 1, 4 and 5, 25 km for well 2 and 28 km for well 3. (b) Crustal thickness beneath the North Sea as determined by seismic-refraction experiments. (From Wood and Barton (1983).)

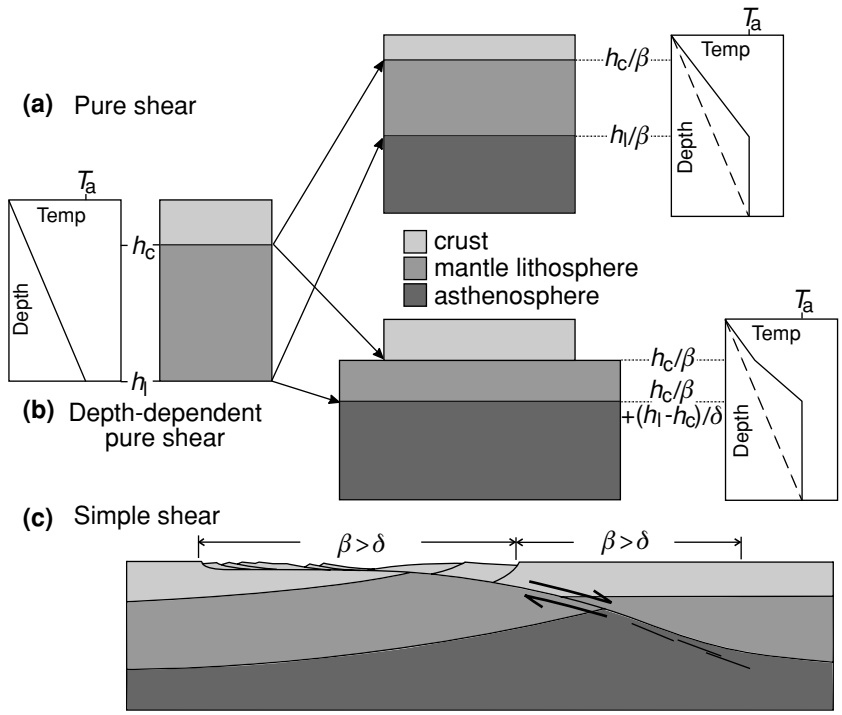


Figure 10.48. Various models of lithospheric extension. (a) Pure shear extension, as in Fig. 10.44. (b) Depth-independent pure-shear extension. The crust is extended by factor β, the mantle by a factor $\delta (\delta > \beta)$. The solid line denotes temperature profiles immediately after extension; the dashed line, the final equilibrium temperature profile. (c) Simple shear extension. A detachment surface, or fault, extends right through the crust and mantle. For this model, the extension factors β and δ vary continually across the structure, in contrast to pure-shear extension, for which they are constant across the structure. (After Quinlan (1988) and Wernike (1985).)

To explain this subsidence satisfactorily, the original crustal thickness beneath the central graben must have been less than that on either side. This suggests that there may have been an earlier stretching event in the graben and is consistent with a Triassic rifting event. Other factors that may have affected subsidence in the North Sea include the thermal effect of the Iceland 'hotspot' and the presence of faults that are too small to be resolved by seismic methods (Sect. 4.4).

More complex modifications of this continental-extension model involve depth-dependent extension – that is, more extension in the lower, more ductile part of the lithosphere than in the upper crustal part (Fig. 10.48(b)) – dyke intrusion or melt segregation and lateral variation of stretching (Fig. 10.48(c)). These complex models have been developed to explain why some continental margins and rift systems apparently exhibit no initial subsidence but some uplift or doming, and to explain why thermal contraction is insufficient to account for the maximum depth of the ocean basins. However, the simple one-dimensional model described here is a reasonable initial approximation to the formation of many continental margins and basins.

10.3.7 Compressional basins

A striking feature of the gravity field of central Australia is the 600-km sequence of east–west anomalies with a north–south wavelength of about 200 km. These Bouguer anomalies range from −150 to +20 mgal (Fig. 10.49). This part of

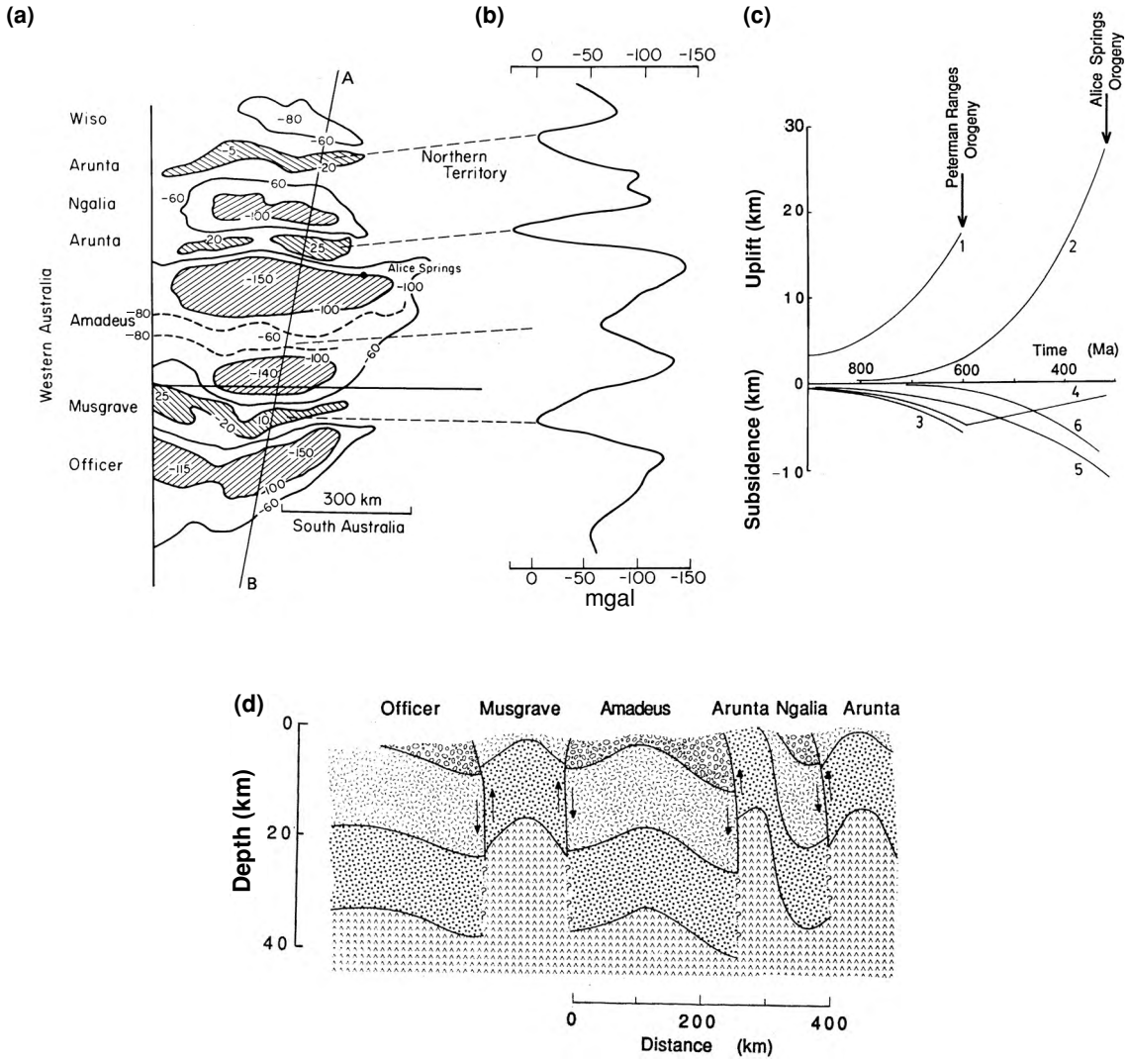


Figure 10.49. (a) A map of the regional Bouguer gravity field in central Australia. \\\\, anomalies less negative than −20 mgal; ////, anomalies more negative than −100 mgal. (b) The Bouguer gravity anomaly along line AB. (c) Predicted uplift of (1) the central part of the Musgrave block and (2) the southern Arunta block; and subsidence of (3) the southern Amadeus Basin, (4) the central Amadeus Basin, (5) the northern Amadeus Basin and (6) the central Ngalia Basin. Note that, except for the central Amadeus Basin, all subsidence and uplift rates have increased with time. (d) The predicted cross section of the crust along line AB after the Alice Springs orogeny (about 320 Ma). The original crust was assumed to be 30 km thick and divided into 15-km-thick upper- and lower-crustal layers. (After Lambeck (1983).)

central Australia comprises a series of east–west-trending intra-continental basins and arches, with the gravity highs corresponding to the arches and the lows to the sedimentary basins. The crust beneath this region apparently has an average thickness of about 40 km, but the Moho is depressed by up to 10 km beneath the basins and is similarly elevated beneath the arches. However, there is no indication of faulting at the bases of the basins. There has been no plate-boundary activity in the region since the late Proterozoic. Subsidence of the basins started about 1000 Ma ago and continued for some 700 Ma, and the subsidence rate has increased with time (Fig. 10.49(c)). These facts, taken together, indicate that neither an extensional nor a thermal model is appropriate for these basins and arches. It has been proposed that they formed instead as a result of compression of the lithosphere.

Let us initially consider a simple problem: an elastic plate, of flexural rigidity D, subjected only to a constant horizontal force H per unit width. The deformation of this plate w satisfies Eq. (5.56):

$$D\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} + H\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} = 0 \tag{10.13}$$

The solution to this equation is obtained by integrating twice, giving

$$D\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + Hw = c_1 x + c_2 \tag{10.14}$$

where c_1 and c_2 are constants of integration.

If we assume the plate to be of a finite length l with $\mathrm{d}^2w/\mathrm{d}x^2 = 0$ and $w = 0$ both at $x = 0$ and at $x = l$ (i.e., the plate is fixed at 0 and l), then both c_1 and c_2 must be zero. The solution to Eq. (10.14) is then sinusoidal:

$$w = c_3 \sin\left(\sqrt{\frac{H}{D}}x\right) + c_4 \cos\left(\sqrt{\frac{H}{D}}x\right) \tag{10.15}$$

where c_3 and c_4 are constants. Because the plate is fixed at $x = 0$, c_4 must equal zero. The condition that w must also equal zero at $x = l$ is then possible only when c_3 is zero (in which case there is no deformation at all) or when

$$\sqrt{\frac{H}{D}}l = n\pi \quad \text{for } n = 1, 2, \ldots \tag{10.16}$$

The smallest value of H for which deformation occurs is therefore given by $n = 1$. This critical value of the horizontal force is $\pi^2 D/l^2$. For horizontal forces less than this value, there is no deformation. At this critical value, the plate deforms into a sine curve given by

$$w = c_3 \sin\left(\frac{\pi x}{l}\right) \tag{10.17}$$

However, this simple calculation is not directly applicable to the lithospheric plates (or to layers of rock) because the lithosphere is hydrostatically supported by the underlying mantle. A hydrostatic restoring force (see Eqs. (5.58) and (5.59)) must be included in Eq. (10.13) in order for us to be able to apply it to the

lithosphere. In this case, Eq. (10.13) becomes

$$D\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} + H\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} = -(\rho_\mathrm{m} - \rho_\mathrm{w})gw \tag{10.18}$$

The sine function $w = w_0 \sin(2\pi x/\lambda)$ is a solution to this equation for values of λ given by

$$D\left(\frac{2\pi}{\lambda}\right)^4 - H\left(\frac{2\pi}{\lambda}\right)^2 = -(\rho_\mathrm{m} - \rho_\mathrm{w})g \tag{10.19}$$

(To check this, differentiate the expression for w and substitute into Eq. (10.18).) Since Eq. (10.19) is a quadratic equation in $(2\pi/\lambda)^2$, the solution is

$$\left(\frac{2\pi}{\lambda}\right)^2 = \frac{H \pm \sqrt{H^2 - 4D(\rho_\mathrm{m} - \rho_\mathrm{w})g}}{2D} \tag{10.20}$$

Because $(2\pi/\lambda)^2$ must be real, not imaginary, the term under the square-root sign must not be negative:

$$H^2 \geq 4D(\rho_\mathrm{m} - \rho_\mathrm{w})g \tag{10.21}$$

The smallest value of H for which there is a real solution is given by

$$H = \sqrt{4D(\rho_\mathrm{m} - \rho_\mathrm{w})g} \tag{10.22}$$

For values of the horizontal force less than this value, there is no deformation; but, at this critical value, the plate deforms into a sine curve. The wavelength of the deformation for this critical force is then obtained by substituting the critical value for H from Eq. (10.22) into Eq. (10.20), which gives

$$\left(\frac{2\pi}{\lambda}\right)^2 = \frac{\sqrt{4D(\rho_\mathrm{m} - \rho_\mathrm{w})g}}{2D} \tag{10.23}$$

Upon reorganization, Eq. (10.23) yields

$$\lambda = 2\pi\left(\frac{D}{(\rho_\mathrm{m} - \rho_\mathrm{w})g}\right)^{1/4} \tag{10.24}$$

For an elastic plate with flexural rigidity of 10^{25} N m, a value that may be appropriate for the lithosphere, the critical compressive force as given by Eq. (10.22) is therefore 10^{15} N m^{-1}, which corresponds to a critical horizontal compressive stress of 10^{10} N m^{-2} (10 GPa). Even a flexural rigidity of 10^{24} N m corresponds to a critical compressive stress of more than 6 GPa. Such values of the compressive stress are much greater than reasonable failure limits of the lithosphere. Buckling would not occur in reality: failure by the formation of faults would take place first.

However, if the lithosphere is modelled as a viscoelastic plate subjected to some irregular normal load, it can be shown that, under constant compression, the initial deflections due to this load are magnified and increase with time. Such deformations occur for compressive forces an order of magnitude less than the critical buckling forces. With time, failure of the crust by thrust faulting presumably also occurs. Figure 10.49(d) shows the cross section of the crust

predicted by such model. All the main geological and geophysical features of the region are correctly predicted.

Oil as a metamorphic product

The main reason for the great commercial interest in sedimentary basins is, of course, the deposits of oil, gas and coal which they may contain. Organic deposits must undergo metamorphism – elevated temperatures and pressures for considerable times – before they become hydrocarbons. The subsidence history of a sedimentary basin combined with oil-maturation history will show where in the basin the oil is likely to be found. In view of the immense economic importance of hydrocarbons, a short discussion of their formation is included here.

Organic remains deposited in a sedimentary basin are gradually heated and compacted as the basin subsides. These organic deposits are called kerogens (Greek *keri*, 'wax' or 'oil'). There are three types. *Inert kerogen*, which is contained in all organic material, transforms into graphite; *labile kerogen*, which is derived from algae and bacteria, transforms into oil, though a small proportion transforms directly into gas; and *refractory kerogen* is derived from plants and transforms into gas. At elevated temperatures, oil also transforms into gas by a process called oil-to-gas cracking. These are complex organic chemical reactions with time and temperature controlling factors. Of course, none of the reactions could take place if the organic material were not buried in sediment and protected from oxidation. In the Guaymas basin in the middle of the Gulf of California, the planktonic carbon-rich silts have been heated to such a degree by the hydrothermal systems (Section 9.4.4) that the kerogens have been transformed into hydrocarbons. The sediments there smell like diesel fuel.

The rates at which the chemical reactions proceed are described mathematically by

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -kC \tag{10.25}$$

where C is the concentration of the reactant (i.e., the kerogen) and k is the rate coefficient of the reaction. The Arrhenius equation (see also Eq. (6.24)) defines the temperature dependence of k as

$$k = A\mathrm{e}^{-E/(RT)} \tag{10.26}$$

where A is a constant (sometimes called the Arrhenius constant), E the activation energy, R the gas constant and T the temperature.

Data on the laboratory and geological transformation of kerogens into oil and gas are shown in Fig. 10.50(a), which demonstrates the time dependence of Eq. (10.25), showing the difference between heating labile and refractory kerogens at geological (natural) rates and heating at a fast rate in the laboratory. The calculations were performed using $A = 1.58 \times 10^{13}\,\mathrm{s}^{-1}$ and $E = 208\,\mathrm{kJ\,mol}^{-1}$ for labile kerogens and $A = 1.83 \times 10^{18}\,\mathrm{s}^{-1}$ and $E = 279\,\mathrm{kJ\,mol}^{-1}$ for refractory kerogens. For the kerogen-to-hydrocarbon reactions to take place within a

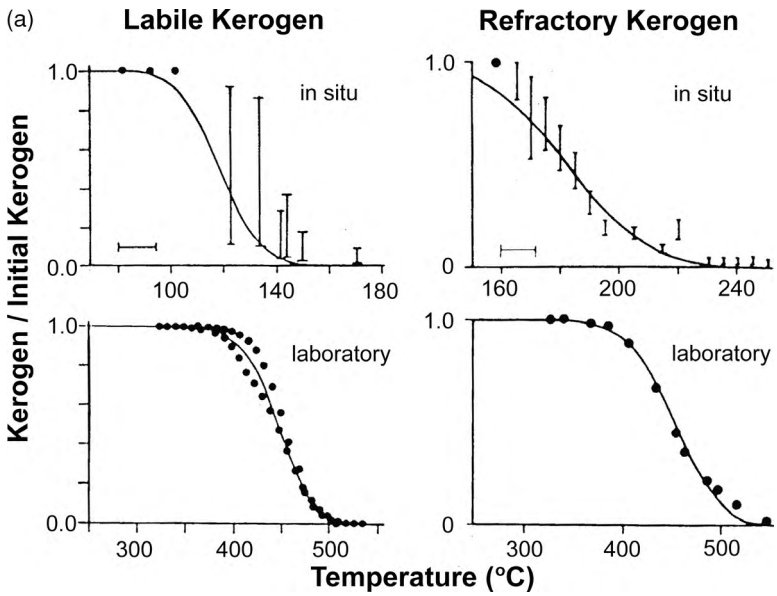


Figure 10.50. (a) Relative concentrations of kerogen for labile kerogen (left) and refractory kerogen (right) as a function of maximum temperature attained: measured (dots and bars) and calculated (solid line). Upper graphs are geological measurements and represent the actual thermal history of the samples; estimates of average heating of these *in situ* kerogens are $1\,^{\circ}\mathrm{C\,Ma^{-1}}$ for the labile kerogen and $6\,^{\circ}\mathrm{C\,Ma^{-1}}$ for the refractory kerogen. Lower graphs are for laboratory-heating measurements carried out at $25\,^{\circ}\mathrm{C\,min^{-1}}$. The fit between measured and calculated relative concentrations is good. (From Quigley and McKenzie (1988).)

reasonable time, very much higher temperatures have to be attained in the laboratory than are necessary in the Earth. (As an illustration, contemplate cooking a turkey in an oven at 50, 100, 150, 200 or 250 °C.) The range of temperatures corresponding to the range of plausible geological heating rates is fairly small. Figure 10.50(b) shows an estimate of the effect of temperature on the time taken for oil to be transformed into gas. The oil half-life is the time necessary for half the oil to transform into gas. At a temperature of 160 °C, the predicted half-life is less than 10 Ma, whereas at 200 °C the half-life is less than 0.1 Ma. These times are short on the geological scale.

In summary, mathematical predictions based on geological and laboratory data indicate that temperatures of 100–150 °C are necessary for labile kerogens to transform into oil, temperatures of 150–190 °C are necessary for the cracking of oil to gas, and temperatures of 150–220 °C are necessary for refractory kerogens to transform into gas.

A standard empirical relationship between temperature and time and the hydrocarbon maturity is called the *time–temperature index* (TTI). This relationship states that the reaction rate doubles for each rise of 10 °C in temperature. The total maturity of a hydrocarbon, or its TTI, is defined as

$$\mathrm{TTI} = \sum_{j=n_{\min}}^{n_{\max}} \Delta t_j \, 2^j \tag{10.27}$$

where Δt_j is the time in millions of years that it takes for the temperature of the material to increase from $100 + 10j$ to $100 + 10\,(j+1)\,^{\circ}\mathrm{C}$, and $n_{\min}$ and $n_{\max}$ are the values of j for the lowest and highest temperatures to which the organic material was exposed. This empirical approach is generally appropriate for chemical reactions on laboratory timescales but has been extended to

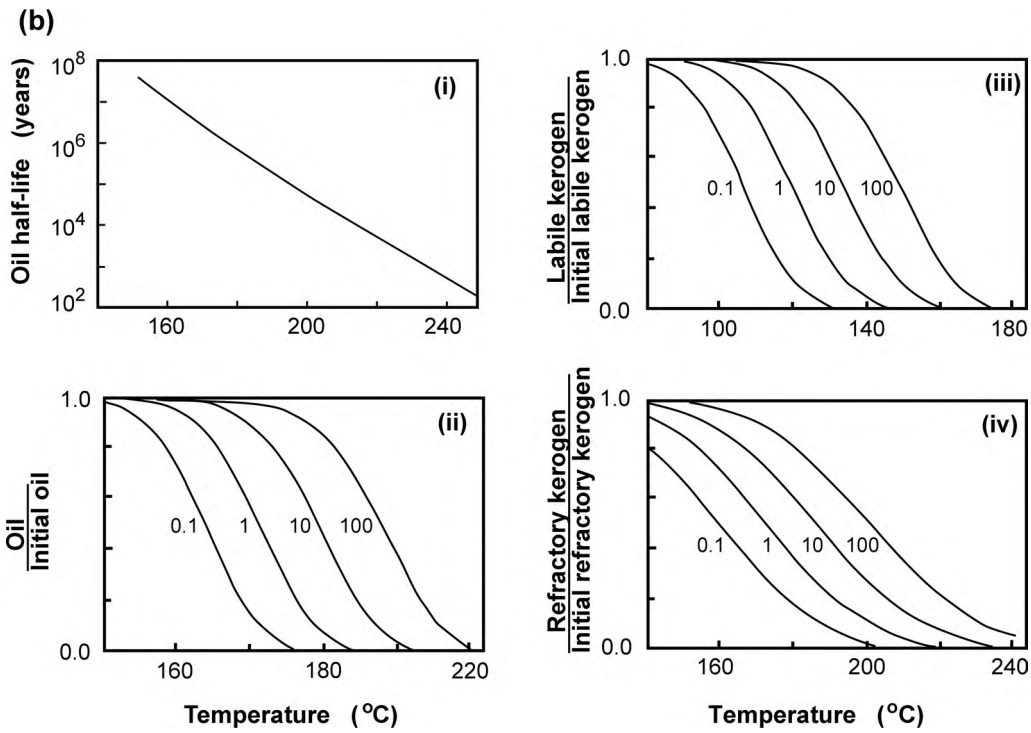


Figure 10.50. (b) Calculated time and temperature dependences of kerogen reactions. (i) Time taken to convert half a given mass of oil into gas at a given temperature. (ii) The relative concentration of oil as a function of the maximum temperature for heating rates of 0.1, 1, 10 and 100 °C Ma^{-1}. (iii) As in (ii) but for labile kerogen. (iv) As in (ii) but for refractory kerogen. (From Quigley and McKenzie (1988).)

geological timescales. However, the data and predictions of transformation rates summarized here suggest that this widely used TTI approach may overestimate the importance of time and underestimate the importance of temperature on hydrocarbon maturation.

Slumps and slides on the continental slope

Continental rifting produces continental shelves and slopes. In the plate cycle, volcanism and deformation at plate boundaries create relief. Erosion then reduces the relief, depositing piles of debris on the continental shelf and slope. Often the final stage is catastrophic failure of the piles, transporting the material down to the deep ocean basin and so eventually to the subduction zones.

Evidence for catastrophic collapse of sediment piles is widespread. The Amazon fan has repeatedly failed in giant submarine landslides, as have many other large river fans (e.g. Nile, Rhône). Deltaic deposits are often rich in organic material and from this gas develops, that in turn seeps up to accumulate in methane hydrates (an ice-like material) below which pools of free gas are trapped. The

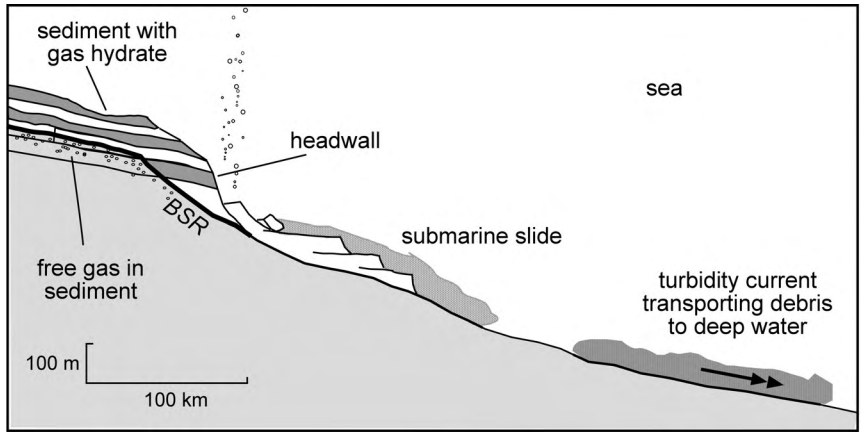


Figure 10.51. A schematic illustration of a submarine landslide on a continental margin slipping on the bottom-simulating reflector (BSR). The sediment is transported to deep water in a turbidity current. Significant volumes of gas would be released in such a failure.

hydrate layers are identifiable seismically as bottom-simulating reflectors (Figs. 4.43 and 4.44) and they may act as planes of failure in landslips (Fig. 10.51).

During glaciation, sediment is bulldozed onto the continental shelf by ice, or accumulates there as a result of down cutting at times of low sea level (125 m at the peak of the most recent ice age). When the ice age ends, this sediment can be unstable. Isostatic uplift of deglaciated land can induce earthquakes, triggering landslips. About 6200 B.C., a large earthquake (magnitude ~8.2), with a 160-km fault break and uplift of 5–15 m, may have shaken Norway. Offshore, the Storegga slide occurred, spreading 3300 km^3 of debris across the floor of the Norwegian Sea and producing a 20-m high tsunami on Shetland. It is likely that methane hydrates were involved in the failure. Since the Storegga slide, new gas has seeped up from older rocks deeper in the sequence, and has created the Ormen Lange gas field, one of Europe's largest.

10.4 Continental rift zones

10.4.1 Introduction

Some of the continental rift zones which are active today have not yet, and perhaps may never, become active mid-ocean ridges (refer to Section 10.3.6). However, some features are common to all continental rift zones:

1. a rift or graben structure with a rift valley flanked by normal faults;
2. negative Bouguer gravity anomalies;
3. higher than normal surface heat flow;
4. shallow, tensional seismicity; and
5. thinning of the crust beneath the rift valley

These features are in agreement with those expected for the early stages of extensional rifting.

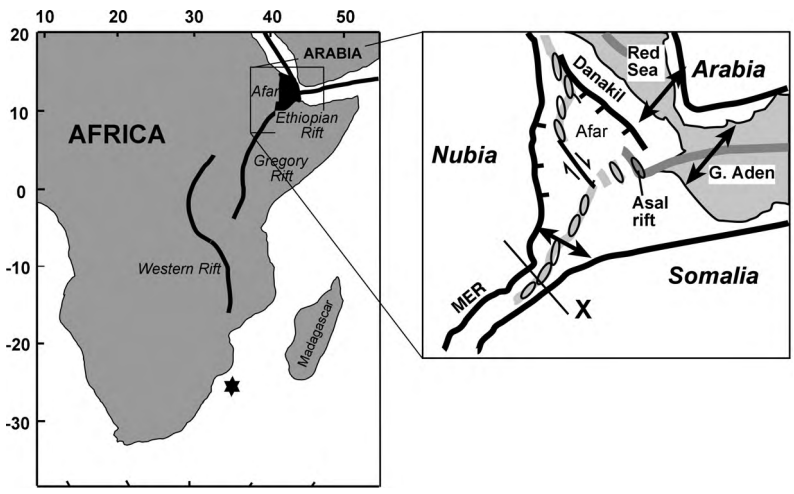


Figure 10.52. The East Arican Rift system. The African plate is slowly splitting along the rift system. The star indicates the pole of motion between the Nubian plate and Somalian plate. The separation is very slow – 6 mm/yr^{-1} in the north and 3 mm/yr^{-1} in the south. Below: detail of the Afar region showing the connection between the extension in the Main Ethiopian Rift and the mid-ocean ridges in the Gulf of Aden and the Red Sea. Arrows show the direction of spreading. The Danakil horst has moved southeastwards away from Nubia – the intervening material is oceanic. Solid lines: main border faults; grey shading, locus of extension; black ovals, magmatic segments. X, location of section shown in Fig. 10.55(b). (Lower part after Hayward and Ebinger (1996).)

Two of the best-known rift zones are the East African Rift and the Rio Grande Rift, though there are others such as the Rhine Graben in Europe and the Baikal Rift in Asia. The Keweenawan Rift is a North American example of an ancient extinct continental rift.

10.4.2 The East African Rift

This long rift system stretches over 3000 km from the Gulf of Aden in the north towards Zimbabwe in the south (Fig. 10.52). In the Gulf of Aden it joins, at a triple junction, the Sheba Ridge and the Red Sea. Along this rift system, uplifting, stretching, volcanism and splitting of the African continent are in progress. The African plate is currently moving as two plates – the main Nubian plate and an eastern Somalian plate. The rotation pole for these two plates is just off the east coast of southern Africa (Fig. 10.52).

The crustal and upper-mantle structure of the rift system has been determined from seismic and gravity data. Figure 10.53 shows data and models from Kenya, where the long-wavelength Bouguer gravity anomaly and the earthquake data can be explained by invoking the presence of anomalous low-velocity, low-density

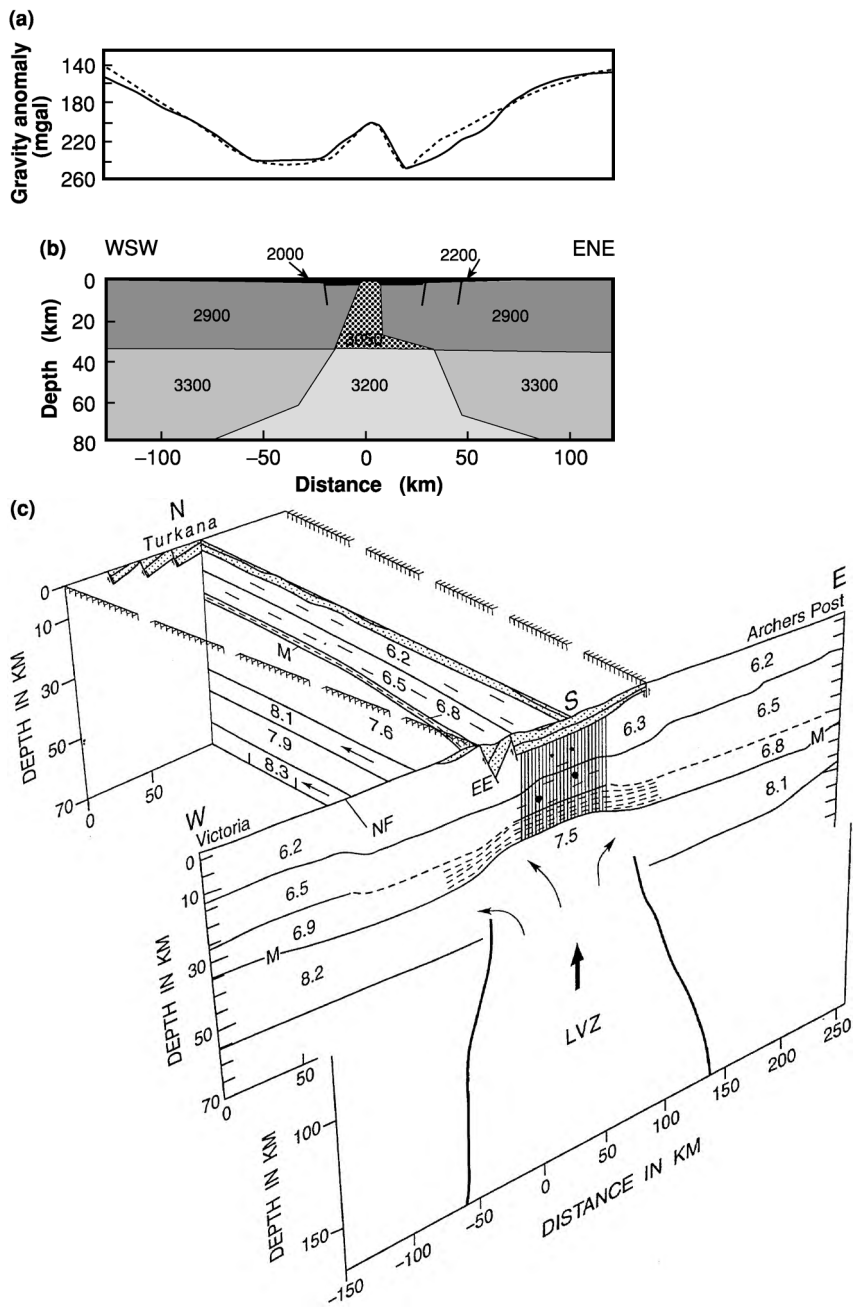


Figure 10.53. Cross section over the East African Rift in Kenya. (a) Observed (solid line) and model (dotted line) Bouguer gravity anomalies. (b) The density model for model anomalies shown in (a). (c) Seismic P-wave velocity structure as determined from refraction data (crust and uppermost mantle) and teleseismic delay times (mantle low-velocity zone). The northern section is at Lake Turkana and the southern section is at Lake Baringo (~300 km apart). The rift widens from 100 to 175 km. The dyke injection zone (vertical lines) is 40 km wide in the south. Stipple, rift infill; M, Moho; EE Eigyo escarpment; NF, Nandi fault, separating the Tanzanian craton (west) from the Mozambique Proterozoic belt (east). Arrows indicate possible flow of mantle rock upwards and away from the Kenya Dome. (After Baker and Wohlenberg (1971). Reprinted from *Tectonophysics*, **278**, Mechie, J. *et al.*, A model for the structure, composition and evolution of the Kenya Rift, 95–119, Copyright 1997, with permission from Elsevier.)

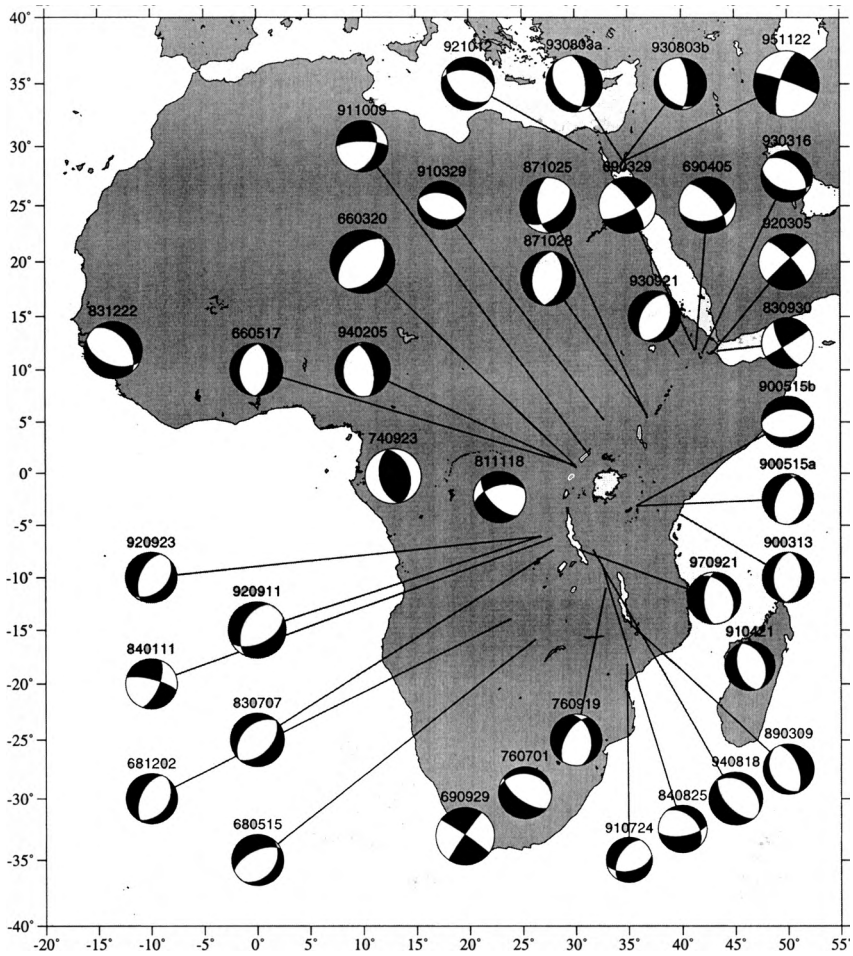


Figure 10.54. Epicentres and focal mechanisms for African earthquakes. Dates are in the format year month day. (From Foster and Jackson 1998.)

material in the upper mantle. The region is approximately in isostatic equilibrium, with evidence for up to 15% dynamic support of the topography. Earthquake and seismic-refraction data show that the crust on either side of the rift valley has a simple, typical shield structure: upper and lower crustal layers with velocities of 6.0–6.2 and 6.5–7.0 km s^{-1} and a total thickness of 35–45 km overlying a normal upper mantle with velocity 8.0–8.3 km s^{-1}. Within the rift valley in northern Kenya the crust is 20 km thick. A positive Bouguer gravity anomaly lies immediately over the eastern rift in Kenya, which is interpreted as being due to a zone of denser, molten material.

The largest body-wave magnitudes for earthquakes occurring along the rift system are about 7.2–7.5, but such events are very rare. Earthquakes along the rift system are normal-faulting events (Fig. 10.54), which are generally consistent with the expected relative motions. The focal depths extend down to 35 km, with seismicity taking place throughout the upper and lower crust. Generally, in areas

of continental extension, earthquakes nucleate in the upper 15 km of the crust (the seismogenic layer). The implication of this thick seismogenic layer in eastern Africa is that the crust is strong and thick. This is consistent with the low heat flow and low temperature gradients measured over cratons (Chapter 7).

The uplift and the volcanism which started in northeast Africa ~40 Ma ago were caused by a mantle plume. Global seismic tomographic images reveal extensive low-velocities in the mantle beneath east Africa (Plate 10). The hotter-than-normal mantle is providing dynamic support for the elevation of the whole region – the African superswell. About 31 Ma ago there was an outpouring of flood basalts across a ~1000-km-wide area and after this, as Arabia moved north-eastwards away from Africa, stretching began in the Red Sea and Gulf of Aden. Seafloor spreading had started in the Gulf of Aden by 10 Ma ago and in the Red Sea by 4 Ma ago. The East African Rift exhibits all stages of the break- up of a continent along its length. As continental stretching starts, normal faults develop and the lithosphere thins. To the south, where the continental rift is young and extension is not great, the rift is characterized by border faults. At some stage in the stretching process the continental lithosphere reaches the point of 'break-up' – and a new ocean basin forms. Magmatic processes control the resultant oceanic spreading whereas faulting controls the earlier continental rifting. This transition from continental to oceanic rifting is currently taking place along the northern Ethiopian Rift (Fig. 10.55). There the extension (geodetic data show ~80% of the strain) is confined to a narrow zone within the rift valley rather than being accommodated on the normal faults that define the ~100-km-long rift valley. The planform of the volcanic activity is oceanic – the segmentation is that of a slow-spreading ridge (Fig. 9.34 and Table 9.6). In the extreme north towards Afar (i.e., furthest from the rotation pole) continental break-up has already taken place and seafloor spreading is effectively taking place, but along the rest of the rift system the continental lithosphere is still undergoing extension. In Afar the maximum extension may be as much as 70–100 km, but south of Afar the geological estimate of the maximum extension which has taken place is 30 km. Thus the extension rates decrease from north to south as the rotation pole is approached. Volcanism along the rift is rather alkaline, which is normal for continental volcanism in relatively undisturbed lithosphere.

10.4.3 The Rio Grande Rift

The Rio Grande Rift is a much smaller feature than the East African Rift system. Visually, the two rift systems are very similar, with platform-like rift blocks rising in steps on each side of the central graben. Volcanism in the Rio Grande Rift began 27–32 Ma ago in the Precambrian Shield as a northeast–southwest rift opened. Subsequently extension 5–10 Ma ago resulted in a north–south rift characterized by a thermal anomaly and crustal thinning. The present–day lithospheric and mantle anomaly is primarily the result of westnorthwest–eastnortheast extension

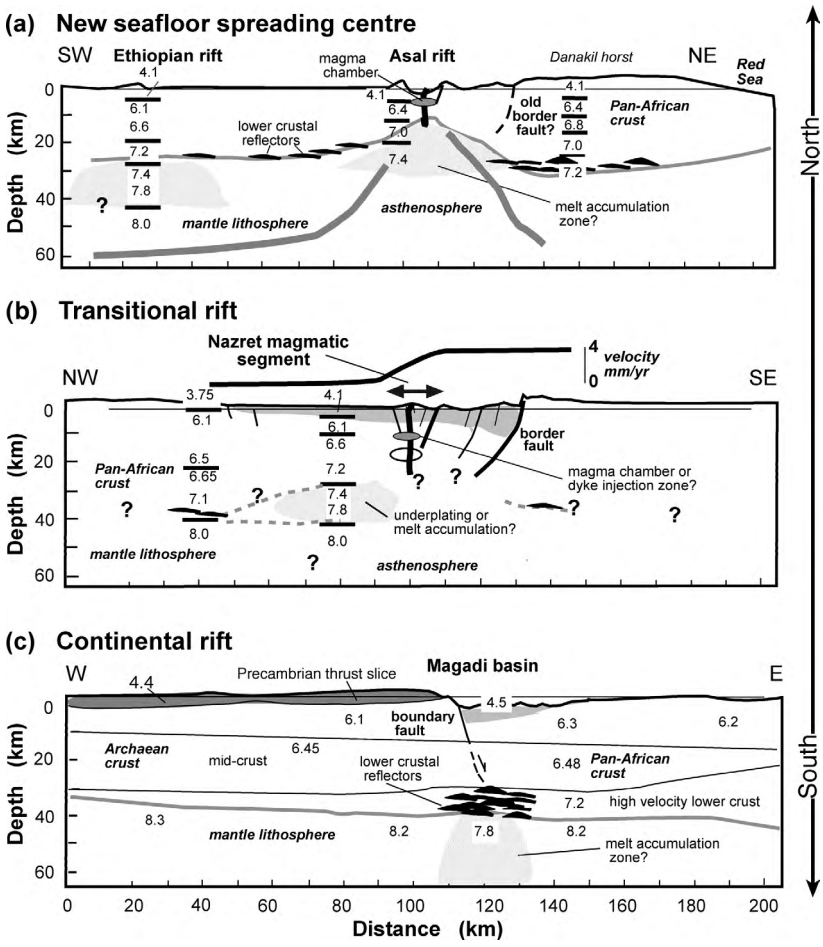


Figure 10.55. Sections across the East African Rift system, showing the changes in crustal structure on going from (c) continental rifting in Kenya (at 1–2°S in Fig. 10.52) to (b) transitional rifting in Ethiopia (location in Fig. 10.52 inset), where the extension is confined to a narrow zone in the centre of the rift valley rather than on the border faults and magmatism is organized in oceanic-style segments, to (a) seafloor spreading (Asal rift in Fig. 10.52 inset). (C. J. Ebinger, personal communication 2003.)

over the last 5 Ma. This is all a contrast to the East African Rift, where the narrow, shallow and wider, deep anomalies are aligned.

Seismic-refraction data indicate that the crust in the central part of the rift is about 30 km thick, which is some 20 km thinner than the crust beneath the Great Plains and 10–15 km thinner than the crust beneath the Colorado Plateau. The upper-mantle P-wave velocity beneath the rift is only 7.7 km s^{-1} (Fig. 10.56). Results of teleseismic time-delay studies of the upper mantle show that P- and S-wave velocities down to 145 km beneath the rift are 7%–8% lower than those beneath the Colorado Plateau and the Great Plains. The presence of melt within the upper mantle is consistent with much of the seismic data but is not required in order to explain it – the temperatures are presumed to be close to the solidus. The gross crustal structure for the rift is simple, having just two layers. The discontinuity between the upper and lower crustal layers gives rise to a strong reflection in the seismic-refraction data. The amplitude of this reflection has

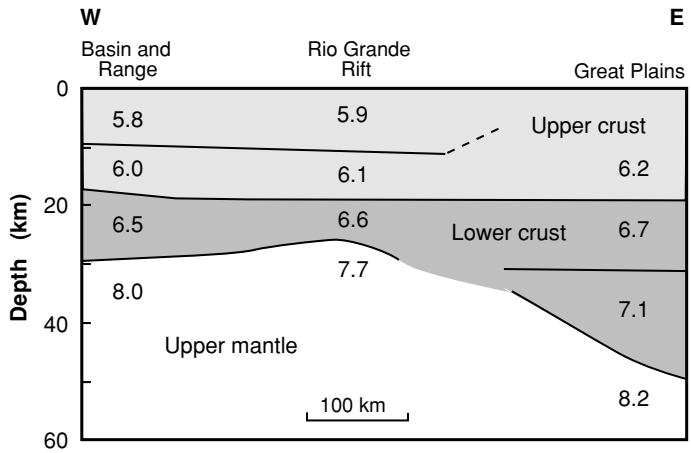


Figure 10.56. A cross section of crust and upper-mantle structure across the Rio Grande Rift. Velocities are based on refraction and surface-wave results. (From Sinno *et al.* (1986).)

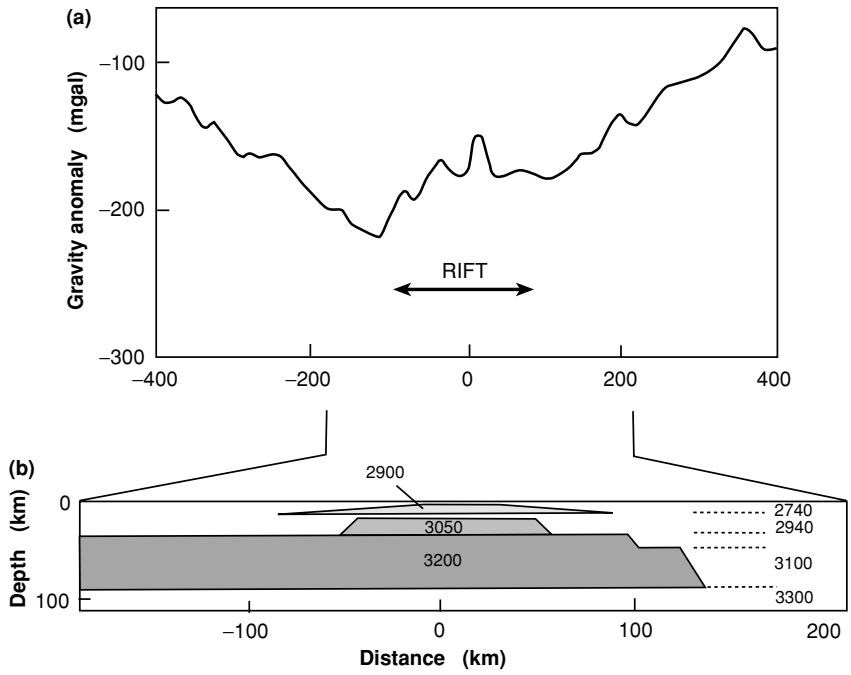


Figure 10.57. The Rio Grande Rift. (a) The Bouguer gravity anomaly on a profile at approximately 33°N. (After Cordell (1978).) (b) Interpretation of the Bouguer gravity anomaly after corrections for the shallow structure have been made. (After Ramberg (1978).)

been modelled with synthetic seismogram programs. Provided that the upper few kilometres of the lower crustal layer have a low S-wave velocity, the amplitudes of the synthetic seismograms are in agreement with the data.

Figure 10.57 shows the gravity anomaly along a profile crossing the rift. A small positive Bouguer gravity anomaly is superimposed on a wide, low (−200 mgal) anomaly. The interpretation of these gravity data is ambiguous, but, like the East African Rift, the broad, low anomaly appears to be caused

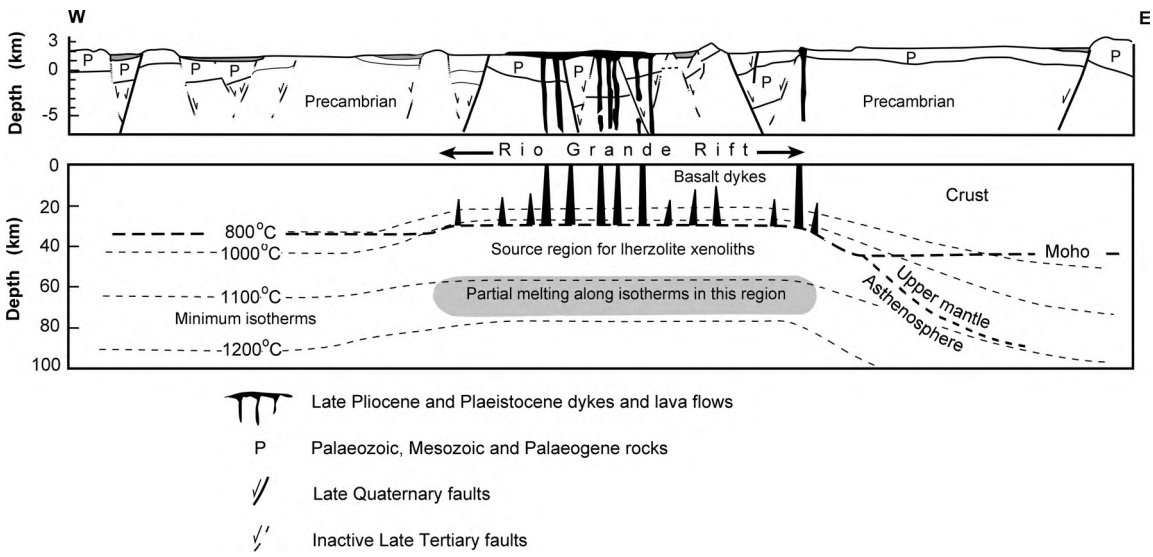


Figure 10.58. Cross sections of the (a) geological and (b) thermal structure beneath the Rio Grande Rift. (From Seager and Morgan (1979).)

by thinning of the lithosphere beneath the rift, and the small positive anomaly beneath the rift is caused by shallow, dense intrusions at the rift itself.

Heat flow along the rift is high, about 120–130 mW m^{-2}, with local values of up to 400 mW m^{-2}. Results from studies of xenoliths (fragments of rock brought up from depth) erupted from volcanoes indicate that their source is at a temperature greater than 1000 °C. Figure 10.58 illustrates the temperature field beneath the rift. The uppermost mantle is also the location of significant electrical-conductivity anomalies, with high conductivity (low resistivity) beneath the rift. This is yet another piece of evidence that the temperatures in uppermost mantle beneath the rift are very high.

Since detailed instrumental studies began in 1962, the seismicity in the rift has not been high by western American standards; on average there have been only two earthquakes per year with magnitude greater than 2. Several areas of concentrated microseismic activity exist in the rift and are associated with magma bodies in the middle and upper crust. An aseismic region has particularly high heat-flow values (up to 200 mW m^{-2}), indicating that the lack of seismic activity is due to high temperatures in the crust. Fault-plane solutions for the microearthquakes in the rift show that the focal mechanisms are predominantly normal faulting with some strike–slip faulting.

The microearthquake data have also been used to delineate the top of the mid-crustal magma body beneath the rift. The seismograms for earthquakes in the region of Socorro, New Mexico, U.S.A., show pronounced secondary energy arriving after the direct P- and S-waves (Fig. 10.59). These arrivals, S_zP and S_zS,

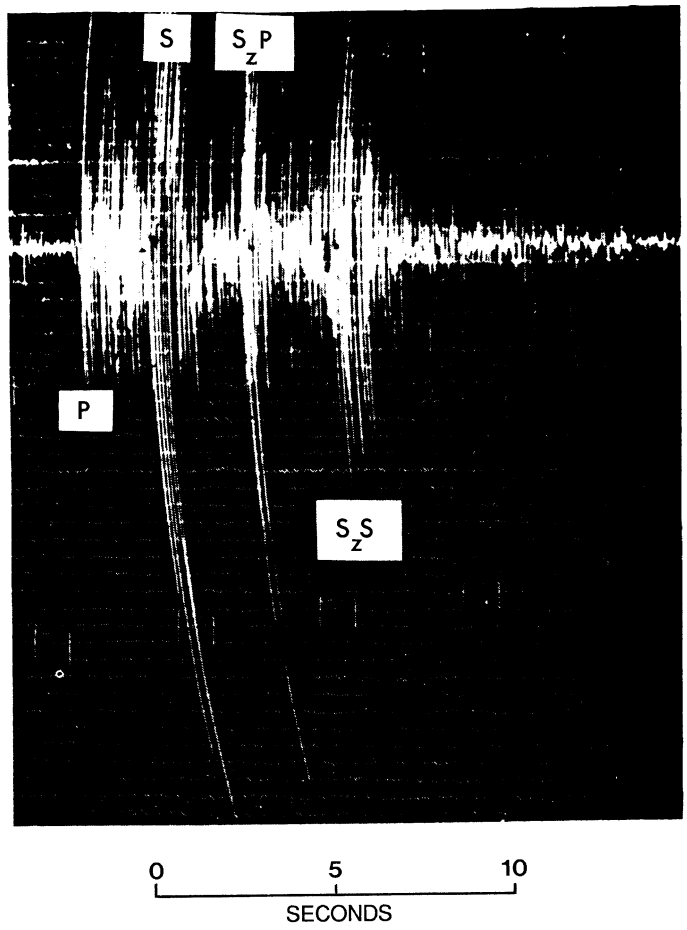


Figure 10.59. A microearthquake seismogram from the Rio Grande Rift. P and S denote first P- and S-wave phases; S_zS and S_zP are S-to-S and S-to-P reflections from the upper surface of an extensive 20-km-deep magma body. (From Rinehart *et al.* (1979).)

are an S-to-P reflection and an S-to-S reflection from a seismic discontinuity at 20 km depth. Ratios among the amplitudes of the various phases have been used to determine the nature of the material immediately beneath the reflecting horizon. A solid–liquid interface can account for the observations. However, this magma body cannot be more than about 1 km thick (if it is completely molten) because it does not cause observable delays in P-wave teleseismic or refraction arrivals. The body lies beneath 1700 km^2 of the central part of the rift.

A series of 24-fold, deep-seismic-reflection lines was shot by COCORP across the Rio Grande Rift near Socorro in the region where this magma sill is located (Figs. 4.56 and 10.60). The shallow reflections show normal faults, some having an offset of more than 4 km. The extensional origin of the rift shows very clearly. A clear, rather complicated P-wave reflector at 7–8 s two-way time corresponds to the magma body. The time at which the reflections from the Moho arrive is consistent with the predictions from the surface-wave and refraction data.

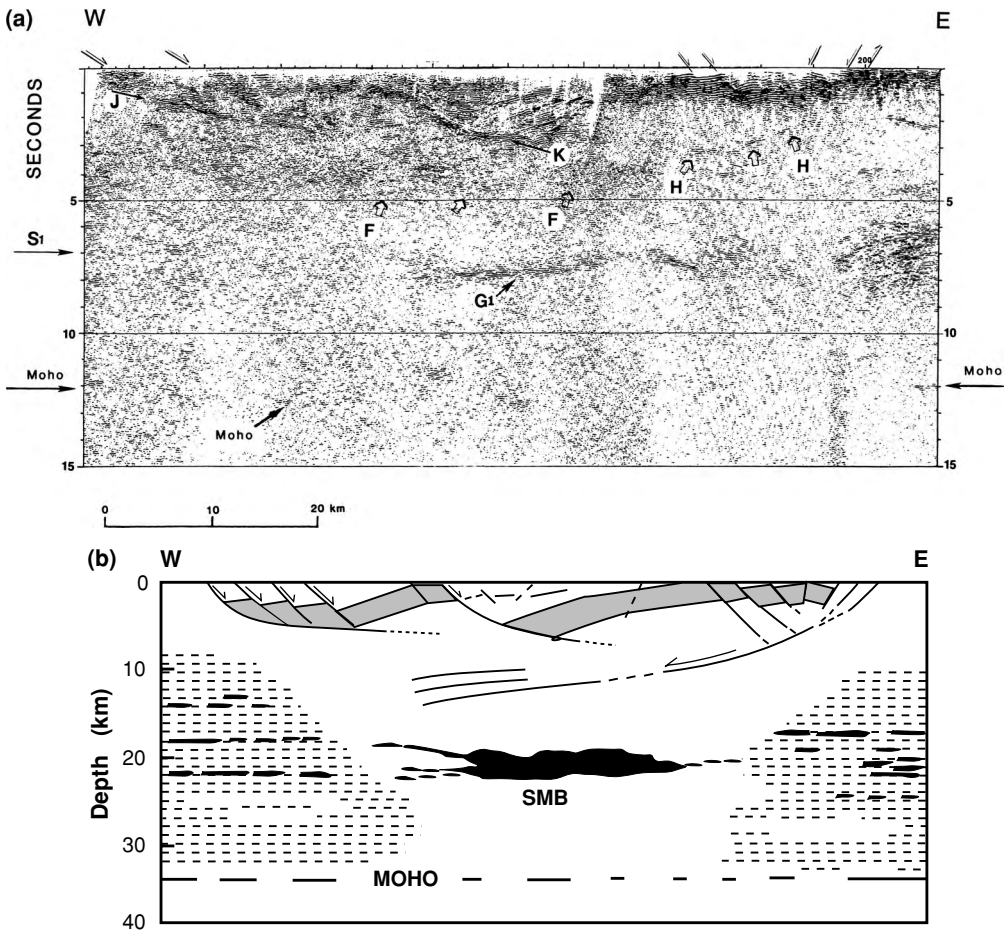


Figure 10.60. (a) COCORP deep-seismic-reflection sections crossing the Rio Grande Rift near Socorro, New Mexico, U.S.A., at 34°N. Layered reflections in the top 2 s include syn-rift deposits, offset by normal faults. F and H are reflections from a major crustal fault. GI is a reflection from the top of the Socorro magma body. (b) Interpretation of the COCORP deep-reflection lines across the Rio Grande Rift. Dotted blocks represent pre-rift sedimentary strata. Note the deep fault that penetrates to mid-crustal depths. Horizontal shading represents the horizontal compositional/deformational fabric of the lower part of the crust. SMB is the Socorro magma body. True scale. (From de Voogt *et al.* (1988).)

Ideally, it is possible to determine the presence of a solid–liquid interface by studying the polarity of its reflections. Unfortunately, it was not possible to determine unequivocally the polarity of reflections recorded on these COCORP lines, but they are consistent with a thin layer of magma in solid material. The complexity of the reflections, however, indicates that the reflector is not a simple continuous sill but may be layered in some way and/or discontinuous.

10.4.4 The Keweenawan Rift system

The Keweenawan or Mid-continent Rift system is a 100-km-wide, 2000-km-long, extinct (~1100 Ma old) rift system extending from Kansas to Michigan in the U.S.A. (see Fig. 10.39). Beneath Lake Superior, the rift bends by 120°; it has been suggested that this is the location of an ancient triple junction. The rift is delineated by the high gravity and magnetic anomalies associated with the thick sequence of basaltic lavas it contains. Seismic-refraction data from the rift

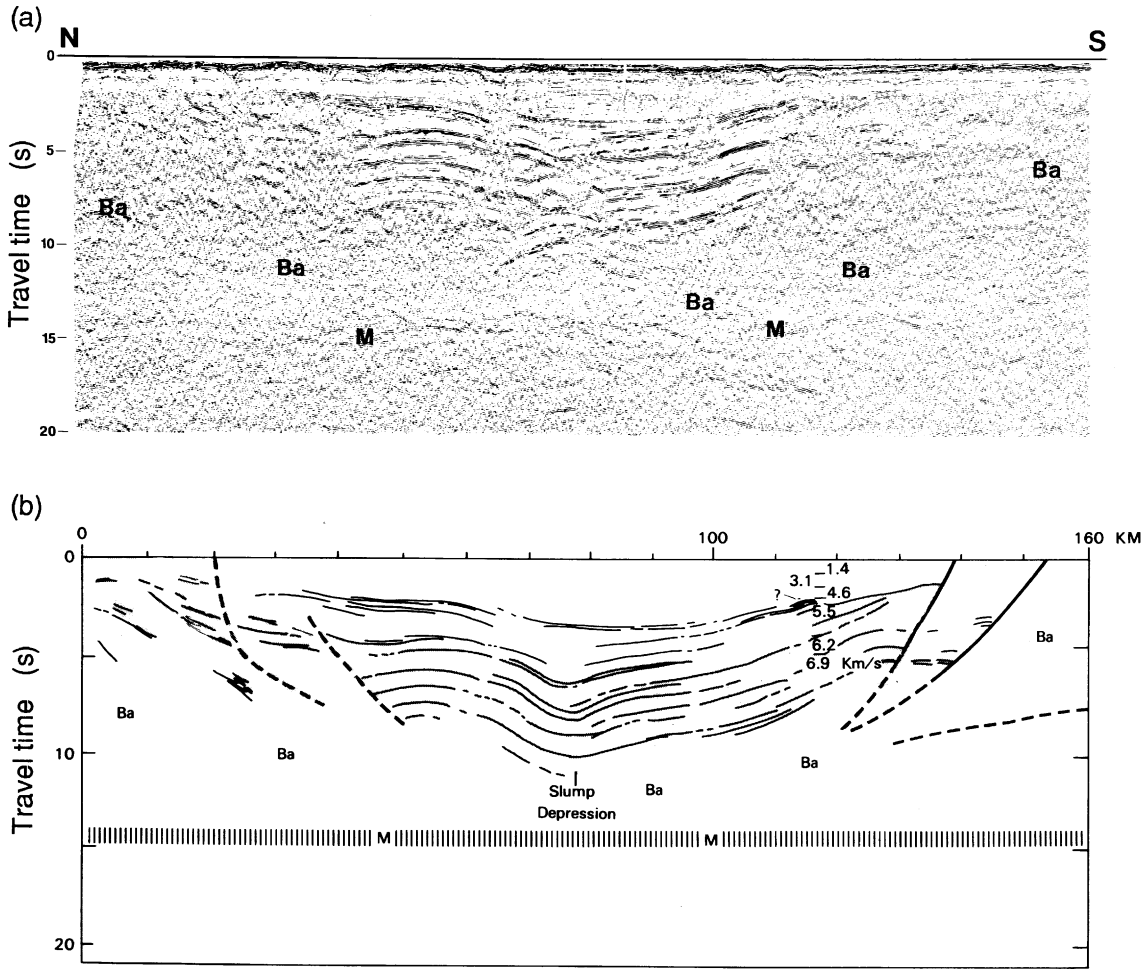


Figure 10.61. (a) An unmigrated seismic-reflection record section from Lake Superior, crossing the Keweenawan Rift, shown to approximately true scale for 6 km s^{-1} material. Ba denotes reflection from pre-rift basement; M, reflection from crust-mantle boundary. (b) A line drawing of the migrated version of the reflection record section shown in (a). Vertical lines indicate the crust-mantle boundary. (From Behrendt *et al.* (1988).)

indicate that the basalt deposits are very thick and that the crust is about 50 km thick beneath the rift, compared with a more typical 35–45 km for neighbouring regions.

A 24–30-fold, deep-seismic-reflection line across this rift in Lake Superior is shown in Fig. 10.61. The rift is very clear indeed. On its northern and southern margins, the rift is bounded by normal faults. The major basin reflectors are believed to be lavas with some interlayered sediments; they extend downwards to almost 10 s two-way time (about 30 km depth). These reflections may originate either from the sediment–lava contacts or from the contacts between lavas of differing compositions. Similar strong reflections (SDR) observed on Atlantic continental margins are thought to be associated with basaltic lavas that were erupted, at elevated temperature or near a hot spot, when the continent split apart.

The reflections labelled Ba are interpreted as the pre-rift basement and those labelled M as the crust–mantle boundary. These mantle reflections, occurring at 13–15 s, indicate that the crust in this region is nearly 50 km thick. In the central part of the rift, the present thickness of crust between the rift deposits and the Moho gives only about 4 s two-way time. This corresponds to a thickness of about 12–14 km, which is about one-third of the normal crustal thickness and therefore about one-third of the pre-rifting crustal thickness. Thus, if the simple stretching model in Section 10.3.610.3.6 is assumed valid, the crust was extended by a factor $\beta = 3$ during the rifting. Such a value implies that complete separation of the crust may have occurred. The assumption that the present M reflector was the ancient as well as the present Moho is, of course, open to debate. The lowermost crust beneath the rift could easily be intrusive material, and the M reflection could be a new post-rifting Moho. In that case, the value of β would be considerably greater than 3; so we can take 3 as a minimum value.

Whatever the final interpretation of the details of the Keweenawan structure, it is a major, old intercontinental rift filled with an incredibly thick sequence of lavas and sediments.

10.5 The Archaean

The Earth's history has four aeons. The Hadean is the time from accretion until about 4 Ga ago (this boundary is as yet undefined). The Archaean is from around 4 Ga ago to 2.5 Ga ago, comprising about one third of the Earth's lifetime. One of the problems facing geologists and geophysicists who are studying the Archaean is that the *uniformitarian* assumption, loosely stated as 'the present holds the key to the past', may be only partly correct. *Aktualism*, 'the present is the same as the past', is certainly not true. The Earth may have behaved very differently in the beginning, with a different tectonic style, so our interpretations of structures, rocks and chemistry may be ambiguous.

When did the continents form? As was mentioned in Section 6.10, the oldest rocks are about 3.8–4.0 Ga old. The oldest known terrestrial material is some

zircon crystals from Western Australia, which have been dated as up to 4.4 Ga old. These crystals are held in a younger (but still early Archaean) meta-sedimentary rock. A handful of zircon does not make a continent, but this material and the 3.8-Ga rocks suggest that some sort of continent was in existence at that time. Isotopic evidence from the zircons suggests that subduction occurred, and deep oceans existed, even in the Hadean. The cratons which form the cores of the present continents are, for the most part, rafts of Archaean granitoids and gneisses, formed in a complex assortment of events from 3.5 to 2.7 Ga. Infolded into the granitoid gneiss cratons are belts of supracrustal lavas and sediment, including komatiitic lavas. These are highly magnesian lavas, formed from melts with up to 29% MgO. Experimental melting has shown that, if dry, such lavas must have erupted at higher temperatures than did modern basalts. Young (less than 100 Ma old) komatiite does occur, with MgO content about 20%, but it is very rare. To produce such hot lavas in abundance, the mantle may have been hotter in the Archaean than it is today. It is possible that in the Hadean and Archaean some plume-derived lavas arrived at the surface at temperatures as high as 1580 °C, implying temperatures of 1800–1900 °C or more at their source.

Various questions can be asked about the Archaean Earth. What was the continental crust like? Could plate tectonics have operated in the Archaean? Was there oceanic crust, and, if so, what was it like?

10.5.1 Archaean continental crust

Two tectonic accidents have resulted in exposures of Archaean crust. The Vredefort Dome in South Africa is a structure some 50 km in diameter in which a section of the Archaean crust aged 3.0–3.8 Ga is exposed. The Dome is thought to have formed at about 2.0 Ga as the result of deformation from within the Earth,

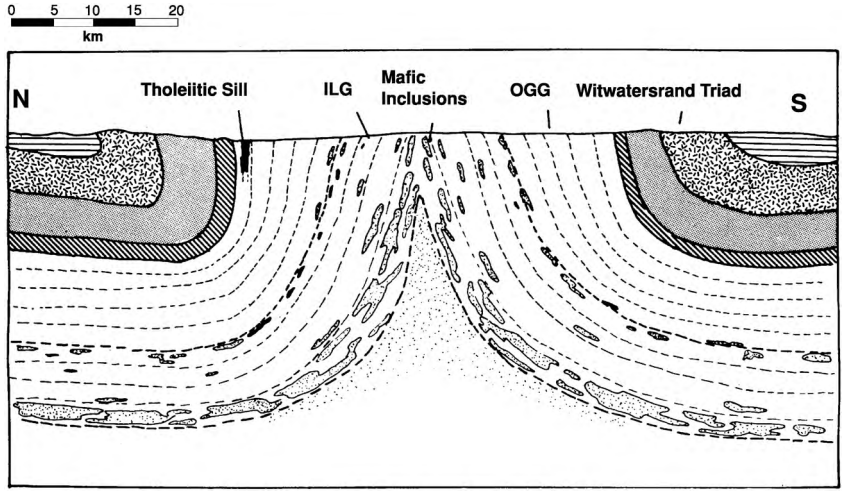


Figure 10.62. A geological cross section across the Vredefort Dome structure in South Africa. OGG, outer granite gneiss; ILG, Inlandsee Leeucogranofels felsic rocks. (From Nicolaysen *et al.* (1981).)

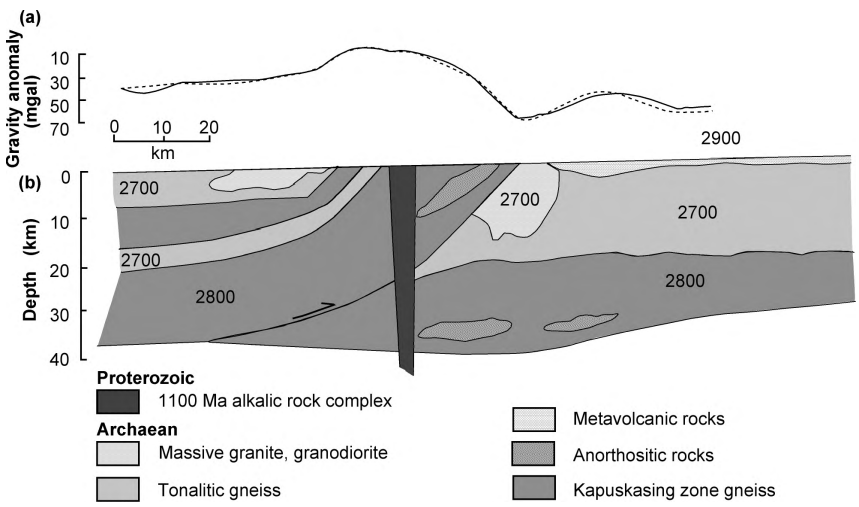


Figure 10.63. (a) Bouguer gravity across the Kapuskasing zone: observed (solid line) and calculated (dashed line). (b) A crustal model based on geology and gravity (densities are in kg m^{-3}). (From Percival *et al.* (1983).)

perhaps an explosive intrusion, though some authors have suggested that it may be a meteorite-impact structure. A cross section (Fig. 10.62) through the Dome shows that the sedimentary layers were underlain by a granite–gneiss upper crust and granulites in the middle crust. Approximately the upper 20 km of the crust is exposed here.

Another exposure of Archaean crust is in the Superior geological province of Canada. Figure 10.63 shows a generalized west–east cross section through the Kapuskasing zone. It appears that, in this case, a major thrust resulted in the uplifting of the deep crustal rocks. The upper crust is granitoid and the lower crust gneiss. The total thickness of exposed crust is about 25 km. This interpretation has been supported by results of deep-seismic-reflection studies.

The total thickness of the continental crust towards the end of the Archaean was probably about 35 km or more, similar to today's value. Problem 9.3 offers an insight into the consequences of this conclusion. The radioactive heat generation in the crust can be estimated by extrapolating backwards in time from the modern content of radiogenic elements. From these estimates, together with a knowledge of the metamorphic facies (and hence temperature and pressure) attained by Archaean rocks, Archaean equilibrium geotherms can be calculated. Some such models are shown in Fig. 10.64. Geotherm 1 implies that the thermal base of the lithosphere (1600–1700 °C) was at about 80 km. However, in North America there is strong isotopic evidence that Archaean diamonds existed and they have even been mined from Archaean conglomerates in South Africa. The stability of diamonds is a major constraint on the thermal structure beneath the continents because they crystallize at about 150 km and 1150 °C. These comparatively low temperatures indicate that the assumptions in the calculation of geotherm 1 need

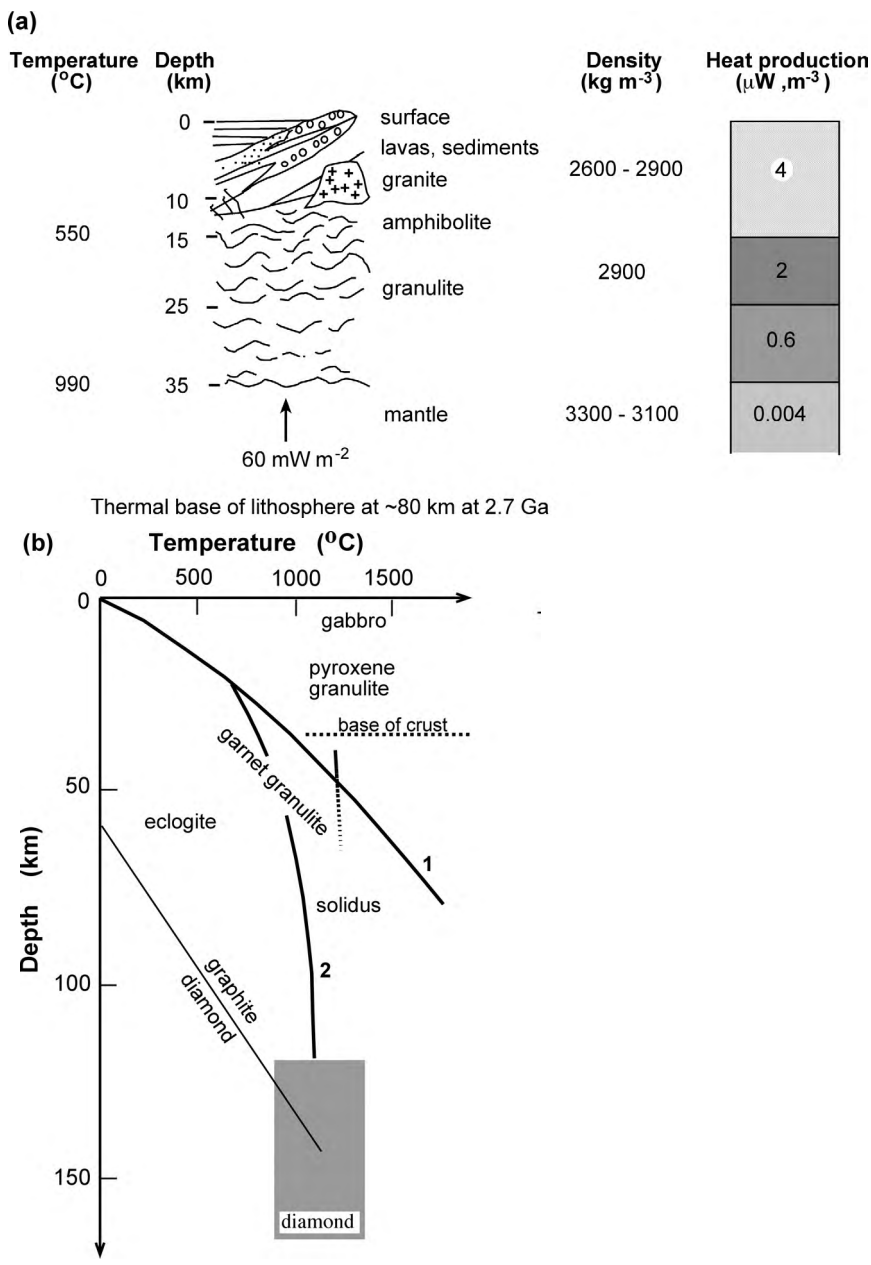


Figure 10.64. (a) A model of 2.8-Ga Archaean continental crust. (b) Model equilibrium Archaean continental geotherms based on (a). For geotherm 1, the heat flow into the base of the crust is 63 mW m^{-2}, and the conductivity is 3.3 W m^{-1} °C^{-1}; this geotherm has $T = 550$ °C at 15 km depth, as determined from the metamorphic assemblages. Geotherm 2 is a possible geotherm in an old cold continent as implied by the existence of Archaean diamonds. The box shows the pressure-temperature field inferred from mineral inclusions in diamonds. The dashed line shows the position of the crustal solidus. (After Nisbet (1984; 1987).)

not have been valid everywhere. Some regions may have been relatively hot, others relatively old and cool (geotherm 2). These temperatures can be interpreted as differences between young (hot) and old (cold) continents. The hot, newly formed continents may have had a lithosphere 80 km or less in thickness, while coexisting colder continental regions may have had a 150–200-km-thick lithosphere.

10.5.2 Archaean tectonics and ocean crust

Thermal models are important in attempts to model Archaean tectonics. For the modern Earth (see Section 7.4), about 65% of the heat loss results from the creation and destruction of plates and about another 17% is from radioactive heat produced in the crust. The heat flow from the mantle into the crust is about 29×10^{-3} W m^{-2}. Most of the heat that is lost comes from the mantle as the Earth cools. The Archaean Earth had much higher rates of radioactive heat generation than does the modern Earth. At 3 Ga, the internal heat production was 2.5–3.0 times its present value (see Table 7.2). It has been shown that, if plate tectonics had not been operating in the Archaean, and if all this heat had been lost from the asthenosphere and had flowed through the lithosphere by conduction, then the equilibrium heat flow at the base of the lithosphere would be roughly 140×10^{-3} W m^{-2}. When this value is used to calculate temperatures in the Archaean continents, geothermal gradients of about 50 °C km^{-1} are obtained. The temperature at the base of the crust would have been high enough to melt it. Indeed, if this model is correct, at 3.5 Ga the heat flow into the base of the lithosphere would have been about 190×10^{-3} W m^{-2}, with a temperature of 800 °C at 10 km depth. These results are contrary to the metamorphic record of deep crustal rocks preserved from the Archaean (Fig. 10.64(b)). The continental crust is clearly self-stabilizing: heat production is moved to the surface by geochemical processes such as partial melting. The problems remain, however, what to do with the heat, and how it was dissipated.

Massive volcanism on a large scale – in other words, spreading centres or mid-ocean ridges – could provide a solution. However, the heat problem is not neatly resolved. To dissipate such large amounts of heat, spreading rates need to be very high, which in turn means, assuming that the Earth did not expand, that the destruction or subduction rates must also have been very high. The dilemma is that young, hot oceanic lithosphere does not subduct easily. What would drive the system? Would it not heat up until a different tectonic pattern was attained? A possible model for Hadean tectonics is that mid-ocean ridges created komatiitic crust. If so, then subduction might have taken place because a komatiitic crust would be denser than a basaltic crust. Such a crust would have been considerably thicker and denser (approximately 15 km and 3.23×10^3 kg m^{-3}) than modern oceanic crust. However, like the modern oceanic crust, the Archaean oceanic crust would probably have had a layered structure with lavas overlying dykes overlying cumulates. A thermal model for such early oceanic lithosphere, which is based on the assumption that lithosphere is cooled mantle, is shown in Fig. 10.65(b). A schematic representation of Archaean plate tectonics is shown in Fig. 10.66.

One unresolved controversy about the structure of the Archaean mantle is particularly interesting because it is in strong contrast to today's mantle. At depths greater than 250 km, it is possible that olivine was less dense than melt. This has

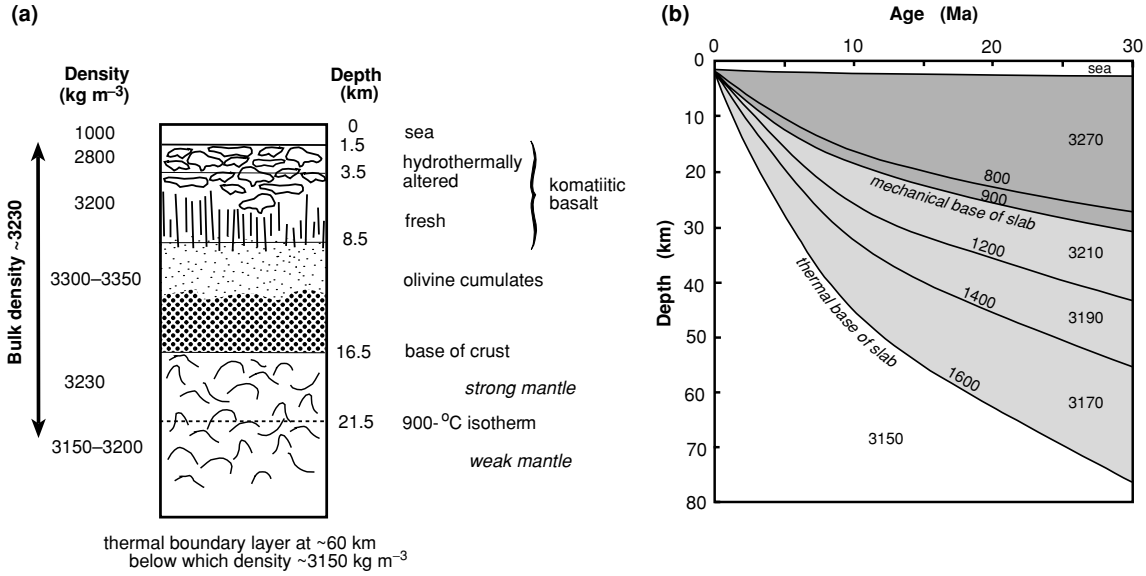


Figure 10.65. (a) Komatiitic Hadean oceanic crust, 15 km thick, assuming an asthenospheric temperature of 1700 °C. Upper crust would be komatiitic basalt with dykes and pillow lavas, and lower crust would be cumulates. (b) A cooling model of the Archaean oceanic lithosphere. The density of Archaean mantle at 1700 °C was assumed to be 3150 kg m^{-3}. The mechanical base of the plate was arbitrarily chosen as 900 °C for refractory mantle. (After Nisbet and Fowler (1983).)

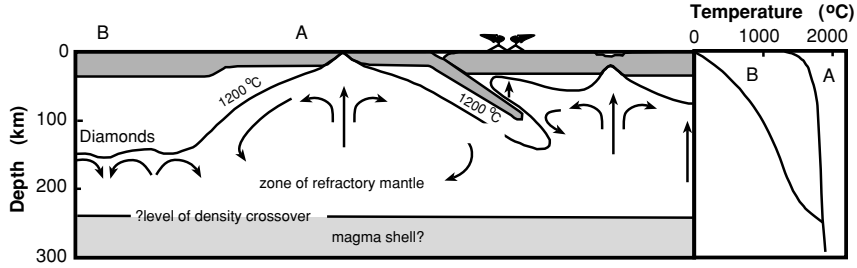


Figure 10.66. A diagram of some speculations about the Archaean upper mantle. Geotherm A at the right-hand side is for mid-ocean ridges; geotherm B is for cool continents. (From Nisbet (1985).)

produced the fascinating speculation that olivine may have floated above a buried magma ocean of melt (on the modern Earth, melts, less dense everywhere than crystal residue, rise). Such a gravitationally stable deep-magma 'ocean' has been nicknamed the LLLAMA (large laterally linked Archaean magma anomaly). If such a density difference existed in the late Hadean to earliest Archaean, a magma shell could have surrounded the Earth, as in Fig. 10.66. It would have been overlain by a layer of olivine. All of this is the subject of debate, but it illustrates how very

different the internal structure of the Archaean Earth may have been. Of course, there is an analogous structure in the modern Earth: the liquid outer core.

Finally, in this discussion of the diversity of continents and their history, it should be noted that the other planets have different tectonics. Even Venus, which is so similar to the Earth, seems to have evolved in quite a different way. However, that subject is planetology, and each planet deserves a book for itself, matching geophysics with aphroditophysics, aresophysics and even plutophysics, puzzles for the next generation of geophysicists.

Problems

1. Calculate the elastic thickness of the subducting Indian plate beneath (a) the Lesser Himalayas and (b) the Greater Himalayas.
2. Fault-plane solutions similar to those shown in Fig. 10.67 were obtained for earthquakes on the North Anatolian Fault in Turkey. Can you give a simple explanation for them?
3. Calculate how long it would take for three-quarters of a given mass of oil to convert to gas at the following temperatures: (a) 160 °C, (b) 180 °C and (c) 220 °C.
4. If oil that had apparently been heated to 160 °C were found, what would you infer about the tectonic setting of the host sediments?
5. How much gas would you expect to find associated with an oil deposit that had been heated to a maximum temperature of 150 °C?
6. (a) Calculate the thickness of sediment with density 2.1×10^3 kg m^{-3} that would be deposited in a subaqueous depression 0.5 km deep.
   (b) What would happen if a Precambrian ironstone sediment with density 4×10^3 kg m^{-3} were deposited in the basin?
7. Assume that the asthenosphere behaves as a viscoelastic material. What is its viscous-relaxation time? (Use 70 GPa for Young's modulus.) Does this value seem reasonable to you?
8. An elastic plate 1000 km long, with a flexural rigidity of 10^{25} N m, is fixed at each end.
   (a) Calculate the critical value of the horizontal compressive force for this plate.
   (b) Calculate the critical stress associated with the compressive force (stress = force per unit area).
   (c) Comment on the magnitude of your answers. What do they indicate about the behaviour of the lithospheric plates?
9. Calculate the critical value of the compressive stress for a 0.5-km-thick rock layer that is isostatically supported by the underlying lithosphere. What is the wavelength of the initial deformation?
10. Calculate the effective elastic thicknesses for the three model lithospheres shown in Fig. 10.35(a). Which one would you intuitively expect to be appropriate for the North American plate?

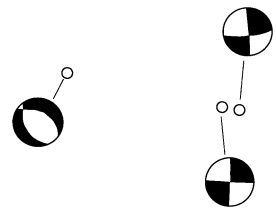


Figure 10.67. Earthquake focal mechanisms on the North Anatolian Fault in Turkey. (After Jackson and McKenzie (1984).)

11. Derive an equation for the initial subsidence of an instantaneously stretched lithosphere, when no water fills the surface depression.
12. Assume that the earthquake shown in Fig. 10.59 took place in the upper crust ($\alpha = 6\,\mathrm{km\,s^{-1}}$, $\beta = 3.5\,\mathrm{km\,s^{-1}}$).
    (a) Use the P- and S-wave arrival times to calculate the distance from the focus to the seismometer.
    (b) What estimates can you make about the height of the focus above the reflecting horizon and the depth of that horizon beneath the surface?
13. What can deep seismic profiling reveal about the structure of the continental lithosphere? (Cambridge University Natural Sciences Tripos II, 1986.)
14. A 35-km-thick continental crust is heated, resulting in an instantaneous increase in temperature of 500 °C.
    (a) Calculate the resulting elevation of the surface.
    (b) Calculate the thickness of sediments which could finally be deposited if 500 m of crust were eroded while the surface was elevated. (Let ρ_s, ρ_c and ρ_m be 2.3×10^3, 2.8×10^3 and $3.3 \times 10^3\,\mathrm{kg\,m^{-3}}$, respectively; α, $3 \times 10^{-5}\,^\circ\mathrm{C}^{-1}$.)
15. Assume that continental lithosphere, thickness 125 km, undergoes instantaneous extension. What is the minimum value of the stretching factor β necessary for asthenospheric material to break through to the surface? For this value of β, what is the total amount of subsidence that would eventually occur? (Let the crustal thickness be 35 km; water, crustal and mantle densities, 1.03×10^3, 2.8×10^3 and $3.35 \times 10^3\,\mathrm{kg\,m^{-3}}$, respectively; coefficient of thermal expansion, $3 \times 10^{-5}\,^\circ\mathrm{C}^{-1}$; asthenosphere temperature, 1350 °C.)
16. Calculate the amount of initial subsidence that would result from instantaneous extension of the continental lithosphere by factors of two and ten. Make reasonable assumptions for the thickness, density and temperature of the crust and mantle.
17. What evidence has been used to confirm the importance of a stretching-and-cooling mechanism in the formation of some sedimentary basins? (Cambridge University Natural Sciences Tripos IB, 1983.)
18. (a) Using the information available to you in this chapter, estimate the value of the stretching factor β for (i) the East African Rift and (ii) the Rio Grande Rift. Assume that both formed as a result of uniform stretching of the lithosphere.
    (b) Using the values of β estimated in (a), calculate initial and final subsidences for these two rifts. Do these values appear reasonable? (Remember that these rifts are not subaqueous; assume that $\rho_w = 0$.)
    (c) Now assume that the sea breaks through in the Gulf of Aden and floods the small portion of the rift valley that is below sea level there. What might happen?
    (d) Now assume that the drainage systems change and the entire East African Rift fills with water. What would happen in this eventuality?

References and bibliography

Ahern, J. L. and Ditmars, R. C. 1985. Rejuvenation of continental lithosphere beneath an intercratonic basin. *Tectonophysics*, **120**, 21–35.

Allenby, R. J. and Schnetzler, C. C. 1983. United States crustal thickness. *Tectonophysics*, **93**, 13–31.

Allègre, C. J. 1985. The evolving Earth system. *Terra Cognita*, **5**, 5–14.

1988. *The Behavior of the Earth*. Cambridge, Massachusetts: Harvard University Press.

Allègre, C. J. *et al*. 1984. Structure and evolution of the Himalaya–Tibet orogenic belt. *Nature*, **307**, 17–22.

Alsdorf, D. and Nelson, D. 1999. Tibetan satellite magnetic low: evidence for widespread melt in the Tibetan crust? *Geology*, **27**, 943–6.

Anderson, R. N., DeLong, S. E. and Schwarz, W. M. 1978. Thermal model for subduction with dehydration in the downgoing slab. *J. Geol.*, **86**, 731–9.

1980. Dehydration, asthenospheric convection and seismicity in subduction zones. *J. Geol.*, **88**, 445–51.

Baker, B. H. and Wohlenberg, J. 1971. Structure and evolution of the Kenya Rift Valley. *Nature*, **229**, 538–42.

Barton, P. J. and Wood, R. 1984. Tectonic evolution of the North Sea basin: crustal stretching and subsidence. *Geophys. J. Roy. Astr. Soc.*, **79**, 987–1022.

Beaumont, C. 1978. The evolution of sedimentary basins on a visco-elastic lithosphere: theory and examples. *Geophys. J. Roy. Astr. Soc.*, **55**, 471–97.

1981. Foreland basins. *Geophys. J. Roy. Astr. Soc.*, **65**, 291–329.

Beaumont, C., Ellis, S., Hamilton, J. and Fullsack, P. 1996. Mechanical model for subduction–collision tectonics of Alpine-type compressional orogens. *Geology*, **24**, 675–8.

Beaumont, C., Ellis, S. and Pfiffner, A. 1999. Dynamics of subduction–accretion at convergent margins: short-term modes, long-term deformation and tectonic implications. *J. Geophys. Res.*, **104**, 17 573–601.

Beaumont, C., Quinlan, G. M. and Hamilton, J. 1987. The Alleghanian orogeny and its relationship to the evolution of the eastern interior, North America. In C. Beaumont and A. J. Tankard, eds., *Sedimentary Basins and Basin-forming Mechanisms*. Vol. 12 of Canadian Society of Petroleum Geologists Memoirs. Calgary: Canadian Society of Petroleum Geologists, pp. 425–46.

Behrendt, J. C., Green, A. G., Cannon, W. F., Hutchinson, D. R., Lee, M. W., Milkereit, B., Agena, W. F. and Spencer, C. 1988. Crustal structure of the mid-continent rift system: results from GLIMPCE deep seismic reflection profiles. *Geology*, **16**, 81–5.

Beloussov, V. V. *et al*. 1980. Structure of the lithosphere along the deep seismic sounding profile: Tien Shan–Pamirs–Karakorum–Himalayas. *Tectonophysics*, **70**, 193–221.

Bickle, M. J. 1978. Heat loss from the Earth: a constraint on Archaean tectonics from the relation between geothermal gradients and the rate of plate production. *Earth Planet. Sci. Lett.*, **40**, 301–15.

Bilham, R. *et al*. 1997. Indo-Asian convergence rates in the Nepal Himalaya. *Nature*, **386**, 61–4.

Bingham, D. K. and Klootwijk, C. J. 1980. Paleomagnetic constraints on Greater India's underthrusting of the Tibetan plateau. *Nature*, **284**, 336–8.

Birt, C. S., Maguire, P. K. H., Kan, M. A., Thybo, H., Keller, G. R. and Patel, J. 1997. The influence of pre-existing structures on the evolution of the southern Kenya rift valley – evidence from seismic and gravity studies. *Tectonophysics*, **278**, 211–42.

Black, P. R. and Braille, L. W. 1982. Pn velocity and cooling of the continental lithosphere. *J. Geophys. Res.*, **87**, 10 557–69.

Bott, M. H. P. 1982. *The Interior of the Earth*, 2nd edn. London: Edward Arnold.

Boultier, R. and Keen, C. E. 1999. Small-scale convection and divergent plate boundaries. *J. Geophys. Res.*, **104**, 7389–403.

Bourne, S. J., England, P. C. and Parsons, B. 1998. The motion of crustal blocks driven by flow of the lower lithosphere and implications for slip rates of continental strike–slip faults. *Nature*, **391**, 655–9.

Braille, L. W. and Smith, R. B. 1975. Guide to the interpretation of crustal refraction profiles. *Geophys. J. Roy. Astr. Soc.*, **40**, 145–76.

Brandon, C. and Romanowicz, B. 1986. A 'no lid' zone in the central Chang-Thang Platform of Tibet: evidence from pure path phase velocity measurements of long-period Rayleigh waves. *J. Geophys. Res.*, **91**, 6547–64.

Brewer, J., Smithson, S., Oliver, J., Kaufman, S. and Brown, L. 1980. The Laramide orogeny: evidence from COCORP deep crustal seismic profiles in the Wind River Mountains, Wyoming. *Tectonophysics.*, **62**, 165–89.

Brown, C. and Girdler, R. W. 1980. Interpretation of African gravity and its implication for the breakup of the continents. *J. Geophys. Res.*, **65**, 287–304.

Brown, L. D. *et al*. 1979. COCORP seismic reflection studies of the Rio Grande Rift. In R. E. Riecker, ed., *Rio Grande Rift: Tectonics and Magmatism.* Special Publication 23 of the American Geophysical Union. Washington: American Geophysical Union, pp. 169–84.

Brown, L. D. *et al*. 1996. Bright spots, structure and magmatism in southern Tibet from INDEPTH seismic reflection profiling. *Science*, **274**, 1688–91.

Burchfiel, B. C. 1983. The continental crust. *Sci. Am.*, **24** (3), 55–67.

Burchfiel, B. C., Oliver, J. E. and Silver, L. T., eds., 1980. *Studies in Geophysics: Continental Tectonics.* Washington: National Academy of Sciences.

Burov, E., Jaupart, C. and Mareschal, J. C. 1998. Large-scale heterogeneities and lithospheric strength in cratons. *Earth Planet. Sci. Lett.*, **164**, 205–19.

Carslaw, H. S. and Jaeger, J. C. 1959. *Conduction of Heat in Solids*, 2nd edn. New York: Oxford University Press.

Cattin, R., Martelet, G., Henry, P., Avouac, J. P., Diament, M. and Shakya, T. R. 2001. Gravity anomalies, crustal structure and thermo-mechanical support of the Himalaya of central Nepal. *Geophys. J. Int*., **147**, 381–92.

Chapin, C. E. 1979. Evolution of the Rio Grande Rift – a summary. In R. E. Riecker, ed., *Rio Grande Rift: Tectonics and Magmatism*. Special Publication 23 of the American Geophysical Union. Washington: American Geophysical Union, pp. 1–6.

Cheadle, M. J., McGeary, S., Warner, M. R. and Matthews, D. H. 1987. Extensional structures on the western U.K. continental shelf: a review of evidence from deep seismic profiling. In M. P. Coward *et al*., eds., *Continental Extension Tectonics*. Geological Society of London. Special Publication **28**. London: Geological Society of London, pp. 445–65.

Chen, W.-P. and Molnar, P. 1981. Constraints on the seismic wave velocity beneath the Tibetan plateau and their tectonic implications. *J. Geophys. Res.*, **86**, 5937–62.

1983. Focal depths of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere. *J. Geophys. Res.*, **88**, 4183–214.

Chu, D. and Gordon, R. 1998. Evidence for motion between Nubia and Somalia along the Southwest Indian Ridge. *Nature*, **398**, 64–7.

Cloetingh, S., Burov, E. and Poliakov, A. 1999. Lithosphere folding: primary response to compression? (from central Asia to Paris basin), *Tectonics*, **18**, 1064–83.

Conrad, V. 1925. Laufzeitkurven des Tauernbebens vom 28. November 1923. *Mitt. Erdb. Komm. Wiener Akad. Wiss.*, **59**, 1–23.

Cordell, L. 1978. Regional geophysical setting of the Rio Grande Rift. *Geol. Soc. Am. Bull.*, **89**, 1073–90.

Davies, J. H. 1999. The role of hydraulic fractures and intermediate-depth earthquakes in generating subduction zone magmatism. *Nature*, **398**, 142–5.

Decker, E. R. and Smithson, S. B. 1975. Heat flow and gravity interpretation in southern New Mexico and West Texas. *J. Geophys. Res.*, **80**, 2542–52.

de Voogt, B., Serpa, L. and Brown, L. 1988. Crustal extension and magmatic processes: COCORP profiles from Death Valley and the Rio Grande Rift. *Geol. Soc. Am. Bull.*, **100**, 1550–67.

Dewey, J. F., Helman, M. L., Turco, E., Hutton, D. H. W. and Knott, S. D. 1989. Kinematics of the western Mediterranean. In M. P. Coward, D. Dietrich and R. G. Park, eds., *Alpine Tectonics*. Geological Society Special Publication 45. London: Geological Society of London, pp. 265–83.

Dietrich, V. J. 1976. Evolution of the eastern Alps: a plate tectonics working hypothesis. *Geology*, **4**, 147–52.

Ebinger, C. and Casey, M. 2001. Continental break-up in magmatic provinces: an Ethiopian example. *Geology*, **29**, 527–9.

Ebinger, C. J. and Sleep, N. H. 1998. Cenozoic magmatism throughout E Africa resulting from impact of a single plume. *Nature*, **395**, 788–91.

Elliott, T., Plank, T., Zindler, A., White, W. and Bourden, B. 1997. Element transport from slab to volcanic front at the Mariana arc. *J. Geophys. Res.*, **102**, 14 991–15 019.

England, P. C. 1983. Constraints on the extension of continental lithosphere. *J. Geophys. Res.*, **88**, 1145–52.

England, P. and Molnar, P. 1997a. The field of crustal velocity in Asia calculated from Quaternary rates of slip on faults. *Geophys. J. Int*., **130**, 551–82.

1997b. Active deformation of Asia: from kinematics to dynamics. *Science*, **278**, 647–50.

Ernst, W. G. 1999. Metamorphism, partial preservation, and exhumation of ultrahigh pressure belts. *The Island Arc*, **8**, 125–53.

Ernst, W. G. and Peacock, S. M. 1996. A thermotectonic model for preservation of ultrahigh-pressure phases in metamorphosed continental crust. In G. E. Bebout, D. W. Scholl, S. H. Kirby and J. P. Platt eds., *Subduction Top to Bottom*. Geophysical Monograph 96. Washington: American Geophysical Union, pp. 171–8.

Foster, A. N. and Jackson J. A. 1998. Source parameters of large African earthquakes: implications for crustal rheology and regional kinematics. *Geophys. J. Int.*, **134**, 422–48.

Fowler, C. M. R., Ebinger, C. J. and Hawkesworth, C. J. eds. 2002. The early Earth: physical, chemical and biological development. *Geol. Soc. Lond. Spec. Publ*., 199.

Fowler, C. M. R. and Nisbet, E. G. 1985. The subsidence of the Williston Basin. *Can. J. Earth Sci.*, **22**, 408–15.

Fowler, S. R., White, R. S. and Louden, K. E. 1985. Sediment dewatering in the Makran accretionary prism. *Earth Planet. Sci. Lett.*, **75**, 427–38.

Giese, P., Nicolich, R. and Reutter, K. J. 1982. Explosion seismic crustal studies in the Alpine–Mediterranean region and their implications to tectonic processes. In H. Berkehemer and K. Hsü, eds., *Alpine–Mediterranean Geodynamics*. Geodynamics Series Vol. 7. Washington: American Geophysical Union; and Boulder, Colorado: Geological Society of America, pp. 39–74.

Giese, P., Prodehl, C. and Stein, A., eds., 1976. *Explosion Seismology in Central Europe*. Berlin: Springer-Verlag.

Goto, K., Suzuki, Z. and Hamaguchi, H. 1987. Stress distribution due to olivine–spinel phase transition in descending plate and deep focus earthquakes. *J. Geophys. Res.*, **92**, 13 811–20.

Hacker, B. D., Ratschbacher, L., Webb, L., McWilliams, M. O., Ireland, T., Calvert, A., Dong, S., Wenk, H.-R. and Chateigner, D. 2000. Exhumation of ultrahigh-pressure continental crust in east central China: Late Triassic–Early Jurassic tectonic unroofing. *J. Geophys. Res.*, **105**, 13 339–64.

Haq, B. U., Hardenbol, J. and Vail, P. R. 1987. Chronology of fluctuating sea levels since the Triassic. *Science*, **235**, 1156–67.

Hayward, N. and Ebinger, C. J. 1996. Variations in the along-axis segmentation of the Afar rift system. *Tectonics*, **15** (2), 244–57.

Hildreth, W. and Moorbath, S. 1988. Crustal contributions to arc magmatism in the Andes. *Contrib. Mineral. Petrol*., **98**, 455–89.

Hoffman, P. F. 1988. United plates of America, the birth of a craton: Early Proterozoic assembly and growth of Laurentia. *Ann. Rev. Earth Planet. Sci.*, **16**, 543–603.

1989. Speculations on Laurentia's first gigayear (2.0 to 1.0 Ga). *Geology*, **17**, 135–8.

Holbrook, W. S., Mooney, W. D. and Christensen, N. I. 1992. The seismic velocity structure of the deep continental crust. In D. M. Fountain, R. Arculus and R. W. Kay, eds., *Continental Lower Crust*. Developments in Geotectonics No. 23. Amsterdam: Elsevier, pp. 1–43.

Holliger, K. and Kissling, E. 1992. Gravity interpretation of a unified 2-D acoustic image of the central Alpine collision zone. *Geophys. J. Int.*, **111**, 213–335.

Houseman, G. A., Mckenzie, D. P. and Molnar, P. 1981. Convective instability of a thickened boundary layer and its relevance for the thermal evolution of continental convergent belts. *J. Geophys. Res.*, **86**, 6115–32.

Hsui, A. T. and Toksöz, M. N. 1979. The evolution of thermal structures beneath a subduction zone. *Tectonophysics*, **60**, 43–60.

Hurley, P. M. and Rand, J. R. 1969. Pre-drift continental nuclei. *Science*, **164**, 1229–42.

Hyndman, R. D. 1988. Dipping seismic reflectors, electrically conductive zones and trapped water in the crust over a conducting plate. *J. Geophys. Res.*, **93**, 13 391–405.

Jackson, J. and McKenzie, D. 1984. Active tectonics of the Alpine–Himalayan belt between western Turkey and Pakistan. *Geophys. J. Roy. Astr. Soc.*, **77**, 185–264.

1988. The relationship between plate motions and seismic moment tensors, and the rates of active deformation in the Mediterranean and Middle East. *Geophys. J.*, **93**, 45–73.

Jarvis, G. T. and McKenzie, D. P. 1980. Sedimentary basin evolution with finite extension rates. *Earth Planet. Sci. Lett.*, **48**, 42–52.

Jeffreys, H. 1926. On near earthquakes. *Mon. Not. Roy. Astron. Soc. Geophys. Suppl.*, **1**, 385.

Jin, Y., McNutt, M. K. and Zhu, Y. S. 1996. Mapping the descent of the Indian and Eurasian plates beneath the Tibetan Plateau from gravity anomalies. *J. Geophys. Res.*, **101**, 11 275–90.

Jobert, N., Journet, B., Jobert, G., Hirn, A. and Sun, K. Z. 1985. Deep structure of southern Tibet inferred from the dispersion of Rayleigh waves through a long period seismic network. *Nature*, **313**, 386–8.

Kahle, H.-G., Cocard, M., Peter, Y., Geiger, A., Reilinger, R., Barka, A. and Veis, G. 2000. GPS-derived strain field rate within the boundary zones of the Eurasian, African and Arabian plates. *J. Geophys. Res.*, **105**, 23 353–70.

Kaila, K. L. *et al*. 1979. Crustal structure along Kavali–Udipi profile in Indian Peninsula shield from deep seismic sounding. *J. Geol. Soc. India*, **20**, 307–33.

Kaila, K. L., Roy Choudhury, K., Krishna, V. G., Dixit, M. M. and Narian, H. 1982. Crustal structure of Kashmir Himalaya and inferences about the asthenosphere layer from DSS studies along the international profile, Qarrakol–Zorkol–Nanga Parbat–Srinagar–Pamir. Himalaya monograph. *Bull. Geofis. Teor. Appl.*, **25**, 221–34.

Kaila, K. L., Tripathi, K. M. and Dixit, M. M. 1984. Crustal structure along Wulan Lake–Gulmarg–Naoshera Profile across Pir Panjal range of the Himalayas from deep seismic soundings. *J. Geol. Soc. India*, **75**, 706–19.

Karner, G. D. and Watts, A. B. 1983. Gravity anomalies and flexure of the lithosphere at mountain ranges. *J. Geophys. Res.*, **88**, 10 449–77.

Kastens, K. *et al*. 1988. ODP Leg 107 in the Tyrrhenian Sea: insights into passive margin and back-arc basin evolution. *Geol Soc. Am. Bull.*, **100**, 1140–56.

Khan, M. A. *et al*. 1999. The lithospheric structure of the Kenya Rift as revealed by wide-angle seismic measurements. In C. MacNioall and P. D. Ryan, eds., *Continental Tectonics*, Geological Society Special Publication **164**. London: Geological Society of London, pp. 257–269.

Keen, C. E. and Cordsen, A. 1981. Crustal structure, seismic stratigraphy and rift processes of the continental margin off eastern Canada: ocean bottom seismic refraction results off Nova Scotia. *Can. J. Earth Sci.*, **18**, 1523–38.

Keleman, P. and Holbrook, S. 1995. Origin of thick, high-velocity igneous crust along the U.S. east coast margin. *J. Geophys. Res.*, **100**, 10 077–94.

Keller, G. R., Braile, L. W. and Schlue, J. W. 1979. Regional crustal structure of the Rio Grande rift from surface wave dispersion measurements. In R. E. Riecker, ed., *Rio Grande Rift*: *Tectonics and Magmatism*. Special Publication 23. Washington: American Geophysical Union, pp. 115–26.

Kirby, S., Engdahl, E. R. and Denlinger, R. 1996. Intermediate-depth earthquakes and arc volcanism as physical expressions of crustal and upper mantle metamorphism in subducting slabs. In G. E. Bebout, D. W. Scholl, S. H. Kirby and J. P. Platt, *Subduction Top to Bottom*. Geophysical Monograph 96. Washington: American Geophysical Union, pp. 195–214.

Klingelé, E. and Oliver, R. 1980. La nouvelle carte gravimétrique de la Suisse. *Beitr. Geol. Karte Schweiz, Ser. Geophys.*, 20.

Knox, R., Nyblade, A. and Langston, C. 1999. Upper mantle S velocities beneath Afar and western Saudi Arabia from Rayleigh wave dispersion. *Geophys. Res. Lett.*, **25**, 4233–6.

Kohlstedt, D. L., Evans, B. and Mackwell, S. J. 1995. Strength of the lithosphere: constraints imposed by laboratory experiments, *J. Geophys. Res.*, **100**, 17 587–602.

Kopp, H., Flueh, E. R., Klaeschen, D., Bielas, J. and Reichert, C. 2001. Crustal structure of the central Sunda margin at the onset of oblique subduction. *Geophys. J. Int.*, **147**, 449–74.

Korenaga, J., Holbrook, S., Kent, G., Keleman, P., Detrick, R., Larsen, H.-C., Hopper, J. and Dahl-Jensen, T. 2000. Crustal structure of the southeast Greenland Margin from joint refraction and reflection seismic tomography. *J. Geophys. Res.* **105**, 21591–614.

Kurtz, R. D., DeLaurier, J. M. and Gupta, J. C. 1986. A magnetotelluric sounding across Vancouver Island detects the subducting Juan de Fuca plate. *Nature*, **321**, 596–9.

Lambeck, K. 1983. Structure and evolution of the intercratonic basins of central Australia. *Geophys. J. Roy. Astr. Soc.*, **74**, 843–86.

Lambeck, K., Burgess, G. and Shaw, R. D. 1988. Teleseismic travel-time anomalies and deep crustal structure in central Australia. *Geophys. J.*, **94**, 105–24.

LASE Study Group 1986. Deep structure of the U.S. east coast passive margin from large aperture seismic experiments (LASE). *Mar. Petrol. Geol.*, **3**, 234–42.

Latham, T. S., Best, J., Chaimov, T., Oliver, J., Brown, L. and Kaufman, S. 1988. COCORP profiles from the Montana plains: the Archean cratonic crust and a lower crustal anomaly beneath the Williston Basin. *Geology*, **16**, 1073–6.

Le Pichon, X. and Sibuet, J. C. 1981. Passive margins, a model of formation. *J. Geophys. Res.*, **86**, 3708–20.

Lèpine, J.-C., Hirn, A., Pandey, M. R. and Tater, J. M. 1984. Features of the P-waves propagated in the crust of the Himalayas. *Ann. Geophys.*, **2**, 119–22.

Lewis, T. J., Bentkowski, W. H., Davis, E. E., Hyndman, R. D., Souther, J. G. and Wright, J. A. 1988. Subduction of the Juan de Fuca plate: thermal consequences. *J. Geophys. Res.*, **93**, 15 207–25.

Lithgow-Bertelloni, C. and Silver, P. G. 1998. Dynamic topography, plate driving forces and the African Superwell. *Nature*, **395**, 269–72.

Lithoprobe, http://www.litho.ucalgary.ca/atlas/index1.html.

Long, R. E. and Backhouse, R. W. 1976. The structure of the western flank of the Gregory Rift. Part II. The mantle. *Geophys. J. Roy. Astr. Soc.*, **44**, 677–88.

Lucas, S. B., Green, A., Hajnal, Z., White, D., Lewry, J., Ashton, K., Weber, W. and Clowes, R. 1993. Crustal architecture of an Early Proterozoic orogen: implications for continental growth. *Nature*, **363**, 339–342.

Lucas, S. B. *et al.* 1994. Three-dimensional collisional structure of the Trans-Hudson orogen, Canada. *Tectonophysics*, **232**, 161–78.

Luetgert, J. H. and Meyer, R. P. 1982. Structure of the western basin of Lake Superior from cross structure refracting profiles. In J. R. Wold and W. J. Hinze, eds., *Geology and Tectonics of the Lake Superior Basin*. Vol. 156 of Geological Society of America Memoirs. Boulder, Colorado, pp. 245–56.

Lyon-Caen, H. 1986. Comparison of the upper mantle shear wave velocity structure of the Indian Shield and the Tibetan Plateau and tectonic implications. *Geophys J. Roy. Astr. Soc.*, **86**, 727–49.

Lyon-Caen, H. and Molnar, P. 1985. Gravity anomalies, flexure of the Indian plate and the structure support and evolution of the Himalaya and Ganga Basin. *Tectonics*, **4**, 513–38.

Makovsky, Y., Klemperer, S. L., Ratschbacher, L. and Alsdorf, D. 1999. Midcrustal reflector on INDEPTH wide-angle profiles: an ophiolitic slab beneath the India–Asia suture in southern Tibet? *Tectonics*, **18**, 793–808.

McDonough, W. F. and Sun, S.-s. 1995. The composition of the Earth. *Chem. Geol.*, **120**, 223–53.

McGeary, S. and Warner, M. 1985. Seismic profiling the continental lithosphere. *Nature*, **317**, 795–7.

McKenzie, A. S. and McKenzie, D. P. 1983. Isomerization and aromatization of hydrocarbons in sedimentary basins formed by extension. *Geol. Mag.*, **120**, 417–528.

McKenzie, D. P. 1969. Speculations on the consequences and causes of plate motions. *Geophys. J. Roy. Astr. Soc.*, **18**, 1–32.

1978. Some remarks on the development of sedimentary basins. *Earth Planet. Sci. Lett.*, **40**, 25–32.

McKenzie, D. P. and Bickle, M. J. 1988. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.*, **29**, 625–79.

McLennan, S. M. and Taylor, S. R. 1996. Heat flow and the chemical composition of continental crust. *J. Geol.*, **104**, 369–77.

Mechie, J., Keller, G. R., Prodehl, C., Kahn, M. A. and Gaciri, S. J. 1997. A model for the structure, composition and evolution of the Kenya Rift. *Tectonophysics*, **278**, 95–119.

Meissner, R. 1986. *The Continental Crust: A Geophysical Approach.* Orlando, Florida: Academic Press.

Menzies, M. A., Klemperer, S. L., Ebinger, C. J. and Baker, J. 2003. *Characteristics of Volcanic Rifted Margins*. Geological Society of America Special Paper 362. Boulder, Colorado: Geological Society of America, pp. 1–14.

Molnar, P. 1984. Structure and tectonics of the Himalaya: constraints and implications of geophysical data. *Ann. Rev. Earth Planet. Sci.*, **12**, 489–518.

Molnar, P. and Chen, W.-P. 1983. Seismicity and mountain building. In K. J. Hsü, ed., *Mountain Building Processes*. London: Academic Press, pp. 41–57.

Molnar, P., England, P. and Martinod, J., 1993. Mantle dynamics, uplift of the Tibetan plateau and the Indian monsoon. *Rev. Geophys.*, **31**, 357–96.

Mooney, W. D., Laske, G. and Masters, T. G. 1998. CRUST 5.1: a global crustal model at $5^\circ \times 5^\circ$. *J. Geophys. Res.*, **103**, 727–47.

Mueller, S. 1977. A new model of the continental crust. In J. G. Heacock, ed., *The Earth's Crust*. Vol. 20 of Geophysical Monographs of the American Geophysical Union. Washington: American Geophysical Union, pp. 289–317.

1983. Deep structure and recent dynamics in the Alps. In K. J. Hsü, ed., *Mountain Building Processes*. London: Academic Press, pp. 181–200.

Mueller, S., Ansorge, J., Egloff, R. and Kissling, E. 1980. A crustal cross-section along the Swiss Geotraverse from the Rhinegraben to the Po Plain. *Eclogae Geol. Helv.*, **73**, 463–83.

Najman, Y., Pringle, M., Godin, L. and Oliver, G. 2001. Dating of the oldest continental sediments from the Himalayan foreland basin. *Nature*, **410**, 194–7.

Nelson, K. D., Baird, D. J., Walters, J. J., Hauck, M., Brown, L. D., Oliver, J. E., Ahern, J. L., Hajnal, Z., Jones, A. G. and Sloss, L. L. 1993. Trans-Hudson orogen and Williston basin in Montana and North Dakota: new COCORP deep-profiling results. *Geology*, **21**, 447–50.

Nelson, K. D. *et al.* 1996. Partially molten middle crust beneath southern Tibet: synthesis of project INDEPTH results. *Science*, **274**, 1684–7.

Ni, J. and Barazangi, M. 1984. Seismotectonics of the Himalayan collision zone: geometry of the underthrusting Indian plate beneath the Himalayas. *J. Geophys. Res.*, **89**, 1147–63.

Nicolaysen, L. O., Hart, R. J. and Gale, N. H. 1981. The Vredefort radioelement profile extended to supracrustal strata at Carltonville, with implications for continental heat flow. *J. Geophys. Res.*, **86**, 10 653–61.

Nisbet, E. G. 1984. The continental and oceanic crust and lithosphere in the Archaean: isostatic, thermal and tectonic models. *Can. J. Earth Sci.*, **21**, 1426–41.

1985. Putting the squeeze on rocks. *Nature*, **315**, **541**.

1986. Archaean mantle models. *Nature*, **320**, 306–7.

1987. *The Young Earth.* London: Allen and Unwin.

Nisbet, E. G. and Fowler, C. M. R. 1983. Model for Archaean plate tectonics. *Geology*, **11**, 376–9.

Nisbet, E. G. and Walker, D. W. 1982. Komatiites and the structure of the Archaean mantle. *Earth Planet. Sci. Lett*, **60**, 105–13.

Nisbet, E. G., Cheadle, M. J., Arndt, N. T. and Bickle, M. J. 1993. Constraining the potential temperature of the Archaean mantle: a review of the evidence from komatiites, *Lithos*, **30**, 291–307.

Nowroozi, A. A. and Mohajer-Ashjai, A. 1985. Fault movements and tectonics of eastern Iran: boundaries of the Lut plate. *Geophys. J. Roy. Astr. Soc.*, **83**, 215–37.

Nyblade, A. A., Owens, T. J., Gurrola, H., Ritsema, J. and Langston, C. A. 2000. Seismic evidence for a deep mantle thermal anomaly beneath east Africa. *Geology*, **28**, 599–602.

Olsen, K. H., Keller, G. R. and Stewart, J. N. 1979. Crustal structure along the Rio Grande Rift from seismic refraction profiles. In R. E. Riecker, ed., *Rio Grande Rift: Tectonics and Magmatism*. Special Publication 23 of the American Geophysical Union. Washington: American Geophysical Union, pp. 127–43.

Owens, T. J. and Zandt, G. 1997. Implications of crustal property variations for models of Tibetan plateau evolution. *Nature*, **387**, 37–42.

Panza, G. F. and Mueller, St. 1979. The plate boundary between Eurasia and Africa in the Alpine area. *Mem. Sci. Geol.*, **33**, 43–50.

Parsons, T., Toda, S., Stein, R. S., Barka, A. and Dieterich, J. H. 2000. Heightened odds of large earthquakes near Istanbul and interaction-based probability calculation. *Science*, **288**, 661–5.

Pavoni, N. 1979. Investigation of recent crustal movements in Switzerland. *Schweiz. Mineral. Petrol. Mitt.*, **59**, 117–26.

Peacock, S. M. 1996. Thermal and petrologic structure of subduction zones. In G. E. Bebout, D. W. Scholl, S. H. Kirby and J. P. Platt, eds., *Subduction Top to Bottom*. Geophysical Monograph 96. Washington: American Geophysical Union, pp. 113–33.

Peacock, S. M. and Wang, K. 1999. Seismic consequences of warm versus cool subduction metamorphism: examples from southwest and northeast Japan, *Science*, **286**, 937–9.

Pearce, J. A. 1983. Role of the sub-continental lithosphere in magma genesis of active continental margins. In C. J. Hawkesworth and M. J. Norry, eds., *Continental Basalts and Mantle Xenoliths*. Nantwich: Shiva, pp. 230–49.

Percival, J. A., Card, K. D., Sage, R. P., Jensen, L. S. and Luhta, L. E. 1983. The Archaean crust in the Wawa–Chapleau–Timmins region. In L. D. Ashwal and K. D. Card, eds., *Workshop on Cross-section of Archaean Crust*. Lunar and Planetary Institute Technical Report 83–03. Houston, Texas: Lunar and Planetary Institute, pp. 99–169.

Pfiffner, O. A. 1992. A continent revealed. In D. J. Blundell, R. Freeman and S. Müller, eds., *The European Geotraverse*. Cambridge: Cambridge University Press, pp. 180–90.

Pfiffner, O. A., Lehner, P., Heitzmann, P., Mueller, S. and Steck, A. 1997. *The Deep Structure of the Swiss Alps*. NRP 20. Basel: Birkhäuser.

Phipps-Morgan, J. and Chen, J. 1993. Dependence of ridge-axis morphology on magma supply and spreading rate. *Nature*, **364**, 706–8.

Pickup, S. L. B., Whitmarsh, R. B., Fowler, C. M. R. and Reston, T. J. 1996. Insight into the nature of the ocean–continent transition off west Iberia from a deep multichannel seismic reflection profile. *Geology*, **24**, 1079–82.

Planke, S., Symonds, P. A., Alvestad, E. and Skogseid, J. 2000. Seismic volcano-stratigraphy of large-volume basaltic extrustive complexes on rifted margins. *J. Geophys. Res.*, **105**, 19 335–51.

Platt, J. P. and England, P. C. 1993. Convective removal of lithosphere beneath mountain belts: thermal and mechanical consequences. *Am. J. Sci.*, **293**, 307–36.

Press, F. 1966. Seismic velocities. In S. P. Clark, ed., *Handbook of Physical Constants*. Vol. 97 of Geological Society of America Memoirs. Boulder, Colorado: Geological Society of America, pp. 195–218.

Quigley, T. M. and McKenzie, A. S. 1988. The temperatures of oil and gas formation in the sub-surface. *Nature*, **333**, 549–52.

Quigley, T. M., McKenzie, A. S. and Gray, J. R. 1987. Kinetic theory of petroleum generation. In B. Doligez, ed., *Migration of Hydrocarbons in Sedimentary Basins*. Paris: Technip, pp. 649–65.

Quinlan, G. 1987. Models of subsidence mechanisms in intracrationic basins and their applicability to North American examples. In C. Beaumont and A. J. Tankard, eds., *Sedimentary Basins and Basin-forming Mechanisms*. Vol. 12 of Canadian Society of Petroleum Geologists Memoirs. Calgary: Canadian Society of Petroleum Geologists, pp. 463–81.

1988. The thermal signatures of tectonic environments. In E. G. Nisbet and C. M. R. Fowler, eds., *Heat, Metamorphism and Tectonics*. Short Course Vol. 14. Saskatoon: Mineralogical Association of Canada. 213–57.

Quinlan, G. H. and Beaumont, C. 1984. Appalachian thrusting, lithospheric flexure, and the Paleozoic stratigraphy of the Eastern Interior of North America. *Can. J. Earth Sci.*, **21**, 973–96.

Ramberg, I. B., Cook, F. A. and Smithson, S. B. 1978. Structure of the Rio Grande Rift in southern New Mexico and west Texas based on gravity interpretation. *Geol. Soc. Am. Bull.*, **89**, 107–23.

Ratschbacher, L., Frisch, W., Liu, G. and Chen, C. 1994. Distributed deformation in southern and western Tibet during and after the Indo-Asia collision. *J. Geophys. Res.*, **99**, 19 917–45.

Reilinger, R. E., McClusky, S. C., Oral, M. B., King, R. W., Toksöz, M. N., Barka, A. A., Kinik, I., Lenk, O. and Sanli, I. 1997. Global positioning system measurements of present-day crustal movements in the Arabia–Africa–Eurasia plate collision zone. *J. Geophys. Res.*, **102**, 9983–99.

Rinehart, E. J., Sanford, A. R. and Ward, R. M. 1979. Geographic extent and shape of an extensive magma body at mid crustal depths in the Rio Grande Rift near Socorro, New Mexico. In R. E. Riecker, ed., *Rio Grande Rift: Tectonics and Magmatism*. Special Publication 23 of the American Geophysical Union. Washington: American Geophysical Union, pp. 237–51.

Ritsema, J. and van Heijst, H. 2000. New seismic model of the upper mantle beneath Africa. *Geology*, **28**, 63–6.

Rosendahl, G. R. 1987. Architecture of continental rifts with special reference to East Africa. *Ann. Rev. Earth Planet Sci.*, **15**, 445–503.

Rudnick, R. L. and Fountain, D. M. 1995. Nature and composition of the continental crust: a lower crustal perspective. *Rev. Geophys.*, **33**, 267–309.

Rybach, L., Mueller, St., Milnes, A. G., Ansorge, J., Bernoulli, D. and Frey, M. 1980. The Swiss Geotraverse Basel–Chiasso – a review. *Eclogae Geol. Helv.*, **73**, 437–62.

Sanford, A. R. 1983. Magma bodies in the Rio Grande Rift in central New Mexico. In *34th Field Conference Guidebook*. New Mexico: New Mexico Geological Society, pp. 123–5.

Sanford, A. R., Mott Jr, R. P., Shuleski, P. J., Rinehart, E. J., Caravella, F. J., Ward, R. M. and Wallace, T. C. 1977. Geophysical evidence for a magma body in the crust in the vicinity of Socorro, N.M. In J. G. Heacock, ed., *The Earth's Crust*. Vol. 20 of American Geophysical Union Monographs. Washington: American Geophysical Union, pp. 385–403.

Schelling, D. and Arita, K. 1991. Thrust tectonics, crustal shortening and the structure of the far-eastern Nepal Himalaya. *Tectonics*, **10**, 851–62.

Sclater, J. G., Parson, B. and Jaupart, C. 1981. Oceans and continents: similarities and differences in the mechanisms of heat loss. *J. Geophys. Res.*, **86**, 11 535–52.

Seager, W. R. and Morgan, P. 1979. Rio Grande Rift in southern New Mexico, West Texas and northern Chihuahua. In R. E. Riecker, ed., *Rio Grande Rift: Tectonics and Magmatism*. Special Publication 23 of the American Geophysical Union. Washington: American Geophysical Union, pp. 87–106.

Searle, R. C. 1970. Evidence from gravity anomalies for thinning of the lithosphere beneath the Rift Valley in Kenya. *Geophys. J. Roy. Astr. Soc.*, **21**, 13–21

Searle, R. C. and Gouin, P. 1972. A gravity survey of the central part of the Ethopian Rift Valley. *Tectonophysics*, **15**, 41–52.

Shatsky, V. S., Jagoutz, E., Sobolev, N. V., Kozmenko, O. A., Parkhomenko, V. S. and Troesch, M. 1999. Geochemistry and age of ultrahigh pressure metamorphic rocks from the Kokchetav massif (Northern Kazakhstan). *Contrib. Mineral. Petrol.*, **137**, 185–205.

Shen, Z.-K., Wang, M., Li, Y., Jackson, D. D., Yin, A., Dong, D. and Fang, P. 2001. Crustal deformation along the Altyn Tagh fault system western China from GPS. *J. Geophys. Res.*, **106**, 10 607–21.

Sibson, R. 1992. Implications of fault-valve behaviour for rupture, nucleation and recurrence. *Tectonophysics*, **211**, 283–93.

Sinno, Y. A., Daggett, P. H., Keller, G. R., Morgan, P. and Harder, S. H. 1986. Crustal structure of the southern Rio Grande Rift determined from seismic refraction profiling. *J. Geophys. Res.*, **91**, 6143–56.

Slack, P. D., Davis, P. M., Baldridge, W. S., Olsen, K. H., Glahn, A., Achauer, U. and Spence, W. 1996. The upper mantle structure of the central Rio Grande rift region from teleseismic P and S wave travel time delays and attenuation. *J. Geophys. Res.*, **101**, 16 003–23.

Sleep, N. H. 1971. Thermal effects of the formation of Atlantic continental margins by continental break-up. *Geophys. J. Roy. Astr. Soc.*, **24**, 325–50.

1976. Platform Subsidence Mechanisms and 'eustatic' sea-level changes. *Tectonophysics*, **36**, 45–56.

Sleep, N. H., Nunn, J. A. and Chou, L. 1980. Platform basins. *Ann. Rev. Earth Planet. Sci.*, **8**, 17–34.

Sleep, N. H. and Sloss, L. L. 1978. A deep borehole in the Michigan Basin. *J. Geophys. Res.*, **83**, 5815–19.

Sleep, N. H. and Snell, N. S. 1976. Thermal contraction and flexure of mid-continent and Atlantic marginal basins. *Geophys. J. Roy. Astr. Soc.*, **45**, 125–54.

Smith, A. G. and Woodcock, N. H. 1982. Tectonic synthesis of the Alpine–Mediterranean region: a review. In H. Berkehemer and K. Hsü, eds., *Alpine–Mediterranean*

Geodynamics. Vol. 7 of Geodynamics Series. Washington: American Geophysical Union; and Boulder, Colorado: Geological Society of America, pp. 15–38.

Steckler, M. S. and Watts, A. B. 1978. Subsidence of the Atlantic type margin off New York. *Earth Planet. Sci. Lett.*, **41**, 1–13.

Stein, S. R., Barka, A. and Dieterich, H. J. 1997. Progressive failure on the North Anatolian Fault since 1939 by earthquake stress triggering. *Geophys. J. Int.*, **128**, 594–604.

Stockmal, G. S. 1983. Modelling of large-scale accretionary wedge deformation. *J. Geophys. Res.*, **88**, 8271–87.

Stockmal, G. S. and Beaumont, C. 1987. Geodynamic models of convergent margin tectonics: the southern Canadian Cordillera and the Swiss Alps. In C. Beaumont and A. J. Tankard, eds., *Sedimentary Basins and Basin-forming Mechanisms*. Vol. 12 of Canadian Society of Petroleum Geologists Memoirs. Calgary: Canadian Society of Petroleum Geologists, pp. 393–412.

Tapponier, P., Peltzer, G., Le Dain, A. Y., Armijo, R. and Cobbold, P. 1982. Propagating extrusion tectonics in Asia: new insights from simple experiments with plasticine. *Geology*, **10**, 611–16.

Tapponier, P. *et al.* 1981. The Tibetan side of the India–Eurasia collision. *Nature*, **294**, 404–10.

Tatsumi, Y., Sakuyama, M. Fukuyama, H. and Kushiro, I. 1983. Generation of arc basalt magmas and thermal structure of the mantle wedge in subduction zones. *J. Geophys. Res.*, **88**, 5815–25.

Taylor, S. R. and McLennan, S. M. 1985. *The Continental Crust: Its Composition and Evolution*. Oxford: Blackwell Scientific.

1996. The evolution of continental crust. *Sci. Am.*, January, 76–81.

Turcotte, D. L. and Schubert, G. 2002. *Geodynamics*, 2nd edn. Cambridge: Cambridge University Press.

Turner, F. J. 1981. *Metamorphic Petrology: Mineralogical, Field and Tectonics Aspects.* New York: McGraw-Hill.

Vail, P. R. and Mitchum Jr, R. M. 1979. Global cycles of relative changes of sea-level from seismic stratigraphy. In J. S. Watkins, L. Montadert and P. W. Dickerson, eds., *Geological and Geophysical Investigations of Continental Margins*. American Association of Petroleum Geologists Memoir 29. Tulsa, Oklahoma: American Association of Petroleum Geologists.

Van den Beukel, J. and Wortel, R. 1987. Temperatures and shear stresses on the upper part of a subduction zone. *Geophys. Res. Lett.*, **14**, 1057–60.

1988. Thermomechanical modelling of arc–trench regions. *Tectonophysics*, **154**, 177–93.

Van der Voo, R., Spakman, W. and Bijwaard, H. 1999. Tethyan subducted slabs under India, *Earth Planet. Sci. Lett.*, **171**, 7–20.

Van Schmus, W. R. and Hinze, W. J. 1985. The mid-continent rift system. *Ann Rev. Earth Planet. Sci.*, **13**, 345–83.

Waldhauser, F., Kissling, E., Ansorge, J. and Mueller, St. 1998. Three-dimensional interface modelling with two-dimensional seismic data: the Alpine crust–mantle boundary. *Geophys. J. Int.*, **135**, 264–78.

Walpes, D. W. 1980. Time and temperature in petroleum formation: application of Lopatin's method to petroleum exploration. *Bull. Am. Assoc. Petrol. Geol.*, **64**, 916–26.

Warner, M. and McGeary, S. 1987. Seismic reflection coefficients from mantle fault zones. *Geophys. J. Roy. Astr. Soc.*, **89**, 223–30.

Watts, A. B. 2001. *Isostasy and Flexure of the Lithosphere*. Cambridge: Cambridge University Press.

Watts, A. B. and Ryan, W. B. F. 1976. Flexure of the lithosphere and continental margin basins. *Tectonophysics*, **36**, 25–44.

Wernicke, B. 1985. Uniform-sense simple shear of the continental lithosphere. *Can. J. Earth Sci.*, **22**, 108–25.

Wernicke, B. and Axen, G. J. 1988. On the role of isostasy in the evolution of normal fault systems. *Geology*, **16**, 848–51.

White, R. S. 1984. Active and passive plate boundaries around the Gulf of Oman, north-west Indian Ocean. *Deep Sea Res.*, **31**, 731–45.

White, R. S. and Louden, K. E. 1982. The Makran continental margin: structure of a thickly sedimented convergent plate boundary. In J. S. Watkins and C. L. Drake, eds., *Studies in Continental Margin Geology*. Vol. 34 of American Association of Petroleum Geologists Memoirs. Tulsa, Oklahoma: American Association of Petroleum Geologists, pp. 499–518.

White, R. S. and McKenzie, D. P. 1989. Magmatism at rift zones: the generation of volcanic continental margins and flood basalts. *J. Geophys. Res.*, **94**, 7685–729.

White, R. S., Spence, G. D., Fowler, S. R., McKenzie, D. P., Westbrook, G. K. and Bowen, A. N. 1987. Magmatism at rifted continental margins. *Nature*, **330**, 439–44.

White, R. S., Westbrook, G. K., Fowler, S. R., Spence, G. D., Barton, P. J., Joppen, M., Morgan, J., Bowen, A. N., Prestcott, C. and Bott, M. H. P. 1987. Hatton Bank (Northwest U.K.) continental margin structure. *Geophys. J. Roy. Astr. Soc.*, **89**, 265–72.

Whitmarsh, R. B., Manatschal, G. and Minshull, T. 2001. Evolution of magma-poor continental margins from rifting to seafloor spreading. *Nature*, **413**, 150–4.

Windley, B. F. 1995. *The Evolving Continents*, 2nd edn. New York: Wiley.

Wood, R. and Barton, P. J. 1983. Crustal thinning and subsidence in the North Sea. *Nature*, **302**, 134–6.

Wyllie, P. J. 1979. Magmas and volatile components. *Am. Mineral.*, **64**, 469–500.

1981. Experimental petrology of subduction, andesites and batholiths. *Trans. Geol. Soc. S. Afr.*, **84**, 281–91.

Wyss, M., Hasegawa, A. and Nakajima, J. 1999. Source and path of magma for volcanoes in the subduction zone of northeastern Japan, *Geophys. Res. Lett.*, **28**, 1819–22.

Yogodzinski, G. M., Lees, J. M., Churikova, T. G., Dorendorf, F., Wöerner, G. and Volynets, O. N. 2001. Geochemical evidence for the melting of subducting oceanic lithosphere at plate edges. *Nature*, **409**, 500–4.

Zandt, G., Gilbert, H., Owens, T. J., Ducea, M., Saleeby, J. and Jones, C. H. 2004. Active foundering of a continental arc root beneath the southern Sierra Nevada in California, *Nature*, **431**, 41–6.

Zhao, W., Nelson, K. D. and Project INDEPTH Team 1993. Deep seismic reflection evidence for continental underthrusting beneath southern Tibet. *Nature*, **366**, 557–9.

Zhao, W. *et al.* 2001. Crustal structure of Tibet as derived from project INDEPTH wide-angle seismic data. *Geophys. J. Int.*, **145**, 486–98.

Appendix 1

Scalars, vectors and differential operators

Scalars and vectors

A scalar is a quantity that just has a magnitude. For example, the temperature outside today could be $+10\,^{\circ}\mathrm{C}$.

A vector is a quantity that has a magnitude and a direction. For example, the wind velocity in your city today could be 20 km hr^{-1} and due east. The speed (magnitude of the velocity) would be measured by an anemometer and the direction by a wind vane.

A vector is indicated in print by a boldface character such as **x**. Its magnitude is indicated by the same character in italic type, x, as is a scalar. A vector or a scalar can be either a constant or a function of some variable, which can itself be either a scalar or a vector. When the scalar (or vector) is a function of a variable, it is called a scalar (or vector) field. For example, the temperature at midday across the province of British Columbia, Canada, is a scalar field. That is, the temperature at each place depends on its position in the province and thus is written $T(x, y, z)$, where x, y and z are geographic and height coordinates within British Columbia. The wind velocity **V** across British Columbia at midday depends on geographic position and so is a vector field, written $\mathbf{V}(x, y, z)$. Note that, if the coordinate system is changed, the scalar is unaffected, but the components of the vector must be recalculated.

Products of scalars and vectors

Many physical relationships are best expressed by using the products of scalars and vectors. The product of two scalars is another scalar, and everyone is well accustomed to the process called multiplication, learned laboriously in elementary school. The product of a scalar s and a vector $\mathbf{V} = (V_x, V_y, V_z)$ is another vector, $s\mathbf{V}$. In Cartesian coordinates (x, y, z), the product is simply

$$s\mathbf{V} = s(V_x, V_x, V_z) = (sV_x, sV_y, sV_z) \tag{A1.1}$$

Thus, the scalar multiplies each component of the vector.

When two vectors are involved, multiplication becomes more complicated. There are two products of vectors: one, called the *scalar product*, is a scalar; the other, called the *vector product*, is a vector.

The scalar product of two vectors **U** and **V** is written $\mathbf{U} \cdot \mathbf{V}$ and defined as

$$\mathbf{U} \cdot \mathbf{V} = UV \cos\theta \tag{A1.2}$$

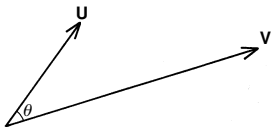


Figure A1.1.

where θ is the angle between the two vectors and U and V are the magnitudes of the vectors (Fig. A1.1). The scalar product is also known as the dot product. If **U** and **V** are parallel, then $\theta = 0$, $\cos\theta = 1$, and so $\mathbf{U} \cdot \mathbf{V} = 1$. However, if **U** and **V** are perpendicular, then $\theta = 90^\circ$ and $\mathbf{U} \cdot \mathbf{V} = 0$. Thus, the scalar product of two perpendicular vectors is zero. In Cartesian coordinates (x, y, z), the scalar product is

$$\mathbf{U} \cdot \mathbf{V} = U_x V_x + U_y V_y + U_z V_z \tag{A1.3}$$

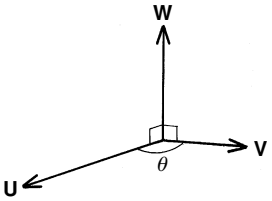


Figure A1.2.

As an example of a scalar product, consider a force **F** acting on a mass m and moving that mass a distance **d**. The work done is then $\mathbf{F} \cdot \mathbf{d}$: work is a scalar.

The vector product of two vectors **U** and **V** is written $\mathbf{U} \wedge \mathbf{V}$ or $\mathbf{U} \times \mathbf{V}$ and defined as

$$\mathbf{U} \wedge \mathbf{V} = \mathbf{W} \tag{A1.4}$$

where **W** is a vector perpendicular both to **U** and to **V** (Fig. A1.2), with magnitude

$$W = UV \sin\theta \tag{A1.5}$$

The vector product is also known as the *cross product*. The vector product of two parallel vectors is zero since $\sin\theta = 0$ when $\theta = 0$. In Cartesian coordinates (x, y, z), the vector product is expressed as

$$\mathbf{U} \wedge \mathbf{V} = (U_y V_z - U_z V_y, U_z V_x - U_x V_z, U_x V_y - U_y V_x) \tag{A1.6}$$

As an example of a vector product, consider a rigid body rotating about an axis with angular velocity $\boldsymbol{\omega}$ (the Earth spinning about its north–south axis if you like). The velocity **v** of any particle at a radial position **r** is then given by

$$\mathbf{V} = \boldsymbol{\omega} \wedge \mathbf{r} \tag{A1.7}$$

Compare this with Eq. (2.3); rotation of the plates also involves the vector product.

Gradient

The gradient of a scalar T is a vector that describes the rate of change of T. The component of grad T in any direction is the rate of change of T in that direction. Thus, the x component is $\partial T/\partial x$, the y component is $\partial T/\partial y$, and the z component is $\partial T/\partial z$. grad T is an abbreviation for 'the gradient of T':

$$\text{grad}\, T \equiv \nabla T \tag{A1.8}$$

defines the vector operator ∇. The notations grad T and ∇T are equivalent and are used interchangeably. In Cartesian coordinates (x, y, z), ∇T is given by

$$\nabla T = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right) \tag{A1.9}$$

grad T is normal (perpendicular) to surfaces of constant T. To show this, consider the temperature T at point (x, y, z). A small distance $\delta\mathbf{r} = (\delta x, \delta y, \delta z)$ away, the temperature is $T + \delta T$, where

$$\begin{aligned} \delta T &= \frac{\partial T}{\partial x}\,\delta x + \frac{\partial T}{\partial y}\,\delta y + \frac{\partial T}{\partial z}\,\delta z \\ &= (\nabla T) \cdot \delta\mathbf{r} \end{aligned} \tag{A1.10}$$

On the surface $T =$ constant, $\delta T = 0$. This means that the scalar product $(\Delta T) \cdot \delta\mathbf{r}$ is zero on a surface of constant T and hence that ∇T and $\delta\mathbf{r}$ are perpendicular. Since $\delta\mathbf{r}$ is parallel to the surface and $T =$ constant, ∇T must be perpendicular, or normal, to that surface.

Divergence

The divergence of a vector field $\mathbf{V}$, div $\mathbf{V}$, is a scalar field. It is written $\boldsymbol{\nabla} \cdot \mathbf{V}$. and is defined as

$$\text{div}\,\mathbf{V} \equiv \boldsymbol{\nabla} \cdot \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \tag{A1.11}$$

where the components of $\mathbf{V}$ in Cartesian coordinates are (V_x, V_y, V_z). The divergence represents a net flux, or rate of transfer, per unit of volume. If the wind velocity is $\mathbf{V}$ and the air has a constant density ρ, then

$$\boldsymbol{\nabla} \cdot (\rho\,\mathbf{V}) = \rho\,\boldsymbol{\nabla} \cdot \mathbf{V} \tag{A1.12}$$

represents the net mass flux of air per unit volume. If no air is created or destroyed, then the total mass flux entering each unit volume is balanced by that leaving it, so the net mass flux is zero:

$$\rho\,\boldsymbol{\nabla} \cdot \mathbf{V} = 0 \tag{A1.13}$$

A vector field for which $\boldsymbol{\nabla} \cdot \mathbf{V} = 0$ is called *solenoidal*.

Curl

The curl of a vector field $\mathbf{V}$, curl $\mathbf{V}$, is a vector function of position. It is written $\boldsymbol{\nabla} \wedge \mathbf{V}$, or $\boldsymbol{\nabla} \times \mathbf{V}$, and is defined in Cartesian coordinates as

$$\text{curl}\,\mathbf{V} \equiv \boldsymbol{\nabla} \wedge \mathbf{V} = \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}, \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}, \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) \tag{A1.14}$$

It is related to rotation and is sometimes called *rotation*, or rot. For example, the differential expression of *Ampère's law* for the magnetic field $\mathbf{H}$ due to a current $\mathbf{J}$ is $\boldsymbol{\nabla} \wedge \mathbf{H} = \mathbf{J}$. Alternatively, consider a body rotating with constant angular velocity $\boldsymbol{\omega}$. Equation (A1.7) expresses the velocity at $\mathbf{r}$ in terms of the angular velocity:

$$\mathbf{V} = \boldsymbol{\omega} \wedge \mathbf{r}$$

Now, take the curl of $\mathbf{V}$:

$$\boldsymbol{\nabla} \wedge \mathbf{V} = \boldsymbol{\nabla} \wedge (\boldsymbol{\omega} \wedge \mathbf{r}) \tag{A1.15}$$

Since ω is a constant, this equation can be simplified to

$$\begin{aligned} \boldsymbol{\nabla} \wedge \mathbf{V} &= \boldsymbol{\omega}(\boldsymbol{\nabla} \cdot \mathbf{r}) - (\boldsymbol{\omega} \cdot \boldsymbol{\nabla})\mathbf{r} \\ &= 3\boldsymbol{\omega} - \boldsymbol{\omega} \\ &= 2\boldsymbol{\omega} \end{aligned} \tag{A1.16}$$

Thus, the curl of the velocity is twice the angular velocity. A vector field for which $\nabla \wedge \mathbf{V} = 0$ is called *irrotational*.

The Laplacian operator

In Cartesian coordinates, the Laplacian operator ∇^2 is defined by

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{A1.17}$$

which is the divergence of the gradient. It is a scalar operator:

$$\nabla^2 T = \nabla \cdot \nabla T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \tag{A1.18}$$

To define a Laplacian operator ∇^2 for a vector, it is necessary to use the identity

$$\nabla \cdot (\nabla \mathbf{V}) = \nabla(\nabla \cdot \mathbf{V}) - \nabla \wedge (\nabla \wedge \mathbf{V}) \tag{A1.19}$$

In Cartesian coordinates, this is the same as applying the Laplacian operator to each component of the vector in turn:

$$\nabla^2 \mathbf{V} = (\nabla^2 V_x, \nabla^2 V_y, \nabla^2 V_z) \tag{A1.20}$$

However, in curvilinear coordinate systems this is not true because, unlike the unit vectors (1, 0, 0), (0, 1, 0) and (0, 0, 1) in the Cartesian coordinate system, those in curvilinear coordinate systems are not constants with respect to their coordinate system. The calculation of the Laplacian operator applied to a vector in cylindrical and spherical polar coordinates is long and is left to the reader as an extracurricular midnight activity. (Hint: use Eqs. (A1.19), (A1.22)–(A1.24) and (A1.28)–(A1.30).)

Curvilinear coordinates

In geophysics it is frequently advantageous to work in curvilinear instead of Cartesian coordinates. The curvilinear coordinates which exploit the symmetry of the Earth, and are thus the most often used, are *cylindrical polar coordinates* and *spherical polar coordinates*. Although not every gradient, divergence, curl and Laplacian operator is used in this book in each of these coordinate systems, all are included here for completeness.

Cylindrical polar coordinates (r, ϕ, z)

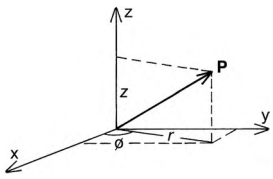


Figure A1.3.

In cylindrical polar coordinates (Fig. A1.3), a point P is located by specifying r, the radius of the cylinder on which it lies, ϕ, the longitude or azimuth in the x–y plane, and z, the distance from the x–y plane to the point P, where $r \geq 0$, $0 \leq \phi \leq 2\pi$ and $-\infty < z < \infty$. From Fig. A1.3 it can be seen that

$$\begin{aligned} x &= r\cos\phi \\ y &= r\sin\phi \\ z &= z \end{aligned} \tag{A1.21}$$

In these cylindrical polar coordinates (r, ϕ, z), the gradient, divergence, curl and Laplacian operators are

$$\nabla T = \left(\frac{\partial T}{\partial r}, \frac{1}{r}\frac{\partial T}{\partial \phi}, \frac{\partial T}{\partial z}\right) \tag{A1.22}$$

$$\boldsymbol{\nabla} \cdot \mathbf{V} = \frac{1}{r}\frac{\partial}{\partial r}(rV_r) + \frac{1}{r}\frac{\partial V_\phi}{\partial \phi} + \frac{\partial V_z}{\partial z} \tag{A1.23}$$

$$\boldsymbol{\nabla} \wedge \mathbf{V} = \left(\frac{1}{r}\frac{\partial V_z}{\partial \phi} - \frac{\partial V_\phi}{\partial z}, \frac{\partial V_r}{\partial z} - \frac{\partial V_z}{\partial r}, \frac{1}{r}\frac{\partial}{\partial r}(rV_\phi) - \frac{1}{r}\frac{\partial V_r}{\partial \phi}\right) \tag{A1.24}$$

$$\boldsymbol{\nabla}^2 T = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 T}{\partial \phi^2} + \frac{\partial^2 T}{\partial z^2} \tag{A1.25}$$

$$\boldsymbol{\nabla}^2 \mathbf{V} = \left(\nabla^2 V_r - \frac{V_r}{r^2} - \frac{2}{r^2}\frac{\partial V_\phi}{\partial \phi}, \nabla^2 V_\phi + \frac{2}{r^2}\frac{\partial V_r}{\partial \phi} - \frac{V_\phi}{r^2}, \nabla^2 V_z\right) \tag{A1.26}$$

Spherical polar coordinates (r, θ, ϕ)

In spherical polar coordinates (Fig. A1.4), a point P is located by specifying r, the radius of the sphere on which it lies, θ, the colatitude, and ϕ, the longitude or azimuth, where $r \geq 0$, $0 \leq \phi \leq 2\pi$, $0 \leq \theta \leq \pi$. From Fig. A1.4 it can be seen that

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned} \tag{A1.27}$$

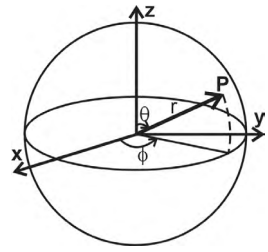


Figure A1.4.

In spherical polar coordinates (r, θ, ϕ) the gradient, divergence, curl and Laplacian operators are

$$\nabla T = \left(\frac{\partial T}{\partial r}, \frac{1}{r}\frac{\partial T}{\partial \theta}, \frac{1}{r\sin\theta}\frac{\partial T}{\partial \phi}\right) \tag{A1.28}$$

$$\boldsymbol{\nabla} \cdot \mathbf{V} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 V_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta V_\theta) + \frac{1}{r\sin\theta}\frac{\partial V_\phi}{\partial \phi} \tag{A1.29}$$

$$\begin{aligned} \boldsymbol{\nabla} \wedge \mathbf{V} = \Bigg(&\frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta V_\phi) - \frac{1}{r\sin\theta}\frac{\partial V_\theta}{\partial \phi}, \frac{1}{r\sin\theta}\frac{\partial V_r}{\partial \phi} \\ &-\frac{1}{r}\frac{\partial}{\partial r}(rV_\phi), \frac{1}{r}\frac{\partial}{\partial r}(rV_\theta) - \frac{1}{r}\frac{\partial V_r}{\partial \theta}\Bigg) \end{aligned} \tag{A1.30}$$

$$\boldsymbol{\nabla}^2 T = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial T}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 T}{\partial \phi^2} \tag{A1.31}$$

$$\begin{aligned} \boldsymbol{\nabla}^2 \mathbf{V} = \Bigg(&\nabla^2 V_r - \frac{2}{r^2}V_r - \frac{2}{r^2\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta V_\theta) - \frac{2}{r^2\sin\theta}\frac{\partial V_\phi}{\partial \phi}, \\ &\nabla^2 V_\theta + \frac{2}{r^2}\frac{\partial V_r}{\partial \theta} - \frac{V_\theta}{r^2\sin^2\theta} - \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial V_\phi}{\partial \phi}, \\ &\nabla^2 V_\theta + \frac{2}{r^2\sin\theta}\frac{\partial V_r}{\partial \phi} + \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial V_\theta}{\partial \phi} - \frac{V_\phi}{r^2\sin^2\theta}\Bigg) \end{aligned} \tag{A1.32}$$

Appendix 2
Theory of elasticity and elastic waves

When a fixed solid body is subjected to an external force, it changes in size and shape. An *elastic solid* is a solid that returns to its original size and shape after the external deforming force has been removed. For small deformations and on a short timescale (minutes not millions of years), rocks can be considered to be elastic.

Stress

Stress is defined as a force per unit area. When a deforming force is applied to a body, the stress is the ratio of the force to the area over which it is applied. If a force of one newton (1 N) is applied uniformly to an area of one square metre, the stress is 1 N $m^{-2} \equiv$ one pascal (1 Pa). If the force is normal (perpendicular) to the surface, then the stress is termed a *normal stress*; if it is tangential to the surface, the stress is termed a *shearing stress*. Usually, the force is neither entirely normal nor tangential but is at some arbitrary intermediate angle, in which case it can be resolved into components normal and tangential to the surface; so the stress is composed both of normal components and of shearing components. The sign convention is that tensional stresses are positive and compressional stresses negative.

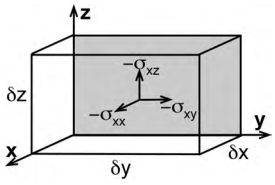


Figure A2.1.

Now consider a small parallelepiped with sides δx, δy and δz (Fig. A2.1) and imagine that it is being stressed by some external force. On each face, the stresses can be resolved into components in the x, y and z directions. The stresses acting on the shaded face are $-\sigma_{xx}$, $-\sigma_{xy}$ and $-\sigma_{xz}$. The notation is that σ_{xy} refers to the stress σ acting in the y direction on the face which is perpendicular to the x axis. The *normal stress* is thus $-\sigma_{xx}$ and the *shearing stresses* are $-\sigma_{xy}$ and $-\sigma_{xz}$.

If the parallelepiped is to be in *static equilibrium* (not moving), then the stresses on opposite faces must balance, and there must be no net couple that would rotate the parallelepiped. This requires that the stresses on opposite faces be equal in magnitude and opposite in direction. The shearing stresses on opposite faces of the parallelepiped (e.g., $-\sigma_{xy}$ and σ_{xy} on the back and front faces as shown in Fig. A2.1) provide a couple that will rotate the parallelepiped. Since the parallelepiped must not rotate, this couple must be balanced by the couple provided by the shearing stresses $-\sigma_{yx}$ and σ_{yx} acting on the two side faces. This means that σ_{xy} must equal σ_{yx}. The same conditions apply to the other shearing stresses: $\sigma_{xy} = \sigma_{yx}$, $\sigma_{xz} = \sigma_{zx}$ and $\sigma_{yz} = \sigma_{zy}$ (i.e., the stress tensor must be symmetrical).

Strain

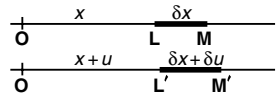


Figure A2.2.

When a body is subjected to stresses, the resulting deformations are called *strains*. Strain is defined as the relative change (i.e., the fractional change) in the shape of the body. First, consider a stress that acts in the x direction only on an elastic string (Fig. A2.2). The point L on the string moves a distance u to point L′ after stretching, and point M moves a distance $u + \delta u$ to point M′. The strain in the x direction, termed e_{xx}, is then given by

$$\begin{aligned} e_{xx} &= \frac{\text{change in length of LM}}{\text{original length of LM}} \\ &= \frac{\text{L}'\text{M}' - \text{LM}}{\text{LM}} \\ &= \frac{\delta x + \delta u - \delta x}{\delta x} \\ &= \frac{\delta u}{\delta x} \end{aligned} \tag{A2.1}$$

In the limit when $\delta x \to 0$, the strain at L is

$$e_{xx} = \frac{\partial u}{\partial x} \tag{A2.2}$$

To extend the analysis to two dimensions x and y, we must consider the deformation undergone by a rectangle in the x–y plane (Fig. A2.3).

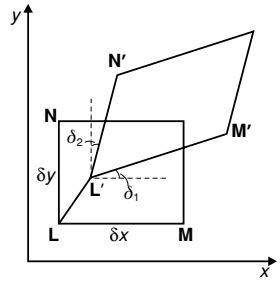


Figure A2.3.

Points L, M and N move to L′, M′ and N′ with coordinates

$$\begin{aligned} &\text{L} = (x, y), && \text{L}' = (x + u, y + \text{v}) \\ &\text{M} = (x + \text{d}x, y), && \text{M}' = \left(x + \delta x + u + \frac{\partial u}{\partial x}\delta x, y + v + \frac{\partial v}{\partial x}\delta x\right) \\ &\text{N} = (x, y + \text{d}y) && \text{N}' = \left(x + u + \frac{\partial u}{\partial y}\delta y, y + \delta y + v + \frac{\partial v}{\partial y}\delta y\right) \end{aligned}$$

The strain in the x direction, e_{xx}, is given by

$$\begin{aligned} e_{xx} &= \frac{\text{change in length of LM}}{\text{original length of LM}} \\ &= \frac{\delta x + \dfrac{\partial u}{\partial x}\delta x - \delta x}{\delta x} \\ &= \frac{\partial u}{\partial x} \end{aligned} \tag{A2.3}$$

Likewise, the strain in the y direction is

$$\begin{aligned} e_{yy} &= \frac{\text{change in length of LN}}{\text{original length of LN}} \\ &= \frac{\partial v}{\partial y} \end{aligned} \tag{A2.4}$$

These are called the *normal strains*, the fractional changes in length along the x and y axes. For three dimensions, $e_{zz} = \partial w/\partial z$ is the third normal strain.

As well as changing size, the rectangle undergoes a change in shape (Fig. A2.3). The right angle $\widehat{\text{NLM}}$ is reduced by an amount $\delta_1 + \delta_2$ called the *angle of shear*, where

$$\delta_1 + \delta_2 = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \tag{A2.5}$$

(We assume that products of $\partial u/\partial x$, $\partial v/\partial x$ and so on are small enough to be ignored, which is the basis of the *theory of infinitesimal strain*.) The quantity which measures the change in

shape undergone by the rectangle is called the *shear component of strain* and is written e_{xy}. In three dimensions, there are six shear components of strain:

$$e_{xy} = e_{yx} = \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \tag{A2.6}$$

$$e_{xz} = e_{zx} = \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) \tag{A2.7}$$

$$e_{zy} = e_{yz} = \frac{1}{2}\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right) \tag{A2.8}$$

Note that the angle of shear is equal to twice the shear component of strain. As well as undergoing a change in shape, the whole rectangle is also rotated anticlockwise by an angle $\frac{1}{2}(\delta_1 - \delta_2)$, termed θ_z, where

$$\begin{aligned}\theta_z &= \frac{1}{2}(\delta_1 - \delta_2)\\ &= \frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\end{aligned} \tag{A2.9}$$

θ_z is an anticlockwise rotation about the z axis.

Extending the theory to three dimensions, the deformation $(\delta u, \delta v, \delta w)$ of any point $(\delta x, \delta y, \delta z)$ can be expressed as a power series, where, to first order,

$$\begin{aligned}\delta u &= \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y + \frac{\partial u}{\partial z}\delta z\\ \delta v &= \frac{\partial v}{\partial x}\delta x + \frac{\partial v}{\partial y}\delta y + \frac{\partial v}{\partial z}\delta z\\ \delta w &= \frac{\partial w}{\partial x}\delta x + \frac{\partial w}{\partial y}\delta y + \frac{\partial w}{\partial z}\delta z\end{aligned} \tag{A2.10}$$

Alternatively, Eqs. (A2.10) can be split into symmetrical and antisymmetrical parts:

$$\begin{aligned}\delta u &= e_{xx}\delta x + e_{xy}\delta y + e_{xz}\delta z - \theta_z\delta y + \theta_y\delta z\\ \delta v &= e_{xy}\delta x + e_{yy}\delta y + e_{yz}\delta z + \theta_z\delta x - \theta_x\delta z\\ \delta w &= e_{xz}\delta x + e_{yz}\delta y + e_{zz}\delta z - \theta_y\delta x + \theta_x\delta y\end{aligned} \tag{A2.11}$$

where[1]

$$\begin{aligned}\theta_x &= \frac{1}{2}\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right)\\ \theta_y &= \frac{1}{2}\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right)\\ \theta_z &= \frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\end{aligned} \tag{A2.12}$$

In more compact matrix form, Eq. (A2.11) is

$$(\delta u, \delta v, \delta w) = \begin{pmatrix} e_{xx} & e_{xy} & e_{xz}\\ e_{xy} & e_{yy} & e_{yz}\\ e_{xz} & e_{yz} & e_{zz}\end{pmatrix}\begin{pmatrix}\delta x\\ \delta y\\ \delta z\end{pmatrix} + \begin{pmatrix} 0 & -\theta_z & \theta_y\\ \theta_z & 0 & -\theta_x\\ -\theta_y & \theta_x & 0\end{pmatrix}\begin{pmatrix}\delta x\\ \delta y\\ \delta z\end{pmatrix} \tag{A2.13}$$

Strain is a dimensionless quantity. Generally, in seismology, the strain caused by the passage of a seismic wave is about 10^{-6} in magnitude.

[1] The curl of the vector (u, v, w), $\boldsymbol{\nabla} \wedge \mathbf{u}$, is equal to twice the rotation $(\theta_x, \theta_y, \theta_z)$ as discussed in Appendix 1.

The fractional increase in volume caused by a deformation is called *cubical dilatation* and is written Δ. The volume of the original rectangular parallelepiped is V, where

$$V = \delta x\ \delta y\ \delta z \tag{A2.14}$$

The volume of the deformed parallelepiped, $V + \delta V$, is approximately

$$V + \delta V = (1 + e_{xx})\ \delta x\ (1 + e_{yy})\ \delta y\ (1 + e_{zz})\ \delta z \tag{A2.15}$$

The cubical dilatation Δ is then given by

$$\begin{aligned}\Delta &= \frac{\text{change in volume}}{\text{originalvolume}} \\ &= \frac{V + \delta V - V}{V} \\ &= \frac{(1 + e_{xx})(1 + e_{yy})(1 + e_{zz})\ \delta x\ \delta y\ \delta z - \delta x\ \delta y\ \delta z}{\delta x\ \delta y\ \delta z}\end{aligned} \tag{A2.16}$$

Therefore, to first order (recall that assumption of infinitesimal strain means that products of strains can be neglected), the cubical dilatation is given by

$$\Delta = e_{xx} + e_{yy} + e_{zz}$$

or

$$\Delta = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$$

or

$$\Delta = \boldsymbol{\nabla} \cdot \mathbf{u} \tag{A2.17}$$

The relationship between stress and strain

In practice, in a given situation, we want to calculate the strains when the stress is known. In 1676, the English physicist Robert Hooke proposed that, for small strains, any strain is proportional to the stress that produces it. This is known as *Hooke's law* and forms the basis of the theory of perfect elasticity. In one dimension x, Hooke's law means that

$$\sigma_{xx} = c e_{xx}$$

where c is a constant. Extending the theory to three dimensions gives thirty-six different constants:

$$\begin{aligned}\sigma_{xx} &= c_1 e_{xx} + c_2 e_{xy} + c_3 e_{xz} + c_4 e_{yy} + c_5 e_{yz} + c_6 e_{zz} \\ \ldots &\qquad\qquad \ldots \\ \sigma_{zz} &= c_{31} e_{xx} + c_{32} e_{xy} + c_{33} e_{xz} + c_{34} e_{yy} + c_{35} e_{yz} + c_{36} e_{zz}\end{aligned} \tag{A2.18}$$

If we assume that we are considering only isotropic materials (materials with no directional variation), the number of constants is reduced from thirty-six to two:

$$\begin{aligned}\sigma_{xx} &= (\lambda + 2\mu) e_{xx} + \lambda e_{yy} + \lambda e_{zz} \\ &= \lambda\Delta + 2\mu e_{xx} \\ \sigma_{yy} &= \lambda\Delta + 2\mu e_{yy} \\ \sigma_{zz} &= \lambda\Delta + 2\mu e_{zz} \\ \sigma_{xy} &= \sigma_{yx} = 2\mu e_{xy} \\ \sigma_{xz} &= \sigma_{zx} = 2\mu e_{xz} \\ \sigma_{yz} &= \sigma_{zy} = 2\mu e_{yz}\end{aligned} \tag{A2.19}$$

The constants λ and μ are known as the Lamé elastic constants (named after the nineteenth-century French mathematician G. Lamé). In suffix notation, Eqs. (A2.19) are written as

$$\sigma_{ij} = \lambda\Delta\delta_{ij} + 2\mu e_{ij} \quad \text{for} \quad i, j = x, y, z \tag{A2.20}$$

where the Kronecker delta

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

The Lamé elastic constant μ (where $\mu = \sigma_{xy}/(2e_{xy})$ from Eq. (A2.19)) is a measure of the resistance of a body to shearing strain and is often termed the *shear modulus* or the *rigidity modulus*. The shear modulus of a liquid or gas is zero.

Besides the Lamé elastic constants, other elastic constants are also used: *Young's modulus E, Poisson's ratio* σ (no subscripts) and the bulk modulus *K*.

Young's modulus

E is the ratio of tensional stress to the resultant longitudinal strain for a small cylinder under tension at both ends. Let the tensional stress act in the *x* direction on the end face of the small cylinder, and let all the other stresses be zero. Equations (A2.19) then give

$$\begin{aligned} \sigma_{xx} &= \lambda\Delta + 2\mu e_{xx} \\ 0 &= \lambda\Delta + 2\mu e_{yy} \\ 0 &= \lambda\Delta + 2\mu e_{zz} \end{aligned} \tag{A2.21}$$

and

$$0 = e_{xy} = e_{xz} = e_{yz} \tag{A2.22}$$

Adding Eqs. (A2.21) gives

$$\sigma_{xx} = 3\lambda\Delta + 2\mu\Delta \tag{A2.23}$$

Substituting Eq. (A2.23) into Eq. (A2.21) gives

$$e_{xx} = (\lambda + \mu)\frac{\Delta}{\mu} \tag{A2.24}$$

Hence, Young's modulus is

$$E = \frac{\sigma_{xx}}{e_{xx}} = \frac{(3\lambda + 2\mu)\Delta\mu}{(\lambda + \mu)\Delta} = \frac{(3\lambda + 2\mu)\mu}{(\lambda + \mu)} \tag{A2.25}$$

Poisson's ratio

σ (named after the nineteenth-century French mathematician Siméon Denis Poisson) is defined as the negative of the ratio of the fractional lateral contraction to the fractional longitudinal extension for the same small cylinder under tension at both ends. Using Eqs. (A2.23) and (A2.21), Poisson's ratio is given by

$$\sigma = -\frac{e_{zz}}{e_{xx}} = \frac{\lambda\Delta}{2\mu}\frac{\mu}{(\lambda + \mu)\Delta} = \frac{\lambda}{2(\lambda + \mu)} \tag{A2.26}$$

Consider a small body subjected to a hydrostatic pressure (i.e., the body is immersed in a liquid). This pressure causes compression of the body. The ratio of the pressure to the

resulting compression is called the *bulk modulus* or *incompressibility K* of the body. For hydrostatic pressure p, the stresses are

$$\begin{aligned} \sigma_{xx} &= \sigma_{yy} = \sigma_{zz} = -p \\ \sigma_{xy} &= \sigma_{xz} = \sigma_{yz} = 0 \end{aligned} \tag{A2.27}$$

Equations (A2.19) then give

$$\begin{aligned} -p &= \lambda\Delta + 2\mu e_{xx} \\ -p &= \lambda\Delta + 2\mu e_{yy} \\ -p &= \lambda\Delta + 2\mu e_{zz} \end{aligned} \tag{A2.28}$$

and

$$0 = e_{xy} = e_{xz} = e_{yz} \tag{A2.29}$$

Adding Eqs. (A2.28) gives

$$-3p = 3\lambda\Delta + 2\mu\Delta \tag{A2.30}$$

Finally, the bulk modulus is given by

$$\begin{aligned} K &= \frac{\text{pressure}}{\text{compression}} = \frac{\text{pressure}}{-\text{dilatation}} \\ &= \frac{p}{-\Delta} \\ &= \lambda + \tfrac{2}{3}\mu \end{aligned} \tag{A2.31}$$

Using these relations (Eqs. (A2.25), (A2.26) and (A2.31)) amongst the five elastic constants, we can write Eq. (A2.19) or Eq. (A2.20) in terms of any pair of the constants.

Poisson's ratio is dimensionless, positive and less than 0.5 (it is exactly 0.5 for a liquid since then $\mu = 0$). Young's modulus, the Lamé constants and the bulk modulus are all positive and (together with stress and pressure) are all quoted in units of N m^{-2} (1 Pa $\equiv$ 1 N m^{-2}). (For rocks, E, K, λ and μ are generally 20–120 GPa). The two Lamé constants have almost the same value for rocks, so the approximation $\lambda = \mu$ is sometimes made. This approximation is called *Poisson's relation*.

Equations of motion

Let us assume that the stresses on opposite faces of the small parallelepiped illustrated in Fig. A2.1 do not exactly balance, so the parallelepiped is not in equilibrium: motion is possible. In this case (Fig. A2.4), although the stresses on the rear face are $(-\sigma_{xx}, -\sigma_{xy}, -\sigma_{xz})$, the stresses on the front shaded face can be written as $(\sigma_{xx} + \delta\sigma_{xx}, \sigma_{xy} + \delta\sigma_{xy}, \sigma_{xz} + \delta\sigma_{xz})$. The additional stress $(\delta\sigma_{xx}, \delta\sigma_{xy}, \delta\sigma_{xz})$ can be written as $[(\partial\sigma_{xx}/\partial x)\,\delta x, (\partial\sigma_{xy}/\partial x)\,\delta x, (\partial\sigma_{xz}/\partial x)\,\delta x]$.

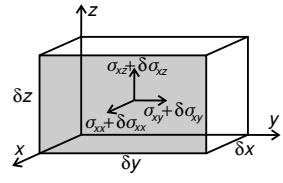


Figure A2.4.

Thus, the net force (stress multiplied by area) acting on the two faces perpendicular to the x axis is

$$\begin{aligned} &\left(-\sigma_{xx} + \sigma_{xx} + \frac{\partial\sigma_{xx}}{\partial x}\,\delta x, -\sigma_{xy} + \sigma_{xy} + \frac{\partial\sigma_{xy}}{\partial x}\,\delta x, -\sigma_{xz} + \sigma_{xz} + \frac{\partial\sigma_{xz}}{\partial x}\,\delta x\right)\delta y\,\delta z \\ &\quad = \left(\frac{\partial\sigma_{xx}}{\partial x}, \frac{\partial\sigma_{xy}}{\partial x}, \frac{\partial\sigma_{xz}}{\partial x}\right)\delta x\,\delta y\,\delta z \end{aligned} \tag{A2.32}$$

and similarly for the other two pairs of faces. The total force acting on the parallelepiped in the x direction thus is

$$\left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z}\right) \delta x\ \delta y\ \delta z$$

and similarly for the force in the y and z directions. Using Newton's second law of motion (force = mass × acceleration), we can write

$$\left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z}\right) \delta x\ \delta y\ \delta z = \rho\ \delta x\ \delta y\ \delta z\ \frac{\partial^2 u}{\partial t^2} \tag{A2.33}$$

where ρ is the density of the parallelepiped and u is the x component of the displacement. (We assume that all other body forces are zero; that is, gravity does not vary significantly across the parallelepiped.) This equation of motion, which relates the second differential of the displacement to the stress, can be simplified by expressing stress in terms of strain from Eqs. (A2.19) and strain in terms of displacement from Eqs. (A2.3), (A2.6) and (A2.7). Substituting for the stress from Eqs. (A2.19) into Eqs. (A2.33) gives

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}(\lambda \Delta + 2\mu e_{xx}) + \frac{\partial}{\partial y}(2\mu e_{xy}) + \frac{\partial}{\partial z}(2\mu e_{xz}) \tag{A2.34}$$

Substituting for the strains from Eqs. (A2.3), (A2.6) and (A2.7) into Eq. (A2.34) gives

$$\begin{aligned} \rho \frac{\partial^2 u}{\partial t^2} = {} & \frac{\partial}{\partial x}\left(\lambda \Delta + 2\mu \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left[\mu\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right] \\ & + \frac{\partial}{\partial z}\left[\mu\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right)\right] \end{aligned} \tag{A2.35}$$

Assuming λ and μ to be constants, we can write

$$\begin{aligned} \rho \frac{\partial^2 u}{\partial t^2} &= \lambda \frac{\partial \Delta}{\partial x} + 2\mu \frac{\partial^2 u}{\partial x^2} + \mu \frac{\partial^2 u}{\partial y^2} + \mu \frac{\partial^2 v}{\partial x \partial y} + \mu \frac{\partial^2 w}{\partial x \partial z} + \mu \frac{\partial^2 u}{\partial z^2} \\ &= \lambda \frac{\partial \Delta}{\partial x} + \mu \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) \\ &= \lambda \frac{\partial \Delta}{\partial x} + \mu \frac{\partial \Delta}{\partial x} + \mu\, \boldsymbol{\nabla}^2 u \end{aligned} \tag{A2.36a}$$

where $\boldsymbol{\nabla}^2$ is the Laplacian operator $\equiv \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ (see Appendix 1). Likewise, the y and z components of the forces are used to yield equations for v and w:

$$\rho \frac{\partial^2 v}{\partial t^2} = (\lambda + \mu)\frac{\partial \Delta}{\partial y} + \mu\, \boldsymbol{\nabla}^2 v \tag{A2.36b}$$

$$\rho \frac{\partial^2 w}{\partial t^2} = (\lambda + \mu)\frac{\partial \Delta}{\partial z} + \mu\, \boldsymbol{\nabla}^2 w \tag{A2.36c}$$

These three equations are the equations of motion for a general disturbance transmitted through a homogeneous, isotropic, perfectly elastic medium, assuming that we have infinitesimal strain and no body forces. We can now manipulate these equations to put them into a more useful form.

First, if we differentiate the u, v and w equations with respect to x, y and z, respectively, and add the results, we obtain

$$\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)=(\lambda+\mu)\frac{\partial^2\Delta}{\partial x^2}+\mu\,\nabla^2\left(\frac{\partial u}{\partial x}\right)$$
$$+(\lambda+\mu)\frac{\partial^2\Delta}{\partial y^2}+\mu\,\nabla^2\left(\frac{\partial v}{\partial y}\right)$$
$$+(\lambda+\mu)\frac{\partial^2\delta}{\partial z^2}+\mu\,\nabla^2\left(\frac{\partial w}{\partial z}\right) \qquad \text{(A2.37)}$$

or

$$\rho\frac{\partial^2\Delta}{\partial t^2}=(\lambda+\mu)\,\boldsymbol{\nabla}^2\Delta+\mu\,\boldsymbol{\nabla}^2\Delta$$
$$=(\lambda+2\mu)\,\boldsymbol{\nabla}^2\Delta \qquad \text{(A2.38)}$$

This is a *wave equation* for a dilatational disturbance transmitted through the material with a speed

$$\alpha=\sqrt{\frac{\lambda+2\mu}{\rho}} \qquad \text{(A2.39)}$$

In seismology, as discussed in Chapter 4, this type of wave involves only dilatation and no rotation and is termed the *primary wave* or *P-wave*.

Second, we can differentiate Eq. (A2.36a) with respect to y and Eq. (A2.36b) with respect to x:

$$\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial u}{\partial y}\right)=(\lambda+\mu)\frac{\partial^2\Delta}{\partial x\,\partial y}+\mu\,\nabla^2\left(\frac{\partial u}{\partial y}\right) \qquad \text{(A2.40)}$$

and

$$\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial v}{\partial x}\right)=(\lambda+\mu)\frac{\partial^2\Delta}{\partial x\,\partial y}+\mu\,\nabla^2\left(\frac{\partial v}{\partial x}\right) \qquad \text{(A2.41)}$$

Subtracting Eq. (A2.41) from (A2.40) gives

$$\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial u}{\partial y}-\frac{\partial v}{\partial x}\right)=\mu\,\nabla^2\left(\frac{\partial u}{\partial y}-\frac{\partial v}{\partial x}\right) \qquad \text{(A2.42a)}$$

By differentiating and subtracting derivatives, we obtain the other two equations:

$$\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial u}{\partial z}-\frac{\partial w}{\partial x}\right)=\mu\,\nabla^2\left(\frac{\partial u}{\partial z}-\frac{\partial w}{\partial x}\right) \qquad \text{(A2.42b)}$$

$$\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial v}{\partial z}-\frac{\partial w}{\partial y}\right)=\mu\,\nabla^2\left(\frac{\partial v}{\partial z}-\frac{\partial w}{\partial y}\right) \qquad \text{(A2.42c)}$$

However, since $\partial u/\partial y-\partial v/\partial x$ and so on are the components of curl $\mathbf{u}$ (or $\boldsymbol{\nabla}\wedge\mathbf{u}$; see Appendix 1), these three equations can be written

$$\rho\frac{\partial^2}{\partial t^2}(\text{curl }\mathbf{u})=\mu\,\nabla^2(\text{curl }\mathbf{u}) \qquad \text{(A2.43)}$$

This is a vector wave equation for a rotational disturbance transmitted through the material with a speed

$$\beta=\sqrt{\frac{\mu}{\rho}} \qquad \text{(A2.44)}$$

In seismology, as discussed in Chapter 4, this type of wave involves only rotation and no change in volume and is called the *secondary wave* or *S-wave*.

Displacement potentials

We can use the method of Helmholtz to express the displacement $\mathbf{u}$ as the sum of the gradient of a scalar potential ϕ and the curl of a vector potential $\boldsymbol{\psi}$. The divergence of the vector potential must be zero: $\nabla \cdot \boldsymbol{\psi} = 0$. The displacement is then expressed as

$$\mathbf{u} = \nabla\phi + \nabla \wedge \boldsymbol{\psi} \tag{A2.45}$$

The two potentials ϕ and $\boldsymbol{\psi}$ are called the *displacement potentials*. Substituting Eq. (A2.45) into Eqs. (A2.38) and (A2.43) and using the vector identities $\nabla \cdot (\nabla \wedge \mathbf{V}) = 0$, $\nabla \wedge (\nabla S) = 0$ and $\nabla \wedge (\nabla \wedge \mathbf{V}) = \nabla(\nabla \cdot \mathbf{V}) - \nabla^2\mathbf{V}$, where S is a scalar and $\mathbf{V}$ a vector, gives

$$\frac{\partial^2}{\partial t^2}(\nabla^2\phi) = \left(\frac{\lambda + 2\mu}{\rho}\right)\nabla^2(\nabla^2\phi) \tag{A2.46}$$

and

$$\frac{\partial^2}{\partial t^2}(\nabla^2\boldsymbol{\psi}) = \frac{\mu}{\rho}\nabla^4\boldsymbol{\psi} \tag{A2.47}$$

The potentials therefore satisfy the wave equations

$$\frac{\partial^2\phi}{\partial t^2} = \left(\frac{\lambda + 2\mu}{\rho}\right)\nabla^2\phi \tag{A2.48}$$

and

$$\frac{\partial^2\boldsymbol{\psi}}{\partial t^2} = \frac{\mu}{\rho}\nabla^2\boldsymbol{\psi} \tag{A2.49}$$

Equation (A2.48) is thus an alternative expression of Eq. (A2.38), the wave equation for P-waves, and Eq. (A2.49) is an alternative expression of Eq. (A2.43), the wave equation for S-waves.

Plane waves

Consider the case in which ϕ is a function of x and t only. Then Eq. (A2.48) simplifies to

$$\begin{aligned}\frac{\partial^2\phi}{\partial t^2} &= \frac{\lambda + 2\mu}{\rho}\frac{\partial^2\phi}{\lambda x^2} \\ &= \alpha^2\frac{\partial^2\phi}{\partial x^2}\end{aligned} \tag{A2.50}$$

Any function of $x \pm at$, $\phi = \phi(x \pm at)$ is a solution to Eq. (A2.50), provided that $\partial\phi/\partial x$, $\partial^2\phi/\partial x^2$, $\partial\phi/\partial t$ and $\partial^2\phi/\partial t^2$ are continuous. The simplest harmonic solution to Eq. (A2.50) is

$$\phi = \cos[\kappa(x - \alpha t)] \tag{A2.51}$$

where κ is a constant termed the *wave number*. Equation (A2.51) describes a plane wave travelling in the x direction with velocity α. The displacement of the medium due to the passage of this wave is given by Eq. (A2.45):

$$\begin{aligned}u &= \nabla\phi \\ &= \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right) \\ &= (-\kappa \sin[\kappa(x - at)], 0, 0)\end{aligned} \tag{A2.52}$$

The velocity at any point $\partial \mathbf{u}/\partial t$ is then given by

$$\frac{\partial \mathbf{u}}{\partial t} = (\alpha\kappa^2 \cos[\kappa(x - \alpha t)], 0, 0) \tag{A2.53}$$

The wavelength λ, angular frequency ω, frequency f and period T of this wave are given by

$$\begin{aligned} \lambda &= \frac{2\pi}{\kappa} \\ \omega &= \kappa\alpha \\ f &= \frac{\omega}{2\pi} \\ T &= \frac{\lambda}{\alpha} = \frac{2\pi}{\omega} = \frac{1}{f} \end{aligned} \tag{A2.54}$$

Appendix 3

Geometry of ray paths and inversion of earthquake body-wave time–distance curves

To be able to use the travel-time–distance curves for teleseismic earthquakes (Fig. 4.16) to determine the internal structure of the Earth, it is necessary to devise equations relating seismic velocity and depth to travel time and distance.

Initially, consider an Earth assumed to consist of spherically symmetrical shells, each shell having constant seismic velocity. Consider part of the particular seismic ray (Fig. A3.1) which traverses three of these layers. Applying Snell's law (Section 4.3.2) to interface 1 gives

$$\frac{\sin i_1}{v_1} = \frac{\sin j_1}{v_2} \tag{A3.1}$$

and applying it to interface 2 gives

$$\frac{\sin i_2}{v_2} = \frac{\sin j_2}{v_3} \tag{A3.2}$$

However, from the right-angled triangles OP_1Q and OP_2Q, we can write

$$OQ = OP_1 \sin j_1 = r_1 \sin j_1 \tag{A3.3}$$

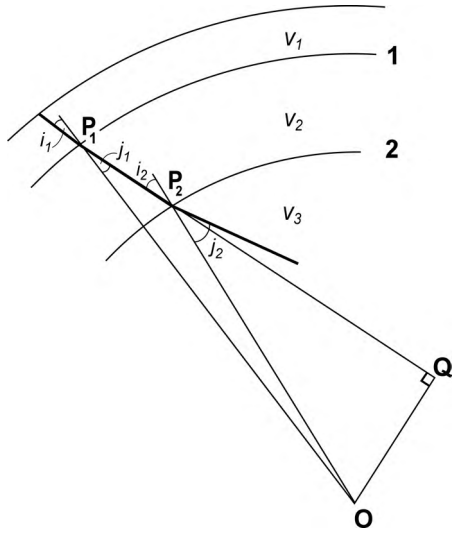


Figure A3.1.

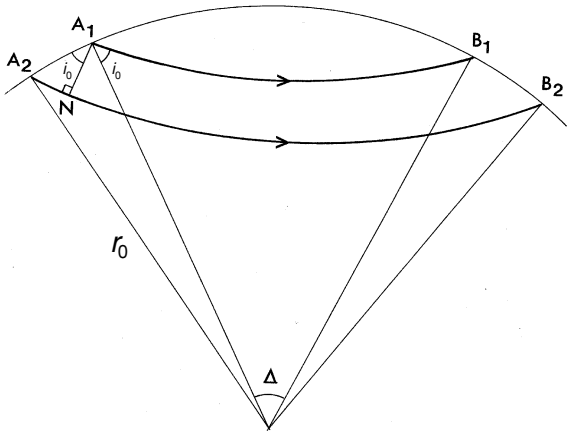


Figure A3.2.

and

$$\mathrm{OQ} = \mathrm{OP}_2 \sin i_2 = r_2 \sin i_2 \tag{A3.4}$$

where $\mathrm{OP}_1 = r_1$ and $\mathrm{OP}_2 = r_2$. Thus, on combining Eqs. (A3.3) and (A3.4), we have

$$r_1 \sin j_1 = r_2 \sin i_2 \tag{A3.5}$$

Multiplying Eq. (A3.1) by r_1 and Eq. (A3.2) by r_2 and using Eq. (A3.5) means that

$$\frac{r_1 \sin i_1}{v_1} = \frac{r_1 \sin j_1}{v_2} = \frac{r_2 \sin i_2}{v_2} = \frac{r_2 \sin j_2}{v_3} \tag{A3.6}$$

At this point we define a parameter p as the *ray parameter:*

$$p = \frac{r \sin i}{v} \tag{A3.7}$$

where r is the distance from the centre of the Earth O to any point P, v is the seismic velocity at P and i is the angle of incidence at P. Equation (A3.6) shows that p is a constant along the ray. At the deepest point to which the ray penetrates (the *turning point*), i is $\pi/2$, so Eq. (A3.7) becomes

$$p = \frac{r_{\min}}{v} \tag{A3.8}$$

where $r_{\min}$ is the radius of the turning point and v the velocity at the point. The value of the ray parameter p is different for each ray.

Now consider two adjacent rays (Fig. A3.2). The shorter ray $\mathrm{A_1B_1}$ subtends an angle Δ at the centre of the Earth, and the longer ray $\mathrm{A_2B_2}$ subtends $\Delta + \delta\Delta$. The travel time for ray $\mathrm{A_1B_1}$ is t, and the travel time for ray $\mathrm{A_2B_2}$ is $t + \delta t$. In the infinitesimal right triangle $\mathrm{A_1NA_2}$, the angle $\widehat{\mathrm{A_2A_1N}}$ is i_0 and

$$\sin i_0 = \frac{\mathrm{A_2N}}{\mathrm{A_2A_1}} \tag{A3.9}$$

Assuming that the surface seismic velocity is v_0,

$$\mathrm{A_2N} = \tfrac{1}{2} v_0 \, \delta t \tag{A3.10}$$

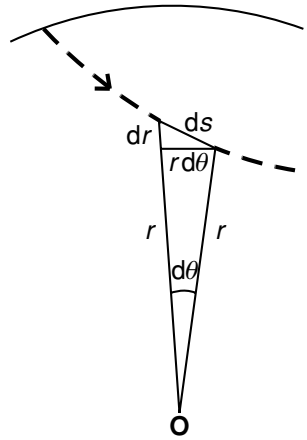


Figure A3.3.

and

$$A_2A_1 = \tfrac{1}{2} r_0 \, \delta\Delta \tag{A3.11}$$

Substituting Eqs. (A3.10) and (A3.11) into Eq. (A3.9) gives

$$\sin i_0 = \frac{v_0 \, \delta t}{r_0 \, \delta\Delta} \tag{A3.12}$$

Comparison with Eq. (A3.7) means that, in the limit when δt, $\delta\Delta \to 0$,

$$p = \frac{\mathrm{d}t}{\mathrm{d}\Delta} \tag{A3.13}$$

The ray parameter p is therefore the slope of the curve of travel time versus epicentral angle (Fig. 4.16) and so, for any particular phase, is an observed function of the epicentral angle.

Let ds be the length of a short segment of a ray, as shown in Fig. A3.3. Then, using Pythagoras' theorem on the infinitesimal triangle, we obtain

$$(\mathrm{d}s)^2 = (\mathrm{d}r)^2 + (r \, \mathrm{d}\theta)^2 \tag{A3.14}$$

However, from Eq. (A3.7) we have

$$p = \frac{r \sin i}{v} = \frac{r}{v} r \, \frac{\mathrm{d}\theta}{\mathrm{d}s} \tag{A3.15}$$

Eliminating ds from Eqs. (A3.14) and (A3.15) gives an expression for dθ,

$$\frac{r^4 (\mathrm{d}\theta)^2}{p^2 v^2} = (\mathrm{d}r)^2 + r^2 (\mathrm{d}\theta)^2 \tag{A3.16}$$

which, upon rearranging, becomes

$$\mathrm{d}\theta = \frac{p \, \mathrm{d}r}{r \, (r^2/v^2 - p^2)^{1/2}} \tag{A3.17}$$

Integrating this equation between the surface ($r = r_0$) and the deepest point ($r = r_{\min}$) gives an expression for Δ:

$$\Delta = 2p \int_{r=r_{\min}}^{r_0} \frac{\mathrm{d}r}{r \, (r^2/v^2 - p^2)^{1/2}} \tag{A3.18}$$

Eliminating $\mathrm{d}\theta$ from Eqs. (A3.14) and (A3.15) yields an expression for $\mathrm{d}s$,

$$\mathrm{d}s = \frac{r\,\mathrm{d}r}{v(r^2/v^2 - p^2)^{1/2}} \tag{A3.19}$$

The travel time $\mathrm{d}t$ along this short ray segment $\mathrm{d}s$ is $\mathrm{d}s/v$. Integrating this along the ray between the surface ($r = r_0$) and the deepest point ($r = r_{\min}$) gives an expression for t, the total travel time for the ray path:

$$t = 2\int_{r=r_{\min}}^{r_0} \frac{\mathrm{d}s}{v} = 2\int_{r=r_{\min}}^{r_0} \frac{r\,\mathrm{d}r}{v^2(r^2/v^2 - p^2)^{1/2}} \tag{A3.20}$$

Sometimes for convenience another variable η, defined as

$$\eta = \frac{r}{v} \tag{A3.21}$$

is introduced. When this substitution is made, Eqs. (A3.18) and (A3.20) are written as

$$\Delta = 2p\int_{r=r_{\min}}^{r_0} \frac{\mathrm{d}r}{r(\eta^2 - p^2)^{1/2}} \tag{A3.22}$$

$$t = 2\int_{r=r_{\min}}^{r_0} \frac{\eta^2\,\mathrm{d}r}{r(\eta^2 - p^2)^{1/2}} \tag{A3.23}$$

These two integrals can always be calculated: the travel times and epicentral distances can be calculated even for complex velocity–depth structures involving low-velocity or hidden layers.

In order to use the t–Δ curves to determine seismic velocities, it is necessary to change the variable in Eq. (A3.22) from r to η, which is possible only when η decreases monotonically with decreasing r:

$$\Delta = 2p\int_{\eta=\eta_{\min}}^{\eta_0} \frac{1}{r(\eta^2 - p^2)^{1/2}} \frac{\mathrm{d}r}{\mathrm{d}\eta}\mathrm{d}\eta \tag{A3.24}$$

The limits of integration are $\eta_0 = r_0/v_0$ and $\eta_{\min} = r_{\min}/v(r = r_{\min})$. However, since by Eq. (A3.8), $p = r_{\min}/v(r = r_{\min})$, the lower limit of integration $\eta_{\min}$ is in fact equal to the ray parameter p for the ray emerging at epicentral angle Δ.

Now, at $r = r_1$, where r_1 is any radius for which $r_0 \geq r_1 > r = r_{\min}$, let η and v have values η_1 and v_1, respectively. Assume that there is a series of turning rays sampling only the region between r_0 and r_1 with values of p between η_0 ($\eta_0 = r_0/v_0$), which is the ray travelling at a tangent to the Earth's surface and hence having $\Delta = 0$, and η_1 ($\eta_1 = r_1/v_1$), which is the ray whose turning point is r_1. Multiplying both sides of Eq. (A3.24) by $1/(p^2 - \eta_1^2)^{1/2}$ gives

$$\frac{\Delta}{(p^2 - \eta_1^2)^{1/2}} = \frac{2p}{(p^2 - \eta_1^2)^{1/2}}\int_{\eta=p}^{\eta_0} \frac{1}{r(\eta^2 - p^2)^{1/2}} \frac{\mathrm{d}r}{\mathrm{d}\eta}\mathrm{d}\eta \tag{A3.25}$$

Now integrate Eq. (A3.25) with respect to p between the limits η_1 and η_0:

$$\int_{p=\eta_1}^{\eta_0} \frac{\Delta}{(p^2 - \eta_1^2)^{1/2}}\mathrm{d}p = \int_{p=\eta_1}^{\eta_0} \frac{2p}{(p^2 - \eta_1^2)^{1/2}} \left(\int_{\eta=p}^{\eta_0} \frac{1}{r(\eta^2 - p^2)^{1/2}} \frac{\mathrm{d}r}{\mathrm{d}\eta}\mathrm{d}\eta\right)\mathrm{d}p \tag{A3.26}$$

It is mathematically permissible to change the order of integration on the right-hand side of Eq. (A3.26) from η first and p second to p first and η second:

$$\int_{p=\eta_1}^{\eta_0} \frac{\Delta}{(p^2 - \eta_1^2)^{1/2}}\mathrm{d}p = \int_{\eta=\eta_1}^{\eta_0} \left(\int_{p=\eta_1}^{\eta} \frac{2p}{r(p^2 - \eta_1^2)^{1/2}(\eta^2 - p^2)^{1/2}} \frac{\mathrm{d}r}{\mathrm{d}\eta}\mathrm{d}p\right)\mathrm{d}\eta \tag{A3.27}$$

Integrating the left-hand side of Eq. (A3.27) by parts gives

$$\int_{p=\eta_1}^{\eta_0} \frac{\Delta\,\mathrm{d}p}{(p^2 - \eta_1^2)^{1/2}} = \left[\Delta\cosh^{-1}\left(\frac{p}{\eta_1}\right)\right]_{p=\eta_1}^{\eta_0} - \int_{p=\eta_1}^{\eta_0} \frac{\mathrm{d}\Delta}{\mathrm{d}p}\cosh^{-1}\left(\frac{p}{\eta_1}\right)\mathrm{d}p \tag{A3.28}$$

since

$$\int \frac{\mathrm{d}x}{(x^2-1)^{1/2}} = \cosh^{-1}(x)$$

The first term on the right-hand side of Eq. (A3.28) is zero because the epicentral angle Δ is zero when $p = \eta_0$, and, when $p = \eta_1$, $\cosh^{-1}(p/\eta_1)$ is zero. The second term simplifies to

$$-\int_{\Delta=\Delta_1}^{0} \cosh^{-1}\left(\frac{p}{\eta_1}\right) \mathrm{d}\Delta$$

or

$$\int_{\Delta=0}^{\Delta_1} \cosh^{-1}\left(\frac{p}{\eta_1}\right) \mathrm{d}\Delta$$

where Δ_1 is the value of Δ for the ray with parameter η_1 (which has its deepest point at $r = r_1$) and $\Delta = 0$, for the ray with parameter η_0. Thus the left-hand side of Eq. (A3.27) is simplified to

$$\int_{p=\eta_1}^{\eta_0} \frac{\Delta\,\mathrm{d}p}{\left(p^2-\eta_1^2\right)^{1/2}} = \int_{\Delta=0}^{\Delta_1} \cosh^{-1}\left(\frac{p}{\eta_1}\right) \mathrm{d}\Delta \tag{A3.29}$$

The right-hand side of Eq. (A3.27) is handled by first performing the p integration:

$$\int_{p=\eta_1}^{\eta} \frac{p}{\left(p^2-\eta_1^2\right)^{1/2}(\eta^2-p^2)^{1/2}} \mathrm{d}p$$

On making the substitution $x = p^2$, we obtain

$$\begin{aligned}\frac{1}{2}\int_{x=\eta_1^2}^{\eta^2} \frac{\mathrm{d}x}{\left(x-\eta_1^2\right)^{1/2}(\eta^2-x)^{1/2}} &= \left[\tan^{-1}\left(\frac{x-\eta_1^2}{\eta^2-x}\right)^{1/2}\right]_{x=\eta_1^2}^{\eta^2} \\ &= \tan^{-1}(\infty) - \tan^{-1}(0) = \frac{\pi}{2}\end{aligned} \tag{A3.30}$$

(Reference works such as the *Standard Mathematical Tables*, edited by S. M. Selby, Chemical Rubber Company, are invaluable in solving integrals such as these. Alternatively, if we make the substitution $p^2 = \eta_1^2 \sin^2\theta + \eta^2 \cos^2\theta$ the integral simplifies to $\int_{\theta=0}^{\pi/2} \mathrm{d}\theta$.)

The solution to Eq. (A3.27) is now provided by Eqs. (A3.29) and (A3.30):

$$\begin{aligned}\int_{\Delta=0}^{\Delta_1} \cosh^{-1}\left(\frac{p}{\eta_1}\right) \mathrm{d}\Delta &= \int_{\eta=\eta_1}^{\eta_0} \frac{2}{r}\frac{\mathrm{d}r}{\mathrm{d}\eta}\frac{\pi}{2}\mathrm{d}\eta \\ &= \pi \int_{r=r_1}^{r_0} \frac{\mathrm{d}r}{r} = \pi[\log_\mathrm{e} r]_{r=r_1}^{r_0} = \pi \log_\mathrm{e}\left(\frac{r_0}{r_1}\right)\end{aligned} \tag{A3.31}$$

This equation now allows the velocity at any depth to be evaluated from the t–Δ curves, provided that certain conditions are met. As was shown in Eq. (A3.13), p is the slope of the t–Δ curve, and $\mathrm{d}t/\mathrm{d}\Delta$ is a function of Δ. For chosen values of Δ_1 and η_1 (the value of $\mathrm{d}t/\mathrm{d}\Delta$ at Δ_1), the integral on the left-hand side of Eq. (A3.31) can be evaluated and r_1 determined. Repeating the calculations for all possible values of η_1 means that r_1 is determined as a function of η_1. Recalling from Eq. (A3.21) that $\eta = r/v$, this determination means that the seismic velocity has been determined as a function of radius.

Such an inversion (due to Herglotz, Wiechert, Rasch and others and dating from 1907) has been invaluable in enabling us to evaluate the seismic structure of the interior of the Earth. It is generally called the *Herglotz–Wiechert inversion*. The main limitations of the method stem from the mathematical restriction that $\eta = r/v$ must decrease with depth (i.e., increase with increasing radius). Thus Eq. (A3.31) cannot be used in situations in which r/v increases with

depth, which in practice means in low-velocity regions. Within the Earth, therefore, the method fails for those parts of the upper mantle where there are low-velocity zones and at the core–mantle boundary. Other difficulties occur because of the lack of exact spherical symmetry within the Earth and the fact that the time–distance curves are not completely error-free (this means that the S-wave structure is less well determined than the P-wave structure since S-wave arrival times are more difficult to pick).

Appendix 4
The least-squares method

In geophysics it is often useful to be able to fit straight lines or curves to data (e.g., in radioactive dating and seismology). Although the eye is a good judge of what is and is not a good fit, it is unable to give any numerical estimates of errors. The method of least squares fills this need.

Suppose that $t_1, \ldots, t_n$ are the measured values of t (e.g., for travel times in seismology) corresponding to values $x_1, \ldots, x_n$ of quantity x (e.g., distance). Assume that the x values are accurate but the t values are subject to error. Further assume that we want to find the particular straight line

$$t = mx + c \tag{A4.1}$$

which fits the data best. If we substitute the value $x = x_i$ into Eq. (A4.1), the resulting value of t might not equal t_i. There may be some error e_i:

$$e_i = mx_i + c - t_i \tag{A4.2}$$

In the least-squares method, the values of m and c are chosen so that the sum of the squares of the errors e_i is least. In other words, $\sum_{i=1}^{n} e_i^2$ is minimized, where

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (mx_i + c - t_i)^2 \tag{A4.3}$$

To minimize this sum, it must be partially differentiated with respect to m, the result equated to zero and the process repeated for c. The two equations are then solved for m and c:

$$\begin{aligned} 0 &= \frac{\partial}{\partial m}\left(\sum_{i=1}^{n} e_i^2\right) \\ &= \frac{\partial}{\partial m}\left(\sum_{i=1}^{n} (mx_i + c - t_i)^2\right) \\ &= \sum_{i=1}^{n} 2x_i(mx_i + c - t_i) \\ &= 2m\sum_{i=1}^{n} x_i^2 + 2c\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} x_i t_i \end{aligned} \tag{A4.4}$$

and

$$\begin{aligned} 0 &= \frac{\partial}{\partial c}\left(\sum_{i=1}^{n} e_i^2\right) \\ &= \frac{\partial}{\partial c}\left(\sum_{i=1}^{n} (mx_i + c - t_i)^2\right) \\ &= \sum_{i=1}^{n} 2(mx_i + c - t_i) \end{aligned} \tag{A4.5}$$

Rearranging Eqs. (A4.4) and (A4.5) gives

$$\sum_{i=1}^{n} x_i t_i = m \sum_{i=1}^{n} x_i^2 + c \sum_{i=1}^{n} x_i \tag{A4.6}$$

and

$$\sum_{i=1}^{n} t_i = m \sum_{i=1}^{n} x_i + nc \tag{A4.7}$$

Equations (A4.6) and (A4.7) are simultaneous equations, which are solved to give m and c:

$$m = \frac{n\left(\sum_{i=1}^{n} x_i t_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} t_i\right)}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{A4.8}$$

$$c = \frac{\left(\sum_{i=1}^{n} t_i\right)\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i t_i\right)}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{A4.9}$$

The standard errors in these values of m and c are δm and δc (these are one-standard-deviation, 1σ, errors), given by

$$(\delta m)^2 = \frac{n\left(\sum_{i=1}^{n} e_i^2\right)}{(n-2)\left[n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2\right]} \tag{A4.10}$$

and

$$(\delta c)^2 = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} e_i^2\right)}{(n-2)\left[n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2\right]} \tag{A4.11}$$

Equations (A4.8)–(A4.11) can easily be programmed. Two-standard-deviation, 2σ, errors are generally quoted in geochronology.

The least-squares method can be applied also to curve fitting in exactly the same way as is shown here for straight lines. However, it becomes more difficult to solve the simultaneous equations when more than two coefficients need to be determined.

Appendix 5
The error function

The error function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{y=0}^{x} \mathrm{e}^{-y^2}\, \mathrm{d}y \tag{A5.1}$$

It is apparent that

$$\operatorname{erf}(-x) = -\operatorname{erf}(x) \tag{A5.2}$$

and

$$\operatorname{erf}(0) = 0 \tag{A5.3}$$

and

$$\operatorname{erf}(\infty) = 1 \tag{A5.4}$$

The complementary error function $\operatorname{erfc}(x)$ is defined as

$$\begin{aligned} \operatorname{erfc}(x) &= 1 - \operatorname{erf}(x) \\ &= \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} \mathrm{e}^{-y^2}\, \mathrm{d}y \end{aligned} \tag{A5.5}$$

The error function is shown in Fig. 7.5 and tabulated in Table A5.1. An easily programmable approximation to the error function is

$$\operatorname{erf}(x) = 1 - \left(a_1 t + a_2 t^2 + a_3 t^3\right)\mathrm{e}^{-x^2} + \varepsilon(x) \tag{A5.6}$$

where $t = 1/(1 + 0.470\,47x)$, $a_1 = 0.348\,024\,2$, $a_2 = -0.095\,879\,8$ and $a_3 = 0.747\,855\,6$. The error in this approximation is $\varepsilon(x) \leq 2.5 \times 10^{-5}$. (C. Hastings, *Approximations for Digital Computers*, Princeton University Press, Princeton, 1955.)

In this text, the error function appears in solutions of the heat-conduction equation (see Section 7.3.6). In more detailed thermal problems, the solutions may include repeated integrations or derivatives of the error function. For example,

$$\int_{x}^{\infty} \operatorname{erfc}(y)\, \mathrm{d}y = \frac{1}{\sqrt{\pi}} \mathrm{e}^{-x^2} - x \operatorname{erfc}(x)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}(\operatorname{erf}(x)) = \frac{2}{\sqrt{\pi}} \mathrm{e}^{-x^2}$$

Table A5.1 *The error function*

x	erf(x)
0.05	0.056 372
0.10	0.112 463
0.15	0.167 996
0.20	0.222 703
0.25	0.276 326
0.30	0.328 627
0.35	0.379 382
0.40	0.428 392
0.45	0.475 482
0.50	0.520 500
0.55	0.563 323
0.60	0.603 856
0.65	0.642 029
070	0.677 801
0.75	0.711 156
0.80	0.742 101
0.85	0.770 668
0.90	0.796 908
0.95	0.820 891
1.00	0.842 701
1.1	0.880 205
1.2	0.910 314
1.3	0.934 008
1.4	0.952 285
1.5	0.966 105
1.6	0.976 348
1.7	0.983 790
1.8	0.989 091
1.9	0.992 790
2.0	0.995 322
2.5	0.999 593
3.0	0.999 978

(For more values see H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, 2nd edn, Oxford University Press, Oxford, 1959.)

Appendix 6
Units and symbols

Conversion factors

Time

1 day $= 1.44 \times 10^3$ minutes (min) $= 8.64 \times 10^4$ seconds (s)
1 year (a) $= 8.76 \times 10^3$ hours (h) $= 5.26 \times 10^6$ min $= 3.16 \times 10^7$ s
1 Ma $= 3.16 \times 10^{13}$ s
1 Ga $= 10^3$ Ma $= 10^9$ yr (a)

Length

1 metre (m) $= 100$ cm $= 10^3$ millimetres (mm) $= 10^6$ micrometres (μm) $= 10^8$ ångström units (Å)
1 kilometre (km) $= 10^3$ m
1 fathom $= 6$ ft $= 1.8288$ m
1 nautical mile $= 1.852$ km
60 nautical miles $= 1^\circ$ latitude

Area

1 $m^2 = 10^4$ cm^2
1 $km^2 = 10^6$ m^2

Volume

1 $m^3 = 10^3$ litres
1 $km^3 = 10^{12}$ litres

Velocity

1 m $s^{-1} = 3.6$ km h^{-1}
1 km $s^{-1} = 10^3$ m $s^{-1} = 3.6 \times 10^3$ km h^{-1}

Angle

1 radian (rad) $= 57.30^\circ = 57^\circ 18'$
$1^\circ = 0.017\,45$ rad

Mass

1 kilogram (kg) = 1000 grams (g)

Force

1 newton (N) = 1 kg m s^{-2} = 10^5 dynes = 10^5 g cm s^{-2}

Pressure

1 pascal (Pa) = 1 N m^{-2} = 1 kg m^{-1} s^{-2} = 10^{-5} bar
= 10^{-8} kilobars (kbar)
1 MPa = 10^6 Pa = 10^6 N m^{-2}
1 GPa = 10^9 Pa = approximate pressure at the base of a 30-km-high column of rock
1 atmosphere (atm) = pressure at the base of a 76-cm-high column of mercury
= 1.013×10^5 Pa

Energy, work, heat

1 joule (J) = 1 kg m^2 s^{-2} = 10^7 ergs = 0.2389 calories (cal)
= 2.389×10^{-4} kcal
1 kcal = 4185 J

Power

1 watt (W) = 1 joule/second (J s^{-1}) = 0.2389 cal s^{-1}
= 2.389×10^{-4} kcal s^{-1}
1 kilowatt (kW) = 1000 W
1 kW h = 3.6×10^6 J

Heat-flow rate across a surface

1 W m^{-2} = 2.389×10^{-5} cal cm^{-2} s^{-1}
1 cal cm^{-2} s^{-1} = 4.18×10^4 W m^{-2}
1 heat-flow unit (hfu) = 10^{-6} cal cm^{-2} s^{-1} = 4.18×10^{-2} W m^{-2}

Heat-generation rate

1 W kg^{-1} = 7.54×10^3 cal g^{-1} a^{-1}
1 W m^{-3} = 2.389×10^{-7} cal cm^{-3} s^{-1}
1 cal cm^{-3} s^{-1} = 4.18×10^6 W m^{-3}
1 heat-generation unit (hgu) = 10^{-13} cal cm^{-3} s^{-1} = 4.18×10^{-7} W m^{-3}
= 0.418 μW m^{-3}

Thermal conductivity

$1\ \mathrm{W\ m^{-1}\ {}^{\circ}C^{-1}} = 2.389 \times 10^{-3}\ \mathrm{cal\ cm^{-1}\ s^{-1}\ {}^{\circ}C^{-1}}$
$1\ \mathrm{cal\ cm^{-1}\ s^{-1}\ {}^{\circ}C^{-1}} = 4.18 \times 10^{2}\ \mathrm{W\ m^{-1}\ {}^{\circ}C^{-1}}$

Specific heat

$1\ \mathrm{J\ kg^{-1}\ {}^{\circ}C^{-1}} = 2.389 \times 10^{-4}\ \mathrm{cal\ g^{-1}\ {}^{\circ}C^{-1}}$
$1\ \mathrm{cal\ g^{-1}\ {}^{\circ}C^{-1}} = 4.18 \times 10^{3}\ \mathrm{J\ kg^{-1}\ {}^{\circ}C^{-1}}$

Latent heat

$1\ \mathrm{J\ kg^{-1}} = 2.389 \times 10^{-4}\ \mathrm{cal\ g^{-1}}$
$1\ \mathrm{cal\ g^{-1}} = 4.18 \times 10^{3}\ \mathrm{J\ kg^{-1}}$

Diffusivity

$1\ \mathrm{m^2\ s^{-1}} = 10^{4}\ \mathrm{cm^2\ s^{-1}}$
$1\ \mathrm{cm^2\ s^{-1}} = 10^{-4}\ \mathrm{m^2\ s^{-1}}$

Temperature

degrees Kelvin (K) = degrees Celsius (°C) + 273.16

Density

$1\ \mathrm{kg\ m^{-3}} = 10^{-3}\ \mathrm{g\ cm^{-3}}$
$1\ \mathrm{g\ cm^{-3}} = 10^{3}\ \mathrm{kg\ m^{-3}}$

Dynamic viscosity

1 pascal second (Pa s) = $1\ \mathrm{N\ m^{-2}\ s}$
$1\ \mathrm{Pa\ s} = 10\ \text{poise} = 10\ \mathrm{g\ cm^{-1}\ s^{-1}}$

Kinematic viscosity

$1\ \mathrm{m^2\ s^{-1}} = 10^{4}\ \mathrm{cm^2\ s^{-1}}$

Frequency

1 hertz (Hz) = 1 cycle per second

Magnetic induction

1 tesla (T) = $1\ \mathrm{kg\ A^{-1}\ s^{-2}} = 10^{4}$ gauss $= 10^{9}$ gamma (γ)

Table A6.1 *Symbols*

Symbol	Name	Units	First equation
A	Activity	s^{-1}	(6.6)
A	Radioactive heat generation rate per unit volume	$W\,m^{-3}$	(7.7)
A	Arrhenius constant		(10.26)
A	Amplitude		(4.13)
A	Area		(4.20)
a	Gravitational acceleration	$m\,s^{-2}$	(5.3)
a	Area	m^2	(7.6)
$\mathbf{B}$	Magnetic field	T	(3.2)
B	Amplitude		(4.56)
b	Radius	m	(5.7)
C	Concentration of reactant		(10.25)
c_P	Specific heat at constant pressure	$J\,kg^{-1}\,{}^{\circ}C^{-1}$	(7.9)
D	Declination	degrees	(3.19)
D	Distance	m	(5.7)
D	Compensation depth	m	(5.28)
D	Flexural rigidity	N m	(5.56)
D	Number of daughter atoms		(6.9)
D	Diffusion coefficient	$m^2\,s^{-1}$	(6.24)
d	Depth to sediment bed	m	(10.5)
d	Depth	m	(4.100)
d	Ocean depth	km	(7.57)
d_s	Sediment thickness	m	(10.1)
d_w	Water depth	m	(10.1)
D_I	Number of decays by induced fission		(6.58)
D_R	Number of radioactive decays		(6.56)
D_S	Number of decays by spontaneous fission		(6.57)
E	Energy	$kg\,m^2\,s^{-2}$	(4.27)
E	Young's modulus	Pa	(5.57)
E	Activation energy	$J\,mol^{-1}$	(6.24)
e	Angle		(4.55)
e	Elevation	m	(8.38)
e	Strain		(A2.1)
F	Force	N	(5.1)
F_{RP}	Ridge-push force per unit length	$N\,m^{-1}$	(8.38)

(*cont.*)

Table A6.1 (*cont.*)

Symbol	Name	Units	First equation
F_{SP}	Slab-pull force per unit length	$N\ m^{-1}$	(8.39)
f	Frequency	s^{-1}	(4.7)
f	Angle		(4.55)
f	Ellipticity		(5.17)
G	Gravitational constant	$m^3\ kg^{-1}\ s^{-2}$	(5.1)
g	Gravitational acceleration	$m\ s^{-2}$	(5.18)
ΔG^*	Free energy of activation	$J\ mol^{-1}$	(10.7)
g_e	Gravitational acceleration at the equator	$m\ s^{-2}$	(5.19)
g_{rot}	Gravitational acceleration of a rotating sphere	$m\ s^{-2}$	(5.18)
H	Horizontal component of the Earth's magnetic field	T	(3.15)
H	Horizontal force per unit length	$N\,m^{-1}$	(5.56)
h	Focal depth		(4.12)
h	Height	m	(5.23)
Δh	Geoid height anomaly	m	(5.48)
H_a	Enthalpy of activation	$J\ mol^{-1}$	(10.7)
I	Angle of inclination	degrees	(3.16)
i	Angle		(4.33)
j	Angle		(A3.1)
K	Bulk modulus or incompressibility	$Pa = N\,m^{-2}$	(4.5)
k	Thermal conductivity	$W\,m^{-1}\,{}^\circ C^{-1}$	(7.1)
k	Reaction rate coefficient		(10.25)
L	Skin depth	m	(7.41)
L	Thickness of the lithosphere	m	(7.63)
l	Length	m	(10.16)
M	Induced magnetization	T	(3.21)
M	Earthquake magnitude		(4.13)
M	Mass of a sphere	kg	(5.15)
M	Horizontal bending moment per unit length	N	(5.64)
m	Dipole moment	$A\ m^2$	(3.1)
m	Mass	kg	(5.1)
m_b	Body-wave magnitude		(4.17)
M_E	Mass of the Earth	kg	(5.15)
M_o	Seismic moment	N m	(4.20)
M_r	Mass of the Earth within a sphere of radius r	kg	(8.5)

Table A6.1 (*cont.*)

Symbol	Name	Units	First equation
M_S	Surface-wave magnitude		(4.14)
M_w	Moment magnitude		(4.21)
N	Number		(4.24)
N	Number of parent atoms		(6.1)
n	Neutron dose	cm^{-2}	(6.58)
N_I	Number of induced fission tracks		(6.59)
N_S	Number of spontaneous fissions		(6.59)
Nu	Nusselt number		(8.34)
P	Pressure	$Pa = N\,m^{-2}$	(4.30)
p	Seismic ray parameter	s degree^{-1}	(A3.7)
Pe_t	Peclet number		(8.36)
Q	Königsberger ratio		
Q	Quality factor		(8.21)
Q	Rate of flow of heat per unit area	$W\,m^{-2}$	(7.2)
R	Radius of the Earth	m	(2.3)
R	Gas constant	J mol^{-1} °C^{-1}	(6.24)
r	Radius	m	(3.1)
r	Depth of root	m	(5.23)
Ra	Rayleigh number		(8.31)
Re	Reynolds number		(8.33)
R_e	Equatorial radius of the Earth	m	(5.17)
S	Entropy	J kg^{-1} °C^{-1}	(7.86)
T	Temperature	°C, K	(6.24)
T	Age of the Earth		(6.61)
T	Period	s	(4.13)
t	Time	s	(4.1)
t	Thickness	m	(5.31)
T_p	Potential temperature	°C	(7.95)
$T_{1/2}$	Half-life		(6.7)
TTI	Time–temperature index		(10.27)
U	Group velocity	km s^{-1}	(4.7)
u	Displacement	m	(4.3)
u	Velocity	m s^{-1}	(7.18)
V	Magnetic potential	A (amp)	(3.1)
V	Phase velocity	km s^{-1}	(4.7)
V	Volume	m^3	(4.36)
V	Gravitational potential	$m^2\,s^{-2}$	(5.2)

(*cont.*)

Table A6.1 (*cont.*)

Symbol	Name	Units	First equation
V	Vertical force per unit length	N m^{-1}	(5.56)
v	Relative velocity	cm yr^{-1}	(2.1)
u	Seismic velocity	km s^{-1}	(4.6)
v_p	P-wave velocity	km s^{-1}	
v_s	S-wave velocity	km s^{-1}	
w	Width	m	(4.100)
w	Vertical deflection	m	(5.56)
x	Horizontal distance	m	(2.21)
y	Horizontal distance	m	(2.22)
Z	Inward radial component of the Earth's magnetic field	T	(3.14)
z	Depth	m	(2.23)
α	P-wave velocity	km s^{-1}	(4.1)
α	Flexural parameter	m	(5.61)
α	Coefficient of thermal expansion	$^{\circ}$C^{-1}	(7.89)
β	Angle	degrees, radians	(2.8)
β	S-wave velocity	km s^{-1}	(4.2)
β	Stretching factor		(10.8)
Δ	Angular distance	degrees, radians	(4.13)
Δ	Cubical dilatation		(A2.16)
δ	Dip		(4.42)
η	Dynamic viscosity	Pa s	(8.28)
θ	Angle	degrees	(2.3)
θ	Colatitude	degrees	
κ	Thermal diffusivity	m^2 s^{-1}	(7.43)
κ	Wavenumber		(4.25)
λ	Latitude	degrees	(2.4)
λ	Wavelength		
λ	Radioactive decay constant		(6.1)
λ	Lamé elastic constant	Pa	(A2.19)
μ	Shear modulus or Lamé elastic constant	Pa	(4.4)
μ_0	Magnetic permeability of free space	T m A^{-1}	(3.2)
ν	Kinematic viscosity	m^2 s^{-1}	(8.31)
ρ	Density	kg m^{-3}	(4.4)
σ	Poisson's ratio		(5.57)
σ	Neutron-capture cross section		(6.58)

Table A6.1 (*cont.*)

Symbol	Name	Units	First equation
σ	Stress	Pa	(A2.18)
τ	Temperature difference	°C	(8.18)
ϕ	Longitude	degrees	(2.6)
ϕ	Seismic parameter		
ϕ	Phase angle		(7.42)
Φ	Seismic scalar displacement potential	m^2	(4.1)
χ	Magnetic susceptibility		(3.18)
Ψ	Seismic vector displacement potential	m^2	(4.2)
ω	Angular velocity	10^{-7} degrees yr^{-1}	(2.3)
ω	Angular frequency		(4.25)

Table A6.2 *Multipliers for powers of ten*

n	nano-	10^{-9}	k	kilo-	10^3
μ	micro-	10^{-6}	M	mega-	10^6
m	milli-	10^{-3}	G	giga-	10^9

Table A6.3 *The Greek alphabet*

Alpha	A	α	Nu	ν	ν
Beta	B	β	Xi	ξ	ξ
Gamma	Γ	γ	Omicron	O	o
Delta	Δ	δ	Pi	Π	π
Epsilon	ϵ	ε	Rho	ρ	ρ
Zeta	ζ	ζ	Sigma	Σ	σ
Eta	η	η	Tau	τ	τ
Theta	Θ	θ	Upsilon	ϒ	ν
Iota	ι	ι	Phi	Φ	ϕ
Kappa	κ	κ	Chi	X	χ
Lambda	Λ	λ	Psi	ψ	ψ
Mu	M	μ	Omega	Ω	ω

Appendix 7
Numerical data

Physical constants

Gravitational constant, G	6.673×10^{-11} m^3 kg^{-1} s^{-2}
Gas constant, R	8.3145 J mol^{-1} $^\circ$C^{-1}
Permeability of free space (vacuum) μ_0	$4\pi \times 10^{-7}$ kg m A^{-2} s^{-2}

The Earth

Age of the Earth, T	4550 Ma
Angular velocity of the Earth	7.292×10^{-5} rad s^{-1}
Mean distance to the Sun	1.5×10^{11} km
Average velocity around the Sun	29.77 km s^{-1}
Length of solar day	8.64×10^4 s
Length of year	3.1558×10^7 s
Equatorial radius, R_{eq}	6378.14 km
Polar radius, R_p	6356.75 km
Polar flattening, f	1/298.247
Radius of outer core	3480 km
Radius of inner core	1221 km
Volume of the Earth	1.083×10^{21} m^3
Volume of crust	approximately 10^{19} m^3
Volume of mantle	9.0×10^{20} m^3
Volume of core	1.77×10^{20} m^3
Mass of the Sun	1.99×10^{30} kg
Mass of the Moon	7.35×10^{22} kg
Mass of the Earth, M_E	5.97×10^{24} kg
Mass of the oceans	1.4×10^{21} kg
Mass of the crust	2.8×10^{22} kg
Mass of the mantle	4.00×10^{24} kg
Mass of the core	1.94×10^{24} kg
Mean density of the Earth	5.52×10^3 kg m^{-3}
Mean density of the mantle	4.5×10^3 kg m^{-3}
Mean density of the core	1.1×10^4 kg m^{-3}

Equatorial gravity at sea level, g_e	9.780 318 5 m s^{-2}
Polar gravity at sea level, g_p	9.832 177 3 m s^{-2}
Surface area	5.10×10^{14} m^2
Area of continents and continental shelves	2.01×10^{14} m^2
Area of oceans and ocean basins	3.09×10^{14} m^2
Mean depth of the oceans	3.8 km
Mean height of land	0.84 km

Appendix 8
The IASP91 Earth Model

Depth z (km)	Radius r (km)	P-wave velocity α (km s^{-1})	S-wave velocity β (km s^{-1})
5153.9–6371	0–1217.1	$11.240\,94 - 4.096\,89x^2$	$3.564\,54 - 3.452\,41x^2$
2889–5153.9	1217.1–3482	$10.039\,04 + 3.756\,65x - 13.670\,46x^2$	0
2740–2889	3482–3631	$14.494\,70 - 1.470\,89x$	$8.166\,16 - 1.582\,06x$
760–2740	3631–5611	$25.1486 - 41.1538x + 51.9932x^2 - 26.6083x^3$	$12.9303 - 21.2590x + 27.8988x^2 - 14.1080x^3$
660–760	5611–5711	$25.969\,84 - 16.934\,12x$	$20.768\,90 - 16.531\,47x$
410–660	5711–5961	$29.388\,96 - 21.406\,56x$	$17.707\,32 - 13.506\,52x$
210–410	5691–6161	$30.787\,65 - 23.254\,15\,x$	$15.242\,13 - 11.085\,52x$
120–210	6161–6251	$25.413\,89 - 17.697\,22x$	$5.750\,20 - 1.274\,20x$
35–120	6251–6336	$8.785\,41 - 0.749\,53x$	$6.706\,231 - 2.248\,585x$
20–35	6336–6351	6.50	3.75
0–20	6351–6371	5.80	3.36

Note: The variable $x = r/6371$ is the normalized radius.

Source: From Kennett, B. L. N., Engdahl, E. R. and Buland, R. 1995. Constraints on seismic velocities in the Earth from travel times. *Geophys. J. Int.*, **105**, 429–65.

Appendix 9

The Preliminary Reference Earth Model, isotropic version – PREM

Radius (km)	Depth (km)	α (km s^{-1})	β (km s^{-1})	Density (kg m^{-3})	Pressure (GPa)	Gravity (m s^{-2})
0	6371	11.2622	3.6678	13088.5	364.0	0.00
100	6271	11.2606	3.6667	13086.3	363.7	0.37
200	6171	11.2559	3.6634	13079.8	363.0	0.73
300	6071	11.2481	3.6579	13068.9	361.8	1.10
400	5971	11.2371	3.6503	13053.7	360.2	1.46
500	5871	11.2230	3.6404	13034.1	358.0	1.82
600	5771	11.2058	3.6284	13010.1	355.4	2.19
700	5671	11.1854	3.6141	12981.8	352.3	2.55
800	5571	11.1619	3.5977	12949.1	348.8	2.91
900	5471	11.1352	3.5790	12912.1	344.8	3.27
1000	5371	11.1054	3.5582	12870.8	340.4	3.62
1100	5271	11.0725	3.5352	12825.0	335.5	3.98
1200	5171	11.0364	3.5100	12775.0	330.2	4.33
1221.5	5149.5	11.0283	3.5043	12763.6	329.0	4.40
1221.5	5149.5	10.3557	0.0	12166.3	329.0	4.40
1300	5071	10.3097	0.0	12125.0	324.7	4.64
1400	4971	10.2496	0.0	12069.2	318.9	4.94
1500	4871	10.1875	0.0	12009.9	312.7	5.25
1600	4771	10.1229	0.0	11946.8	306.3	5.56
1700	4671	10.0558	0.0	11879.9	299.5	5.86
1800	4571	9.9856	0.0	11809.0	292.3	6.17
1900	4471	9.9121	0.0	11734.0	284.9	6.47
2000	4371	9.8350	0.0	11654.8	277.1	6.77
2100	4271	9.7540	0.0	11571.2	269.1	7.07
2200	4171	9.6687	0.0	11483.1	260.8	7.37
2300	4071	9.5788	0.0	11390.4	252.2	7.66
2400	3971	9.4841	0.0	11190.7	234.2	8.23
2600	3771	9.2788	0.0	11083.4	224.9	8.51
2700	3671	9.1676	0.0	10970.9	215.4	8.78

(*cont.*)

Radius (km)	Depth (km)	α (km s^{-1})	β (km s^{-1})	Density (kg m^{-3})	Pressure (GPa)	Gravity (m s^{-2})
2800	3571	9.0502	0.0	10853.2	205.7	9.04
2900	3471	8.9264	0.0	10730.1	195.8	9.31
3000	3371	8.7958	0.0	10601.5	185.7	9.56
3100	3271	8.6581	0.0	10467.3	175.5	9.81
3200	3171	8.5130	0.0	10327.3	165.2	10.05
3300	3071	8.3602	0.0	10181.4	154.8	10.28
3400	2971	8.1994	0.0	10029.4	144.3	10.51
3480	2891	8.0648	0.0	9903.4	135.8	10.69
3480	2891	13.7166	7.2647	5566.5	135.8	10.69
3500	2871	13.7117	7.2649	5556.4	134.6	10.66
3600	2771	13.6876	7.2657	5506.4	128.8	10.52
3630	2741	13.6804	7.2660	5491.5	127.0	10.49
3700	2671	13.5960	7.2340	5456.6	123.0	10.41
3800	2571	13.4774	7.1889	5406.8	117.4	10.31
3900	2471	13.3608	7.1442	5357.1	111.9	10.23
4000	2371	13.2453	7.0997	5307.3	106.4	10.16
4100	2271	13.1306	7.0552	5257.3	101.1	10.10
4200	2171	13.0158	7.0105	5207.2	95.8	10.06
4300	2071	12.9005	6.9653	5156.7	90.6	10.02
4400	1971	12.7839	6.9195	5105.9	85.5	9.99
4500	1871	12.6655	6.8728	5054.7	80.4	9.97
4600	1771	12.5447	6.8251	5003.0	75.4	9.95
4700	1671	12.4208	6.7760	4950.8	70.4	9.94
4800	1571	12.2932	6.7254	4897.9	65.5	9.93
4900	1471	12.1613	6.6731	4844.3	60.7	9.93
5000	1371	12.0245	6.6189	4789.9	55.9	9.94
5100	1271	11.8821	6.5624	4734.6	51.2	9.94
5200	1171	11.7336	6.5036	4678.5	46.5	9.95
5300	1071	11.5783	6.4423	4621.3	41.9	9.96
5400	971	11.4156	6.3781	4563.1	37.3	9.97
5500	871	11.2449	6.3108	4503.8	32.8	9.99
5600	771	11.0656	6.2404	4443.2	28.3	10.00
5701	670	10.7513	5.9451	4380.7	23.8	10.02
5701	670	10.2662	5.5702	3992.1	23.8	10.02
5771	600	10.1578	5.5160	3975.8	21.1	10.01
5771	600	10.1578	5.5159	3975.8	21.1	10.01
5800	571	10.0094	5.4313	3939.3	19.9	10.00
5900	471	9.4974	5.1396	3813.2	16.0	9.99
5971	400	9.1339	4.9325	3723.7	13.4	9.97
5971	400	8.9052	4.7699	3543.3	13.4	9.97
6000	371	8.8495	4.7496	3525.9	12.3	9.96

Radius (km)	Depth (km)	α (km s^{-1})	β (km s^{-1})	Density (kg m^{-3})	Pressure (GPa)	Gravity (m s^{-2})
6100	271	8.6571	4.6796	3466.2	8.9	9.93
6151	220	8.5589	4.6439	3435.8	7.1	9.91
6151	220	7.9897	4.4189	3359.5	7.1	9.91
6200	171	8.0200	4.4370	3364.8	5.5	9.89
6291	80	8.0762	4.4705	3374.7	2.5	9.86
6291	80	8.0762	4.4705	3374.7	2.5	9.86
6300	71	8.0818	4.4738	3375.7	2.2	9.86
6346	24.4	8.1106	4.4910	3380.7	0.6	9.84
6346	24.4	6.8000	3.9000	2900.0	0.6	9.84
6356	15	6.8000	3.9000	2900.0	0.3	9.84
6356	15	5.8000	3.2000	2600.0	0.3	9.84
6368	3	5.8000	3.2000	2600.0	0.0	9.83
6368	3	1.4500	0.0	1020.0	0.0	9.83
6371	0	1.4500	0.0	1020.0	0.0	9.82

Source: From Dziewonski, A. M. and Anderson, D. L. 1981. Preliminary Reference Earth Model. *Phys. Earth Planet. Inter.*, **25**, 297–356.

Appendix 10

The Modified Mercalli Intensity Scale (abridged version)

I. Not felt except by a very few under especially favourable conditions.

II. Felt only by a few persons at rest, especially on upper floors of buildings.

III. Felt quite noticeably by persons indoors, especially on upper floors of buildings. Many people do not recognize it as an earthquake. Standing motor cars may rock slightly. Vibrations similar to the passing of a truck.

IV. Felt indoors by many, outdoors by few during the day. At night, some awakened. Dishes, windows, doors disturbed; walls make cracking sound. Sensation like heavy truck striking building. Standing motor cars rocked noticeably.

V. Felt by nearly everyone; many awakened. Some dishes, windows broken. Unstable objects overturned.

VI. Felt by all, many frightened. Some heavy furniture moved; a few instances of fallen plaster. Damage slight.

VII. Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable damage in poorly built or badly designed structures; some chimneys broken.

VIII. Damage slight in specially desinged structure; considerable damage in ordinary substantial buildings with partial collapse. Damage great in poorly built structures. Fall of chimneys, factory stacks, columns, monuments, walls. Heavy furniture overturned.

IX. Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings shifted off foundatins. Underground pipes broken.

X. Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations. Rails bent.

XI. Few, if any (masonry) structures remain standing. Bridges destroyed. Rails bent greatly.

XII. Damage total. Waves seen on ground surface. Lines of sight and level are distorted. Objects thrown into the air

Glossary

This is a compilation of some of the technical terms used in this book. For more formal definitions, refer to J. A. Jackson, ed., 1997. *Glossary of Geology*, 4th edn, American Geological Institute, Falls Church, Virginia.

a abbreviation for year

abyssal plain deep, old ocean floor; well sedimented.

accreted terrain (terrane) terrane that has been accreted to a continent.

active margin continental margin characterized by volcanic activity and earthquakes (i.e., location of transform fault or subduction zone).

adiabat pressure–temperature path of a body that expands or contracts without losing or absorbing heat.

aeon (eon) longest division of geological time; also sometimes used for 10^9 years.

alpha decay radioactive decay by emission of an alpha (α) particle.

alpha particle nucleus of a helium atom (two protons and two neutrons).

altered rocks rocks that have undergone changes in their chemical or mineral structure since they were formed.

amphibolite intermediate-grade metamorphic rock; temperature attained above 400-450 °C; characterized by amphibole minerals such as hornblende.

andesite extrusive igneous rock, usually containing plagioclase and mafic phases(s); about 55% SiO_2. Usually associated with subduction zones.

anticline a fold, convex upwards, whose core contains stratigraphically older rocks.

Archaean (Archean) that division of geological time prior to ~2500 Ma ago.

aseismic region region with very infrequent earthquakes.

asthenosphere region beneath the lithosphere where deformation is dominantly plastic and heat is transferred mainly by convection; now sometimes means the entire upper mantle beneath the lithosphere. Literally, the 'sick' or 'weak' sphere.

atomic number number of protons in the nucleus of an atom.

back-arc basin basin on the overriding plate behind the volcanic arc of a subduction zone.

band-pass filter filters a signal to retain only those frequencies within the required range, e.g., 5–40 Hz.

basalt mafic igneous rock.

basement rock continental crust that provides the substrate for later deposition.

basin depression in which sediments collect.

batholith large body of igneous rock, several kilometres thick and extending over areas up to thousands of square kilometres.

bathymetry depth of the seabed.

beta decay radioactive decay by emission of an electron.

beta particle electron. An elementary particle with a charge of -1.

blueschist low-grade metamorphic rock; formed at lower temperatures and higher pressures than greenschist; characterized by blue minerals.

body wave seismic wave that travels through the interior of the Earth; P-waves are longitudinal body waves; S-waves are transverse body waves. Body waves are short-period ($\sim$0.1–20 Hz), short-wavelength ($<$100 km) waves for which ray-theory approximations are valid.

body-wave magnitude magnitude of an earthquake as estimated from the amplitude of body waves.

bulk modulus (K) bulk property of a material; equal to the pressure acting on a sample divided by the resultant fractional decrease in volume of that sample.

cation positively charged ion.

centripetal acceleration acceleration of a body with a circular motion, towards the centre of that circle; proportional to the square of the body's velocity and inversely proportional to the radius of the circle; depending on one's point of view, also called centrifugal acceleration.

CHUR chondritic uniform reservoir (see Section 6.8).

compensation depth depth at which the overlying rocks are assumed to exert a constant pressure; below this depth, there are no large lateral variations in density.

compressional waves another name for P-waves. See body waves.

conduction transfer of heat by molecular collisions.

continental rise part of the continental margin between the continental slope and the abyssal plain; slopes generally 1 : 40 to 1 : 2000.

continental shelf part of the continental margin between the coast and the continental slope; slopes about 0.1° (1 : 500).

continental slope part of the continental margin between the continental rise and the continental shelf; slopes about 3–6° (1 : 10 to 1 : 20).

convection transfer of heat by the physical movement of molecules from one place to another; hot, less-dense fluid rises and cool, denser fluid sinks.

cordillera mountains of western North America from the Rocky mountains to the Pacific coast.

core iron-rich centre of the Earth, 2889–6371 km below the surface.

cosmic rays atomic muclei, largely protons, travelling at or near relativistic speeds.

craton (1) large stable part of a continent, which has not been subject to deformation for a very long time (e.g., since the Precambrian); (2) distinct, tectonically coherent, large region of granitoid crust.

cumulate term applied to rocks formed by the accumulation (e.g., by precipitation) of crystals.

Curie point or temperature temperature above which a mineral cannot be permanently magnetized.

declination horizontal angle between geographic north and magnetic north.

decollement zone detachment (unsticking) zone between strata, due to deformation such as folding or thrusting.

dehydration loss of water.

density ratio of the mass of a material to its volume; usual symbol, ρ.

depleted mantle mantle that has been depleted of some elements by processes such as partial melting; residue after extraction of crust.

dextral fault *see* right-lateral fault.

diapir body of light material (e.g., salt, magma) that pierces upwards into overlying strata.

differentiate (1) mathematical term; (2) rock formed by magmatic differentiation (e.g., precipitation of crystals).

dilatation change in volume per unit volume.

dip–slip fault fault on which the movement is parallel to the dip of the fault; thrust fault or normal fault.

dunite rock dominantly composed of olivine.

dyke (dike) small igneous body that has intruded into fissures that cut across the existing rock strata.

earthquake sudden violent movement within the crust or upper mantle.

earthquake epicentre point on the Earth's surface immediately above the earthquake focus.

earthquake focus location of the earthquake within the Earth.

eclogite type of dense rock formed by metamorphism of basalt.

elastic limit maximum stress a body can withstand without being permanently deformed.

electron capture absorption of an electron by a nucleus.

endothermic reaction chemical reaction that requires heat in order to take place.

epicentre location of an earthquake specified by latitude and longitude. Full location including depth is termed the focus or hypocentre.

epicentral distance distance from the epicentre of an earthquake to some other point on the Earth's surface, measured along a great circle in degrees or kilometres.

erosion process by which rock is worn away and the material removed.

extrusion eruption of magma.

exothermic reaction chemical reaction that releases heat.

facies character of a rock; can be applied to the appearance, composition or physical environment in which the rock originated.

fault fracture in a rock body, along which motion has occurred or is still occurring.

feldspar mineral family $(XAl)(AlSi)_3O_8$, where X = K, Na, Ca; plagioclase feldspars have X = Na, Ca; alkali feldspars have X = Na, K.

fore-arc basin sedimentary basin on the trench side of the volcanic arc of a subduction zone.

foreland basin sedimentary basin on the trench side of a continent–continent collision zone.

fractionation separation of components in a magma (e.g., by precipitation of crystals).

free oscillation vibration of the whole Earth after a major earthquake.

frequency number of oscillations per unit time; unit is Hertz (Hz), which equals one cycle per second.

Ga 10^9 years, giga-annum.

gabbro coarse-grained intrusive equivalent of basalt; composed of calcic plagioclase, clinopyroxene, with or without olivine, and/or orthopyroxene.

gamma decay radioactive decay by emission of a gamma ray or photon (a short-wavelength electromagnetic wave).

geotherm temperature–depth curve within the Earth.

gneiss foliated metamorphic rock rich in quartz and feldspar; can be derived from igneous or sedimentary rocks.

Gondwanaland former continent that comprised present-day South America, Africa, India, Australia and Antarctica; started to break up about 150 Ma ago.

granite type of plutonic rock with quartz together with plagioclase-feldspar, the plagioclase making up between 10% and 65% of total feldspar.

granitoid family of granite-like rocks.

granodiorite type of plutonic rock with quartz together with plagioclase-feldspar, the plagioclase making up between 65% and 90% of total feldspar.

granulite high-grade (anhydrous) metamorphic rock; formed at temperatures of 700 °C or more and depths of 10 km or more, sometimes much more.

gravimeter instrument used to measure variations in the gravitational field.

gravity anomaly difference between the observed (measured) value of gravity and the theoretical value (e.g., the Bouguer anomaly, free-air anomaly, isostatic anomaly).

great circle the line which marks the intersection of a plane through the centre of the sphere with the surface of that sphere. Thus lines of longitude are great circles. Distance along a great circle is often quoted in degrees or radians.

greenschist low-grade metamorphic rock; temperature attained 300 °C or more, characterized by presence of green minerals.

greenstone belt belt of rocks usually including volcanic rocks and sediments at low-to-moderate metamorphic grade, often surrounded by granite or gneiss.

Gutenberg discontinuity discontinuity in seismic velocity that marks the boundary between the core and the mantle; named after seismologist Beno Gutenberg.

half-life time required for half of a given number of atoms of a radioactive element to decay.

harzburgite peridotite composed of olivine and orthopyroxene.

heat-producing element U, K, Th in the modern Earth, also Al and Pu in the early Earth.

high-pass filter filters a signal to remove frequencies below a given frequency, passes only the high frequencies.

hotspot localized region characterized by volcanism, high heat flow and uplift; believed to result from hot material rising from depth in the mantle.

hydration incorporation of water into the mineral structure.

hydrothermal circulation of hot water in rock (e.g., around a hot igneous body or a volcano).

igneous rocks rocks that were once molten but have since cooled and solidified.

inclination angle at which magnetic field lines dip.

incompatible element element that enters the melt during partial melting (e.g., Ti, Zr, Y).

inner core central solid region of the Earth's core, probably mostly iron; radius about 1221 km, discovered by Inge Lehmann in 1936.

instantaneous velocity velocity of a body at a specific instant.

interplate earthquake earthquake with its focus on a plate boundary.

intra-continental sedimentary basin sedimentary basin that formed within a continent.

intraplate earthquake earthquake with its focus within a plate.

intrusion movement of magma into country rock.

intrusive rock igneous rock that solidified within the Earth.

ion atom that is not electrically neutral; an atom with fewer than the normal number of electrons is a positive ion; an atom with more than the normal number of electrons is a negative ion.

island arc chain of islands above a subduction zone (e.g., Japan, Aleutians).

isostasy the way in which the lithosphere 'floats' on the asthenosphere.

isotherm line or surface of constant temperature.

isotopes forms of an element with differing numbers of neutrons in their nucleus. For hydrogen, for example, the nucleus of hydrogen, ^{1}H, consists of one proton; the nucleus of the hydrogen isotope deuterium, ^{2}H, consists of one proton and one neutron; the nucleus of tritium, ^{3}H, has one proton and two neutrons.

kbar kilobar, 0.1 GPa of pressure.

kerogen organic material that can be transformed into oil and gas.

kimberlite type of peridotite that erupts from great depth in a so-called pipe and occasionally contains diamonds.

kinematic viscosity dynamic viscosity of a fluid divided by its density; unity of kinematic viscosity, $m^2\ s^{-1}$; symbol, ν.

kinetic energy energy that something has by virtue of its motion.

komatiite highly magnesian lava.

large-ion lithophile (LIL) elements elements such as Rb, K, Ba, Th, Sr that have large ionic radii.

latent heat or heat of fusion the amount of heat which must be supplied to a unit quantity of a given solid material at its melting point to change it into a liquid; conversely, the (same) amount of heat released when a unit quantity of the given liquid material at its melting temperature is changed into a solid.

latitude angular distance north or south of the equator of a point on the Earth's surface.

Laurasia former continent that comprised present-day North America, Greenland, Europe and part of Asia; broke up during the Cretaceous, about 90 Ma ago.

lava molten rock when it is erupted at the Earth's surface.

left-lateral fault, sinistral fault fault on which the displacement is such that, to an observer on the ground, the opposite side is displaced to the left.

lherzolite peridotite composed of olivine, clinopyroxene and orthopyroxene.

liquidus surface in temperature–pressure space above which the system is totally liquid.

lithology physical character of rocks.

lithophile element element usually found in silicate minerals.

lithosphere the rigid outer skin of the Earth, which includes the crust and the uppermost mantle; its base is defined either mechanically by strength, or thermally as the level above which heat is dominantly transferred by conduction.

LLLAMA large laterally linked Archaean magma anomaly; see Section 10.5.

longitude angular distance east or west of the prime meridian (which passes through Greenwich, U.K.) of a point on the Earth's surface.

Love wave surface wave for which the particle motion is transverse and orthogonal to the direction of propagation of the wave.

low-pass filter filters the signal to remove frequencies above a given frequency, passes only the low frequencies.

Ma million years, mega-annum.

mafic rock rock rich in magnesium (*ma*) and iron (*fic*) minerals.

magma molten rock when it is within the Earth.

magmatism development and movement of magma.

magnetic anomaly difference between observed (measured) and theoretical values of the magnetic field.

magnetic poles ends of a permanent bar magnet; by convention, lines of force leave the north pole of the magnet and enter its south pole.

magnetic variation *see* declination.

magnetometer instrument that measures the Earth's magnetic field.

major elements the major components of the earth's crust: O, Si, Al, Fe, Ca, Na, K and Mg; in geochemical work, also Ti and P.

mantle solid shell of the Earth extending from the crust to the core; divided into the upper mantle (from the Moho down to 670 km depth) and the lower mantle (from 670 to 2891 km depth, 2891 km being the depth of the core–mantle boundary).

metamorphic grade recorded temperature of metamorphism; sometimes, but less correctly, used to imply pressure.

metamorphic rocks rocks that have been changed from their original state, usually by temperature or pressure.

metamorphism changes in mineralogy and texture of a rock caused by temperature and pressure.

minerals crystals that make up rocks.

Moho seismic boundary between the crust and the mantle; named after A. Mohorovičić.

molasse sedimentary rocks that are deposited in a fore-arc basin in front of a thrust-up mountain belt; often includes thick sandstones, conglomerates and shales.

moment magnitude magnitude of an earthquake as estimated by using the seismic moment; usual symbol M_O.

momentum linear momentum of a body, the product of its mass and velocity; angular momentum of a body, the product of its moment of inertia and angular velocity.

neutrino particle with no mass and no charge, but which can possess both energy and momentum; emitted in the decay of some elementary particles.

nuclear fission splitting of a nucleus into fragments, with release of energy.

obduction process that involves breaking off a piece of a subducting plate and thrusting it upwards onto the overriding plate; ophiolites were obducted into their present positions.

offset horizontal distance.

olivine mineral $(Mg,Fe)_2Si_2O_4$.

ophiolite fragment of crust and upper-mantle rocks, now exposed on land, analogous to oceanic crust and mantle.

orogen region or belt of rocks that have been deformed together at a particular time; literally, mountain creation.

orthopyroxene mineral $(Mg,Fe)_2Si_2O_6$.

outer core outer liquid shell of the Earth's core, probably iron with some oxygen; inner radius, 1221 km; outer radius, 3480 km.

Pangaea the single supercontinent of the Permian and Triassic, which broke up during the Jurassic.

partial melting melting of part of a rock; because a rock is composed of different minerals, each with its own melting behaviour, melting does not take place at one temperature (as for ice at 0 °C) but takes place over a range of temperatures; melting starts at the solidus temperature and continues, non-linearly, as the temperature increases to the liquidus temperature, at which the rock is totally molten.

passive margin continental margin formed during initial rifting apart of continents to form an ocean; frequently has thick sedimentary deposits.

peridotite rock with over 90% mafic minerals, usually dominated by olivine.

period the time taken for one complete wave to pass any point; the inverse of the frequency.

phase change change of minerals from one crystallographic structure to another more compact form at the greater temperatures and pressures within the Earth, or the reverse.

pillow lava lava with pillow-like features, erupted under water.

plate tectonics the 'carpentry' or 'architecture' of the Earth's surface; the system of large lithospheric plates that move across the Earth's surface as spherical caps; most igneous and tectonic activity occurs along the boundaries between, rather than within, the plates.

pluton body of magma that has solidified within the crust.

potential energy energy a body has by virtue of its position.

polarization of shear waves vertically polarized S-waves, termed SV, have their particle motion in the plane containing the source, the receiver and the centre of the Earth. Horizontally polarized S-waves, termed SH, have their particle motion orthogonal to this plane.

primary waves *see* body waves.

P-waves *see* body waves.

quartz crystals of silicon dioxide, SiO_2.

radioactive decay spontaneous decay of nucleus of an atom.

range horizontal distance between a shot point and receiver.

Rayleigh wave surface wave for which the particle motion is in the vertical plane that includes the direction of propagation of the wave.

REE rare-earth element; the fifteen metallic elements with atomic numbers 57–71: La, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb and Lu.

refraction bending of rays (light or sound) when they pass from one medium to another.

refractory mineral or element that remains solid during partial melting.

rift region where the crust has split apart, usually marked by a rift valley (e.g., East African Rift, Rhine graben).

right-lateral fault, dextral fault fault on which the displacement is such that, to an observer on the ground, the opposite side is displaced to the right.

scalar quantity that has a magnitude but not a direction.

secondary waves another name for S-waves. *See* body waves.

sedimentary rocks rock made up of fragments of other rocks, usually resulting from the action of wind, water or ice.

seismic discontinuity surface within the Earth at which the seismic P- and/or S-wave velocities change.

seismicity distribution of earthquakes in space and time.

seismograph, seismometer instrument that measures motion of the ground.

seismology study of earthquakes and the passage of seismic waves through the Earth.

serpentine group of minerals in the family $(Mg,Fe)_3Si_2O_5(OH)_4$, formed by the hydration of ferromagnesium-rich silicates (e.g., olivine and pyroxene).

serpentinization alteration of olivine to serpentine by hydration.

shale fine-grained sedimentary rock.

silicate large class of minerals containing silicon and oxygen as SiO_4; they make up most of the Earth's crust.

sill small igneous body that has been intruded into and parallel with the existing rock strata.

sinistral fault *see* left-lateral fault.

solidus surface in pressure–temperature space below which the system is wholly solid.

specific heat amount of heat necessary to raise the temperature of 1 kg of a substance by 1 °C; measured in $J\,kg^{-1}\,°C^{-1}$; symbol, c_P.

strain deformation of a body expressed as change in dimension (e.g., length) divided by original value of that dimension.

stratigraphy classification of strata in sedimentary rocks.

streamline line marking the position of a particle in a flowing fluid at successive times.

stress force acting on a body, expressed as force per unit area or pressure.

strike–slip fault fault on which the movement is horizontal and parallel to the strike of the fault; *see* left-lateral and right-lateral faults.

subduction process by which a lithospheric plate sinks into the mantle.

subduction zone region where one lithospheric plate descends beneath another into the mantle.

subsidence sinking of part of the surface of the Earth; possible causes include loading and thermal contraction during cooling.

surface wave seismic wave that travels in the surface layers of the Earth; Love waves and Rayleigh waves are surface waves. The amplitude of the wave decreases with depth.

surface-wave magnitude magnitude of an earthquake as estimated from measurements of the amplitude of surface waves.

suspect terrain (terrane) terrane that is suspected of having originated far from its present location.

S-waves *see* body waves.

syncline a fold, concave upwards, whose core contains stratigraphically younger rocks.

tectonics geological process that involves the movement of solid rock; the carpentry of the Earth.

terrain (terrane) geological unit bounded by faults and differing markedly in structure or stratigraphy from its neighbours; *see* accreted terrane, suspect terrane.

tesla the unit of magnetic field; one tesla (^{1}T) is equal to $1\,kg\,A^{-1}\,s^{-2}$.

thermal expansion change in volume of a sample of material divided by its original volume for an increase in temperature of 1 °C; units are $°C^{-1}$, symbol, α.

tholeiitic family name of silica-saturated basalts.

tonalite granitoid rock with quartz together with plagioclase-feldspar, the plagioclase making up more than 90% of total feldspar.

trace element element that occurs at a concentration of less than 1%.

trench long, narrow arcuate depression in the seabed, which results from the bending of the lithospheric plate as it descends into the mantle at a subduction zone.

triple junction point where three plates meet.

ultramafic rock rock ultra-rich in magnesium (*ma*) and iron (*fic*) minerals such as olivine and pyroxenes.

unconformity rock strata above an unconformity, which were deposited or extruded at a much later time, or after erosion or deformation, upon the strata beneath; marks a break or gap in the geological record.

vector quantity having both magnitude and direction.

velocity vector quantity that indicates both the speed and the direction in which a body is moving.

viscoelastic material a material that can behave as an elastic solid on a short timescale and as a viscous fluid on a long timescale.

viscosity resistance to flow within a fluid; a measure of the internal friction of the fluid, also known as dynamic viscosity; the unit of viscosity is $\mathrm{N\,s\,m^{-2}} = \mathrm{Pa\,s}$; symbol, η; *see also* kinematic viscosity.

volatiles chemical components that go into a vapour phase during igneous and metamorphic processes (e.g., water, CO_2).

volcanic rock igneous rock that solidified at the Earth's surface.

volcanism geological process that involves the eruption of molten rock.

wavefront imaginary surface or line that joins points at which the waves from a source are in phase (e.g., all at a maximum or all at a minimum).

wavelength distance between one crest of a wave and the next; usual symbol, λ.

Index

absolute age 234
absolute plate motion 32–6
abyssal plain 393, 459, 655, *see also* ocean basin
Acasta gneiss 262
accreted terrane 482, 516, 551, 655
accretion, Earth 265, 285
accretionary wedge, *see* subduction zone
accuracy 234
achondritic meteorite, *see* meteorite
acid rocks 396
actinium series 247
active margin 394, 655, *see also* continental margin
activity, *see* radioactive decay
adakite, formation 526
Adams–Williamson equation 331, 333, 334
adiabat 303–7, 334, 343, 350, 524, 655
adiabatic expansion 304
Adriatic block 541, 542
advection 269, 270, 274, 291, 358, 462, *see also* heat
Aegean
  mantle structure 338
  relative motion 542
  subduction zone 344, 461, 541, 542
aeon 655
Africa 88
  rifting 585
  seismic velocity 339
Africa–Eurasia, motion, seismicity 440
African plate 8, 31, 36, 70, 541
aftershock, *see* earthquake aftershock
afterslip, *see* earthquake afterslip
age of Earth, *see* Earth
age of Sun 234
air gun 141, 157
Airy, Sir George 202
Airy's hypothesis 203–4, 209, 210, 215, 216, 397
ak135 326
aktualism 595
Alaska subduction zone 465
Alberta Basin 559, 561, 562, 563, 566
Alboran Sea 542
Aleutian subduction zone 6, 77, 461, 465, 470, 471, 526
Aleutian Trench 8, 86, 475, 477
alkali basalt, radioactive elements 276
alkaline rock 397
alkali-olivine basalt 397
allanite, closure temperatures 243, 261
alpha particle 655
Alpine fault, New Zealand 319
Alps 88, 312, 541–50
  crustal thickness 548
  earthquakes 464
  gravity 545, 546
  seismic structure 546, 547–8, 549
  seismicity 548
  shortening 545
  uplift 550
Altai Shan 535
altered rock 655
Altyn Tagh Fault 535
Amadeus Basin 578
Amazonia 88
Amitsôq gneiss 246
Ampère's law 617
amphibolite 655
Amundsen Basin 422
Anatolian plate 541, 542
Andes 202, 310, 552
  crustal structure 479, 482, 485
  gravity 475, 479
  Sr ratio 486
  topography 479
andesite 395, 396, 397, 485, 525, 655
andesitic basalt 397
Andrew Bain fracture zone 414
angle of shear 621
angular momentum, Earth 348
angular velocity 14, 19, 648
  present-day 15
  relative 18
anisotropy 346–8
  oceanic mantle 408
Antarctica 8, 23, 71, 88
anticline 173, 655
anticrack faulting, olivine-spinel 472
apatite, closure temperatures 243, 261
apatite, fission track 318, 319
Apennines 542
Appalachian Basin 561, 562–5, 566
Appalachians 302, 516, 564
Ar–Ar dating, *see* radioactive dating
Ar ratio, non-radiogenic atmospheric 254
Ar
  loss 252, 254
  natural isotopic abundance 252
  radioactive decay 238, 365
Arabian plate 8, 31, 542
  subduction 526
Arabian Sea, magnetic anomaly 76
Archaean 308, 595–601, 655
  continental crust 280, 596–8
  diamonds 597
  geotherm 598, 600, *see also* geotherm
  heat 298, 371, 599
  initial $^{87}Sr/^{86}Sr$ ratio 245
  lithosphere 597, 598
  magma anomaly 600
  magnetic field 381
  mantle 366, 371, 600
  oceanic crust 599, 600
  plate tectonics 285, 370–1, 599
  temperature 280, 285
Archimedes' principle 202

Arctic 70, 422
arc–trench gap 459
argon, *see* Ar
Arrhenius equation 242
aseismic region 655
aseismic slip 114, 468, 469, 483
Asia, tectonics 535, 536
aspect ratio 359–60
asteroids, origin 263
asthenosphere 5, 108, 202, 297, 655
  viscosity 225, 228
Atlantic MORB 414
Atlantic Ocean
  opening 53, 70, 88, 545
  sediment thickness 400
  spreading rate 60
Atlantis Bank 398, 402, 450
Atlantis II transform fault, Southwest Indian Ridge 398, 441, 450, 452
Atlantis transform fault 420–2, 424, 441
atmosphere 47
atomic number 655
attenuation, *see* quality factor
Aurora 47
Australia 71, 76, 262
Australia, gravity 578, 579
Australia, rifting from Antarctica 74, 88
Australian–Antarctic Discordance 414
auxiliary plane, focal mechanism 134, 137, 139
axial magnetization 67
axial volcanic ridge 419
Azores Triple Junction 8, 33, 440

Babylonians 1
back-arc basin 394, 402, 459, 460, 486, 526, 655
backstripping 557
Baffin Bay 70
Baikal Rift 535, 536, 585
Baja California 82
balloon mode 110
Baltica 88, 302
Banggong Nujiang suture 531, 532
basalt 395, 396, 397, 656
  melting 518, 526
  metamorphic reactions 518
  P-wave velocity with *P*, *T* 413
  radioactive elements 242, 276
  seismic velocity 513
basaltic andesite 397
basement rock 656
basic rocks 396
basin 656
Basin and Range 299, 300, 302, 511, 559, 566, 590
Bateman relations 239
batholith 656
bathymetry 656
  Chain transform fault 443
  Clipperton transform fault 442
  East Pacific Rise 442
  Hawaii 219
  Mid-Atlantic Ridge 392, 411
  mid-ocean ridge 442, 443
  New Hebrides Trench 392
  Pacific–Antarctic Ridge 392
  Peru–Chile Trench 392
  Romanche Transform Fault 443
  Southwest Indian Ridge 392
  St Paul Transform Fault 443
Bay of Bengal, magnetic anomaly 71, 76
Bay of Biscay 542, 574
Bay of Islands ophiolite 405
b.c.c., *see* iron
Becquerel, Henri 233
Benioff, Hugo 464
beta particle 656
Betics 542
biotite, closure temperature 243, 261
Birch's law 103, 104
BIRPS 177, 178, 550
Black Sea 88, 542
black smokers, *see* hydrothermal circulation
Blake Spur fracture zone 450, 452
Blanco transform fault 441
blocking temperature 51, 242, 318, *see also* closure temperature
blueschist 523, 656
body wave 101–4, 656
  amplitude 104, 152
  magnitude, *see* earthquake
  seismic velocity 104
Boltwood, Bruce 233, 262
bottom-simulating reflector, *see* BSR
Bouguer anomaly 206, *see also* gravity
Bouguer correction 206
Bouguer, Pierre 202
boundary-layer model, *see* plate models
Boussinesq approximation 355, *see also* convection
Bouvet Triple Junction 8, 33, 71
bow shock 47
Bowie hotspot 363
Brazilide Ocean 88
breathing mode 110
British Institutions Reflection Profiling Syndicate, *see* BIRPS
brittle–ductile transition 514
Bruhnes epoch 60
Bruhnes, Bernard 60
BSR, bottom-simulating reflector 167–8, 527
bulk Earth, isotope ratio 361
bulk modulus 102–4, 331, 332, 625, 656
bulk modulus, versus depth 335
Bullard, Sir Edward (Teddy) 14, 374
b-value, *see* earthquake, *b*-value

c 126
^{14}C decay 238
Calaveras Fault 124
calcic rock 397
Caledonian orogeny 550–1
California
  motion 18
  seismicity 124
Cambrian 88
Canada Basin 422
Canadian Shield 300, 339
Canyon Diablo meteorite 265
Capricorn plate 76
carbonaceous chondrite 263
Caribbean plate 70
Carlsberg Ridge 8, 70, 74
Carpathians 88, 542
Cascadia subduction zone 31, 81, 83, 167, 461
  earthquakes 468, 469
  heat flow 482
  seismic structure 168, 483
  seismic hazard 121
  seismicity 482
  uplift 473
Caspian Sea 88, 542
cation 656
Caucasus 542

CDP reflection method 158, 159, 160, 162, *see also* stacking
Cenozoic
Central America subduction zone 465
Central Indian Ridge 8, 70, 71, 74, 75, 418
centrifugal acceleration 197
centripetal acceleration 656
Chagos–Maldive–Laccadive Ridge 71, 74, 75
Chain Transform Fault 443, 458
Challenger Deep 391
Challenger, HMS 391
characteristic distance 281
characteristic time 281
Charlie Gibbs transform fault 452, *see* transform fault
chemical remanent magnetization 51
chemosynthesis 430
Chile Trench
  bathymetry 479
  crustal structure 482
Chile, composition magma 522
China 88
chondritic meteorite 254, *see* meteorite
CHUR (chondritic uniform reservoir) 256–7, 264, 656
Clapeyron curve, olivine 349, 464
classical ray theory 156
clinopyroxene 349
Clipperton transform fault
  bathymetry 442
  crust 441, 453
closed system, *see* radioactive dating
closure temperature 242, 243, 244, 252, 261, 262
CMB, *see* core–mantle boundary
COCORP 176, 177, 510, 553, 567, 592, 593
Cocos Nazca Ridge, *see* Galapagos
Cocos plate 5, 8, 31, 79, 81, 83
coefficient of thermal expansion 305, 307, 379
colatitude 50
Colorado Plateau 566, 589
common-depth-point stacking 158, *see* stacking
common-offset stacking 158, *see* stacking
compensation depth 207, 210, 656
compensation 203, 205, 210, 289, 453, *see also* isostasy
  ocean and continent 204, 209, 397
complementary error function 282
composition
  core 350–3
  crust 348, 513–17
  Earth 348–53
  mantle 348–50
compressional basins 577–81
compressional fault 131
compressional wave 656
  equation 102
concordant dates, *see* radioactive dating
concordia, *see* radioactive dating
conduction 5–10, 269, 271, 656, *see also* heat
conductive heat flow, timescale 281
conductivity, *see* thermal conductivity
Congo 88
Conrad discontinuity 511
conservative plate boundary 6
Consortium for Continental Reflection Profiling, *see* COCORP
constructive plate boundary, *see* divergent boundary
contact metamorphism 310
continent
  ages 262, 263, 515, 516, 595
  seismic structure 109, 538, 550, 593, 594
continental crust
  Archaean 597, 598
  area 516, 649
  classification 512
  composition 513–17
  density 10
  geotherm 279, 280, 303
  heat generation 300–2, 313, 525
  melting subduction zone 525
  oldest 262, 263
  radioactive elements 276
  seismic-velocity structure 326, 511–13
  thickness 511
continental drift 3, 10, 14
continental drillhole 510
continental flood basalt 569–70, *see* flood basalt
continental geotherm 303
continental growth rate 556
continental heat flow 286, 298–303
  province 299, 300
continental lithosphere, strength envelope 514–15
continental margin
  bathymetry 394
  non-volcanic 568, 570
  passive 393
  sediment 398, 560
  seismic velocity 576
  subsidence 575
  transform fault 442
  volcanic 568, 569, 576
continental reconstructions 88
continental rifting
  Africa and Antarctica 71
  Africa and India 74
  Africa and North America 70
  Australia and Antarctica 74
  early geometry 443–4
  Eurasia and North America 70
  Greenland and Eurasia 70
  Greenland and North America 70
  India and Australia 71
  Madagascar and Antarctica 71
continental rifts 70, 124, 570, 584–95
continental rise 393, 394, 656
continental shelf 286, 393, 394, 656
continental shield, roots 339, 347
continental slope 393, 394, 656
continent–continent collision 75, 362, 530–55
convection, *see also* mantle *and* heat
  bimodal 354, 355
  Boussinesq approximation 355
  cell dimensions 357, 359–60, 365, 366
  coupling core and mantle 379
  critical Rayleigh number 356–7, 358
  dimensionless numbers 356, 357, 358
  equations governing 355
  flow lines 360, 362, 367
  hexagonal 354, 355
  horizontal roll 354
  laminar flow 357
  mantle 3, 288, 297, 358–67
  models
    computer 360, 362, 364
    laboratory 361
    three-dimensional 365–6

ocean-depth variation 218
planform 354, 364, 367
Rayleigh number 356, 357, 358, 366
Rayleigh–Bénard 353–5
separate upper and lower mantle 359, 363–5
spoke 355
temperature 304, 360, 362, 364, 367
turbulent flow 357
convergent boundary 6, 15, *see* subduction zone
cooling history and closure temperature 242–3
cordillera 516, 657, *see also* Cascadia
core 1, 2, 657
attenuation 335
carbon 351
coefficient of thermal expansion 379
composition 350–3, 371, 372
convection 375, 379
density 335, 352, 648
differentiation 263, 265
Ekman number 378
elastic moduli 335
electrical current 373, 379
flow at CMB 380
formation 381
gravity 335
heat 286, 304, 371
inner
anisotropy 347
discovery of 2, 330
physical properties 378, 379
rotation of 347
structure 330
magnetic field at CMB 380
mass 648
melting temperature 372
outer 2, 102, 601, 662
composition 351
convection, driving force 380–1
density, mass, volume 336, 379
flow 378, 379
radius 2, 648
Rayleigh number 381
physical properties 378, 379
pressure 335, 352
Q 335
Roberts number 378
Rossby number 378
seismic parameter 352
seismic velocity 330–1
seismic waves 126
temperature 306, 307–8, 330, 373
volume 648
core–mantle boundary (CMB) 338, 353
boundary layer 306, 329, 346
temperature 306, 307, 329
Coriolis force 376, 378
corrugated slip surface, Mid-Atlantic Ridge 424
co-seismic deformation 472, 483
cosine formula 20
cosine formula, spherical triangle 52
cosmic rays 258, 657
Costa Rica Rift, oceanic crust 401
coverage, *see* seismic reflection
craton 262, 298, 515, 551, 596, 657
Cree Lake fold belt 552
Cretaceous 63, 88
Crete 541
critical angle 142, 151
critical distance 143, 146, 151, 157
CRM, *see* chemical remanent magnetization
cross product 616
crossover distance 144
CRUST 5.1 512
crust 1, 2, *see also* continental, oceanic
composition 348
continental growth curve 246, 556
heat 275, 276, 286
mass 648
radioactive elements 275, 276
seismic structure 326, 512
thickness, world map 512
volume 648
cubical dilatation 623
cumulate 657
Curie temperature or point 51, 374, 539, 657
curl 617–18
curvilinear coordinates 618–19
cylindrical polar coordinates 618–19
Cyprus 542

D′ 329
D″ 330, 346
anisotropy 347
dacite 395
Darwin, Charles 2, 3
dating
absolute age 234
relative age 234
daughter atom 236
Dead Sea fault 31, 542
decay constant 234
Deccan Traps 33, 74, 75, 532
declination 50, 52, 657
London 13
decollement zone 657
decompaction 558
decompression melting 75, 414, 568
deconvolution 157, 165
Deep Sea Drilling Project 59
deflection, elastic plate 220, 222
deformation 5, 101, 222
deformation front 528
dehydration, oceanic lithosphere on subduction 517–19, 522–4
dehydration reaction 517, 520, 521, 657
density 102–4, 331, 332, 642, 657
and temperature 292
and seismic velocity 103
self-compression model 304
whole Earth 331, 335
with depth 333, 334, 335
depleted mantle 257, 657, *see also* mantle
deposition, affect on geotherm 313–18
depositional remanent magnetization 51
destructive plate boundary, *see* convergent boundary
detachment fault 445
detachment surface 420–2
detrital remanent magnetization 51
Deval, *see* mid-ocean ridge
Devonian 88
dextral fault 12, 25, 443, 657
diamond, Archaean 597, 598, 600
diamond-anvil experiments 351, 371
diapir 657
differentiation 657
mantle and core 263, 265
diffraction 169–72, 173
diffuse plate boundary 70, 76
diffusion coefficient 242, 244
diffusion equation 273, 282
diffusivity 642
dilatation 657
dilational wave, *see* P-wave

dimensionless numbers, convection 356, 357, 358
Dinarides 542
diorite 395, 396
dip pole 45
dipole 45, 47, 48, 373
  magnetic field 49, 373
dipping layer
  seismic reflection 172
  seismic refraction 146
dip–slip fault 131, 657
direct wave 142, 144, 145
discordant dates, *see* radioactive dating
discordia 248, 249, *see* radioactive dating
dispersion curve 107, 108, 110
dispersion, *see* surface wave
displacement potential 628
diurnal variation of magnetic field 47
divergence 617
divergent boundary 6, 15, 391, *see also* mid-ocean ridge
diving wave 153
Dora Maira massif, Alps 554
dot product 616
double seismic zone, *see* subduction zone
down-dip shooting 146
drilling, crust 399, 510
driving force, plate tectonics 368, 369
DRM, *see* depositional remanent magnetization
DSDP (Deep Sea Drilling Project) 398, 401
Du Toit, Alex 3, 4
dunite 349, 395, 408, 513, 657
dyke 657
  cooling 283
dynamo
  Parker–Levy 377
  reversal of 376, 377, 378
  see also geodynamo *and* geomagnetic dynamo 374–5
  self-exciting 374–5, 377

Earth
  accretion 265, 285
  age 233–4, 262–5, 430, 648
  angular momentum 348
  bulk modulus 335
  cooling 288
  density 335, 336, 648
  early ^{244}Pu 259, 275
  early ^{26}Al 275, 285
  early differentiation 263, 265
  geoid height-anomaly map 199
  global heat loss 429
  gravity 200, 335
  heat engine 5
  heat loss 285–8, 371
  mass 333, 336, 648
  origin of magnetic field 373–9
  pressure 335
  Q 335
  radius 648
  rigidity 335
  seismic structure 327, 338
  shape 198–202
  surface area 6, 649
  temperature 269, 286–8, 306
  topographic map 200
  velocity around Sun 648
  volume 336, 648
earthquake 657
  aftershock 113–15
  afterslip 114
  Alaska 120, 470, 471
  Aleutian 1957 130
  Alps 464
  Armenia 1988 120
  aseismic slip 469
  body-wave magnitude 117, 118, 122, 125, 656
  Bolivia 108, 475
  Buller New Zealand 2, 330
  b-value 123–4, 475
  California
    Landers 1992 131–2
    Loma Prieta 1989 120
    San Fernando 1971 120
    San Francisco 1906 120
    San Francisco rate 123
  Cascadia 26 January 1700 469
  Chile 108, 120, 127, 130
  China
    Haicheng 1975 122
    Tangshan 1976 120, 122
  Concepción 1835 3
  continental rift 124
  co-seismic deformation 472
  damage 115, 120
  deep 113, 117, 124, 465, 471–5
  depth 113, 114, 458, 464–5
  descriptors 115
  East African Rift 135, 136, 587
  East Pacific Rise 435
  elastic-rebound theory 130
  energy 122, 124–6, 269, 464
  epicentre 111–13, 657
  fault length and area 113, 119, 130
  fault-plane solution 16, 17, 130–9, 140, 180
    Africa 587
    Anatolia 542
    Himalayas 535
    mid-ocean ridge 443, 458
    subduction zone 470–1
  first motion 122, 132, 133, 180
  Flores Sea 1996 331
  focal depth 113, 114, 116
  focus 111, 112, 133, 657
  foreshock 113–15
  Gakkel Ridge 421, 422
  great 469, 470–1
  Hawaii 487, 489
  hazard 120, 121
  Himalaya 464, 534, 535
  hydrofracture 521
  India Gujarat 2001 120
  intensity 115
  interplate 130
  inter-seismic deformation 472, 473, 483
  intraplate 130, 131, 223, 297
  Iran 464
  Japan 115, 120, 521
  Jeffreys–Bullen travel-time curves 127, 327
  Landers interferogram 132
  Macquarie Ridge 135, 136
  magnitude 114, 115–17, 126, 410
  magnitude–frequency relationship 123–4
  magnitude, local scale 115, 118, 122
  magnitude maximum expected 123
  magnitude with depth 114
  Mercall intensity scale 115
  Mexico 120, 470
  Mid-Atlantic Ridge 181, 456, 458
  mid-ocean ridge 124, 410, 420–2
  Missouri, New Madrid 1811–1812 130
  moment magnitude 118
  Nicaragua 127

prediction 121–2
records 127
recurrence time 123, 468
Richter magnitude scale 115, 122
Rio Grande Rift 591–2
seismic moment 117–18, 119, 122, 123
seismology 111–40
Škopje 1963 120
slip vector 135, 139
Spain 1954 475
stress drop 118, 119
strike–slip 133, 134
subduction zone 464–75, 476
surface-wave generation 104, 122
surface-wave magnitude 116, 117–18, 119, 120, 122, 123
swarm 113–15, 124, 420, 458
S-wave generation 122
Tonga 475, 476
travel-time curves 127, 128, 130
Turkey 120, 464, 545
volcanic 124
East African Rift 511, 585–8, 589, 590
earthquake 136, 587
gravity 586
seismic structure 586, 589
East Anatolian fault 31, 545
East Pacific Rise 8, 58, 83, 391, 417, 419
bathymetry 442
crustal melt (OSC) 448
deep seismic structure 339
emplacement zone 57
hydrothermal vents 429
magma chamber 427, 428, 434–6, 437, 439, 440
seismic-reflection profile 437
seismic velocity 432, 438
transform faults 441
Easter Island Triple Junction 1, 8, 28
eclogite 462, 520, 657
Egyptians 1
Ekman number 378
elastic deformation 222
elastic limit 657
elastic lithosphere 572, 579–80
elastic moduli, whole Earth 331, 335
elastic plate, flexure 220, 222, 562
elastic-rebound theory 130
elastic thickness, lithosphere 218, 219, 220, 223, 224, 297, 541
elastic wave, equation of motion 620, 625–7
electron capture 658
Elizabeth I, Queen 45, 381
ellipticity 197
Elsasser, W. M. 374
Eltannin transform-fault system 441
Emperor–Hawaiian seamount chain 1, 32, 33, 35, 86, 394
endothermic reaction 272, 349, 658
energy 641
entropy 304
Eocene 88
epicentral distance 112, 116, 658, *see also* earthquake
epicentre 111, 658, *see also* earthquake
equal-area projection/net 137, 180
equipotential 198
equivoluminal wave, *see* S-wave
erosion 658
erosion, affect on geotherm 313–18
error function 282, 638–9
error, two standard deviations 637
Euler, fixed-point theorem 14
Eurasian plate 8, 36, 70
Everest, Mount 391
Everest, Sir George 202
exact ray theory, *see* generalized ray theory
Exmouth Plateau 71, 86
exothermic reaction 272, 349, 658
Explorer Ridge 55
explosive cord 158–62
extensional basins 567–77
extrusion 658

facies 658
FAMOUS, *see* Mid-Atlantic Ridge
fan shooting 140
Farallon plate 24, 25, 77, 79, 81, 82, 86, 87
remnant in mantle 339, 344, 480, 482
fault 658
Anatolian E, N, S 31, 542
Dead Sea 542
displacement 117
normal 131
right lateral 17, *see also* dextral
slip vector 131
strike–slip 17, 131, 132, 664
throw resolvable 175
thrust 131
fault plane 130, 131, 137, 139, *see also* earthquake
determination 139
dip 131, 134
strike 134
fault-plane solution, *see* earthquake
fayalite 348
f.c.c., *see* iron
Fe, abundance in cosmic rays 258
feldspar 658
felsic rocks 395, 396
Fennoscandia, uplift 226, 227
Fennoscandian ice sheet 225
fertile mantle 414
Fiji Islands 469
fission branching ratio 260
fission
induced 260
spontaneous 258
thermonuclear 234
^{238}U 258, 259, 260
fission track
apatite 261, 318, 319
closure temperatures 243, 261
sphene 261, 262
temperature dependence 260–2
zircon 261, 318, 319, 320
fission-track dating 258, 259, 260–2
FitzRoy, Capt. 3
Flannan Thrust 178, 550, 551
flat subduction 526
flexural parameter 220
flexural rigidity 219, 562
flexure
Hawaii 217, 219, 221
lithosphere 218–24, 562
Mariana Trench 222
Pacific plate 219, 222
Tonga trench 222
flood basalt 75, 569
focal depth 113, *see also* earthquake focal depth
focal mechanism, *see* earthquake fault-plane solution
focal sphere 133
focus 111, *see also* earthquake focus
football mode 110
fore-arc basin 459, 526, 658

foreland basin 563, 658
fosterite 348
fractionation 658
fracture zone 56, 77, 441, 445, 446, *see also* transform fault
  Andrew Bain 414
  Atlantis II 450
  Blake Spur 450
  Kane 432
  Kurchatov 441
free oscillation modes 108–11, 327, 658
free-air anomaly 206, *see also* gravity
free-air correction 205, *see also* gravity
frequency 629, 642, 658
Fresnel zone 171, 174, 175–6
frictional heat, *see* shear-stress heat
frontal fold 459, 527, 528

G, *see* gravitational constant
Ga 658
gabbro 395, 396, 513, 658
gabbro-eclogite, transformation rate 565–7
Gakkel Ridge 8, 70, 407, 417, 422
  crust 407, 416
  seismicity 421
gal 197
Galapagos spreading centre 83, 410
Galapagos Triple Junction 8, 28, 31, 33
galena 263
Galileo 1, 196
gamma decay 658
Ganga basin, gravity 534, 540, 541
Garrett transform fault 434
gas constant 648
gas exploder 157
gas hydrate 167, 168, 527
  Makran
Gauss epoch 60
Gauss, Carl Friedrich 60, 373
GDH1, *see* plate models
generalized ray theory 156
geochronology, *see* radioactive dating
geographic pole 13
geoid height anomaly 214–17, 218, 477
  global map 199
geoid 198–201, 214, 217
geological time scale 234
geomagnetic, *see also* magnetic
geomagnetic dynamo 374, 380–1
  Parker–Levy 376
geomagnetic equator 13, 45, 47
geomagnetic field 43
geomagnetic N and S poles 13, 45, 47
geomagnetic reversals 63, 234
geometrical ray theory 142
geophone 157
geotherm 275, 279, 280, 303, 658
  Archaean 280, 597, 598, 599, 600
  continental 279, 303
  equilibrium 275–9, 280
  erosion 313–18
  oceanic 303
  one-layer crust 278, 279
  two-layer crust 279–80
Gilbert epoch 60
Gilbert, William 45, 60, 373, 381
glacial rebound, uplift 224, 226, 227, *see also* isostatic rebound
glaciation, loading 224
Global Seismic Hazard map 121
Global Seismograph Network 131
Glomar Challenger 399
gneiss 658
Godthaab, Greenland 246
Gofar transform fault 441
Gondwanaland 77, 88, 545, 658
Gorda Ridge 55
GPS 17, 114
gradient (grad) 616–17
granite 141, 276, 310–11, 395, 396, 513, 525, 658
  ice-rafted 398
granitoid 242, 658
granodiorite 396, 514, 525, 658
granulite 514, 525, 659
gravimeter 202, 659
gravitational acceleration 193–6
  spherical shell 196
gravitational constant 193, 648
gravitational potential 193–6, 198
gravity 196, 335
gravity unit 202
gravity anomaly 202, 659
  Aleutian Trench 477
  Alps 546
  Australia 579
  Bouguer 206, 208, 209–10
  buried sphere 212
  calculation 205–6
  cylinder 213
  equatorial 648
  fault 131
  free-air 206, 208
  global map 200
  Hawaii 208
  Himalaya 540
  horizontal sheet 213
  isostasy 207–10
  mantle Bouguer 206, 419
  mantle convection 216–17, 218
  Mid-Atlantic Ridge 207, 213, 411
  mountains 209–10
  polar 648
  reference gravity formula 197, 205
  Rio Grande Rift 590
  sphere 212, 213
  spherical shell 332
  Tibet 540
  U.K. continental margin 207
great circle 15, 16, 138, 659
Great Dyke, Zimbabwe 315
great earthquake, *see* earthquake descriptor
Great Magnetic Bight 77, 79, 86, *see also* magnetic bight
Great Plains 589, 590
Green River Basin 510
Greenland 70, 88, 246, 262, 422
greenschist 659
greenschist–amphibolite, reaction 520
greenstone belt 308, 659
Grenville province 516
ground roll 104
group velocity 106, 107, 109
Guadalupe Island, crust 402, 403
Guadalupe plate 81, *see* Cocos plate
Gulf of Aden 70, 585
Gulf of Bothnia 225, 226
Gulf of Corinth, extension 541
Gutenberg discontinuity 330, 659
Gutenberg, Beno 2, 4
Gutenberg–Richter law 123
Guymas Basin 581

Hadean
haematite 51
half-life 236, 659, *see also* radioactive decay

half-space cooling model, *see* plate models
Harve Trough 469
harzburgite 349, 395, 659
Hatton Bank 207
Hawaii 1, 32, 33, 87, 394, 397
  density 488
  flexure 217, 221
  gravity anomaly 217
  magma 487, 490
  seismic structure 487–9
  seismicity 487–9
  topography 217, 391
Hawaiian lineations 86
Hawaiian Ridge 1, 86
Hayward Fault 124
h.c.p., *see* iron
head wave 141, 142, 143, 144, 145, 146, 152, 153
heat 641
  convection 269, *see also* convection
  from Sun 269
  radiation 269, 270
  rate of transfer 269
heat conduction 270–4
  crust 308–18
  one-dimensional equation 273
  three-dimensional equation 273, 291
heat flow 286, 287, 641
  age 289, 290, 291, 299
  Archaean 298, 299
  continental 286, 287, 298, 299, 591
  mantle contribution 298–302
  oceanic 285, 286, 287, 288–90, 296
  Precambrian 286, 299
  Proterozoic 298, 299
heat-flow province 299, 300
heat-flow unit 271, 641
heat generation 299, 358, 525, 599, 641
  crust 279, 299, 300–2
  radioactive 233, 237, 274–5, 276, 288, 659
  with time 277, 599
heat-generation unit 271, 641
heat loss
  Archaean 599
  continents 286
  Earth 269, 271, 285–8
heat of fusion 660
heat production, *see* heat generation
heat-producing element 659
Hellenic arc 541
Herglotz–Wiechert inversion 327, 634
Hess, Harry H. 3, 4, 57
hidden layer, seismic refraction 145
High Himalaya 531, 533, 534
Himalaya 6, 88, 202
  collision 461, 525, 530–41
  convergence rate 534–7
  cross section 531, 537
  crustal thickness 530, 537
  earthquake 464, 535
  gravity 539–41
  shortening 533, 534
  tectonic evolution 530, 535
  uplift 550
Hiroshima, energy release 125
Holmes, A. 4, 263
Holmes–Houtermans equation, model 263
hominoid remains 252
Hooke's law 623
horizontal stress, in plates 36
hornblende, closure temperature 243
hotspot 32, 33, 88, 217, 339, 394, 487, 659
  global plate velocity 35
  reference frame 35
  track 1, 35, 36, 71, 75, 76
HS thermal model, *see* plate models
Hudson Bay 225, 227
Hutton 2
Huygens' principle 169–72
hydration 659
hydrocarbon formation 581–3
hydrofracture 521
hydrogen, in core 351
hydrophone 140, 141
hydrostatic pressure 332
hydrothermal circulation 288–9, 309, 453, 581, 659
  heat loss 286
  life 430–1
  oceanic crust 428–31
  salinity 233
  subduction zone 431
  vents 288, 429
  volume water 429
hyperthermophile bacteria 430
hypocentre, *see* earthquake focus
I 126
i 126
Iapetus Ocean 88, 551
IASP91 Earth model 650
iasp91 326, 327, 330
Iberian margin 570–1
IC (inner corner), *see* transform fault
ice, seismic velocity 513
ice age 285
ice caps 224
Iceland 33, 70, 288, 569
  crust thickness 415, 417
Iceland–Faroe Ridge 394
ice-rafted granite 398
igneous rock
  classification 395–7, 659
  seismic velocity 104
IGRF (International Geomagnetic Reference Field) 43, 44, 54
IGS (International GPS) 17
Illinois basin 559, 560, 561, 564, 565, 566
ilmenite–perovskite 473
impedance 152, 167, 168
inclination 50, 52, 659
incompatible element 414, 659
incompressibility 625, *see* bulk modulus
India 88, 299, 302
India–Eurasia collision 531, 532, 533, 535, 536, 538
Indian Ocean, geological history 60, 70–7, 88
Indian Ocean triple junction 8, 70, 71, 74, 418
Indian plate 8, 36, 76
  effective elastic thickness 541
  subduction 530–41
Indonesia, Sr ratio 486
induced fission, *see* fission
induced magnetization 51
Indus–Tsangpo suture 530, 531, 532, 535, 541
infinitesimal strain, theory 621
inhomogeneity 346
initial ratio 245, *see also* radioactive dating
  Archaean 245
  meteorite 245
  MORB 245
  Nd CHUR 256

inner core 2, 659, *see also* core
density 336, 379
iron phases 351–3
mass 336, 379
radius 648
seismic waves 126
volume 336, 379
instantaneous velocity 659
Insubric Line 544, 545, 548
interference head wave 153
intermediate rocks 395, 396
International Association of Geodesy 197
International Geomagnetic Reference Field, *see* IGRF
interplate earthquake 130, 659
interseismic deformation 472, 483
interval velocity 165, 166, 167
inter-volcano gap, mid-ocean ridge 447
intra-continental sedimentary basin 659
intraplate earthquake 130, 131, 223, 297, 659
intraplate volcanoes 32
intra-volcano gap, mid-ocean ridge 447
intrusion 659
intrusive rock 660
inversion, body-wave travel times 108, 326, 633–5
ion 660
Iran, subduction zone 88, 461, 464
iron meteorite, *see* meteorite
iron, b.c.c., f.c.c. and h.c.p. in core 347, 351, 353
in core 353, 371
melting 353, 372
irrotional wave, *see* P-wave 618
ISC (International Seismological Centre) 123
island arc 6, 363, 459, 486, 522, 660, *see also* subduction zone *or* volcanic arc
island chain 32, *see also* hotspot track
isochron, *see* radioactive dating
isostasy 202–5, 292, 660
Airy hypothesis 204
determination of whether compensated 204, 207–10, 215–16
Pratt hypothesis 204
isostatic anomaly 208, 209, 210
isostatic equilibrium 203, 207, 208, 210, 214
ocean and continent 397
isostatic rebound 201, 218, 224–8, 270
isotherm 660
isotope 660
abundance with time 277
isotope ratio
bulk Earth 361
oceanic basalt 361
oceanic-island basalt 361
isotopic abundance
Ar 252
K 251
Nd 254
Sr 244
U 247
Isua, Greenland 262
Ivrea body 545
Izanagi plate 86, 87
Izu–Bonin subduction zone 461, 465, 474, 477

J 126
Jack Hills 262
Jaili Fault 535
Jalisco, Mexico 468
James I, King 45
Japan 6, 344
crust 478, 485
earthquakes 467, 478, 521
gravity 475, 478
heat flow 478
seismic velocity 338, 465, 467, 477, 480
subduction zone 459, 464–5, 466
uplift 473
Japanese lineations 86
Jaramillo polarity event 60
Java Trench 8, 71, 76
Jeffreys, Sir Harold 1, 2, 4, 326
Jeffreys–Bullen travel-time curves, model 112, 127, 327, 330
Job 1
JOIDES (Joint Oceanographic Institutions for Deep Earth Sampling) 398
Juan de Fuca plate 8, 31, 79, 81, 83, 167, 482, 483
Juan de Fuca Ridge 55, 58, 81
magnetic anomaly 59, 77
seismic velocity 432
spreading rate 393
transform faults 441
Jurassic 63, 88

K 126
^{40}K, radioactive decays 238, 240–1
Kaena polarity event 60
Kalahari 88
Kamchatka subduction zone 32, 465
Kanakee Arch 564, 566
Kane transform fault 420–2, 432, 441, 450, 452, 456
Kangmar dome, Tibet 538
Kapuskasing zone 597
K–Ar closure temperatures 243
K–Ar dating, *see* radioactive dating
Karakorum Fault 535
Karoo 567
Kazakhstan 88
kbar 660
keel, *see* continental shield, roots
Kelvin, Lord 233, 324, 325
Kepler, Johann 198
Kerguelen 33, 71
Kermadec subduction zone 8, 465, 469, *see also* Tonga
kerogen 581, 582, 583, 660
Keweenawan Rift 585, 594–5
Kilauea 487, 490
kimberlite 660
kinematic viscosity 660
kinetic energy 660
Knipovitch Ridge 70
Kokchetav complex 554
Kola drillhole 8, 510
Kolbeinsey Ridge 8, 70
komatiite 395, 396, 596, 600, 660
Königsberger ratio 52
KTB drillhole 510
Kula plate 24, 79, 86, 87
Kula triple junction 79
Kunlun block 530, 532
Kunlun fault 535
Kurchatov transform fault 441
Kurile subduction zone 8, 465

Lake Bonneville 224
Lamé elastic constants 624, 625
laminar flow, *see* convection
Lamont seamounts 442
land bridges 3
land, mean height 649
Laptev Sea 416, 422
Laplacian operator 618
large igneous province (LIP) 569–70
laser-probe dating 252
latent heat 272, 642, 660
latitude, symbol and sign convention 19, 660
Lau back-arc basin 469
  crust 402
  magma chamber 433, 486
  seismic structure 480
Laurasia 88, 545, 660
Laurentia 88
lava 660
layering, oceanic crust 401, 402
lead, *see* Pb
leaky transform fault 440, *see also* transform fault
least-squares method 636–7
left-lateral fault 660
Lehmann discontinuity 330
Lehmann, Inge 2, 4, 330
Leonardo 2
Lesser Himalaya 531, 533, 534
Lhasa block or terrane 530, 531, 532
lherzolite 395, 660
LIL, large-ion lithophile 522, 660
limestone, seismic velocity 513
linearized inversion, *see* inversion
liquidus 660
lithology 660
lithophile element 660
Lithoprobe 551, 553
lithosphere 2, 5, 202, 661
  extension 571–4, 577
  flexure 218–24, 557–8, 561–5, 577
  rigidity 562
  temperature 292, 293, 294
  thermal time constant 560
  thickness 292, 294, 297, 409, 457, 598
  yield strength 515, 600
LLLAMA 600, 661
loading by sediments 557–8
Loihi 32, 487
Lomonosov Basin 422
longitude 19, 661
Louisville Ridge 469
Love, A. E. H. 104
Love wave 104, 105, 110, 112, 126, 127, 661, *see also* surface wave
  dispersion 106, 107, 109
  particle motion 105
  travel time 127, 128
lower focal hemisphere 133, 137
lower mantle, *see* mantle
low-velocity layer 145, 146
LQ, *see* Love wave
LR, *see* Rayleigh wave
Lu–Hf dating 238, 258

Ma 661
Macquarie Ridge 8, 136
Madagascar 71, 74, 88
mafic rock 395, 396, 661
magma 661
magma chamber
  mid-ocean ridge 423–6, 427–8, 432, 433, 434–6, 440
magma chemistry, subduction zone 485–6
magma ocean 515, 600, 661
magmatic segments, mid-ocean ridge 439
magmatism 281, 661
magnesiowüstite 350
magnetic anomaly 54, 661
  Pacific 55, 56, 59, 80
  marine 55, 57
  Mesozoic 60
  oldest Atlantic 70
  oldest Indian Ocean 71
  Philippine Basin 459
  South Fiji Basin 459
  Tibet 539
magnetic-anomaly map, first published 1961 55
magnetic bight 86, *see also* Great Magnetic Bight
magnetic dipole 45
magnetic equator 13, 45
magnetic field 44, *see also* geomagnetic field
  at CMB 380
  dipole models 373
  diurnal variation 47
  Earth 43
  equations governing 375–6
  origin 348
  reversal 43, 57, 374, 375, 377, 378
  self-exciting dynamo 374, 377
  westward drift 374
magnetic induction 642
magnetic latitude 50, 52
magnetic permeability of free space 48
magnetic pole 43, 45, 53, 661
Magnetic Quiet Zone 60, 63, 74, 77, 86
magnetic storm 47
magnetic susceptibility 51
magnetic variation 50, 661
magnetic, *see also* geomagnetic
magnetite 51
magnetization, basalt 67
magnetization of rocks 51–2
magnetohydrodynamics 375
magnetometer 54, 67, 661
magnetopause 47
magnetosheath 47
magnetosphere 47
magnetotail 47
magnitude, *see* earthquake
Main Boundary Thrust 531, 532, 534, 535
Main Central Thrust 531, 532, 533, 534, 535
Main Ethiopian Rift 585
Main Frontal Thrust 535
Main Himalayan Thrust 531, 533–7
major element 661
Makran subduction zone 459, 526–8, 542
Maksyutov complex, Urals 554
Mammoth polarity event 60
mantle 1, 661
  adiabatic temperature gradient 305, 306, 307, 343, 350
  Archaean 596, 600
  composition 348–50
  convection 216–17, 218, 233, 288, 353–67, 370
    three-dimensional models 365–6
    geoid height anomaly 216–17
    gravity anomaly 216–17
    planform 217
  convection and core 379
  convection and plate tectonics 370
  density 10, 335, 336, 352, 648
  depleted 257

mantle (*cont.*)
differentiation 263, 265
eclogite 177
elastic moduli 335
fertile 414
growth curve 246, *see also* radioactive dating
heat 275, 276, 279, 304, 379
K abundance 275
low-velocity zone 327
mass 336, 648
melting 32, 303–7, 372
at subduction zone 519, 521, 522–4
Nusselt number 357
olivine 335, 348–9, 350, 366
Péclet number 358
peridotite 177
phase changes 334–5, 349–50, 363, 364, 366, 367
Prandtl number 358
pressure 335, 352
P-wave-velocity maps 338
Q 335, 344
radioactive elements 276
Rayleigh number 358, 359
Reynolds number 357
rigidity 108
seismic parameter 352
seismic velocity 338, 343, 408–9, 480
seismic waves 126
spinel–post-spinel 335, 349, 363, 366
temperature 303–7, 330, 334, 343, 350, 416
Th abundance 275
transition zone 2, 328, 334, 349–50
U abundance 275
ULVZ (ultra-low-velocity zone) 346, 347
undepleted 275, 276
upper and lower boundaries 306, 328, 344
viscosity 5, 214, 218, 219, 224–8, 270, 358, 364, 365
volume 336, 648
mantle wedge 524, *see also* subduction zone
marginal basin 394, 459, *see also* back-arc basin
Mariana arc, eruption time 523
Mariana subduction zone 459, 465, 477, 485, 523
Mariana Trench 8, 222, 391
marine magnetic anomaly, calculation of 61–7
marine sediments 285, 289, 398, 399, 400, 401
MARK, Mid-Atlantic Ridge 421
Mascarene Plateau 75
Maslov integral method 156
mass spectrometer 237
Matthews, D. H. 56, 58
Matuyama epoch 60
Matuyama Motonori 60
Mauna Kea 391
Mauna Loa 487
maximum compressive/tensional stress 515
Maxwell's thermodynamic relation 305
m_b versus M_S 118
McCullagh's formula 205
mechanical boundary layer 298, 304
median valley, mid-ocean ridge 417–20
Mediterranean Sea 70, 88
Mediterranean tectonics 517, 541–5
melt fractionation, subduction zone 525
melting temperature, iron 353
melting
attenuation 413
basalt 427
mantle 307, 350
mid-ocean ridges 414–17
seismic velocity 413
Mendocino transform fault 31, 441
Mendocino triple junction 8, 28, 31, 81, 82, 83
Mercalli intensity scale, *see* earthquake
metamorphic belts, paired 316
metamorphic dates 318–21
metamorphic facies series 315, 316
metamorphic geotherm 313–18
metamorphic grade 661
metamorphic rock 104, 661
metamorphic terrain 319–21
metamorphism 661
burial 308–9, 314–15
high grade 312, 314, 316
intrusion 309–11, 315–16
progression/retrogression 312
meteorite 255, 263
achondrite 263–5
age 263
Canyon Diablo 265
chondrite 254, 263–5
composition 351
dating 254, 264
initial Sr ratio 245
iron 263–5
origin 263
^{244}Pu 259
stony 263, 264
Mexico, earthquake 120, 470
Michigan Basin 560, 561, 564, 565, 566
microbial life, oceanic crust 430
microcline, closure temperature 243
microplate 81
Mid-Atlantic Ridge 8, 70, 391
Atlantis Transform Fault 424
bathymetry 392, 411
deformation zone 420
density 410, 411
detachment surface 420–2, 424
earthquake 181, 410, 443, 458
emplacement zone 57, 420
FAMOUS 421, 441, 442, 455
gravity 207, 411
hydrothermal vents 429
Kane transform fault 452
magma chamber 426, 428
median valley 417, 424
microearthquakes 432
Oceanographer Transform Fault 452
segmentation 455
seismic-velocity structure 154, 431, 432
spreading rate 60, 70, 393
transform fault 441, 451, 452
Middle America subduction zone 83, 461
mid-ocean ridge 3, 6, 10, 57, 363, 435
Arctic 70
attenuation seismic waves 410
axial topography 417–23
axial volcanic ridge (AVR) 419
bathymetry 391–3, 411
crustal magma chamber 423–6, 427–8
deep structure 339, 409–17
density 410, 411
detachment surface 420–2, 423, 424, 445

Deval (deviation in axial linearity) 446, 447
earthquake 114, 124, 181, 410, 422, 432, 443, 458
emplacement zone 57, 420, 423
eruption volume 420
fast 439, 454
axial structure 418, 419, 434
bathymetry 418, 442
characteristics 420
transform fault 445–6, 453
gravity 207, 411
^{3}He 363
heat flow 286
hydrothermal circulation 428–31
intermediate
characteristics 420
topography 418
isostatic equilibrium 409–10
magma chamber 426, 428, 435, 439, 448
mantle S-wave velocity 410
median valley 417–20, 424
melting 414–17, 423
neovolcanic zone 420
oblique spreading 27, 444
overlapping spreading centre (OSC) 447, 448, 450
propagating 79
Q structure 410
segmentation 439, 443, 447–50, 455
seismic structure 154, 412, 427, 431, 432, 437, 438, 439
slow 420, 439, 454
axial structure 419, 431–2
topography 418
transform fault 442–3
spreading rate 16, 57, 60, 74, 393, 418, 420
thermal structure 426–8
transform-fault crust 439
ultra-fast, axial structure 434–6
upper-mantle seismic velocity 431–4, 440
volcanic emplacement zone 57
migration 172–3
seismic reflection 550, 551, 553
Milne, John 1, 4, 202
Minch Basin 178
minerals 661
Miocene 88
Mohns Ridge 8, 70, 444, *see also* mid-ocean ridge
Moho 2, 152, 178, 326, 406, 408–9, 453, 513, 661
reflection 176, 210, 211, 405, 406
Mohorovičić, Andrya 2, 4
molasse 661
moment magnitude 118, 662
moment of inertia, Earth 334
momentum 662
monazite, closure temperatures 243
monsoon, role of Tibet 539
Moon, orbit 193
MORB 245, 361, 414, 435
Morley, L. W. 56
mountain belt, thermal state 281, 318
mountain building, *see* orogeny
mountain, root 203, 208, *see also* isostasy
mountains, mass deficiency 202
moveout 160, *see also* seismic reflection
Mozambique Ocean 88
multiple reflection 141, 164
muscovite, closure temperature 243, 261
Musgrave block 578
mylonite 177

NACP 552
Nafe–Drake curve 103, 104
Nansen Basin 422
Nashville Dome 564, 566
Nazca plate 5, 8, 21, 23, 31, 79, 83
Neo-Tethys Ocean 545
neovolcanic zone, mid-ocean ridge 420
neutrino 239, 662
New Hebrides subduction zone 8, 392, 461, 465, 517
New Zealand
Southern Alps 319, 320
subduction zone 465, 469
transform fault 440
Newton, Isaac 1, 193, 198
Newton's second law 626
Newtonian fluid/viscosity 224, 353
Ni, in core 351, 353, 381
Ninety-East Ridge 33, 71, 74, 76
NMO 160, 161, 163, 166, *see also* seismic reflection
NMO correction 161
nodal basin, *see* transform fault
nodal plane 134, 139, *see also* earthquake
non-dipole field 45
non-Newtonian fluid 224
non-volcanic margin 568–71
Nootka Fault 78
normal fault 131
normal fault, earthquake fault-plane solution 135, 136, 181
normal modes 108, 326
normal moveout 160, *see also* seismic reflection
normal strain 621
normal stress 620
North America 53, 88
internal deformation 81, 82
west coast evolution 167, 482–5, 517, 551
North American Central Plains, conductivity anomaly 552
North American plate 8, 31, 36, 70, 79, 81, 83, 87
North Anatolian fault 120, 541, 601
North China 88
North Lewis basin 178
north magnetic pole 13
North Sea 8, 511, 559, 575, 576
Northern lights, *see* Aurora
Nubian plate 585
nuclear explosion
energy 125
magnitude 118, 122
surface-wave generation 104
nuclear fission 662
nuclear reactor, natural 260
Nusselt number 357
NUVEL-1A 17, 18, 35

O, in core 381
Oahu 219
obduction 662
oblate spheroid 197, 198, 199
oblique shear zone, mid-ocean ridge 447
oblique subduction 14, 319, 320, 465, 469
OC (outer corner) 445, *see also* transform fault
ocean basin 286, 289, 393, 459
ocean-island basalt 363

ocean
area 649
average depth 391
depth variation, mantle convection 214, 218
depth 288–90, 292, 293, 295, 296, 297, 391–4, 649
heat flow 285, 286, 287, 288–90, 296
mass 648
salinity 233
oceanic basalt, isotope ratio 361
oceanic basement 401
oceanic crust 397–407
Archaean 599, 600
back-arc basin 402
composition 397–8, 514
decarbonation 518
formation 405–7, 416
layers 399, 400–2, 439
melting of 518, 519
metamorphic facies 518
microbial life 430
radioactive elements 276
Raitt–Hill layering 400
remanent magnetization 62
seismic structure 326, 398–405, 406
thickness 415–17
transformation to eclogite 462
oceanic-island basalt 361–3
oceanic islands 487–9
oceanic layer 399, *see also* oceanic crust layers
oceanic lithosphere
age 64
cooling 285, 298
dating 54–67
density 289, 292
effective elastic thickness 223, 297
geotherm 303
oldest 60, 67
seismic structure 109, 223, 409, 412, 431–40
strength envelope 514–15
thermal structure 223, 296, 297, 298
oceanic sediment, *see* marine sediments
Oceanographer transform fault 441, 450, 452, 453, 456
ODP (Ocean Drilling Project) 398, 399, 401, 402, 450
Officer Basin 578
offset 662
ohmic decay time 374
oil formation 581–3
oil–gas reaction 583
Oklo, uranium deposit 260
oldest continental material 595
Oldham, R. D. 2, 4
Olduvai Gorge 252
Olduvai polarity event 60
olivine 349, 350, 662
Archaean 600
mantle 346–9, 350
olivine-spinel
mantle 335, 349–50, 366, 369, 464, 472–5
subduction zone 366, 369, 462, 463, 464, 472–5, 476
olivine tholeite 397
Oman, Gulf of 526–8
Ontong Java 33
open system 237, *see also* radioactive dating
ophiolite 88, 405, 535, 545, 662
comparison with oceanic crust 425–6
global map 407
seismic structure 405, 408
orbit, satellite 198
orogen 662
orogeny 10, 88, *see also* continent–continent collision
Orozco transform fault, crust 453
orthopyroxene 349, 662
OSC (overlapping spreading centre), *see* mid-ocean ridge
osmium, *see* Os
outer-arc high 459, 526
outer core, *see* core
Outer Isles Thrust 178, 551
overlapping spreading centre, *see* mid-ocean ridge
overthrust, metamorphism 312–13, 316–18
Owen Fracture Zone 8, 74
oxygen, in core 351

P 126, 149, 150, 151
P, travel time 127, 128
Pacific MORB 414
Pacific Ocean 32, 88
plate-tectonic history 60, 77–88
ring of fire 464
Pacific plate 5, 8, 23, 24, 31, 36, 81, 83, 87
elastic thickness 224
Pacific–Antarctic Ridge 56, 392, 410, 417, 441
Pacific–Farallon Ridge 77, 83, 86
paired metamorphic belts 316
palaeocontinental reconstructions 88
palaeolatitude 52
palaeomagnetic pole position 52, 54
Paleocene 88
Pangaea 88, 545, 662
Pannonian Basin 542
Parana 33, 75
parent atom 236
Parker–Levy dynamo 376, 377
Parkfield, California 121
partial melting 662
partial ray 156
particle motion
body wave 101, 149
surface wave 105
passive continental margins 394, 567–71, 662, *see also* continental margin
Pb isotopes 263
Pb–Pb isochron 249, 251, 264
PcP, travel time 127, 128
Péclet number 358
pegmatites 313
Penninic nappes 548
peridotite 395, 396, 472, 662
melting 518
radioactive elements 276
seismic velocity 413, 513
period 629, 662
peripheral upwarp 563
permeability 289
permeability of free space 648
Permian 88
perovskite 349, 350
Persian Gulf 542
Peru, degree of latitude 202
Peru–Chile subduction zone 8, 21, 120, 392, 394, 461, 465
P_g 141, 511
Phanerozoic
phase change 662

phase velocity 106, 107, 109
Philippine Basin 459
Philippine plate 8
Phoenix Islands 86
Phoenix lineations 86
Phoenix–Pacific Ridge 86
photographic emulsion, track detector 258
Piemont Ocean 544, 545
Pilbara 262
pillow lava 662
PKIKP 126, 127, 129, 330, 331
PKiKP 127, 130, 330
PKJKP 126, 130, 331
PKP 126, 127, 128, 129, 327
PKS 127
plane wave, theory 628–9
planets, origin 263
plastic deformation 222
plate, *see* lithosphere
plate boundaries, types 6
plate formation 285, 290–8
plate models 294–7, 458
  half-space cooling model 291–3, 294, 295, 296
  boundary-layer model 293–4, 298, 304
  GDH1 223, 289, 290, 294, 295, 296, 297
  PSM 223, 289, 290, 294, 295, 296, 297, 303, 304
plate motion
  absolute 32–6
  reconstruction 67–88
plate tectonics 1, 3, 5, 67, 285, 662
  Archaean 370–1, 599
  basic concept 5
  development of transcurrent fault 442
  driving forces 368, 369, 464, 515
  history 2–3
  mantle convection 370
  resistive forces 369–70
PLATES project 88
Pliny 2
plume 343, 363, 366
pluton 662
P_mP 141, 152
P_n 141, 152
Po Basin 544
Poisson's ratio 624–5
Poisson's relation 625
polar flattening 197, 648
polar wander path 53, 54
polarization of shear waves 663
porosity 289, 453
postglacial uplift, *see* isostatic rebound
potassium 274, 275
  natural isotopic abundance 251
potassium–argon dating, *see* radioactive dating
potential energy 663
potential temperature 307, 416
potential, scalar/vector displacement 102
power law, strain rate 353
power 641
pP 113, 126, 127
PP 126, 127, 128
Prandtl number 358
Pratt hypothesis, isostasy 204–5, 210, 211, 216
Pratt, J. H. 202
Precambrian
precision 234
precritical reflection 143
PREM 326, 327, 330, 651–3
present-day plate motion see Section 2.4
pressure wave, *see* P-wave
pressure 641
pressure, versus depth 335, 371
pressure-release melting 75
primary wave, *see* P-wave 663
primordial heat 380
Prince Edward Fracture Zone 71
principal transform displacement zone (PTDZ) 445, 455
propagating ridge 79
Proterozoic 88
Proterozoic heat flow 298, 299
PS 127
pseudofault 77
PSM, *see* plate models
PTDZ (principal transform displacement zone) 445, 455
Puerto Rico Trench 70
pure-shear extension 571–4, 595
push–pull wave, *see* P-wave
P-wave 101, 157, 663, *see also* body wave
  equation of motion 627
  particle motion 101–2
  reflection and transmission 149, 150, 151
  velocity 102, 413, 513
^{244}Pu, spontaneous fission 259
pyrolite 348
pyroxene-garnet, mantle 335
pyroxenite 349

Q, *see* quality factor
Qaidam Basin 531
Qiangtang block or terrane 530, 531, 532
Qinling–Dabie–Sulu belt, China 554
quality factor, *Q* or attenuation 156, 337, 339
  Earth 335
  core 379
  mantle 344
  melting 413
  seismic waves 337
  subduction zone 466–7
quartz 663
Quaternary
Queen Charlotte transform fault 78, 83, 471
Queen Charlotte triple junction 31
Querwellen, *see* Love wave

R 104, *see also* Rayleigh wave
radar interferogram, Landers earthquake 132
radiation 269, *see also* heat
radioactive dating
  Ar–Ar 252–4, 255
  choice of method 241
  closed system 237, 241–2, 318
  concordant dates 237
  concordia 247–9, 251
  crustal-growth curve 246
  discordant age 237, 248, 249
  first developed 234
  isochron 245, 246, 249, 251, 254, 264, 265, 266
  K–Ar 57, 251–2, 319
  K–Ca 241
  Lu–Hf 258
  mantle-growth curve 246
  metamorphism 318–21
  open system 237
  Pb–Pb 249, 264

radioactive dating (*cont.*)
precision 234
Rb–Sr 244–7, 265, 266
Re–Os 258
Sm–Nd 254
Th–Pb 249–51
U–Pb 234, 247–9
zircon 247, 248
radioactive decay 286, 663
activity 236
Ar 238, 365
branched decay 240–1
^{14}C 238
constant 234, 237
daughter 236, 237
general theory 234–44
half-life 236, 237
K 238, 240–1
Lu 238
parent 236
Rb 238
Re 238
secular equilibrium 240
series 238, 239–40
Sm 238
Th 239
U 238, 239, 247
radioactive elements
concentration 242, 276
continental crust 313
radioactive heat generation 233, 237, 272, 274–5, 288, 659, *see also* heat generation
undepleted mantle 275
radioactive isotope, decay schemes 238, 239–40
radioactivity, discovery 233
radiogenic lead 241
Rae/Hearne craton 552, 554
Raitt–Hill layering, oceanic crust 400
rare-earth elements (REE) 254–8
ray parameter 150, 631
ray path, spherical Earth 630–3
ray series 156
ray theory 156
geometrical 142
Rayleigh, Lord 1, 104, 233
Rayleigh number 356, 357, *see also* convection
convection planform 354, 364, 367
mantle 358, 366
Rayleigh wave 104, 110, 116, 126, 127, 663, *see also* surface wave
amplitude 105
dispersion 106, 109, 127, 128, 409
particle motion 105
phase velocity 409
travel time 127, 128
Rb, isotopes 238, 244
Rb–Sr dating, *see* radioactive dating
Red River Fault 535
Red Sea 8, 31, 542, 585
reduction velocity, record section 148
REE 414, 663
reference gravity formula 197, 205, *see also* gravity
reference spheroid 199
reflected wave 101, 141, 142, 144, 151
reflection coefficient 149–52, 157, 167, 168, 177
reflection seismology 157–77
field method 157–8
multi-layer model 162–5
source 157
two-layer model 158–62
reflectivity method 154, 156
refracted wave 141, *see also* head wave
amplitude 146, 153
refraction 663
refraction line 140, 146
refraction seismology
hidden layer 146
method 140–1, 145
refractory kerogen 581
Reindeer zone 552
relative motion between plates 13, 14, 17, 19–21
relative rotation vector 18, 22, 23
relic spreading centre 445
remanent magnetization, oceanic crust 62
Re–Os dating 238, 258
Réunion hotspot 71, 74, 75
Réunion Island 33, 74, 75
Réunion polarity event 60
reverse fault 131
Reykjanes Ridge 8, 70, 288, 431, 433, 436, 444
crest 417
crust thickness 417
hydrothermal vents 429
seismic structure 432
Reynolds number 357, 368
Rheic Ocean 88
rhenium, *see* Re
Rhine Graben 511, 542
rhyolite 395, 396
Richter magnitude scale, *see* earthquake
Richter, C. F. 115
ridge-push force 368, 369, 370, 371
rift 663
right-lateral fault 663
rigid-body displacement 14
rigidity modulus, *see* shear modulus
ring of fire 464
ringwoodite 350
Rio Grande Ridge 75
Rio Grande Rift 566, 588–93
Rio Grande Rise 394
Rivera plate 83
Rivera subduction zone 468, 470
Rivera triple junction 28, 83
RMS velocity 163, 165, 166
RN corner, *see* transform fault
Roberts number 378
Rocky Mountains 559, 562, 566
Rodinia 88
Rodriguez Triple Junction, *see* Indian Ocean Triple Junction
Romanche transform fault 441, 443, 458
Rossby number 378
rotation (rot) 617
finite 22
instantaneous 15, 22
rotation axis 14, 15
rotation pole 14–15, 16, 18, 19, 24, 440
changes in 77
rotation vector 14–15, 22
rotational wave equation 102
rotational wave, *see* S-wave
RRR, *see* triple junction
RT corner, *see* transform fault
RTI, *see* transform fault
Rutherford, Lord 1, 4, 233, 234, 262, 263
Ryukyu–Philippine subduction zone 8, 461, 477

s 126
S 126, 127, 128

$_0S_0$, $_0S_2$ and $_0S_3$ 110
S20RTS 338
salinity, ocean 233
salt, seismic velocity 513
Samoa Islands 469
San Andreas Fault 6, 31, 81, 83, 120, 121, 441
  seismic hazard 121
  seismicity 124
  slip 18
San Francisco, seismicity 124, *see also* earthquake
sandstone, seismic velocity 513
SAR (synthetic-aperture radar) 131
satellite, orbit 198, 200
SB4L18 338
scalar 615, 663
scalar displacement potential 102
scalar field 615
scalar product 615–16
Schmidt projection/net 180
Scotia arc 459, 461
Scotia plate 5, 8
ScS, travel time 127
seabed depth, *see* ocean depth
seafloor spreading 3, 6, 57, *see also* mid-ocean ridge
sea level 198
  Phanerozoic 558
sealing age 289, 429
seamount chain 32, 487
SEASAT 200
seaward-dipping reflector 569, 571
secondary wave, *see* S-wave 663
secular equilibrium 240
secular variation 48, 379
sediment
  loading 557–8, 563
  marine 285, 289, 398, 399, 400, 401
sedimentary basin 557–83
  compressional 577–81
  extensional 567–77
  flexural 561–5
  thermal 560–1
sedimentary rock 663
segmentation, mid-ocean ridge 447–50
seismic discontinuity 663
seismic gap 468, 469, 470
seismic hazard
  global map plate 3
  U.S.A. map 121
seismic moment 117–18, 122, 123, *see also* earthquake
seismic parameter 333, 352
seismic phases, coding 126–30
seismic-ray paths, mantle 126
seismic reflection
  amplitude 167–8
  anticline 173
  calculation of velocity and thickness 166–7
  CDP field layout 159
  coverage 158
  deep profiling 176–7
  dipping layer 171–2
  fault 174
  Fresnel zone 174
  interval velocity 165
  migration 173, 550, 551
  multiple 164, 165
  resolution 173–6
  Rio Grande Rift 177
  signal-to-noise ratio 162
  stacking velocity 161, 164
  syncline 173
  velocity analysis 161, 162, 163, 164
seismic refraction 140–57
  determination of velocity–depth 144–5
  dipping layer 146–8
  direct wave 145
  field layout 140
  hidden layer 145
  low-velocity layer 145
  multilayer model 145–6
  oceanic crust 398–405
  one-layer model 144–5
  record sections 148–9
  time–distance plot 142
seismic tomography 337–46
seismic velocity
  anisotropy 140
  density 103
  mantle 338, 343, 408–9, 480
  melting 413
  metamorphic rocks 104
  porosity 453
  temperature 339, 343
  values 104, 513
  whole Earth 326–31
Seismic wave, *see also* body wave and surface wave
  amplitude 115, 149, 152–3, 336, 337
  attenuation 337
  polarization 102
seismicity 5, 7, 131, 542, 663, *see also* earthquake
  Africa–Eurasia 440
  Alaska 471
  Alps 548
  Cascadia 482
  Gakkel Ridge 422, plate19
  Japan 115, 120, 467, 521
  Mid-Atlantic Ridge 443
  mid-ocean ridge 420–2
  San Francisco region 124
  subduction zone *see* subduction zone
  Tonga 466, 469, 480
  transform fault see transform fault
seismograph 663
seismology 100–77, 663
seismometer 140, 663
  broad-band 125
  Wood–Anderson 115
self-compression model 304, 333, 334
self-exciting dynamo 374–5, 377
Semail ophiolite 405, 408, 428
sepentinization 663
serpentine 518, 663
serpentinite–peridotite, reaction 518, 520, 521
S_g 141
SH 102, 149, 347
shadow zone 146, 327, 328, 329, 437
shake wave, *see* S-wave, or body wave
shale 242, 663
Shatsky Rise 33
shear, angle 621
shear component of strain 622
shear heating 272
shear modulus 102–4, 117, 331, 624
shear stress 462, 519, 620
shear wave, *see* S-wave
Sheba Ridge 585
shock-wave experiments 351, 352, 371
SHRIMP 243
Shumagin gap 468, 470
Siberia 88

Sierra Nevada 31, 300, 302, 566
silicate 663
silicon, in core 351
sill 664
Sinai 32
sine formula 20
  spherical triangle 53
sinistral fault 12, 131, 443, 664
skin depth 284
SKJKP 331
SKS, travel time 127, 128
slab graveyard 339, 362, *see also* mantle
slab-pull 368, 369, 371
slip vector
  Africa–Eurasia 440
  Alaska/Aleutian 465, 471
  Anatolia 542
  California 18, 124
  earthquake 131, 135, 137, 139
SLR 17
small circle 15, 16
Sm–Nd dating, *see* radioactive dating
S_mS 141
S_n 141, 410
Snell's law 101, 142, 146, 150, 630
snowball Earth 88
Soddy, F. 233
solar day length 648
solar nebula, composition 264
solar wind 47
solenoidal field 617
solid-state track detector 259
solidus 664
Solomon–New Hebrides subduction zone 8, 461, 465, 517
Somali Basin 71
Somalian plate 585
Songpan–Ganzi Terrane 531
sound wave 101
South America 88
South American plate 8, 21, 70
South Chile subduction zone 465
South Fiji Basin 459, 469
south magnetic pole 13
South Sandwich–Scotia subduction zone 70, 461
Southeast Indian Ridge 8, 70, 74, 76, 418
  spreading rate 74
Southern Alps 319, 320
Southern Lights, *see* Aurora
Southwest Indian Ridge 8, 70, 71, 74, 418
  bathymetry 392
  crust 407, 416
  earthquake 410
  spreading rate 71, 393, 407
  transform faults 441, 452
space weather 47
specific heat 271, 272, 305, 642, 664
  core 379
  lithosphere 294
sphene, closure temperatures 243
spherical polar coordinates 619
spheroid 197, 198, 199, 214
spheroidal oscillations, *see* free oscillations
spinel–perovskite phase change 473
spinel–post-spinel phase change 335, 349, 350, 363, 366, 462, 464
spontaneous fission 258, 259
spreading centre, *see* mid-ocean ridge
spreading rate, *see* mid-ocean ridge
Sr isotopes 244, 246
Sr ratio of sea water 486
sS 113
SS 127
sSP 130
St Paul's transform fault 443, 454
stacking 158–62, 163, 165, 166
stacking velocity 161, 164
standard error 637
stony meteorite, *see* meteorite
strain 621–3, 664
strain rate 353, 514
stratigraphy 664
streamline 664
strength envelopes 514–15
stress 620, 623, 664
  World Stress Map 36
strike–slip fault, *see* fault
Strutt R. J. 233, *see also* Rayleigh
subduction 664
subduction zone 6, 10, 12, 459, 461, 664
  56 Ma ago 345
  accretionary wedge 459, 519, 526–30
  Aegean 461
  Alaska 465, 470
  Aleutian 461, 465, 470, 471
  attenuation 466–7, 468
  bathymetry 222, 394
  Cascadia 31, 461, 470, 482–5
Central America 465
crustal contamination 524
deep earthquake 471–5
deformation front 528
density 345, 528
dip 461, 519, 521
double seismic zone 474, 476
earthquake depth 464–5
earthquake mechanism 470–1, 473, 478
eclogite 462
flat 526
gravity 475–6, 478, 528
heat flow 478, 482
Himalaya 461
hydrofracture 521
induced mantle flow 524
Iran 461
Izu–Bonin 461, 465, 474, 477
Japan 86, 459, 464–5, 466, 467, 478
Kermadec *see* Tonga
Kurile 465
length 461
LIL elements 522
locked 472, 473, 483
magma chemistry 485–6
Makran 459, 526–8
Mariana 465, 477, 485
melting 485–6, 523, 524
metastable wedge 463
Middle America 83, 461
New Hebrides 8, 392, 461, 465, 517
New Zealand 465, 469
oceanic crust melting 519
olivine-spinel 462, 463, 464, 476
Peru 461, 465
Q structure 466–7
Rivera 468, 470
water 431, 462, 465, 466, 482, 485–6, 517–19, 522–4
Ryukyu–Philippine 461, 477
sediment melting 519, 522
seismic gap 467, 468
seismic profile 168, 528
seismic velocity 339, 465, 466–7, 477–85, 526–30
seismicity 328, 463, 464–75
shear stress 462, 519
Shumagin gap 468
slab depth 345, 465

Solomon *see* New Hebrides
South Sandwich–Scotia 461
subduction rate 461
Sunda–Java–Sumatra–Indonesia 461, 528–30
thermal parameter 461, 462, 474, 475
thermal structure 281, 283–5, 460–4, 520
Tonga–Kermadec 460, 461, 464, 465, 466, 468, 469, 474, 485
uplift 470, 473
volcanic line 465, 485
Sub-Himalaya 533, 534
subsidence 664
Bay of Biscay 574
central Australia 578
Illinois Basin 561, 564
Michigan Basin 561, 564
North Sea 576
Nova Scotia margin 575
sulphur, in core 351
Sumatra Trench, seismic reflection 528
Sun 47, 648
abundance of elements 351
age 233, 234
radiation 270
temperature 288
Sunda–Java–Sumatra–Indonesia subduction zone 461, 528–30
sunspot 47
superadiabatic gradient 334, 364
Superior craton 299, 300, 552, 553, 554, 597
surface wave 104–8, 664
amplitude 104, 152
dispersion 106–8, 109, 116, 297, 326, 327, 409
magnitude 122, 664, *see also* earthquake
group velocity 106, 107
particle motion 105
phase velocity 106, 107
resolution 105
suspect terrane 517, 664
SV 102, 149, 150, 347
Svalbard 416
swarm, *see* earthquake
S-wave 101, 664, *see also* body wave
equation of motion 627
particle motion 101
propagation through liquid 102
velocity 102, 327
mantle 338, 350
reflection/transmission 149, 151
syncline 173, 664
syn-rift basin 568
synthetic refraction seismogram 153–7, 403
synthetic seismogram 154

$_0T_2$, $_1T_2$ 110
TAG, Mid-Atlantic Ridge 421
Tamayo transform fault 441, 453
Tarim basin 535, 537
tear fault 131
tectonics 664
temperature 642
centre of Earth 287, 288, 371
lithosphere 287, 292, 293, 294
periodic variation at surface 281, 283–5
Sun 288
tensional fault 131
terrain correction 206
terrane or terrain 664
accreted 482, 516
suspect 517
Tertiary
tesla 664
Tethys Ocean 88, 405, 545
Tethys slab 339, 344
theoretical seismogram, *see* synthetic seismogram
thermal boundary layer 298, 304, 306, 359
thermal capacity, *see* specific heat
thermal conductivity 270, 271, 279, 294, 642
thermal diffusivity 273
core 379
lithosphere 294
thermal expansion 294, 305, 307, 664
thermal models 223, 296, 298, 455, 460–4, 519–21
thermal parameter 462, 463, 474, *see also* subduction zone
thermal sedimentary basins 560–1
thermal time constant, lithosphere 560
thermonuclear fission 234
thermoremanent magnetization 51, 52
tholeite 397, 664
tholeitic basalt, radioactive elements 276
Thompson belt 552, 553
Thompson, William *see* Kelvin
thorium 274, 275
Th–Pb, dating 249–51, *see also* radioactive dating
thrust fault 131, *see also* fault
earthquake fault-plane solution 135, 136
Tian Shan 537
Tibet
crustal thickness 530, 531, 537, 539, 541
gravity 540
isostatic model 540
partial melt 537–9
lithosphere 531, 539
monsoon 539
seismic structure 537, 538
tectonic evolution 530
tides 2
Tien Shan 535
timescale 234
time–temperature index (TTI) 582
time–distance graph 144
tomography 337–46
tonalite 395, 525, 664
Tonga subduction zone 461, 464, 465, 466, 468, 469, 474, 485
detached slab 466
lava chemistry 486
seismic structure 480
seismicity 480
thermal parameter 463, 474
Tonga Trench 8, 222
topography, global map 200, plate8
torsional oscillation 108, *see* toroidal oscillations *or* free oscillations
total-fusion method 252
trace element 665
track detectors 258, 259
Trans-Hudson Orogen 552
transcurrent fault 131, 442, *see also* fault
transform domain 444, *see also* transform fault
transform fault 6, 10, 12, 16, 363, 394, 448
age contrast 454
Atlantis II 398, 441
Blanco 441

transform fault (*cont.*)
characteristics 441
Charlie Gibbs 441, 456, 486
Clipperton 441, 442, 453
crust 439, 444, 450–3, 455–6
Dead Sea 31
dextral 443
earthquake fault-plane solution 134, 136, 181
Eltannin 441
FAMOUS 441
fast-spreading ridge 445–6, 447
fracture zone 441, 446
Garrett 434
geometry 440–4
Gofar 441
gravity 453
IC (inner corner) 445
Kane 441, 450, 456
Kurchatov 441
leaky 440
lithosphere thickness 454, 457
Mendocino 31, 441
microearthquakes 458
Mid-Atlantic Ridge 441, 451, 452
mid-ocean-ridge segmentation 447–50
New Zealand 440
nodal basin 445, 488
OC (outer corner) 445
Oceanographer 441, 450, 453
Orozco 453
plate thickness contrast 454
principal transform displacement zone (PTDZ) 445, 455
Queen Charlotte 471
resistive force 370
RN corner 445
Romanche 441
RT corner 445
RTI nodal basin 445
San Andreas, *see* San Andreas
seismic reflection 451
seismicity 456–8, 471
sinistral 443
slow-spreading ridge 442–3, 445, 447, 452, 457
St Paul's 454
strike 16
Tamayo 441, 453
temperature 455, 457
topography 444, 445, 454
transcurrent fault 442
transform domain 444
transform tectonized zone (TTZ) 444, 445
transpression 440
Udintsev 441
Vema 441, 450, 451, 458
Wilkes 441
transformational faulting, olivine-spinel 472, 475
Trans-Hudson Orogen 551–4, 565
transition zone, *see* mantle
transmission coefficient 101, 149–52
transpressional fault 440
transverse ridge 445
trench 6, 83, 86, 363, 458, 459–60, 665, *see also* subduction zone
Triassic
triple junction 14, 25, 26–32, 665
evolution 81, 82
FFF 27
FFR 30
FFT 31
FRR 27, 39
Indian Ocean 8, 70, 71, 74, 418
RRR 27, 31, 70, 79, 86, 418
stability 26, 27, 28, 81
TTT 27, 28, 29
Tristan hotspot 75
TRM, *see* thermoremanent magnetization
troilite 264, 265
TTI, *see* time–temperature index
TTZ, *see* transform fault
Turkey, earthquakes 120, 464, 542

U
natural isotopic abundance 247
radioactive decay 238
fission 258, 260
Udintsev transform fault 441
UHP, *see* ultra-high-pressure belt
Ukranian Shield 300
ultra-high-pressure belt (UHP) 554–5
ultrabasic rock 396
ultramafic rock 242, 349, 395, 396, 514, 665
ULVZ (ultra-low-velocity zone) 346, 347, *see also* mantle
unconformity 665
undepleted mantle 275, 276
uniformitarianism 595
up-dip shooting 147
uplift
Gulf of Bothnia
Hudson Bay 225
Southern Alps 319, 320
upper mantle, *see* mantle
Urals 554
uranium mineral deposit 260
uranium series 247
uranium universal ratio 247
U–Pb closure temperatures 243
U–Pb dating 247–9, 251, 262, *see also* radioactive dating

Valu Fa Ridge, crust 433, 436, 486
Van Allen radiation belt 47
Vancouver plate 80, 81
variation, *see* declination
vector 615, 665
vector addition 13, 21–3
vector displacement potential 102
vector field 615
vector product 615–16
vector triangle 14
velocity 665
velocity analysis, *see* seismic reflection
Vema Transform Fault 441, 450, 451, 458, *see also* transform fault
vibrator 157
Vine, F. J. 56
Vine–Matthews hypothesis 58
viscoelastic lithosphere 561–5, 577, 580
viscoelastic material 665
viscoelastic relaxation time 561
viscosity 224, 665
core 377, 379
mantle 365, 371, 539
Newtonian 224
oil/syrup/water 359
VLBI 17
volatiles 665
volcanic arc 459, 460, 517, 526
volcanic layer, oceanic crust 399, 401, 402
volcanic line 465, 485
volcanic margins 568–71
volcanic rock 665

volcanism 1, 32, 517–26, 588, 665, *see also* subduction zone *and* mid-ocean ridge
Vredefort Dome 596

Wadati, Kiyoo 464
Wadati–Benioff zone 464
wadsleyite 350
Walvis Ridge 394
Wathaman batholith 552, 553
wave equation 102
wavefront 665
wavelength 629
wavenumber 628
Wegener, Alfred 3, 4
wehrlite 349
Western Gneiss Norway 554
Western Transverse Ranges 82
westward drift 379
Wharton Basin 71
white smokers, *see* hydrothermal circulation
wide-angle reflection 151, 398–405, 406
Wiechert, Emil 1, 4
Wilkes transform fault 441
Williston Basin 565–7
Wind River Thrust 510
Wisconsin ice sheet 225, 227
WKBJ ray method 156
Wood–Anderson seismometer 115
work 641
World Stress Map 36
wrench fault 131

Yarlung–Zangbo suture 530
year 648
Yellowstone 33, 87
yield strength of lithosphere 514–15
Yilgarn, Australia 262, 299
Young's modulus 624, 625

Zagros 542
Zimbabwe 315
zircon
  closure temperatures 243, 261
  dating 247, 248, 318
  oldest 595
Zoeppritz equations 150

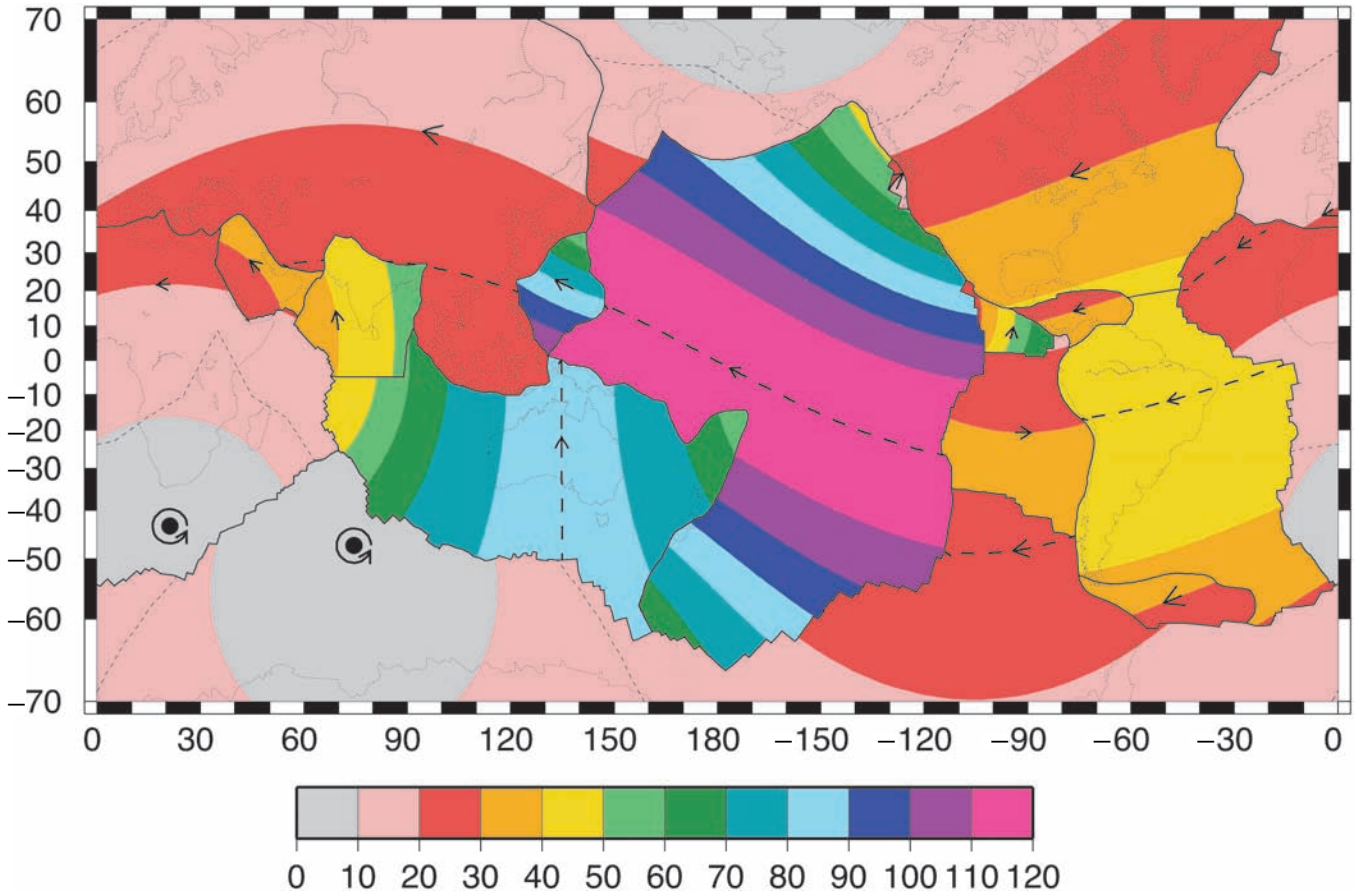


Plate 1. Plate velocities with reference to the hotspots: HS3–NUVEL-1A. Solid black circles, Africa and Antarctica rotation poles, which are each located within their plate; dashed lines, 'equators' for the poles of rotation. Grey version Fig. 2.20. (From Gripp and Gordon (2002).)

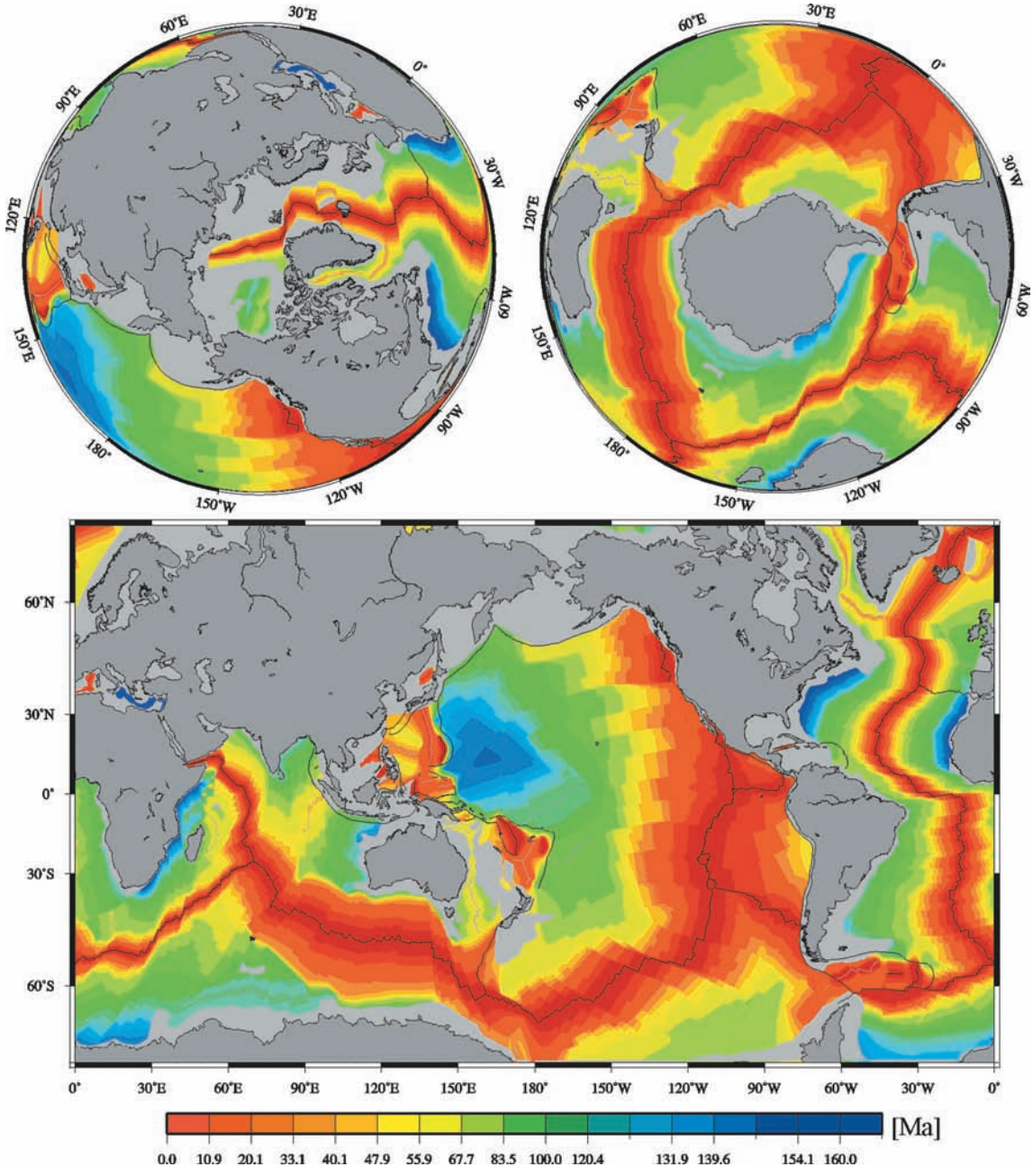


Plate 2. Age of the oceanic lithosphere. Grey version Fig. 3.15. (R. D. Müller, personal communication 2004).

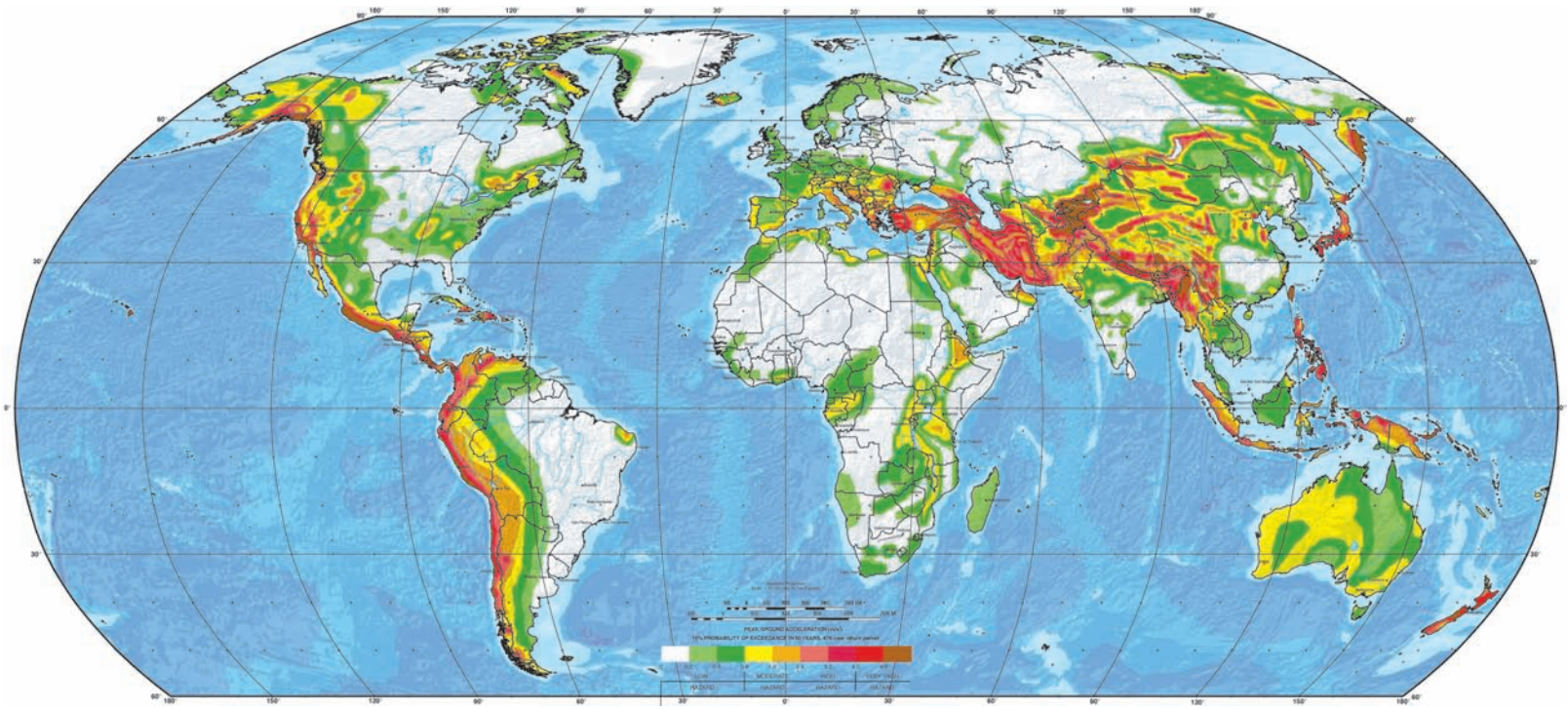


Plate 3. Seismic hazard, probable level of ground shaking associated with earthquakes. The map shows the likely level of short-period ground motion from earthquakes within a fifty-year time window. Hazard levels: white/green; low (0%–8% gravity); yellow/orange, moderate (8%–24% gravity); pink/red, high (24%–40% gravity); dark red/brown, very high (>40% gravity). Greatest seismic hazard is in areas with largest plate-boundary earthquakes. Seismic risk is a combination of seismic hazard with local factors (type/age of buildings/infrastructures, population density, land use, date/time of day). Frequent, large earthquakes in remote areas result in high hazard but pose no risk, whereas moderate earthquakes in densely populated areas entail small hazard but high risk. (From Global Seismic Hazard Assessment Program, http://seismo.ethz.ch/GSHAP/.)

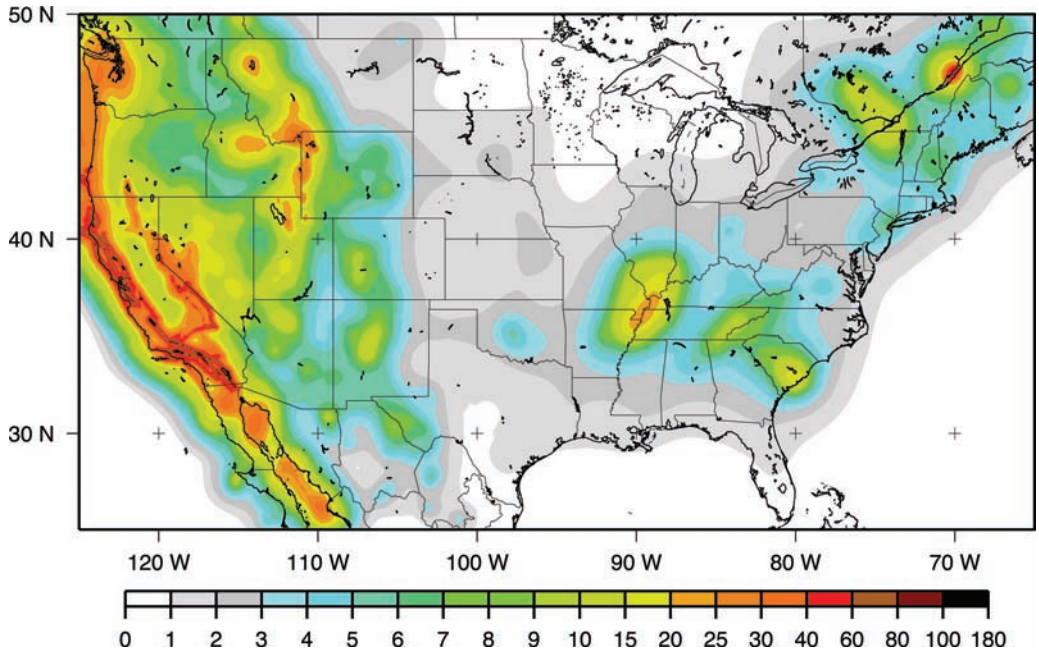


Plate 4. Seismic hazard for the U.S.A. (see Plate 3). Grey version Fig. 4.13.

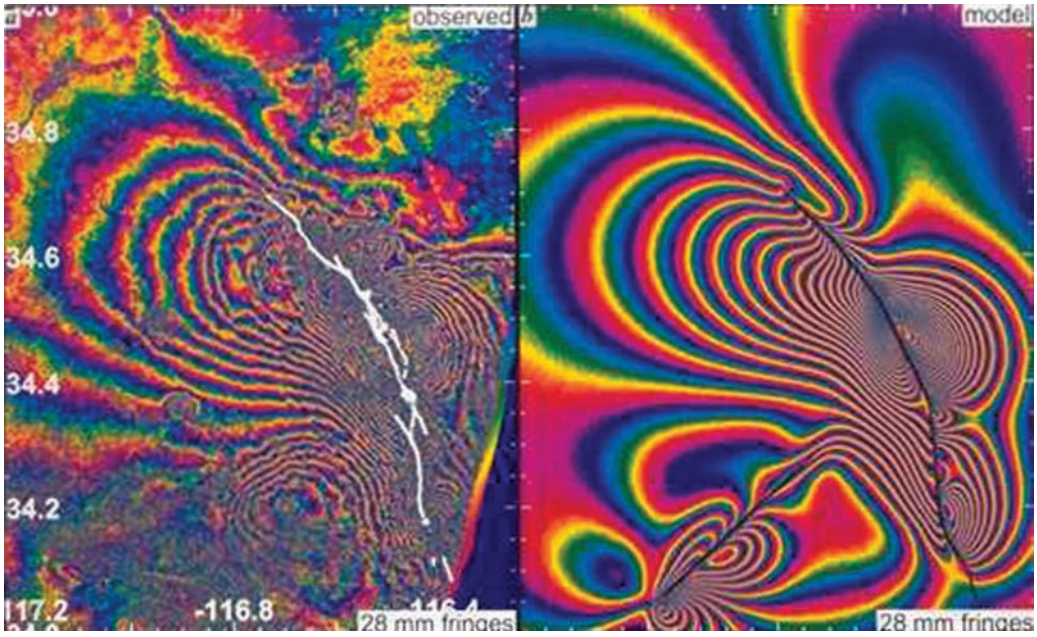


Plate 5. (a) A radar interferogram showing the change in distance between a satellite and the ground surface resulting from the Landers earthquake. Images used to construct the interferogram were taken on 24 April 1992 and 7 August 1992. The $M_W = 7.3$, strike–slip event on 28 June 1992 ruptured 85 km along a set of faults (white). (b) A synthetic interferogram calculated with faults modelled as eight planar segments (black) which rupture from the surface to 15 km depth. Images, 90 km $\times$ 110 km. One cycle of the colour scale represents, 28 mm change in distance. Grey version Fig. 4.21. (Massonnet *et al.*, personal communication 2001.)

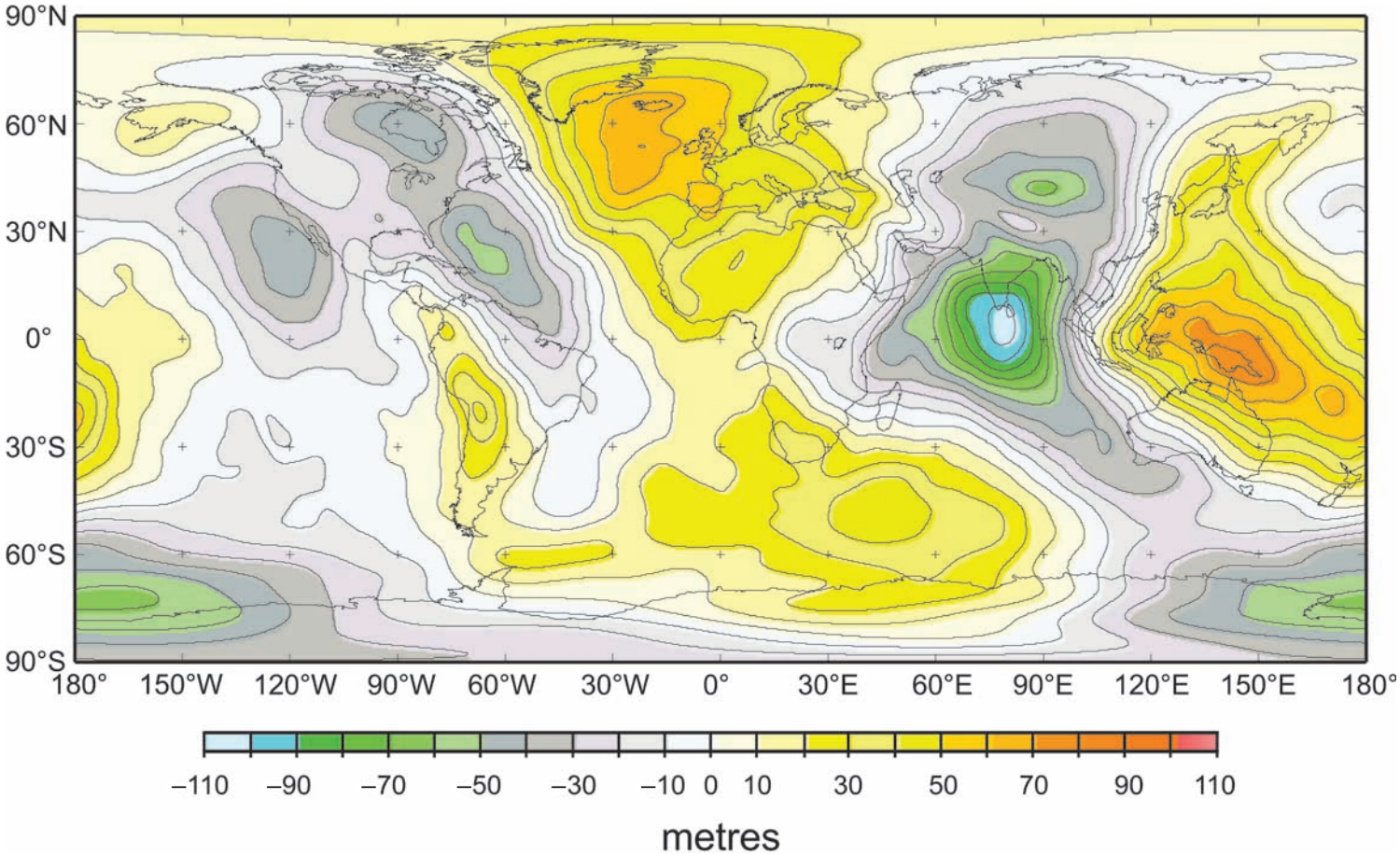


Plate 6. The geoid height anomaly – the height of the geoid above (positive) or below (negative) the spheroid in metres. Grey version Fig. 5.3(a). (Bowin, C., Mass anomaly structure of the Earth, *Rev. Geophys.*, **38**, 355–87, 2000. Copyright 2000 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

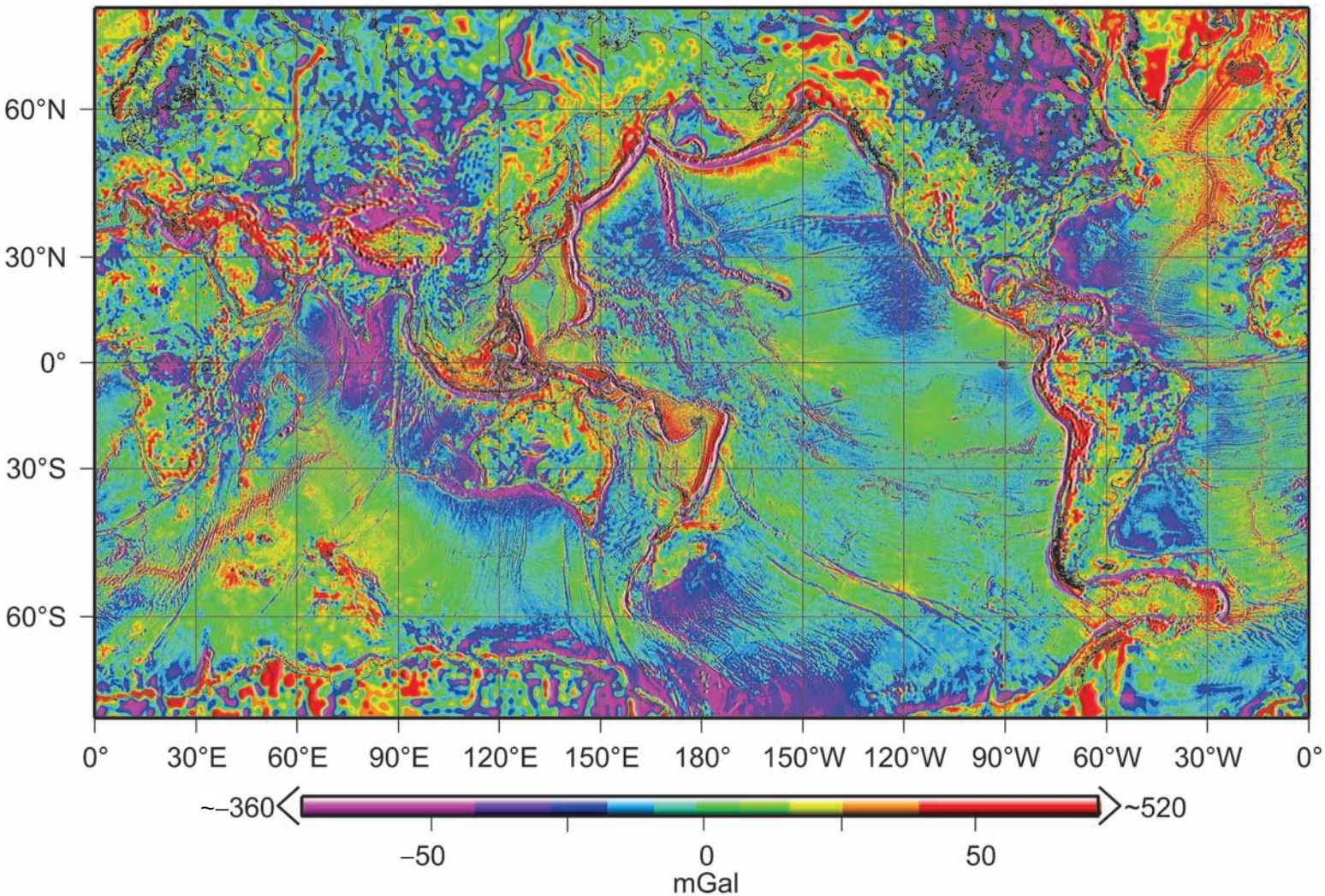


Plate 7. A colour-shaded relief image of the Earth's gravity anomaly field. Grey version Fig. 5.4(a).

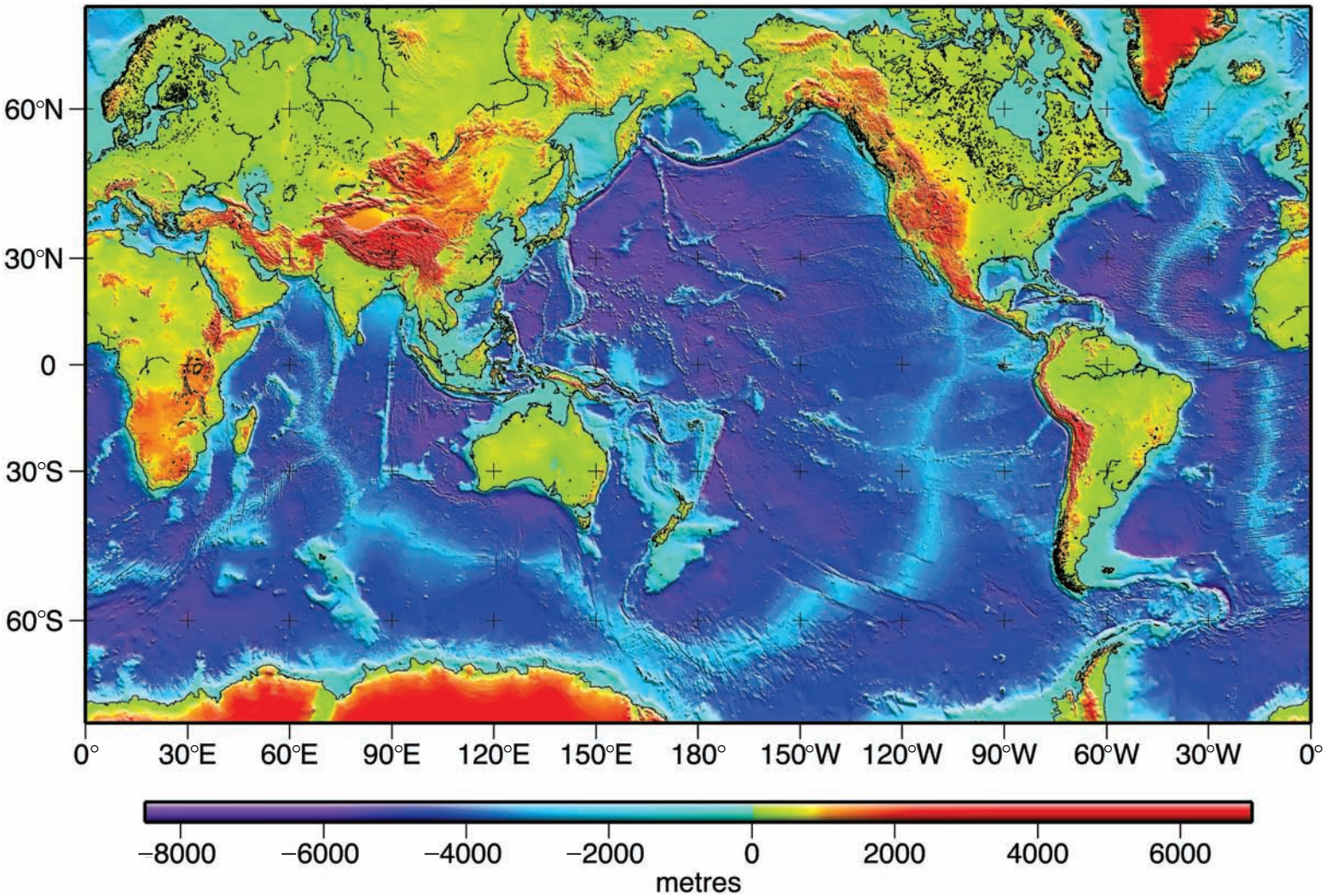


Plate 8. Global topography. Grey version Fig. 5.4(b).

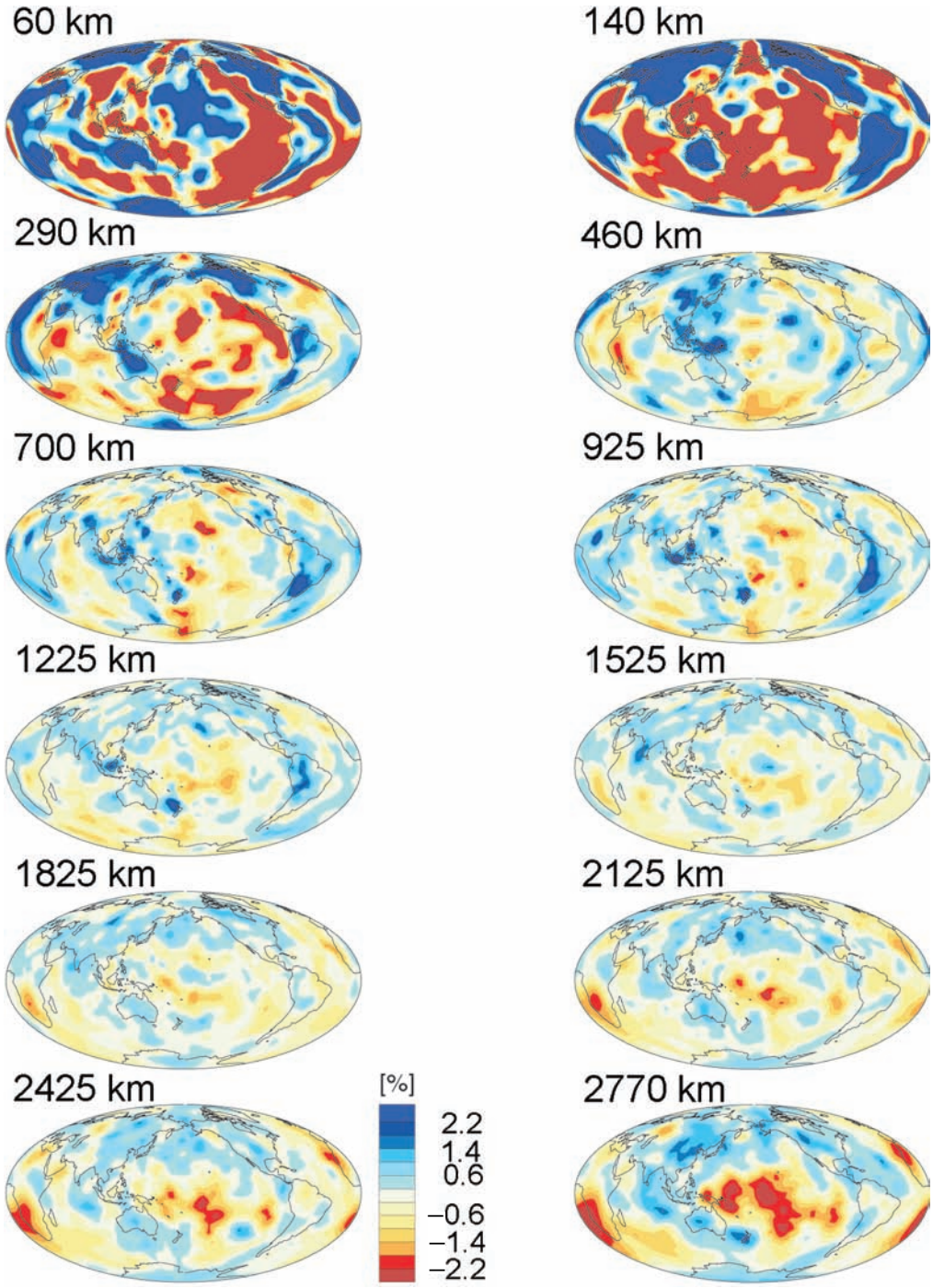


Plate 9. Long-wavelength perturbations of S-wave velocity from a standard whole-Earth model at depths shown. Model SB4L18 used surface-wave phase-velocity maps, free-oscillation data and long-period body-wave travel times. Gradation in shading shows increasing perturbation from the standard model. Maximum deviations decrease from $\geq$2% at the top of the upper mantle to $\pm$1% in the lower mantle and then increase to $\pm$2% just above the CMB. Grey-scale version Fig. 8.6(a). (From Masters *et al.* The relative behaviour of shear velocity, bulk sound speed and compressional velocity in the mantle, *Geophys. Monog.*, **117**, 2000. Copyright 2000 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

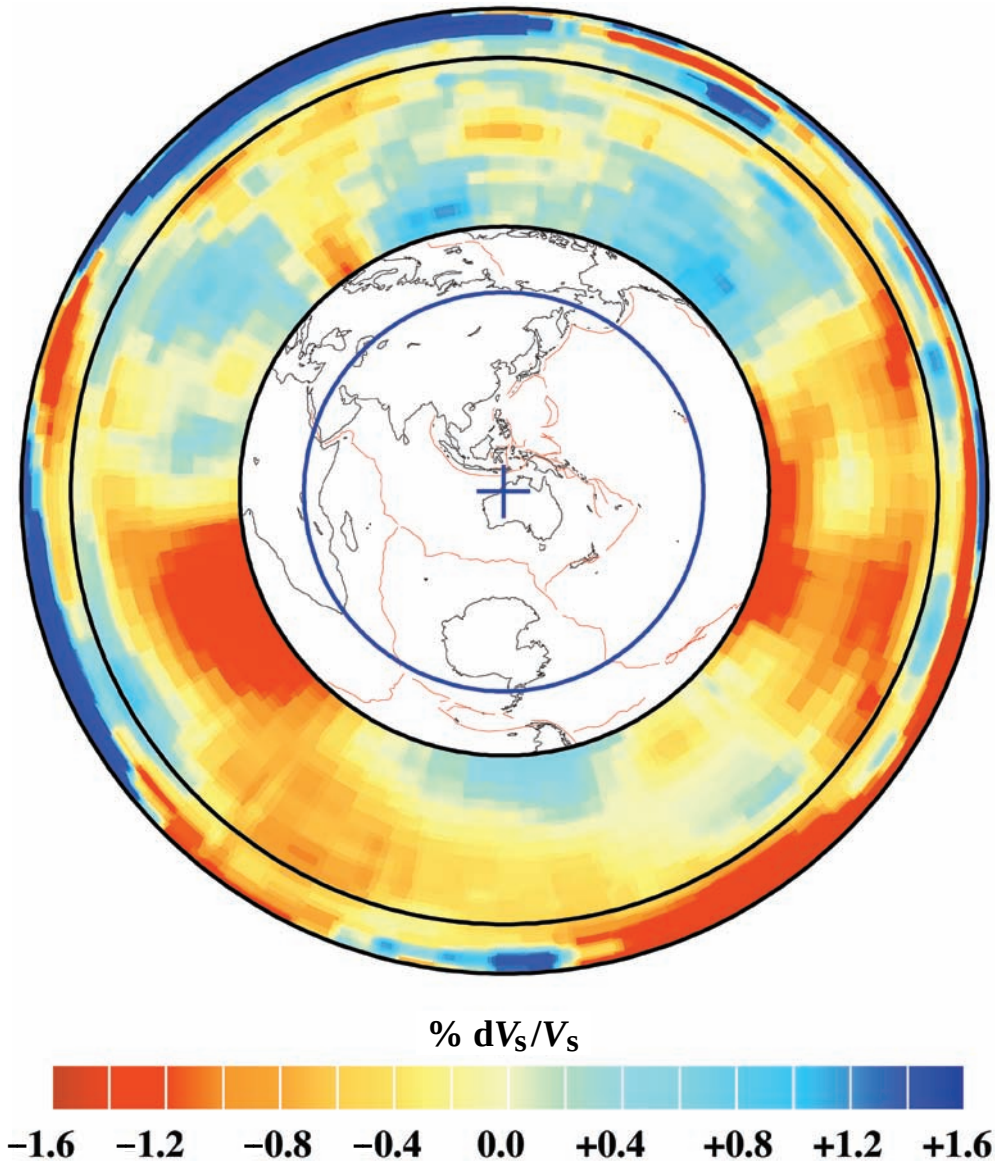
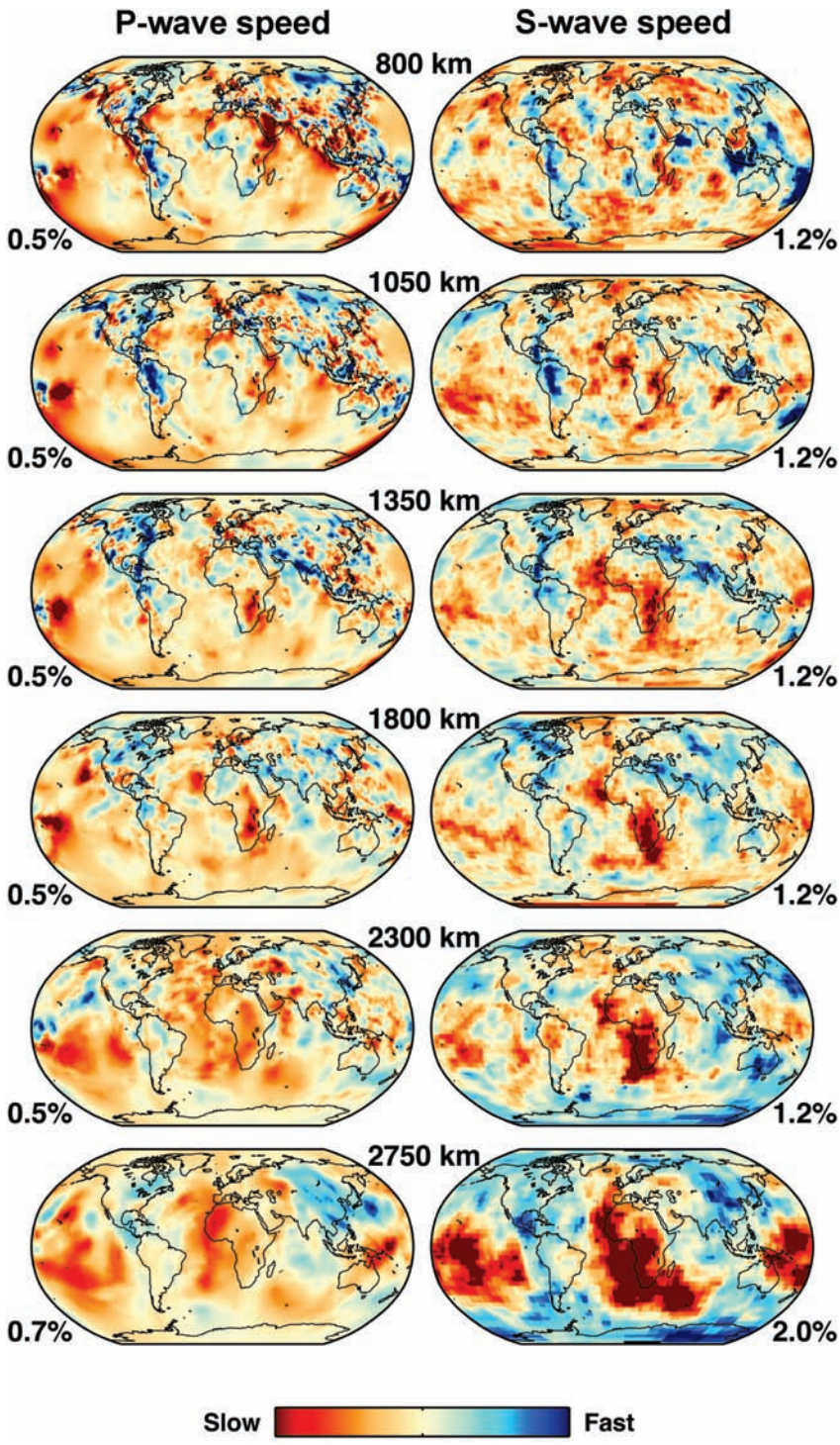


Plate 10. S-wave velocity perturbation model viewed as a slice through the centre of the Earth along the great circle shown as the blue line in the central map. Note two slow (red) anomalies beneath Africa and the central Pacific that start at the CMB, the base of the lower mantle. Black circle, boundary between upper and lower mantle. Grey-scale version Fig. 8.6(b). (Model SB4L18 from Masters *et al*. The relative behaviour of shear velocity, bulk sound speed and compressional velocity in the mantale, *Geophys. Monog.*, **117**, 2000. Copyright 2000 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

Plate 11. Comparison of P- and S-wave velocity models for the lower mantle at depths shown. Models are shown as perturbations from a standard whole-Earth model. The P-wave model was calculated using 7.3 million P and 300 000 pP travel times from ~80 000 well-located teleseismic earthquakes that occurred between 1964 and 1995 (van der Hilst *et al*. 1997). The S-wave model used 8200 S, ScS, Ss, SSS and SSSS travel times (Grand 1994). The maximum percentage difference from the standard model for each image is indicated beside each map. White, insufficient data sampling. Grey-scale version Fig. 8.6(c). (S. Grand and R. van der Hilst, personal communication 2002, after Grand *et al*. 1997.)



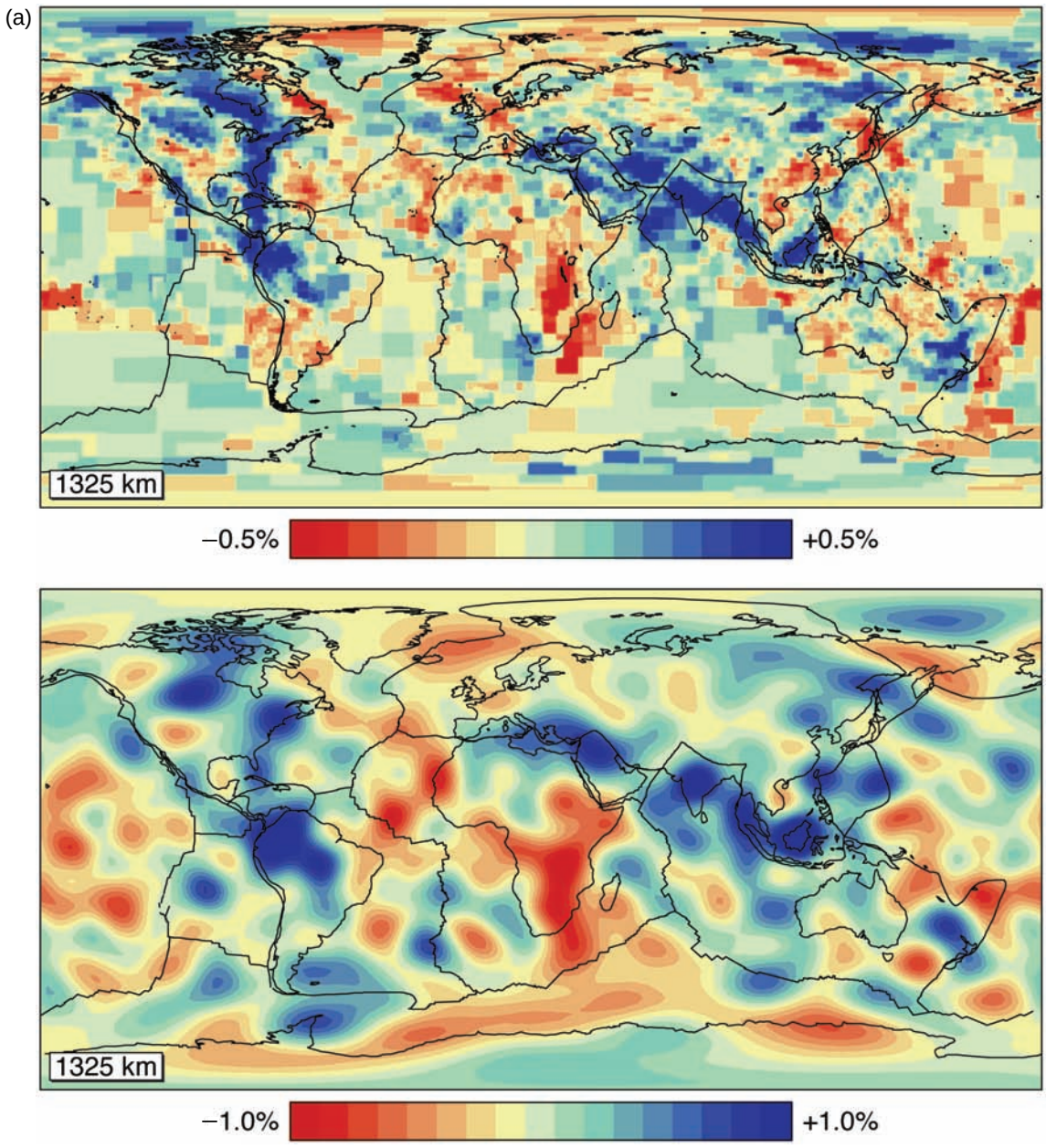


Plate 12. Comparisons of perturbations from a standard whole-Earth model (a) at 1325 km depth and (b) along a section between the Aegean and Japan. Upper panels, P-wave models based on travel times from Bijwaard *et al.* (1998). Lower panels, degree-20 S-wave models based on Rayleigh-wave dispersion (Model S20RTS from Ritsema and van Heijst 2000). Note the different perturbation scales and the different resolutions for the images. Grey-scale versions Figs. 8.6(d) and (e). (W. Spakman, personal communication 2003.)

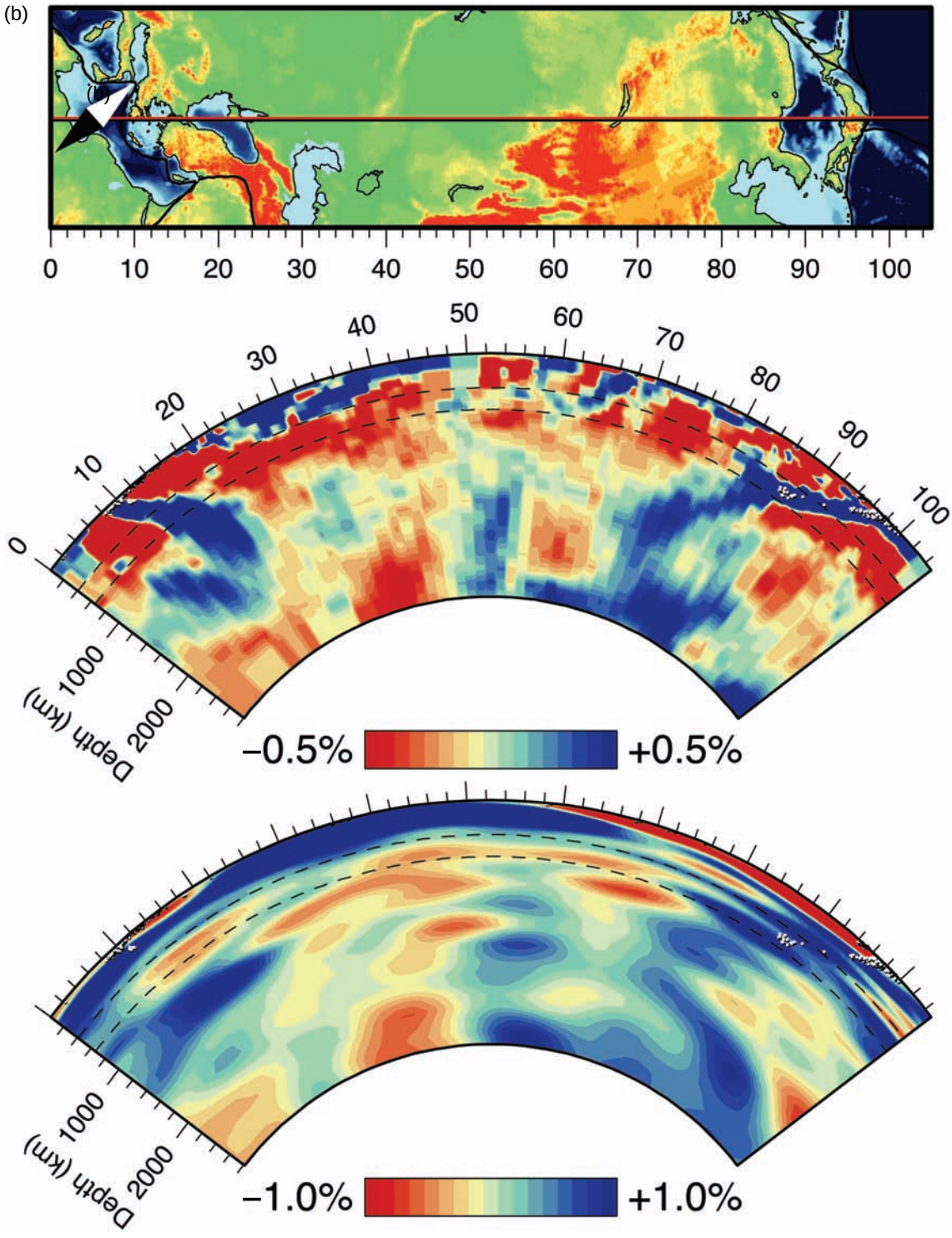


Plate 12. (*cont.*)

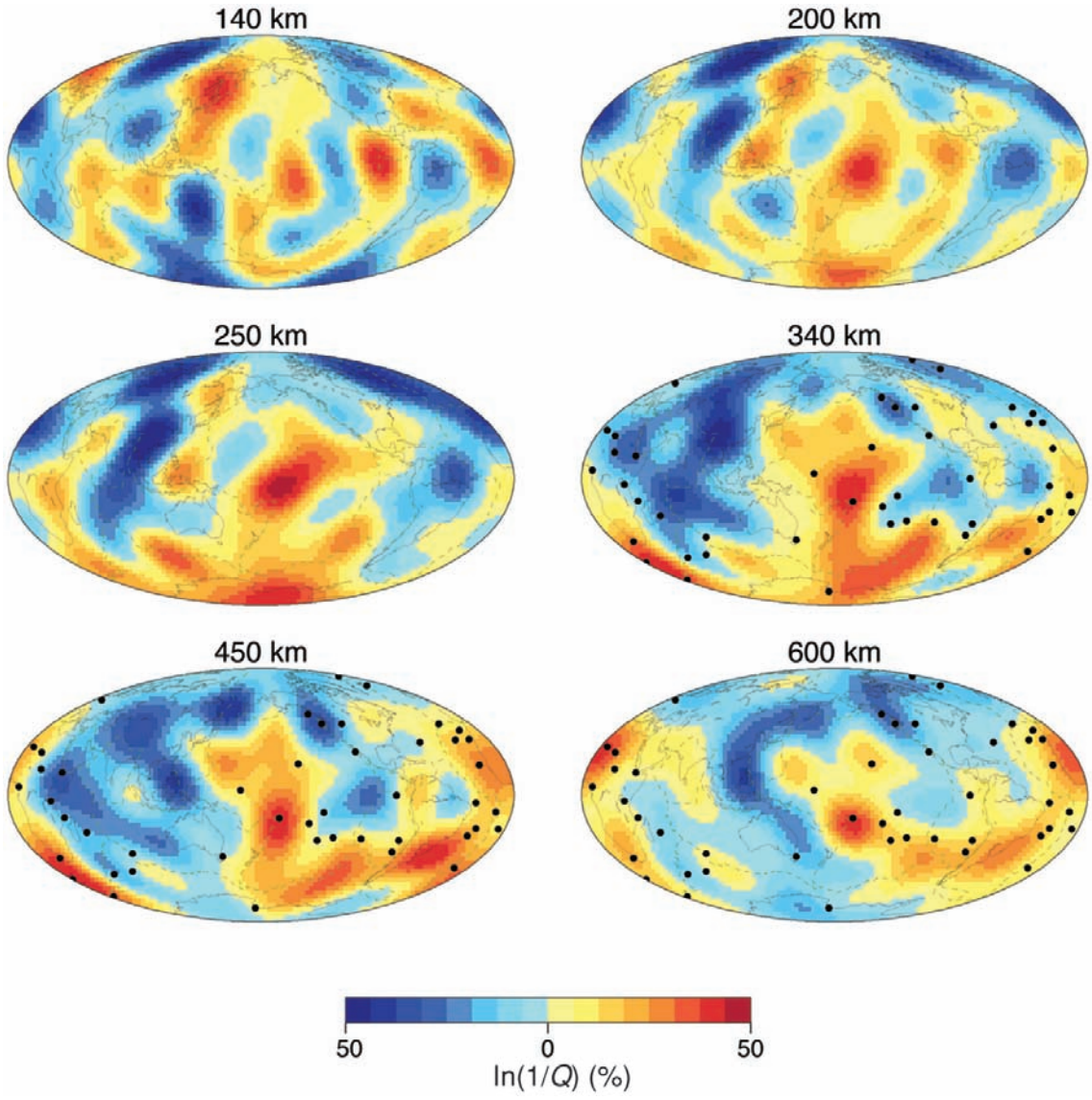


Plate 13. Perturbations in the quality factor *Q* of the upper mantle. Blue, lower than average attenuation (higher *Q*); orange, higher than average attenuation (lower *Q*). Grey version Fig. 8.8. (From Gung and Romanowicz, 2004).

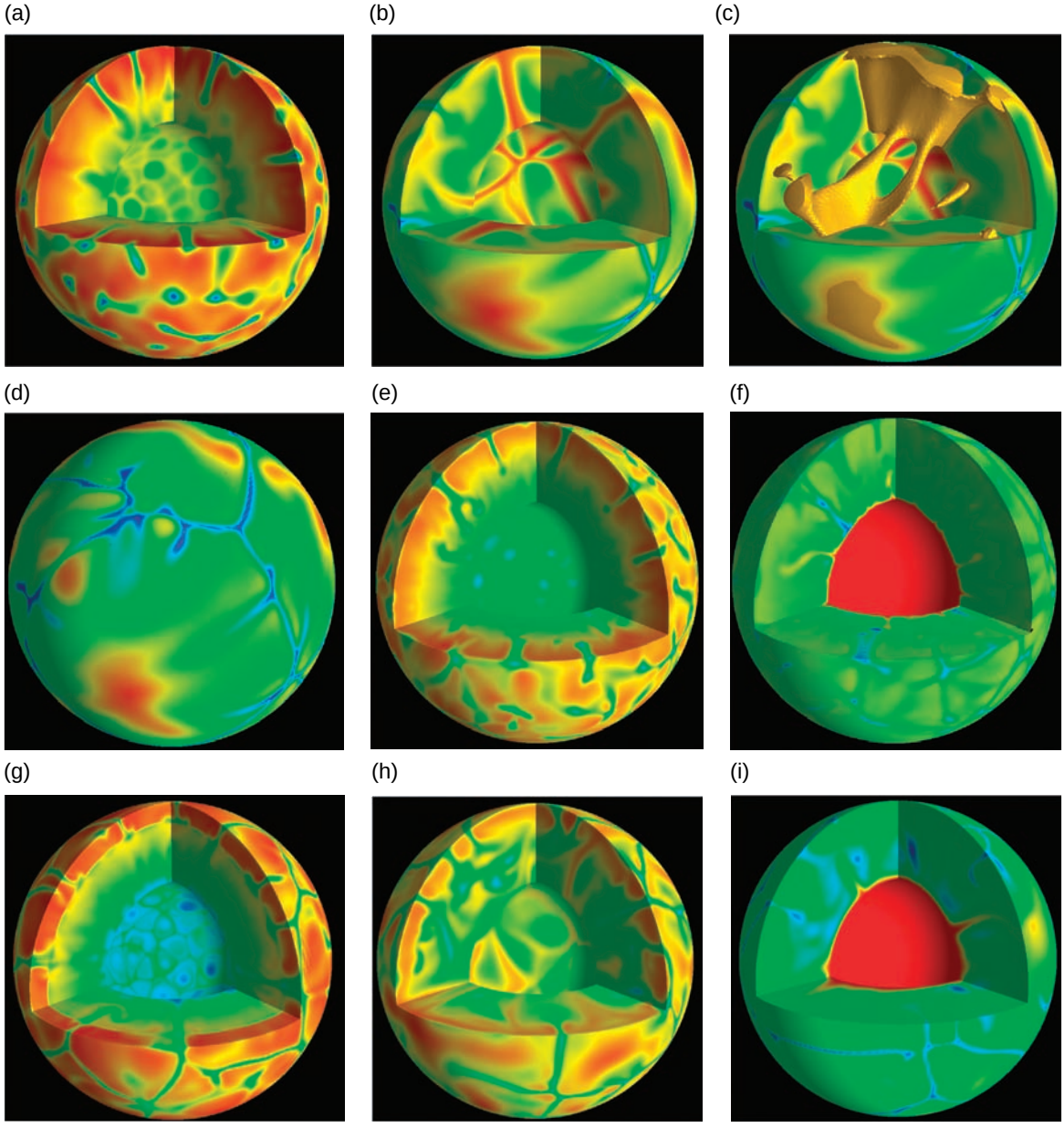


Plate 14. Spherical three-dimensional convection models of the mantle (uppermost 200 km boundary layer not shown). Superadiabatic temperatures: red, hot; blue, cold. (a) Incompressible mantle, constant viscosity, internal heating only, $Ra = 4 \times 10^7$. (b) As (a) but with the viscosity of the lower mantle thirty times the viscosity of the upper mantle. (c) As (b) but showing the isosurface. (d) As (b) but showing the planform. (e) Compressible mantle of constant viscosity, $Ra = 10^8$. (f) As (e) but with 38% of heating from the core. (g) As (e) but with an endothermic phase change of -4 MPa K^{-1} at 670 km depth. (h) As (e) but with the viscosity of the lower mantle thirty times the viscosity of the upper mantle. (i) As (h) and with 38% of heating from the core. Grey version Fig. 8.18. (H.-P. Bunge, personal communication 2003. After Bunge *et al.*) (1996) and Bunge *et al.* A sensitivity study of three-dimensional spherical mantle convection at 10^8 Rayleigh number: effect of depth-dependent viscosity, heating mode, and an endothermic phase change, *J. Geophys. Res.*, **102**, 11 991–12 007, 1997. Copyright 1997 American Geophysical Union. Reprinted by permission of American Geophysical Union.)

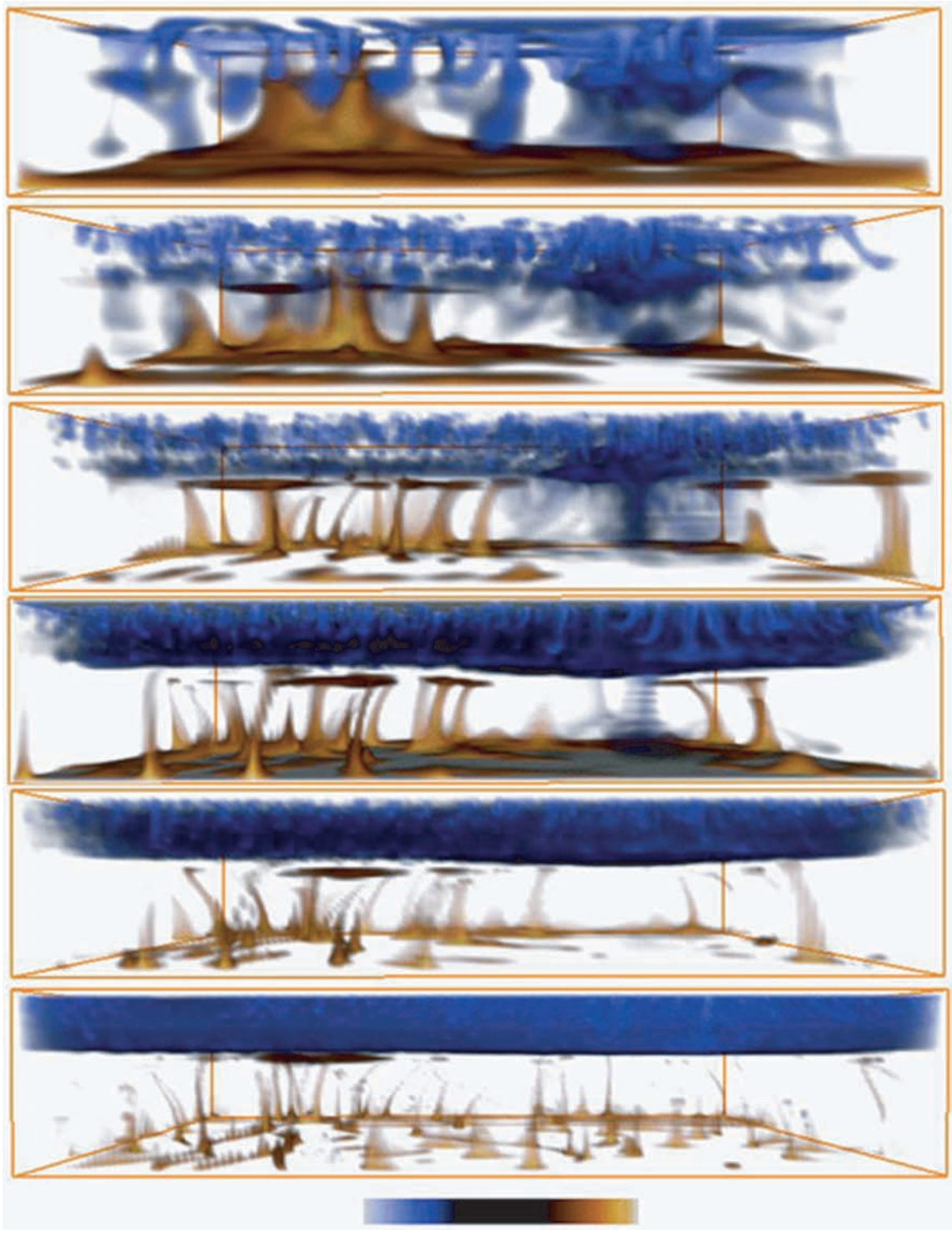


Plate 15. The effect of Rayleigh number on convection in a compressible three-dimensional rectangular model mantle in which viscosity increases with depth and includes phase changes at 400 and 660 km. *Ra*: (a) 2×10^6, (b) 1×10^7, (c) 4×10^7, (d) 6×10^7, (e) 1×10^8, (f) 4×10^8. Red, hot; blue, cold. Grey version Fig. 8.19. (Reprinted from *Phys. Earth Planet. Interiors*, **86**, Yuen *et al*. Various influences on three-dimensional mantle convection with phase transitions, 185–203, Copyright 1994, with permission from Elsevier.)

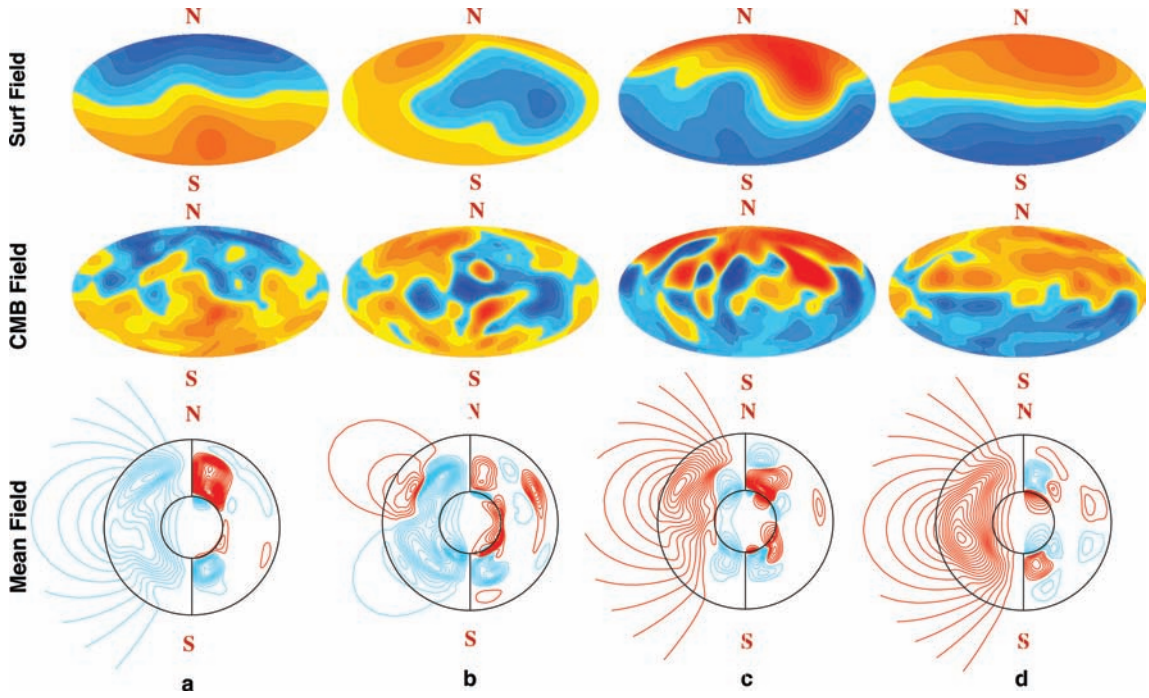


Plate 16. A numerical dynamo model during a reversal of the magnetic field. Columns (a)–(d), images every 3000 years. Top, map view of the radial field at the Earth's surface. Middle, map view of the radial field at the CMB. Orange, outward field; blue, inward field. Note that the magnitude of the surface field is displayed $\times 10$. Bottom, longitudinally averaged magnetic field through the core. Outer circle, CMB; inner circle, IC. RHS, contours of the toroidal field direction and intensity (red, eastward; blue, westward). LHS, magnetic field lines for the poloidal field (red, anticlockwise; blue, clockwise). Grey-scale version Fig. 8.27. (Reprinted with permission from *Nature*, Glatzmaier *et al*. 1999, *Nature*, **401**, 885–890. Copyright 1999 Macmillan Magazines Limited.)

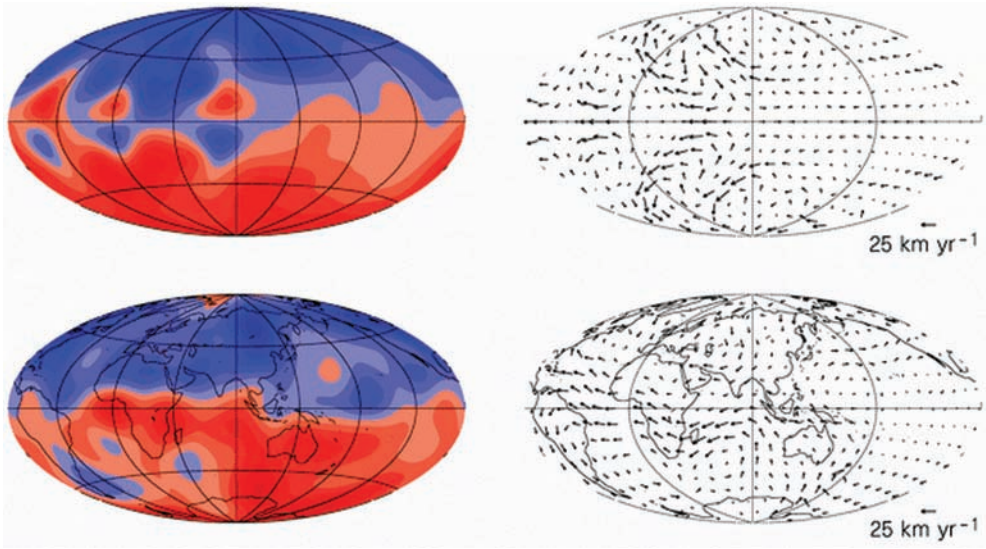


Plate 17. Radial component of the magnetic field (left) and the fluid flow at the CMB (right). Upper, numerical model of the geodynamo with viscous stress-free boundary conditions at the rigid boundaries; lower, the Earth's field averaged between 1840 and 1990. Grey-scale version Fig. 8.28. (Reprinted with permission from *Nature*, Kuang and Bloxham 1997, *Nature*, 389, 371–4. Copyright 1997 Macmillan Magazines Limited.)

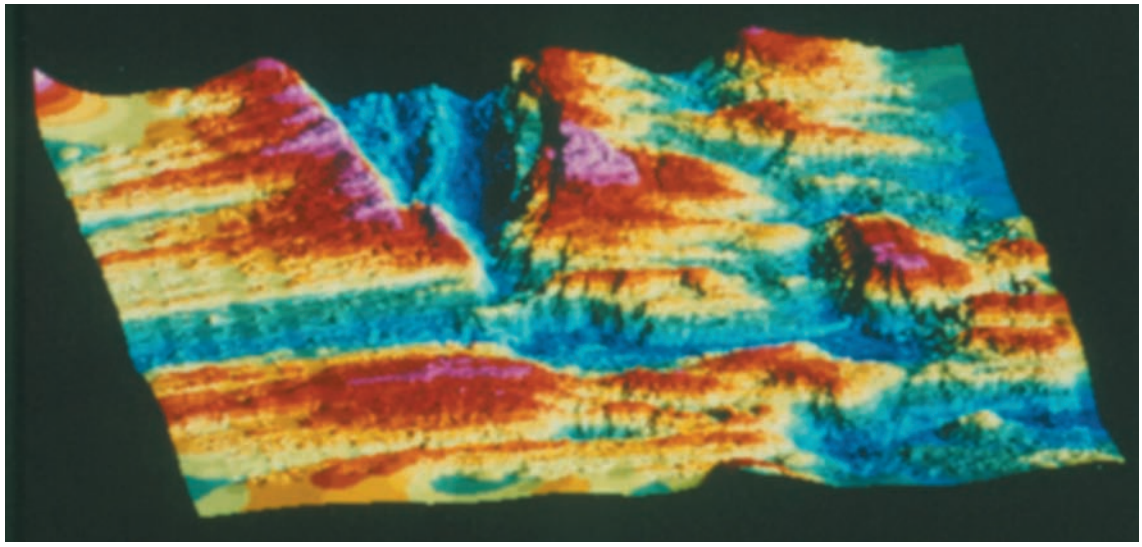


Plate 18. A three-dimensional bathymetric image of the Indian Ocean triple junction where the Antarctic, Australian and African plates meet (Figs. 2.2 and 2.16), looking westwards. The slow-spreading Southwest Indian Ridge (top) has a deep rift valley and uplifted rift flanks; the two intermediate-spreading ridges, the Southeast Indian Ridge (left) and the Central Indian Ridge (right), are both broad and regular and strike almost north-south. Image ~90 km × 90 km. Grey version Fig. 9.15(b). (N. J. Mitchell, personal communication.)

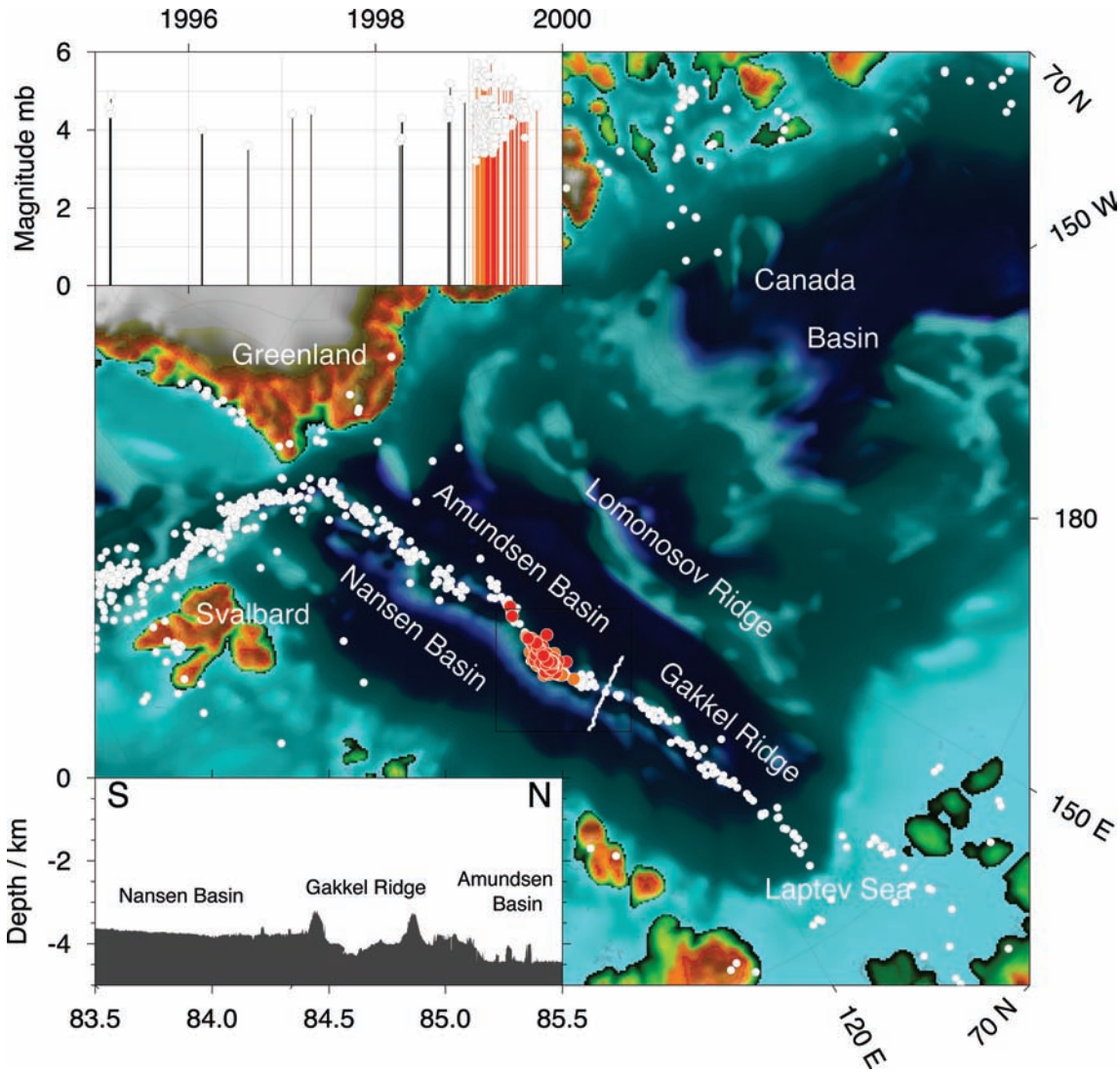


Plate 19. The Gakkel Ridge, the slowest-spreading mid-ocean ridge, runs across the Arctic Ocean. White dots, locations of earthquakes that occurred in 1999. Inset: magnitudes and occurrence of earthquakes on the Gakkel Ridge, 1995–2000. Grey version Fig. 9.18. (From Müller and Jokat, Seismic evidence for volcanic activity at the eastern Gakkel ridge, *EOS Trans. Am. Geophys. Un.*, **81** (24), 265, 2000. Copyright 2000 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

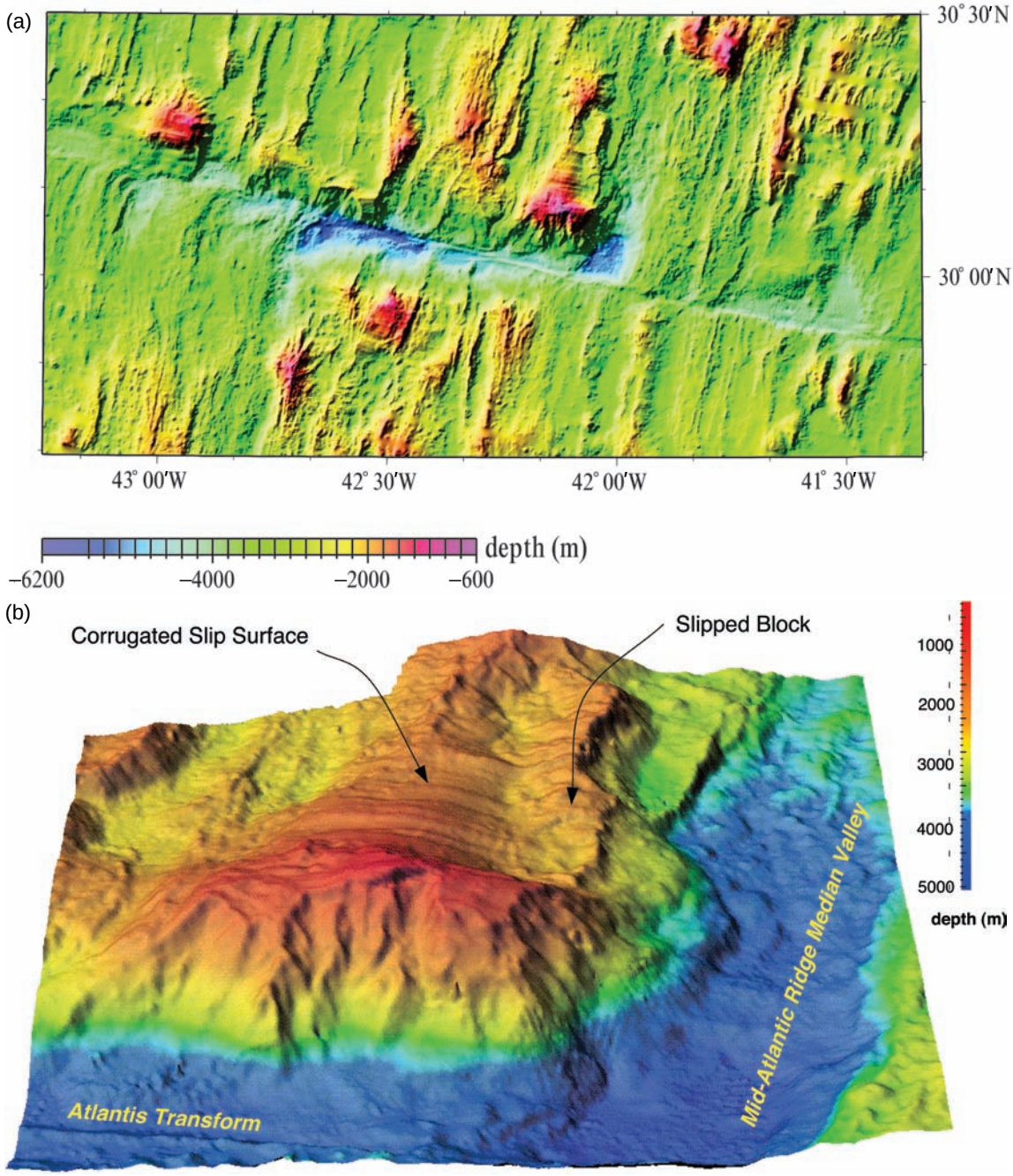


Plate 20. (a) Bathymetry of a 90 km × 180 km area on the Mid-Atlantic Ridge, showing the Atlantis transform fault. The fracture zone continues as a clear linear feature trending 103° both on the African plate and on the North American plate. Three topographic highs show the corrugations suggestive of detachment surfaces: 30° 10′N, 42° 05′W, 30° 15′N, 43° 00′W and 29°55′N, 42°30′W. Grey version and expanded caption Fig. 9.19(a). (From Blackman *et al*. Origin of extensional core complexes: evidence from the Mid-Atlantic Ridge Atlantis fracture zone, *J. Geophys. Res.*, **103**, 21 315–33, 1998. Copyright 1998 American Geophysical Union. Reproduced by permission of American Geophysical Union.) (b) A three-dimensional shaded relief image of the active inside corner shown in (a). The image is viewed from the south and illuminated from the northwest. Grey version and expanded caption Fig. 9.19(b). (Reprinted with permission from Nature, Cann *et al*. 1997, *Nature*, **385**, 329–32. Copyright 1997 Macmillan Magazines Limited.)

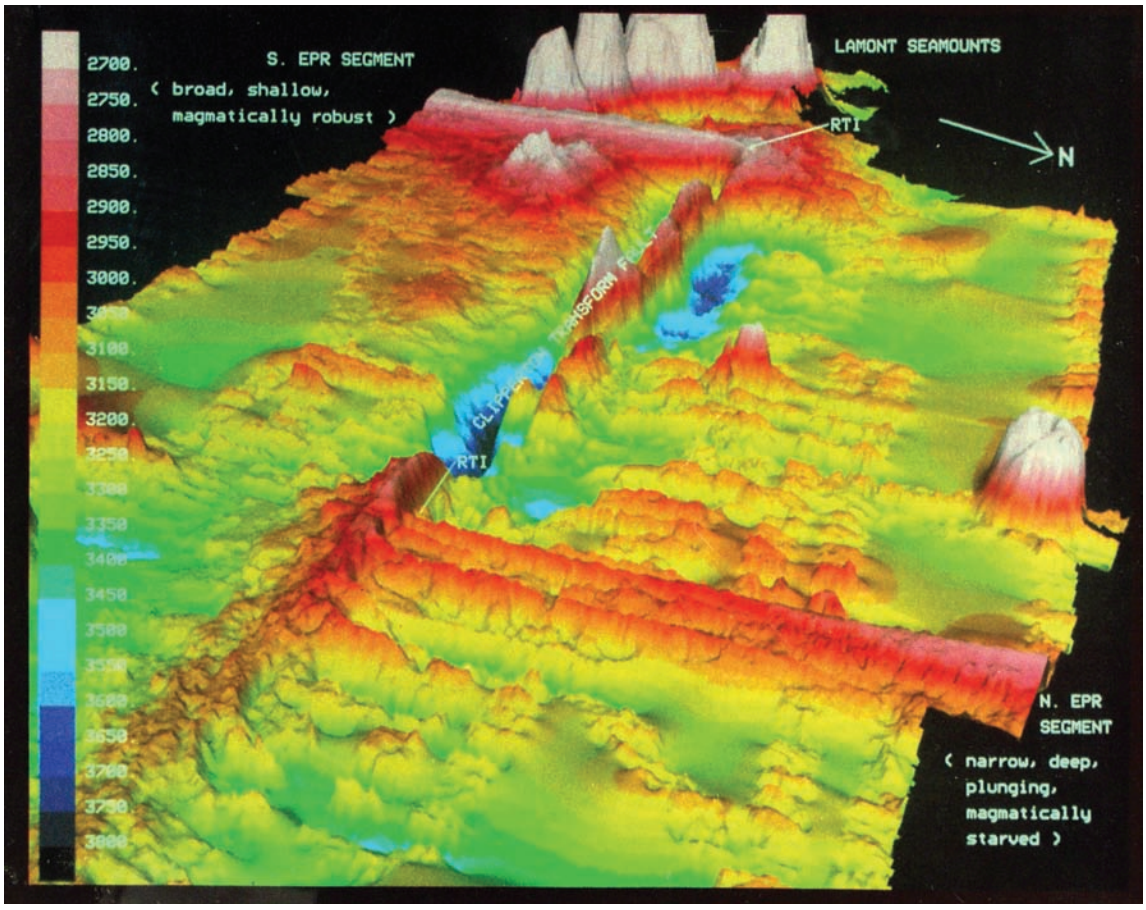


Plate 21. Bathymetry of the Clipperton transform fault on the East Pacific Rise. Image, 100 km (north–south), 175 km (east–west). Transform fault offset, 85 km. The northern ridge segment is starved of magma and deepens towards the transform fault; the southern segment is elevated and is underlain by a magma chamber up to the transform fault. Grey version Fig. 9.31. (K. MacDonald, personal communication.)

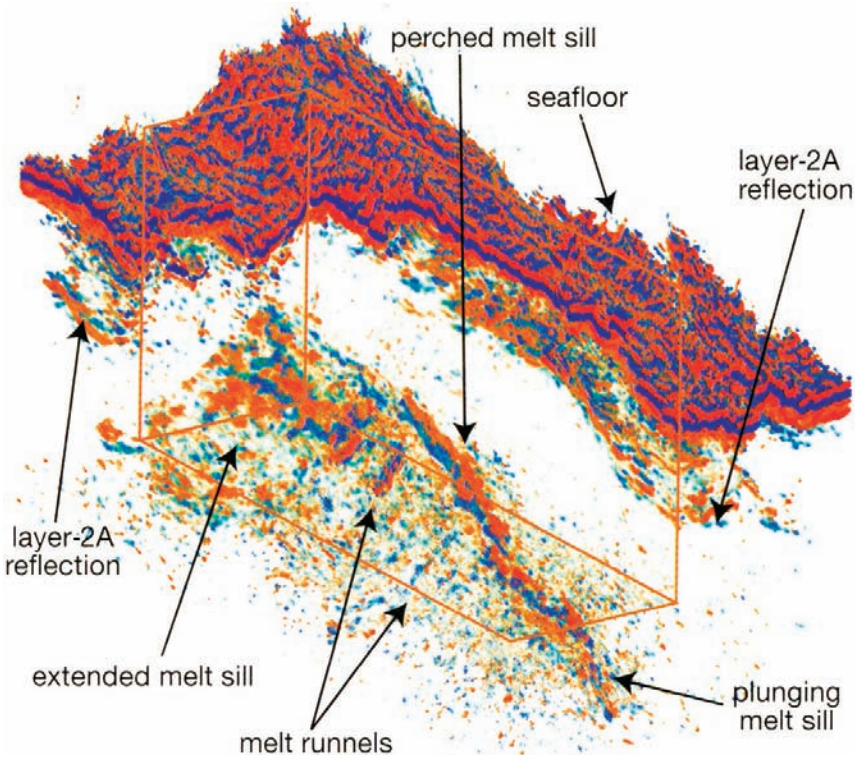


Plate 22. A three-dimensional seismic-reflection survey across a limb of the overlapping spreading centre (OSC) on the East Pacific Rise at 9° N. A cut-away box (orange frame) reveals the reflectivity. Top surface (red), seafloor. Reflections from layer 2A, blue. Reflections from magma chamber, orange and blue–green. Image ~8 km × 14 km. Two-way time from seafloor to base of image ~1.5 s. Fine details of melt transport are complex, but there must be a robust vertical magma supply from the underlying mantle. Grey version and expanded caption Fig. 9.35(b). (Reprinted with permission from *Nature*, Kent *et al*. 2000, *Nature*, **406**, 614–18. Copyright 2000 Macmillan Magazines Limited.)

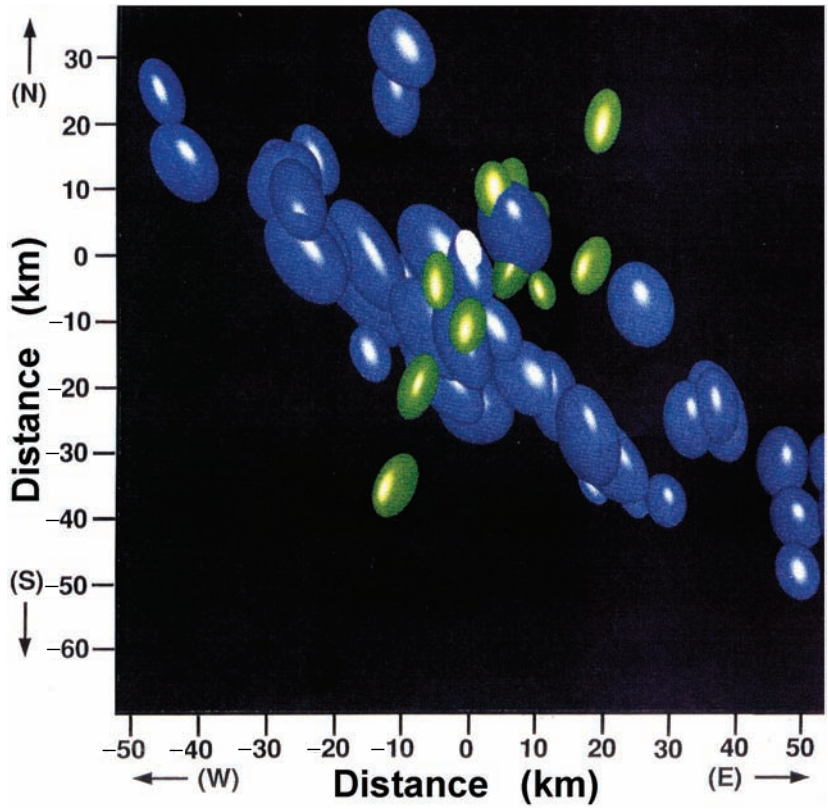


Plate 23. A map view of all well-located earthquakes in the Tonga subduction zone (1980–1987) between 525 and 615 km depth (blue), with the 9 June 1994 earthquake (white) and its best located aftershocks (green) shown as 95% confidence ellipsoids. Grey version and extended caption, Fig. 9.56. (D. Wiens, personal communication.)

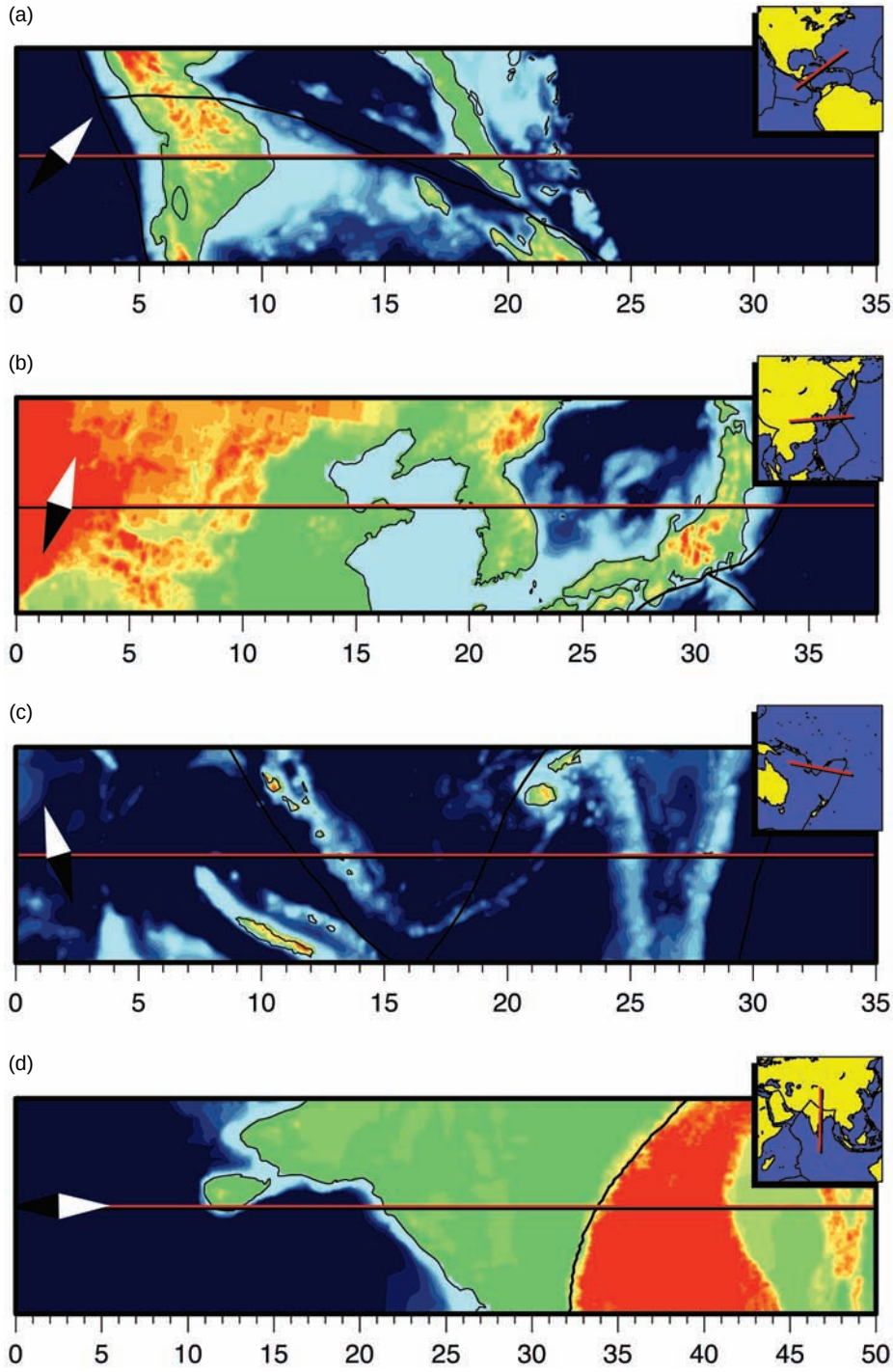


Plate 24. Location maps for the cross sections shown in Plates 25(a)–(d). (W. Spakman, personal communication 2003.)

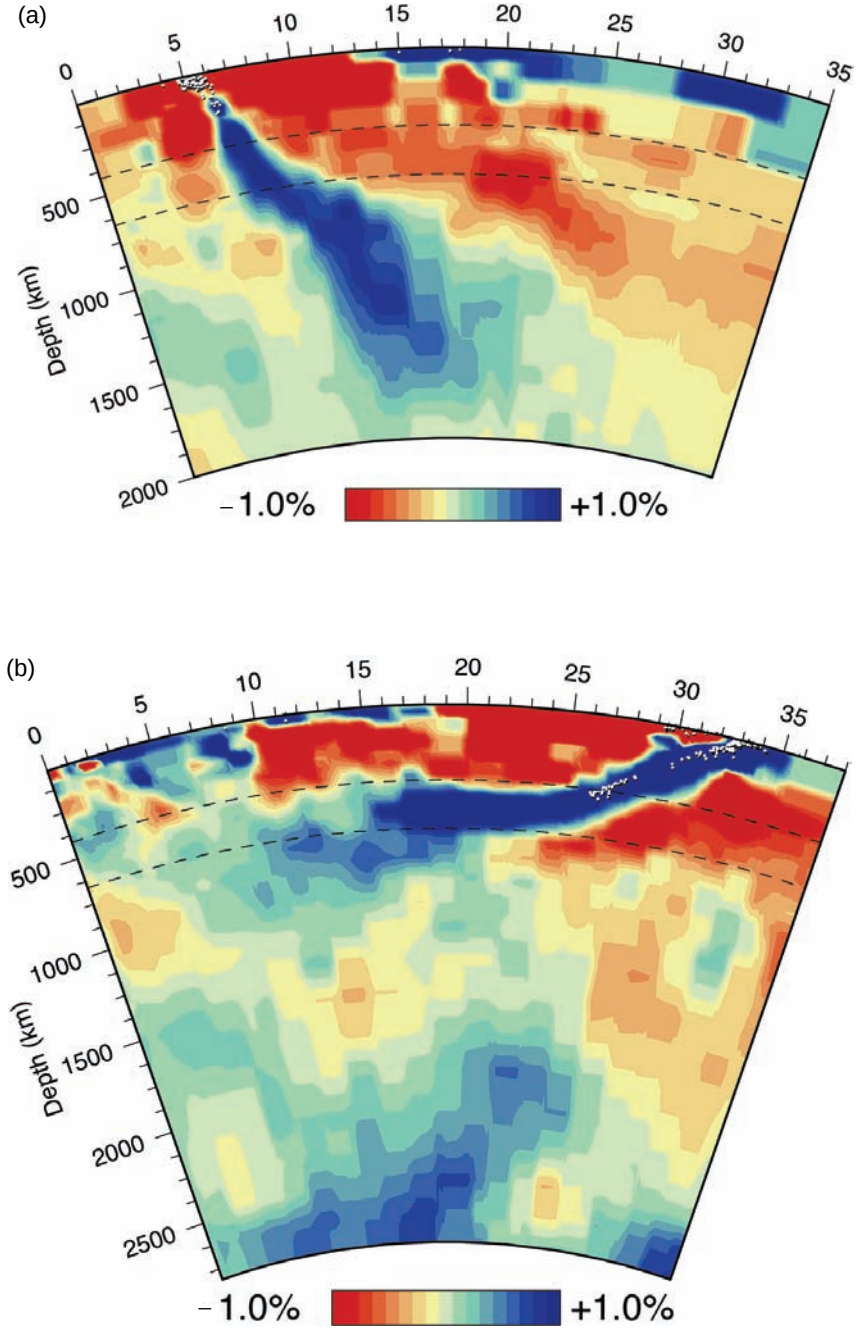
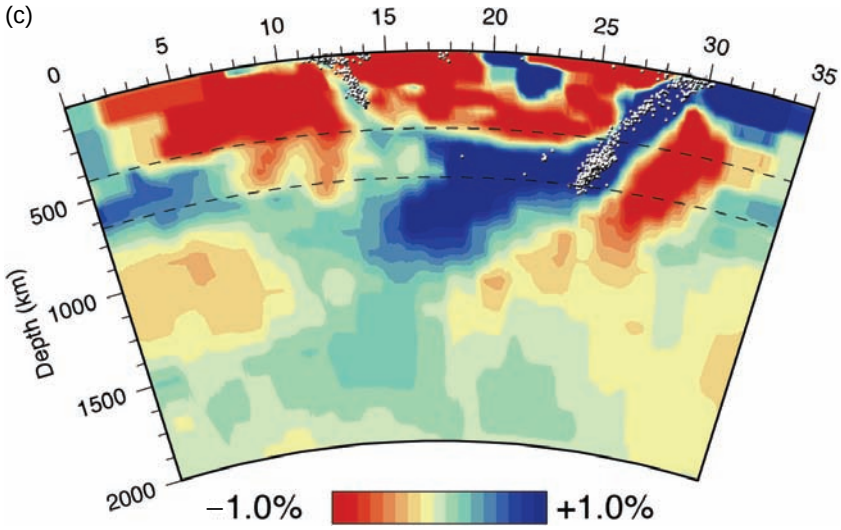


Plate 25. Subduction zones: deviation of seismic velocity from a standard model as determined from P-wave travel times. The dipping blue (fast) structures are interpreted as subducting plates. Location maps shown in Plate 24. (a) Farallon plate, of which only the Juan de Fuca plate remains – the rest has been subducted over the last ~100 Ma. (b) Japan and (c) Tonga. (d) India–Eurasia continental collision. Note the high velocities (blue) of the cold subducting plate. The earthquakes are white circles. Grey-scale version Figs. 9.60 and 10.13 (W. Spakman, personal communication 2003.)



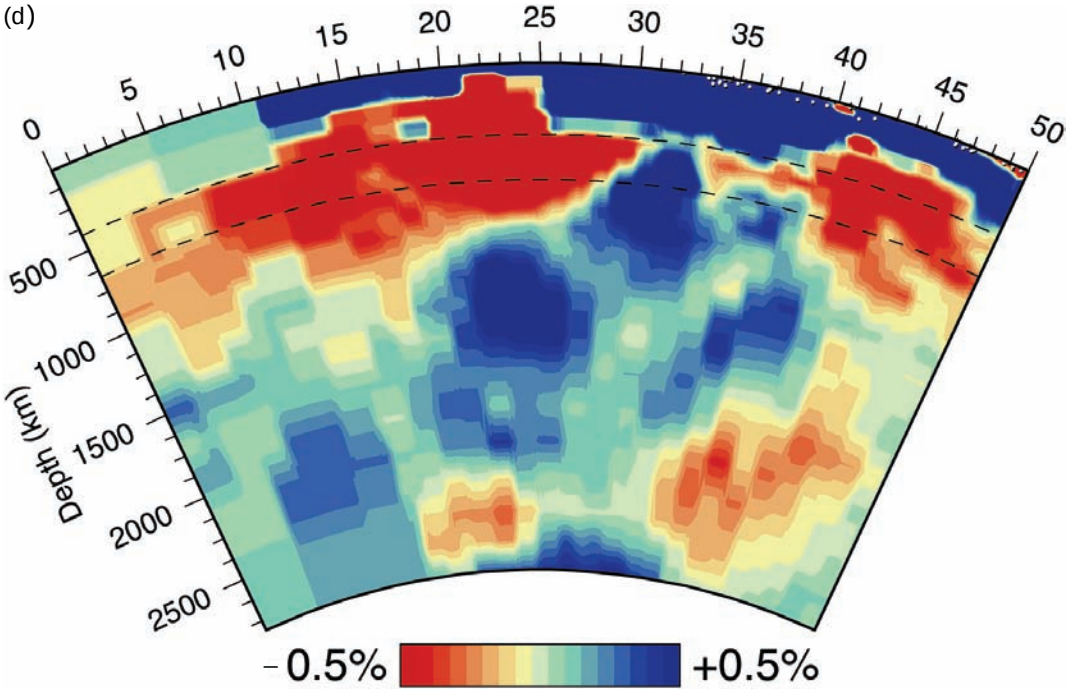
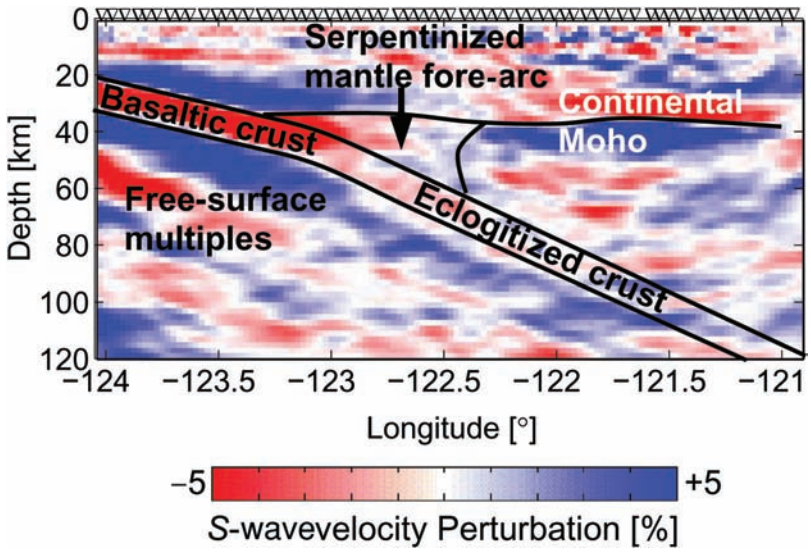


Plate 25. (*cont.*)

Plate 26. Perturbation in S-wave velocity across the Cascadia subduction zone. Red, low S-wave velocities; blue, high S-wave velocities. Grey version Fig. 9.62(c). (From Bostock *et al*. 2002.)



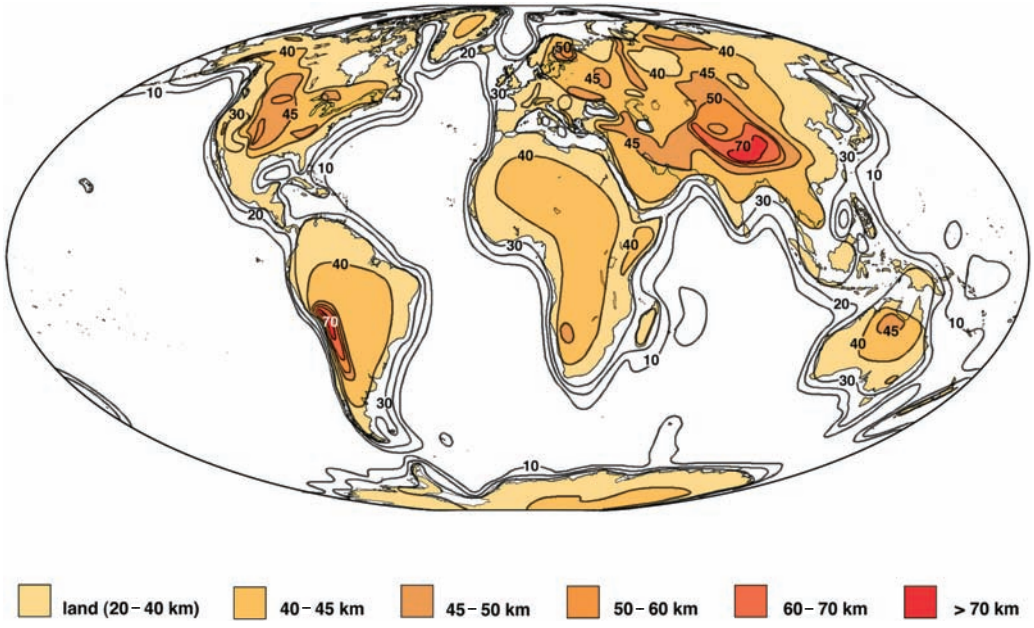


Plate 27. Continental crustal thickness in kilometres, CRUST 5.1. Grey version Fig. 10.2 (From Mooney *et al*., CRUST 5.1: a global crustal model at 5° × 5°, *J. Geophys. Res.*, **103**, 727–47, 1998. Copyright 1998 American Geophysical Union. Reproduced by permission of American Geophysical Union.)

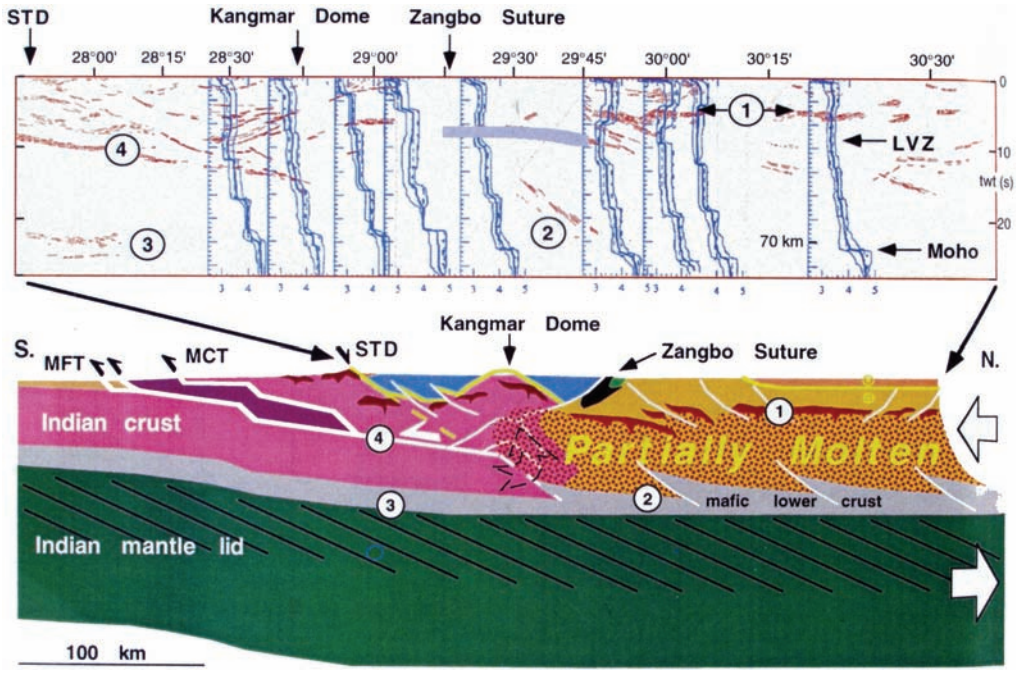


Plate 28. Sections across southern Tibet. S-wave velocity models derived from waveform modelling of broad-band data (blue line with error bars); grey stipple, wide-angle reflection beneath and north of the Zangbo suture; red, reflection data. A schematic interpretation of the India–Eurasia collision zone based on seismic and structural data. Grey version and expanded caption Fig. 10.18. (Reprinted with permission from Brown *et al.* 1996, Bright spots, structure and magmatism in southern Tibet from INDEPTH seismic reflection profiling, *Science*, **274**, 1688–91; and Nelson *et al.* 1996. Partially molten middle crust beneath southern Tibet: synthesis of project INDEPTH results, *Science*, **274**, 1684–7. Copyright 1996 AAAS.)