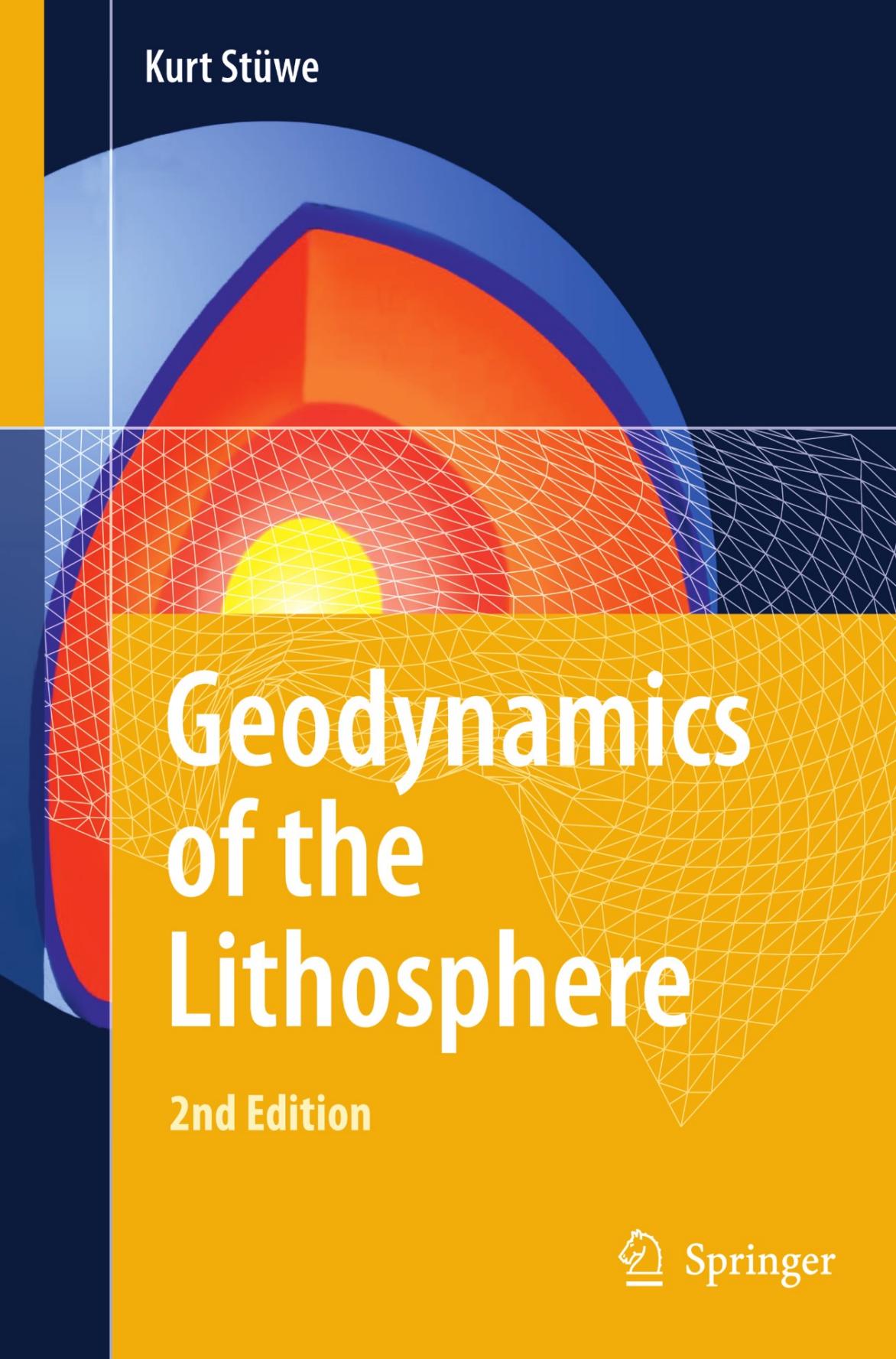


Kurt Stüwe



# Geodynamics of the Lithosphere

2nd Edition

 Springer

Kurt Stüwe

**Geodynamics of the Lithosphere**

An Introduction

Kurt Stüwe

# Geodynamics of the Lithosphere

An Introduction

Second Edition

with 270 Figures, 16 in color

Kurt Stüwe  
University of Graz  
Institute of Earth Science  
Heinrichstr. 26  
8010 Graz  
Austria

*Email:* [kurt.stuewe@uni-graz.at](mailto:kurt.stuewe@uni-graz.at)

Library of Congress Control Number: 2007921988

ISBN 978-3-540-71236-7 Springer Berlin Heidelberg New York  
ISBN 978-3-540-41726-2 (1st. Edition) Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media  
[springer.com](http://springer.com)  
© Springer-Verlag Berlin Heidelberg 2007

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: deblik, Berlin  
Production: Almas Schimmel  
Typesetting: camera-ready by the author

Printed on acid-free paper 30/3141/as 5 4 3 2 1 0

I dislike very much to consider any quantitative problem set by a geologist. In nearly every case, the conditions given are much too vague for the matter to be in any sense satisfactory, and a geologist does not seem to mind a few millions of years in matters relating to time...

John Perry, 1895

(In the paper in which he calculated the age of the earth – previously estimated by Lord Kelvin to be 90 my – to the correct age of 4.5 Gy.)

## Foreword to 2nd edition

I was invited by SPRINGER publishers to write a 2nd edition of this textbook days before departing for the EGU conference in Vienna in spring 2006. Naturally, the idea roamed around my mind whilst sitting through a variety of great talks which presented new results from the field of geodynamics. Many talks presented incredibly realistic results from advanced numerical models and I quickly got the feeling that – in order to remain up to date – a new edition should contain much more on some of these sophisticated models. Quite clearly, geodynamics is a rapidly developing field that makes more and more use of computing skills to cope with an ever growing number of processes that scientists want to consider simultaneously to explain large data sets.

However, as exciting as this development is, it also bears a danger: Shouldn't a good model provide a weighted balance between simplicity and realistic description? Do we really help the field geologists, our undergraduates or even ourselves by producing models that look more and more like nature but that are increasingly difficult to follow intuitively and need graduate knowledge in physics and advanced programming skills to understand?

Personally I feel that the pendulum has swung a bit too far towards the “realistic description” side. What use is a model that extracts an awed gasp from the impressed viewer, but fails to provide an intuitive feel what is going on or – worse still – might fail on first principles behind the curtain of colorful animations? I felt myself confirmed when I sat in a talk where 2nd order morphological features of a drainage system were modeled numerically with fantastic reality; a talk that was followed by a discussion in which the 1st order features and their implementation were hotly debated.

I have therefore refrained from diverging too far from the simplistic style of text, maths and figures of the 1st edition. New sections are written in the same vein as in the 1st edition and in fact, some pages I deemed too complicated have been omitted. I still feel that – even in today's sophisticated world – there is room for a simplistic view that reminds us of the basics and that helps the quantitatively inclined field geologist to enter the wonderful world of geodynamic modeling.

## Preamble to 1st edition

Field geologists typically interpret their data in terms of tectonic models that are consistent with their observations in a given terrain, but that often lack an independent test. Such models can be strengthened considerably if they are supported by independent estimates of the magnitude of the implied geodynamic processes. For example, estimates of the thermal energy budget of a metamorphic terrain are an invaluable aid for the tectonic interpretation of metamorphic isograds mapped in the field; estimates for the orogenic force balance of a certain nappe staking geometry interpreted from structural mapping are a fantastic way to test its mechanical plausibility.

This book was written because there appears to be a strong bi-modality in the nature of text books dealing with such problems. Books that introduce the reader to the modeling of geodynamic processes often require a relatively high maths background. On the other hand, books that deal with basic maths usually lack any connection to geology. This book was written in an attempt to bridge this gap. It is the aim of this book to introduce field based geologists to the power of the quantitative treatment of their field data. Because of this, the emphasis of this book lies on the interpretation of data that are typically collected by structural geologists, petrologists and geochronologists in the field, rather than those collected by seismologists or geophysicists.

As an introductory text, little mathematical knowledge is required. All calculations are discussed in detail without omitting steps in the derivations and an extensive appendix on mathematical tools is provided. All computer codes used to calculate the figures are available from the author. They may also be downloaded from the address <http://wegener.uni-graz.at>.

# Contents

<b>1. Introduction</b>	1
1.1 What is a Model?	1
1.2 Spatial Dimension of Geological Problems	6
1.2.1 Reducing Spatial Dimensions	10
<b>2. Plate Tectonics</b>	15
2.1 Historical Development	15
2.2 Working on a Spherical Surface	20
2.2.1 ... or is the Earth Flat After All?	21
2.2.2 Geometry on a Sphere	23
2.2.3 Kinematics on a Sphere	25
2.2.4 Mechanics on a Sphere	25
2.3 Map Projections	28
2.3.1 The Mercator Projection	29
2.3.2 National Grids	30
2.4 The Layered Structure of Earth	33
2.4.1 Crust and Lithosphere	35
2.4.2 The Lithospheric Plates	39
2.4.3 The Plate Boundaries	39
2.4.4 The Wilson Cycle	46
2.5 Problems	47
<b>3. Energetics: Heat and Temperature</b>	51
3.1 Principles of Heat Conduction	52
3.1.1 The Heat Conduction Equation	52
3.1.2 The Laplace Equation	60
3.1.3 The Error Function	61
3.1.4 Time Scales of Diffusion	62
3.2 Principles of Heat Production	64
3.2.1 Radioactive Heat Production	66
3.2.2 Mechanical Heat Production	68
3.2.3 Chemical Heat Production	76
3.3 Principles of Heat Advection	80
3.3.1 Heat Advection by Magma	81

3.3.2	Heat Advection by Solid Rock .....	82
3.3.3	Heat Advection by Fluids .....	85
3.3.4	The Peclet Number .....	86
3.4	Heat in the Continental Lithosphere .....	87
3.4.1	Stable Geotherms: The Relevant Equation .....	90
3.4.2	The Contribution of Radioactivity .....	92
3.4.3	Geotherms with Heat Flow Boundary Condition .....	93
3.4.4	Geotherms with Fixed Basal Boundary Condition .....	98
3.5	Heat in the Oceanic Lithosphere .....	101
3.5.1	Aging Oceanic Lithosphere .....	102
3.5.2	Subduction Zones .....	106
3.6	Thermal Evolution of Intrusions .....	109
3.6.1	Step-shaped Temperature Distributions .....	110
3.6.2	One-dimensional Intrusions .....	112
3.6.3	Two-dimensional Intrusions .....	118
3.6.4	Modeling Realistic Intrusions .....	120
3.7	Selected Heat Transfer Problems .....	125
3.7.1	Periodic Temperature Fluctuations .....	125
3.7.2	Folded Isotherms .....	126
3.7.3	Isotherms and Surface Topography .....	128
3.7.4	Temperature Distribution Around Faults .....	132
3.8	Problems .....	134
4.	<b>Kinematics: Morphology and Deformation .....</b>	139
4.1	Strain - The Basics .....	139
4.2	Reference Levels .....	142
4.2.1	Sphere, Spheroid and Geoid .....	143
4.2.2	Lagrangian and Eulerian Reference Frames .....	144
4.2.3	The Undeformed Reference Lithosphere .....	147
4.3	Uplift and Exhumation .....	149
4.3.1	Definition of Uplift and Exhumation .....	149
4.3.2	Vertical Motion in Active Orogens .....	152
4.3.3	Exhumation Processes .....	153
4.3.4	Modeling Vertical Kinematics .....	154
4.4	Isostasy .....	160
4.4.1	Hydrostatic Isostasy .....	161
4.4.2	Flexural Isostasy .....	170
4.5	Geomorphology .....	178
4.5.1	Erosion Models on Orogenic Scale .....	181
4.5.2	Short Range Transport .....	184
4.5.3	Long Range Transport: Drainages .....	193
4.5.4	Discontinuous Landscape Formation .....	200
4.5.5	The Shape of Volcanoes .....	200
4.5.6	Fractals .....	202

4.6	Problems . . . . .	205
<b>5.</b>	<b>Mechanics: Force and Rheology . . . . .</b>	<b>209</b>
5.1	Stress - The Basics . . . . .	210
5.1.1	The Stress Tensor . . . . .	210
5.1.2	Deformation Laws . . . . .	220
5.2	Rheology of the Lithosphere . . . . .	233
5.2.1	Rheology of the Continental Lithosphere . . . . .	234
5.2.2	Rheology of the Oceanic Lithosphere . . . . .	242
5.3	Forces Applied to Lithospheric Plates . . . . .	245
5.3.1	Transmission Mechanisms . . . . .	245
5.3.2	Forces in Oceanic Lithosphere . . . . .	253
5.3.3	Forces in Continental Plates . . . . .	257
5.4	Problems . . . . .	260
<b>6.</b>	<b>Dynamic Processes . . . . .</b>	<b>265</b>
6.1	Continents in Extension . . . . .	265
6.1.1	Basin Subsidence Mechanisms . . . . .	268
6.1.2	Basin Types . . . . .	269
6.1.3	Subsidence Analysis . . . . .	271
6.1.4	Models of Continental Extension . . . . .	276
6.2	Continents in Collision . . . . .	284
6.2.1	Thermal Evolution of Collisional Orogens . . . . .	285
6.2.2	Mechanical Description of Colliding Continents . . . . .	294
6.2.3	Orogenic Wedges . . . . .	308
6.3	Selected Geodynamic Processes . . . . .	315
6.3.1	Flood Basalts and Mantle Plumes . . . . .	315
6.3.2	Delamination of the Mantle Lithosphere . . . . .	318
6.3.3	Low Pressure - High Temperature Metamorphism . . . . .	321
6.3.4	High Pressure Metamorphism . . . . .	323
6.3.5	Tectonic Overpressure . . . . .	326
6.3.6	Feedback and Episodicity . . . . .	329
6.4	Problems . . . . .	334
<b>7.</b>	<b>Metamorphic Processes . . . . .</b>	<b>337</b>
7.1	Introduction . . . . .	337
7.1.1	What Exactly are <i>P-T</i> - and <i>P-T-t-D</i> -Paths? . . . . .	338
7.1.2	The Metamorphic Facies . . . . .	339
7.2	Basic Principles of Petrology . . . . .	342
7.2.1	Equilibrium Information: Thermobarometry . . . . .	346
7.2.2	Non-equilibrium Information: Kinetics . . . . .	349
7.3	Classification of <i>P-T</i> -Paths . . . . .	353
7.3.1	Slope and Curvature of <i>P-T</i> -Paths . . . . .	355
7.4	Interpretation of <i>P-T-t-D</i> -Relationships in Orogens . . . . .	357
7.4.1	Interpreting Temporal Relationships . . . . .	359

7.4.2 Interpreting Spatial Relationships . . . . .	360
7.5 Problems . . . . .	364
<b>A. Topographic Maps . . . . .</b>	<b>369</b>
<b>B. Mathematical Tools . . . . .</b>	<b>393</b>
B.1 What is a Differential Equation? . . . . .	393
B.1.1 Terminology Used in Differential Calculus . . . . .	395
B.2 The Finite Difference Method . . . . .	399
B.2.1 Stability and Accuracy . . . . .	401
B.2.2 Implicit and Explicit Finite Difference Methods . . . . .	402
B.2.3 Approximation of the Transport Equation . . . . .	404
B.2.4 Grids and Boundary Conditions . . . . .	407
B.2.5 Dealing with Irregular Grid Boundaries . . . . .	408
B.2.6 Recommended Reading . . . . .	409
B.3 Scalars, Vectors and Tensors . . . . .	410
B.4 Using Fourier Series . . . . .	413
B.5 Selected Numerical Tricks . . . . .	416
B.5.1 Integrating Differential Equations . . . . .	416
B.5.2 Analytically Unsolvable Equations . . . . .	416
B.5.3 The Least Squares Method . . . . .	417
B.5.4 Basic Statistical Parameters . . . . .	419
B.6 Problems . . . . .	421
<b>C. Maths Refresher . . . . .</b>	<b>423</b>
<b>D. Symbols and Units . . . . .</b>	<b>429</b>
<b>E. Answers to Problems . . . . .</b>	<b>439</b>
<b>F. Internet Addresses . . . . .</b>	<b>455</b>
<b>References . . . . .</b>	<b>463</b>
<b>Index . . . . .</b>	<b>485</b>

**Table 0.1.** List of Maps in this book

Maps in text	Figure no.	page no.
Topographic highs in the world	Fig. 1.1	p. 3
Sea floor topography (global)	Fig. 2.1	p. 16
Intraplate stress field (global)	Fig. 2.2	p. 17
Palaeogeographic map (global reconstruct.)	Fig. 2.3	p. 18
Age of the ocean floor (global)	Fig. 2.4	p. 19
Seismicity (global)	Fig. 2.5	p. 21
Pacific Rim (slab contours)	Fig. 2.6	p. 22
Map projections (global)	Fig. 2.12	p. 31
Plate tectonic division (global)	Fig. 2.20	p. 40
San Andreas Fault (Dragonsback)	Fig. 2.24	p. 44
African Rift System (cartoon)	Fig. 2.26	p. 45
Heat flow map (global)	Fig. 3.1	p. 53
Strain rate map (global)	Fig. 4.2	p. 142
Geoid map (global)	Fig. 4.5	p. 145
South African Great Escarpment	Fig. 4.26	p. 178
Styrian Basin (Neotectonics)	Fig. 4.28	p. 180
Precipitation (global)	Fig. 4.29	p. 181
Chinese Loess Plateau	Fig. 4.46	p. 203
Sediment Thickness (global)	Fig. 6.3	p. 267
Altiplano and Tibetan Plateau (fault plane sol.)	Fig. 6.25	p. 302
Large Igneous Provinces (global)	Fig. 6.34	p. 316
Barrow and Buchan Zones of Scotland	Fig. 7.2	p. 342

**Atlas Section**

Map key	Fig. A.1	p. 369
The India - Asia collision zone	Fig. A.2	p. 376
South America	Fig. A.3	p. 377
Europe and the Mediterranean Sea	Fig. A.4	p. 378
Arabian Peninsula and Northern Africa	Fig. A.5	p. 379
Southern Africa	Fig. A.6	p. 380
North Atlantic	Fig. A.7	p. 381
Eastern North America	Fig. A.8	p. 382
Western North America	Fig. A.9	p. 383
Western Pacific Ring of Fire	Fig. A.10	p. 384
Eastern Pacific	Fig. A.11	p. 385
Indian Ocean	Fig. A.12	p. 386
Australia	Fig. A.13	p. 387
New Zealand and South Pacific	Fig. A.14	p. 388
Scotia Plate and South America	Fig. A.15	p. 389
The Arctic	Fig. A.16	p. 390
The Antarctic	Fig. A.17	p. 391

## Acknowledgements to 2nd edition

This book is a reflection of my enthusiasm for the quantitative treatment of geological problems. However, I am originally a field geologist and over the years I had to rely on the help of many who are more numerically literate than myself to teach me the wonders of quantitative geodynamics. Many of those are acknowledged in the first edition of this book. Of those who helped me with this 2nd edition, my special thanks go to S Hergarten and J Robl. Stefan's background in mathematics has been a continuous source of enlightenment for me and his patience to teach me over many lunches has helped to make the physics in this edition much more soundly founded – certainly in my mind and hopefully visibly on these pages. Jörg Robl helped me with the handling of the digital data sets that form the basis for many of the figures. As a former student of mine, his art in handling GMT, awk and bash scripts has now long outgrown my own expertise. Many other colleagues and friends have helped to improve this 2nd edition with help that ranged from a brief email to answer a specific question to providing photos or proof reading entire sections. In particular I would like to thank: JP Avouac, M Bevis, M Brown, A Camacho, D Coblenz, L Estey, T Gerya, W Gosnold, B Grasemann, E Hartz, B Heberer, G Houseman, K Hudnut, N Mancktelow, S Matthai, F Neubauer, H Pollack, R Powell, G Schubert, B Tucholke and S Willett. It was a joy to find you all so open and helpful when I asked for assistance.

Sadly, the number of mistakes  $F_2$  in a book – even in its 2nd edition – will not be zero. It follows roughly the relationship:  $F_2 = F_1 e^{-xt} + St$ , where  $F_1$  is the number of mistakes in the first edition,  $t$  is correction time during writing of the 2nd edition,  $x$  is an unknown decay constant and  $S$  is a source term which describes the rate of appearance of new errors during the writing of new text. I now know that  $F_1$  was about 200. The time  $t$  I have taken to work on the 2nd edition is about 1 year. The magnitude of  $x$  and the source term remains unknown. I hope that they are about:  $x = 1 \times 10^{-7} \text{ s}^{-1}$  and  $S = 6 \times 10^{-7} \text{ s}^{-1}$ . I leave it as the first exercise for the interested reader to determine the number of remaining mistakes. Clearly, these are my own fault. Some of them will be obvious to me within hours of sending this manuscript into print, about others I invite observant readers to inform me. As an excuse I can only say that I have written this book as a field geologist and not as a theoretical geophysicist. I therefore hope, that the advantages of a text book written on the basis of many years of field observation outweigh the disadvantages of potential mathematical shortcomings. Good luck!

December 2006 Kurt Stüwe

# 1. Introduction

The large scale structure of the earth is caused by *geodynamic processes* which are explained using *energetic*, *kinematic* and *dynamic* descriptions. While “geodynamic processes” are understood to include a large variety of *processes* and the term is used quite loosely, the methods of their *description* involve well defined fields. *Energetic* descriptions are involved with distribution of energy in our planet, typically expressed in terms of heat and temperature. *Kinematic* descriptions describe movements using velocities, strains and strain rates. *Dynamic* descriptions indicate how stresses and forces behave.

In the field, we document only the *consequences* of geological processes. The underlying *causes* are much harder to constrain directly. Nevertheless, if we want to explain the tectonic evolution of our planet, we need to interpret these causes or: “*driving forces*”. For this, we need to find a *dynamic description* of geological processes that is consistent with our observations. Our descriptions relate *causes* and *consequences* – *tectonic processes* with *field observations*. In many cases, we will use equations as a concise form to describe processes and observations in nature. As we will be dealing mostly with large scale tectonic questions, the observations that we shall use are also on a large scale. For example, we shall use observations on the elevation (Fig. 1.1) and heat flow of mountain ranges, the thickness of continents and the water depth of the oceans.

As the processes we seek to describe are *changing* with time or space, many of the descriptive equations will be *differential equations*. However, it is *not* the aim of this book to confuse the reader with high level mathematics. Rather it is the aim to introduce field geologists to the beauty and simplicity of descriptions with equations. Thus, all equations will be explained from basic principles so that an intuitive understanding can be achieved. While I prefer the term “*description*”, many colleagues would refer to the contents of this book as “*modeling*”. Thus, it appears useful to begin this book with a discussion what we mean when we use the word “*model*”.

## 1.1 What is a Model?

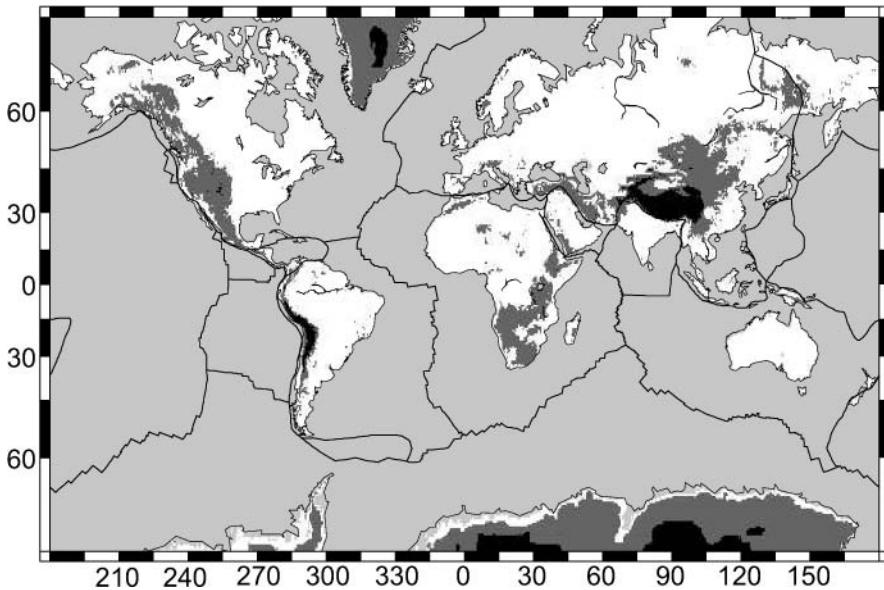
Models are tools that we use to describe the world around us in a simplified way so that we can understand it better. Sadly, many geologists misunder-

stand the word model and think of it only as something complicatedly mathematical that has little to do with field work (s. Greenwood 1989). In doing so, many field geologists overlook the fact that field mapping itself is also a form of modelling:

**Geological Mapping: A Typical Modeling Exercise.** A map is a transformation of reference frames; for example the projection of the geographic position of field locations onto a piece of paper. However, a geological map that is a mere representation of field data in a new reference frame (i.e. our piece of paper) may be a good map but is usually a poor *geological* map (Passchier et al. 1990). It may still be useful to find a given outcrop of a given lithology, structure or metamorphic grade, but as scientists we are usually unsatisfied by simply documenting what is there. Usually we want to go beyond that and map to clarify field relationships, for example when we infer where a lithological contact is underneath a meadow. We make maps so that we can explain some features of nature to a colleague geologist without him or her having to do the same amount of field work we invested to produce the map in the first place. In order to achieve this aim, the geological map must illustrate field relationships in a *simplified* and *interpreted* manner. This forces the field geologist to a constant decision-making process. First of all the geologist has to decide *what* is to be mapped. Is it topography? Is it structure? Is it metamorphic isograds or is it lithology? Which of these (and many others) is to be mapped depends on the question with which we go into the field (Fig. 1.1). Then, the geologist has to decide on the *scale* on which the map should be produced. This decision is not trivial! The scale of the map depends on the problem to be solved. Once the scale is decided upon many more decisions are to be made. Which observations are too small to be mapped and should be neglected? Which ones should be drawn into the map? Which ones are to be emphasized by lines? Can a contact seen in two outcrops be mapped as a line, even in the paddock separating the two outcrops? The geologist is modeling!

If the map is good, then it helps the reader - like any other good model - to understand nature quickly and easily. It also helps to make predictions how the geology may look at different places that were not mapped yet. For example, constructing profiles across our map helps us (to a certain degree) to explain how the geology looks underground. In numerical-, analogue-, conceptual- or thought-models this process is identical. Mathematical models consist of a series of rules that determine which observations in nature are to be neglected and which ones are to be emphasized. The former will *not* appear in an equation, the latter *will* appear as a parameter in the equation. As such, a mathematical model is no different from the field work of a geologist.

Every model can be considered as a tool that can be used to make predictions about observations in nature. Just as a geological map can be used to construct cross sections and thus predict the geology underneath the surface, a numerical model can be used to make predictions about temperatures,



**Figure 1.1.** Topographic map of the continents showing the large mountain ranges which always have inspired geologists to geodynamic models explaining their origin. Elevations higher than 1 000 m above sea level are shown as grey, elevations higher than 3 000 m are shown in black. The map bears all the characteristics of a good model. It captures the essence of the planet's distribution of high mountain ranges within only about  $0.01 \text{ m}^2$  of this page, which is only possible by making a range of simplifying assumptions, for example on resolution or choice of cut-off elevation. As such, elevations above 3 000 m in the Azores or Japan escape the scale. Also – while the model largely succeeds in its aim – it is also “wrong” in some places. For example, much of the land surface of (ice covered) Greenland and Antarctica is actually below sea level

forces or velocities which cannot be observed directly because of their enormous time scale or depth.

If the choice of parameters that we consider in our model (and the rules that relate them to each other) are good, then our model is good and it will predict many new observations which will be proven to be correct by future observations. If our choice of parameters and rules is bad, then our model may explain the one or other field observations, but it will predict many other features that will be proven wrong by future observations. Modeling is therefore an iterative back and forth between the choice of parameters and rules that are to be considered or neglected, new observations in nature and finally improvement of the model based on the new observations. Good models are *consistent* with a large number of observations (s. p. 5), but models are hardly ever *unique* in fitting those observations.

**Consistent versus Unique.** The difference between consistent and unique is an important one that is often not recognized, even by modelers them-

selves. *Unique* means - as the word says - that the model is the very only explanation for a set of observations. *Consistent* means that the proposed model *does* explain a set of observations, but that other models may also explain the same set of observations. The largest majority of models are consistent but by no means unique. For example, the heliocentric Copernican model for the solar system (which states that the planets rotate around the sun as initially suggested by Aristarchus of Samos 310 - 230 B.C., but only brought to wide acceptance by Copernicus in the early 16th century) is a model which is consistent with our observations of when the sun rises and sets and so on. However, a *geocentric* model (initially designed by Ptolemy) in which the planets and the sun rotate around the earth is possible too. The geocentric model is amazingly more complicated than the heliocentric model and it involves weird planetary motions including epicycles and cycloid curves. However, it also is consistent with our observations on planetary motions. Neither the heliocentric, nor the geocentric model are therefore unique. When viciously defending a model in a discussion it is always sobering to remind oneself that practically all models are only consistent (at best).

**Good versus Accurate.** The difference between *good* (or possibly: “adequate”) and *accurate* models is related to the difference between consistent and unique models, but it is not quite the same. Here, it is important to realize that the *best* model must not be the most *accurate* model! The best model is the one that finds the best balance between accurate description and simplicity. A good example for this is given by a comparison of Newton’s law of gravitation and relativity theory when applied to the description of planetary motion (e.g. Hawkins 1988). Newtonian theory states that the gravitational attraction,  $F$ , between two masses is directly proportional to the masses of the two bodies  $m_1$  and  $m_2$ , and inverse proportional to the square of the distance  $r$  between the two bodies. This model is incredibly simple and may be described by a simple equation:

$$F = G \frac{m_1 m_2}{r^2} . \quad (1.1)$$

The constant of proportionality is called the gravitational constant,  $G$ . Its value may be found in Table D.4 in the Appendix of this book. This model describes the elliptic motions of the planets (that were discovered by Kepler in order to improve the Copernican model) extremely well. However, very detailed measurements early this century showed that the motions of some planets differ a bit from those described by eq. 1.1. These differences may be explained with the model of general relativity, which describes the planetary motions more exactly than Newton’s law. Indeed, general relativity shows that many concepts of Newtonian physics are “wrong”, for example the constancy of mass (general relativity states that mass depends on speed). Thus, one might consider Newton’s model to be superseded by general relativity and use this new model from now on.

However, general relativity is much harder to grasp than the intuitive understanding of eq. 1.1. It is therefore often not very practical to use. In fact, for the largest majority of purposes – for example to find a planet with a telescope in the sky, or for the interpretation of gravity anomalies by a geologist – Newton's model (eq. 1.1) is sufficient. Thus, for most purposes Newton's model is better (because simpler), albeit less accurate. In short, a good model should find a good balance between *simplicity* of the model and *accuracy* in describing a set of observations and this balance depends on the problem to be solved.

**Accuracy versus Precision.** In the last paragraph we have used the word *correct* to describe a very good correlation between model description and observation in nature. In general this is the same what is meant by the word *accurate*. However, *precision* is something different. Precision describes how good a model or an experiment can be repeated with the same result (s. p. 419). Let us illustrate this with an example. A radiochemical analysis may indicate that a rock formed 100 my ago. This analysis is very precise if every time we perform it, we arrive at the same age of 100 my. This applies to errors as well. The analysis is still called very precise if we come up with an answer of  $100 \pm 50$  my, if that answer is reproducible with the identical error limits and we know these error bars very well. However, the radiometric age above is not very accurate. In fact, even analyses with very small analytical errors may be not very accurate at all. It could be that a precise but inaccurate age of  $100 \pm 0.1$  my was obtained for a rock that actually formed 150 my ago, which in itself bears important information, for example, why the chosen radiometric system re-equilibrated 50 my later.

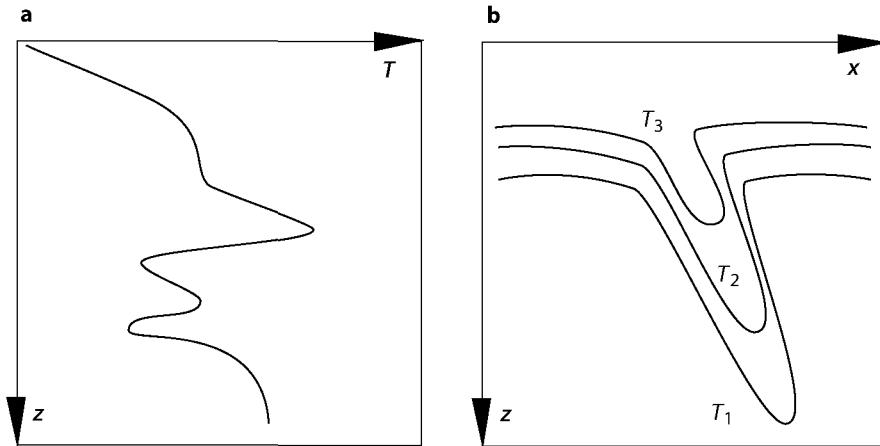
In conclusion of this section, let us define a good model. A model is a description of nature that has the three following properties:

- It should describe a *large* set of observations with a comparably *small* set of parameters.
- A good model must be useable as a tool to make predictions about fact that have not been observed yet.
- It must be possible to test a good model by making new experiments or observations.

Note that none of these three requirements includes *accuracy*. The deciding factor for a good model remains the balance between accuracy and simplicity. Accurate description of nature is a virtue that remains reserved to (explanation-free) collections of measured accurate data. (Although, remember that many measured data may not be accurate either.)

Geodynamics describes the dynamic evolution of earth through space and time. This dynamic evolution occurs on time scales of up to many hundreds of million years and spatial scales of up to thousands of kilometers. Direct observation is therefore often difficult. Geodynamics is therefore a science which

relies to a much larger extent on the model tool than many other sciences. Mapping in the field, analogue modeling in the laboratory and programming on the computer are all three modeling techniques that are of equal importance in this process. Integrated use of various modeling techniques is the most elegant way to arrive efficiently at a good description of the nature around us.



**Figure 1.2.** Examples of illustrations of one- and two-dimensional models. **a** shows an example of temperature evaluated as a function of a single spatial coordinate,  $z$ . The axes of the surface of this page are used up for this one-dimensional description. The strange geotherm is uniquely defined: Every depth corresponds uniquely to a single temperature. **b** shows a two-dimensional model, for example temperatures in a subduction zone as a function of depth  $z$  and horizontal distance  $x$ . The surface of this page is used up for the two spatial dimensions of the model and the evaluated variable must be portrayed by the contours

## 1.2 Spatial Dimension of Geological Problems

When a geological process is to be described by a simple model, one of the first decisions that has to be made is often the number of spatial dimensions that are to be considered. This decision has to be made according to the requirements of a good model, as discussed above. That is, it should be tried to describe the problem in question with as few spatial dimensions as possible, without loosing the essence of the problem. Fortunately, a very large number of geological problems can be considered one-dimensionally or even without any spatial dimension. Before we go on to discuss some problems that require one-, two-, or three-dimensional consideration, it should be emphasized that “dimension” need not always be the “spatial dimension”. For many problems

it is important to consider time or some other independent variable as the dimension which may then become an axis on a diagram which we use to illustrate some feature of our description. In this book we will stick to (unless otherwise noted) the meaning of a *spatial* dimension when we talk of the dimension of a problem and use the SI *units*, when we use another variable as our “dimension”.

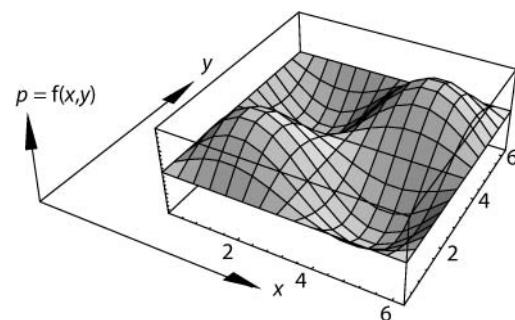
**One-dimensional Problems.** A simple example of a one-dimensional problem is the description of temperature in the lithosphere as a function of depth (Fig. 1.2a). A one-dimensional description of this problem is sufficient when the lateral extent of the plate is large compared to the thickness of the lithosphere and there is no other lateral changes in physical parameters or structure. Continental lithosphere is often of the order of 100 km thick, but continents are usually many hundreds or even thousands of kilometers in lateral extent. The temperature as a function of depth can therefore be well-illustrated in a diagram in which one axis denotes depth and the other temperature. The depth axis is the dimension of the problem, the temperature axis is the evaluated variable. Similar logic may be applied to the description of temperatures around magmatic dikes.

One-dimensional problems need not be in Cartesian coordinates, but can also occur in spherical or cylindrical coordinates, for example temperatures around a spherical intrusion or compositional zoning profiles in a cylindrical crystal. One-dimensional descriptions are useful when there is little or no variation of the variable of interest in the other spatial directions.

**Two-dimensional Problems.** An example of a typical two-dimensional problem is the temperature distribution in *subduction zones*. There, the subduction angle as well as the plate thickness are both critical for the shape of the isotherms and a vertical ( $z$ ) *plus* a horizontal spatial coordinate ( $x$ ) are therefore needed for a meaningful description (Fig. 1.2b). Fortunately, the third spatial dimension – the direction parallel to the trench – can often be neglected because trenches are usually long and the subduction direction is often at roughly right angles to the trench. Calculated temperatures can be illustrated as contours on a diagram with two spatial dimension of the axes.

**Figure 1.3.** A

“three-dimensional” illustration of the two-dimensional function  $p = \sin(x) \times \sin(y)$  shown on the two-dimensional paper of this book page

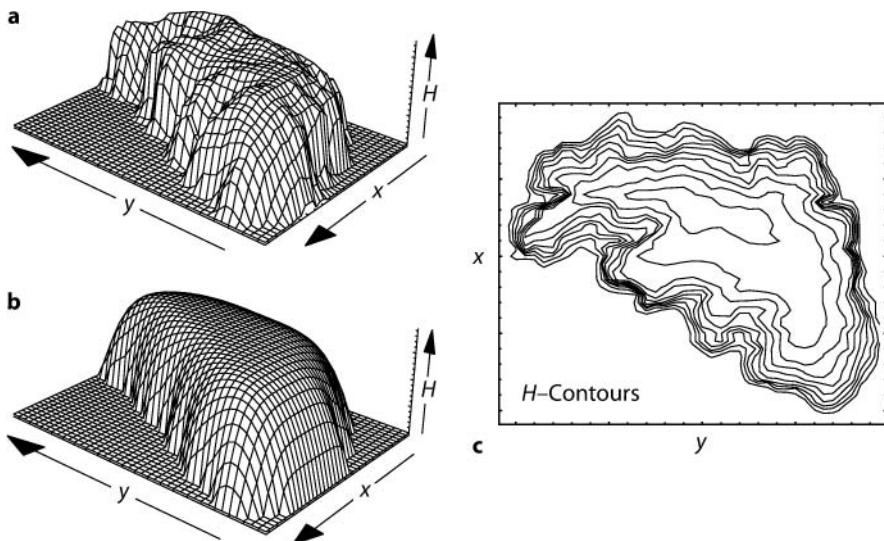


The same two-dimensional problem can also be illustrated by using a perspective sketch where the two spatial coordinates *and* the temperature axis are all shown in one diagram. Fig. 1.3 illustrates such a “three-dimensional” illustration of a two-dimensional model using the abstract function  $f(x, y) = \sin(x) \times \sin(y)$  as an example. “Three-dimensional” is shown here in quotation marks, firstly because the third dimension ( $p$ , or  $T$  in Fig. 1.2b) is not really a spatial dimension but the evaluated variable, and secondly because a perspective sketch is still drawn on the two dimensions of this page. Thus, such drawings remain two-dimensional on paper (unless they are built of wire or something else) and they only appeal to the three dimensional imagination of the reader. For two-dimensional models that are used to evaluate temperatures “three-dimensional” illustrations are not common practice, but for geomorphological models and many geophysical questions “three-dimensional” illustrations of two-dimensional model results are quite instructive and are often used.

However, be careful! The fact that two-dimensional models are often shown as surfaces over a grid of two spatial coordinates should not be mistaken for a real three-dimensional model. The third dimension is only the evaluated variable! For Fig. 1.2b we could write  $T=f(x, z)$ : temperature is a function of  $x$  and  $z$  only. In landscape models (e. g. Fig. 1.4) this is often quite confusing as the evaluated variable (surface elevation) has the same units (meters) as the two spatial coordinates on which it was evaluated. To stop confusion it is often useful to use the expression *potential surface*. Just like geophysicists evaluate gravitational- or electromagnetic- potential as a function of two spatial coordinates, so can surface elevation or temperature be viewed as a *potential surface* overlying the plane defined by the two spatial coordinates.

**Three-dimensional Problems.** Three-dimensional models are not only very difficult to design, their results are also hard to understand, to explain or to show graphically. For many geological problems – for example those that involve the modeling of stress or strain – three-dimensional models involve tensors algebra, which makes it often hard to follow their results intuitively. Three-dimensional models should therefore only be used if the problem that is to be solved cannot be simplified in its spatial dimensions (s. p. 7). Examples of important geological problems that are inherently three-dimensional are mantle convection or oblique subduction. Such problems are only solvable with three dimensional models and brave earth scientists use modern methods of calculation and illustration to tackle such difficult problems (e. g. Braun and Beaumont 1995; Platt 1993a). Practically, three-dimensional model results can only be illustrated on a computer screen or with series of contoured diagrams or videos.

**Zero-dimensional Problems.** A very large number of geological problems can be solved quite elegantly without considering any spatial dimensions. For example, the surface elevation of mountain belts in isostatic equilibrium or the influence of heat production on temperature of rocks can be estimated



**Figure 1.4.** **a, b** “Three-dimensional” illustration of a two-dimensional model shown on the two dimensions of this page. **c** Two-dimensional illustration of the same model results. The coordinates  $x$  and  $y$  are the spatial coordinates of the model and  $H$  is the evaluated variable - or potential surface. In other models the vertical axis may be the gravitational potential (in  $\text{m s}^{-2}$ ), temperature of a metamorphic terrain or the concentration of an element in a crystal. However, in the shown example,  $H$  is of the same units as the spatial coordinates - the model is a landscape model. **a** and **c** are illustrations of field data of the shape of Ayers Rock in Central Australia and **b** shows the modeled shape according to Stüwe (1994)

without spatial considerations and can still give use enormous insight into the nature of many tectonic processes. When neglecting all spatial dependence of a problem, this gives us the great freedom to evaluate the model results graphically as a function of two variables simultaneously. Both axes that may be drawn on paper can be used to evaluate the influence of two different variables against each other.

Zero-dimensional problems should not be confused with *dimensionless variables* or *normalized variables*. In many problems of this book we will encounter variables that are normalized to some standard value and therefore have no units. This is called a *dimensionless variable*. Dimensionless variables are great to get a feeling what a certain number means in relationship to some better known value. For example, the elevation of Mt Blanc,  $H_{(\text{MtBlanc})}$ , can be either given in meters, or it could be described in terms of its proportional elevation  $h$  (in %) relative to, say, Mt Everest:

$$h = \frac{H_{(\text{MtBlanc})}}{H_{(\text{MtEverest})}} = \frac{4807 \text{ m}}{8848 \text{ m}} = 0.543 \quad . \quad (1.2)$$

$h = 0.543$  is the dimensionless elevation of Mt Blanc relative to Mt Everest. The use of the dimensionless elevation in eq. 1.2 may not be immediately clear. The next example will illustrate the usefulness of this approach better.

When describing the thermal evolution of contact metamorphic aureoles, (sect. 3.6) we will often encounter dimensionless temperatures of the following form:

$$\theta = (T - T_b) / (T_i - T_b) \quad (1.3)$$

There,  $T_i$  and  $T_b$  are the temperatures of an intrusion and the background temperature of the host rock, for example:  $T_i = 900^\circ\text{C}$  and  $T_b = 300^\circ\text{C}$ .  $T$  is the variable temperature which may change as a function of distance from the intrusion or time. When  $T = 600^\circ\text{C}$ , then this in itself is not very instructive. However, when expressed as  $\theta=0.5$  we can see that this temperature is exactly half way between the host rock and the intrusion temperature. For many questions this is much more instructive.

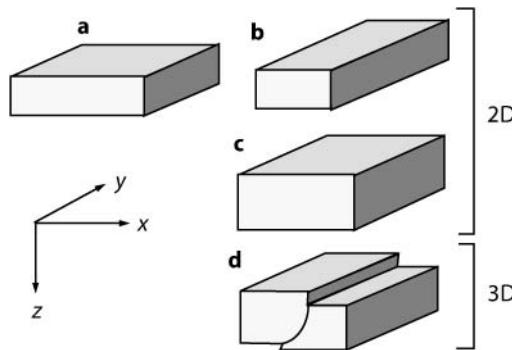
In other examples we will encounter even more complicated formulations of dimensionless variables. For example, in diffusion problems it is often useful to evaluate temperature as a function of the dimensionless variable:

$$T = f \left( \frac{\kappa t}{l^2} \right) \quad (1.4)$$

This may appear quite confusing, but it has also the purpose of simplifying the results, similar to what we did in eq. 1.3. Eq. 1.4 shows that the variables  $\kappa$  (diffusivity),  $t$  (time) and  $l$  (size) are coupled in the particular form of this equation (“ $f()$ ” in eq. 1.4 means “function of”). Using dimensionless variables is not only useful for the better illustration of the meaning of a result, but is also a great aid in differential calculus.

### 1.2.1 Reducing Spatial Dimensions

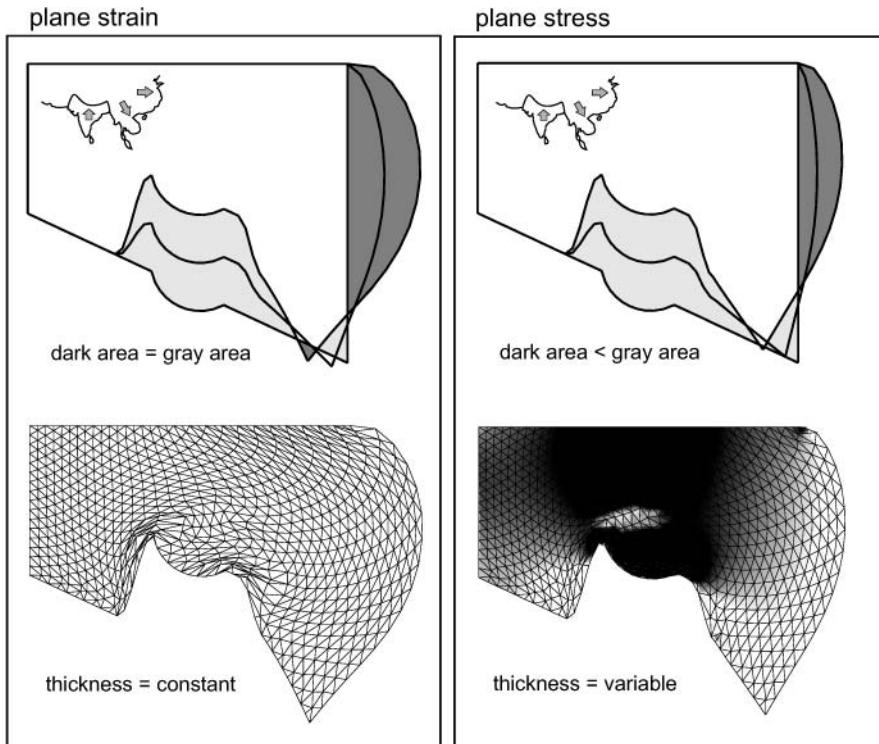
Deformation of lithospheric plates is – in the most general case – a three-dimensional problem (however: s. sect. 2.2.2). In three dimensions, stress, strain and strain rate are described by tensors and any fully three-dimensional description of continental deformation does therefore involve tensor calculations (s. sect. 5.1 and B.3). In order to reduce the complications that arise in such calculations, it is useful to see if the number of considered dimensions can be reduced. When describing lithospheric deformation, there are various well-established simplifications that allow us to neglect some or even all of the components of a tensor. Of course, whether or not these simplifications should be used depends on the nature of the problem. Two of these simplifications are important enough for plate tectonic modeling so that they are mentioned already in this first chapter of the book.



**Figure 1.5.** Illustration of two- and three-dimensional deformation models (abbreviated with 2D and 3D in the figure). The shortening of the block shown in **a** in the  $x$  direction is compensated in **b** only by stretching in the  $y$  direction. The thickness of the block remains constant. This is plane strain deformation. In **c** the shortening of the block is compensated by both stretching in the  $y$  and  $z$  directions. The thickening or stretching in the  $z$  direction is homogeneous. This corresponds to the two-dimensional thin sheet approximation. In **d** the shortening in the  $x$  direction is compensated by stretching in the  $y$  direction and also by inhomogeneous thickening in the  $z$  direction. This kind of deformation can only be described with a fully three-dimensional model. In **b** and **c** the stretching in the  $y$  and  $z$  directions is no function of  $x$  (in contrast to Fig. 1.6). This is no requirement of the plane strain or thin sheet approximations. The illustrated geometries are therefore special cases which could also be described one-dimensionally. Stretching in  $y$  and  $z$  direction could both be evaluated as a function of the shortening in  $x$  only

**Plane Strain Approximation.** The plane strain approximation helps to reduce three-dimensional problems to two dimensions. It assumes that all deformation is strictly two-dimensional so that all strain and displacement occurs in plane and no strain perpendicular to this plane. In plate tectonic modeling we normally consider only the case of no volume change. Then, the total amount of shortening in one spatial direction must be compensated by stretching in the other (Fig. 1.5b; Fig. 1.6). No area change occurs. Tapponier (e.g. 1982) has made great advances in our understanding of continental deformation using this assumption in his descriptions of the India-Asia collision. Plane strain deformation may be viewed as the deformation of a thin film of material that deforms between two fixed parallel plates. It is not *plane stress*, as the normal stresses on the surfaces of the confining plates will vary, depending on where deformation concentrates. Plane strain modeling is a good approximation when the lateral extent of deformation is much larger than the extent in the direction normal to the two modeled dimensions, e.g. crustal thickening or thinning when modeling continental deformation in plan view.

**Thin Sheet Approximation.** When lithospheric shortening in one horizontal direction is compensated both by stretching in the second horizontal direction and by thickening (or thinning) in the vertical, the lithospheric deformation becomes three-dimensional. However, such scenarios can still



**Figure 1.6.** A comparison of modeling in two dimensions assuming plane strain (left) and plane stress or a thin sheet (right hand row). The top two diagrams show the outline of the model region at three different time steps of the runs. The bottom two diagrams show the mesh at the final time step. The shading indicates plate thickness. The initial geometry and boundary conditions are chosen to resemble the India-Asia collision zone (shown as inset at the top) (Houseman and England 1996). In plane strain, the area of the plate remains constant throughout deformation: The light shaded area is identical in size to the dark shaded area. The area loss in the right hand row is compensated by thickening in the third dimension. The boundaries along the top and left hand sides are force to remain fixed. The right hand boundary is stress free in both tangential and normal direction. The bottom boundary is tangentially stress free and is forced to move northwards at the constant rate that ends up in the shape of the plate boundary between Indian and Asian plate. The model region is assumed to have a viscous rheology. Model calculated with the finite element package BASIL of Barr and Houseman (1996)

be modeled in two dimensions assuming the *thin sheet* approximation (s. e.g. Houseman and England 1986a; s. however: Braun 1992). The thin sheet model is based on the assumption that the normal stresses at the surfaces of the plate are constant and that there are no shear stresses on horizontal planes. Thus, the thin sheet approximation is also called the *plane stress* approximation. As a consequence, the plate may thicken or thin in the vertical

direction as to maintain the surface stresses constant but there are no vertical strain rate gradients (Fig. 1.5; 1.6) (England and McKenzie 1982, England and Jackson 1989). Using  $z$  for the vertical spatial coordinate and  $\dot{\epsilon}$  for strain rate, this can be described by:

$$\frac{d\dot{\epsilon}}{dz} = 0 \quad . \quad (1.5)$$

The thin sheet approximation is a good approximation for the description of lithospheric scale deformation when:

1. The shear stresses at the surface and the base of the lithosphere are negligible.
2. If the topographic gradients at these two surfaces are small.

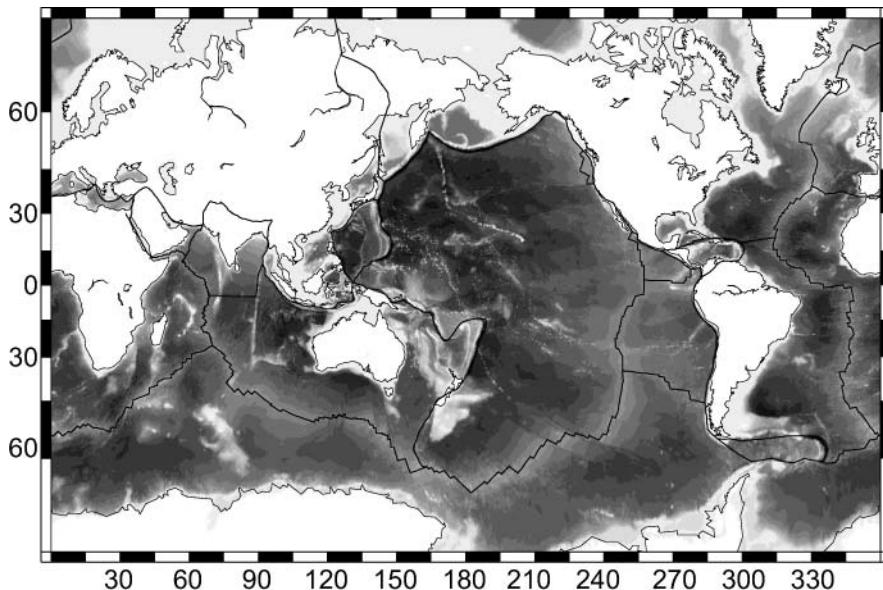
Both are usually given on the scale of whole lithospheric plates (s. more detailed discussion in sect. 5.3 and 6.2). Note however that this description can only describe homogeneous thickening of the lithosphere. Thus, this thickening strain can be evaluated as a variable without the need for a third spatial dimension. Extended formulations have been devised that can help to overcome some of these problems (Medvedev and Podladchikov 1999).

## 2. Plate Tectonics

In this chapter we repeat basic aspects of the theory of plate tectonics. In the first part of the chapter a summary of the history of the plate tectonic model is presented and some basic principles how to describe it are discussed. Some arguments that suggest why plate tectonics may *not* work are also presented. In the second part of this chapter we discuss the layered structure of the earth and the geographic distribution of lithospheric plates. We will also use this chapter to introduce the terminology that is used in the remainder of the book. As such, the chapter is meant as a basis for all following chapters.

### 2.1 Historical Development

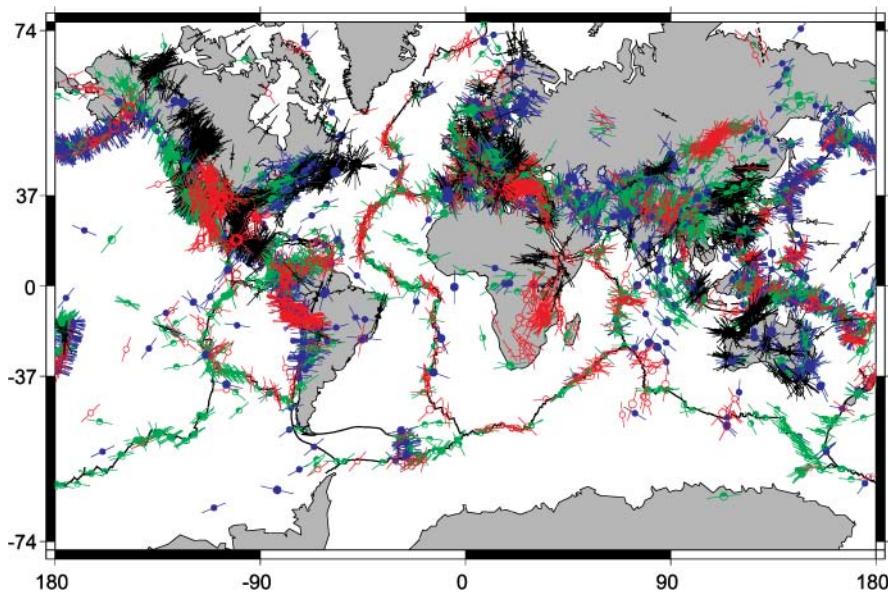
Observations pertaining to the theory of plate tectonics are at least 500 years old. In the late 16th century, Sir Francis Bacon observed that the coast lines of the American and African continents have a matching shape. At Darwin's times the connections between the two continents were already well-established. However, it was only Alfred Wegener who presented the first synthesis explaining these similarities with a theory of plate motion. In part, his synthesis was based on his own observations on the climatological connections between the two continents. While Wegener's synthesis was ultimately proven wrong by the first detailed bathymetric surveys of the oceans in the middle of the past century, his publications (and those of others around his time, e.g. Taylor 1910) are still viewed as the basic foundation of plate tectonic theory (Wegener 1912a,b; 1915). At the time of Wegener, the significance of mid-oceanic ridges and subduction zones were still unknown. However, the deepest point in the world (the Mariana Trench with 11.5 km below sea level) and the eastern Pacific rise had already been discovered by the research vessel H.M.S. Challenger around 1875. Also, mantle convection was already established as the driving force for plate motion (Holmes 1929; Griggs 1939). Nevertheless, Wegener had no model for the processes on the ocean floors and thought of plate motion as some "plough-like" motion of the continents through the oceans. From the time of this theory, only the names of the ancient super continents *Gondwana* and *Laurasia* (du Toit 1937) are still being used.



**Figure 2.1.** Topographic map of the ocean floor. Most regions where the water depth is less than 200 m (nearly white regions) are made of continental lithosphere. Note that in some regions substantial portions of continental lithosphere are actually below sea level, for example around New Zealand, between North-America and Siberia, in the Mediterranean or east of southern South America

The real break through of plate tectonics came not until the mid-twentieth century when the first detailed bathymetric surveys of the Atlantic were performed around 1950 to 1965 (Heezen 1962; Menard 1964). During these surveys, the gigantic mountain ranges and valleys of the ocean floors were discovered that are now known as mid-oceanic ridges and subduction zones. Earlier on, however, these valleys and ridges were interpreted by all kinds of theories, for example the *expanding earth theory* (Carey 1976; King 1983). However, since the mid-sixties it is well-established that mid-oceanic ridges are areas of lithosphere production, while subduction zones are areas where lithosphere is being consumed (*destructive plate margins*). The chains of volcanoes that produce new oceanic lithosphere along the mid-oceanic ridges had already been predicted by some authors, but they were only discovered by the submarine research vessel Alvin (s. Edmond and Damm 1983). Subduction zones and mid oceanic ridges are in volumetric balance so that there is no need for an expanding earth theory (Hess 1961; Vine and Mathews 1963). Today our knowledge of mid-oceanic ridges and subduction zones is well-established as the basic foundation of plate tectonic theory (sect. 2.4.3) (e. g. Morgan 1968).

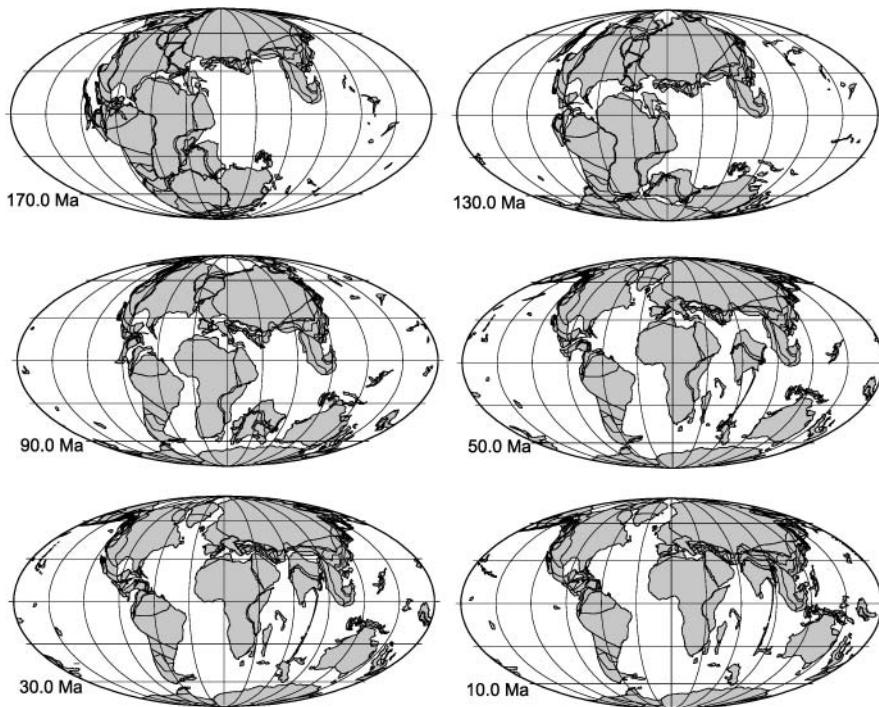
**Modern Understanding of Plate Tectonics.** Our modern understanding of plate tectonic theory is largely based on observations from the ocean floors (Le Pichon et al. 1976). The total length of mid-oceanic ridges is of the order



**Figure 2.2.** The intraplate stress field of the world. The different symbols indicate different methods of determination including earthquake focal mechanisms, bore-hole breakouts and geological indicators. The different shadings indicate different deformation regimes: darkest are thrust faults, medium gray are normal faults and light shading are strike slip faults. For detailed resolution of this figure see the original CASMO facility on the world stress map home page, which was used to create this map (Müller et al. 2000)

of  $60\,000$  km. Most of them are located on the ocean floors (s. sect. 2.4.3). The average rifting rate at these ridges is about  $4$  cm per year, (Table 2.3) which implies that the total production rate of new surface on earth is about  $2\text{ km}^2\text{ y}^{-1}$ . This surface production is balanced by surface wasting at the subduction zones. The total surface of the oceans is about  $A_0 = 3.1 \cdot 10^8\text{ km}^2$  which means that all oceanic lithosphere is being renewed about every 155 my. This is a geologically short time span and means that oceanic lithosphere is one of the younger features of this globe. It is therefore ironic that plate tectonic theory – which is now mostly being applied to our very detailed observations in the continents – has its origin *not* on the continents, but at the deepest points of the ocean floors. Cox (1972) has summarized the revolution of plate tectonic theory. He considers the plate tectonic theory to be based on four independent data sets:

- The topographic maps of the ocean floors (Fig. 2.1).
- The magnetic maps of the ocean floors.
- The age dating of the magnetic maps (Fig. 2.4).
- Detailed maps of the epicenters of global earthquakes (Fig. 2.5).



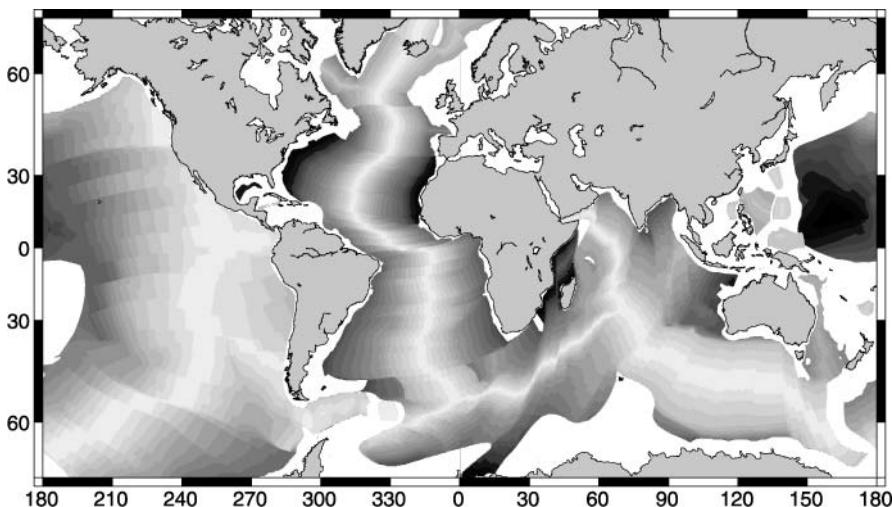
**Figure 2.3.** Plate reconstructions for the last 170 my. Produced with the ODSN home page (s. p. 455). Maps are in Mollweide projection reconstructed relative to the magnetic reference frame

With the advance of GPS measurements of active plate motions over the last 10 years (e.g. Kreemer et al. 2003), it may be said that a 5th data set must be added to the list:

- Direct measurement of plate motions.

Since the fundamental break through of the sixties, plate tectonic theory has made dramatic and rapid advances. It was soon discovered that many observations can be explained by astoundingly simple physical models, all within the plate tectonic concept. For example, the startlingly simple quadratic relationship between space- and time scales of diffusion processes (sect. 3.1.4) has now been used to explain an enormous variety of processes including the water depth of the oceans (sect. 4.4.1), the duration of metamorphic events, the magnitude of contact metamorphic aureoles (sect. 3.6.2) or the shape of chemical zoning profiles in minerals (sect. 7.2.2).

Such amazing success of simple physical models has lead in the past 30 years to an unparalleled development of plate tectonic theory. Much of this development has been characterized by the application of simple analytical



**Figure 2.4.** The age of the ocean floor (Müller et al. 1997). Shading intervals are every 10 my from 0 (white) to 160 my (black). Ocean floor older than 160 my is black. Areas with no data are white. These regions are both on continental and on oceanic lithosphere. Landmasses are grey. Oldest known parts of oceanic lithosphere are around 180 my in the western Pacific just east of the Mariana Trench, between Madagascar and Africa and in the westernmost Atlantic just east of the US. Compare this map also with the topography of the ocean floor (Fig. 2.1) and note the similarities.

models to geological problems. Many of these models are introduced in this book.

**Are We Sure that it Works?** Despite the wide spread acceptance of the plate tectonic model, there are observations and ideas that indicate that it may not work after all. Most important among these observations are the occurrence of zircons with continental signatures found in oceanic islands, for example in Mauritius or Iceland (e.g. Schaltegger et al. 2002). It also has been argued that there is a discrepancy between the observed long narrow zones of axial volcanism at the mid ocean ridges and the geometry of upwelling mantle plumes (McKenzie 2001). As an alternative, it is being discussed if mid ocean ridges may reflect zones of asthenospheric downwelling and that the magnetic striping of the oceanic plates is the consequence of narrowing of formerly much more expanded zones. Within this view substantial parts of the oceanic plates may represent the ultimate end member of an extremely thinned continental margin. Even if arguments can be found that refute such models, some of these hypothesis remind us not to accept the plate tectonic model uncritically.

Nevertheless, there is no debate about the fact that plates move and that the surface of earth includes zones of divergence and convergence that separate zones of less internal strain which we often refer to as cratonic interiors

or plates (Fig. 4.2). More than calling the plate tectonic concept “wrong” it could be said that it becomes slowly outdated in view of our much better understanding of the processes that govern plate motion and deformation.

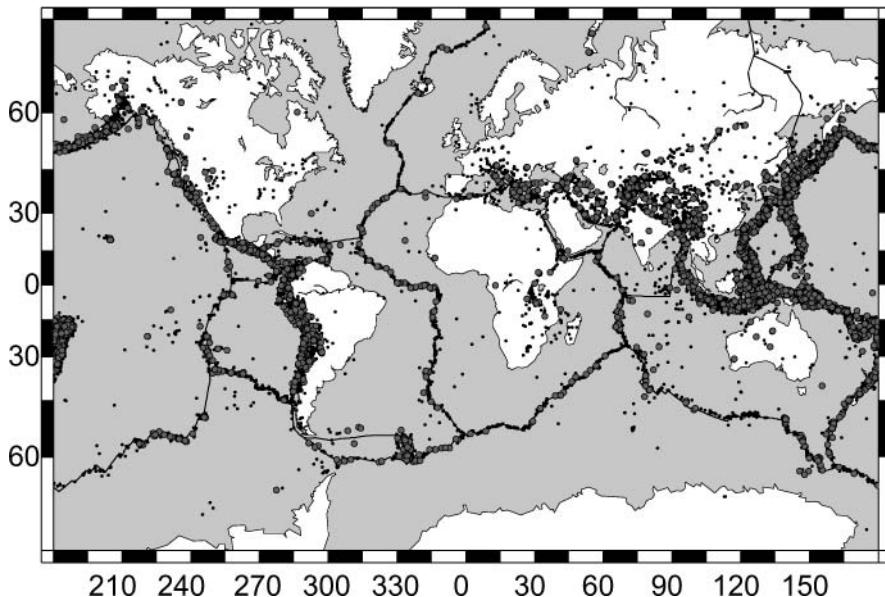
**Recent Developments and Future.** Although the global distribution of earthquakes shows that *oceanic* plates act according to the principles of plate tectonic theory, continental plates do not: Oceanic plates are large and flat and their margins are narrow compared to their size and seismicity is confined to the rims of these plates. *Continents* behave differently. They are characterized by deformation and seismicity which reaches wide into the continents themselves, their thickness varies dramatically and some seismicity even occurs in their centers. Their plate boundaries are diffuse. It has therefore been suggested to use the term “cheese tectonics” as a superceding term for “plate tectonics”, because continents seem to deform rather like soft camembert than like rigid plates. Like many good jokes this one has a true core. It has been known since the turn of the century that rocks behave like fluids on geological time scales (s. summary by Gordon 1965 or England 1996). However, only in the past 30 years geologists have begun to actually use the theory of deformation of viscoelastic materials to describe the dynamic evolution of continents (England and McKenzie 1982; England and Jackson 1989). In part these new descriptions were triggered by studies like those of Goetze (1978) or Brace and Kohlstedt (1980) who provided us with the first simple models of lithospheric rheologies.

Because of such models it has now been possible to describe largest scale tectonic processes with simple dynamic models. Even coupled thermomechanical approaches can now be performed with startling simplicity (e.g. Sonder and England 1986) and have found their use even by non-geophysically oriented earth scientists.

The current development of plate tectonics is going more and more towards the use of numerical models and further away from simple analytical models. Digital data sets of global observations – for example the global digital elevation model SRTM3 or the global sea floor topography of Smith and Sandwell (1997) – make it now possible to tackle problems that can only be solved using large numbers of data. Such models can now be used to explain problems that go beyond single observations and pertain to the whole globe. The future will surely be characterized by an increasing use of large data sets and numerical models.

## 2.2 Working on a Spherical Surface

The earth is nearly a sphere (s. sect. 4.2) and many aspects of the geometry and the mechanics on a sphere are different from its Cartesian equivalent. In this section we discuss some aspects of spherical coordinates that may need to be considered when solving geodynamic problems on very large scales. We



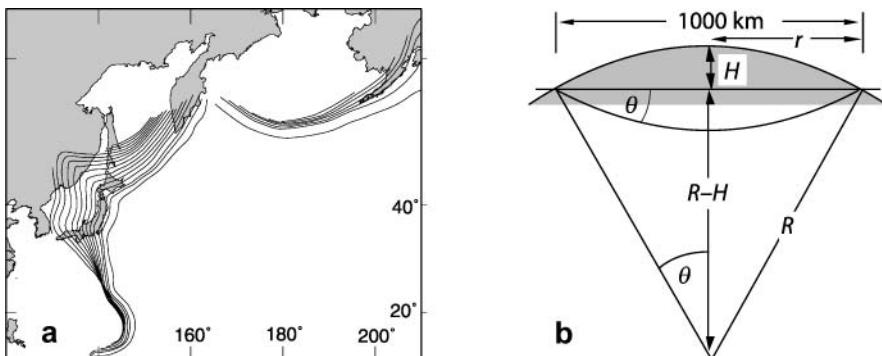
**Figure 2.5.** Global seismicity. All recorded earthquakes above magnitude 6 since 1973 are shown. Note that most earthquakes along oceanic plate boundaries (e.g. the Mid-Atlantic Ridge) occur very near the plate boundary, while earthquakes associated with plate boundaries in continental lithosphere cover much more diffuse areas (e.g. the India-Asia collision zone). See Fig. 2.20 and Fig. 4.2 for more details about the plates and strain distribution

also discuss some ball park estimates that can be used to judge whether it is sufficient to neglect the spherical geometry.

### 2.2.1 ... or is the Earth Flat After All?

By far the largest number of geodynamic problems can be described assuming a flat earth. Fig. 2.6 shows that over a line two thousand kilometers in length, the curvature of the earth causes about 80 km deviation from a straight line. This is only about 5% of the extent of the feature. For geological features of some tens of kilometers extent, the deviation is only some tens of meters, which is about 0.1% of the extent of the feature. We can conclude that only for geodynamic problems that have a length scale of substantially more than thousand kilometers the curvature of earth must be considered.

A famous example for a problems that can not be described on a flat earth is the shape of long subduction zones. On a flat earth, the trace of a subduction zone should be linear, just like the linear trace of the curvature of a sheet of paper hanging off the edge of a table. In contrast, the trace of many deep ocean trenches is curved along the surface of earth. For example, the Aleutian, Kurile, Japan, Izu Bonin and Mariana trenches from northeast to southwest) along the Pacific ring of fire, the South Sandwich Trench east



**Figure 2.6.** Illustration of the ping pong model. **a** Slab contours for the Aleutian slab and the Kuril-, Japan- and Mariana slabs. Depth contours are shown every 50 km from 0 km to 700 km (in the Japan Trench). Data from Gudmundsson and Sambridge (1998), (<http://rses.anu.edu.au/seismology/projects/RUM>). **b** The difference between a flat and a curved surface of the earth. The maximum deviation of a curved surface from a flat surface,  $H$ , is given from geometric relationships by  $H = R - \sqrt{R^2 - r^2}$ . For  $r = 1000$  km as shown here and the value for  $R$  from Table D.3 the deviation is  $H \approx 80$  km

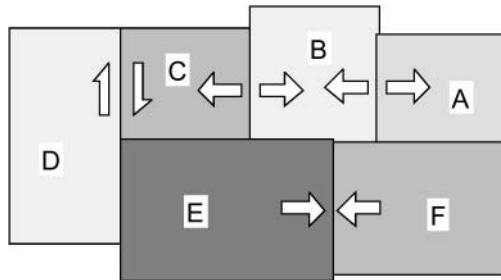
of South Georgia or the Java (Sunda) Trench south of Sumatra and Java are all curved concave towards the subduction direction in plan view. This is often explained with the ping pong ball model. This model compares the curvature of trenches on the surface of the earth with the curvature of the indentation edge on a dented ping pong ball. If the indented part of a ping pong ball is not deformed in itself, then this edge forms a small-circle on the surface of the ball. Exactly this is observed in subduction zones. In fact, the model can be used to predict the subduction angle ( $2\theta$ ), which should be given (according to Fig. 2.6) by:

$$\sin(\theta) = \frac{r}{R} \quad (2.1)$$

where  $r$  is the small circle radius and  $R$  is the radius of earth. Most of the small circle radii of subduction zones on earth correspond well with the subduction angle predicted by eq. 2.1 (s. Isacks and Barazangi 1977). However, the subduction angles may also depend on a large number of other parameters, for example whether subduction occurs *in* or *against* the direction of convection in the asthenosphere (Doglioni 1993). In fact, it has been suggested that the earth's rotation causes a westward drag between lithosphere and asthenosphere which also influences the direction and steepness of subduction zones (Doglioni et al. 1999). Other plate scale examples for the influence of the curvature on plate motions are the transform faults in the oceanic lithosphere (Fig. 2.9).

*Geometrical* problems can be described on a flat earth using the familiar Cartesian coordinate system. In this book we use  $x$  and  $y$  for the coordinates

**Figure 2.7.** Six plates on a flat surface. The relative motions of some plates are shown by the arrows. However, the relative motions of plates C and E, D and E, B and E, B and F as well as that between A and F are completely unconstrained by the shown relative motions



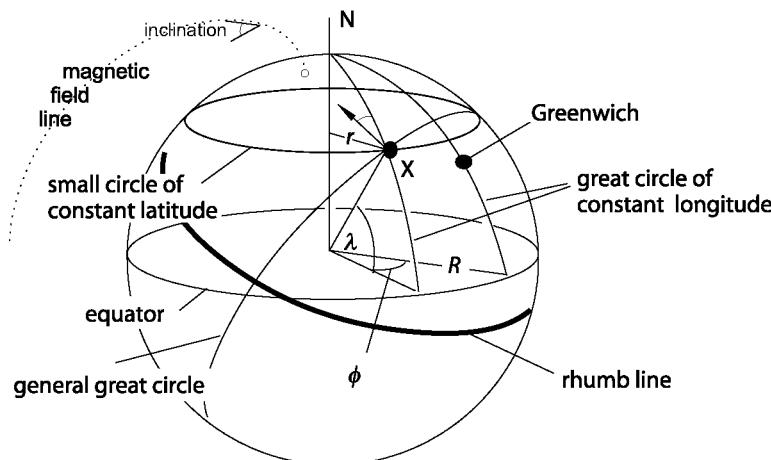
oriented parallel to the surface of the earth and  $z$  for the vertical direction. In section 2.3 we discuss the relationship between those coordinates and those of other coordinate systems. In section 4.2 we discuss some confusions that may arise from mixing different reference frames.

*Kinematic* and *dynamic* problems on a flat earth can be described by motion and forces in the horizontal and vertical directions. Velocities and forces are described by vectors. This means that they have a direction and a magnitude and can be split up into vector components that are parallel to the axes of a Cartesian coordinate system. For combined paths we can therefore use the rules of vector algebra. Be careful to note, however, that all motions (and forces) are relative! This is important to realize as an observed motion in *one* place of the globe must not imply that this motion is elsewhere the same (for example China is moving towards Tibet, but the region *between* China and Tibet is under extension). Fig. 2.7 illustrates some more examples. All this is quite different on a spherical surface.

### 2.2.2 Geometry on a Sphere

On a spherical surface the position of a point is described by its longitude  $\phi$ , and latitude  $\lambda$  (Fig. 2.8). As with time, spherical geometry is one of the few branches in science where the duo-decimal system is still in use: a right angle has 90 degrees and longitude and latitude around the globe are divided into 360 degrees. (The use of 100 degrees for a right angle was attempted by the introduction of “new degrees” but has not found footing in science). Every degree of longitude is described by a great circle which goes through the geographic poles. These great circles are called *Meridians*. Great circles are lines on the surface of a sphere that are defined by the intersection of a planar surface through the center of the sphere, with the surface of that sphere. Meridians are therefore a special kind of great circle, namely one that goes through the poles. Small circles are defined as intersections of all other planar surfaces with the surface of a sphere. 180 of the 360 Meridians are numbered west of Greenwich and the other 180 east of Greenwich, which has been internationally agreed upon to be the reference for longitude. Each degree of latitude is defined by a small circle parallel to the Equator and at right angles to the axis that connects the poles. 90 degrees of latitude

are north of the equator and 90° are south. Note that there is a total of 360 Meridians, but only 180 degrees of latitude. The spacing of the degrees of latitude is chosen so that they divide the Meridians into 360 sections of equal length. Thus, the distance (along the surface of the earth) between degrees of latitude is constant everywhere on the globe, while the distance between degrees of longitude is largest at the equator and zero at the poles. For more detailed description of locations on a spherical surface, every degree is divided into 60 arc minutes and every arc minute into 60 arc seconds. Just to make things worse, the duo-decimal system is often coupled with the decimal system: Geographic locations are often described by degrees and decimals. That is, tenth and hundredths of degrees are given, rather than arc minutes and arc seconds.



**Figure 2.8.** Definitions of important lines and angles on a spherical surface. The geographic longitude  $\phi$  of point X is *west* of Greenwich. The white dot is the magnetic pole. The angle between magnetic and geographic north (labeled at point X) is the declination and the plunge angle of the magnetic field lines is the inclination

The circumference of a great circle on earth is about  $2R\pi \approx 40\,000$  km. (If it were exactly 40 000 km, then the radius of the earth would have to be  $R = 6\,366.2$  km; in reality the equatorial radius is 6 378.139 km and the polar radius is 6 356.75 km). In fact, one meter was long defined as the 1/40 000 000 part of the circumference of earth. One degree of longitude at the equator (and *all* degrees of latitude) is therefore about  $40\,000/360 \approx 111$  km. On small circles north and south of the equator, the distance between full degrees of longitude,  $l$ , decreases with the cosine of the latitude:

$$l \approx \cos(\lambda) \cdot 111 \quad . \quad (2.2)$$

In eq. 2.2 we have used the approximate value for one degree of longitude at the equator. Correspondingly, the small circle radius of each small circle of

constant latitude shrinks with the cosine of the latitude:  $r = \cos(\lambda) \times R$ , where  $R$  is the radius of earth. One arc minute of latitude is defined as one nautical mile which is  $\approx 1.8$  km. Along the equator, distances between degrees of longitude and latitude are of equal length.

Other important lines on spherical surfaces are *rhumb lines* (also called *loxodromes*). These are lines that intersect degrees of latitude and longitude at constant angles. Rhumb lines are easy to follow, for example when setting constant course on a ship, but they form spiral-shaped curves on a sphere and they are *not* the shortest connection between two points (Fig. 2.8). The angle between magnetic north and the lines of longitude (geographical north) is called the magnetic declination. The vertical angle between the normal to the Geoid surface and the magnetic field lines is called the magnetic inclination.

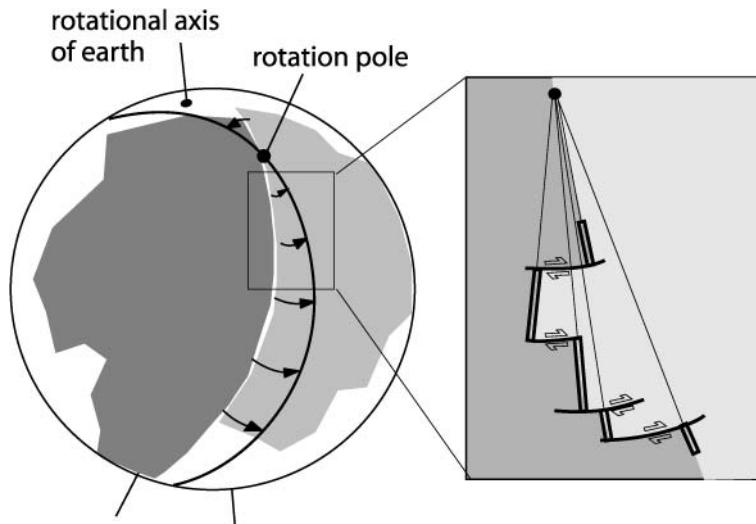
All distances and angles on a spherical earth can be calculated with simple combinations of the trigonometric functions. Throughout the book it is assumed that the use of those, as well as some sound spatial imagination is familiar to the reader. As a reminder, some of the relationships we will need are summarized in Appendix C, Tables C.4, C.6, C.7, and Fig. C.1.

### 2.2.3 Kinematics on a Sphere

On a flat surface, velocity  $v$  and speed have the units of  $\text{m s}^{-1}$ . Velocity is a vector and speed is a scalar quantity. For example, the Indian Plate has a *speed* of  $0.05 \text{ m s}^{-1}$ , but a *velocity* of  $0.05 \text{ m s}^{-1}$  *moving north*. The equivalent to velocity on a spherical surface is the angular velocity  $w$ .  $w$  has the units of radian per time, which is  $\text{s}^{-1}$ . The axis that is perpendicular to the planar surface swept over by angular motion is called the pole of rotation or *Euler pole* (Fig. 2.9). The velocity that corresponds to a given angular velocity depends on the distance of the angular motion from the pole of rotation. Acceleration in a straight line is the change of velocity over time and has the units  $\text{m s}^{-2}$ . Correspondingly, the angular acceleration has the units of  $\text{s}^{-2}$ . The differences in units between linear velocity and angular velocity has lead to a lot of confusion in the literature. For example, a *constant* rate of plate motion with a constant angular velocity will cause differences in the rates of relative plate motions along the plate margin. The relative plate motion rate depends on the small circle radius of the velocity vector (Fig. 2.9). In fact, even qualitative changes from divergent plate motion to convergent plate motion may occur along a plate margin because of this (compare Fig. 2.7 and 2.9). This is spectacularly illustrate by the transform faults on the ocean floors which solve the space problem caused by the angular rotation.

### 2.2.4 Mechanics on a Sphere

Plate tectonic forces are often described in the literature as “torques” (e.g. Sandiford et al. 1995). For example, ridge “push” is a force, while many authors rather use the term ridge “torque”. Strictly speaking, we should



**Figure 2.9.** Illustration showing the meaning of rotation poles. The arrows are vectors showing the direction and magnitude of relative motion of the two plates (shaded regions). The thick line connecting the arrow heads is the new plate margin after some time. The axis of the earth is only shown to emphasize that it has nothing to do with the rotation pole of plate motion. The enlargement illustrates the origin of transform faults

only use torques to understand the mechanics that cause plate motion on the earth's surface as plates do not move in a linear direction but rather around a rotation pole (the center of earth). In calculating a force or torque balance of a mountain belt, where every point in the belt is virtually the same distance from the pole of rotation, the distance to the rotation pole cancels out and torque balances and force balances are practically equivalent (s. p. 216).

Force  $F$  is given in Newtons [N] and:  $1 \text{ N} = 1 \text{ kg m s}^{-2}$ . Force is a vector with a magnitude and direction. Horizontal forces are therefore tangential to the globe. The equivalent on a spherical surface is torque. Torque (which is different from angular momentum!) is the turning moment which is exerted by a force about an axis. It is given by the product of force and the distance from the axis about which the torque acts. Torque has the units of Nm or  $\text{kg m}^2 \text{s}^{-2}$ . A force of  $10^{12}$  Newton that acts in direction of a great circle on the earth's surface, corresponds to a torque of  $6.37 \cdot 10^{18}$  Nm. The torque changes along a plate margin, as the normal distance of the plate margin to the rotation pole changes. The units of torque can be read as "Newton times meter of leverage", where the "meters of leverage" are the normal distance to the rotation pole. In the literature, "forces" are often given in Newton per meter, meaning for example, that the force building a mountain range is normalized "per meter length of orogen". It is important not to confuse this

**Table 2.1.** Important kinematical and mechanical parameters and their units. Each parameter is given with both, the linear and the spherical equivalents

physical parameter	unit
velocity	$\text{m s}^{-1}$
angular velocity	$\text{s}^{-1}$
acceleration	$\text{m s}^{-2}$
angular acceleration	$\text{s}^{-2}$
force	$\text{kg m s}^{-2}$
torque	$\text{kg m}^2 \text{s}^{-2}$
mass	$\text{kg}$
moment of inertia	$\text{kg m}^2$
linear momentum	$\text{kg m s}^{-1}$
angular momentum	$\text{kg m}^2 \text{s}^{-1}$

with torques, which have the units of Newton times meters, there the meters are the distance to the rotation pole.

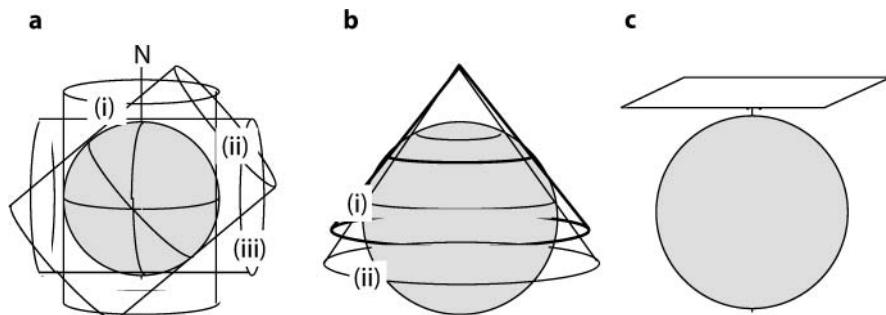
Force = mass  $\times$  acceleration ( $F = m \times dv/dt$ ) and similarly torque = mass  $\times$  angular acceleration. In plate tectonics the changes in velocity and angular velocity occur over very long time periods, so that accelerations and angular acceleration are negligible (s. p. 218). Thus, the common assumption is that the sum of the torques or the net torque acting on a plate is zero or, correspondingly, that the sum of the forces or net force acting on a smaller region such as a mountain belt is zero.

Aside from force and torque, there are some other important mechanical parameters that we will need in this book. The linear momentum  $I$  is the product of mass and velocity:  $I = mv = \text{kg m s}^{-1}$ . Just as momentum is given by  $I = mv$ , the *change* of momentum is given by:  $\Delta I = m\Delta v$ , mass times the *change* in velocity. Considering that force has the units of the product of mass and acceleration (the *change* in velocity) we can write that:  $F = mdv/dt$ . Thus, the *change* of momentum is:  $\Delta I = Ft$ , force acting over a given time. On a spherical surface, the angular momentum is analogous to the linear momentum. The angular momentum  $D$  is the product of the moment of inertia  $J$  and the angular velocity  $w$ :  $D = Jw$ . The moment of inertia is the ratio of torque and angular velocity and has the units of  $J = \text{kg m}^2$ . Angular momentum has the units of  $D = \text{kg m}^2 \text{s}^{-1}$ . In plate tectonics, changes in momentum and angular momentum are ignored because changes in velocity and angular velocity of plates takes place over very large time periods (s. sect. 5.1.1). In most problems then, the net forces or torques acting on a given plate must balance or add up to zero (s. p. 216). However, momentum

and angular momentum help to understand why forces *or* velocities can form the boundary conditions for orogenic processes.

## 2.3 Map Projections

Map projections are mathematical or geometrical models that portray the features of the spherical surface of our globe on the two dimensions of paper. The map projection model serves the purpose to make these features easier to look at and pays for that by accepting some distortion. Usually, the process of projection involves the conversion of latitude  $\lambda$  and longitude  $\phi$  into Cartesian coordinates (Robinson et al. 1984). Map projections have become an increasingly important part of geology since global data sets are used. Map projections can be performed geometrically or they can be purely mathematical, without any apparent geometrical equivalent. Among projections *with* a geometrical equivalent we discern three important types (Fig. 2.10):



**Figure 2.10.** Schematic illustration of three examples of commonly used geometrical map projections. **a** *Cylindrical projection* with examples of (i) normal or equatorial, (ii) transverse and (iii) oblique orientations of the projection surface; **b** *conical projection* with examples of (i) tangential orientation (touching the sphere along the thick line) and (ii) secant orientation (intersecting the globe) of the projection surface; **c** *Azimuth* projection with polar orientation of the projection surface

- Cylindrical projections,
- conical projections,
- azimuth projections.

These three types of projections can be imagined as projections of the earth's surface from an imaginary light source (usually assumed to be located at the center of earth) onto some enveloping surface which is then rolled open to form a planar surface. In these three projections, the enveloping surface is a cylinder, a cone or a planar surface. The enveloping surface can touch the globe along one line (tangential orientation) or transect it (sectant orientation) so that it has two lines along which there is no distortion. In total,

there are several hundreds of known projections that serve a large variety of purposes. A large number of other projections are explained by Snyder (1987) and Snyder and Voxland (1989). All their projections are usable with the software PROJ by Evenden (1990) and there is great overlap with the projections incorporated in GMT and described therein (Wessel and Smith 1995). A bit more detail was also given in the first edition of this book. Most known projections are *not* geometrical but are only described by a *projection function* that relates spherical and planar coordinates. Just about all map projections are either:

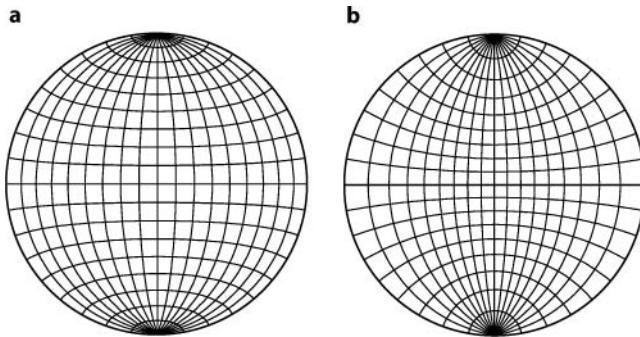
- Conformal (orthomorphic) projections,
- equivalent (equal-area) projections,
- azimuthal projections.

*Conformal* projections represent angular relationships the way they are on the spherical surface. *Equivalent* projections render the same areas as those on the spherical surface. *Azimuthal* projections retain azimuthality, that is, the directional relationships along great circle bearings are the same on the map as they are on the sphere. No projections on two-dimensional paper can retain more than one of these relationships from a sphere at the same time. However, many projections aim to find a useful compromise between different relationships, on the expense of being neither true in angle nor true in area or great circle bearing.

Most geologists are familiar with at least one conformal and one equivalent projection, even if they have never used maps of the globe. These are the nets of Schmidt (also known as *Lambert's equivalent projection*) and Wulff (also known as *stereographic projection*) that are used by crystallographers for the presentation of angular relationships of crystal faces and by structural geologists for the equivalent illustration of planar structures, respectively (Fig. 2.11). Note that lines of constant latitude and longitude need not look like they do on Fig. 2.11. If the orientation of the projection surface is oblique to the axis of the globe, then these lines may have very different shapes as shown on Fig. 2.12.

### 2.3.1 The Mercator Projection

The *Mercator projection* is one of the most familiar projection and is one of several known cylindrical projections where the projection surface is a cylinder as shown in Fig. 2.10a. It was first used by G. Mercator in 1569 for navigation on the oceans. For this projection, the projection surface is a cylinder a cylinder that touches the globe along the equator and has its axis therefore parallel to the axis of earth (Fig. 2.10a). Every point of the earth's surface is projected onto this cylinder from an imaginary light source at the center of earth. The imaginary cylinder is then rolled open. The map is ready! There are two great advantages of this projection:



**Figure 2.11.** **a** The equal-area net by Schmidt. **b** The conformal net of Wulff

1. There is *no* distortion along the equator and the distortion in equator-near regions is negligible.
2. Lines of constant longitude, latitude and rhumb lines appear as straight lines.

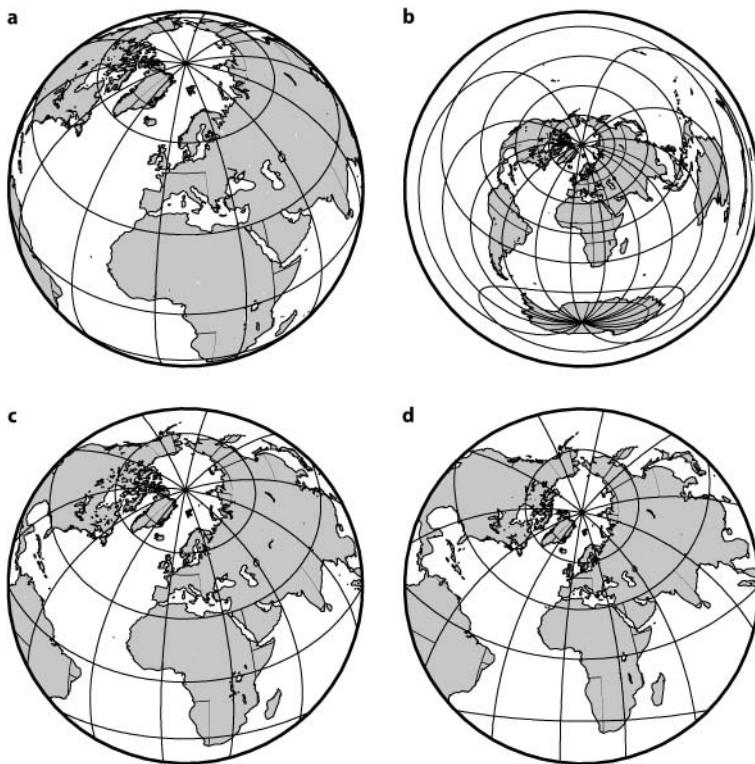
Because rhumb lines are easy to navigate along, Mercator projections used to be common for shipping (before navigation software made more direct routing much easier). However, note that the shortest connection between two points on the globe – a great circle – appears on the Mercator projection as a curved line. The biggest disadvantage of the Mercator projection is that the areal distortion increases with latitude. Areas near the poles appear too large and the poles themselves cannot be represented at all. For example, Greenland appears to be larger than south America although it has only an eighth part of its size.

This problem can be avoided by using the *transversal Mercator projection*. There, the cylinder touches the globe along one of the Meridians and its axis is at right angles to the axis of the earth. Accordingly, there is no distortion along this Meridian and the distortion goes towards infinity on the equatorial points of the Meridian that is  $90^\circ$  from the Meridian where the cylinder touches. Both lines of constant longitude *and* latitude are curved on this projection.

**The UTM Projection.** The Universal Transversal Mercator projection (UTM) is defined by a total of 60 different Meridians (all separated by 6 degrees of longitude) all of which are touching Meridians for 60 assumed projection cylinders (Fig. 2.13). In other words, it is a combination of 60 transversal-Mercator projections. This avoids serious distortion of any point on the surface of the globe.

### 2.3.2 National Grids

Using latitude and longitude is not always the best way to define a point on a map. Lines of constant latitude and longitude may be curved lines or – even

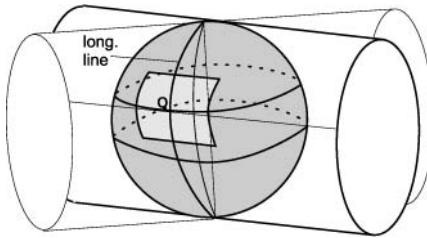


**Figure 2.12.** Four examples of map projections with circular outlines. **a** orthographic projection, **b** equidistant azimuthal projection, **c** Lambert's equal area projection and **d** stereographic projection. All shown examples are for inclined orientations of the rotation axis of earth relative to the projection direction. If **c** and **d** were projected parallel to the axis of earth, then lines of constant longitude and latitude in **c** would correspond to those of Fig. 2.11a and lines of constant longitude and latitude in **d** would correspond to those of Fig. 2.11b

if they are straight lines – their spacing is unlikely to be constant. Thus, in order to simplify the use of maps, *national grids* are used. National grids are orthogonal grids with equal spacing in two horizontal directions  $x$  and  $y$  that are superimposed on projected maps (Figs. 2.13; 2.14). Different countries use different grids for their national mapping but many of them are based on the UTM projection or slight modifications thereof.

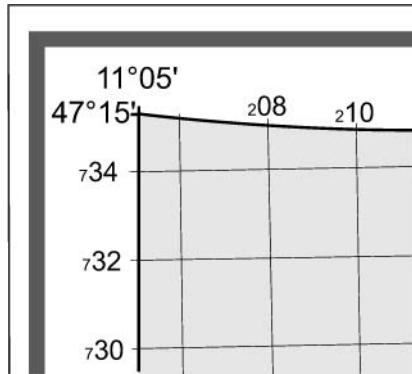
For example, topographic maps in Austria are based on the Gauss Krüger projection which differs from the UTM projection by having transversally oriented cylinders touching earth (described by a Bessel ellipsoid) every 3 degrees (not 6), starting at Ferro (instead of Greenwich) and going east (for details on the Bessel ellipsoid and its difference to the WGS84 system see sect. 4.2.1). Within this projection, the axes of the superimposed orthogonal national grid are assumed to be the equator and the touching Meridian (which are called M28, M31 and M34 in Austria).  $x$  and  $y$  values on the orthogonal

**Figure 2.13.** Illustration of the UTM projection. A series of transversely oriented cylinders touch the reference ellipsoid along longitudinal lines. A given point along this line is then chosen as the origin of an orthogonal national grid with northing and easting values.

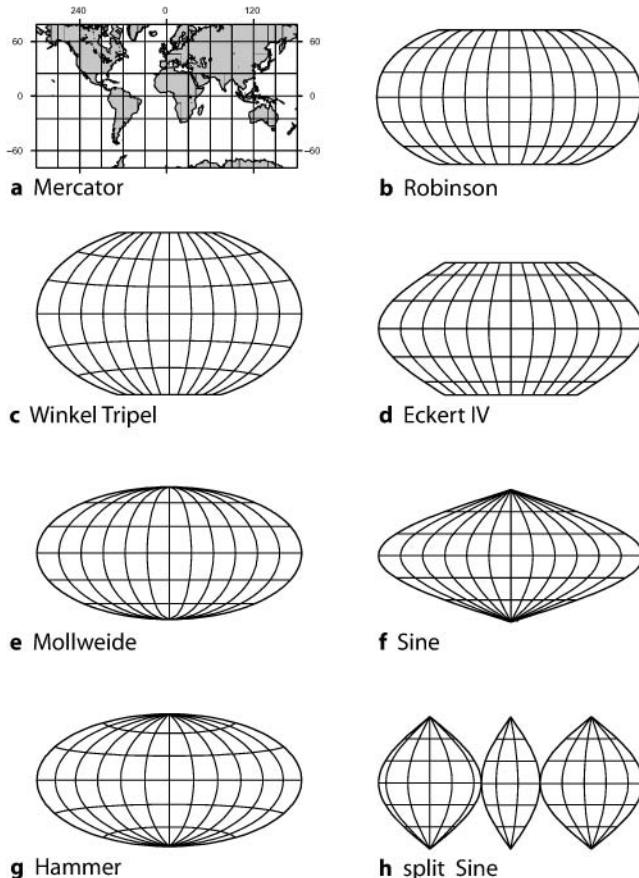


grid are then counted eastwards from the touching Meridian in kilometers and northwards from the equator, respectively. Just to make it a bit more complicated surveyors have decided that – in Austria – the  $y$  values get too large if counted from the equator and that negative  $x$  values west of the touching Meridian look ugly. To resolve this, the origin at the equator is assigned an arbitrary  $x$  and  $y$  value ( $x=150$  km and  $y = -5000$  km for the M28) so that the numbers at  $45^\circ$  have reasonable values and eastings don't get negative.

**Figure 2.14.** Example of the top left corner of a topographic map sheet. Typically, referenced map sheets are bound by lines of constant latitude and longitude. The “lat” and “long” values are indicated at the corners. Because these may be curved (depending on the projection), the corners are usually not parallel to the edges of the paper. The superimposed national grid is usually slightly oblique to the lat and long lines and is labelled in kilometer. The fact that the first digit is printed small has no significance other than to save some space



Similarly, the National Imagery and Mapping Agency (NIMA) of the United States has divided the United States into 10 zones (numbered from 10 through 19), each of which is 6 degrees wide and which lie between  $126^\circ$  west and  $66^\circ$  west. Within each zone, a transversely oriented cylinder is assumed to touch a central Meridian and northing and eastings on the superimposed orthogonal grid are counted in kilometers from 10 000 km south of the equator and 500 km west of the touching Meridian. The shifting of the grid origin for each zone to 10 000 km south and 500 km to the west is done to avoid negative numbers south of the equator or on western halves of zones, because the US use this system for maps world wide.



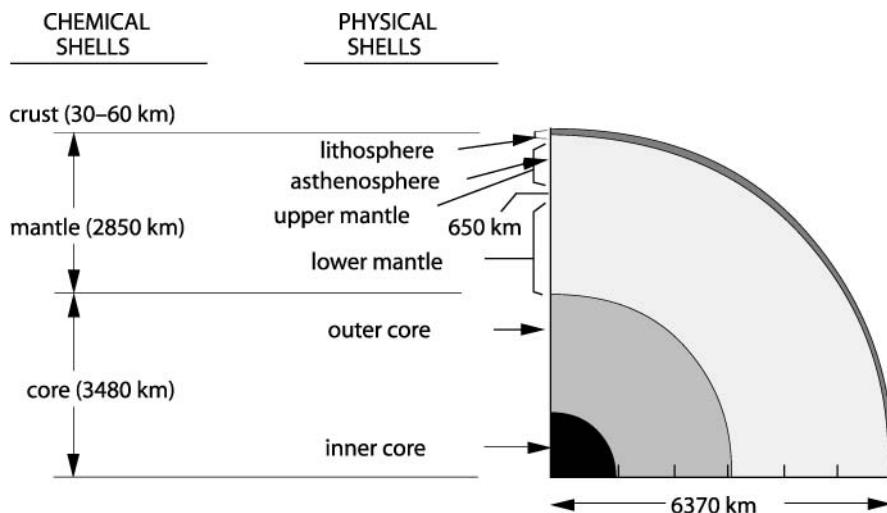
**Figure 2.15.** Eight examples of projections commonly used to portray the entire globe. **a** Mercator projection shown up to 75 degrees north. **b** The Robinson projection is pseudocylindrical and is neither conformal nor equivalent but “looks right”. **c** Winkel tripel projection. **d** The Eckert IV projection is pseudocylindrical and equivalent. It is common in Atlases. **e** The Mollweide projection is also pseudocylindrical and equivalent with longitudes being straight lines and Meridians parts of ellipses. **f** The equivalent sine projection is often used for south America. **g** The Hammer projection is also equivalent. **h** A split version of the sine projection shown above

## 2.4 The Layered Structure of Earth

Earth can be divided into layers according to:

- different materials,
- different physical properties.

When considering the layers made of *different materials*, there are three: a crust, a mantle and a core (layers above the surface of solid earth like the



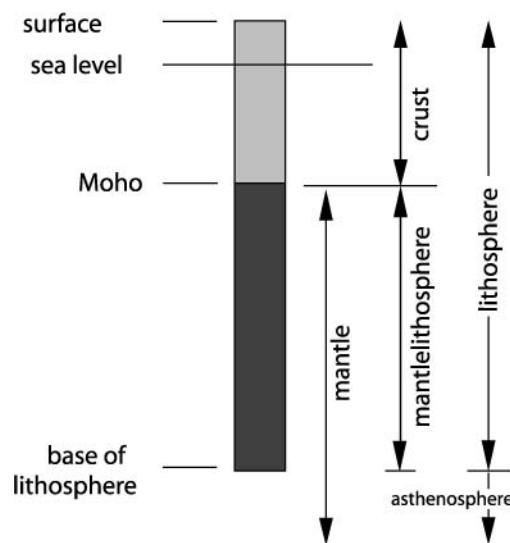
**Figure 2.16.** The layered structure of earth

hydrosphere, biosphere or atmosphere are not considered here). The crust is the uppermost layer. In its normal state, it is between some 5 to 7 and 30 km thick, depending on whether we deal with oceanic or continental crust. In regions of deformed continental crust – for example underneath the Tibetan Plateau or the Alps – continental crust can get up to 80 km thick. Chemically, the crust is highly-differentiated and very heterogeneous, but many of its mechanical and physical properties (e.g. density, conductivity or rheology) can be well-approximated with those of quartz. The mantle is largely made up of olivine and – at larger depths – its high-pressure breakdown products. The seismically clearly visible contact between crust and mantle is called the Mohorovičić-discontinuity (short: Moho). From the Moho the mantle reaches down to a depth of about 2900 km. The core consists mainly of iron and nickel.

When considering the *physical properties*, the layered structure is quite different. Then, the outermost layers of earth are the lithosphere and the asthenosphere. The lithosphere is solid and acts like rock on geological time scales. Therefore its name. It involves both a crustal and a mantle part. The asthenosphere consists of the soft mantle that underlies the lithosphere. Some authors call the entire upper mantle underneath the lithosphere the asthenosphere. Others use the term only for the mantle section that lies above the point where the adiabatic melting curve comes nearest to the temperature profile (Fig. 3.11). According to Ringwood (1988) the mantle can be divided into three zones:

1. The upper mantle, which reaches down to about 400 km and is characterized by a seismic p-wave velocity of about  $8.1 \text{ km s}^{-1}$ .
2. A transition zone from about 400 km to the 650 km discontinuity.

**Figure 2.17.** Nomenclature of different parts of the outer shells of the earth



3. The lower mantle which reaches from the 650 km discontinuity to the core-mantle boundary at 2 900 km depth.

Below that is the core. The outer core is liquid and the inner core is solid.

#### 2.4.1 Crust and Lithosphere

The lithosphere is the outer solid layer of the earth (s. sect. 3.4). As for the whole globe, the lithosphere can be divided according to its physical properties or according to its chemical (material) properties. Because there is overlap between layers distinguished on the basis of different properties it is crucial for the understanding of this book to be familiar with the nomenclature illustrated in Fig. 2.17. However, for many geodynamic questions it is not necessary to consider these subdivisions, as the lithosphere acts as a whole.

When considering its chemical properties, the lithosphere consists of a crust and a mantle part. The crust consists of highly-differentiated partial melts from the mantle. The mantle part of the lithosphere is largely made up of the similar material to that of the underlying asthenospheric mantle, but it acts like a solid, because of its lower temperature. However, we note that chemical differences between the mantle lithosphere and the underlying asthenosphere do exist and account for example for unusually thick, but apparently mechanically stable mantle lithosphere, underneath southern Africa. Modern research has been able to document much detail of the compositional variation within the uppermost mantle, both on chemical grounds (e.g. McDonough and Rudnick 1998) and based on seismic velocities (Jordan 1981a;

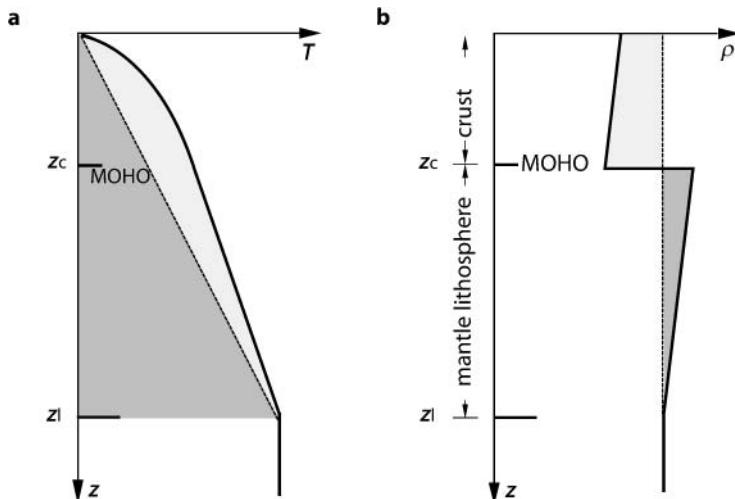
1989). Nevertheless we take in this book the simple-minded view that density variations between mantle lithosphere and asthenosphere may be largely attributed to differences in temperature (e.g. on p. 165).

A schematic but characteristic thermal and density profile of the lithosphere is shown in Fig. 2.18. A very large number of the geodynamic processes discussed in this book are a function of the fundamental shape of the curves on this figure.

Within the crust the temperature profile is curved, because of radioactive heat production. Within the mantle part of the lithosphere, the thermal profile is linear (in a steady state). The base of the thermal lithosphere is defined by the point where the temperature profile intersects the 1200°C or 1300°C isotherm (sect. 3.4). At higher temperatures, mantle material begins to flow rapidly on geological time scales and any temperature gradients will be eliminated by convection. Thus, temperature and density are constant below the depth  $z_1$  on the scale of Fig. 2.11. Both curves of Fig. 2.18 will be useful help throughout this book.

**Definition of the Lithosphere.** The term “lithosphere” comes from the Greek *lithos* = rock and was introduced by Suess (1885). The term was later used by Barrell (1914) and ultimately defined by Isacks et al. (1968) as a *near surface layer of strength* of earth. Even today it remains difficult to find a more precise definition than this. Most of the physical parameters, for example temperature or density, change continuously underneath the Moho and the transition from the rigid outer shell of the earth (the mechanical boundary layer) into the more viscous hot asthenosphere (from the Greek *asthenia* = soft) is also continuous. This transition zone is called *thermal boundary layer* (Fig. 2.19; sect. 6.3.2; Parsons and McKenzie 1978; McKenzie and Bickle 1988). However, even on the definition of the term “thermal boundary layer” there is no clear consensus in the literature. Some authors refer with this term only to the transition zone between lithosphere and asthenosphere and others to the entire lithosphere as being a thermal boundary layer to earth (s. p. 89, p. 64 and Fig. 2.19).

One thing can be said with certainty: the definition of the lithosphere depends on the question that is being asked. For example, it can be shown that the thickness of the lithosphere is a function of the observed time scale. Seismic motion, isostatic uplift and ductile deformation occur on time scales of seconds,  $10^4$  y and  $> 10^6$  y, respectively. The larger the time scale of the process, the smaller the thickness of the lithosphere. Seismically, the lithosphere is of the order of 200 km thick, while the elastic thickness of the lithosphere is only some tens of kilometers thick. Very generally the lithosphere may be defined *mechanically* as the outer part of the earth in which stresses can be transmitted on geological time scales (s. McKenzie 1967). According to a somewhat different mechanical definition the thickness of continental lithosphere may be defined as the layer that is in isostatic equilibrium with the mid-oceanic ridges (Cochran 1982). This is meaningful, because the mid-

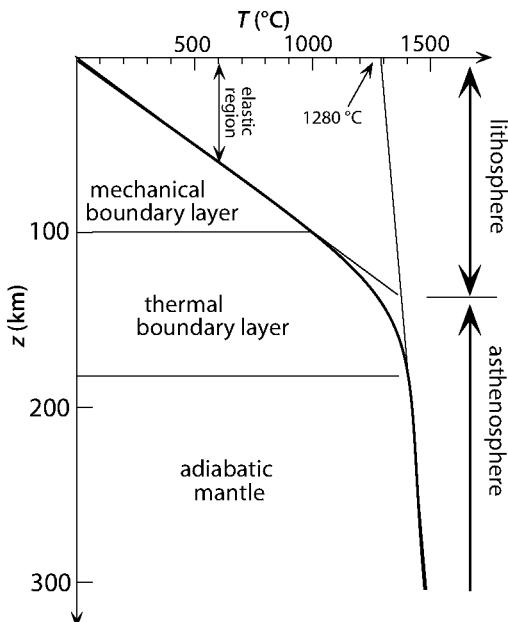


**Figure 2.18.** **a** Temperature and **b** density of the continental lithosphere as a function of depth. The depth of the Moho below surface is  $z_c$ , that of the whole lithosphere is  $z_l$ . **a** The curvature of the geotherm within the continental crust is caused by radioactive heat production. The light shaded area corresponds to the heat content of the lithosphere that can be attributed to radioactive decay in the crust. The dark shaded region corresponds to the heat content conducted into the lithosphere from the asthenosphere. The relative contributions of the radioactive and the mantle heat flow will be discussed in sect. 3.4.2 and sect. 6.2.1. Here it may be seen that the heat content of the crust consists to roughly equal part of mantle heat and of radioactive heat. **b** The slope of the density profile within the crust and within the mantle lithosphere is a function of the thermal expansion. A comparison of the shaded areas shows that the density deficiency in the crust (light shaded area) is comparable to the density excess in the mantle lithosphere (dark shaded area) – both relative to the asthenosphere. Within the asthenosphere convective flow equalizes all density and temperature heterogeneities. Both curve are therefore vertical

oceanic ridges may be interpreted as manometers of the upper mantle (s. p. 166, Turcotte et al. 1977).

According to a *thermal* definition the lithosphere is the part of earth in which thermal energy is largely transferred by heat conduction, in contrast to the asthenosphere, where heat is transferred by convection (s. sect. 3.4 for more detail). In some ways the thermal definition encompasses the mechanical definition because many of the mechanical properties of rocks depend on the ratio of their temperature to their melting temperature. In stable continental lithosphere, thermal and mechanical definitions indicate thicknesses of 100–150 km. The thickness of the crust and its content in radioactive minerals is crucial to the thickness of the lithosphere, because they strongly influence the Moho-temperature (s. Fig. 3.20). Many studies that are concerned with the rheology of the lithosphere avoid to use a number for lithospheric thickness. Rather, the lithosphere is defined via the Moho-temperature (s. sect. 6.2.2).

**Figure 2.19.** Definition of the mechanical and thermal boundary layer as well as of the lithosphere and asthenosphere according to Parsons and McKenzie (1978) and McKenzie and Bickle (1988). The thick line is the geotherm. Its curvature in the upper crust is ignored. However, note that other definitions for thermal boundary layer are also in use in the literature s. p. 89, p. 64)



**Types of Lithosphere.** There are two fundamentally different types of lithosphere on earth: *oceanic* and *continental* lithosphere. Despite these names, the correlation of oceanic lithosphere with the geographic area of the ocean and vice versa is only very approximate and there are substantial parts of continental lithosphere that lie under water and vice versa. A very rough indicator for the line separating continental from oceanic lithosphere is about 500 m water depth. On Fig. 2.1 most of the white colored regions within the oceans are actually continental lithosphere.

- *Oceanic lithosphere.* Oceanic lithosphere begins its life at the mid oceanic ridges. There, it consists only of an about 7 km thick oceanic crust, which is made up of crystallized partial melts from the uppermost mantle. The thickness of the mantle part of the oceanic lithosphere is zero near the mid-oceanic ridges. With increasing age – that is: with increasing distance from the ridge – the thickness of the mantle part of the oceanic lithosphere increases as the asthenosphere successively freezes to the base of the cold crust. In the oldest parts of known oceanic lithosphere the thickness of the oceanic mantle lithosphere is almost as thick as continental mantle lithosphere. However, oceanic lithosphere is being produced and consumed at all times, so that there is hardly any oceanic lithosphere on earth that is much older than about 150 my. Because of the young age (and therefore low thickness) the mean oceanic heat flow is significantly higher than that of the continents. Although there is practically no radiogenic heat production in oceanic crust

the mean oceanic heat flow is  $101 \text{ mWm}^{-2}$  (Pollack et al. 1993; Wei and Sandwell 2006) (s. p. 17, Figs. 3.1, 2.4).

- *Continental lithosphere.* In contrast to the oceanic lithosphere, the total area of continental lithosphere has remained largely constant in the entire Phanerozoic. Thus, the present day continents consist largely of Proterozoic continental lithosphere, which has been reworked in many places. Nevertheless there are enough places around the world where Archaean and Proterozoic continental crust is preserved giving us a glimpse of tectonic processes in the Precambrian (e.g. Greenland, Canadian Shield, East Antarctic Shield, Kongo Craton, Arunta Block, Yilgarn block and many others). Continental crust is chemically highly-differentiated, it has a high content in radioactive elements and in its stable state it is about 30–50 km thick. According to thermal and mechanical definitions, the mantle part of the lithosphere is 70–100 km thick, so that the total thickness of stable continental lithosphere is of the order of 100–150 km. However, in old shield areas this thickness can be much more, probably due to a different chemical composition of the mantle lithosphere. The mean heat flow of the continental lithosphere is about  $65 \text{ mWm}^{-2}$  (Pollack et al. 1993; Wei and Sandwell 2006).

#### 2.4.2 The Lithospheric Plates

The surface of the earth can be divided into seven major lithospheric plates plus a number of smaller plates (Fig. 2.20, Table 2.2). Not all major plates correspond to the seven continents and it is of some coincidence that the number of continents equals that of the major plates. Most major plates consist of both continental *and* oceanic lithosphere.

Among the seven major plates, the Antarctic Plate and the African Plate have a special position as they are surrounded on just about all sides by mid-oceanic ridges. Both plates increase therefore permanently in size and they form a good example where plates go from a compressive state into an extensional state purely as a function of their increasing age (Sandiford and Coblenz 1994; sect. 5.3.1).

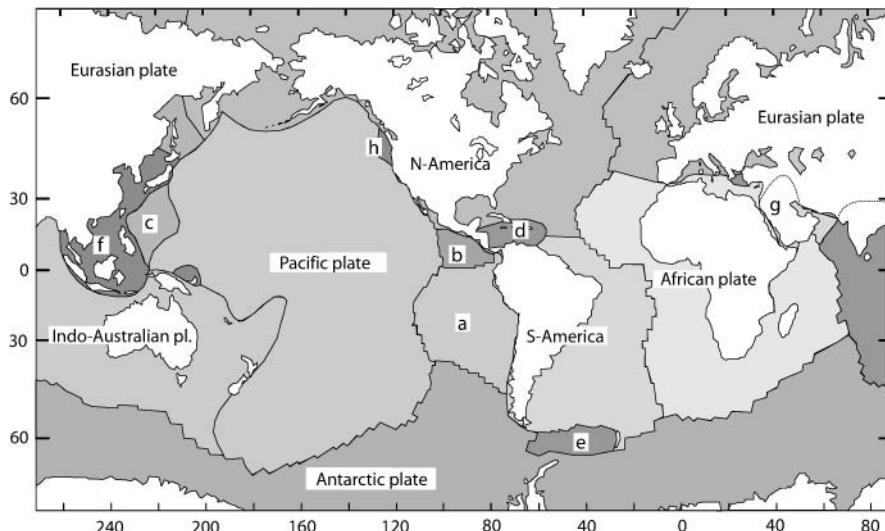
#### 2.4.3 The Plate Boundaries

Most geodynamically interesting processes occur along the plate boundaries. These boundaries can be divided according to:

- their kinematics,
- the types of plates that are in contact.

When choosing a division according to the types of bounding plates, we can discern:

- plate boundaries between two continental plates,
- plate boundaries between two oceanic plates,



**Figure 2.20.** Plate tectonic division of the surface of earth. Continents are white and oceans are shaded. The difference between continental and oceanic lithosphere is not shown. For that, see Fig. 2.1. Note that the plate boundaries coincide only in a few places with the coast lines that delineate the continents (s. Table 2.2). The seven major plates are labeled with their names. The most important minor plates are labeled with letters. They are: *a* Nazca Plate; *b* Cocos Plate; *c* Philippine Plate; *d* Caribbean Plate; *e* Scotia Plate; *f* Chinese Subplate; *g* Arabic Plate; *h* Juan-de-Fuca Plate

- plate boundaries between a continental and an oceanic plate.

When choosing a division according to the kinematics, we can discern between *convergent*, *divergent* and *transform* plate boundaries (s. Tables 2.3; 2.4). Passive margins are formerly divergent plate boundaries between two continental plates which now consist of a passive contact between oceanic- and continental lithosphere. They are often listed as its own type of plate margin.

**Divergent Plate Boundaries.** Divergent plate boundaries are regions where two plates move into opposite directions or where one plate is splitting into two. Divergent plate margins exist only between two continental plates (e. g. central African rift system) or between two oceanic plates (e. g. Mid-Atlantic Ridge). The passive seams between continental and oceanic lithosphere are mechanically very strong and it would be a great coincidence if a divergent plate margin would form exactly along them. However, there are places on the globe where divergent plate margins *cross* passive margins. The Sheba Ridge in the Gulf of Aden and the Carlsberg Ridge in the Indian Ocean are examples. Divergent plate boundaries on the continents are called rifts. The best known examples (in order of progressive rift development) are the Rheingraben, the Central African rift system and the Red Sea. Divergent

**Table 2.2.** Approximate parts of oceanic and continental lithosphere of each of the major plates

plate	% oceanic lith.	% continental lith.
<i>major plates</i>		
Pacific Plate	100	0
North-American Plate	30	70
South-American Plate	50	50
Eurasian Plate	30	70
Antarctic Plate	50	50
African Plate	50	50
Indo-Australian Plate	40	60
<i>important minor plates</i>		
Nazca Plate	100	0
Cocos Plate	100	0
Juan-de-Fuca Plate	100	0
Scotia Plate	100	0
Philippine Plate	100	0
Caribbean Plate	100	0
Arabic Plate	10	90

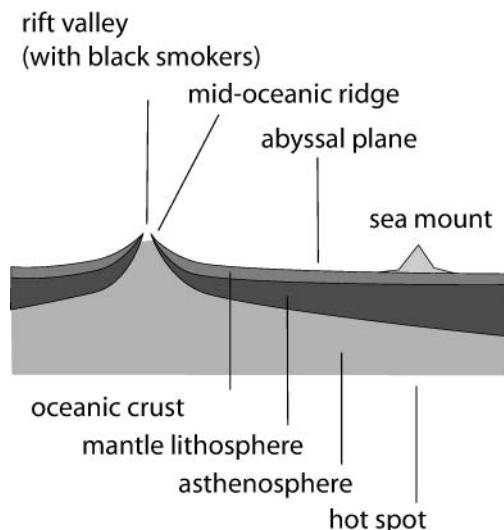
**Table 2.3.** The twelve most important relative motions of plates (after DeMets et al. 1990). 1° corresponds to about 110 km

plate boundary	rotation pole longitude	latitude	angular velocity · 10 <sup>-70</sup> /y
Africa – Antarctica	5.6°N	39.2°W	1.3
Africa – Eurasia	21.0°N	20.6°W	1.3
Africa – North-America	78.8°N	38.3°E	2.5
Africa – South-America	62.5°N	39.4°W	3.2
Australia – Antarctica	13.2°N	38.2°E	6.8
Pacific – Antarctica	64.3°S	96.0°E	9.1
South-America – Antarctica	86.4°S	139.3°E	2.7
India – Eurasia	24.4°N	17.7°E	5.3
Eurasia – North-America	62.4°N	135.8°E	2.2
Eurasia – Pacific	61.1°N	85.8°W	9.0
Pacific – Australia	60.1°S	178.3°W	11.2
North-America – Pacific	48.7°N	78.2°W	7.8

plate margins between two oceanic plates (mid-oceanic ridges) are – in most cases – the last stage of a rift (sect. 2.4.4, Fig. 2.21).

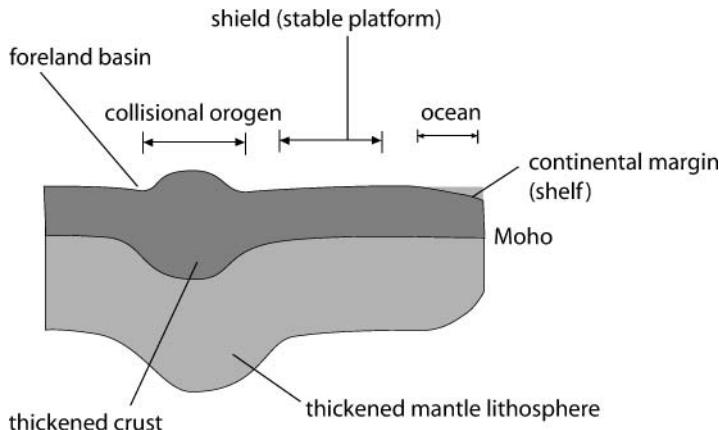
**Convergent Plate Boundaries.** Convergent plate boundaries may form between two continental plates (Fig. 2.15), between two oceanic plates or between a continental and an oceanic plate. In convergent plate boundaries

**Figure 2.21.** Nomenclature of some of the most important parts of oceanic lithosphere



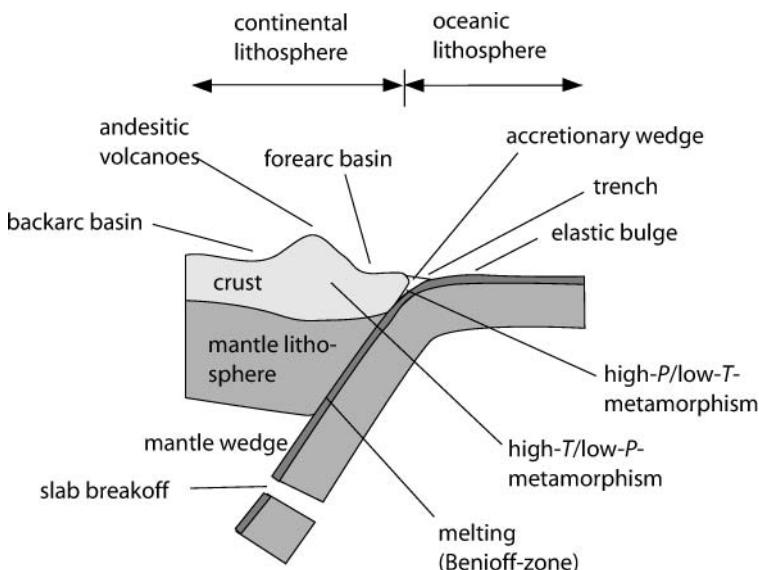
between continental and oceanic plates, the oceanic plate dives beneath the continent, because of its higher density. This is called subduction and its surface expression is a trench (Fig. 2.23). The most famous example for subduction is the subduction of the Nazca Plate underneath the south-American continent along the Peru-Chile Trench. Subduction leads to high pressure metamorphism in the hanging wall of the subducted plate. This metamorphism is associated with dehydration and partial melting of the plate of the oceanic plate in the *Benioff zone*. Fluids that rise through the overlying mantle wedge react with the wedge material in endothermic reactions (Fig. 3.28). However, additional heat input by convection in the wedge leads to partial melting (Hoke et al. 1994) and ultimately to the development of volcanic arcs on the surface.

The kinematics of subduction zones is complicated. The forces and velocities with which subducting plates sink into the asthenosphere are comparable to the forces exerted by mid-oceanic ridges onto the plate. Subduction zones can therefore move *backwards* (towards the mid-ocean ridge) if the downward velocity of the subducting plate is *larger* than the rifting rate at the ridge (e. g. South-Georgia, Scotia Plate). They can move *forward* (towards the continent) if the rifting rate at the ridge is *larger* than the downward velocity (e. g. Pacific Plate – Alaska). In other words, the distance between the trench and the continent in the far field hinterland increases, decreases or remains constant. Depending on details of the force and velocity field in subduction zone environments, forearc- or backarc basins may develop. In some cases of collision between oceanic and continental lithosphere, parts of the oceanic plate are welded onto the continental plate or even thrust over it. This is called *obduction*.



**Figure 2.22.** Nomenclature of important parts of continental plates

In contrast to subduction zones – where one of the two plates usually dives underneath the other – the convergence of two continental plates leads to a



**Figure 2.23.** Nomenclature of important parts of subduction zones. Forearc- and backarc basins are even more clearly developed during subduction of oceanic lithosphere underneath another oceanic plate. Good descriptions of various phenomena on this figure can be found for slab break off by: Blanckenburg and Davies (1995); about the mantle wedge by: Spiegelman and McKenzie (1987) (sect. 3.5.2); about metamorphism by: Miyashiro (1973); about accretionary wedges in: sect. 6.2.3; about the elastic fore bulge: sect. 4.4.2

head-on collision of both plates. The reason for this is because continental lithosphere is (i) much thicker, (ii) less dense and (iii) much softer than oceanic lithosphere. Why the continental lithosphere is much softer than the oceanic lithosphere although it is much thicker will be discussed in sect. 5.2. This leads to the formation of the collisional mountain ranges that form most of the topographically high mountain belts of our globe (Fig. 1.1). During this head on collision the crust gets typically pervasively deformed, while the mantle parts of the lithosphere override each other (a bit like shown schematically on the cover of this book).

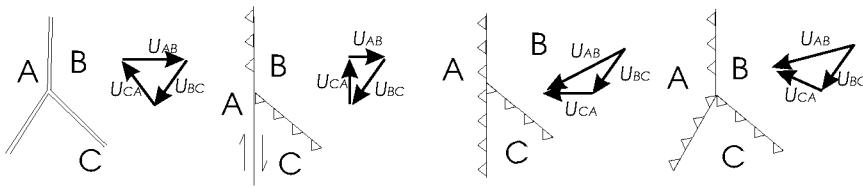
When two oceanic plates converge, no collision occurs and subduction zones form, similar to those that form when two plates of different kind collide. In contrast to the collision between two continental plates, *no* collision occurs between two oceanic plates because they are thinner, much stronger and because they are much denser and can therefore dive easier into the upper mantle (sect. 5.2.2). Because little internal deformation of the plates occurs, island arcs form that are clearly defined in space. Two beautiful examples for this are the subduction of the Pacific Plate underneath the Philippine Plate along the Mariana Trench or the subduction of the Pacific Plate underneath North-America along the Aleutes.

**Figure 2.24.** Topography of the Dragonsback Ridge segment of the San Andreas Fault in California. The shown segment is about 3 km long and the resolution is less than 0.5 m. The fault is oriented northwest - southeast (top left to bottom right). Image was obtained by the B4 Project using Airborne Laser Swath Mapping (Hudnut and Bevis, pers. comm., 2006)



**Transform Plate Boundaries.** When two plates glide past each other without much convergence or divergence, their contact is called a transform plate boundary. No topographic features as significant as rift valleys or mountain ranges form. However, transform plate boundaries are well-known because they form some of the most important zones of seismicity on the globe. The best known examples are the San-Andreas-fault zone (Fig. 2.24), (probably because it crosses one of the most densely populated parts of North-America) or the Alpine fault in New Zealand.

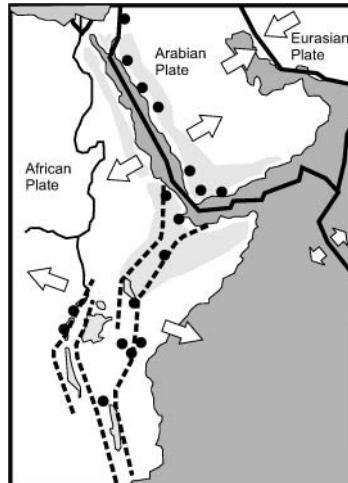
**Triple Junctions.** The spherical geometry of the earth requires that three or more plates touch each other in some places. This is called a triple junction.



**Figure 2.25.** Cartoons of four types of triple junction and vector diagrams showing their relative velocities. The shown examples are all stable because the sum of the vectors results in zero velocity

Triple junctions have a *stable* configuration if the relative motions of the bounding plates can maintain the geometry through time. They have an *unstable* configuration if their geometry is transient. Places where four or more plates touch are always unstable and will quickly resolve into two or more triple junctions. Such an area occurs currently west of New Guinea, where the Philippine, Australian, Eurasian and Pacific Plates meet. However, detailed mapping shows that this area may be divided into a series of micro plates and triple junctions and that the touching of more than three plates has only occurred transiently. Depending on the kinematics of plate boundaries, we can discern between a large number of different triple junctions. Using “R” (as in Ridge) for divergent plate boundaries, “T” (as in Trench) for convergent plate boundaries and “F” (as in Fault) for transform boundaries, we can describe RRR-, TTT-, FFF-, RTF- and a number of other triple junctions (McKenzie and Morgan 1969). Four examples are shown in Fig. 2.25. The examples shown there are all stable because the sum of the velocity vectors ( $U$ ) is zero in all cases.

**Figure 2.26.** Schematic map of the most famous triple junction on earth: the meeting point of the East African rift system (dashed) with the Red Sea rift. The black dots are volcanoes, the light shaded areas show the rift flank uplift



**Table 2.4.** Different types of plate boundaries divided according to their kinematics and according to the plate type ( $O$  = oceanic lithosphere,  $C$  = continental lithosphere)

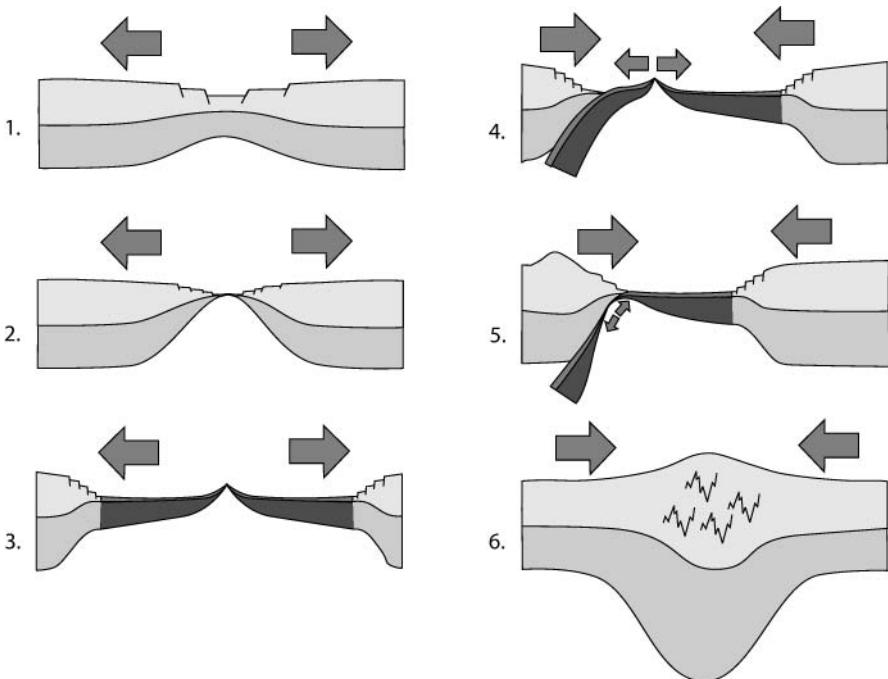
relative motion	type	plate tectonic feature	example
convergent	O-O	island arc	Philippines
	O-C	subduction zone	west coast of
	C-C	trench	South-America
divergent	O-O	collisional mountain belt	Himalaya
	C-C	mid-oceanic ridge	Atlantic
passive	O-C	continental rift	east African rift
		passive plate margin	eastern Australia

One of the most famous triple junction is the RRR triple junction between the Gulf of Aden, the Red Sea and the Central African rift system (Fig. 2.26). With this, as with other RRR triple junctions, the three different branches occur in various stages of development and, in fact, one arm may stop rifting leaving behind a failed rift and a single plate boundary. A beautiful example of this is given by the failed Lambert Graben rift in Antarctica which stopped rifting in the Cretaceous while the ridge in the southern Indian Ocean continued to separate Australia from Antarctica.

#### 2.4.4 The Wilson Cycle

The Wilson-cycle is a model that brings the individual processes that we discussed in the last sections into an imaginary cycle (Fig. 2.27). This cycle was first suggested by Wilson (1972) and supersedes the old terms of “geosynclines”.

The Wilson-cycle begins with extension of a continent. This stage of the Wilson-cycle can be currently observed in the Rheingraben. In the second stage, the continent breaks up and a spreading center forms in its middle. Oceanic lithosphere begins to form. This is the stage of the Red Sea. The third stage involves the development of a classic ocean with passive continental margins on both sides and a mid-oceanic ridge in its center. This is currently being observed in the Atlantic. The following stages describe processes that occur during opposite relative plate motion. These stages may therefore occur independent of the first three stages. The fourth stage describes the commencement of subduction of an oceanic plate underneath a continental plate as it is currently observed along the west coast of South-America. The fifth stage involves the subduction of a mid-oceanic ridge underneath the continental plate. A present day example for this is the subduction of the



**Figure 2.27.** The Wilson cycle. The arrows indicate the relative plate motions

ridges bounding the Juan de Fuca Plate underneath North-America. During the sixth and final stage the collision between two continents occurs. Clearly, the India-Asia collision is the most spectacular example for that.

## 2.5 Problems

### Problem 2.1. Small circles and subduction zones (p. 22):

Make a very rough estimate of the small circle radius of the Aleute arc (from Fig. 2.5 or any atlas) and estimate the subduction angle of the Pacific Plate underneath Alaska using the model of Fig. 2.6. This subduction angle is actually known. It is  $\approx 45^\circ$ . Discuss possible reasons for the difference between your estimate and this value. Compare your estimate with an estimate of the subduction angle of the Indoaustralian Plate underneath the Eurasian Plate along the Java Trench.

### Problem 2.2. Understanding longitude and latitude (p. 23):

Find a point on the surface of earth (outside the north pole!) where the following experiment is possible: You walk 1 km south then 1 km west and then 1 km north and you are where you started from. (It is said that this

question was asked by Sir Ernest Shackleton in 1908 to applicants for his south pole expedition.)

**Problem 2.3.** *Understanding longitude and latitude (p. 23):*

The city of Vienna is located at  $16^\circ$  east and  $48^\circ$  north; the city of Munich is located at  $11^\circ$  east and the same latitude as Vienna. How much earlier does the sun rise in Vienna?

**Problem 2.4.** *Understanding arc-minutes in decimals (p. 23):*

Improve your result of Problem 2.3 using the following values: Vienna:  $16^\circ 21'$  east,  $48^\circ 11'$  north; Munich:  $11^\circ 39'$  east and the same latitude as Vienna.

**Problem 2.5.** *Estimating the curvature of earth (p. 24):*

a) How far is it from Vienna to Munich going directly from east to west along the surface of the globe? (use eq. 2.2). b) How far is it from Vienna to Munich going the shortest distance along a great circle? (use the formula for angular separation of 2 points in Table C.4). c) How far is it from Vienna to Munich if you could go along a tunnel connecting the two cities in a straight line through the earth? (use the result from b) and the logic illustrated in Fig. 2.6). Use the longitudes and latitudes given in Problem 2.3 and a perfectly spherical earth with radius  $R= 6\,370$  km.

**Problem 2.6.** *Spherical kinematics (p. 25):*

What is the west-east velocity of a plate that has rotation pole that coincides with the geographic north pole and rotates relative to another plate with an angular velocity of  $10^{-15}\text{s}^{-1}$ ; a) at the equator; b) at  $48^\circ$  north? (Use  $R= 6\,370$  km).

**Problem 2.7.** *Understanding torque (p. 25):*

A north-south striking mid-oceanic ridge has a west-directed force of  $10^{12}$  N. At which latitude is the torque of this ridge around the axis of the earth  $4 \cdot 10^{18}$  Nm? (The radius of earth is:  $R= 6\,370$  km)

**Problem 2.8.** *The Mercator projection (p. 29):*

Derive the function that is used in the Mercator projection to convert longitude  $\phi$  and latitude  $\lambda$  into the Cartesian coordinates  $x$  and  $y$ . Detailed study of Fig. 2.10 is helpful to answer this problem.

**Problem 2.9.** *Physics of the lithosphere (p. 37):*

Put numbers on the axes of the diagrams on Fig. 2.18. (This problem requires the use of parameter values that have not been discussed up to this chapter yet. It is meant as a self-test for the interested reader.)

**Problem 2.10.** *Modern and ancient plate boundaries (p. 40):*

a) Draw relative velocity vectors between the plates onto Fig. 2.20. b) Draw three future new plate boundaries and three plate boundaries that will disappear within the next 50 my on Fig. 2.20. c) Draw three ancient plate boundaries onto Fig. 2.20.

**Problem 2.11.** *Understanding triple junctions (p. 44):*

How many different types of triple junctions exist, using the three possible relative motions R, T and F? Not all of the triple junctions you should come up with are stable. Draw an example of a stable and an unstable triple junction.

### 3. Energetics: Heat and Temperature

In this chapter we discuss geodynamic processes that may be described with the units of energy or temperature. This general theme is an obvious starting point in geodynamics, as so many properties of rocks that may be observed in the field are a strong function of temperature, for example the formation of metamorphic parageneses or the mechanisms with which rocks deform.

Energy has the unit Joule [J], which is equivalent to mass  $\times$  velocity<sup>2</sup> ( $1 \text{ J} = 1 \text{ kg m}^2 \text{s}^{-2}$ ), or volume  $\times$  pressure ( $1 \text{ J} = 1 \text{ m}^3 \times \text{Pa} = 1 \text{ N m}$ , because  $1 \text{ N} = 1 \text{ kg m s}^{-2}$ , is also  $1 \text{ kg m}^2 \text{s}^{-2}$ ). These conversions give us a first indication that thermal and mechanical energy are often hard to separate. *Thermal* energy may be converted into temperature using heat capacity and density which are parameters that we shall discuss in some detail on the next pages.

The production and redistribution of heat in the lithosphere is done by three fundamentally different processes:

- heat conduction,
- heat advection (or convection) and
- heat production.

The relevance of these three processes for regional metamorphism is summarized at the start of sect. 3.4 in Table 3.4. However, in the first part of this chapter we will present some basic methods how to estimate if and how important each of these three processes may be for the thermal budget of the lithosphere. In this context we also discuss the basics of their mathematical description. The importance of such estimates is enormous: it will enable the reader to critically evaluate in his or her metamorphic study terrain to which degree heat conduction (e. g. because of burial), heat production (e. g. friction heat or radioactivity) or active transport of heat (e. g. by fluid or magma) may have played a role during metamorphism. Thus, the information on the following pages provides a powerful tool for the interpretation of heat sources for field geologists.

## 3.1 Principles of Heat Conduction

### 3.1.1 The Heat Conduction Equation

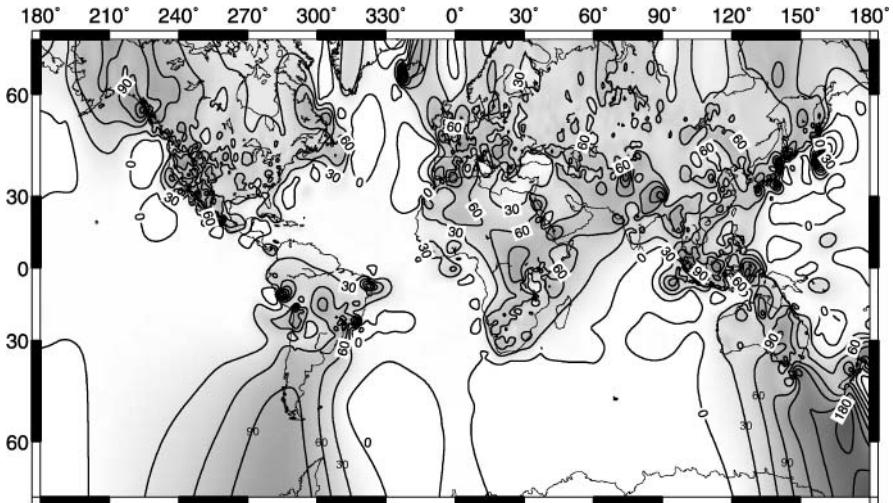
The heat conduction equation - more commonly known as the diffusion equation - is fundamental for the understanding of the transport of heat in the lithosphere. We will also show in other chapters that the very same equation cannot only be applied to the transport of thermal energy, but also to the diffusion of mass. It finds therefore application in many other fields, for example geomorphology, metamorphic petrology or hydrology (s. p. 185; p. 343). Thus, the diffusion equation is the first equation of this book that we will discuss in some detail. The fact that it is a second order partial differential equation should not scare us off. We will show that it is possible to understand it quite intuitively. There is also some explanations on how to read differential equations in section B.1.

**Fourier's Law of Heat Conduction.** Fourier's 1. law is the basic law underlying the diffusion equation. This law states that the flow of heat  $q$  is directly proportional to the temperature gradient (Fourier 1816). This statement can easily be formulated in an equation:

$$q = -k \frac{dT}{dz} . \quad (3.1)$$

In this equation  $q$  is short for heat flow,  $T$  stands for temperature and  $z$  for a spatial coordinate, for example depth in the crust. The ratio  $dT/dz$  is the change of temperature in direction  $z$ . We call this ratio the temperature gradient.  $k$  is the proportionality constant between the gradient and the flow of heat. In order to understand this law better (and understand the units of  $k$ ), let us consider a more familiar analogue: the flow of water in a river. The same law applies. In a river the flow of water can be described by the volume of water passing per unit of time and per area of cross section of the river (in SI-units:  $\text{m}^3 \text{s}^{-1} \text{m}^{-2} = \text{m s}^{-1}$ ). This is called the volumetric flow. When normalized only to the width of the river and not to the cross sectional area of the river, the volumetric flow has the units of  $\text{m}^2 \text{s}^{-1}$  (sect. 4.5). In contrast, the flow of mass has the units  $\text{kg s}^{-1} \text{m}^{-2}$ . Fourier's law – applied to our example of water flow – states that the flow of water is proportional to the topographic gradient of the river. This corresponds well to our observations in nature: The steeper a river bed, the faster the flow of water in the river (per square meter of cross sectional area). Fourier's law seems to be a good model description for this observation. This simple example also explains why there is a negative sign in eq. 3.1. The flow is against the gradient: it is *positive* in the *downwards* direction of the gradient.

In the theory of heat conduction, the flow of heat has obviously not the units of volume per time and area, but *energy* per time and area. (in SI units:  $\text{J s}^{-1} \text{m}^{-2} = \text{W m}^{-2}$ ). The thermal gradient now replaces the topographic gradient of the river. Because of historical reasons heat flow is sometimes given

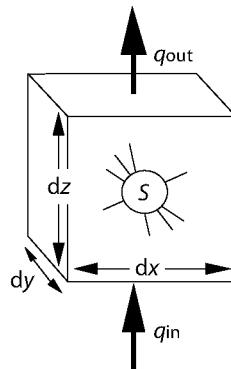


**Figure 3.1.** Map of global surface heat flow. The map is contoured in milli  $\text{W m}^{-2}$  and data above  $300 \text{ mW}^{-2}$  are omitted so that only a conductive response is shown. Nevertheless, because of the highly irregular distribution of the data, the contouring on a global scale is strongly dependent on the contouring algorithm and is associated with large uncertainties in areas of low data density. (made with the Global Heat Flow Data Base of the International Heat Flow Commission, W. Gosnold, Custodian pers. comm. 2006). As a consequence, this map is exclusively a representation of the data base with no consideration of the data distribution, its density or their reliability. As such, it appears (wrongly) that the heat flow on the oceans is lower than that of the continents. For a more considered heat flow map see Pollack et al. (1993) or Wei and Sandwell (2006)

in heat flow units, or hfu. One hfu corresponds to  $10^{-6} \text{ cal s}^{-1} \text{ cm}^{-2}$  and can easily be converted into  $\text{W m}^{-2}$  (s. Problem 3.2, Table D.8). The units of the proportionality constant  $k$ , in eq. 3.1, follows now easily from the units of the other components of the equation: Because temperature has the units of K (or  $^{\circ}\text{C}$ ) and  $z$  has the unit m,  $k$  must have the units  $\text{J s}^{-1} \text{ m}^{-1} \text{ K}^{-1}$  so that the equation is consistent in its units. The constant  $k$  is called *thermal conductivity*. We can now try to read eq. 3.1. We can see that the flow of heat trends to zero if the conductivity is very low, regardless of the thermal gradient. Correspondingly, if the conductivity is very large, the flow of heat becomes large, even if the thermal gradient is very low. The equation may therefore be understood quite intuitively.

Would the thermal gradient be constant everywhere, we could write it as  $\Delta T / \Delta z$ . However, in geological problems this gradient is never constant. Thus, we use the derivative  $dT/dz$ , which states that we want to be careful and consider our thermal gradient only to be constant within each infinitely small section of the thermal profile. If the gradient changes along the  $z$  direction, then eq. 3.1 states that the heat flow must also change.

**Figure 3.2.** The flow of heat in a unity volume of rock. The heat production inside this volume  $S$ , is not considered until we discuss eq. 3.23



**Energy Balance.** The second part of the diffusion equation (often called Fourier's 2. law) describes an energy balance. This energy balance relates heat and temperature and the change of heat flow with change in temperature. This relationship may be established independently from eq. 3.1 and may be written as:

$$\frac{\partial T}{\partial t} \propto -\frac{\partial q}{\partial z} \quad (3.2)$$

This equation states that the rate of temperature change of a rock must be proportional to the rate with which its heat content changes ( $\propto$  is the symbol for "proportional to"). The rate with which the heat content of a rock changes ( $\partial q / \partial z$ ) is given by the difference between the flow of heat *into* the rock and the flow of heat *out of* the rock (Fig. 3.2). If the heat flow into the cube of Fig. 3.2 is larger than the flow of heat out of it, then the heat content of the cube will rise and its temperature will increase. If the heat flow into the volume is just as large as that that flows out, the temperature will remain constant. If more heat flows out of the cube than into it, then its temperature will decrease.

In the last sentences we have begun mixing the terms "temperature" and "heat". However, we have to remain careful no to confuse them as the rate of temperature change is not the same as the rate of heat content change. They relate by:

$$H = T \rho c_p \quad (3.3)$$

where  $H$  is the volumetric heat content in  $\text{J m}^{-3}$  (s. sect. 3.6.4). The rate, with which the temperature will change for a given change in heat content depends on another material specific proportionality constant. This is the *specific heat capacity*  $c_p$ . The specific heat capacity or short "specific heat" has the units of  $\text{J kg}^{-1} \text{K}^{-1}$  and defines how many Joules are required to heat the mass of one kg of rock by one degree Kelvin. The most common abbreviation for specific heat is  $c$ . The subscript  $p$  symbolizes the condition that the specific heat is measured at constant pressure (s. sect. 3.2.2). If the

specific heat of a rock is large, we need many Joules to heat the rock and even a rapid increase of its heat content will lead to slow temperature increase and vice versa. Specific heat is formulated in terms of the *mass* that is heated. Considering that the energy balance in eq. 3.2 is formulated in terms of the *spatial* coordinate  $z$ , and heat capacity is formulated in terms of *mass*, we need to multiply  $c_p$  with the density  $\rho$ , so that the relationship between the spatial change of *heat flow* and the temporal change of *temperature* is consistent with the units. We can write the proportionality of eq. 3.2 as:

$$\rho c_p \frac{\partial T}{\partial t} = - \frac{\partial q}{\partial z} . \quad (3.4)$$

It should now be straight forward to understand eq. 3.4 intuitively using Fig. 3.2. The negative sign arises because the temperature *increase* when  $\partial q = q_{\text{out}} - q_{\text{in}}$  is negative, that is, more heat flows into the rock volume than out of it. You may have noticed that the step from eq. 3.2 to eq. 3.4 was accompanied by the change from total- to partial differentials. This was necessary, because different parts of this equation are now differentiated with respect to different parameters (s. sect. B.1.1).

**The Diffusion Equation.** If we substitute Fourier's law of heat conduction (eq. 3.1) into the thermal energy balance of eq. 3.4, we arrive at:

$$\rho c_p \frac{\partial T}{\partial t} = \frac{\partial (k \frac{\partial T}{\partial z})}{\partial z} . \quad (3.5)$$

Eq. 3.5 is the general form of the one-dimensional diffusion or heat conduction equation. If  $k$  is independent of  $z$  (e.g. if we consider heat conduction in an area without lithological contrasts), it is possible to simplify eq. 3.5 significantly.  $k$  can then be taken out of the differential and we can write:

$$\rho c_p \frac{\partial T}{\partial t} = k \frac{\partial \partial T}{\partial z \partial z} \quad \text{or:} \quad \frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} . \quad (3.6)$$

The constants  $k$ ,  $\rho$  and  $c_p$  are now summarized to  $\kappa = k/(\rho c_p)$ .  $\kappa$  is called thermal diffusivity. Eq. 3.6 can also be understood intuitively, without following the detailed derivation given above. Eq. 3.6 may be formulated in words as:

- *The rate of temperature change is proportional to the spatial curvature of the temperature profile.*

If you do not understand the relationship between this sentence and eq. 3.6, then remember that the first differential of a function describes its slope (or: "gradient", or: "rate") and the second its curvature (s. sect. B.1, Fig. B.3, B.2).

Figure 3.3 illustrates this graphically. In our daily lives we encounter many examples that are described by this equation. Think for example that a piece of toast cools much quicker on its corners than along the edges or in its middle. This is because the *spatial curvature* of the isotherms in the toast

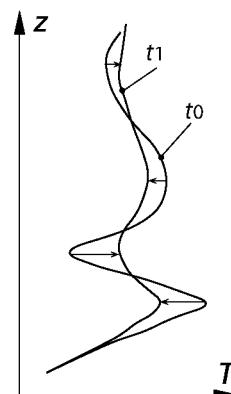
is the largest at the corners! The same is true for the rapid cooling of the tip of a needle, the rapid erosion of ragged mountain tops and countless other examples in nature, all the way down to the rapid chemical equilibration of fine grained rocks in comparison with coarse grained rocks.

If we want to use eq. 3.6 we must solve it. For this we need boundary- and initial conditions. We also need some mathematical knowledge so that we can integrate this equation. Various methods how to go about this are discussed in sect. B.1.1. A large part of this chapter will deal with various solutions of this equation. In this context we will often meet the terms “*boundary conditions*” and “*initial conditions*”. Make sure you understand what they mean (s. p. 398).

- *The magnitude of  $\kappa$ .* A quantitative application of eq. 3.6 requires the knowledge of  $\kappa$  and therefore the knowledge of  $k$ ,  $\rho$  and  $c_p$ . The specific heat of rocks is about  $c_p = 1\,000 - 1\,200 \text{ J kg}^{-1} \text{ K}^{-1}$  (Oxburgh 1980). For most rocks  $c_p$  does not vary by more than 20 % around this value. Thus, the nice and even value of  $c_p = 1\,000 \text{ J kg}^{-1} \text{ K}^{-1}$  is a sound assumption that can be used for many thermal problems. The density of many crustal rocks is of the order of  $2\,750 \text{ kg m}^{-3}$  and varies also not all that much around this value. However, thermal conductivity, varies by the factor 2 or 3 between different rocks types (Table 3.1). Fortunately, it is between 2 and  $3 \text{ Js}^{-1} \text{ m}^{-1} \text{ K}^{-1}$  for many rock types. For  $k = 2.75 \text{ Js}^{-1} \text{ m}^{-1} \text{ K}^{-1}$  and the values for specific heat and density from above the diffusivity is:  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ . Because this value is easy to remember it is commonly used in the literature. Note, however, that  $\kappa$  may also be twice- or half as large if the thermal conductivity of rocks is twice or half as large.

**Heat Refraction.** If rocks of different thermal conductivity are in contact, the phenomenon of *heat refraction* may occur. What this is, is easily explained with eq. 3.1. In thermal equilibrium, the flow of heat in two adjacent rocks must be equal. Following from eq. 3.1 we can formulate:

**Figure 3.3.** The thermal equilibration of a random temperature profile. The temperature profile is drawn at two different time steps  $t_0$  and  $t_1$ . Note that the largest change in temperature between the two time steps has occurred in those places of the profile where the curvature of the profile is the largest (s. eq. 3.4). Where the curvature of the profile is zero (at the inflection points) the temperature does not change at all



$$-q = k_1 \frac{\Delta T_1}{\Delta z} = k_2 \frac{\Delta T_2}{\Delta z} , \quad (3.7)$$

where the subscripts  $1$  and  $2$  denote two different rocks as shown in Fig. 3.4. We can see in this equation that, if the conductivities  $k_1$  and  $k_2$  are different, the temperature gradient in the rock with the higher conductivity must be lower and vice versa (Fig. 3.4). This is called heat refraction. Eq. 3.7 can also be written in differential form. This means, the temperature gradient must not change abruptly, but can also change continuously, if there are continuous changes in thermal conductivity.

Let us illustrate the phenomenon with an example. A rock with extremely high thermal conductivity, for example an iron ore body, will be practically isothermal, even if it stretches over many vertical kilometers in the crust. Its high conductivity will cause it to adapt some average temperature. Thus, the upper part the body may have a significantly higher temperature than its surroundings while its lowest part is colder than its surroundings. As a consequence, it is conceivable that the process of heat refraction will even leads to some kind of contact metamorphism (s. Problem 3.3).

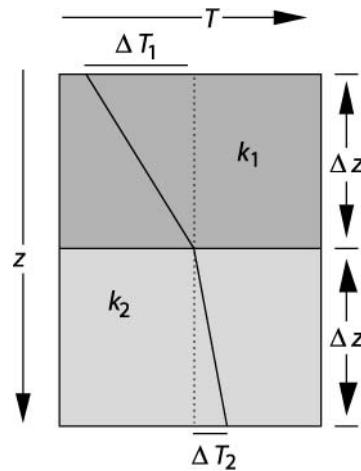
Jaupart and Provost (1985) have noticed that there are some important differences in thermal conductivity between the sediments of the Tethys zone and the high Himalayan crystalline complex. They suggested that the process of heat refraction may have been of relevance in connection with the melting of the Himalayan leucogranites. The process has also been discussed as the cause for high grade metamorphism in several Australian provinces (e.g. Mildren and Sandiford 1995) and as the trigger for a range of tectonic processes (Sandiford 1999). In fact, several exploration companies currently explore for geothermal energy sources in Australia by looking for regions where rocks of low thermal conductivity insulate underlying rocks of high thermal conductivity (Hillis et al. 2004).

If we want to describe the process of heat refraction quantitatively, we can not assume the simplification that we have made in going from eq. 3.5 to eq. 3.6. We must stick with eq. 3.5 to describe conductive equilibration. If we form the derivative of the right side of eq. 3.5 using the rules of differentiation of products (Table C.1) we get:

**Table 3.1.** Thermal conductivities and heat capacities of some rocks and common materials.  $k$  is given in  $\text{J s}^{-1} \text{m}^{-1} \text{K}^{-1}$  and  $c_p$  in  $\text{J kg}^{-1} \text{K}^{-1}$ . The change of thermal conductivity as a function of pressure and temperature are negligible at geologically relevant temperatures in the crust (Cull 1976; Schatz and Simmons 1972). Nevertheless, the numbers given here are only approximate

rock type	$k$	$c_p$
sandstone	1.5-4.2	920
gneiss	2.1-4.2	800
amphibolite	2.5-3.8	840
granite	2.4-3.8	790
ice	2.2	1 800
water	0.58	4 200
salt	5.4-7.2	880
iron	73	460

**Figure 3.4.** Illustration of the process of heat refraction. The flow of heat in the dark and the light shaded bodies is the same. However, the temperature gradient in the dark shaded body is larger, because its thermal conductivity  $k_1$  is smaller. The subscripts 1 and 2 denote the dark and the light shaded body, respectively



$$\rho c_p \frac{\partial T}{\partial t} = \frac{\partial (k \frac{\partial T}{\partial z})}{\partial z} = \frac{\partial k}{\partial z} \frac{\partial T}{\partial z} + k \frac{\partial^2 T}{\partial z^2} . \quad (3.8)$$

In this form the heat conduction equation can be used for the description of many problems with variable conductivity.

**Heat Conduction in Several Spatial Dimensions.** Eq. 3.6 is a linear differential equation (s. p. 396). This means that heat conduction in two or three dimensions can be described by the sum of heat conduction in the individual directional components. In general, we can write:

$$\frac{\partial T}{\partial t} = \kappa \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) \quad (3.9)$$

if the thermal diffusivity is the same in all three spatial directions. In the literature eq. 3.9 is often written as:

$$\frac{\partial T}{\partial t} = \kappa \nabla^2 T \quad (3.10)$$

The symbol  $\nabla$  is called the “Nabla-” or: “Del”-operator and is defined in eq. B.32.  $\nabla^2$  describes the same thing for the second partial derivative (s. sect. B.3). Partial differentials (or derivatives) are discussed in detail in section B.1.1. An important property of the diffusion equation is that it contains an energy balance. This means that no energy can be gained or lost by diffusion processes. If a rock cools by conduction, then it is by heat loss at the model boundaries, not because of the conduction itself.

- *Complicated geometries.* The diffusion equation we presented above is linear in temperature if all the material constants are independent of temperature and time. Therefore, it is possible to describe diffusion problems of complicated geometries simply by a linear superposition of the solutions for

more simple problems. We will use this approach extensively when we describe the cooling history of intrusions in sect. 3.6. For the same reason is it possible to describe time dependent problems by a superposition of a steady state solution and a time dependent component.

**Heat Conduction in Polar Coordinates.** Many heat conduction problems in the earth sciences are much better described in cylindrical or spherical coordinates. In *cylindrical* coordinates, eq. 3.6 adapts the form:

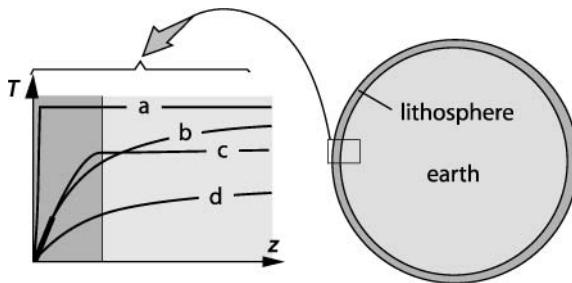
$$\frac{\partial T}{\partial t} = \kappa \left( \frac{\partial^2 T}{\partial r^2} + \frac{1}{r} \frac{\partial T}{\partial r} \right) . \quad (3.11)$$

If we consider heat conduction in *spherical* coordinates (and restrict ourselves to heat flow in the radial direction), then the heat conduction equation adapts the form:

$$\frac{\partial T}{\partial t} = \kappa \left( \frac{\partial^2 T}{\partial r^2} + \frac{2}{r} \frac{\partial T}{\partial r} \right) . \quad (3.12)$$

In eqs. 3.11 and 3.12  $r$  is the distance from the coordinate origin on the cylinder axis or at the sphere center, respectively. Detailed derivations of these equations will not be described here, but they can be found - among many others - by Carslaw and Jaeger (1959); Crank (1975) or Smith (1985). Examples for spherical conduction problems that can be described with eq. 3.12 are the Kelvin model for the cooling of the earth or the chemical diffusion of elements in garnet crystals. Note that eqs. 3.11 and 3.12 are one-dimensional and their results can therefore be directly compared with eq. 3.6 (e. g. Fig. 3.32 and 3.37).

**The Kelvin Model for the Cooling of Earth.** The most famous example for the application of the heat conduction equation is the estimate of the age of the earth by Lord Kelvin (1864). However, we note that very similar estimates were already performed by Fourier himself in 1820. Both physicists realized that - in principle - one could estimate the age of the earth from the present day thermal gradient at the surface (surface heat flow, s. Table D.3) if the following assumptions are made (Fig. 3.5): 1. The whole earth was at the time of its formation of constant temperature (which was assumed by Kelvin to be 4 000 °C); 2. the surface temperature has remained constant ever since and 3. heat conduction is the principle process of cooling. Using eq. 3.12, Kelvin concluded that the earth must be around 100 my old (he used a very similar solution to the one we introduce in sect. 3.6.1). Today we know that his calculation was wrong for two important reasons. The first (and most quoted) reason is that there are heat producing elements in the crust that prolong the cooling of earth. Radioactivity was unknown at Kelvin's time. The second reason is that convection in the upper mantle rises the isotherms. While this process actually leads to a faster cooling of earth, both *radioactivity* and *convection* cause the present day geothermal gradient in the lithosphere to be steeper than it would be without these two processes.



**Figure 3.5.** Schematic illustration of the model of Lord Kelvin for the estimation of the age of earth. The small  $T$ - $z$ -diagram is an enlarged section of outer 300 km of the globe. The shaded area is the lithosphere.  $a$  is the initial thermal profile of the earth at the time of its formation.  $b$  is the present day thermal profile that was calculated by Kelvin based on the measured surface heat flow  $q_s$  (thick drawn part on the thermal profile) and on a spherical heat conduction model. However, because of radioactivity (unbeknown to Kelvin), the lower lithosphere is much hotter than Kelvin thought possible. Moreover, in the asthenosphere convection destroys a conductive thermal gradient, which is very low in this part of earth. (convection was known at Kelvin's time but ignored in his estimates).  $d$  is the thermal profile that we should measure today, if the earth had only cooled by conduction since its formation at  $\approx 4.5 \cdot 10^9$  y. The surface heat flow would be much lower than what we measure today

This leads to a massive underestimate of the age of earth. Interestingly, the error that is caused in Kelvin's estimate by him neglecting convection is much larger than that caused by neglecting radioactivity. In 1895 John Perry improved Kelvin's estimates drastically, by assuming a "convective mantle conductivity" which he assumed to be ten times larger than the conductivity of the crust (s. p. V). With this assumption he arrived at an almost correct age.

Why Kelvin ignored convection in the mantle, although it had been known for some time, we do not know. However, in defense of his estimate it should be said that he *did* mention in his discussion that heat production, for example by chemical reaction, was not considered in his model but might change the results. The model of Kelvin remains a fantastic example for a conduction model that turned out to be completely wrong because other heat transport mechanisms were not considered. It should serve us as a reminder when interpreting the heat sources of a metamorphic terrain.

### 3.1.2 The Laplace Equation

Equations 3.6 and 3.10 describe the evolution of temperature as a function of time. Thus, we can use them to describe heating and cooling curves of rocks. However, for many geological questions we are not all this interested in the *temporal* changes of the temperature, but rather in the steady state shape of a temperature profile, for example the shape of a stable continental

geotherm. For all those problems, where it is reasonable to assume that there is *no* temporal change of a temperature profile, we can write:

$$\frac{\partial T}{\partial t} = 0 \quad . \quad (3.13)$$

Any geological situation where this equation applies is called the *steady state*. When we make this assumption,  $\kappa$  can be cancelled out of eq. 3.6 and eq. 3.10. We are left with:

$$\nabla^2 T = 0 \quad . \quad (3.14)$$

Eq. 3.14 is called Laplace-equation. Just like the diffusion equation, it is an extremely important equation for many geological problems. In this chapter we will need it when we consider stable geotherms, but it also has applications in many other branches of the earth sciences.

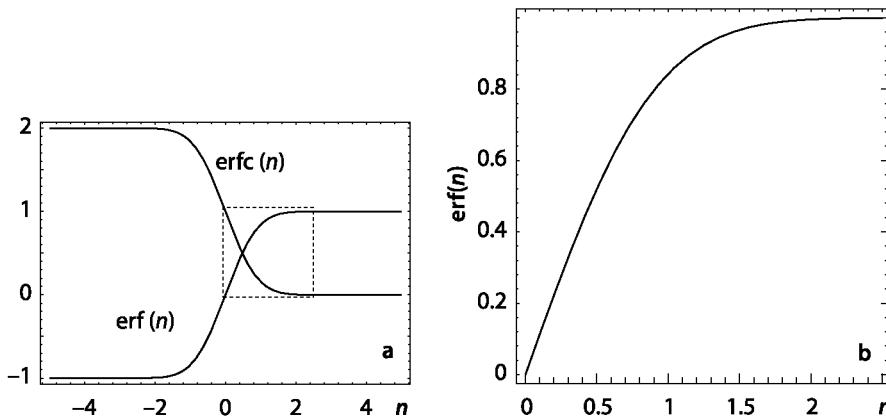
### 3.1.3 The Error Function

If we want to use eq. 3.6 to describe a time dependent conduction problem, we must solve it for a given set of boundary- and initial conditions. If we try this, we would quickly realize that this is only possible for a very few boundary- and initial conditions. Periodic problems are some of those for which there are “real” solutions of this equation (sect. 3.7.1). For most problems there are simply no solutions of eq. 3.6 possible. For example, for many geological problems we will see that it is useful to assume that the boundary conditions lie at infinity (at distances that are far away compared to the scale of the problem). In all such problems, the results of integrating eq. 3.6 will contain a term of the form:

$$\frac{2}{\sqrt{\pi}} \int_0^n e^{-n^2} dn = \text{erf}(n) \quad (3.15)$$

This integral cannot be solved. However, because it occurs so often in solutions of the heat flow equation, it has its own name: the *error function*. The values of the error function for different values of  $n$  have been determined numerically and can be looked up on tables, or it can be calculated with some numerical approximation (s. Table C.9). Fig. 3.6 shows the shape of the error function. In many solutions of eq. 3.6 the variable  $n$  from eq. 3.15 has the form  $n = z/\sqrt{4\kappa t}$ . There, time  $t$ , and distance  $z$  are inside the error function and they are in a quadratic relationship to each other. Most solutions that we will use for the description of contact metamorphism contain error functions of this form (sect. 3.6). We will need the term  $n = z/\sqrt{4\kappa t}$  in sect. 3.1.4, 4.4.1 and others. The *complementary error function*  $\text{erfc}$  is defined as:

$$\text{erfc}(n) = 1 - \text{erf}(n) \quad . \quad (3.16)$$



**Figure 3.6.** The error function and the complementary error function. The dashed frame in **a** shows the part of  $\text{erf}(n)$  that is shown enlarged in **b**

### 3.1.4 Time Scales of Diffusion

In the last sections we have seen that heat conduction approaches thermal equilibrium asymptotically. Even after a very long time the new equilibrium is not reached a 100 % and isotherms will still be curved, even if ever so slightly. Thus, if we want to define a duration of thermal equilibration we need to define some arbitrary point along the equilibration curve which we use as a time scale, for example when half or 90 % or 99 % of the new equilibrium is reached. We call such a number along an asymptotically evolving process *characteristic time scale*, in this case: *time scale of diffusion* or: *thermal time constant*,  $t_{eq}$ . We can use  $t_{eq}$  as a scaling factor for the duration of a thermal event. It is a fantastic aid for an enormous number of estimates, even when in the field. It can be used to estimate the width of contact aureoles, the duration of metamorphic events, the chemical zoning profile of crystals and much more (s. Problem 7.8). The thermal time constant  $t_{eq}$  is given by the relationship

$$t_{eq} \propto \left( \frac{l^2}{\kappa} \right) . \quad (3.17)$$

There,  $l$  is the spatial size, or: *characteristic length scale* of a thermal event (e. g. the diameter of an intrusion, of a hydrothermal vein or a metamorphic terrain).  $\kappa$  is the thermal diffusivity and  $\propto$  means “proportional to”. Eq. 3.17 tells us that the duration of diffusive equilibration is proportional to the square of the size of the equilibrating body. In other words, if for example there are two granitic plutons with one being *twice* the size of the other, then the duration of cooling will be *four* times longer for the larger body. The proportional relationship from eq. 3.17 adapts various forms, depending on

the problem solved and on how exactly we want to define the thermal time constant. Usually it is defined as:

$$t_{eq} = \left( \frac{l^2}{2\kappa} \right) \quad \text{or :} \quad t_{eq} = \left( \frac{l^2}{(2.32)^2 \kappa} \right) . \quad (3.18)$$

The different formulations are based on different assumptions about the magnitude of the scaling factor for  $t_{eq}$ . In eq. 3.20 we will discuss this in some more detail. Table 3.2 shows some numerical values for different sizes of cooling bodies.

The value of  $t_{eq}$  can be used either way. It can be used to estimate the duration of a diffusion process *or* it can be used to estimate the length scale of a thermally influenced region after a given time scale. We should try to remember that the typical duration of thermal equilibration of length scales of 10 m, 100 m, 1 km, 10 km and 100 km is of the order of 1 y, 100 y, 10 000 y, 1 my and 100 my, respectively. Table 3.2 gives a number of examples of thermal time constants for different length scales.

**The Meaning of the “Time Scale” of Diffusion.** We now want to define a bit more clearly what is meant by the terms “length scale” and “time scale” which we have introduced in the last section. For this, we have to use some information from sect. 3.6. There, we shall see that many solutions of the heat conduction equation contain an error function of the form:  $\text{erf}(l/\sqrt{4\kappa t})$ . For example, the solutions discussed on page p. 111 or in eq. 3.80 are all of the form:

$$T(t) = a + b \left( \text{erf} \left( \frac{l}{\sqrt{4\kappa t}} \right) \right) \quad (3.19)$$

where  $a$  and  $b$  are constants. In eq. 3.19 the expression  $(l/\sqrt{4\kappa t})$  corresponds to the variable  $n$  from eq. 3.15. The length scale  $l$  and time  $t$  are both contained within the error function and do not appear elsewhere in the solution. The shape of the error function in Fig. 3.6 shows that it reaches asymptotically 1 as  $n$  get very large. Correspondingly, from eq. 3.19, the term inside

**Table 3.2.** Different values of the thermal time constant  $t_{eq}$  for a series of geologically relevant length scales  $l$ , and calculated for one of the two proportionalities given in eq. 3.18 as well as for one used much later in this book (p. 279) for the same purpose

$l$	$t_{eq} = l^2/2\kappa$	$t_{eq} = l^2/\pi^2 \kappa$
10 m	$5 \cdot 10^7 \text{ s} \approx 1.58 \text{ y} \approx 10^0 \text{ y}$	$1.01 \cdot 10^7 \text{ s} \approx 16 \text{ weeks} \approx 10^{-1} \text{ y}$
100 m	$5 \cdot 10^9 \text{ s} \approx 158 \text{ y} \approx 10^2 \text{ y}$	$1.01 \cdot 10^9 \text{ s} \approx 32 \text{ y} \approx 10^1 \text{ y}$
1 km	$5 \cdot 10^{11} \text{ s} \approx 15 000 \text{ y} \approx 10^4 \text{ y}$	$1.01 \cdot 10^{11} \text{ s} \approx 3200 \text{ y} \approx 10^3 \text{ y}$
10 km	$5 \cdot 10^{13} \text{ s} \approx 1.5 \text{ my} \approx 10^6 \text{ y}$	$1.01 \cdot 10^{13} \text{ s} \approx 320 000 \text{ y} \approx 10^5 \text{ y}$
100 km	$5 \cdot 10^{15} \text{ s} \approx 158 \text{ my} \approx 10^8 \text{ y}$	$1.01 \cdot 10^{15} \text{ s} \approx 32 \text{ my} \approx 10^7 \text{ y}$

the brackets will always remain 1 for very large  $l$  (regardless of  $t$ ), or for very small  $t$  (regardless of  $l$ ). We can also see that – because time is in the denominator inside the error function – *complete* equilibrium is reached only after infinite time (when the term inside the brackets asymptotically approaches zero). In order to define a “duration of equilibration” we may want to arbitrarily choose a point in the equilibration process where the argument of the error function (for which we use  $n$  in eq. 3.15) is 1. This means that:

$$\left( \frac{l}{\sqrt{4\kappa t}} \right) = 1 \quad \text{or :} \quad t = \frac{l^2}{2\kappa} = t_{eq} . \quad (3.20)$$

Figure 3.6 illustrates that for the argument to be 1 (where  $t = l^2/(2\kappa)$ ), the thermal equilibration is 84.3 % complete. Just for simplicity, this arbitrary value is often chosen as a scaling factor for the equilibration history where it may be said that the diffusive equilibration is “largely complete”.

As the value of 84.3 % seems a bit arbitrary, thermal time constants are sometimes formulated using the reverse consideration: What is the thickness of a layer, which has equilibrated to 90 % within a given time? From Fig. 3.6 we can see that  $\text{erf}(n) = 0.9$ , for  $n \approx 1.16$  or  $l \approx 2.32\sqrt{\kappa t}$ . This expression for thickness is also used for the thermal definition of the lithosphere (sect. 3.4). It is what is sometimes called the *thermal boundary layer*.

Both formulations discussed above are chosen in an arbitrary way, and either can be used depending whether one wants a slightly simpler formulation of  $t_{eq}$ , or the value of the percent of thermal equilibration. It is important to note that  $t_{eq}$  gives an approximation of the time it will take for most of the thermal equilibration to take place. In summary, it will be sufficient for most purposes to remember the basic message of eq. 3.17:

- During conductive processes the duration of thermal equilibration increases with the square of the length scale of the equilibrating body.
- For geologically realistic thermal diffusivities, the characteristic time scale of equilibration of a  $10^3$  m length scale is of the order of  $10^4$  y. On the scale of the lithosphere ( $10^8$  m), the time scale of equilibration is of the order of  $10^8$  y.

This means that regional metamorphism of nappe piles that are several tens of kilometers thick should last of the order of several tens of my. We shall discuss the implications of this in much more detail in sect. 6.2.1.

## 3.2 Principles of Heat Production

We discern three fundamentally different geological mechanisms that produce heat:

- radioactive heat production,

- chemical heat production,
- mechanical heat production.

In the next sections we derive the basic equations that are needed to describe these three mechanism and we discuss their respective geological relevance. In general, the rate of temperature change due to heat production may be described by:

$$\frac{dT}{dt} = \frac{S}{\rho c_p} . \quad (3.21)$$

There,  $T$ ,  $t$ ,  $\rho$  and  $c_p$  correspond to temperature, time, density and heat capacity as discussed on p. 53 and  $S$  is the volumetric rate of heat production in  $\text{J s}^{-1} \text{m}^{-3} = \text{W m}^{-3}$ . Heat production rate must be divided by density and specific heat to convert the volumetric heat production rate into a rate of temperature change, just as we have done with heat flow in section 3.1.1. If  $S$  is positive, heat is produced,  $dT/dt$  is positive and rocks heat up. If  $S$  is negative, heat is consumed,  $dT/dt$  is negative and rocks cool. The heat production rate  $S$  can be of radioactive, chemical or mechanical origin so that we can write:

$$S = S_{\text{rad}} + S_{\text{chem}} + S_{\text{mec}} . \quad (3.22)$$

All three of these components may have a significant influence on the thermal evolution of rocks depending on the circumstances and all three have different characteristics that require different methods of description.

**Production or Conduction?** From section 3.1 we remember that the rate of temperature change due to conduction is proportional to the difference between heat flow into and out of a unit rock volume ( $q_{\text{in}} - q_{\text{out}}$ ). However, we now learn that the rate of temperature change also depends on the amount of heat that is produced *inside* this unit volume (Fig. 3.2). This value must be added to the difference  $q_{\text{in}} - q_{\text{out}}$ . Thus, if we adapt eq. 3.6 to formulate a thermal energy balance considering conduction *and* production, we can write:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} + \left( \frac{S}{\rho c_p} \right) . \quad (3.23)$$

Note that we now need to use partial derivatives. Whether or not we need to consider eq. 3.23 to model a given thermal problem, or if we can use the simpler form of eq. 3.21 (neglecting heat production) or eq. 3.21 (neglecting heat conduction) depends on the magnitude and relative rates of heat production and heat conduction. Some heat production rates are very rapid (e.g. friction heat production during an earthquake), others are very slow (e.g. reaction heat production during retrograde metamorphism) but it really is on their rate *relative* to conduction that is critical to consider. A comparison between the diffusion and the heat production term on the right hand side of eq. 3.23

element	U	Th	K
mean mantle concentration (kg/kg)	$31 \times 10^{-9}$	$124 \times 10^{-9}$	$31 \times 10^{-5}$
mean crustal concentration	$1.24 \times 10^{-6}$	$5.6 \times 10^{-6}$	$1.43 \times 10^{-2}$
mantle heat production (W/kg)	$3 \times 10^{-12}$	$3.2 \times 10^{-12}$	$1.1 \times 10^{-12}$
crustal heat production (W/kg)	$1.4 \times 10^{-10}$	$1.5 \times 10^{-10}$	$0.5 \times 10^{-10}$

**Table 3.3.** Concentrations of heat producing elements in the crust and undepleted mantle (after Turcotte and Schubert 2002). In granites, the heat production is about 2–3 times higher than the values listed here. The heat productions are per kg of rock, i.e. the values come from concentrations given in the first two rows multiplied with the heat productions given in the text

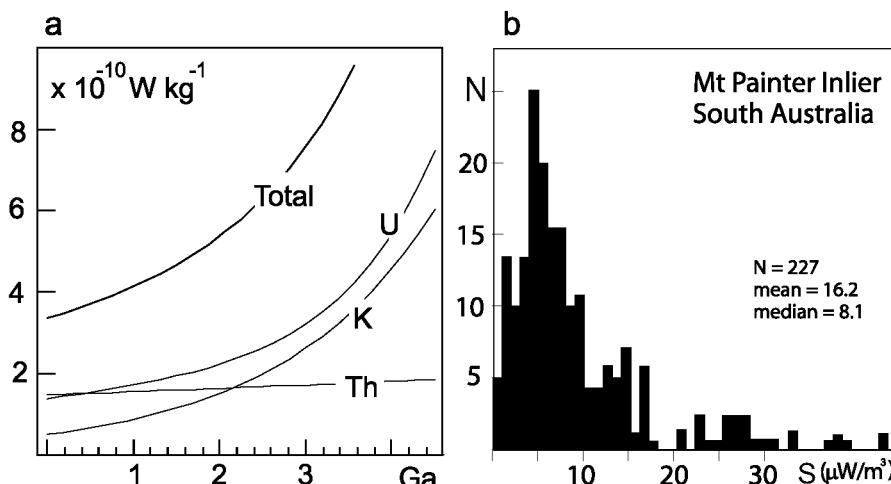
can help us to judge if one of the two processes is much smaller than the other and can therefore be neglected in the problem that is being considered. Such a comparison could be made two ways:

1. *Comparison.* The characteristic time scale of diffusion  $t_{eq}$  of diffusion may be compared with the duration of heat production. If  $t_{eq}$  is much larger than the duration of heat production, then conduction can be neglected and heat production will dominate the temperature change. However, if  $t_{eq}$  is much smaller than the duration of heat production, then all heat that is produced will be conducted away at a much faster rate than it is being produced. Even large amounts of heat production will have a comparably small influence on the temperature and it may be possible to neglect the production term. If heat production- and conduction occur on similar time scales, then the following comparison may be of value.
2. *Comparison.* Consider the case of stable temperature where  $dT/dt=0$ . Then we can see from eq. 3.23 and eq. 3.6 that:  $-S/k = d^2T/dx^2$ . In words: if the ratio  $S/k$  corresponds to the curvature of the temperature profile, then the rate of heat production is balanced by the rate of heat conduction. An important example for such a balance is given by the steady state shape of isotherms in subduction zones (p. 107) and a corresponding example in the theory of mass transfer is discussed on p. 192.

However, for this chapter, let us remain with the case where heat production is much more important than heat conduction and we neglect any diffusion processes in the first instance.

### 3.2.1 Radioactive Heat Production

Radioactive (or: radiogenic) heat is produced in the earth predominantly by the naturally occurring radioactive isotopes  $^{238}\text{U}$ ,  $^{235}\text{U}$ ,  $^{232}\text{Th}$  and  $^{40}\text{K}$ . Of the two naturally occurring uranium isotopes 99.28% is  $^{238}\text{U}$  and only 0.72% is  $^{235}\text{U}$ . All of the naturally occurring thorium is  $^{232}\text{Th}$  and only 0.0119% of



**Figure 3.7.** a Radioactive heat production in the crust through time. Note that the heat production in the Archaean 3 Ga ago was about twice as high as today. Also note that U and Th are the primary heat producing elements today, but it was U and K in the past. b Rates of radiogenic heat production in granites from the Mt Painter province, a low-pressure high-temperature metamorphic terrain in Australia (after Sandiford and Hand 1998b). Heat production rates in other Proterozoic terrains of Australia are similar.  $N$  is the number of data points

the natural potassium is  $^{40}\text{K}$  (Turcotte and Schubert 2002). As pure metals, these 4 isotopes produce the following amounts of heat:  $^{238}\text{U} = 9.46 \times 10^{-5} \text{ W kg}^{-1}$ ;  $^{235}\text{U} = 5.69 \times 10^{-4} \text{ W kg}^{-1}$ ;  $^{232}\text{Th} = 2.64 \times 10^{-5} \text{ W kg}^{-1}$  and  $^{40}\text{K} = 2.92 \times 10^{-5} \text{ W kg}^{-1}$ . Fortunately, the concentrations of these elements in rocks are quite low so that substantially less heat is produced per cubic meter of rock. Table 3.3 lists some average concentrations of the heat producing elements in the continental crust and in the mantle. We can see that the earth's mantle (oceanic crust has comparable values) contains about 2 orders of magnitude less radioactive elements than the crust. These concentrations are still important when considering problems related to cooling of earth as a whole or when thinking about the vigor of mantle convection in the Archaean, but for considerations of the heat budget of the Phanerozoic crust we need not consider radioactivity in the mantle. However, the crustal heat production is significant: The sum of the values listed in this table is about  $3.4 \times 10^{-10} \text{ W kg}^{-1}$ , which corresponds roughly to a heat production rate of about one  $\mu\text{W m}^{-3}$ . Using typical values for heat capacity and density of crustal rocks and eq. 3.21,  $S = 1 \mu\text{W m}^{-3}$  converts to a heating rate of about  $10^\circ\text{C}$  per million years (s. Problem 3.6). So the burial of highly radiogenic bodies by deformation can cause significant heating ! In fact, most granites have substantially higher heat productions than those listed in Table 3.3 and there are many terrains around the world where radiogenic heat

production rates is significantly higher than some  $\mu\text{W m}^{-3}$  (Fig. 3.7b; Sandiford and Hand 1998b). Radiogenic heat production in the continental crust is responsible for about half of the heat flow that we can measure at the surface of the earth (Chapter 2, Fig. 2.18). However, in total, about 80 % of the cooling of earth is related to radioactive decay because even the low radioactivity in the mantle has an important influence on enhancing convective motion in the mantle and therefore on the heat that flows from the mantle into the lithosphere. Radioactive heat production is an important contributor to the heat budget of the crust as a whole (Chamberlain and Sonder 1990; MacLaren et al. 1999; Sandiford et al. 1998).

**Radioactivity in Space and Time.** We know that heat producing elements are not homogeneously distributed across the crust, but that the bulk of the heat producing elements is concentrated in the upper crust. This is because all of the important radiogenic elements occur in granite and the intrusion of granites has been the principle agent of fractionating and redistributing matter from the earth's interior towards the surface. Uranium and thorium are very large atoms and are therefore incompatible and thus contained in granitic melts. Potassium is less radioactive, but there is so much of it in potassium feldspar so that it also has a significant contribution. Over time, granitic melts have transported those elements into the upper crust. Nevertheless, the vertical distribution of heat producing elements in the crust may be extremely variable (e. g. Haack 1983; Lachenbruch and Bunker 1971). In the section on the calculation of geotherms we will spend some time to discuss the influence of the vertical distribution of these elements on the temperature profile of the lithosphere (sect. 3.4.2, see also p. 289).

The temporal variation of radioactivity in the earth's crust through time has been a topic of significant debate, as it is intimately connected to the question if plate tectonics is a modern feature or if it occurred in the Archaean as well. The half lives of the four important radioactive isotopes are known to us. They are:  $^{238}\text{U} = 4.47 \text{ Ga}$ ;  $^{235}\text{U} = 0.704 \text{ Ga}$ ;  $^{232}\text{Th} = 14 \text{ Ga}$  and  $^{40}\text{K} = 1.25 \text{ Ga}$ . As such, the half lives are comparable to the age of earth and it is easy to estimate the radioactivity in the past. Fig. 3.7a shows that the heat production in the Archaean was about twice as high as today. Of course this is true for the mantle as well. A series of authors have discussed the influence of higher heat production in the Archaean for secular changes observed in the style of ore deposits (Groves et al. 2005), the style of metamorphism (Sandiford 1989) and the style of continent formation in general (Abbott and Hoffman 1984).

### 3.2.2 Mechanical Heat Production

The forces that deform rocks can be viewed as mechanical energy that is added to the rock. The work done on the system is the product of force applied to the system times the distance over which it is deformed. This

energy will be taken up by a variety of *mechanical energy sinks*. A part of this energy will be transformed into potential energy (s. sect. 5.3.1), some into dislocation energy in crystal lattices, some in noise and other forms of energy. However, most authors agree that the majority of this mechanically produced energy will be transformed into friction heat. Frictional heating is also often referred to as *shear heating* (because it is produced when rocks are sheared) or *viscous dissipation* (because momentum is dissipated, see details in sect. 5.1.1). We abbreviate this mechanical heat production with  $S_{\text{mec}}$ . The rate of mechanical heat production  $S_{\text{mec}}$  is given by the product of deviatoric stress  $\tau$  and strain rate  $\dot{\epsilon}$ . Stress has the units of Pascal. One Pascal is one Joule per cubic meter ( $1 \text{ Pa} = 1 \text{ J m}^{-3}$ ). Thus, stress can be expressed as energy per volume and energy is stress *times* volume. These conversions between the different units should be straight forward, remembering the well-known relationships:

$$\text{force} = \text{mass} \times \text{acceleration} \quad \text{and} \quad \text{stress} = \frac{\text{force}}{\text{area}} \quad .$$

The units of acceleration are  $\text{m s}^{-2}$  and those of force are therefore:  $\text{kg m s}^{-2}$ . Stress and pressure therefore have the units of  $\text{kg m s}^{-2} \text{ m}^{-2}$  or  $\text{Pa} = \text{kg m}^{-1} \text{ s}^{-2}$  and energy has the units of  $\text{J} = \text{kg m}^2 \text{ s}^{-2}$ . Accordingly, if high deviatoric stresses are required to deform a rock, a lot work is done on the system and the mechanical energy production rate is high, and vice versa. We notice that when we rub our hands together: The harder we press and the faster we rub, the warmer they get. Both deviatoric stress and strain rate are tensors and the rate of mechanical heat production is therefore given by a tensor product. Considering tangential and normal components in three dimensions, friction heat is given by:

$$S_{\text{mec}} = \tau_{xx}\dot{\epsilon}_{xx} + \tau_{yy}\dot{\epsilon}_{yy} + \tau_{zz}\dot{\epsilon}_{zz} + 2(\tau_{xy}\dot{\epsilon}_{xy} + \tau_{xz}\dot{\epsilon}_{xz} + \tau_{yz}\dot{\epsilon}_{yz}) \quad (3.24)$$

(see e.g. Burg and Gerya 2005). The subscripts and notation of the components of deviatoric stress and strain rate should not worry us here and will be discussed in some detail in sect. 5.1.1. For now, we consider only the one-dimensional case (and only normal components, i.e. we neglect shear stresses and shear strain rates). Then we can view the mechanical heat production rate as the simple scalar product:

$$S_{\text{mec}} = \tau\dot{\epsilon} \quad . \quad (3.25)$$

Some of the implicit assumptions in this simplification are discussed in sect. 5.1.1 and 6.2.2. In order to write the temperature change that arises from frictional heating we can write in analogy to eq. 3.21:

$$\frac{dT}{dt} = \frac{\tau\dot{\epsilon}}{\rho c_p} \quad . \quad (3.26)$$

Note that eq. 3.26 is independent of the deformation mechanism. Both brittle and ductile deformation mechanisms will produce the same amount of friction

heat if they support the same deviatoric stresses. We only may need to be careful with the units: Brittle faults do not have a *strain rate* (in  $\text{s}^{-1}$ ) but a *slip rate* in meters per second. The product of slip rate and deviatoric stress does not have the units of heat production per cubic meter, but the units of heat flow (i.e.  $\text{J s}^{-1} \text{m}^{-2}$  normal to the fault surface) which can be converted into a heating rate using the laws of heat conduction discussed in previous sections.

**Temperature Dependent Strength.** In chapter 5 (p. 231) we will show that the deviatoric stresses that rocks can support during viscous deformation are a strong function of temperature. Thus, eq. 3.26 should strictly have been written as:

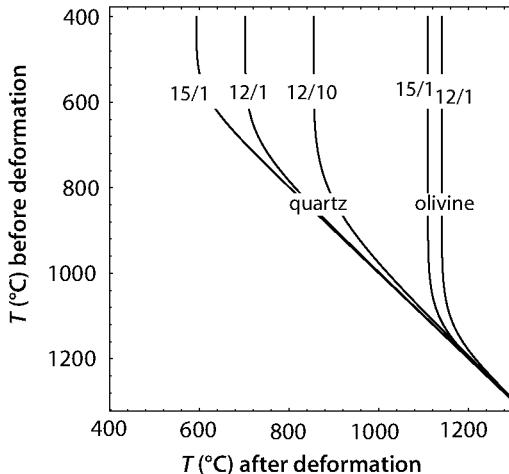
$$\frac{dT}{dt} = \frac{\tau(T)\dot{\epsilon}}{\rho c_p} . \quad (3.27)$$

We can see from this equation that the temperature increase that may occur due to shear heating is in itself a function of temperature. There is a negative feedback between shear heat production and thermal weakening so that any incremental amount of shear heating will instantaneously soften the rocks and subsequent heating becomes more and more difficult as temperature gets higher. Shear heating is self limiting.

For some assumptions about the deformation mechanism it is possible to integrate eq. 3.27, even though temperature occurs both on the left and right hand sides of the equation. For example, for the deformation law described by eq. 5.46, it is possible to integrate eq. 3.27 (Stüwe 1998a). While this analytical solution will not be repeated here, a graph of this solution is shown in Fig. 3.8 for the rheological parameters of quartz and olivine. It may be seen that – if rocks are very soft (e.g. at very high temperatures) – there is no shear heating and temperatures before and after deformation are the same. However, at temperatures below about  $600^\circ\text{C}$  shear heating may be quite significant. Due to the self limiting nature of shear heating, curves become parallel to the vertical axis at even colder temperatures.

Note that Fig. 3.8 can only be used to evaluate the effects of shear heating if conduction of heat may be neglected. The time scale of conductive equilibration of the lithosphere is of the order of hundreds of my, while the duration of continental deformation processes is only 1–10 my. Thus, on the scale of the lithosphere, the duration of mechanical heat production is at least one order of magnitude less than the duration of heat conduction. Thus, conduction of heat may be neglected if shear heating is considered on the scale of the lithosphere (s. comparison on p. 66). Also note that Fig. 3.8 was calculated assuming that the strain rate remains constant. While this describes some geological scenarios, there are many others where *stress* is a constant and the strain rate changes in order to balance it (s. p. 246).

**Shear Zones of Finite Width.** If we consider problems on a small scale, for example shear heating around a small shear zone, then the influence of conduction during heat production needs to be considered. For the solution



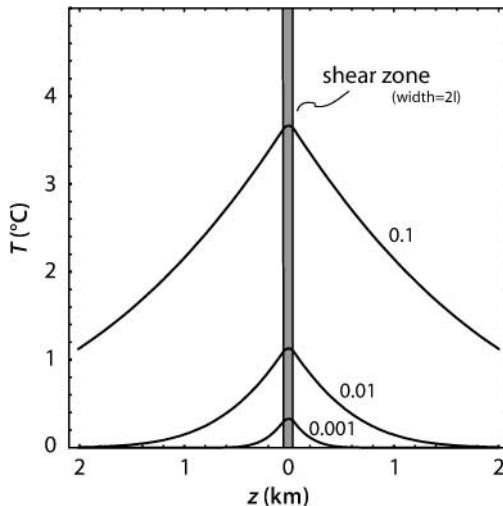
**Figure 3.8.** Temperature of rocks at the start of deformation, plotted against the temperature at the end of deformation. Mechanical heat production is assumed to be the only heating process and heat conduction is neglected. The diagram was calculated with the solution presented by Stüwe (1998a), which was derived from eq. 3.27 and eq. 5.46. The various curves are labeled according to the following syntax: before the dash:  ${}_{10}\log(\dot{\epsilon})$  between  $\dot{\epsilon} = 10^{-12}$  and  $10^{-15} \text{ s}^{-1}$ ; after the dash: elongation from 1–10. Obviously, the increase in temperature is larger for larger strains. However, during crustal shortening, a stretch of 2–3 is rarely exceeded. Thus, the effects of shear heating are likely to be much larger in strike slip regimes where much larger strains are possible than during crustal thickening. Rheological assumptions for quartz are:  $Q = 1.9 \cdot 10^5 \text{ J mol}^{-1}$ ,  $A = 5 \cdot 10^{-6} \text{ MPa}^{-3} \text{ s}^{-1}$ ,  $n = 3$ . Rheological assumptions for olivine are:  $Q = 5.2 \cdot 10^5 \text{ J mol}^{-1}$ ,  $A = 7 \cdot 10^4 \text{ MPa}^{-3} \text{ s}^{-1}$ ,  $n = 3$

of such problems, eq. 3.26 must be enlarged by a term that describes diffusion. The equation we must solve is eq. 3.23, subject to initial and boundary conditions that describe the shear zone geometry. Fortunately, many shear zones have reasonably simple planar geometries that may be described with simple initial and boundary conditions. If we define the width of a shear zone to be  $2l$  and assume a one-dimensional spatial coordinate  $z$  that extends normal to the shear zone and has its origin in the shear zone center, then the boundary conditions for a planar shear zone of constant width may be formulated as follows:

- $T = 0$  at the time  $t = 0$  in the half space  $z > 0$ .
- For all  $t > 0$  there is heat production in the region  $0 < z < l$  at a constant rate  $S$ .
- For all  $t > 0$  the thermal gradient at  $z = 0$  is:  $dT/dz = 0$ .
- At  $z = \pm\infty$  the temperature stays at  $T = 0$  at all times.

You can visualize yourself these initial and boundary conditions by sketching them into a diagram where  $z$  is plotted against  $T$ . Using these initial and boundary conditions it is possible to find solutions of eq. 3.23 that describe the

**Figure 3.9.** Temperature profile across a shear zone of the width  $2l = 50$  m at three different time steps labeled in my. The heat production rate in the shear zone is  $10^{-4} \text{ W m}^{-3}$ ; the thermal conductivity  $k$  is  $2.7 \text{ J s}^{-1} \text{ m}^{-1} \text{ K}^{-1}$ . The figure was calculated with equations given in the 1st edition of this book. For another example where the same model may apply see p. 201



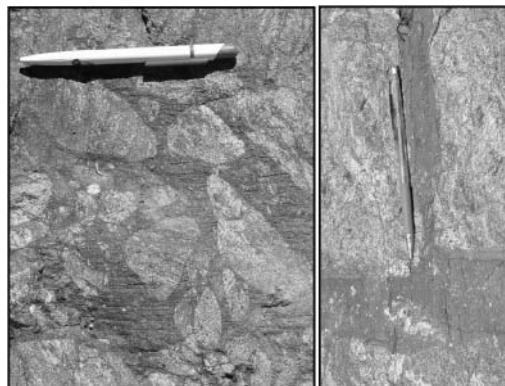
temperatures in and around the shear zone. This and many other solutions of the heat flow equation are discussed by Carslaw and Jaeger (1959) and can be looked up therein. Fig. 3.9 shows a graph of the relevant solution (given in the first edition of this book). The largest error in this figure is introduced by the assumption that the mechanical energy production rate was assumed to be constant. In reality, there is a feed back mechanism between heating and softening in the shear zone. Carslaw and Jaeger (1959) present for this problem some solutions in which the heat production rate can be varied as a function of time. However, for most realistic problems, where shear heat production feeds back on the rheology of the rock (e. g. in Fig. 3.8) it is wiser to use a numerical solution of eq. 3.23.

**Geological Relevance of Shear Heat Production.** A range of authors have discussed the importance of shear heating on a geologically significant scale (e. g. Burg and Gerya 2005; Nabelek and Liu 1999; Brun and Cobbold 1980; Lachenbruch 1980; Scholz 1980; Barton and England 1979; Graham and England 1976). Nevertheless, its importance in many tectonic and metamorphic processes remains contentious. This is because both, deviatoric stresses and strain rates on the scale of the crust are not very well constrained and are among the most discussed geological parameters. We can constrain shear heating to a certain degree using eq. 3.26 to estimate the temperature increase a rock might experience for some realistic deviatoric stresses and strain rates. For this, we neglect heat conduction away from the site of mechanical heat production in the first instance. Our estimates are therefore an upper constraint but may be quite appropriate if the length scale of shear heat production is very large (e.g. deforming nappe piles of several tens of kilometers) and the time scale of heat production is short (e.g. less than a few millions of years).

Methods to measure geological strain rates show an upper limit of  $\dot{\epsilon} = 10^{-12}$  to  $10^{-14} \text{ s}^{-1}$ . These numbers imply that deformation doubles the thickness of a rock package (strain of about 100 %) within 1–10 my. The magnitude of deviatoric stresses is much less constrained. Stress determination experiments are performed at strain rates of  $\dot{\epsilon} = 10^{-6} \text{ s}^{-1}$  and must be extrapolated by six to eight orders of magnitude of strain rate. The relevance of such experimental results remains therefore debated. Moreover, deviatoric stress is strongly temperature dependent. Nevertheless, we know that the order of magnitude of plate tectonic driving forces is between  $10^{12}$  and  $10^{13} \text{ N m}^{-1}$  (s. sect. 5.3) and we will show on p. 300 that this implies a rock strength of 50–100 MPa, averaged over the thickness of the lithosphere. However, the vertical distribution of this strength in the lithosphere is largely unclear. In fact, in section 6.3.5 and 5.2.1 we will see that the middle crust may exceed those values dramatically. In problem 3.7 we use some number to estimate the arising heating rates.

An estimate for total heating can be arrived at even easier: During orogenesis, the crust typically changes its thickness by a factor of two (it doubles in collisional orogens and halves its thickness in many extensional settings). Remember that a stretch of 2 corresponds to a longitudinal strain (or elongation) of 1 ( $\dot{\epsilon} \cdot t = \epsilon = 1$ ; s. eq. 4.1). Assuming all else being constant, it is easy to integrate eq. 3.26. The temperature at the end of deformation for a longitudinal strain of 1 can simply be estimated with:  $T = \tau / (\rho c_p)$ . Using standard values for the density and specific heat ( $\rho = 2700 \text{ kg m}^{-3}$  and  $c_p = 1000 \text{ J kg}^{-1} \text{ K}^{-1}$ ) we can see that a rock that has a shear strength of 100 MPa will be heated by about  $37^\circ\text{C}$ . If rocks are twice as strong, then the temperature increase is twice as high. As the strength of rocks may be several hundreds of MPa under some circumstances, we must conclude that viscous dissipation may be of significant importance to the thermal energy budget of the lithosphere.

**Figure 3.10.** Field photographs of pseudotachylites. The left photo shows network typical of many pseudotachylites, the right photo shows a pseudotachylite with chilled margin and recrystallised center. Both photos are 550 my old pseudotachylites from the Woodroffe Thrust system, Central Australia (Camacho et al. 1995)



Examples where friction heat production has a significant influence on the temperature of rocks are well-known to us from pseudotachylites from all crustal levels (Camacho et al. 1995; 2001; Austrheim et al. 1996) (Fig. 3.10). In those, friction heat was sufficient to even melt the rock. Pseudotachylites form during seismic events where extremely rapid deformation occurred on a very local scale. They are therefore not very appropriate to estimate the influence of friction heat on the thermal evolution of the entire crust where we have to deal with averaged strain rates and averaged stresses (e.g. Kincaid and Silver 1996; Stüwe 1998a). Regardless, even significant amounts of friction heat need not be reflected in significantly increased temperatures. Whether or not shear heating actually becomes geologically significant on a crustal scale depends largely on 2 factors:

- 1. It depends on the relationship between the *length scale* of heat production (which determines how rapidly heat may be conducted away from the site of production) and the *time scale* of heat production. The same considerations that we discussed on p. 66 apply. For example, if a 100 m thick shear zone is active for 1 my, then eq. 3.18 tells us that the characteristic time scale of diffusion of this shear zone is of the order of only 1000 y. Thus, shear heat produced over a time interval of 1 my will be largely conducted away as it is produced. In contrast, if a 15 km nappe pile deforms under the same conditions, then its thermal time constant will be tens of my and all heat produced within 1 my will be largely retained in the pile.
- 2. It depends on the feedback between heating and softening of rocks.

In summary we can say that shear heating *is* a potential candidate for significant heating of rocks and that every argument for or against shear heating should be reduced to an argument about one of the following three points:

- What is the strength of the rocks under consideration?
- What is the strain rate?
- What is the relationship between the *duration* of deformation and the *size* of the deforming rock body?

We can also conclude from the active discussion in the literature that the question is partly open and that it is unwise to take sides without some valued consideration of the problem.

**Adiabatic Processes.** “Adiabatic” means: “without change in heat content” or: “without change in enthalpy”. If a rock is buried, the pressure rises because of the weight of the increasing overburden. The rock is compressed and there is work done to change its volume. The energy of the rock increases by the product of the applied lithostatic force and the distance of shortening during the volume change. It is a type of mechanically produced energy. If no change of the heat content of the system is allowed, then the rock must get warmer. The rock heats *adiabatically*. Correspondingly, an adiabatically heated rock will cool and expand when it is decompressed. We can observe

adiabatic cooling on gas bottles that freeze on their surface when they are opened and adiabatic heating on bicycle tires that warm up when they are pumped up. Rocks are compressible enough so that adiabatic processes is a geologically relevant process, in particular in the earth's mantle. For the consideration of adiabatic processes we define the *isothermal compressibility*  $\beta$  as the relative volume change,  $\partial V$ , per increment of pressure change,  $\partial P$ , at constant temperature. This has the units of  $\text{Pa}^{-1}$  (sect. 5.1.2):

$$\beta = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T . \quad (3.28)$$

We discuss the compressibility in sect. 5.1.2. More commonly we will encounter the thermal expansion coefficient  $\alpha$ . Corresponding to the isothermal compressibility,  $\alpha$  is given by the relative volume change per increment of temperature change at constant pressure. This is given by:

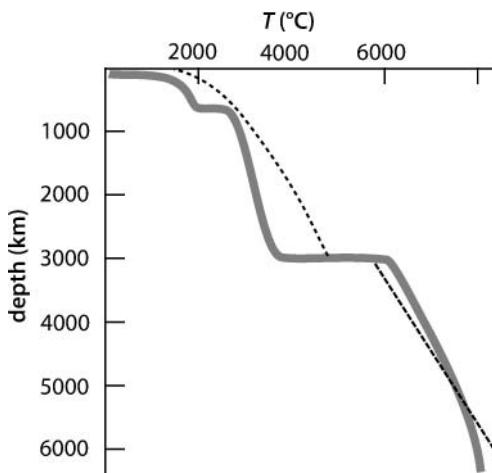
$$\alpha = -\frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_P . \quad (3.29)$$

Rocks have typical values of  $\alpha \approx 3 \cdot 10^{-5} \text{ K}^{-1}$  and  $\beta \approx 10^{-11} \text{ Pa}^{-1}$ . When discussing adiabatic heating processes we need yet another parameter, which is the *adiabatic compressibility* (which is different from the compressibility at constant temperature). This is given by the change in density with change in pressure at constant entropy and should not be confused with the isothermal compressibility discussed in eq. 3.28. The adiabatic compressibility is smaller than the isothermal compressibility.

Whatever the case may be, to the geologist it is relevant to remember that part of the temperature increase with depth in the earth is due to an adiabatic temperature gradient. Without going into the derivation in any detail we state here that this gradient is given by:

$$\frac{dT}{dz} = \frac{\alpha g T}{c_p} . \quad (3.30)$$

The constant  $g$  is the gravitational acceleration,  $T$  is temperature and  $c_p$  is the specific heat. Without further explanation we also want to remember that this gradient is understood to be at constant entropy. Detailed derivations of eq. 3.30 are given by Turcotte and Schubert (1982) and in many other geophysical teaching texts. If we insert realistic numbers we see that the adiabatic temperature gradient in the mantle is about  $0.3\text{--}0.5 \text{ }^\circ\text{C km}^{-1}$ . In other words, the adiabatic temperature change between the surface and the base of a 100 km thick lithosphere is about  $50 \text{ }^\circ\text{C}$ . For most geological processes within the crust this is negligible. However, for geodynamic processes that require the consideration of large vertical length scales in the mantle, the process is of relevance. It is particularly important for melting processes in the upper mantle (Fig. 3.11). In fact, because convection in the asthenosphere is rapid enough to equalize most conductive gradients, the adiabatic gradient is often the only temperature change with depth that there is.



**Figure 3.11.** The temperature profile (thick line) and the melting curve (dashed line) between the earth's surface and the center of the earth in 6 300 km depth (after Jeanloz and Richter 1979; Jeanloz 1988). See Fig. 2.16 for scaling of this figure. In the mantle, the temperature increase is given by the adiabatic temperature gradient (eq. 3.30). The steps in the temperature profile arise from exothermic and endothermic reactions during phase transitions. Note that only the outer core is liquid according to this diagram. However, rapid exhumation of the upper mantle may lead to a cross over of the two curves above 500 km depth and therefore to partial melting. This is what happens underneath mid oceanic ridges or during continental extension

### 3.2.3 Chemical Heat Production

Different rocks are characterized by different internal heat contents defined by the strength of bonding of the atoms in the crystal lattices in the rock-forming minerals. During chemical reaction, the difference in heat content between reactants and products is released or consumed as *latent heat of reaction*. We abbreviate this chemically produced or consumed heat with  $S_{\text{chem}}$ . By far the largest majority of chemical reactions are *endothermic* when the temperature increases. Because of this, temperature rise of rocks may be buffered by the phase transition. Correspondingly, most reactions are exothermic when crossed down temperature. However, most chemical reactions have a positive slope in a pressure-temperature diagram. Thus exothermal reaction can not only be triggered by a *decrease* in temperature, but also by an *increase* in pressure (at constant temperature). In very general terms, we can chemical reactions that produce heat into three groups. In decreasing order of importance these are:

- **Phase transitions:** The chemically produced heat of reactions involving phase transitions is significant to the thermal budget of rocks.

- **Dehydration reactions:** In the solid state, dehydration reactions are the most important producer of reaction heat (Connolly and Thompson 1989; Peacock 1989). In the greenschist facies they produce of the order of  $4 \cdot 10^6$  J per kg of released water. However, rocks contain only of the order of 4%  $\text{H}_2\text{O}$  and this water is being released over quite a large temperature interval. Thus, the heat of reaction is fairly insignificant during regional metamorphism. Connolly and Thompson (1989) estimated that metamorphic reaction produce of the order of  $5 \cdot 10 \cdot 10^{-14}$  W  $\text{cm}^{-3}$ .
- **Solid - solid reactions:** The chemical heat production of solid – solid reaction is negligible for geological problems.

The geologically most important reactions that involve phase transitions are the melting reactions where the *latent heat of fusion* is released or the *latent heat of melting* is consumed. As a consequence, it is important to consider reaction heat when dealing with the thermal energy budget of migmatites and intrusions. A commonly used value for the latent heat of melting of rocks is  $L = 320\,000 \text{ J kg}^{-1}$ . Evaporation and condensation reactions are also strongly exothermic and endothermic respectively, but they are not very important in the geodynamics of the lithosphere.

**Quantitative Description of Chemical Heat Production.** The rate of reaction heat production  $S_{\text{chem}}$  has the same units as any other heat production rate:  $\text{W m}^{-3}$ . It can be described by:

$$S_{\text{chem}} = L \rho \frac{dV}{dt} . \quad (3.31)$$

In this equation  $L$  is the latent heat of reaction in  $\text{J kg}^{-1}$ . Since we think of the chemical heat production rate as a *volumetric* heat production rate, it is necessary to multiply  $L$  by the density  $\rho$  to convert it into a volumetric heat content. The expression  $dV/dt$  is the volumetric proportion of the reaction that occurs per unit time (in  $\text{s}^{-1}$ ). Note that  $V$  has the units of percent and not cubic meters. Thus, the equation determines the part of  $L$  that is freed in every time step of the reaction. Substituting eq. 3.31 into eq. 3.21 we can now formulate the temperature change during chemical heat production to be:

$$\frac{\partial T}{\partial t} = \frac{L}{c_p} \frac{\partial V}{\partial t} . \quad (3.32)$$

If we consider one dimensional diffusion of heat in direction  $z$  as well, but neglect other heat sources we can write:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} + \frac{L}{c_p} \frac{\partial V}{\partial t} , \quad (3.33)$$

where the first term on the right side is the diffusion term discussed at length in eq. 3.6. You will recall from there that  $\kappa = k/\rho c_p$ . In this form of the

equation we also have neglected differences in density between the reactants and the products. It may help you to understand eq. 3.33 by formulating it in terms of energy rather than in terms of temperature. It then takes the form:

$$\frac{\partial H}{\partial t} = k \frac{\partial^2 T}{\partial z^2} + L\rho \frac{\partial V}{\partial t} . \quad (3.34)$$

There,  $H$  is the heat content in  $\text{J m}^{-3}$ . The conversion between heat and temperature is illustrated in Fig. 3.12 and was briefly discussed in eq. 3.3. In principle, it is possible to implement the heat of reaction quantitatively and describe the evolution of  $PT$  paths under consideration of buffering processes. However, the estimates of Thompson and England (1984), Connolly and Thompson (1989), Peacock (1989), Barr and Dahlen (1989) and others shows that the magnitude of reaction heat does not justify to implement such a model in great detail and that it is only important to consider latent heat when rocks melt. Then, reaction heat has indeed quite a significant effect on the thermal evolution of rocks and we will now discuss how to describe it.

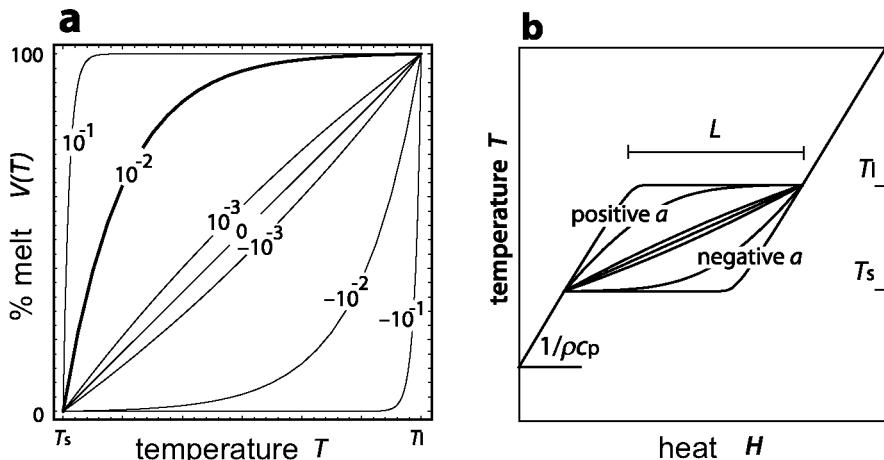
**Thermally Buffered Melting.** Melting during prograde metamorphism in the upper amphibolite and granulite facies is a strongly endothermic process. Thus, the rate with which temperature increases during metamorphism at this grade will be buffered by the melting reactions. At univariant melting reactions, the temperature will remain constant until the phase transition from solid reactants to liquid products is complete. It is the very same reason why we have so much snow slush on our roads in spring: ice and water will both have a temperature of  $0^\circ\text{C}$ , until all ice has melted, even if the air temperature has been above freezing for quite some time. For the same reason water will boil at a constant temperature of  $100^\circ\text{C}$ , regardless of the heat added by the stove, until it all has evaporated. In the buffering interval, the amount of heat added to the rock from the outside is exactly balanced by the amount of heat consumed by the phase transition.

Most rocks consist of many chemical components in complicated chemical systems. As a consequence, they do not melt at a single temperature, but over a melting interval between their solidus (where the first melt appear during temperature increase) and liquidus (where the last remaining piece of rock melts). If we want to describe such rocks, we must find a formulation that allows us to release (or consume) the latent heat over a large temperature interval. For this it is useful to reformulate the rate of volume change from one phase to another from eq. 3.31 into:

$$\frac{\partial V}{\partial t} = \frac{\partial V}{\partial T} \frac{\partial T}{\partial t} . \quad (3.35)$$

In this form it is easier to add the heat of reaction to the time derivative on the left side of eq. 3.33. This reads then as:

$$\left( c_p - L \frac{\partial V}{\partial T} \right) \frac{\partial T}{\partial t} = \frac{k}{\rho} \frac{\partial^2 T}{\partial z^2} \quad (3.36)$$



**Figure 3.12.** a A schematic model for the description of the relationship between melt volume and temperature. The curves are labeled for different values of the constant  $a$  in eq. 3.38. For  $a = 0$ , eq. 3.38 simplifies to the linear relationship  $V(T) = T/(T_l - T_s)$ . The thick drawn line is probably the most realistic curve for the melting of hydrated metapelitic rocks. b The relationship between heat content and temperature in the melting interval of a melting rock (shown between solidus temperature  $T_s$  and liquidus temperature  $T_l$ ). The slope of the curves within the melting interval are for five different assumptions of the melting process according to eq. 3.38; after Stüwe 1995)

or:

$$\frac{\partial T}{\partial t} = \kappa_{\text{mod}} \frac{\partial^2 T}{\partial z^2} . \quad (3.37)$$

There,  $\kappa_{\text{mod}} = k/(\rho c_{\text{mod}})$  is the modified diffusivity and  $c_{\text{mod}}$  is a modified heat capacity:  $c_{\text{mod}} = c_p - L\rho(\partial V/\partial T)$ . We wrote down these equations to illustrate that chemical buffering of thermal processes can be described by modifying the diffusivity. The numerical problems that can arise when doing this were first described by Price and Slack (1954).

**How Do Rocks Melt?** Many migmatite terrains have equilibrated around 700°C, although their peak pressures vary considerably from terrain to terrain. This leads one to suspect that this common peak temperature may be unrelated to the burial depth of the terrain and has other causes. The proximity of 700°C to the solidus temperature of rocks suggests that thermal buffering by the effects of latent heat is a conceivable cause of this common phenomenon (Stüwe 1995). In order to explore this hypothesis, it is necessary to know the amount of melt that is formed at or around 700°C, because only eutectic melts form at the solidus. Petrogenetic studies have rapidly advanced over the last five years and can – in principle – be used to determine the relationship between melt volume and temperature in the melting inter-

val (White et al. 2001; Riesco et al. 2004; Kelsey et al. 2005). However, for a ball park estimate, a useful relationship may be:

$$V(T) = \frac{e^{aT} - e^{aT_s}}{e^{aT_l} - e^{aT_s}} \quad (3.38)$$

In this freely invented equation,  $a$  can be adjusted to explore different relationships (i.e. different melting models) between melt proportion,  $V(T)$  and temperature  $T$  between the solidus temperature  $T_s$  and the liquidus temperature  $T_l$  (Fig. 3.12). If  $a$  has a large positive value, then most of the rock melts near the solidus. If  $a$  adapts large negative values, then most of the melting occurs near the liquidus. For values of  $a \rightarrow 0$ , melting becomes linear in temperature. Many rocks contain hydrated phases at the onset of melting. As water is a great catalyst for melting processes, it is likely that more melting will occur near the solidus than near the liquidus. Realistic values for  $a$  are therefore likely to be positive and possibly around  $a = 0.01$ . If we accept eq. 3.38 as a melting model, then the change of melt volume with temperature  $dV/dT$  can be derived from eq. 3.36 to be the following:

$$\frac{dV}{dT} = \left( \frac{a}{e^{aT_l} - e^{aT_s}} \right) e^{aT} \quad . \quad (3.39)$$

This relationship can be used directly to estimate the influence of melting on the thermal evolution of rocks. Stüwe (1995) showed that it may be large enough to account for the equilibration of parageneses in the low-pressure high-temperature metamorphic environment.

### 3.3 Principles of Heat Advection

Heat can be transported *actively* by the motion of warm rocks. We discern between *advection* and *convection* of heat. *Advection* is generally used if the active transport of heat is only in one direction, for example the transport of heat by an intrusion that moves in the vertical direction. *Convection* is generally used when referring to material transport in a closed loop, for example the convection of mantle material in the asthenosphere, or that of fluids in a hydrothermal system. In this book, we only deal with advection. One-dimensional active transport of heat (for example in the vertical direction  $z$ ), relative to the  $z$  direction may be described by:

$$\frac{\partial T}{\partial t} = u \frac{\partial T}{\partial z} \quad . \quad (3.40)$$

In eq. 3.40,  $u$  is the transport velocity; the derivative  $\partial T / \partial z$  describes the thermal gradient and  $\partial T / \partial t$  is the change of temperature with time. For positive  $u$ , eq. 3.40 describes transport *against* the spatial coordinate  $z$ : transport is from high  $z$  towards lower  $z$  (Fig. B.7 illustrates the sign of  $u$ ). Eq. 3.40 is

also called the transport equation and is equally applicable to the transport of mass, for example during advection of concentration profiles through a crystal lattice. In three dimensions, the heat advection equation is:

$$\frac{\partial T}{\partial t} = u_x \frac{\partial T}{\partial x} + u_y \frac{\partial T}{\partial y} + u_z \frac{\partial T}{\partial z} . \quad (3.41)$$

Although we shall discuss below some examples where eq. 3.40 may be solved analytically, the vast majority of geological problems require numerical solutions of eq. 3.40. Indeed, although eq. 3.40 looks much simpler to solve than eq. 3.6 it is much more complicated to solve it. In numerical solutions of advection problems we are often faced by the problem of *numerical diffusion* (sect. B.2.3, B.2.3). One way to avoid problems with numerical diffusion is by converting the problem from an advection problem into one where there is no advection, but moving boundary conditions; i.e. we describe the problem in a Lagrangian, rather than a Eulerian reference frame (sect. 4.2.2). In Lagrangian descriptions the material is transported through the coordinate system. Eulerian reference frames move with the material.

There are three different important mechanisms by which heat is advected in the lithosphere that require different methods of description. These three mechanisms are:

- advection of heat by magmas, e. g. magmatic intrusion;
- advection of heat by solid rock motion, e. g. erosion or deformation;
- advection of heat by fluids, e. g. during infiltration events.

The difference between these three processes in terms of their mathematical description arises mainly from the relative rates of advective and diffusive processes. These three processes will therefore now be discussed separately.

### 3.3.1 Heat Advection by Magma

During intrusion of magma from deeper into shallower levels in the crust, the heat of the magma is transported to higher crustal levels by the motion of the magma itself. The process of magmatic intrusion is - in general - much faster than most other geological processes, for example the thermal equilibration during contact metamorphism. It is therefore usually not necessary to describe the intrusion process itself by an advection equation. For questions related to the *thermal* evolution of contact metamorphism it usually suffices to assume that intrusion is infinitely rapid and can be described by *instantaneous heating problems* (s. sect. 3.6). Simple examples for good model assumptions to describe intrusions into the crust are given by Jaeger (1964) and for intrusion at the Moho (under plating) are given by Wells (1980) and Huppert and Sparks 1988). Thermal processes related to intrusion will be discussed at length in sect. 3.6. Nevertheless, we want to note that country rocks that heated during contact metamorphism are often referred to as being

“advectively heated”. Strictly speaking this is not correct as it is *conduction* from the intrusion into the country rocks, not *advection* itself that causes contact metamorphism (s. sect. 6.3.3).

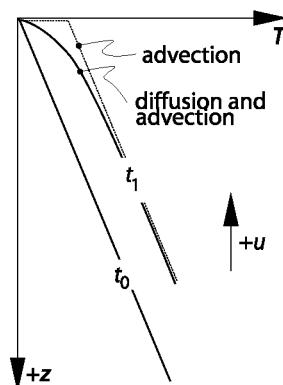
### 3.3.2 Heat Advection by Solid Rock

Any movement and deformation of rocks will carry the heat it contains with it. for example, during exhumation of rocks by erosion, the lithosphere (and its heat) are moved vertically upwards (s. sect. 4.3.4). The column is moved *through* a surface of constant temperature - the surface of earth. Erosion is therefore a heat advection process. In a similar way, any other motion of rocks, for example during thrusting or folding may be interpreted as an advective process (if viewed in an externally fixed Eulerian reference frame, s. p. 144). Here we will only discuss one-dimensional, vertical advection of heat to and from the earth’s surface. The time scale of continental denudation processes is comparable to the time scale of thermal equilibration on the scale of the crust and we can therefore not neglect to consider both processes at the same time. If we want to describe advection and diffusion of heat simultaneously, then we must expand eq. 3.40 by the diffusion term from eq. 3.6. The equation that must be solved becomes:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} + u \frac{\partial T}{\partial z} . \quad (3.42)$$

A schematic illustration how the two processes interact to shape a geotherm during erosion is shown in Fig. 3.13. You may also want to consider to expand this equation by yet another term describing heat production (e.g. the term in eq. 3.21). Then, we would have a complete thermal energy balance to describe any thermal problem. However, it is strongly recommended to perform a careful evaluation of the relative importance of heat conduction, advection and production for a given problem to decide which terms must be considered and which not. Such a comparison may be done similar to that we have shown on p. 66. Here we consider only diffusion and advection.

**Figure 3.13.** Schematic illustration of one-dimensional advection of heat by erosion. The coordinate system is fixed with  $z = 0$  at the earth’s surface. Temperature profiles through the crust are shown for two times: at the onset of erosion  $t_0$ , at which a linear geotherm is assumed and a later time  $t_1$ . The advection rate  $u$  is positive upwards. In the shown time interval the erosion process *adverts* the geotherm by  $u \times t_1$  meters upwards. Simultaneous diffusion causes the curvature of the temperature profile



**Geotherms During Erosion and Sedimentation.** Eq. 3.42 can be used to describe the evolution of a crustal geotherm that is advected through the earth's surface, for example during erosion (upwards advection) or sedimentation (downwards advection) (Benfield 1949) or – on a smaller scale – the evolution of temperature profiles in soil during accumulation or ablation of snow. For simple assumptions about advection rate (e.g. that this is constant) and boundary conditions, analytical solutions of eq. 3.42 may be found. In the following we present some examples of useful boundary conditions for both steady state and time dependent scenarios.

- *Steady state geotherms during erosion.* For fixed boundary conditions a thermal steady state will be reached by geotherms during erosion if the upwards advection of heat is exactly balanced by conductive cooling from the surface. Then, there is no temperature change with time (i.e.  $\partial T / \partial t = 0$ ) and eq. 3.42 simplifies to:

$$\frac{\partial}{\partial z} \left( \frac{\partial T}{\partial z} \right) = -\frac{u}{\kappa} \frac{\partial T}{\partial z} . \quad (3.43)$$

(Spiegel 1968, Mancktelow and Grasemann 1997). Using  $\lambda = u/\kappa$  and  $v = \partial T / \partial z$  this can also be written as:  $v'(z) = -\lambda v(z)$ . This is a very common differential equation and its solution is:  $v(z) = v(0) \exp(-\lambda z)$ . Written in full:

$$\frac{\partial T}{\partial z} = -\frac{\partial T}{\partial z} \big|_{(z=0)} e^{-(uz/\kappa)} . \quad (3.44)$$

The derivation is explained in the appendix on p. 416. There are three geologically relevant boundary conditions with which this equation can be easily integrated a 2nd time to give a closed solution for steady state thermal profiles during advection. In all three examples we assume that one of the two required boundary conditions is given by zero surface temperature ( $T_{(z=0)} = 0$ ), but any surface temperature can simply be added to the given solutions.

- 1. If the other boundary conditions is defined by the mean geothermal gradient  $g$  at  $z = 0$ , then this boundary condition can be directly inserted into eq. 3.44 and integrated a 2nd time. The solution is:

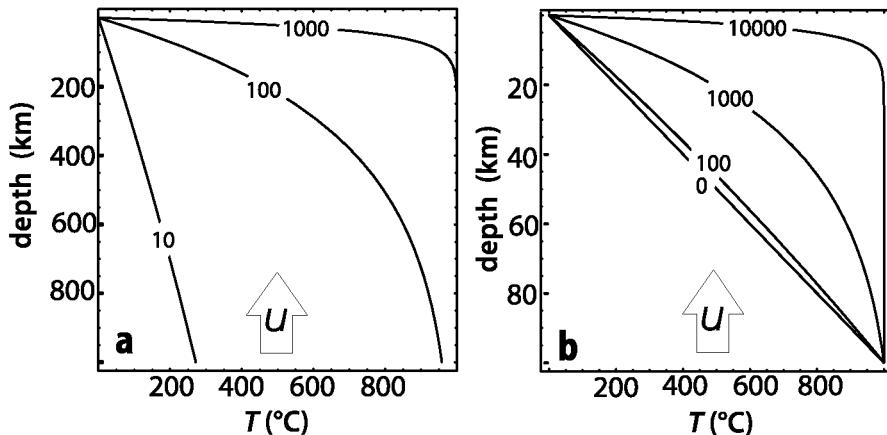
$$T = \frac{g\kappa}{u} \left( 1 - e^{-(uz/\kappa)} \right) . \quad (3.45)$$

- 2. If the lower boundary is fixed with the temperature  $T_\infty$  as the depth goes to infinity ( $T_{(z \rightarrow \infty)} = T_\infty$ ) then a steady state solution of eq. 3.42 is:

$$T = T_\infty \left( 1 - e^{-(uz/\kappa)} \right) . \quad (3.46)$$

- 3. If the lower boundary is fixed at the temperature  $T = T_L$  at a fixed depth  $z = L$ , a steady state the temperature is described by the solution:

$$T = T_L \left( \frac{1 - e^{-(uz/\kappa)}}{1 - e^{-(uL/\kappa)}} \right) . \quad (3.47)$$



**Figure 3.14.** Examples of steady state geotherms during simultaneous upwards advection and diffusion. **a** Geotherms fixed at the surface with  $T = 0$  and at infinite depth at temperature  $T_\infty = 1000^\circ\text{C}$  (calculated with eq. 3.46). **b** Geotherms fixed at the surface with  $T = 0$  and at depth  $L = 100\,000$  with  $T_L = 1000^\circ\text{C}$  (calculated with eq. 3.47). Contours are for different advection rates  $u$  labeled in meters per million years and  $\kappa = 10^{-6}\text{m}^2\text{s}^{-1}$

This latter case is quite a realistic scenario for the lithosphere where it is fair to assume that the temperature at its base is constant (Stüwe et al. 1993b) (Fig. 3.14). More refined solutions for a range of other problems including variably erosion rate and erosion of a heat producing crust are discussed by Mancktelow and Grasemann (1997) as well as Batt and Braun (1997).

- *Time dependent evolutions and cooling paths.* Time dependent solutions may also be found for the two problems discussed above (Mancktelow and Grasemann 1997). Here we present a time dependent evolution for yet another scenario, namely one where there is no lower boundary condition, i.e. the advection of a semi-infinite half space with a constant thermal gradient  $g$  through the surface. For that, a solution of eq. 3.42 is:

$$T = gz - gut + \frac{g}{2} \left( (z + ut) e^{uz/\kappa} \operatorname{erfc} \left( \frac{z + ut}{\sqrt{4\kappa t}} \right) + (ut - z) \operatorname{erfc} \left( \frac{z - ut}{\sqrt{4\kappa t}} \right) \right) \quad (3.48)$$

(Benfield 1949; Carslaw and Jaeger 1959, chapter XV; Mancktelow and Grasemann 1997). While this equation may be used to calculate time dependent geotherms, most geologists do actually not want to know temperatures as a function of depth but the temporal evolution of a given rock. In other words, the evolution of temperature in a Lagrangian reference frame instead of the Eulerian description given by the equations above (s. sect. 4.2.2). Within the reference frames of the geotherm descriptions, rocks move (are advected) through the coordinate system at the rate  $u$ . Thus, if they were at the initial depth  $z_i$  at the onset of erosion, they change their depth to:

$$z = z_i - ut \quad (3.49)$$

at time  $t$ . Thus, if we want to calculate cooling histories of a given rock that are due to its exhumation by erosion, we can simply use eqs. 3.46, 3.47 or 3.48 and substitute eq. 3.49 for  $z$ . Cooling curves that were calculated with this approach and eq. 3.48 are shown in Fig. 3.15.

### 3.3.3 Heat Advection by Fluids

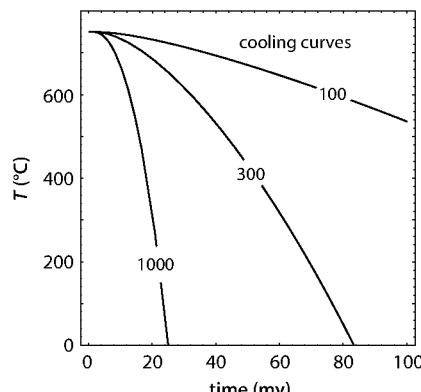
Heat may also be advected by fluids that circulate through rocks. Heating of rocks due to fluid advection is different from the previous examples, because only part of the rock volume is being advected, namely the fluids that fill the pore volume. Thus, when formulating an advective term in an advection-diffusion equation, we need to take care so that we describe only the advection of a fraction of the total rock volume. In a general one-dimensional form eq. 3.40 can be written as:

$$\frac{\partial T}{\partial t} = \phi v_f \left( \frac{\rho_f c_p f}{\rho c_p} \right) \frac{\partial T}{\partial z} , \quad (3.50)$$

(McKenzie 1984). There,  $\phi$  is the porosity of the rock (s. sect. 6.1.3) and  $v_f$  is the fluid flux in  $\text{m}^3 \text{m}^{-2} \text{s}^{-1}$ . The product  $\phi v_f$  is the fluid volume that is transported per unit time and per unit area through the rock. This product has the units of  $\text{m s}^{-1}$ , which corresponds to the standard definition of fluxes (sect. 4.5.3). It is called the *volumetric fluid flux*.  $\rho$  and  $\rho_f$  are the densities and  $c_p$  as well as  $c_{pf}$  are the specific heat capacities, both of the rock and the fluid, respectively. The term inside the brackets is the ratio of the volumetric heat capacities (volumetric heat capacity = heat capacity  $\times$  density) of the fluid to that of the rock.

Eq. 3.50 may be used to describe the thermal effects of fluid advection. However, in many geological processes heat advection by fluids occurs on similar time scales as heat conduction. Thus, it is usually necessary to expand

**Figure 3.15.** Cooling curves of rocks that cool as the consequence of exhumation only. The three curves are for three different erosion rates in  $\text{m my}^{-1}$ . Calculated with eqs. 3.48 and 3.49. Other assumptions are:  $g = 30^\circ \text{C km}^{-1}$  and  $\kappa = 10^{-6} \text{m}^2 \text{s}^{-1}$ . An important feature of the cooling curves is that the cooling rate increases with decreasing temperature. This relationship is opposite to cooling following contact metamorphism



eq. 3.50 by a term describing diffusion as we did in eq. 3.50. The importance of the transport of heat by fluids for the thermal evolution of the crust was discussed by Bickle and McKenzie (1987), Connolly and Thompson (1989) as well as Peacock (1989). These authors agree that the fluid flux that may be caused by metamorphic dehydration reactions is less than about 1 kg fluid per square meter and per year. This is not enough to transport heat very efficiently by fluids. Peacock (1989) estimated that the thermal evolution of rocks can only be influenced by fluids if these are focused into narrow zones from regions as wide as 10 km. We can conclude that heat advection by fluids is insignificant at least when we are interested in thermal budgets of the crust as a whole. Clearly, in special geological settings, for example hydrothermal cells, heat advection by fluids may be the controlling process. Note also that equations similar to eq. 3.50 may be used to describe advection of fluids through rocks in general and are among the important equations governing fluid flow through fractured media.

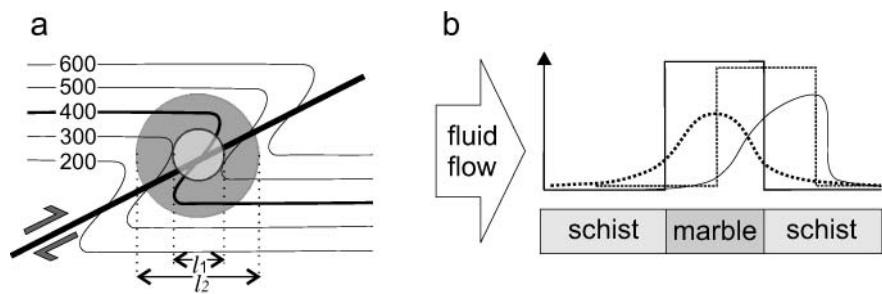
### 3.3.4 The Peclet Number

In many geological processes the diffusion rates and advection rates are of the same order of magnitude. This is true for fluid infiltration processes, for erosion that occurs during thermal equilibration of the crust, thermal profiles around moving faults and many more (sect. 3.7.4). In such processes the Peclet number  $Pe$  is a useful parameter which can be used to estimate the relative influence of diffusive and advective processes. The Peclet number is defined as:

$$Pe = \frac{ul}{\kappa} = \frac{ul\rho c_p}{k} \quad (3.51)$$

where  $u$  is the rate of advection,  $\kappa$  the diffusivity and  $l$  the characteristic length scale of the advection process. The second way to write the Peclet number is only inserted above to remind us that the diffusivity is the ratio of conductivity  $k$ , density  $\rho$  and heat capacity  $c_p$ . If  $Pe$  is about 1, then diffusion and advection are of similar importance to a process. If  $Pe$  is much larger than 1, advection dominates the process. If  $Pe$  is much smaller than 1, then diffusion dominates the process. Eq. 3.51 can be used to derive advection rates. For example, on Fig. 3.16 isotherms have been displaced by a thrust. Consider the 400°C isotherm. On any length scale that is larger than  $l_2$ , this isotherm is simply displaced by the material advection of the hanging wall (i.e.  $Pe > 1$ ). On the length scale of  $l_1$  (dark shaded region) both diffusion and advection have played a role in shaping the curved isotherm (i.e.  $Pe = 1$ ). On the length scale  $l_1$  (light shaded region), diffusion has dominated and the isotherm appears not displaced across the fault (i.e.  $Pe < 1$ ).

Aside from its importance for the description of thermal processes, the Peclet number finds many other applications. For example, Bickle and McKenzie (1987) have used the Peclet number for some fundamental interpretations about the relative importance of diffusive and advective processes during



**Figure 3.16.** **a** Schematic cartoon showing the displacement of isotherms during thrusting. Note that – within the length scale  $l_1$  – diffusion has eliminated any displacement of isotherms across the fault, while above length scale  $l_2$  only displacement, but no curvature is visible. **b** Diffusion and advection of isotopes from a layered sequence. The vertical axis is isotope concentration, the horizontal axis is a profile across the three rock types shown. The continuous step shaped line shows the concentrations at the time of formation. The dotted line after subsequent diffusion ( $Pe < 1$ ), the dashed line after fluid infiltration ( $Pe > 1$ ), but without diffusion ( $Pe = 1$ ). The asymmetric line during both ( $Pe = 1$ )

fluid infiltration of rocks (Fig. 3.16 and Sandiford (2002) has used the Peclet number to argue about the stability of orogens.

### 3.4 Heat in the Continental Lithosphere

In the past sections we have discussed the three fundamental processes that produce and redistribute heat in the lithosphere: *conduction*, *production* and *advection*. In summary from above we can conclude that a full one-dimensional description of the thermal energy balance for the lithosphere has the form:

$$\frac{\partial T}{\partial t} = \left( \frac{k}{\rho c_p} \right) \frac{\partial^2 T}{\partial z^2} + u \frac{\partial T}{\partial z} + \left( \frac{S}{\rho c_p} \right) , \quad (3.52)$$

where the diffusivity is the ratio of conductivity and density  $\times$  heat capacity:  $\kappa = k/\rho c_p$  and the heat production  $S$  may have mechanical, chemical and radioactive contributions (s. eq. 3.22). In this section we will apply our knowledge from the last sections and describe aspects of the thermal structure of the continental lithosphere. For this, not all of the terms in eq. 3.52 may be relevant. Table 3.4 shows (as a summary from the last sections), which thermal energy parameters may be relevant to thermal estimates on the scale of the lithosphere. In order to set the scene, we begin with a brief explanation how the lithosphere may be defined thermally and what “geotherms” actually are.

**Table 3.4.** Summary of possible heat transfer mechanisms in the crust and their potential relevance for the heat budget of metamorphic terrains. The column “*relevance*” makes some very crude suggestions whether the processes listed in the 1st and 2nd column can be of relevance for terrain scale metamorphism. The relationships shown in the 4th column are the key relationships that should be considered when estimating the relevance of a given transfer mechanism in a given terrain. *Conduction* is only relevant if the time scale of conduction,  $t_{eq}$ , is comparable to the ratio of the square of length scale  $l$  and thermal diffusivity  $\kappa$ . *Advection* is only relevant if the Peclet number,  $Pe$ , is larger than 1, where  $u$  is the advection rate,  $l$  length scale of the advective process and  $\kappa$  the diffusivity. *Heat production* is only relevant if the heat production rate  $S$  times its duration  $t$  produces significant amounts of thermal energy (which may then be converted into temperature rise using density and heat capacity). For *radioactivity* this product is generally large on geological time scales, for *chemical* heat production this product is only large for melting reactions and for *mechanical* heat production it is given by the product of stress and strain rate. Both stress and strain rate are not well known and the relevance of mechanical heat production to metamorphism remains a much debated topic in earth sciences. Acronyms for selected key references are: ER77 = England and Richardson (1977); ET84 = England and Thompson (1984); HS88 = Huppert and Sparks (1988); BM87 = Bickle and McKenzie (1987); J64 = Jaeger (1964); L86 = Lux et al. (1986); L70 = Lachenbruch (1970); S98 = Sandiford et al. (1998); C90 = Chamberlain and Sonder (1990); ME90 = Molnar and England (1990a); S98 = Stüwe (1998a); S95 = Stüwe (1995); P89 = Peacock (1989); CT89 = Connolly and Thompson (1989)

transfer mechanism	geological process	geol. relevance	relationship to consider	key ref.	
conduction		large	$t_{eq} = l^2 / \kappa$	ER77; ET84	
advection	by fluids	rare		BM87	
	by magma	- intrusion	large	J64; L86	
		- underplating	large	$Pe = ul / \kappa$	HS88
	by rock	- erosion	large		
		- deformation	large		
production	radioactive	large		L70; S98; C90	
	mechanical	unknown		ME90; S98	
	chemical	- melting	large	$S \times t$	S95
		- dehydration	small		P89; CT89
		- solid-solid	negligible	P89	

**Thermal Definition of the Lithosphere.** The lithosphere may be defined thermally or mechanically (s. sect. 2.4.1). According to the thermal definition, the lithosphere is the outer shell of the earth, in which heat is transported primarily by conduction. In contrast, in the asthenosphere, heat

is transported primarily by convection. Thus, the lithosphere itself is nothing but a *thermal boundary layer* of the earth, although we defined thermal boundary layer somewhat differently in Fig. 2.19. This boundary layer loses heat at all times through the earth's surface into the atmosphere and further – by radiation – into space. The average heat flow through the surface of the continents is  $0.065 \text{ W m}^{-2}$ . The total surface area of the continents is about  $A_c = 2 \cdot 10^8 \text{ km}^2$ . Thus, the total heat loss of earth from the continents is  $1.3 \cdot 10^{13} \text{ J s}^{-1}$ . This heat loss is balanced by radioactive heat production within the lithosphere and by heat flow into the lithosphere from the asthenosphere, so that this thermal boundary layer has a largely constant temperature profile, if it is not disturbed by orogenesis. Thermally stabilized lithosphere has a thickness between 100 and 200 km (Pollack and Chapman 1977).

**Definition of Geotherms.** The function that describes temperature in the lithosphere as a function of depth is what we call a *geotherm*. We discern:

- stable or steady state geotherms,
- transient geotherms.

- *Stable geotherms.* Stable or *steady state* geotherms form by long term thermal equilibration of the lithosphere (sect. 3.4.1). In general, this is understood that the term “steady state” refers to a geotherm in a stationary lithosphere and we shall use it in this way in this section. However, in other reference frames, steady state geotherms may also occur in a moving lithosphere (for example a lithosphere that moves upwards relative to the surface during erosion, s. p. 82). In sect. 3.7.3 and 3.5.2 we discuss examples of steady state geotherms in Eulerian reference frames.

In most geological situations, the temperatures of steady state geotherms increase steadily with depth. Stable geotherms are only found in regions that have had at least about 100 my time for equilibration and have not changed in thickness during this time. The origin of this number is discussed in sect. 3.1.4. Thus, active orogens are *not* characterized by stable geotherms. Regardless, the calculation of steady state geotherms in orogens may help us to estimate the maximum or minimum temperatures that can be attained during an orogenic process at a given depth. This maximum or minimum possible temperature is often called *potential temperature* (e.g. Sandiford and Powell 1990).

- *Transient geotherms.* Transient geotherms are only valid for a particular point in time. In some geological situations, transient geotherms do *not* increase steadily with depth and the change of the geotherm with time can be different in different depths. For example, after rapid stacking of nappes, rocks may simultaneously heat above a major thrust, but cool below it (sect. 7.4.2). In principle it is possible to document such relationships of transient geotherms in space and time with careful observation in the field.

If they can be documented, then such information is invaluable for the interpretation of the nature of a tectonometamorphic process (s. sect. 7.4.2). We discuss the transient evolution of geotherms during orogenic processes in several other chapters (e. g. sect. 3.6 and 6.2.1).

In this section we calculate the quantitative shape of stable continental geotherms as shown schematically in Fig. 2.18.

### 3.4.1 Stable Geotherms: The Relevant Equation

For the stable or steady state case, the heat conduction equation (eq. 3.6) or the full thermal energy balance (eq. 3.52) can be simplified enough so that it is possible to find simple analytical solutions that provide useful tools to understand the thermal structure of the lithosphere, even without a lot of mathematical knowledge. This is therefore a good example to familiarize ourselves with the involved thought process. The equation we must solve is the now familiar heat conduction equation with a heat production term:

$$\frac{\partial T}{\partial t} = \left( \frac{k}{\rho c_p} \right) \frac{\partial^2 T}{\partial z^2} + \frac{S}{\rho c_p} . \quad (3.53)$$

We need the heat production term to account for radioactivity, which is of substantial importance to stable geotherms. However, we can neglect advective terms as we consider only steady state. For *steady state* geotherms, there is no change of the temperature with time (s. sect. 3.1.2). This means:

$$\frac{\partial T}{\partial t} = 0 .$$

Eq. 3.53 simplifies to:

$$\left( \frac{k}{\rho c_p} \right) \frac{d^2 T}{dz^2} + \frac{S}{\rho c_p} = 0 . \quad (3.54)$$

Note that eq. 3.54 is no *partial* differential equation anymore. By canceling out of the constants we get:

$$k \frac{d^2 T}{dz^2} = -S . \quad (3.55)$$

The integration of this equation forms the basis for all calculations of stable geotherms.

**Geotherms Without Radioactivity.** Equation 3.55 may be integrated the easiest, if we neglect radioactive heat production all together. Although we have shown in chap. 2 and sect. 3.2.1 how important radioactivity is for the temperature profiles of the lithosphere (e.g. Fig. 2.18) we can easiest understand the integration of eq. 3.55 by assuming the heat production is negligible:  $S = 0$ . Eq. 3.55 then simplifies to:

$$k \frac{d^2T}{dz^2} = 0 \quad \text{or even:} \quad \frac{d^2T}{dz^2} = 0 \quad . \quad (3.56)$$

As this still is a differential equation of the second order, we must integrate it twice to solve it. A first integration gives:

$$\frac{dT}{dz} = C_1 \quad (3.57)$$

and a second:

$$T = C_1 z + C_2 \quad . \quad (3.58)$$

The two integration constants  $C_1$  and  $C_2$  must be determined by the geological boundary conditions (s. sect. B.1.1). For example, we can assume that we know that the temperature at the surface of earth (at  $z = 0$ ) is constant and has the value  $T = 0$ . Then, for eq. 3.58 to hold,  $C_2$  must be zero so that the temperature is zero at  $z = 0$ . If we assume a thermal definition of the lithosphere, then we can determine the other constant with the assumption that  $T = T_l$  at depth  $z = z_l$  where  $T_l$  and  $z_l$  are the temperature and the depth of the base of the lithosphere. With this assumption  $C_1$  must have the value  $C_1 = T_l/z_l$ . The temperature as a function of depth is therefore described by:

$$T = z \frac{T_l}{z_l} \quad . \quad (3.59)$$

Equation 3.59 describes a linear temperature profile between the surface and the base of the lithosphere. This is not very surprising as we have assumed no heat production and no other reasons why the temperature profile should be anything else but a straight line between the assumptions at the boundaries. With a thermal conductivity of  $k = 2\text{--}3 \text{ W m}^{-1} \text{ }^{\circ}\text{C}^{-1}$ , and  $T_l = 1\,200 \text{ }^{\circ}\text{C}$  as well as a lithospheric thickness of  $z_l = 100 \text{ km}$ , our equation describes a surface heat flow of  $0.024\text{--}0.036 \text{ W m}^{-2}$ . This value is much lower than the average surface heat flow of the continents which is between  $0.04$  and  $0.08 \text{ W m}^{-2}$ . This is one of the proofs of the existence of radioactivity in the lithosphere. We can easily conclude that eq. 3.56 is not a very good model description and that it is wiser to integrate eq. 3.55 using a meaningful function that describes  $S$  as a function of depth. When we do so, we will always assume that  $S = S_{\text{rad}}$ , i.e. there is no other heat production sources but the radiogenic ones. In the steady state mechanical or chemical heat production sources are irrelevant.

### 3.4.2 The Contribution of Radioactivity

The radioactive heat production rate of rocks is of the order of some microwatts per cubic meter (s. Table 3.3). However, unusually high heat productions have been reported from a series of locations around the world, in particular from Australia (McLaren et al. 1999) (Fig. 3.7). A typical value measured from samples at the earth's surface is:  $S = 2-5 \mu\text{W m}^{-3} \equiv 2-5 \cdot 10^{-6} \text{ W m}^{-3}$ . The contribution of this value to the surface heat flow is simply the heat production times its depth extent. For example, if the heat production were constant in the entire crust of 30 km thickness ( $z_c = 30 \text{ km}$ ) then the surface heat flow caused by radioactivity is:

$$q = q_s = S \cdot z_c = 0.03 \text{ W m}^{-2} . \quad (3.60)$$

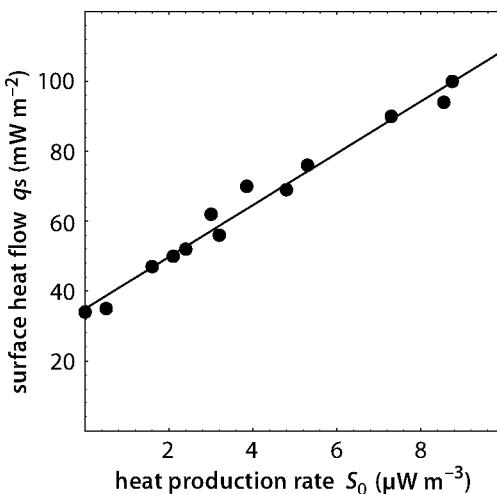
This can be converted into a temperature gradient using eq. 3.1 where we have seen that the thermal gradient has the units of heat flow divided by the thermal conductivity. If the thermal conductivity is  $k = 3 \text{ W m}^{-1} \text{ K}^{-1}$ , then the assumptions from above indicate:  $dT/dz = q/k = 0.05 \text{ }^{\circ}\text{C m}^{-1} = 50 \text{ }^{\circ}\text{C km}^{-1}$ . This geothermal gradient of  $50 \text{ }^{\circ}\text{C}$  per kilometer is only due to the contribution of radioactivity. The mantle heat flow would have to be added to this. Since the resulting thermal gradient would be much higher than just about all thermal gradients measured on earth, we can conclude that the radioactivity of rocks measured at the earth's surface must be higher than that of the rest of the crust (for more information on the radioactive heat flow contribution see p. 289).

**The Distribution of Heat Production.** The considerations above have shown that the radioactive heat production of the crust is unlikely to be that of the surface in the entire crust. Various studies have therefore explored the vertical distribution of heat producing elements (e.g. Cermak and Rybach 1989; Pinet and Jaupart 1987; Pribnow and Hurter 1998; Heier and Brown 1978). The most simple model for a heat source distribution is that the heat production is constant to the depth  $z_{\text{rad}}$  and zero below that. This model depth, to which the crust produces radioactive heat at the same rate as on the surface, has been elegantly determined using the relationship of two independent sets of data that can be measured at the surface: The surface heat flow and the heat production rate at the surface,  $S_0$ . Roy et al. (1968) explored this relationship in the eastern US and its significance was described by Lachenbruch (1968; 1970; 1971). They found a roughly linear relationship between these two parameters (Fig. 3.17). The straight line that fits these data has the form:

$$q_s = q_m + q_{\text{rad}} = q_m + z_{\text{rad}} S_0 . \quad (3.61)$$

In this equation,  $q_s$  is the surface heat flow,  $q_m$  is the mantle heat flow,  $q_{\text{rad}}$  the radiogenically produced heat flow and  $z_{\text{rad}}$  is the thickness of a hypothetical layer in which radioactive heat is produced at the same rate as on the surface.  $q_m$  can be measured from the intersection of the line with the heat flow axis

**Figure 3.17.** Measured data of surface heat flow  $q_s$  and surface heat production  $S_0$  in the eastern US. The best line that fits the data is described by the equation  $q_s = 0.035 + 7413S_0$ . Accordingly, the thickness of the layer that produces heat at a rate  $S$  is  $7413$  m thick and the contribution of mantle heat flow to the total heat flow is  $0.035 \text{ W m}^{-2}$  (after Roy et al. 1968)



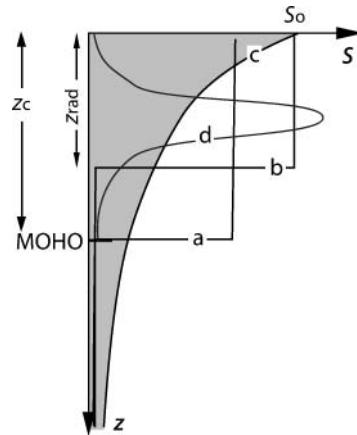
and the value of  $z_{\text{rad}}$  is given by its slope. The data of Roy et al. (1968) show that  $z_{\text{rad}}$  is about 7 km in the eastern US. Similar considerations in other areas indicate thicknesses of 10–15 km. Of course, the crust does not produce heat constantly in this layer and no heat at all below it, but the relationship is useful to estimate the total heat production in the crust. This is given by the product  $z_{\text{rad}}S_0$ . This product corresponds to the area underneath the different model curves in Fig. 3.18.

For the calculation of geotherms in the following sections we will use different model assumptions about the distribution of radioactivity with depth (Fig. 3.18). We will see that these different assumptions have important different implications for the temperatures in both the crust and the mantle part of the lithosphere. So we want to remember that both the *magnitude* and the *distribution* of radioactive heat production are important for the thermal budget of the crust (Sandiford et al. 1998; 2002; Chamberlain and Sonder 1990).

### 3.4.3 Geotherms with Heat Flow Boundary Condition

We begin our calculations of stable geotherms by integrating eq. 3.55 by assuming a somewhat different boundary condition at the base than what we have used to derive eq. 3.59. We will assume as boundary condition that the *heat flow* at the Moho is constant, instead of assuming the *temperature* at the base of the lithosphere. We do this for three reasons: Firstly, because this boundary condition has been the assumption of England and Thompson (1984). Their model has become the classical model to describe Barrovian metamorphism and is therefore the obvious starting point. Secondly – following the argument of England and Richardson (1977) – we want to argue that we know the heat flow at the Moho much better than the temperature

**Figure 3.18.** Four simple models describing the distribution of heat production with depth in the crust (s. Haack 1983). The total heat production of the crust is given by the area underneath the model curves. It is the same for all four models and is shaded for model *c*. *a* Constant concentration in the entire crust and no heat production in the mantle. *b* Constant concentration in the upper crust in a layer with the thickness  $z_{\text{rad}}$  and no heat production below that. *c* Exponential drop off of the heat production with depth. *d* Heat production peaking in the middle crust. Such a situation may occur if a crust with heat source distribution *c* is buried underneath a low heat producing sedimentary pile. Note that the surface heat production is different in the different models, while the total heat production is the same



or depth of the base of the lithosphere. Thirdly, we assume this boundary condition because the integration of eq. 3.55 is quite easy with this boundary condition.

**Constant Heat Production.** If the heat production is independent of depth,  $S$  is a constant. Integration of eq. 3.55 results in:

$$k \frac{dT}{dz} = -Sz + C_1 \quad . \quad (3.62)$$

The left side of this equation has the units of heat flow. The constant of integration  $C_1$  must be derived from the boundary conditions. If we know the mantle heat flow at the Moho, and assume that it is constant through time, we can describe this condition as:  $q_m = \text{constant}$  at  $z = z_c$ . If no details are known about the changes of the heat flow at this depth, then this is the most simple assumption and corresponds therefore with the idea of a good model (sect. 1.1). With this assumption, the constant of integration must be:

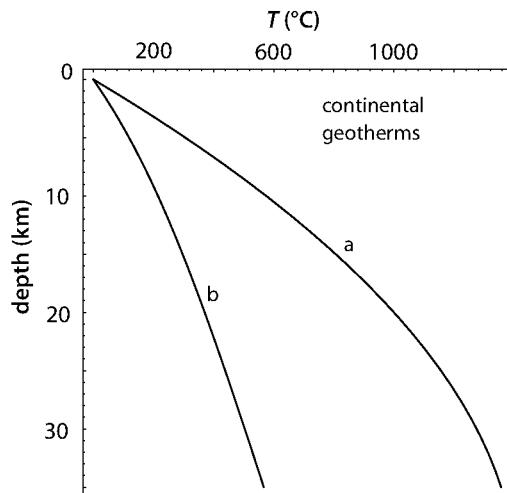
$$C_1 = Sz_c + q_m \quad ,$$

so that at the depth  $z = z_c$  ( $z_c$  is the crustal thickness) we get:  $k \left( \frac{dT}{dz} \right) = q_m$ . After inserting  $C_1$  into eq. 3.62 and integrating a second time we get:

$$kT = -\frac{Sz^2}{2} + Sz_c z + q_m z + C_2 \quad . \quad (3.63)$$

The most meaningful assumption for the second boundary is that the temperature at the surface at  $z = 0$  has the value  $T = 0$ . This gives us that  $C_2 = 0$ . After some simplification we get the following expression:

**Figure 3.19.** Two calculated continental geotherm. It was assumed that the mantle heat flow has the value  $q_m = 0.025 \text{ W m}^{-2}$ . Curve *a* was calculated assuming that the entire crust has a constant heat production. It was calculated with eq. 3.64. Curve *b* was calculated with eq. 3.70 assuming that the heat production decreases exponentially with depth ( $h_r = 10 \text{ km}$ ). For *a* and *b* following other assumptions were made:  $z_c = 35 \text{ km}$ ;  $S = 3 \cdot 10^{-6} \text{ W m}^{-3}$  ( $= S_0$  in eq. 3.70);  $k = 2 \text{ J s}^{-1} \text{ m}^{-1} \text{ K}^{-1}$



$$T = \frac{Sz}{k} \left( z_c - \frac{z}{2} \right) + \frac{q_m z}{k} . \quad (3.64)$$

Equation 3.64 describes the temperature as a function of depth in the crust. It is an *analytical solution* of the differential equation eq. 3.55. An analysis of the units can be used to confirm the internal consistency of this equation. A geotherm calculated with eq. 3.64 is shown in Fig. 3.19a. It is immediately obvious that the temperatures are much too high, although we made meaningful assumptions for mantle heat flow, crustal thickness and surface heat production. However, we also made the incorrect assumption that the heat production is constant in the whole crust and has everywhere the same value as that measured on the surface. This assumption is in contrast to what we argued on p. 93, where we discussed the distribution of heat production with depth. The *real* distribution of the heat producing elements in the crust is not very well known, but we can help ourselves with one of the models from Fig. 3.18. If we make the assumption *b* from that figure (heat is only produced at a constant rate down to the depth  $z_{\text{rad}}$ ), then there is a discontinuity in the heat production with depth. As a consequence, we can integrate eq. 3.55 only to the depth  $z = z_{\text{rad}}$ . We get a result that is very similar to eq. 3.64:

$$T = \frac{Sz}{k} \left( z_{\text{rad}} - \frac{z}{2} \right) + \frac{q_m z}{k} \quad \text{in the region : } z < z_{\text{rad}} . \quad (3.65)$$

For larger depths we can help ourselves with the knowledge that the geothermal gradient must be a straight line (i.e. the heat flow must be constant and is  $q_m$ ) because there is no disturbing influence or heat production anymore (s. eq. 3.59). The temperature at the base of the heat producing layer may be found by evaluating eq. 3.65 at the depth  $z = z_{\text{rad}}$ . Below this depth the temperature  $q_m(z - z_{\text{rad}})/k$  must be added to the temperature at the base of the assumed heat producing layer. It follows that:

$$T = \frac{S z_{\text{rad}}^2}{2k} + \frac{q_m z}{k} \quad \text{in the region: } z > z_{\text{rad}} . \quad (3.66)$$

If you do not understand what we did in the last lines, draw a little sketch that illustrates every written sentence. Equations 3.65 and 3.66 are those that England and Thompson (1984) used for their classic model explaining regional metamorphism (s. sect. 6.2.1). If we insert reasonable values for the parameters, then we get temperatures at the base of the crust between 500 °C and 600 °C. We also get a depth for the base of the lithosphere (the depth of the 1200 °C isotherm) around 100 to 150 km. Both corresponds to a large number of observations from the stable continental shields and the equation provides a good model description of continental geotherms.

**Exponential Heat Production.** A much more elegant model assumption for the distribution of the heat production in the crust is the assumption that there is a continuous exponential drop off in radioactive heat production with depth (model *c* in Fig. 3.18). This model has the great advantage that there is no discontinuity in the heat production in the crust and we do therefore not need several equations to describe a single geotherm as we did above. A continuous distribution of radioactivity with depth also appears intuitively more appropriate. We assume that:

$$S_{(z)} = S_0 e^{(-\frac{z}{h_r})} . \quad (3.67)$$

The variable  $h_r$  is called the *characteristic drop off* or *skin depth* of heat production. According to eq. 3.67, the heat production at depth  $z = h_r$  is only the 1/e part of the heat production at the surface  $S_0$ . Our new starting equation is now:

$$k \frac{d^2 T}{dz^2} = -S_0 e^{(-\frac{z}{h_r})} . \quad (3.68)$$

The integration of this equation is a bit harder than those of the previous section but it still only involves the integration of simple exponential functions (s. Table C.2). A first integration gives:

$$k \frac{dT}{dz} = h_r S_0 e^{(-\frac{z}{h_r})} + C_1 . \quad (3.69)$$

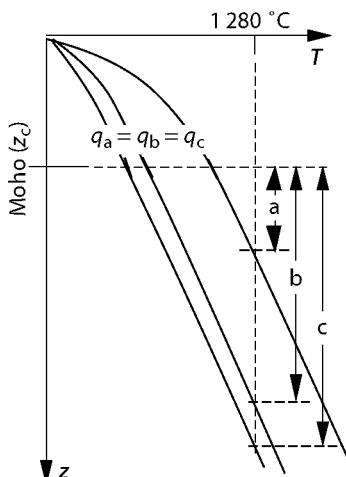
Note that we used the “product rule” for this integration (s. appendix C, Table C.1, C.2). Assuming the same boundary condition that we used in the last section, the constant of integration is:

$$C_1 = -h_r S_0 e^{(-\frac{z_0}{h_r})} + q_m .$$

After inserting this constant into eq. 3.69 and a second integration we get:

$$kT = -h_r^2 S_0 e^{(-\frac{z}{h_r})} - z h_r S_0 e^{(-\frac{z}{h_r})} + q_m z + C_2 . \quad (3.70)$$

**Figure 3.20.** This illustration shows how the thickness of the mantle part of a thermally defined lithosphere is automatically determined by the condition of a constant mantle heat flow at the Moho. All three geotherms on this figure have the same thermal gradient at the Moho and therefore the same heat flow through the Moho. However, the mantle parts of the lithosphere (thickness shown by double arrows) is very different for the three examples

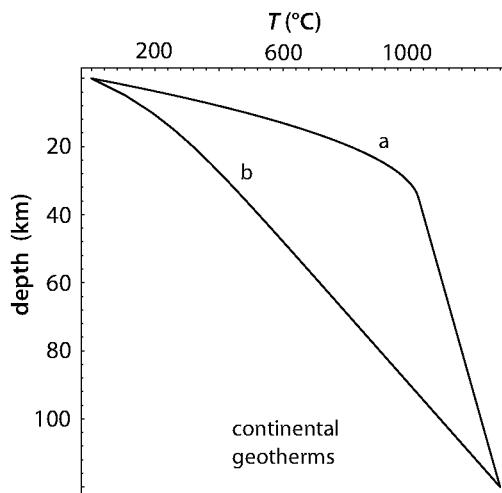


The second constant must be  $C_2 = h_r^2 S_0$  so that the condition  $T = 0$  at the surface ( $z = 0$ ) is fulfilled. Fig. 3.19b shows a geotherm calculated with eq. 3.70 after inserting this integration constant and using reasonable values for all parameters. We see that the geotherm has a realistic shape. Eq. 3.70 is a good model description of stable continental geotherms.

**How do Geotherms Continue at Larger Depth?** In the previous sections we have used a known mantle heat flow as the lower boundary condition for our integration of the steady state heat flow equation. Our calculations were therefore confined to the description of the crust. The thickness of the mantle part of the lithosphere was implicitly determined by the boundary conditions (Fig. 3.20). The depth of the base of the lithosphere is thermally defined by the temperature  $T_l \approx 1200^\circ\text{C}$ , so we only need to find the depth where our geotherm reaches this temperature to determine the thickness of the mantle part of the lithosphere. With a thermal conductivity of  $k = 3$ , and a mantle heat flow at the Moho of  $0.03 \text{ W m}^{-2}$ , the steady state thermal gradient is  $10^\circ\text{C km}^{-1}$  at the Moho. As we assumed no radioactive heat production in the mantle part of the lithosphere, this gradient will be constant in the remainder of the mantle lithosphere. If the Moho-temperature is  $500^\circ\text{C}$  then this gradient implies a thickness of the mantle part of the lithosphere of 70 km (if we define the lithospheric base thermally by the depth where the geotherm intersects the  $1200^\circ\text{C}$  isotherm).

This is quite a reasonable value for the thickness of the mantle lithosphere. Nevertheless, we have only implied this thickness of the mantle part of the lithosphere whereas the equations of the last sections were only designed to describe the crust. Because there are many reasons why we may want to prescribe the thickness of the mantle lithosphere as well (for example because we know it from considerations of surface elevation or lateral buoyancy forces)

**Figure 3.21.** Examples of continental geotherms calculated with a lower boundary condition of a fixed temperature at the base of the lithosphere. Geotherm *a* was calculated assuming constant heat production in the crust and no heat production in the mantle lithosphere. Geotherm *b* was calculated for a continuous, exponentially decreasing heat production using eq. 3.74. The temperature  $T_l$  is assumed to be 1 280 °C; all other parameters are the same as in Fig. 3.19



a lower boundary condition at the base of the lithosphere is often more useful. In the next section we discuss such alternative lower boundary conditions.

#### 3.4.4 Geotherms with Fixed Basal Boundary Condition

If we define the lithosphere thermally, we implicitly state that we know the temperature at its base. An obvious choice for a lower boundary condition may therefore be:  $T = T_l$  at the depth  $z = z_l$ . This choice allows us to describe temperatures in the entire lithosphere. However, we have to pay for this advantage by having to deal with various discontinuities of the parameters at the Moho, for example density. This complicates the integration of eq. 3.55.

**Constant Heat Production.** In a model where we assume constant heat production rate in the crust and no heat production in the mantle part of the lithosphere, density *and* heat production are discontinuous at the Moho. This complicates the integration of eq. 3.55 dramatically. We will not present the equations here and refer the interested reader to the original works of Sandiford and Powell (1990) or Zhou and Sandiford (1992). However, for comparison with the thermal model of England and Thompson (1984) we show an example of a geotherm calculated with these assumptions as curve *a* in Fig. 3.21. We see that this model results in unrealistically high temperatures if we assume the surface heat production rate to be representative for the whole crust.

**Exponential Heat Production.** If we assume a continuous heat production in the whole lithosphere that decreases exponentially with depth, then we can derive from eq. 3.68 an elegant and simple description of stable continental geotherms. After two integrations we get:

$$kT = -h_r^2 S_0 e^{-\frac{z}{h_r}} + C_1 z + C_2 \quad . \quad (3.71)$$

Both constants of integration can only be evaluated after this second integration. The second integration constant  $C_2$  is fairly easy to determine if we assume again that  $T = 0$  at the surface where  $z = 0$ . At  $z = 0$  the exponential term in eq. 3.71 goes to 1 so that  $C_2$  must be:

$$C_2 = h_r^2 S_0 \quad . \quad (3.72)$$

The lower boundary condition of  $T = T_l$  at the depth  $z = z_l$  can be evaluated by rearranging eq. 3.71:

$$C_1 = \frac{kT_l}{z_l} + \frac{h_r^2 S_0 e^{-\frac{z_l}{h_r}}}{z_l} - \frac{C_2}{z_l} \quad . \quad (3.73)$$

After inserting both constants into eq. 3.71 we get:

$$T = \frac{zT_l}{z_l} + \frac{h_r^2 S_0}{k} \left( \left( 1 - e^{-\frac{z}{h_r}} \right) - \left( 1 - e^{-\frac{z_l}{h_r}} \right) \frac{z}{z_l} \right) \quad . \quad (3.74)$$

Curve *b* in Fig. 3.21 is an example of a geotherm calculated with this relationship. Eq. 3.74 provides a realistic and useful description of stable continental geotherms and has been presented and used by a number of authors (Zhou and Sandiford 1992; Stüwe and Sandiford 1995; Mancktelow and Grasemann 1997: eq. 13).

**More General Formulations.** Using a lower boundary condition where we explicitly prescribe the thickness of the mantle part of the lithosphere (as we did in the previous section) is extremely useful to explore the influence of thickness variations of crust and mantle lithosphere on the temperatures in the crust. In order to do this more efficiently, it is useful to introduce two new parameters: the vertical thickening (or thinning) strain of the crust and that of the lithosphere. We call these parameters  $f_c$  and  $f_l$  and will discuss them in some detail in sect. 4.2.3 (s. a. sect. 5.1) (Sandiford and Powell 1990). A value of  $f_c = 2$  means that the crust is twice as thick as in the reference state (Sandiford and Powell 1990; Zhou and Sandiford 1992). Using these parameters, eq. 3.74 can be generalized. All we need to do is multiply the reference crustal and lithospheric thicknesses  $z_c$  and  $z_l$  with their respective thickening strains. We also need to be careful with the thickness of the skin depth of the heat production which also changes with crustal thickening or thinning. Eq. 3.67 becomes:

$$S_{(z)} = S_0 e^{-\frac{z}{h_r f_c}} \quad . \quad (3.75)$$

If we substitute of these generalized formulations for the thicknesses of crust, lithosphere and skin depth into eq. 3.68 and integrate it subject to the same boundary conditions we used for eq. 3.74, we arrive at:

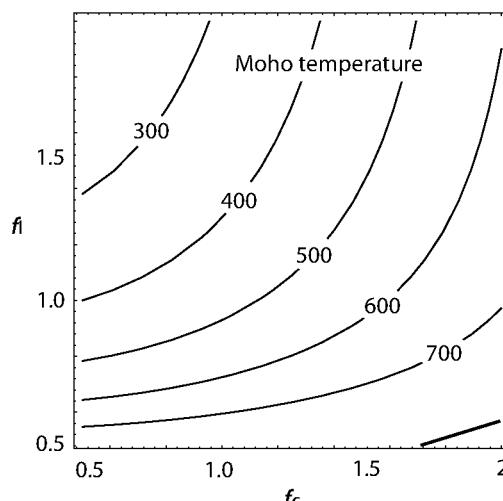
$$T = \frac{zT_l}{f_l z_l} + \frac{f_c^2 h_r^2 S_0}{k} \left( \left( 1 - e^{-\frac{z}{f_c h_r}} \right) - \left( 1 - e^{-\frac{z_l f_l}{f_c h_r}} \right) \frac{z}{f_l z_l} \right) \quad . \quad (3.76)$$

This equation is the most general and most elegant form of a geotherm equation that can be used to calculate temperatures in the entire lithosphere for any thickening geometry of our choice. We only need to keep in mind that the temperatures we calculate are only *potential temperatures* which means that they may not be reached during an orogenic cycle. In the next section we will use this equation to calculate some important temperatures.

- *Moho Temperature from eq. 3.76.* For many thermal and mechanical considerations the Moho-temperature is a meaningful parameter that can be used to characterize the thermal structure of the entire lithosphere (sect. 6.2.2). If we want to calculate this temperature for a series of crustal and lithospheric thicknesses, we must evaluate eq. 3.76 at  $z = z_c$  for a series of  $f_c$  and  $f_l$ . Figure 3.22 shows an  $f_c$ - $f_l$ -diagram of Moho-temperature. It shows that the Moho-temperature does not change very much for homogeneous thickening of the entire lithosphere (which would appear on this figure as a diagonal path from  $f_c = f_l = 1$  towards  $f_c = f_l = 2$ ). This is because the heating effect of the increased thickness of the heat producing crust is almost balanced by the cooling effect of the thickened mantle part of the lithosphere. We have indicated this effect already in the discussion of Fig. 2.18a. However, the strong curvature of the contours on Fig. 3.22 shows that in very thin crusts ( $f_c \ll 1$ ), the Moho-temperature depends largely on the thickness of the mantle part of the lithosphere. Also, in very thick crusts ( $f_c \gg 1$ ), the Moho-temperature depends mostly on crustal thickness. This is because there the Moho-temperature is largely determined by the radioactivity in the crust. Small changes in crustal thickness will lead to quite significant temperature changes, while the thickness of the mantle part of the lithosphere is comparably insignificant.

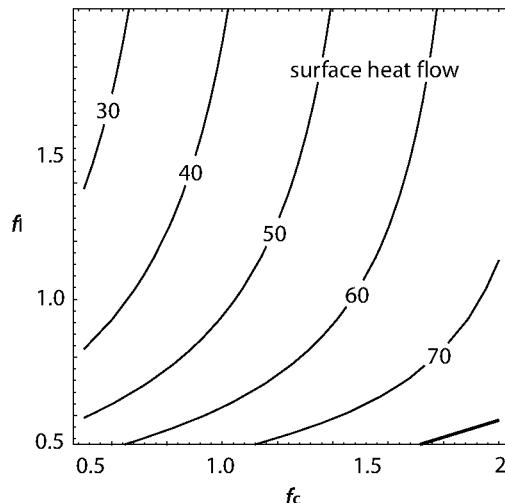
**Figure 3.22.**

Moho-temperatures of continental lithosphere for different crustal thickening strains (expressed by  $f_c$ ) and for different total thickening strains of the lithosphere (expressed by  $f_l$ ). The diagram was calculated with eq. 3.76 assuming  $z = z_c$ . The assumption of the parameters are the same as in Fig. 3.21. The thick line in the bottom right hand corner of the diagram marks the limit of the allowed part of parameter space. It is explained in more detail in the context of Fig. 4.7



- *Surface Heat flow from eq. 3.76.* The surface heat flow  $q_s$  is one of the few important thermal parameters of the lithosphere that can be measured directly. It is therefore a much used parameter to define thermal features in the lithosphere (e.g. Zhou and Sandiford 1992). Within the models discussed above, heat flow is simply given by the derivative of the geotherm equations with respect to depth and divided by the thermal conductivity. The *surface* heat flow  $q_s$  can then be found by evaluating these derivatives at the depth  $z = 0$ . For a boundary condition at the base of the lithosphere (and exponentially decreasing heat production with depth) we get from eq. 3.76:

**Figure 3.23.** The surface heat flow (contoured in  $10^{-3} \text{ W m}^{-2}$ ) for a range of crustal thicknesses (expressed by the vertical thickening strain  $f_c$ ) and for a range of total thicknesses of the lithosphere (expressed by the vertical thickening strain  $f_l$ ). The diagram was calculated with eq. 3.77. The assumptions for all parameters are the same as in Fig. 3.21, curve b



$$q_s = k \frac{T_l}{f_l z_l} + S_0 f_c h_r \left( 1 - \frac{f_c h_r}{f_l z_l} \left( 1 - e^{-\frac{f_l z_l}{f_c h_r}} \right) \right) \quad (3.77)$$

Figure 3.23 shows an  $f_c$ - $f_l$ -diagram, which is contoured for  $q_s$ . We can see that the surface heat flow depends much more on the crustal thickness than the Moho-temperature does (compare Fig. 3.22).

### 3.5 Heat in the Oceanic Lithosphere

Oceanic lithosphere contains practically no radioactive elements. Thus, one could think that it is simple to describe stable oceanic geotherms. In analogy to continental geotherms we might want to formulate the geotherm equation as:

$$k \frac{d^2 T}{dz^2} = 0 \quad . \quad (3.78)$$

After one integration we get:

$$k \frac{dT}{dz} = C_1 \quad .$$

Using the boundary conditions that are well-known to us from sect. 3.4.3 with:  $q = q_m$  at the depth  $z = z_c$ , we get:

$$kT = q_m z + C_2 \quad .$$

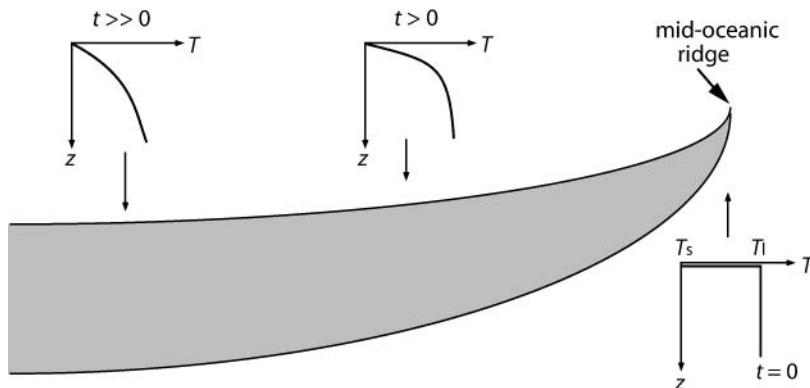
If we also assume that  $T = 0$  at the surface  $z = 0$ , then  $C_2 = 0$  and we can write:

$$T = z \frac{q_m}{k} \quad , \quad (3.79)$$

... if the same boundary conditions as in sect. 3.4.3 are used. However, eq. 3.79 is not a very good model to describe oceanic lithosphere. A simple consideration of the time scale of conductive equilibration will show us why: oceanic lithosphere is produced at the mid-oceanic ridges and it gets its thickness only by its increasing age. The oldest oceanic lithosphere is about 150 my old (Fig. 2.4). However, in sect. 3.1.4 we showed that the characteristic time scale of equilibration for thermally stabilized lithosphere is of the order of 150 my or more! We can conclude that oceanic lithosphere is *not* thermally stabilized. The assumption underlying eq. 3.78 is wrong. There is no thermally stabilized oceanic lithosphere! We can not assume that  $dT / dt = 0$  and we must solve the time dependent diffusion equation (eq. 3.6).

### 3.5.1 Aging Oceanic Lithosphere

The oceanic crust, that is produced from partial mantle melts at the mid-oceanic ridges is only of the order of 5–8 km thick. That is, it is much thinner than the continental crust. At the mid-oceanic ridge the thickness of the entire oceanic lithosphere is that of the crust (Fig. 3.24). The high potential energy of the ridges forces this crust to move away from the ridge. As the oceanic crust ages and moves further and further away from the mid-oceanic ridge, the asthenosphere cools and becomes part of the oceanic mantle lithosphere. It is often said that the mantle successively “freezes” onto the base of the oceanic lithosphere as it ages. While this describes the process quite intuitively, it is somewhat incorrect as the asthenosphere itself is not molten. Regardless, the process of the successive cooling of the aging oceanic lithosphere can be described with the diffusion equation using simple initial and boundary conditions. Indeed, the description of oceanic lithosphere with these boundary conditions has become one of the most successful models of plate tectonic theory (s. a. Sclater et al. 1980). It is called the *half space cooling model* (s. sect. 3.6.1).



**Figure 3.24.** Thickness and thermal profile of oceanic lithosphere at a series of points. The oceanic crust is not drawn separately. On the scale of this sketch it would appear as a very narrow band of constant thickness near the upper boundary of the gray shaded region

**The Half Space Cooling Model.** As any other problem in the theory of heat conduction, the half space cooling model relies on the integration of eq. 3.6, using a set of boundary and initial conditions. These conditions are provided by geological observation: The temperature at the surface of mid-oceanic ridges is that of the water temperature. For simplicity, we assume that it is  $T_s = 0$ . Below the ridge, the mantle temperature is almost constant – convection equalizes all temperature gradients. Thus, we can write a very simple initial condition describing the thermal profile below mid-oceanic ridges:

- $T = T_s$  at the depth  $z = 0$  and:
- $T = T_l$  in all depths  $z > 0$  at time  $t = 0$ .

This initial condition is illustrated in  $T$ - $z$ -diagram on the bottom right corner of Fig. 3.24. For the upper boundary condition it is obvious to assume that the temperature at the ocean floor remains constant. As there is effectively no lower boundary, we assume that it lies at infinity and that the temperature there is  $T = T_l$ . We can write this as follows:

- $T = T_s$  at  $z = 0$  for all  $t > 0$  and:
- $T = T_l$  at  $z = \infty$  for all  $t > 0$ .

(Fig. 3.24). The solution of the heat conduction equation for these boundary conditions is:

$$T = T_s + (T_l - T_s) \operatorname{erf} \left( \frac{z}{\sqrt{4kt}} \right) . \quad (3.80)$$

This solution is already a bit familiar to us from section 3.1.3 and we discuss this solution in some more detail in sect. 3.6.1 (s. a. sect. 3.1.3). Fig. 3.25a

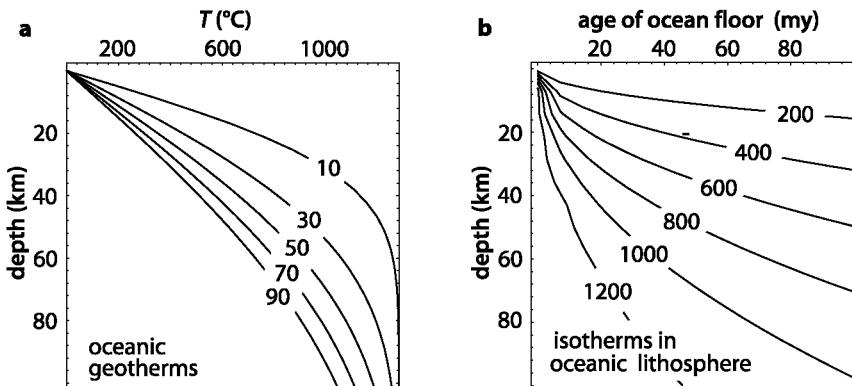
shows temperature profiles through oceanic lithosphere of different ages, that were calculated with eq. 3.80. The curves correspond to the two sketches of thermal profiles in the middle and on the left of Fig. 3.24. Fig. 3.25b shows the depth of a series of isotherms as a function of age.

**Surface Heat Flow: The Test for the Model.** Temperature profiles calculated with this model for the cooling oceanic lithosphere can not be tested directly, as we cannot drill deep enough into the oceanic lithosphere. Our observations are confined to parameters which we can measure near the surface. One of these parameters is easy to measure and very useful to infer the thermal profile: the surface heat flow (Pollack et al. 1993) (Fig. 3.1). The surface heat flow is the product of thermal conductivity and the thermal gradient at  $z = 0$ . This can be calculated from eq. 3.80 and can be compared with measured data in the oceans. To obtain surface heat flow we must differentiate eq. 3.80 with respect to depth and evaluate it at  $z = 0$ . From eq. 3.80 this is:

$$q_s = k(T_l - T_s) \frac{d \left( \operatorname{erf} \left( \frac{z}{\sqrt{4\kappa t}} \right) \right)}{dz} \Big|_{z=0} . \quad (3.81)$$

As the error function itself is an integral (see eq. 3.15), it is easy to differentiate eq. 3.80 (sect. 3.1.3). We get:

$$q_s = k(T_l - T_s) \sqrt{\frac{1}{\pi \kappa t}} . \quad (3.82)$$



**Figure 3.25.** **a** Temperature  $T$  versus depth profiles through oceanic lithosphere at a number of different ages labeled in my. **b** The depth of isotherms (in °C) in oceanic lithosphere as a function of age between 0 and 100 my. The curves on both figures were calculated with eq. 3.80 assuming  $T_l = 1280$  °C. The age can be converted into distance from the mid-oceanic ridge by using  $x = u/t$  where  $x$  is the distance from the ridge and  $u$  is the rifting rate. Compare the curves also with Fig. 3.40

This equation can be generalized for the description of different oceanic plates with different rifting rates. For this, we express the rifting rate  $u$  as  $u = x/t$ . There,  $x$  is the distance from the mid-oceanic ridge and  $t$  is the age. If we replace  $t$  in eq. 3.82 by  $x/u$ , we get:

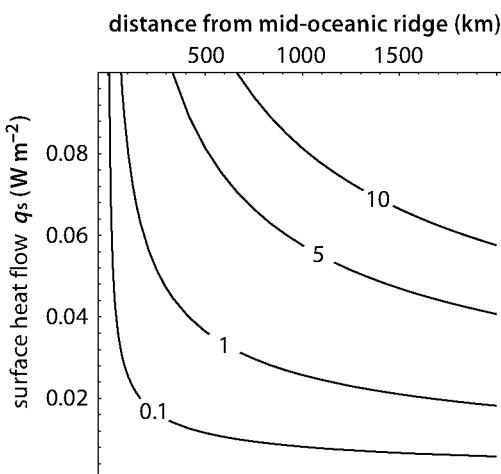
$$q_s = k(T_l - T_s) \sqrt{\frac{u}{\pi \kappa x}} . \quad (3.83)$$

Eq. 3.83 shows us that the surface heat flow as a function of distance from the mid-oceanic ridge is a square root function. Fig. 3.26 shows the surface heat flow in oceanic lithosphere as calculated with eq. 3.83. The heat flow data of Sclater et al. (1980) show that these curves correspond well with heat flow measured in the deep oceans. In sect. 4.4.1 we will show that the half space cooling model is not only a good description for the temperatures and heat flow in oceanic lithosphere, but can also be used to describe the water depth of the oceans. It can even be used to calculate the magnitude of the ridge push force (sect. 5.3.2). The relationship between all these parameters that are described with the half space cooling model is called the *age-depth-heat flow relationship* of oceanic lithosphere. This age-depth-heat flow relationship corresponds fantastically well with our observations up to an age of the oceanic lithosphere of 50 – 80 my. The age-depth-heat flow relationship is generally accepted as one of the greatest successes of plate tectonic theory.

*Alternative Models for the Aging Oceanic Lithosphere.* The half space cooling model is the most famous model for the description of oceanic lithosphere, but there are several reasons why it may not be a perfect description:

- The half space cooling model assumes implicitly that the heat flow at the base of the lithosphere changes through time. It is not clear that this should be so.

**Figure 3.26.** The surface heat flow of oceanic lithosphere as a function of age and therefore as a function of distance from the mid-oceanic ridge as calculated with eq. 3.83. Contours are for different rifting rate labeled in  $\text{cm y}^{-1}$

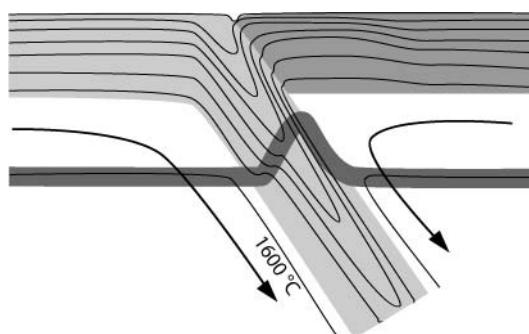


- The model fails to describe the water depth and the surface heat flow of oceanic lithosphere beyond an age of about 80 my.

Moreover, various recent studies have shown that the increase in water depth and changes in heat flow with distance from the ridge deviate a bit from the half space cooling model and that there is a relationship between seafloor subsidence rate and height of the ridge (Dumoulin et al. 2001). One possible alternative model is to assume that the mantle heat flow through the base of the lithosphere is kept constant by small scale convection (Doin and Fleitout 1996). This assumption leads to a description which corresponds much better with the long term evolution of oceanic lithosphere than the half space cooling model. This idea is known by the name of *CHABLIS-model*. This name is an acronym for: **C**onstant **H**eat **f**low **A**ppled to **B**ottom of **L**ithospheric **I**sotherm. Some authors have also argued that the observed seafloor depth and heat flow can be better explained if small amounts of radiogenic heat production of the thin oceanic crust are considered (Hillier and Watts 2005).

### 3.5.2 Subduction Zones

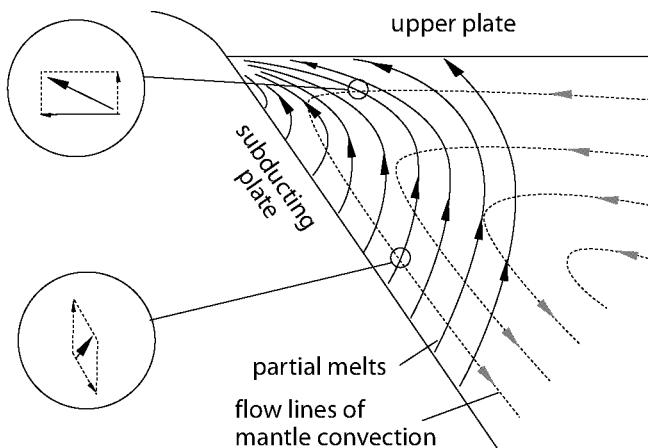
The description of the kinematics, thermal evolution and dynamics of subduction zones is a fundamentally two-dimensional problem. It is the first problem in this book for which we require more than one spatial coordinate to characterize the essence of the problem (Fig. 3.27). Many problems related to subduction zones concern the accretionary wedge that forms near the surface between the surface of the subducting slab and the upper plate. Dynamic and kinematic models for accretionary wedges will be discussed in sect. 6.2.3 and also in sect. 5.3.2. Here we discuss some general aspects about the deep thermal structure of subduction zone environments.



**Figure 3.27.** Schematic illustration of the temperature distribution in subduction zones. The subducted lithosphere is shaded light, the upper plate dark. The thick dark line that follows the  $1600\text{ }^{\circ}\text{C}$ -isotherm outside the subduction zone is the Clapeyron-curve. It marks the olivine-spinel phase transition

**Isotherms in Subduction Zones.** Fig. 3.27 shows, schematically, the shape of isotherms in subduction zones. In the subducting slab, the isotherms will be bent and subducted with the slab. The further they are subducted, the more they merge to the center of the slab as both surfaces of the subducting slab equilibrate with the surrounding mantle temperatures. A thermal steady state will be reached when the curvature of the isotherms is large enough so that the rate of thermal equilibration is balanced by the subducting velocity (s. Fig. 3.3; s. Molnar and England 1995). In this stage, *diffusion* (which leads to the decay of the high curvature of the temperature profile in the tip of the subducting slab) will be balanced by *advection* (which moves the isotherms to larger depths). A steady state is reached that is very similar to the balance discussed in the second comparison on page 3.2 and also equivalent to the steady state that landforms between incising drainages may reach (discussed in Fig. 4.37). The time that is needed by subduction zones to reach this thermal steady state depends on the thickness of the plate and on the subduction rate. It can be estimated with the Peclet number (sect. 3.3.4).

At a temperature of about  $1600^{\circ}\text{C}$ , which is in about 400 km depth, olivine reacts to form spinel. The depth of this phase transition is called the *Clapeyron-Curve*. This reaction is *exothermic* with about  $1.7 \cdot 10^5 \text{ J kg}^{-1}$ . Thus, the isotherms in this depth have a kink. The positive slope of the Clapeyron curve in *P-T*-space causes that the Clapeyron curve is somewhat higher within the subducting slab than it is outside. At a depth of 650 km (about  $1700^{\circ}\text{C}$ ) there is another kink in the geotherm, caused by the phase transition from spinel to oxide (not illustrated on Fig. 3.27). The qualitative



**Figure 3.28.** The motion of melts in and above the Benioff zone according to Spiegelman and McKenzie (1987). The dashed lines show the convective motion in the asthenosphere, the continuous lines are the motion of the partial melts. The enlarged sections show how the velocity field of the partial melts is given by the sum of the upward velocity of the melt and the motion in the mantle wedge

direction of this kink is opposite to that of the Clapeyron curve as the spinel-oxide phase transition is *endothermic*.

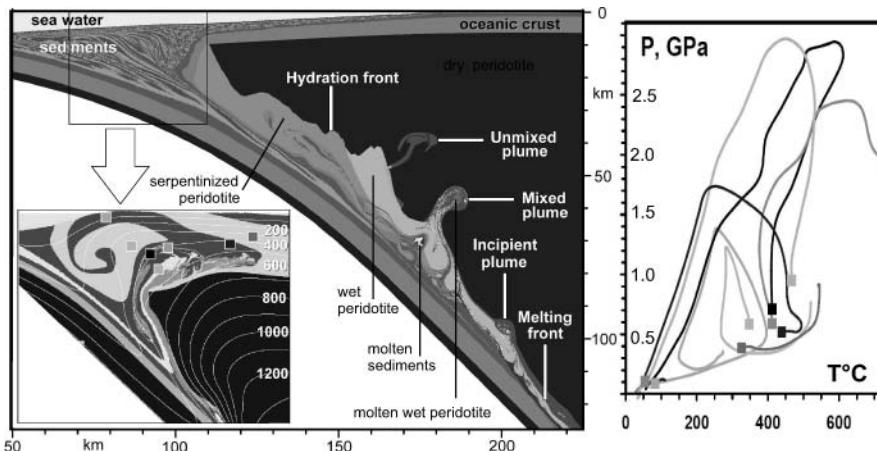
Fig. 3.27 shows schematically that the isotherms within the upper plate are closer to the surface near the subduction zone than they are in the far field. This is because of the dehydration of the subducting plate and the consequential rise of partial melts and other hot fluids. This leads typically to high temperature metamorphism in the rocks of the upper plate and ultimately to the development of magmatic arcs (see next section). This is in contrast to the very *low* temperatures that occur in the subducting plate up to very large depths. This coupled occurrence of low-pressure - high-temperature and high-pressure - low-temperature metamorphism was recognized by Miyashiro (1973) as one of the characteristic features of metamorphic terrains in subduction zone environments. He called this a *paired metamorphic belt*.

**Island Arcs and Subduction Zones.** An interesting observation in the upper plate of subduction zones is that there is volcanic arcs that always form a narrow line that is exactly where the seismically active surface between the subducting slab and the upper plate is about 150 km deep (Isacks and Barazangi 1977). In subduction zones that have a dip of  $45^\circ$  this implies a horizontal distance of the arc from the trench of 100–150 km. If the subduction angle is steeper, then this distance is shorter and vice versa. This observation is true for the distance of the Aleute-volcanoes to the Aleute Trench, for the distance of the Indonesian volcanoes from the Java Trench and many other volcanic arcs around the globe. This observation is not trivial to explain. The volcanics that erupt from these volcanoes are derived from partial melts in the mantle wedge that melted during fluid infiltration of fluids that were derived by dehydration of sediments on the surface of the subducting slab in the Benioff-zone. This zone stretches for several hundreds of kilometers along the surface of the subducting slab and is definitely much wider than the width of the volcanic arcs on the surface. Some authors have suggested that there are important pressure sensitive dehydration reactions that occur in exactly 150 km depth, but there is little petrological evidence for this.

An alternative explanation was suggested by Spiegelman and McKenzie (1987) (Fig. 3.28). Their model describes the motion of partial melts through the mantle wedge as the sum of two vector fields:

1. The motion of the asthenosphere in the mantle wedge. This motion follows the wedge and is illustrated in Fig. 3.28 by the dashed lines.
2. The motion of the partial melts. Partial melt is produced continuously along the surface of the subducting plate and moves vertically upwards.

The sum of the two velocity fields results in curved paths that converge at the tip of the mantle wedge (Fig. 3.28). This elegant model is a beautiful example for a successful model description of fluid flow in deforming rocks.



**Figure 3.29.** Integrated thermomechanical model of Gerya and Yuen (2003) to describe subduction related processes. The shown model also considers metamorphic processes, for example dehydration reactions and is set up to investigate the development of mixed and unmixed plumes in the mantle wedge in response to slab dehydration (Gerya pers. comm. 2006). Different gray shades are different layers. The inset at left shows more detail in the accretionary wedge but is actually from a different model run. The  $P$ - $T$  diagram at right shows  $P$ - $T$  paths of different rocks in the wedge as indicated by the symbols in the inset at left. Note that crustal rocks can easily reach 2.5 GPa by being dragged down along the surface of the subducted slab. Because they are then rapidly exhumed within the same wedge, they have not found time to heat up and remain in the eclogite and blueschist facies (s. Fig. 7.1)

**$P$ - $T$  Paths in Subduction Zones.** Although the model shown in Fig. 3.28 is extremely elegant, a range of observations and new modeling results have made it somewhat redundant. Geologists have noted that arc volcanoes are not as narrow as thought by Isaacks and Barazangi (1977) and our recent understanding of dehydration processes on subducted slabs have indicated that the temperature range of dehydration may be small enough to cause some focusing of fluids without having to appeal to the model of Fig. 3.28. One of the most sophisticated models that have recently been used to describe such processes is the model of Gerya and Yuen (2003) (Fig. 3.29).

### 3.6 Thermal Evolution of Intrusions

Intrusion of magmatic rocks into higher levels in the crust is an important geodynamic process that can be responsible for a large range of thermal, chemical and mechanical changes in the crust. Intrusive rocks, as well as their contact aureoles, are familiar to us from field observations. Thus, their geodynamic interpretations can often be tested directly with structural and petrological data. The process of intrusion itself is a very efficient mechanism

for the transport of heat. Thus, most intrusions do not cool very much on their intrusive path and their temperature may be used to infer the temperature of rocks at the depth of their origin. There are two important reasons, why the quantitative description of the thermal evolution of intrusions is quite simple:

- In comparison to the duration of a contact metamorphic event or the duration of an orogenic cycle, the rate of intrusion is very rapid. Thus, for the thermal modeling of intrusions it is possible to assume that their emplacement was *infinitely* rapid, compared to the time of the subsequent thermal equilibration. This is what is called an *instantaneous heating model*. The thermal equilibration of intrusions with their surroundings can therefore be described with the heat conduction equation (eq. 3.6), assuming the intrusion geometry as an “initial condition”.
- Most intrusions are small if compared to the size of their surroundings, for example the distance to the surface or to the base of the lithosphere. Thus, the boundary conditions that are needed to solve the heat conduction equation can often be assumed to lie at infinity in comparison to the scale of the problem.

With these two assumptions, it is possible to solve the heat conduction equation. The solutions may then be used for the analytical description of the thermal evolution of intrusions. For simple intrusion geometries, for example dikes, the description is very similar to description of the thermal evolution of oceanic lithosphere (sect. 3.5.1). However, before we use such a model, we need formulate two more simplifying assumptions which we will make initially:

- The latent heat of fusion - which actually plays an important role for the thermal energy budget of intrusions - will be neglected. Thus we will analyze in the first instance none of the processes discussed in sect. 3.2.
- The thermal conductivity is assumed to be constant in space. Thus we will also not consider any of the problems discussed in sect. 3.1.1.

With these assumptions, it is possible to integrate eq. 3.6. We begin with some simple examples.

### 3.6.1 Step-shaped Temperature Distributions

The most simple of all model examples describing the cooling of rocks in the direct vicinity of intrusions is given by the thermal equilibration of step-shaped temperature profiles in one dimension. This example is illustrated in Fig. 3.30 and is one of the most useful examples for the understanding of the cooling history of intrusions. We interpret the temperatures on the two sides of the step as the intrusion- and the host rock temperatures;  $T_i$  and

$T_b$ , respectively. The step itself is the intrusion contact. If we choose a one-dimensional coordinate system in which the origin  $z = 0$  is exactly at the contact of the model intrusion, then the initial and boundary conditions of this equilibration problem may be described by:

- Initial condition:  $T = T_i$  for all  $z > 0$  and  $T = T_b$  for all  $z < 0$  at  $t = 0$ .
- Boundary conditions:  $T = T_b$  at  $z = -\infty$  and  $T = T_i$  at  $z = \infty$  for all  $t > 0$ .

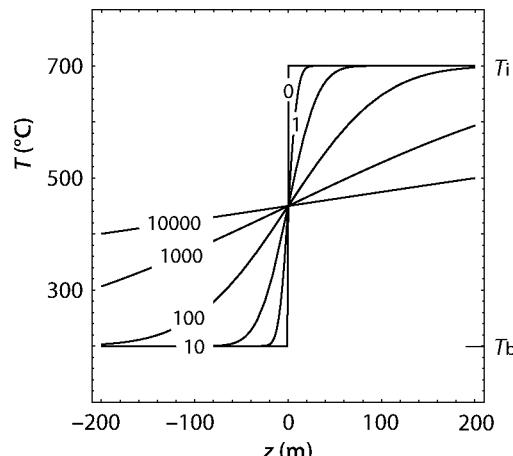
You may have noticed that there is *two* boundary conditions that are located a very long distance from the contact *inside* the intrusion and far away from the contact in the host rock (at  $z = +\infty$  and  $z = -\infty$ ). We need *two* boundary conditions because the equation that is to be solved (eq. 3.6) is a partial differential equation of the *second* order (s. sect. B.1.1). Integration of eq. 3.6 using these boundary and initial conditions gives:

$$T = T_b + \frac{(T_i - T_b)}{2} \left( 1 + \operatorname{erf} \left( \frac{z}{\sqrt{4\kappa t}} \right) \right) . \quad (3.84)$$

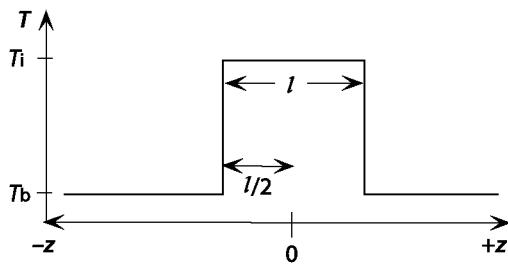
We will not discuss how eq. 3.84 was derived (s. sect. 3.1.3) but when you compare this equation to eq. 3.80 you will see that it is very similar and differs from the most simple form of the half space cooling model only by some shifting and scaling of the error function. Because of the choice of coordinate system we have made to formulate this problem (Fig. 3.30), this solution looks very simple. In another coordinate system in which the coordinate origin is located at a distance  $l$  from the temperature step, the initial condition must be reformulated to:  $T = T_i$  for all  $z > l$  and  $T = T_b$  all  $z < l$  at  $t = 0$ . The boundary conditions remain the same. The solution gets the form:

$$T = T_b + \frac{(T_i - T_b)}{2} \left( 1 + \operatorname{erf} \left( \frac{z - l}{\sqrt{4\kappa t}} \right) \right) . \quad (3.85)$$

**Figure 3.30.** Thermal equilibration of an initially step-shaped temperature distribution. The curves were calculated with eq. 3.84 for  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ ,  $T_i = 700^\circ\text{C}$  and  $T_b = 200^\circ\text{C}$ . The different curves are temperature profiles at different times (in years) after the intrusion event



**Figure 3.31.** Schematic illustration of the initial condition and the variables of the dike cooling problem of eq. 3.88



Again, in comparison to eq. 3.85 you can see that this solution only shifts the temperature profile to the right place. On Fig. 3.30 it may be seen that the thermal evolution on both sides of the mean temperature between  $T_i$  and  $T_b$  develops symmetrically. This is to be expected as the initial and boundary conditions are also symmetric.

- *Cooling of half spaces.* In the example of eq. 3.84, the temperature at  $z = 0$  stays constant in time. It has the value  $T_b + (T_i - T_b)/2$ . Thus, this problem is very similar to the description of the cooling of semi-infinite half spaces (sect. 3.5.1). This thermal problem is very important in the earth sciences and may be described by the following initial and boundary conditions:

- Initial condition:  $T = T_b$  at all  $z > 0$  and  $T = T_i$  at  $z = 0$  at time  $t = 0$ .
- Boundary conditions:  $T = T_i$  at  $z = 0$  and  $T = T_b$  at  $z = \infty$  for all  $t > 0$ .

These conditions are very similar to those of the last problem. Thus, the corresponding solution is also similar to eq. 3.84:

$$T = T_b + (T_i - T_b) \left( 1 - \operatorname{erf} \left( \frac{z}{\sqrt{4\kappa t}} \right) \right) . \quad (3.86)$$

We have met this equation in a slightly different form already in sect. 3.5.1. The result of eq. 3.80 and eq. 3.86 can be simplified if it is expressed as the dimensionless temperature  $\theta = (T - T_b)/(T_i - T_b)$  (sect. 1.2). Then, using the complementary error function, eq. 3.86 simplifies to:

$$\theta = \operatorname{erfc} \left( \frac{z}{\sqrt{4\kappa t}} \right) . \quad (3.87)$$

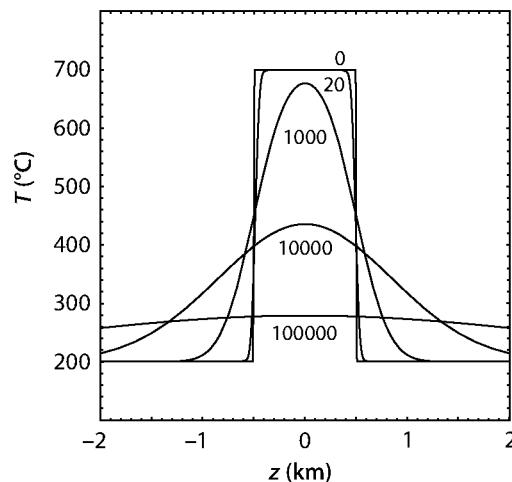
This simplification is shown here to illustrate that cooling curves of this problem have the shape of an error function. However, in the remainder of this section we will not use dimensionless temperatures. We begin with some geologically relevant examples that may be described with this solution.

### 3.6.2 One-dimensional Intrusions

One of the simplest but also most important applications of the equations introduced in the last section is the description of the cooling history of

intrusions (Jaeger 1964). As the solutions shown above are one-dimensional, their application is particularly relevant to the description of the thermal evolution of dikes that are narrow compared with their lateral extent. When using the solutions described above to describe the cooling history of dikes, it is implied that the dike extends “infinitely” in the two spatial directions normal to the coordinate described in the cooling problem. In contrast to the previous sections, we only need to be careful to consider both surfaces of the dike. For a coordinate system with its origin in the center of a dike with the thickness  $l$ , the initial conditions may be described by:  $T = T_i$  for  $-(l/2) < z < (l/2)$  and  $T = T_b$  for  $(l/2) < z < -(l/2)$  (Fig. 3.31). The boundary conditions remain the same as for the step problem. With these conditions, a solution of eq. 3.6 may be found to be:

**Figure 3.32.** Thermal equilibration of one-dimensional intrusions, for example magmatic dikes of large lateral extent (calculated with eq. 3.88 and labeled in years after initial intrusion). All parameters are the same as in Fig. 3.30. Cooling curves of rocks from a range of distance from the dike center are shown in Fig. 3.33a



$$T = T_b + \frac{(T_i - T_b)}{2} \left( \operatorname{erf} \left( \frac{0.5l - z}{\sqrt{4\kappa t}} \right) + \operatorname{erf} \left( \frac{0.5l + z}{\sqrt{4\kappa t}} \right) \right) . \quad (3.88)$$

It may be easily seen, that the solution is made up of descriptions for two opposing step-shaped temperature profiles at  $z = -l/2$  and  $z = l/2$ . Fig. 3.32 shows the thermal evolution described by eq. 3.88. As the diffusion equation is a linear differential equation, the diffusive equilibration of just about any one-dimensional geometry may be described by the summation of solutions for various initial conditions.

In contrast to Fig. 3.30, the temperature at the intrusion contact departs from the temperature  $(T_i + T_b)/2$  after some time in Fig. 3.32. The contact of the dike begins to cool. This is because the dike contact at  $z = +l/2$  begins to follow the thermal effects of the temperature step at  $z = -l/2$ . Correspondingly, the other dike contact at  $z = -l/2$  cools, because it “feels” the cooling at  $z = +l/2$ .

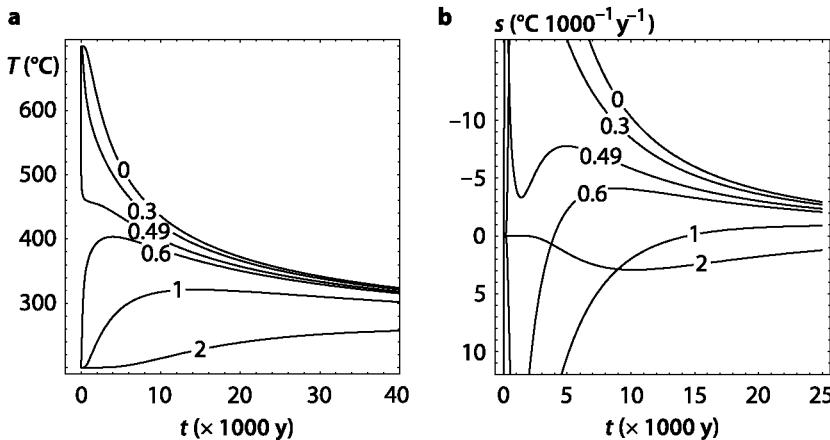
**Cooling History of Simple Intrusions.** In the following paragraphs we will use eq. 3.88 to infer some characteristic features of contact metamorphism. Firstly, eq. 3.88 shows us that – in the absence of other thermal processes – the maximum temperature that may be reached by contact metamorphism is much lower than the intrusion temperature: Only at the very contact of the intrusion the contact metamorphic temperatures may reach the half way mark between the initial host rock and intrusion temperature. We can conclude that field observations of contact metamorphic haloes documenting haloes of considerable width and temperature imply that thermal processes other than conductive equilibration have played a role in their formation (s. sect. 3.6.4).

- *Cooling curves.* In order to interpret heating and cooling curves of rocks in the contact metamorphic environment, it is useful to plot eq. 3.88 in a temperature-time diagram (Fig. 3.33a). This figure illustrates that rocks located at different distances from the intrusion may experience very different cooling curves. For example, it may be seen that some rocks cool, while others heat up, or that rocks cool with different rates. In fact, near the contact of the intrusion, cooling curves have extremely complicated shapes including more than one maximum in the cooling rate (e. g. the 490-m-curve in Fig. 3.33a). Spend some time and think through why these different shapes come about. As simple as the model is, it teaches us much about the equilibration of contact metamorphic aureoles: Many metamorphic processes, for example grain growth or cation diffusion, are strongly dependent on both temperature and cooling rate (e. g. Dodson 1973). The interpretation of these curves is therefore extremely important (s. sect. 7.2).
- *Cooling rates.* The cooling rate  $s$  for a chosen point near the intrusion is described by the time derivative of eq. 3.88. As the error function itself is an integral, it is not too difficult to find this derivative, although we will not explain it in detail here. It is:

$$s = \frac{dT}{dt} = \frac{(T_i - T_b)}{4t\sqrt{\pi\kappa t}} \left( \frac{z - 0.5l}{e^{((0.5l-z)^2/4\kappa t)}} - \frac{z + 0.5l}{e^{((0.5l+z)^2/4\kappa t)}} \right) \quad (3.89)$$

(s. sect. 3.1.3, 3.5.1). Cooling rates as a function of time as calculated with this equation are shown in Fig. 3.33b. They correspond directly to the curves in Fig. 3.33a.

For many petrological questions the cooling rate at a given *temperature* is much more important than the cooling rate at a given *time*. For this, a *parametric* plot of temperature against cooling rate is useful. *Parametric* plots are diagrams in which two independent functions of the same variable (here: time) are plotted against each other. Thus, the parametric plot of  $T$  against  $s$  shown in Fig. 3.34 is a combination of Figs. 3.33a and 3.33b. The time dependence of temperature or cooling rate can not be illustrated on this figure. Thus, Fig. 3.34 may appear a bit confusing at first view. Nevertheless,

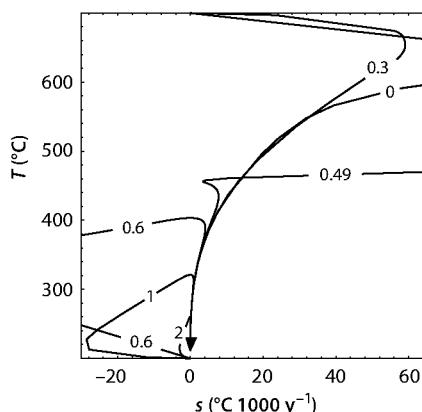


**Figure 3.33.** **a** Temperature-time paths (cooling curves) for a series of rocks within and within the contact aureole of a 1 km thick magmatic dike ( $l=1$  km) shown up to 40 000 years after intrusion (all constants are the same as in Fig. 3.32). The figure was calculated with eq. 3.88. The curves are labeled for distance from the center of the intrusion in km. As the thickness is  $l=1$  km, the first three curves are within the magma and the others in the country rock. **b** Cooling rates for the same points as shown in **a** plotted against time. Curves are calculated with eq. 3.89

such diagrams are crucial for a meaningful interpretation of petrological and geochronological data.

- *Contact Metamorphic Peak.* Aside from cooling curves or cooling rates, there is even more important information on the thermal evolution of intrusions that may be extracted from eq. 3.88. For example, the *time* of the contact metamorphic thermal peak  $t_{T_{\max}}$  for the model of eq. 3.88 may also be found analytically. At the thermal peak the rate of temperature change is zero:  $s|_{t=T_{\max}} = 0$  (read:  $s$  at  $t = T_{\max}$ ). Thus, this time is given by setting eq. 3.89 to zero and solving for time. We get:

**Figure 3.34.** Parametric plot of cooling rate against temperature. The diagram was constructed from Figs. 3.33a and 3.33b. Although the diagram may not appear intuitive at first, it is simply a combination of the curves from Fig. 3.33a,b and is extremely useful for the interpretation of the equilibration of mineral parageneses (sect. 7.2)



$$t_{T_{\max}} = -\frac{lz}{2\kappa \ln \left( \frac{z-0.5l}{z+0.5l} \right)} . \quad (3.90)$$

From this, we can also get the thermal peak temperature as a function of distance from the intrusion. We get this by substituting the time obtained from eq. 3.90 into eq. 3.88 and solving that for temperature (Fig. 3.34; Fig. 3.35). With this diagram we can make some fundamental predictions about the nature of contact metamorphism:

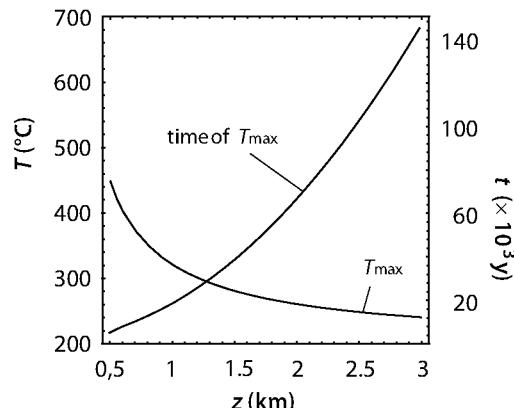
1. The contact metamorphic peak temperature drops rapidly with *increasing* distance from the heat source.
2. The time of peak contact metamorphism increases rapidly with *increasing* distance from the heat source and with *decreasing* metamorphic grade. This predicts that – if contact metamorphism occurred – low grade metamorphic rocks should experience their metamorphic peak later than high grade metamorphic rocks (Den Tex 1963).

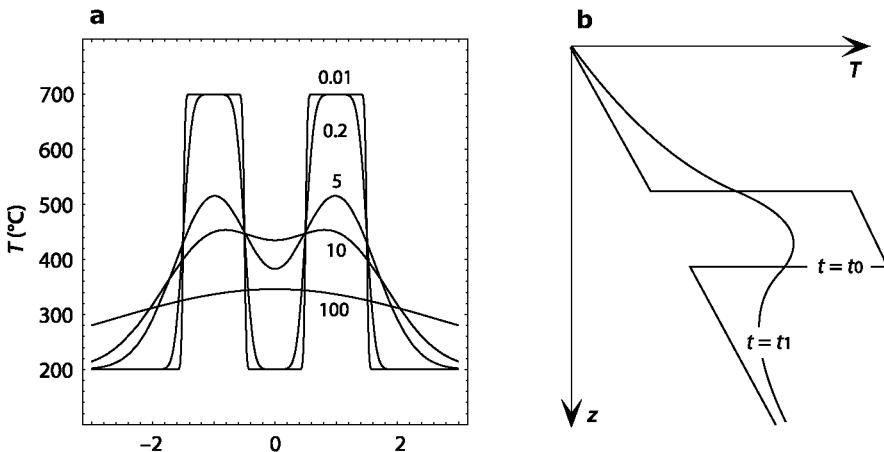
These predictions help to infer heating mechanisms of metamorphic terrains. For example, during regional metamorphism the relationships between grade and timing of metamorphism are reversed (sect. 6.3.3 and sect. 7.4.1). In the chapter on *P-T-t*-paths we will discuss further implications of the timing relationships of various cooling curves.

**More Complicated One-dimensional Geometries.** One of the largest drawbacks of the model discussed in the last section (eq. 3.88 and its derived relationships) is that it predicts much narrower contact metamorphic aureoles than those that are generally observed around intrusions. One of the reasons for this may lie in the geometry of the heat source. Other causes will be discussed in sect. 3.6.4.

- **Dike swarms.** In terrains that are penetrated by many intrusions, the mean contact metamorphic temperature may be much higher than that would be observed around a single intrusion. This may be described schematically by

**Figure 3.35.** Contact metamorphic peak temperature and the time of the contact metamorphic temperature peak of the simple one-dimensional intrusion from Figs. 3.31 and Fig. 3.32. Calculated by substituting eq. 3.90, into eq. 3.88. The thickness of the intrusion  $l$  is 1 km,  $T_i = 700^\circ\text{C}$ ,  $T_b = 200^\circ\text{C}$  and  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$





**Figure 3.36.** Two examples of complicated one dimensional intrusion geometries. **a** Contact metamorphism between two neighboring intrusions. The figure was calculated with eq. 3.91 using  $z_1 = -1000$ ,  $z_2 = 1000$  and  $l_1 = l_2 = 1000$ . Curves are shown in thousands of years (e.g. the line labeled for 0.2 corresponds to 200 years). Note that the contact metamorphic aureole *between* the intrusions is much wider than on their outside. **b** Schematic illustration of the diffusive equilibration of a step-shaped temperature profile (for example a one-dimensional intrusion like a sill) superimposed on a linear geotherm. Temperature profiles are shown at two different time steps  $t_0$  and  $t_1$ . As the diffusion equation is a linear differential equation, the temperatures of eq. 3.88 can simply be added to the geotherm equation (in this case a straight line)

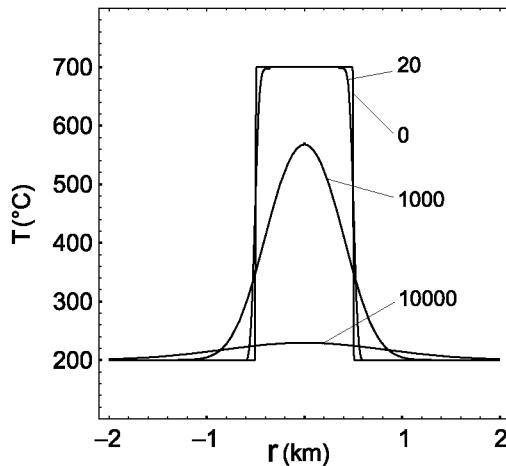
the combination of a series of temperature steps. If there is  $N$  intrusions of the thicknesses  $l_n$ , that intrude at the depths  $z_n$ , then this can be described by the summation of  $2N$  steps in the temperature distribution. The solution is:

$$\theta = \frac{1}{2} \sum_{n=1}^N \left( \operatorname{erf} \left( \frac{z - (z_n - 0.5l_n)}{\sqrt{4\kappa t}} \right) + \operatorname{erf} \left( \frac{(z_n + 0.5l_n) - z}{\sqrt{4\kappa t}} \right) + \dots \right) \quad (3.91)$$

where  $\theta$  is again the normalized temperature  $\theta = (T - T_b)/(T_i - T_b)$ . Fig. 3.36a shows an example of two intrusions calculated with this equation. It may be seen that the contact metamorphic halo between the intrusions is much wider and of a higher temperature than on their outside. Because of this observation, Barton and Hanson (1989) suggested that multiple intrusion may be the principle heat source for low-pressure high-temperature metamorphism (s. sect. 6.3.3). Fig. 3.36b shows another schematic example for a one-dimensional problem that may be solved by the summation of a series of one-dimensional model geometries (s. p. 58).

• *Spherical intrusions.* The thermal evolution of spherical intrusions is also a one-dimensional problem in polar coordinates. The equation we need to solve

**Figure 3.37.** Cooling history of a spherical intrusion. For easy comparison, all chosen parameters and the shown time steps (labeled in years) are the same as those in Fig. 3.32. Calculated with eq. 3.92.



is eq. 3.12. If we want to compare the thermal evolution of spherical intrusions with that described by eq. 3.88, then we formulate the following boundary conditions:  $T = T_b$  at  $r = \infty$  and  $r = -\infty$ , and the initial condition  $T = T_i$  in the region  $-R < z < R$  and  $T = T_b$  in the region  $-R > z > R$  at  $t = 0$ , where  $R$  is the radius of the intrusion and is analogous to  $l/2$  in the Cartesian example discussed in eq. 3.88.  $r$  is the distance from the origin. The solution of eq. 3.12 under these conditions is:

$$T = T_b + \frac{(T_i - T_b)}{2} \left( \operatorname{erf} \left( \frac{R-r}{\sqrt{4\kappa t}} \right) + \operatorname{erf} \left( \frac{R+r}{\sqrt{4\kappa t}} \right) \right) - \frac{(T_i - T_b)}{r} \sqrt{\kappa t / \pi} \left( e^{-\left( \frac{(R-r)^2}{4\kappa t} \right)} - e^{-\left( \frac{(R+r)^2}{4\kappa t} \right)} \right) . \quad (3.92)$$

Temperature profiles across cooling spheres at different times are shown in Fig. 3.37. They may be directly compared with those in Fig. 3.32. It may be seen that spheres cool much faster than dikes, which is intuitively clear as they have a much larger ratio of surface to volume in comparison to one dimensional steps in Cartesian coordinates.

### 3.6.3 Two-dimensional Intrusions

The description of two-dimensional thermal problems is – for most boundary conditions – not much harder than one-dimensional problems because heat conduction in several spatial dimensions can be described as the sum of conduction in the individual directional components (s. eqs. 3.9 to 3.10). In general it can be said that two-dimensional heat conduction problems may be described by the product of the one-dimensional solutions if: 1. the initial conditions may be expressed as the product of two functions  $f(z)$  and  $f(y)$

and 2. the boundary conditions are given by a constant temperature or constant heat flow. Solutions of the two-dimensional heat conduction equation with  $z$  and  $y$  as the spatial coordinates can therefore be expressed as:

$$T = \text{solution in } z\text{-direction} \times \text{solution in } y\text{-direction}$$

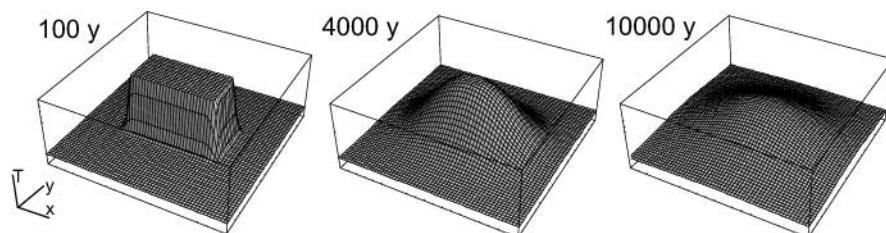
For example, the cooling history of a *corner* is described by the product of the cooling of two initial step-shaped temperature profiles as we discussed them in eq. 3.86:

$$T = T_b + (T_i - T_b) \operatorname{erf} \left( \frac{z}{\sqrt{4\kappa t}} \right) \operatorname{erf} \left( \frac{y}{\sqrt{4\kappa t}} \right) . \quad (3.93)$$

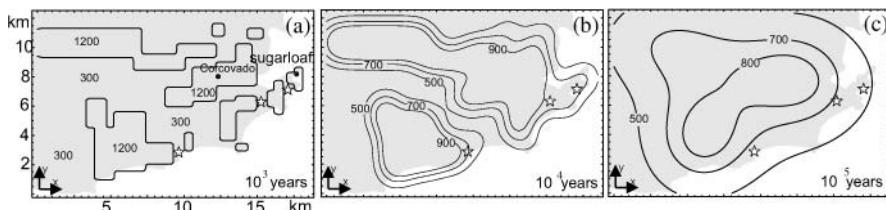
The diffusive equilibration of *squares* may be used to describe the cooling history of a rectangular hot region. The solution for this can be derived directly from eq. 3.88:

$$T = T_b + \frac{(T_i - T_b)}{2} \left( \operatorname{erf} \left( \frac{0.5l - z}{\sqrt{4\kappa t}} \right) + \operatorname{erf} \left( \frac{0.5l + z}{\sqrt{4\kappa t}} \right) \right) \\ \times \left( \operatorname{erf} \left( \frac{0.5l - y}{\sqrt{4\kappa t}} \right) + \operatorname{erf} \left( \frac{0.5l + y}{\sqrt{4\kappa t}} \right) \right) . \quad (3.94)$$

This solution may be used to describe the cooling history of a rectangular hot region (Fig. 3.38) and longer expansions of the same solution can be used to describe the cooling of any two dimensional shape of an instantaneously emplaced heat source (Fig. 3.39). However, it also finds a series of important applications in the fields of geomorphology (sect. 4.5) and petrology (sect. 7.2). In Chapter 4.5 we will see that weathering of granites can be described with the same solution.



**Figure 3.38.** Three dimensional illustration of the cooling of a two-dimensional intrusion (2 times 1 km in size) of 700 °C intruding 300 °C hot rocks. Calculated with eq. 3.94. Temperatures are shown 100, 4 000 and 10 000 years after initial intrusion



**Figure 3.39.** Temperatures around a series of intrusions with complicated shapes  
Modeled with a series of equations like eq. 3.94 after Kühn et al. (2004)

### 3.6.4 Modeling Realistic Intrusions

A series of observations in contact metamorphic aureoles of intrusions show that these are much wider and of a higher temperature than those we have predicted in the last section even as in the two dimensional examples shown. There are two important reasons for this:

- All problems we have discussed so far have been “instantaneous cooling” problems. This means, we have assumed that the cooling history commences at the time of intrusion. This need not be the case. In fact, in dikes through which magma is fed into a pluton, this would be highly unlikely. We need to describe some contact aureoles with a model where the temperature is fixed at the intrusion contact.
- So far we have neglected the latent heat of fusion as part of the cooling history. This latent heat of fusion amounts to about 320 kJ per kg of rock or roughly  $8.64 \cdot 10^8 \text{ J m}^{-3}$ . During crystallization of intrusions this heat is added to the thermal energy budget available to cause temperature change and causes buffering of the cooling history.

In the following we discuss some simple model tools to describe both processes.

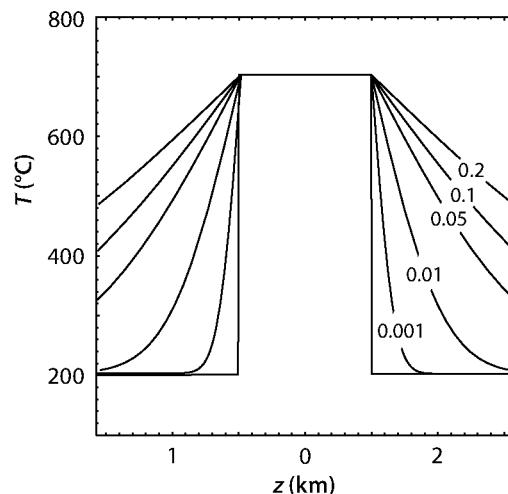
**Fixed Boundary Conditions.** Intrusion near the surface, where the temperature is constant, or intrusions through which magma flows to keep them at constant temperature must be modeled assuming a fixed temperature at the model boundary. If only one of the boundaries is fixed in temperature, then such problems may be described as a *half space problem* (sect. 3.6.1), as we have done when describing the evolution of oceanic lithosphere (see also Fig. 3.40). However, if both boundaries of a one-dimensional problem are fixed in temperature, then none of the solutions we discussed so far can be used. Under fixed boundaries the integration of eq. 3.6 is quite difficult and the solutions are not as simple as those we have discussed in sect. 3.6.2 and 3.6.3. Eq. 3.6 may only be solved by a Fourier transform (s. sect. B.4). Such solutions contain infinite summations. In their most general form, boundary and initial conditions of diffusion problems with fixed boundaries are given by:

- Initial condition: In the region  $0 < z < l$  the temperature is given by  $T = f_1(z)$  at time  $t = 0$ .
- Boundary conditions: At times  $t > 0$  the temperature at  $z = 0$  is given by  $T = f_2(t)$  and at  $z = l$  it is given by  $T = f_3(t)$ .

The function  $f(z)$  can be used to describe the shape of a cooling body and the functions  $f_2$  and  $f_3$  can describe a large number of thermal processes at the model boundaries. The fact that  $f_2$  and  $f_3$  are functions and not fixed values is also a sort of “fixed boundary” (s. sect. B.2.4). In other words, fixed boundary conditions need not be a fixed value of temperature, but they need to be defined at all times by external parameters. Eq. 3.6 can only be solved using Fourier series and such solutions therefore contain infinite summations and trigonometric functions. However, for many special cases of the space and time dependent functions  $f_1(z)$ ,  $f_2(t)$  and  $f_3(t)$  Fourier series solutions are relatively easy to derive and can be looked up in the literature (e.g. Carslaw and Jaeger 1959).

**The Stefan Problem.** The latent heat of reaction is an important part of the heat budget of high grade metamorphic terrains when phase transitions occur (see sect. 3.2.3). As we need about 1000 Joules to heat one kg of rock by one degree ( $c_p = 1000 \text{ J kg}^{-1} \text{ K}^{-1}$ ), the latent heat of fusion is enough to heat a rock by about  $320^\circ\text{C}$  (because  $L = 320 \text{ kJ kg}^{-1}$ ). An intrusion of  $T_i = 700^\circ\text{C}$ , that intrudes host rocks of  $T_b = 200^\circ\text{C}$ , is  $\Delta T = 500^\circ\text{C}$  hotter than its surroundings. This corresponds to  $\Delta T c_p = 500000 \text{ J kg}^{-1}$  additional energy that is brought into the rock by this temperature difference. However, the total *heat content* of the intrusion (including its latent heat) that is brought into the rock is  $\Delta T c_p + L = 820000 \text{ J kg}^{-1}$ . The excess energy is therefore about 1.64 times as large as the excess temperature.

**Figure 3.40.** Temperature profiles around a dike which is kept at constant temperature by the flow of magma. The curves were calculated with eq. 3.86 (curves are labeled in my). Compare this figure with Fig. 3.25 and Fig. 3.30



This means that we have underestimated the cooling history in the previous sections substantially (s. p. 123).

If we want to describe the cooling history properly (rather than being satisfied with the ball park estimate above) we need to consider where and when this crystallization heat is added to the energy budget. To illustrate this, it is useful to imagine the processes involved in the freezing of a lake. During cooling of the air below the freezing point, at first the lakes surface will freeze. The crystallization heat that is freed during this process buffers the further freezing process. During further cooling of the surface, the subsequent thickening of the ice layer will slow down, as the frozen layer insulates the water to the outside and the latent heat of crystallization freed at the ice-water interface remains contained in the water. For this reason it is rare to find pack ice on the polar oceans that is thicker than about 2 m. In fact, it was the observation of ice on the polar oceans that lead Stefan (1891) to describe the problem that now bears his name.

For the one-dimensional case and if the magma of a cooling intrusion crystallizes at a single eutectic temperature, there is an analytical solution that can be used to describe its thermal evolution under consideration of the latent heat of fusion. This is the solution found by Stefan (1891) for the freezing of water. However, most rocks crystallize in divariant reaction over a large temperature interval between a solidus and a liquidus temperature, rather than at a single temperature. Then, numerical solutions of the heat flow equation must be used to consider the effects of latent heat (s. sect. 3.2.3). Regardless, the solution of Stefan (1891) gives important insights into the thermal processes involved. It solves the heat conduction equation under consideration of the latent heat crystallization at the rock- magma interface for the geometry described by:

–  $T = T_b$  at all  $z > 0$  and  $T = T_i$  at  $z < 0$  at time  $t = 0$

and boundary conditions at infinity. It is also assumed that the magma crystallizes at temperature  $T = T_i$ . The solution that describes the temperatures outside the crystallising interface is given by:

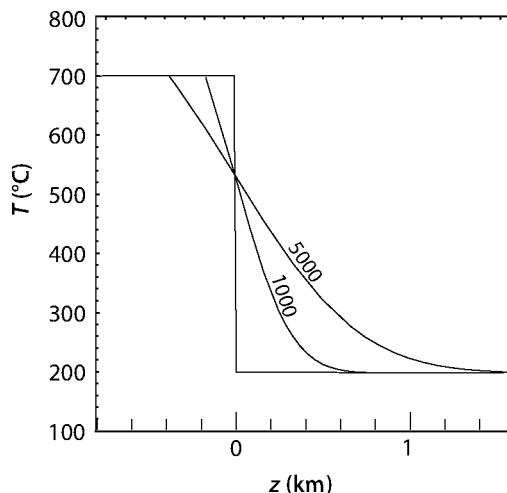
$$T = T_b + (T_i - T_b) \left( \frac{\operatorname{erfc} \left( \frac{z}{\sqrt{4\kappa t}} \right)}{1 + \operatorname{erf}(\lambda)} \right) . \quad (3.95)$$

In this equation,  $\lambda$  is given by the function:

$$\frac{L\sqrt{\pi}}{c_p(T_i - T_b)} = \frac{e^{-\lambda^2}}{\lambda(1 + \operatorname{erf}(\lambda))} \quad (3.96)$$

The derivation of this equation is complicated and of no relevance here. Eq. 3.96 may be solved iteratively or its solutions can be looked up in tables. Values for  $\lambda$  can then be inserted into eq. 3.95. This equation describes the temperature profile around a crystallizing dike margin. Fig. 3.41 shows

**Figure 3.41.** Temperatures in the contact aureole of a crystallizing interface between magma and host rock after 1 000 and 5 000 years, as calculated with eq. 3.95. Note that the contact metamorphic aureole is wider and of higher temperature than when only heat conduction is considered (Fig. 3.32). All parameters assumed here correspond to those used for Fig. 3.32 with the only difference being that latent heat is considered here



some curves calculated with eq. 3.95. The thickness of the dike is no part of this solution. The position  $z = 0$  is the location of the dike margin at the onset of crystallization. The solution may be used up to the time where the crystallization interfaces of the opposite walls of the dike meet at its center. Thereafter numerical solutions must be used to describe the subsequent cooling history.

**Heat Content of Intrusions and Metamorphic Terrains.** If we recall eq. 3.3, there is a simple relationship between heat and temperature. The relating proportionality constants are the heat capacity  $c_p$  and density  $\rho$ . In fact, all problems we have described using temperature, could have also been formulated in terms of heat using:

$$\frac{\partial H}{\partial t} = \frac{\partial T}{\partial t} \rho c_p . \quad (3.97)$$

Instead of going through many calculated examples here, we only want to state that the heat content of an intrusion or a metamorphic terrain can be viewed – in one dimension – as the area underneath a  $T - z$  curve (Fig. 3.42). In essence, this is what eq. 3.3 and eq. 3.97 state (some more detail is given in the first edition of this book). In other words, the heat content of a dike of thickness  $l$  and temperature  $T_i$  relative to its surroundings of temperature  $T_b$  is given in  $\text{J m}^{-2}$  by:  $(T_i - T_b)l\rho c_p$ . Using typical values of  $\rho = 2700 \text{ kg m}^{-3}$  and  $c_p = 1000 \text{ J K}^{-1} \text{ kg}^{-1}$ , the intrusion of Fig. 3.32 has a heat content of  $H = 1.89 \cdot 10^{12} \text{ J m}^{-2}$ . This is the heat content per square meter of dike surface.

In high grade metamorphic terrains containing syn-metamorphic intrusives it is often discussed if the volume of the intrusions is sufficient to contribute significantly to the metamorphism. In other words, it is discussed if the metamorphism is *contact metamorphism* in the widest sense (e.g. Problem 3.14).

In order to do this estimate properly, the heat content of the intrusives must be compared with the heat content of the terrain. If the specific heat capacities of the intrusives and the metamorphic host rocks are the same, then the comparison of energy contents can also be made as a comparison of the temperatures. However, we must consider the latent heat of fusion that is part of the heat of the intrusives, but not of the host rocks. As simple calculation shows that:

$$T_{\max} = T_b + \left( T_i - T_b + \frac{L}{c_p} \right) \frac{V_{\text{intrusives}}}{V_{\text{terrain}}} . \quad (3.98)$$

There,  $T_b$  and  $T_i$  are the temperature of the host rock before metamorphism and that of the intrusion, respectively.  $V_{\text{intrusives}}$  is the volume of the intrusives and  $V_{\text{terrain}}$  is the volume of the entire metamorphic terrain.  $T_{\max}$  is the maximum temperature that can be reached by contact metamorphism.

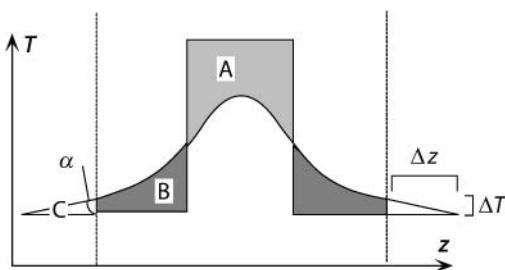
If the *aerial* proportion of intrusives to host rock are representative for the *volumetric* proportion of intrusives in the terrain, then the volumes of eq. 3.98 may be replaced by areas. Using  $T_i = 700^\circ\text{C}$ ,  $T_b = 300^\circ\text{C}$ ,  $L = 320\,000 \text{ J kg}^{-1}$  and  $c_p = 1\,000 \text{ J kg}^{-1} \text{ K}^{-1}$ , eq. 3.98 shows the following: only about 55 % of the terrain must consist of syn-metamorphic granites in order to heat the entire terrain to  $700^\circ\text{C}$ , even if the intrusion temperature itself was only  $700^\circ\text{C}$ . If the intrusives are  $1\,200^\circ\text{C}$  hot mafic magmas, then only 30 % of the terrain must be intrusives in order to heat the terrain to  $700^\circ\text{C}$ .

- *Using heat content as a boundary condition.* In numerical calculations of thermal evolutions on grids of finite extent it is often not possible to assume boundary conditions at infinity. Then, a consideration of the heat content may come in handy as a boundary condition. For example, the boundary conditions for eq. 3.88 were assumed at infinity there, but may also be formulated in terms of the heat content of the intrusion. As no heat can be lost from the system if it cools by conduction only, the area underneath the curves in Fig. 3.32 must stay constant between  $z = +\infty$  and  $z = -\infty$ . This is schematically shown on Fig. 3.42. During cooling of the intrusion, the area  $A$  must always be as large as the sum of  $2B + 2C$ . When solving the heat flow equation numerically on grids of finite extent, for example between the two straight lines on Fig. 3.42, problems are often easier to handle this way than they are if the temperature is assumed to be constant at infinity. In this case the heat flow at the model boundary (given by the angle  $\alpha$ ) could be adjusted so that the areas  $C$  have the correct size. From Fig. 3.42 it may be seen that  $A = 2B + 2C$  and  $\tan(\alpha) = \Delta z / \Delta T$ . From this, the thermal gradient at the model boundary is given by:

$$\alpha = \operatorname{tg}^{-1} ((A - 2B) / \Delta T^2) . \quad (3.99)$$

This relationship is a useful in many numerical descriptions of diffusion problems with boundary conditions at infinity.

**Figure 3.42.** Schematic illustration of the heat content of a dike. The sum of the areas of the two dark-shaded regions  $B$ , plus the white triangles  $C$  must be the same as the light-shaded area  $A$ . Then, the heat content of the total system remains constant



## 3.7 Selected Heat Transfer Problems

In this section we discuss a selection of geologically important heat transfer problems that do not belong directly to any of the past sections but are important to understand geodynamic processes. We discuss mainly problems which are subject to spatially or temporally changing temperatures. Such problems can be described with periodic boundary conditions and have the great advantage that analytical solutions of the heat conduction equation can be found for them. In particular, we have selected three geodynamically relevant sets of problems, which may be described with periodic boundary conditions as well as the temperature distribution around faults. Analytical solutions for these and many others interesting heat transfer problems can be found in the literature (e. g., Carslaw and Jaeger 1959)

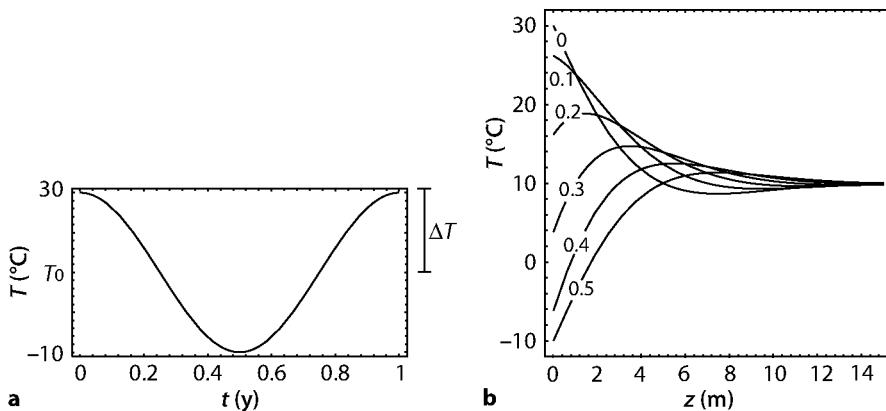
### 3.7.1 Periodic Temperature Fluctuations

The temperatures at the surface of earth are subject to the daily or annually periodically changing temperatures of the atmosphere. Problems where this is relevant, range from understanding the thickness of permafrost soils, to the regulation of temperatures in tunnels and insulation of walls of buildings. Many of these problems can be described with a one-dimensional coordinate system with  $z$  as the coordinate normal to the surface and with boundary conditions that describe a periodic fluctuation of the temperature at the surface. This may be written as:

- Initial condition:  $T = T_0$  at all  $z$  at time  $t = 0$ .
- Boundary condition:  $T = T_0 + \Delta T \cos(ft)$  at  $z = 0$  for all  $t > 0$ , and  $T = T_0$  at  $z = \infty$  for all  $t > 0$ .

There,  $\Delta T$  is half the amplitude of the annual fluctuation,  $t$  is time and  $f$  is the frequency of the periodic temperature cycle (Fig. 3.43 a).  $T_0$  is the mean temperature over one cycle. The time dependent diffusion equation (eq. 3.6) can be integrated using these assumptions. We will not go through this integration here, but the result is amazingly simple. It is given by:

$$T = T_0 + \Delta T e^{\left(-z\sqrt{\frac{f}{2\kappa}}\right)} \cos\left(ft - z\sqrt{\frac{f}{2\kappa}}\right) . \quad (3.100)$$



**Figure 3.43.** The temperature in the upper few meters of the crust as a function of annually changing surface temperature. **a** shows the upper boundary condition: the temperature  $T$  as a function of time  $t$  at  $z = 0$ . The annual mean temperature was assumed to be  $T_0 = 10^\circ\text{C}$ . The variation about this mean was assumed to be  $\Delta T = 20^\circ\text{C}$ . **b** shows six thermal profiles in the ground at six different times through the course of half a year. The curves are labeled for time in years. It may be seen that below depths of about 2 m temperatures never sink below the freezing point

This equation may be used to describe temperature fluctuations at depth as a function of a periodic temperature variation at the surface. It may be seen that this equation contains a trigonometric function and an exponential function. At each time  $t$  it describes a cosine function of temperature which decays in amplitude exponentially with depth. At the depth  $z = (f/2\kappa)^{-0.5}$  the amplitude of the temperature oscillation is that of the surface divided by  $e$ . This depth is often called the *characteristic depth of equilibration* or *skin depth* (s. sect. 3.4.3). Eq. 3.100 is extremely important for many near surface problems, for example temperature profiles in snow, air temperatures in caves and many more (Fig. 3.43b).

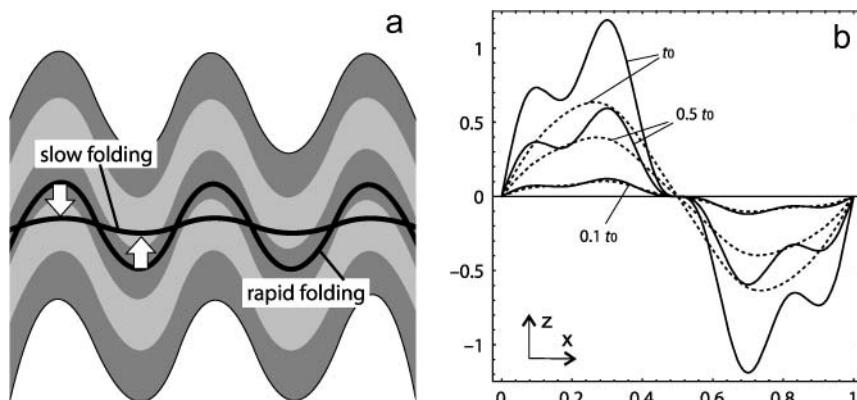
### 3.7.2 Folded Isotherms

A beautiful problem illustrating the relationships between length and time scales of heat conduction and advection (s. sect. 3.1.4, 3.3.4) is given the shape of isotherms during folding. Before we discuss this problem it should be said that – strictly speaking – isotherms can not be folded, as they are not material lines. However, a number of model descriptions use the term “folding” even when talking of isotherms and we use it here as well. Fig. 3.44a shows schematically how isotherms may be folded during deformation of rocks. This process may occur if the axial plane of the folds is not parallel to the isotherms. Whether or not folding of isotherms has a thermal influence on rocks depends on the relationship between the wavelength and amplitude of the fold, on the folding rate and of course on the diffusivity  $\kappa$ . If

the wavelength, the amplitude and the folding rate are large, then isotherms will be folded together with the stratigraphy. If the folding rate is small, then the rate of thermal equilibration is rapid compared with the folding and the isotherms will remain flat. Field observations like the updoming of the Tauern Window in the eastern Alps, together with the updoming of the Alpine metamorphic isograds show that it may be important to estimate the magnitude of this effect. If the estimates show that folding of rocks is rapid enough to fold the isotherms as well, then we should observe that antiforms cool and synforms heat. With the methods of modern petrology it is conceivable that such details of the thermal evolution can be tested (sect. 7.2).

A first estimate of the potential magnitude of an isotherm folding process can be made using the Pecllet number (sect. 3.3.4). For a geologically reasonable shortening rate of  $1 \text{ cm y}^{-1}$  an amplitude of a single antiform of 5 km forms in about 0.5 my. The rate  $u$ , with which the material lines are deformed is therefore  $10^4 \text{ m my}^{-1}$ . Using  $l$  as the amplitude and  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$  we get from eq. 3.51, that  $Pe \approx 1.5$ . This shows that diffusion and advection processes are both relevant for the assumed parameters and that a more detailed investigation of the problem is justified.

A quantitative and very simple model that can be used for the description of this process was developed by Sleep (1979). This model is based on the following assumptions:



**Figure 3.44.** Illustration of folded isotherms. **a** Schematic illustration. The light and dark shaded regions are folded strata, the thick lines are isotherms. During *slow* folding, equilibration is faster than folding and isotherms will not be folded with the rock. During *rapid* folding of the rock conductive equilibration can not keep up with the material advection and the isotherms will be folded with the rock. **b** Folding of stratigraphy and isotherms as calculated with the model of Sleep (1979). The continuous lines show the shape of a layer at three different points in time. The folding is characterized by a principal fold with wave length and amplitude of 1, overprinted by a parasitic fold with wavelength and amplitude 0.25. For this model two sets of eq. 3.101 and eq. 3.102 must be summed up. For details see Sleep (1979) or the first edition of this book

1. Before folding, isotherms are parallel to each other and to the surface of earth. The isotherms have a constant distance to each other so that the initial temperature distribution as a function of depth is:  $T_{(t=0)} = gz$ , where  $g$  is the geothermal gradient and  $z$  is depth.

2. Folding of rocks occurs by shear folding along axial planes that are perpendicular to the isotherms and a sine-shaped fold is produced that has the form:

$$v = v_0 A \sin\left(\frac{2\pi x}{\lambda}\right) \quad \text{or:} \quad z = t v_0 A \sin\left(\frac{2\pi x}{\lambda}\right) . \quad (3.101)$$

In this equation,  $v$  is the displacement rate at each point of the fold,  $v_0$  is the displacement rate of the fold hinge and  $x$  is the horizontal coordinate.  $z$  is the depth of the folded rock at time  $t$ .  $A$  is the amplitude of the fold. This assumption describes shear folding at constant volume and does not consider shortening perpendicular to the axial plane of the folds.  $\lambda$  is the wavelength of folding. Using these assumptions, eq. 3.42 may be integrated. According to Sleep (1979) this gives:

$$T = g \left( z - z_0 \left( 1 - e^{-t/t_0} \right) \right) \sin\left(\frac{2\pi x}{\lambda}\right) . \quad (3.102)$$

There, the constants  $z_0$  and  $t_0$  are:

$$z_0 = \frac{v_0 \lambda^2}{4\pi^2 \kappa} \quad \text{and} \quad t_0 = \frac{\lambda^2}{4\pi^2 \kappa} . \quad (3.103)$$

This equation may – of course – also be reformulated to solve for the depth  $z$  of a chosen isotherm  $T$ . Folding events that last shorter than  $t_0$  will cause a significant deformation of the isotherms. If they last longer, then the additional deformation is negligible. For most field examples of folding, the geometry assumed for the folding process in this model is too simple. However, Sleep (1979) notes that *all* fold geometries that have formed by shear folding can be represented by a summation of sine-functions. Thus, eq. 3.102 could be applied to all sine-shaped components of a random folding geometry and the results could then be summed up. Fig. 3.44b shows an example in which the amplitude and wavelength of two folds, a principal fold and a parasitic fold is considered. There, the amplitude of the overprinting parasitic fold is 1/4 of that of the principal fold.

### 3.7.3 Isotherms and Surface Topography

An important example of a heat conduction - advection problem concerns the influence of the surface topography on isotherms at depth. Rocks inside mountains are thermally insulated, while rocks nearer the surface of an incising valley are cooled by the surface. As a consequence, isotherms follow the earth's surface it in a damped form. Just how the distance of a given

isotherm from the surface varies with amplitude and wavelength of the topography is of large importance for the interpretation of low temperature geochronological data in mountainous regions, for the design of ventilation systems in tunnels and more (Fig. 3.45) (Braun 2002; 2006). In this section we discuss some models that can be used to estimate the magnitude of this effect.

**The Upper Boundary: The Topography.** For simplicity we assume that the surface topography may be described by a sine-function with the wavelength  $\lambda$  and the amplitude  $h_0/2$  where we interpret the wavelength  $\lambda$  as the distance between two parallel valleys and  $h_0$  as the maximum elevation of the peaks above the valleys. Using the coordinate system illustrated in Fig. 3.46a,b, the elevation  $h$  at any point of the topography is described by:

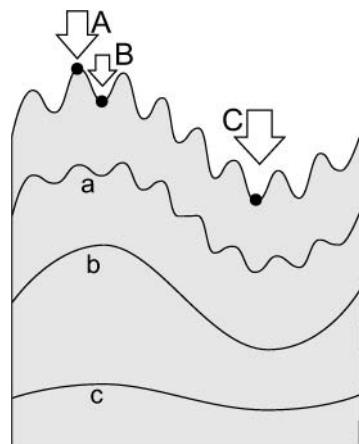
$$z_{(T=0)} = -h = -h_0 \frac{1}{2} \left( 1 + \cos \left( \frac{2\pi x}{\lambda} \right) \right) \quad (3.104)$$

Using such a simple function to describe topography allows us to evaluate the magnitude of the thermal effects as a function of two simple parameters of topography: the wavelength and the amplitude.

We also assume that the temperature along the curve described by eq. 3.104 is the surface temperature  $T_s = 0$  and neglect any atmospheric temperature gradient or seasonal variation. However, even for these simplified assumptions, integration of the two-dimensional diffusion- or diffusion-advection equation (eq. 3.6 or 3.42) under this boundary condition is very difficult (s. sect. 3.6.3). A common way to surround this problem is by substituting this boundary condition of *constant* temperature at a *variable* spatial position, by a boundary condition of *variable* temperature at a *constant* elevation, for example at  $z = 0$ . This would be described by:

$$T(x)_{(z=0)} = \Delta T \frac{1}{2} \left( 1 + \cos \left( \frac{2\pi x}{\lambda} \right) \right) \quad (3.105)$$

**Figure 3.45.** Schematic illustration of isotherms underneath topography. (a), (b) and (c) show three different isotherms at depth. Note that the topographic perturbation of isotherms decreases with depth in proportion to the wavelength of the topography. Using a geochronological system that closes at isotherm (b), a comparison of samples from A and B (vertical age profiling) could be used to determine the erosion rate, while a comparison of samples from A and C may be used to study the landform evolution. Using age elevation relationships at A and B with an isotopic system that closes at isotherm (a) would result in an overestimate of the erosion rate (Braun 2006; Stüwe et al. 1994).

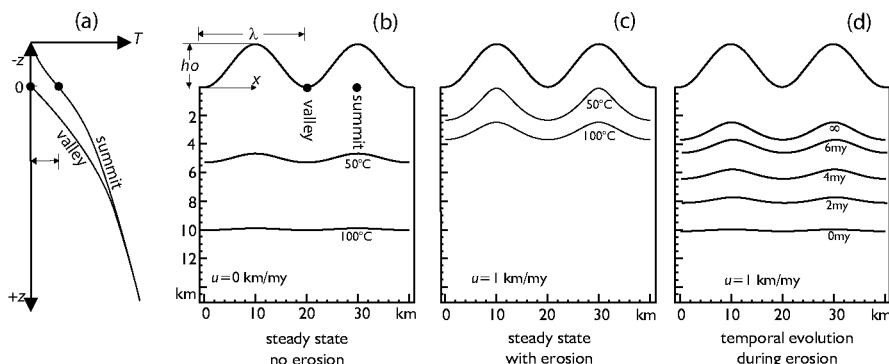


where  $\Delta T = h_0 g$  and  $g$  is the geothermal gradient in the absence of topography. This assumption implies that the thermal gradient inside the mountains is linear.  $\Delta T$  corresponds to  $h_0$  in the proper formulation (s. Fig. 3.46). This approximation is good if the wavelength of the topography is large compared to the amplitude because then the lateral cooling through the sides of the valley can be neglected. Note that this approximation of the topography is actually quite a correct description if the surface temperature on a flat shield *does* vary laterally, for example because of the presence of lakes. Using these assumptions, we can discern three different types of scenarios described by different solutions of the diffusion – advection equation:

**Topography Without Erosion.** If there is no erosion, there is no advection towards the surface. As a consequence, the topographic perturbation effect on isotherms is substantially smaller than it would be in an eroding terrain. To estimate it we need to solve the two dimensional diffusion equation (eq. 3.9) subject to the boundary conditions of eq. 3.105. The solution is:

$$T_{(x,z)} = T(x)_{(z=0)} \times e^{-2\pi z/\lambda} . \quad (3.106)$$

We can see from this solution a fundamental result (shown in Figs. 3.46b): The perturbation of isotherms decreases exponentially with depth and in proportion to the wavelength of the topography. In other words: at a depth  $\lambda/2\pi$  the amplitude of the isotherm is a mere  $e^{-1} \approx 0.3$  of the amplitude of the surface topography. Lets use the Alps and an isotherm relevant for

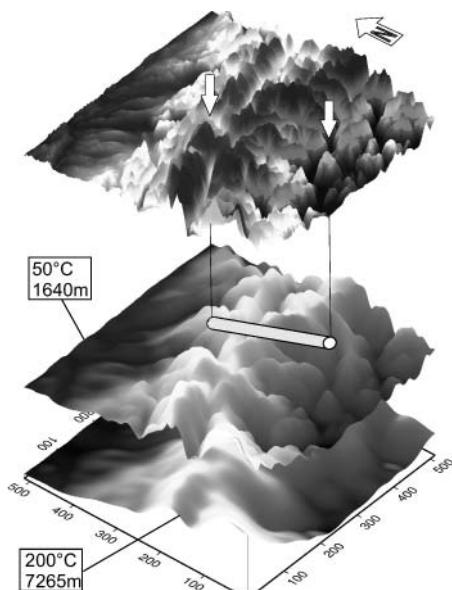


**Figure 3.46.** Isotherms underneath regions of high topographic relief. **b, c** and **d** show two-dimensional profiles through a mountainous topography. The two isotherms in **b** are for a non-eroding topography calculated with eq. 3.105 and adding a linear thermal gradient to the solution to place the different isotherms at the correct depths. **c** shows a thermal steady state case for an eroding topography. In **d** the time dependent evolution underneath an eroding topography is shown for a single isotherm, in this case  $100^\circ\text{C}$ . The two black dots in **a** and **b** are at equal depth, but they have a different temperature. Note that in the chosen coordinate system the surface is located at negative values for  $z$ , but (according to eq. 3.104) at positive values for  $h$  (s. Fig. 4.3)

fission track analysis ( $\approx 100^\circ\text{C}$  corresponding to a depth of  $\approx 3$  km) as an example: The Alps have a topographic amplitude of about 3 km and a series of topographic wavelengths from about  $\lambda \approx 2$  km in the most rugged regions to  $\lambda \approx 200$  km from the northern to the southern Molasse basin. Inserting these numbers into the exponential term of eq. 3.106 shows that the narrow wavelengths are just about invisible for the  $100^\circ\text{C}$  isotherm, while 90% of the longest wavelength is preserved at 3 km depth. This caused Brown (1991) to conclude the interpretation of age – elevation profiles from apatite fission track data may be done without topographic correction.

**Eroding Topography.** During erosion material is advected towards the surface, isotherms are compressed into the topography and the amplitude of a given isotherm is substantially larger than when no erosion occurs. To estimate the perturbation effect quantitatively we need to solve a two dimensional form of the diffusion – advection equation (eq. 3.42). Eqs. 3.45, 3.46 and 3.47 are three examples for geotherms during erosion but without consideration of topography. These equations may be expanded into two dimensions and modified to account for the variable boundary at the top surface. Solutions typically show the same fundamental relationship discussed in connection with eq. 3.106, namely that the effect decreases exponentially with depth and in proportion to the wavelength of the topography. However, depending on the erosion rate, the advection processes may be strong enough to perturb isotherms of geological relevance. Stüwe et al. (1994) found a semi-analytical solution of the two-dimensional diffusion-advection equation to describe this problem (Fig. 3.46c and 3.45). They concluded that at erosion rates above

**Figure 3.47.** Example of a three dimensional conduction-advection model to consider topographic perturbation effects on topography (Hergarten and Stüwe in prep.). For this example, the digital elevation model of the Gotthard region was interpolated onto a three dimensional grid to consider thermal effects important for ventilation in the longest road tunnel on earth: the Gotthard base tunnel. Below the topography, two isotherms are shown (labeled for temperature and depth). Assumed erosion rate was 1 mm per year. North and south entrance of the Gotthard tunnel are shown by the white arrows



500 m  $\text{my}^{-1}$  it becomes important to consider the topographic effects on the interpretation of apatite fission track results.

**Asymmetric Advection, Three-dimensionality and Time Dependence.** Several other processes not considered in the discussion above may be of interest when estimating topographic corrections on isotherms at depth. These include radiogenic heat production, time dependent effects, atmospheric temperature gradients, convective heat transfer by fluids or seasonal temperature variation and asymmetric advection due to variable erosion rate. To account for such processes properly, it is usually necessary to use numerical solutions of the diffusion – advection equation. Then, all these processes can be considered in a much more straight forward manner. A nice examples of the use of a three dimensional model is presented by Craw et al. (2005) and another one is shown in Fig. 3.47. Several three dimensional models are currently in use to describe these effects. Mancktelow and Grasemann (1997) investigated the time dependent effects how this steady state is reached (Fig. 3.46d). However, as the Peclet number of most topographies is small, it is usually not necessary to investigate the time dependent evolution of isotherms. Asymmetric erosion due to differential rain fall on different sides of a mountain range was considered by Stüwe and Hintermüller (2000) and also discussed in the first edition of this book. With the advent of a series of new thermochronological methods (e.g. House et al. 1998; 2000) enough detail of low temperature cooling curves is starting to be resolved to justify careful modeling of all these processes.

### 3.7.4 Temperature Distribution Around Faults

The thermal evolution of rocks around faults is a typical two-dimensional problem in which all three heat transfer mechanisms may play a role: *diffusion* in the foot and hanging wall of the fault, *advection* of heat by the relative motion of the two sides of the fault and *production* of friction heat in the fault itself (s. Fowler and Nisbet 1982). In the following analysis we only discuss diffusion and advection. For the importance of mechanical heat production, the interested reader is referred to the now classic discussions by Molnar and England (1990a), Graham and England (1976) and Pavlis (1986), sect. 3.2.2 and a wealth of modern literature on the subject.

Fig. 3.48 shows a vertical cross section through a package of rock that is transected by a fault (in the shown orientation it is a *normal* fault). The normal fault is inclined with  $60^\circ$  to the left and the initial isotherms are inclined with  $20^\circ$  to the right. The displacement along the fault (10 km in this case) opposes the hot rocks in the foot wall with the relatively cold rocks in the hanging wall. As a consequence, the isotherms are stretched near the fault. The lateral temperature gradient in the vicinity of the fault is *decreased* by the fault displacement (Fowler and Nisbet 1982). If the fault were a reverse fault of the same angle, then the situation would be reversed: the foot wall would be cooled, the hanging wall would be heated and the

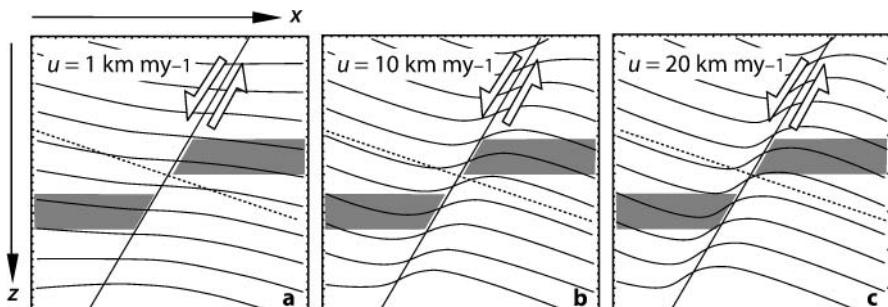
isotherms would be *compressed*. The heating of rocks by reverse faulting is one of several processes that have been made responsible for the formation of inverted metamorphic gradients (s. discussion by England and Molnar 1993).

- *Estimating the thermal influence.* In order to estimate the thermal influence of a fault zone on its surroundings, we can employ two simple tools which we have discussed in previous sections: The thermal time constant (eq. 3.17) and the Peclet number (eq. 3.51). In Fig. 3.48a,b and c, the rate of material advection  $u$  (displacement rate) is 1, 10 and 20 km my<sup>-1</sup>, respectively. From eq. 3.51 we can estimate the length scale  $l$  where both diffusion and advection influence the thermal structure (i.e.  $Pe = 1$ ) by (Fig. 3.16):

$$l = \frac{\kappa}{u} . \quad (3.107)$$

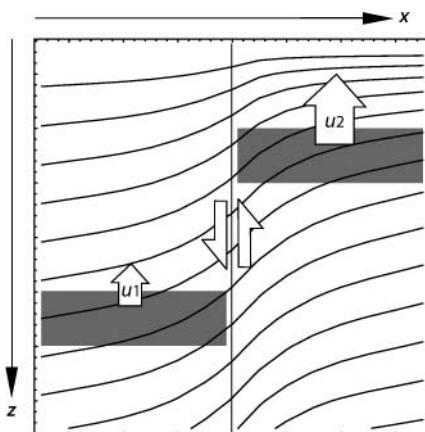
With  $\kappa = 10^{-6}$  m<sup>2</sup> s<sup>-1</sup>, this gives  $l = 30, 3$  and  $1.5$  km, respectively. A comparison with Figs. 3.48a,b and c shows that the region over which isotherms are curved corresponds indeed to these length scales. This suggests that a more detailed description with sophisticated numerical models is not necessary to understand the first order effects.

- *Temperature-time evolution during exhumation in the vicinity of faults.* In many regions of active and ancient mountain belts it is observed that displacement along fundamental structures leads to different rates of exhumation of the foot wall and the hanging wall. This may lead to a complicated thermal evolution of rocks as there are two competing thermal processes:



**Figure 3.48.** Isotherms in a crustal cross section around a normal fault that is inclined with 60 degrees to the left. The shown area is 50 km by 50 km in size. The figures are labeled for the rate of normal displacement  $u$ . a, b and c show the temperatures after 10 my, 1 my and 0.5 my. Thus, the displacement is the same in all three figures. The gray shaded region indicates schematically a lithological layer. The isotherms are shown every 50 °C. The isotherms were inclined with 20 degrees to the right before the fault became active (dashed line). The boundary conditions on all four sides are given by constant heat flow. This implies that the crustal section shown is surrounded on all sides by matter and that the diagram can be rotated arbitrarily. The diagram was calculated with a numerical solution of a two-dimensional form of eq. 3.41. For an analytical solution of related problems see e. g. Voorhoeve and Houseman (1988)

**Figure 3.49.** Isotherms in a vertical cross section through the crust with a vertical fault. Both sides of the fault exhume with different rates  $u_1$  and  $u_2$ . The shown section is 50 km by 50 km. The boundary condition at the top boundary is given by a constant temperature. The boundary conditions on all other sides are given by constant flow of heat. In contrast to Fig. 3.48 all isotherms are preserved, because of the fixed top boundary



1. The heating and cooling of the two respective sides of the fault.
2. The cooling of rocks with proximity to the surface.

Fig. 3.49 illustrates this with a vertical cross section through the crust. This figure differs from Fig. 3.48 in that the upper boundary (the earth's surface) is now defined by a boundary condition of constant temperature. Thus, in contrast to Fig. 3.48, no isotherms intersect the surface. The crustal section shown is divided into two blocks separated by a vertical fault. The two blocks exhume with two different rates,  $u_1$  and  $u_2$ .

In order to understand the thermal evolution of rocks near the fault, consider rocks in the slower exhuming block (left block on Fig. 3.49) near the fault. There, the *cooling* influence of the surface is opposed by the *heating* influence of the block on the right hand side. Because of this, it may happen, that rocks that cool during exhumation experience a late stage heating event caused by the other side of the fault. An example of such a thermal evolution was documented by Grasemann and Mancktelow (1993) at the Simplon-line in the central Alps.

## 3.8 Problems

### Problem 3.1. Converting different units of energy (p. 51):

In nuclear reactions, mass is converted to energy. How long can a 60-W-light globe be lit with 1 g of mass assuming that this mass may be converted completely into energy? Remember: energy = mass  $\times$  speed of light<sup>2</sup>. The speed of light is  $\approx 300\,000\text{ km s}^{-1}$ . A related problem is found in Problem 5.3.

### Problem 3.2. Converting different units of energy (p. 53):

What is the conversion factor between hfu (1 hfu =  $10^{-6}\text{ cal cm}^{-2}\text{ s}^{-1}$ ) and the SI-units  $\text{W m}^{-2}$ .

**Problem 3.3.** *Understanding heat refraction qualitatively (p. 56):*

A spherical iron ore body of 10 km diameter (with extremely high thermal conductivity) lies in the middle of a 30 km thick crust. Draw a cross section through the crust and sketch schematically some isotherms. Discuss the thermal consequences that such a geometry may have for the surrounding rocks.

**Problem 3.4.** *Understanding heat refraction quantitatively (p. 56):*

Calculate the temperature at 10 km depth in the crust using the following assumptions. The crust is made up of schists with a conductivity of  $k = 2 \text{ Js}^{-1} \text{ m}^{-1} \text{ K}^{-1}$ . These schists are intruded by a 2 km thick sill that has a conductivity of  $k = 4 \text{ Js}^{-1} \text{ m}^{-1} \text{ K}^{-1}$ . The upper contact of the sill is in 5 km depth. The thermal gradient at the surface is  $20 \text{ }^{\circ}\text{C km}^{-1}$  and the temperature at the surface is  $0 \text{ }^{\circ}\text{C}$ . The temperature profile is in steady state (there is no exchange of heat between the layers) and there is no heat production in the crust (i.e. the different sections of the geotherm are linear). You can use eq. 3.7.

**Problem 3.5.** *Using thermal time constant (p. 62):*

Two continental plates collide and deform a 35 km thick crust. The deformation rate  $\dot{\epsilon}$  is a)  $10^{-12} \text{ s}^{-1}$  and b)  $10^{-16} \text{ s}^{-1}$ . During and after deformation, the thermal structure of the crust re-equilibrates by diffusion. For the given deformation rates, estimate whether the equilibration of the crust is faster, slower or of similar rate as the deformation? The time scale of the thermal equilibration may be estimated with the thermal time constant (eq. 3.17). What consequences may your results of a) and b) have for structures and parageneses that may be observed in a thin section?

**Problem 3.6.** *The importance of radiogenic heat production (p. 66):*

A roughly spherical radioactive ore body of 10 km diameter produces about  $S_{\text{rad}} \approx 100$  micro Watts of heat per cubic meter. Estimate how hot the center of the body will get after about  $t = 10^5$  years. Use eq. 3.21 but before you do, make an argument why it may be realistic to neglect heat conduction. Use eq. 3.18 and estimates for the necessary parameter values as discussed in the text, for example  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ ,  $c_p = 1000 \text{ J kg}^{-1} \text{ K}^{-1}$  and  $\rho = 2700 \text{ kg m}^{-3}$ .

**Problem 3.7.** *The importance of mechanical heat production (p. 68):*

a) What is the mechanical heat production rate of a 5 km wide shear zone which deforms at a rate of i)  $\dot{\epsilon} = 10^{-13} \text{ s}^{-1}$  or: ii)  $\dot{\epsilon} = 10^{-15} \text{ s}^{-1}$ ? Assume that the deviatoric stress that the shear zone material supports is between 30–300 MPa. Give minimum and maximum values using both strain rates and both shear strengths using eq. 3.25. b) How warm could the shear zone possibly get if deformation lasts for 1 my? ( $c_p = 1000 \text{ J kg}^{-1} \text{ K}^{-1}$ ;  $\rho = 2700 \text{ kg m}^{-3}$ ). c) Which parameters control if (and how much) the shear zone heats up during this process? (Another problem related to mechanical heat production is Problem 5.3)

**Problem 3.8.** *Understanding heat of reaction (p. 76):*

How much mass is converted to energy when burning 5 kg of wood?

**Problem 3.9.** *The importance of latent heat of fusion (p. 78):*

A high grade metamorphic rock contains 30 % partial melts. All of this partial melt was formed by a single melting reaction at a fixed temperature. The rock cools conductively from its metamorphic peak (which was much higher than this melting temperature) with a constant cooling rate of  $100\text{ }^{\circ}\text{C my}^{-1}$ . Estimate how long the rock will be buffered to a constant temperature when this melt crystallizes? Give the result in my. The latent heat of fusion is  $320\,000\text{ J kg}^{-1}$ , the rock has a density of  $\rho = 2\,700\text{ kg m}^{-3}$  and a specific heat capacity of  $c_p = 1\,000\text{ J kg}^{-1}\text{ K}^{-1}$ . Discuss the potential effect of your result on geothermometry that may be planned for this rock.

**Problem 3.10.** *Understanding the Peclet number (p. 86):*

A regional metamorphic event occurred over the whole of a 30 km thick crust. A mountain belt at the surface eroded at the same time and exhumed the metamorphic rocks during this process. Estimate whether the regional thermal evolution can be described by only considering heat conduction (that caused the regional metamorphic event) or if heat advection (due to erosion) must also be considered. Use erosion rates of a)  $100\text{ m my}^{-1}$ ; b)  $1\,000\text{ m my}^{-1}$ ; c)  $5\,000\text{ m my}^{-1}$  and eq. 3.51.

**Problem 3.11.** *Contribution of radioactivity to heat flow (p. 92):*

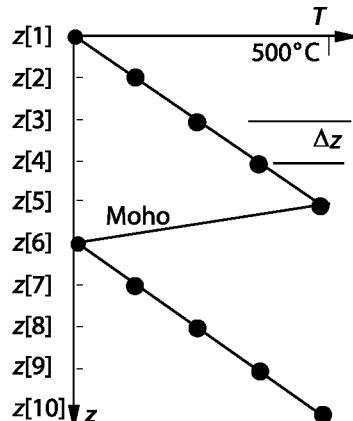
Use eq. 3.61 to estimate how the surface heat flow changes if the radiogenically caused heat flow doubles and the mantle heat flow is decreased by 50%. Note that *both* will occur when the lithosphere is doubled in thickness.

**Problem 3.12.** *Cooling of oceanic lithosphere (p. 101):*

The temperature distribution in oceanic lithosphere may be described with the *half space cooling model* (eq. 3.80). a) Calculate the depth of the  $1\,000\text{ }^{\circ}\text{C}$ -isotherm for an 80 my old oceanic lithosphere. Assume the temperature of the asthenosphere is  $T_l = 1\,200\text{ }^{\circ}\text{C}$  and the temperature at the surface is  $T_s = 0\text{ }^{\circ}\text{C}$ . The diffusivity  $\kappa$  is  $10^{-6}\text{ m}^2\text{ s}^{-1}$ . Use eq. 3.80 and Fig. 3.6 or the approximation in Table C.9 to solve the error function. b) Draw a temperature profile through 10 my old oceanic lithosphere with the same assumptions.

**Problem 3.13.** *Qualitative thermal evolution of intrusions (p. 112 – 116):*

a) Estimate the total duration of cooling of a 50 m wide dike with  $\kappa = 10^{-6}\text{ m}^2\text{ s}^{-1}$ . Use eq. 3.18. b) Assume that the dike has intruded with  $700\text{ }^{\circ}\text{C}$  into host rocks that are  $300\text{ }^{\circ}\text{C}$  hot. What is the maximum contact metamorphic temperature? c) How much additional heat does the dike bring into the rock it intrudes? Use eq. 3.3 and eq. 3.97 and  $\rho = 2\,700\text{ kg m}^{-3}$ ,  $c_p = 1\,000\text{ J kg}^{-1}\text{ K}^{-1}$ . d) Draw a qualitative temperature profile across the dike at 40 years after intrusion. Help yourself for this with the result from a).

**Figure 3.50.** Illustration for Problem 3.15**Problem 3.14.** *Inferring metamorphic heat sources (p. 123):*

Estimate if the following metamorphic terrain could have been heated by contact metamorphism. The terrain has reached a peak temperature of  $600^{\circ}\text{C}$ . About 10% of the area of the terrain are syn-metamorphic mafic intrusions that had an intrusion temperature of  $1100^{\circ}\text{C}$ . Another 30% of the area of the terrain are syn-metamorphic granitoids that had an intrusion temperature of  $700^{\circ}\text{C}$ . The peak metamorphic pressure was 5 kbar, which corresponds to a depth of 18.5 km. Before metamorphism the terrain was at the same depth and had a temperature corresponding to a stable geothermal gradient of  $16.2^{\circ}\text{C km}^{-1}$ . The density of the rocks is  $\rho = 2700 \text{ kg m}^{-3}$  and the specific heat capacity is  $c_p = 1000 \text{ J kg}^{-1} \text{ K}^{-1}$ . Answer the question graphically and algebraically with and without consideration of the latent heat of fusion of the magmas ( $L = 320000 \text{ J kg}^{-1}$ ).

**Problem 3.15.** *Inverted isograds and numerical solutions of the heat flow equation (sect. 3.1.1, B.2):*

Calculate the thermal evolution above and below a thrust that has doubled the entire crust of  $z_c = 40 \text{ km}$ . Assume that the thrusting rate was much faster than the rate of thermal equilibration. Thus, we can assume a “saw tooth” geotherm as our initial condition (Fig. 3.50). The initial geotherm is defined as follows: In the region  $0 < z < z_c$  it is given by  $T = T_{\text{Moho}}(z/z_c)$ ; in the region  $z_c < z < 2z_c$  it is given by  $T = T_{\text{Moho}}((z - z_c - \Delta z)/z_c)$ . ( $T_{\text{Moho}} = 500^{\circ}\text{C}$  and  $z_c = 40 \text{ km}$ ). The distance between discrete points for your calculation is  $\Delta z = 10 \text{ km}$ . Use eq. B.16 and eq. B.18 in order to approximate eq. 3.6. For the mathematical stability of your solution you have to make sure that the constant  $R = (\kappa \Delta t)/(\Delta z^2)$  is smaller than 0.25. This condition gives you the maximum time step  $\Delta t$  that you can use.  $\kappa$  is  $10^{-6} \text{ m}^2 \text{ s}^{-1}$ .

## 4. Kinematics: Morphology and Deformation

In this chapter we discuss the position, shape and the motion of rocks. In short: geodynamic processes measured in *meters*. This includes the discussion of strain and ventures therefore a bit into the field of structural geology. However, mostly we shall discuss processes like the elevation of mountain ranges, and the depth of the oceans, as well as the *change* of such parameters: kinematics. We begin with a consideration of the basics of strain. Our summary remains brief and the interested reader is referred to a number of excellent textbooks in the field of structural geology (Pollard and Fletcher 2006; Ramsay and Huber 1983; 1987; Ramsay and Lisle 2000; Twiss and Moores 1992; Pluijm and Marshack 1997).

### 4.1 Strain - The Basics

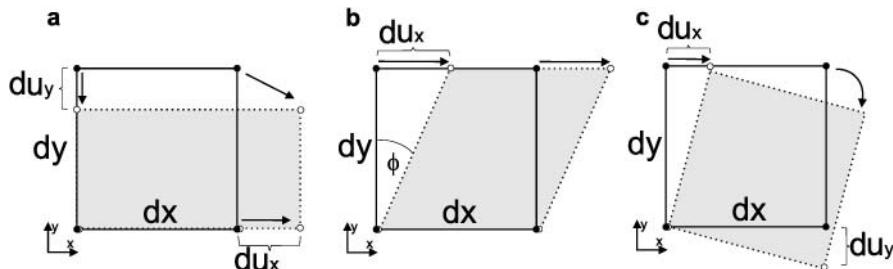
The deformation of rocks can be described by gradients in displacement, (at least for an infinitely small time step). Such displacement gradients may occur *in* the considered direction and *normal* to the considered direction. These two gradients are often referred to as normal strain and shear strain but these descriptions are somewhat imprecise and in fact sometimes wrong (as we will show). It is part of the aim of this section to clarify the meaning of displacement and strain on a simple and intuitive level. Because we want to keep things intuitive, we take a lot of unexplained shortcuts around a rigorous treatment of the subject (for example by assuming volume constancy and many other assumptions common to many rocks, but not explicitly stated here). We will also refrain from the excessive use of general subscripts  $i$  and  $j$  (both being short for  $x$ ,  $y$  and  $z$ ). We will also discuss strain in only two dimensions and use  $u_x$  and  $u_y$  for the displacements in  $x$  and  $y$  direction.

It is noted that most of the tensors related to deformation discussed here are very much analogous to the stress tensor (and its related quantities) discussed in sect. 5.1.1. Thus, you can read through that section and draw analogs to strain. Before venturing into the necessary tensors, we begin with a definition of the most common parameters referred to as “strain”:

- *Normal strain*. Normal- or longitudinal strains are encountered very often in this book, for example when we talk about the thickening parameters  $f_c$

and  $f_l$  (e.g. sect. 4.2.3) or about the stretching factor  $\beta$  in sect. 6.1.4. Normal strain simply relates the length of a body after deformation  $l$  to the length before deformation  $l_0$ . We define the following common terminology: The *stretch* of a rock  $s$  is the ratio of its length after deformation  $l$  to that before deformation  $l_0$ . Its *elongation* is the ratio of the change in length and the original length. We call this  $e$ . We can write the relationship between stretch (uniaxial strain), elongation and length in short:

$$s = \frac{l}{l_0} = 1 + e = 1 + \left( \frac{l - l_0}{l_0} \right) . \quad (4.1)$$



**Figure 4.1.** Deformation of a unity cube for the explanation of strain. **a** Deformation under normal strain also referred to as pure shear. **b** Deformation under shear strain also referred to as simple shear. **c** Rotation of the unity cube showing that displacement gradients can exist without internal strain. Conversely, deformation can usually be split into a rotational and a strain component.

Both  $s$  and  $e$  are often loosely referred to as “strain”. It is therefore important to understand their respective meaning when we try to understand kinematics. With reference to Fig. 4.1a:  $s = (dy + du_y)/dy$ , where  $du_y$  is the displacement  $du$  in direction  $y$ .

- *Shear strain.* is defined as the change in displacement with respect to a direction that is *normal* to this displacement. Shear strain is often referred to as  $\gamma$  and is given by the ratio of  $du_x$  to  $dy$  in Fig. 4.1 b so that:

$$\gamma = \tan \phi = \frac{du_x}{dy} . \quad (4.2)$$

The angle  $\phi$  is called the angular shear strain.

- *General displacement.* In a general state of deformation (for example somewhere between Fig. 4.1a and b) scalar values of  $s$  and  $\gamma$  are insufficient to describe the deformation and we require both shear and normal displacements, both in all considered spatial directions. In total, the deformation of a rock may be described by what is called the displacement gradient tensor, which is a tensor containing all shear and normal gradients in displacement. This tensor is given by:

$$D = \begin{pmatrix} \frac{\partial u_x}{\partial x} & \frac{\partial u_x}{\partial y} \\ \frac{\partial u_y}{\partial x} & \frac{\partial u_y}{\partial y} \end{pmatrix} \quad (4.3)$$

The displacement gradient is also called the *Jacobian matrix* or the *deformation tensor* and its individual components might be abbreviated with  $\partial u_i / \partial j$ . This tensor describes the deformation of a unity cube perfectly well. Clearly, the examples for simple and pure shear discussed above may also be written in terms of this tensor, but several of the terms will be zero.

It is important to note that the displacement gradient tensor does *not* describe strain. For example, the rotating cube in Fig. 4.1c has definitely displacement gradients in both  $x$  and  $y$  direction, but it does not strain. Indeed, even for the simple shear example shown in Fig. 4.1b, the shear strain  $\gamma$  does not describe the strain of the body correctly, as it may be shown that part of the “simple shear” deformation is rigid body rotation. Fortunately (because of its symmetry), the displacement gradient tensor may always be expanded so that it can be resolved into two parts:

$$\frac{\partial u_i}{\partial j} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right) \quad (4.4)$$

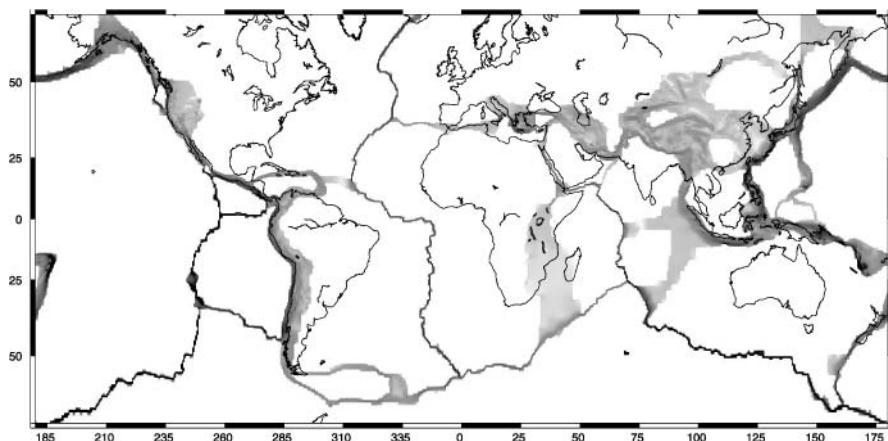
The first term on the right hand side of this equation is called the strain tensor  $\epsilon_{ij}$ , the second part describes the rigid body rotation  $\omega_{ij}$ . Adding a translation of the body  $u$ , we can write the full deformation of a body by:

$$u_i + du_i = u_i + \epsilon_{ij} dx_j + \omega_{ij} dx_i \quad (4.5)$$

Eq. 4.5 is a full description of deformation of rocks including their *translation* (first term), their *strain* (second term) and their *rotation* (third term). Rotational components of deformation are very much the field of structural geology and are not discussed further here. However, the strain tensor and its time derivative, the strain rate tensor, are needed in several parts of this book and we therefore write it out in full as:

$$\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \begin{pmatrix} \frac{\partial u_x}{\partial x} & \frac{1}{2} \left( \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \\ \frac{1}{2} \left( \frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y} \right) & \frac{\partial u_y}{\partial y} \end{pmatrix} \quad (4.6)$$

The strain rate tensor looks identical to eq. 4.6 if  $u$  is considered to be velocity and not displacement. Just like the stress tensor, the strain tensor is symmetric and the invariants of the strain tensor matrix are of some importance, for example when considering flow laws or shear heating. As an example, a map of global strain rate is shown in Fig. 4.2 where the magnitude of the 2nd invariant of the strain rate tensor is shown by the shading (for explanation see p. 304).

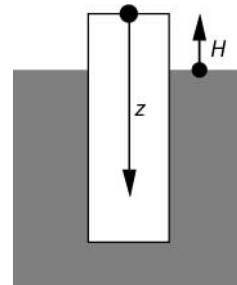


**Figure 4.2.** Global map of strain rate (Kreemer et al. 2003). Data from <http://gsrm.unavco.org/intro/>. Plotted is the 2nd invariant of the strain rate tensor on a logarithmic scale between 0 (white) and  $3 \times 10^{-6} \text{ y}^{-1}$  (black). Invariants of tensors are discussed in more detail in connection with the stress tensor in chapter 5. Compare also with Fig. 2.5

## 4.2 Reference Levels

All geological motion, for example the motion along faults, the uplift of a mountain range or the approaching of two lithospheric plates may only be observed *relatively* (s. Fig. 2.7). Thus, for any kinematic problem it is necessary to define a coordinate system to which the motion may be related. Usually we do this quite intuitively. For example, when talking of a *dextral fault*, we all understand that we mean the motion of *one* side of the fault as seen relative to a reference frame fixed to the *other* side of the same fault. For other examples, in particular those revolving around vertical motions or those around plate motions on a large scale there has been much confusion with reference frames in the literature. For example, in all previous sections (when discussing geotherms), we have fixed our coordinate system to the surface of the crust. This reference level is useful when describing geotherms as it is irrelevant to a thermal evolution whether the entire lithospheric column is being uplifted or remains stationary. However this reference level is not very meaningful when considering vertical motion of this surface itself: the elevation of a mountain range relative to its own surface is always zero. It cannot be described in a reference frame that is fixed to the surface. Thus, for problems dealing with surface elevation, we need to change our reference frame and fix it to sea level or some other externally fixed reference frame (Fig. 4.3). In sect. 4.3.1 we will show in some detail how careless handling of reference levels may lead to grave misinterpretations. However, we begin by discussing some important reference levels for geodynamic problems.

**Figure 4.3.** Schematic sketch of two different vertical axes that are used by earth scientists to describe vertical motions in the lithosphere. The surface elevation  $H$  is usually described relative to sea level or the reference lithosphere and is generally assumed to extend positively upwards. This reference frame is used by geomorphologists. The depth of rocks in the crust is generally measured relative to the surface and is measured positively downwards. This reference frame is used by petrologists. Note that both the direction and the origin of the axes are different for the two reference levels.

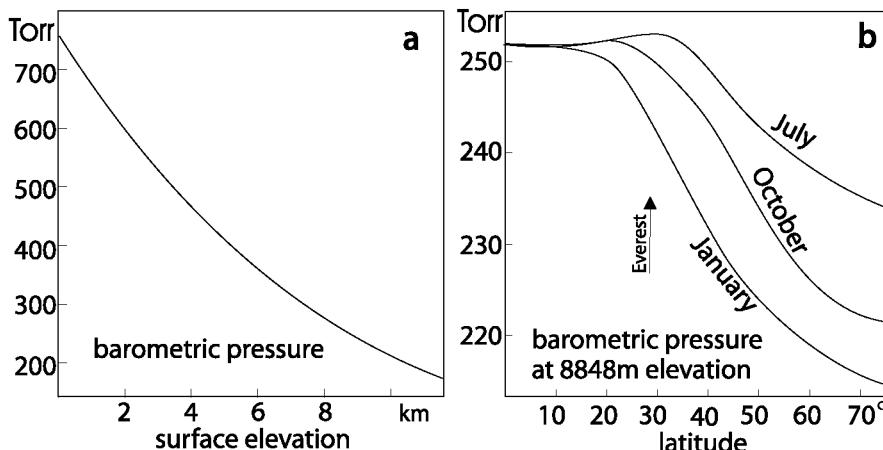


#### 4.2.1 Sphere, Spheroid and Geoid

When considering geological problems for which it is *not* possible to make the two-dimensional approximation that the earth is flat (s. sect. 2.2.1), it is necessary to make another approximation: we must find a good model description of the shape of our three-dimensional globe. The best first order approximation is that the earth is a *sphere*. For a long time one meter was defined as the 40 000 000th fraction of the circumference  $U$  of the globe. According to this definition, the radius  $R$  of a spherical earth would be:  $R = U/(2\pi) \approx 6366$  km. This approximation of the earth's radius is sufficient for most plate tectonic problems, for example those that consider the curvature of mountain belts or subduction zones, or those that consider the torques exerted by mid-oceanic ridges (s. sect. 2.2.2). However, the “real” shape of the earth deviates from a sphere by being flattened in the direction of the rotation axis: the centrifugal forces that arise because of the rotation of the globe cause the earth to have more the shape of an *ellipsoid*. This flattening of the globe at the poles is much more pronounced in the atmosphere than it is in the solid part of earth. In fact in the atmosphere it is pronounced enough so that Mt Everest could not be climbed if it were at any higher latitudes (Fig. 4.4). The polar radius of earth is  $R_P = 6356.75$  km. The equatorial radius is  $R_A = 6378.139$  km. Because the difference between polar and equatorial radius is very small (only about 20 km which is  $\approx 0.3\%$  of the radius), the ellipsoid of earth is also often referred to as a *spheroid*. The ellipsoid defined by these radii is the ellipsoid used by the World Geodetic System WGS84. This is the geodetic system in which the GPS measurements are based. The fractional difference between the two radii of an ellipsoid is given by:

$$f = \frac{R_A - R_P}{R_A} \approx 0.0034 \approx 1 : 298 \quad (4.7)$$

and is called the *ellipticity*. Using this ellipticity, the radius of earth at any location is often described by:



**Figure 4.4.** **a** The drop of atmospheric pressure with elevation and **b** the change of atmospheric pressure at the elevation of the top of Mt Everest with latitude. The pressure is given in Torr which corresponds to millimeters Hg. The conversion to SI units is 1 Torr = 133.32 Pa (s. Table D.8)

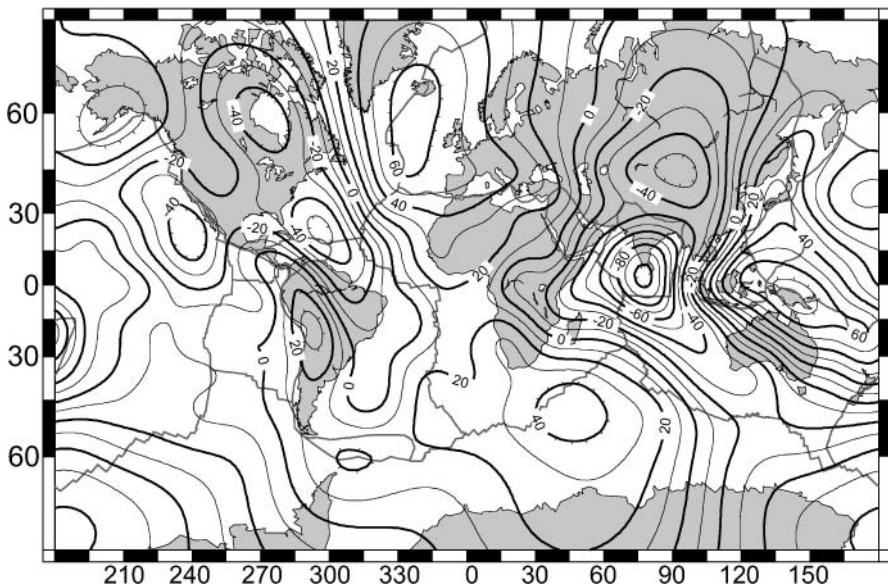
$$R = R_A (1 - f \sin^2 \lambda) \quad (4.8)$$

Another commonly used reference ellipsoid is the Bessel ellipsoid which differs in both radii by about 700 meters. The Bessel ellipsoid provides a better fit for many places, in particular for the northern hemisphere and is therefore used as the basis for many national mapping grids (sect. 2.3.2).

The *geoid* is the surface of constant gravitational potential energy (Fig. 4.5). In areas of high density, the gravitational acceleration is relatively high and the surface of the geoid lies low and vice versa (as to compensate the high acceleration with a lower mass of the column). *Geoid anomalies* are the differences between the equipotential surface of earth and a reference shape that is somewhat more complicated than an ellipsoid (Marsh et al. 1990).

#### 4.2.2 Lagrangian and Eulerian Reference Frames

Most problems we have discussed in the previous chapters of this book were discussed in fixed reference frames. That is, temperatures change and rocks move relative to the coordinate system. Such a coordinate system is called a *Eulerian reference frame*. The description of processes in Eulerian reference frames is referred to as *Eulerian* or *spatial* description. For many problems in the earth sciences, it is more useful to choose a coordinate system that moves with the rock. Such coordinate systems are called *Lagrangian reference frames*. Lagrangian descriptions are also referred to as *material description*. Fig. 4.6 illustrates both reference frames. Fig. 4.6a illustrates the evolution of geotherms during erosion. From the point of view of a Eulerian observer (who stands on the eroding surface) rocks (and the geotherm) are moved *through*



**Figure 4.5.** Geoid map of the world. The largest geoid anomaly (labeled in meters) is in the Indian Ocean where the geoid has an 80 m deep hole relative to the reference spheroid

the surface. From the point of view of a Lagrangian observer, the coordinate system is fixed to the rock and the location of the surface boundary condition changes its location downwards in the coordinate system.

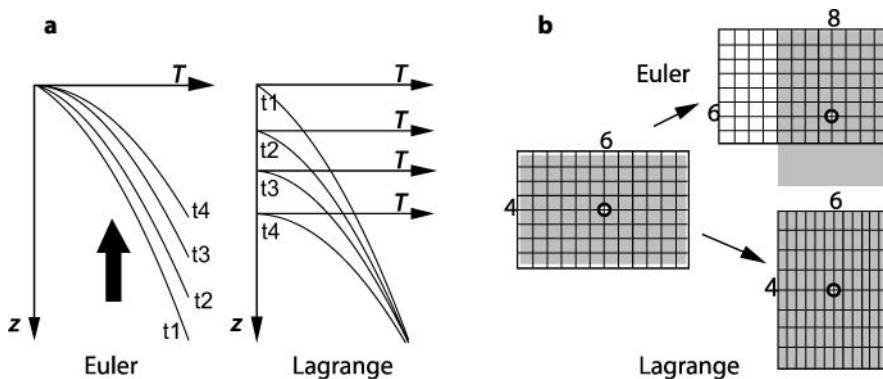
Both reference frames have advantages and disadvantages and it depends on the nature of a problem in terms of which reference frame it should be dealt with. The *Eulerian* description has the advantage that it is easy to describe the motion of material relative to an unmoved area, for example relative to the reference lithosphere. One advantage of *Lagrangian* descriptions is it that the history of a given rock is much easier to track. For example, in numerical calculations using a Lagrangian reference frame, the time dependent changes of a variable at a given grid node will always describe the evolution of a rock at this location, while in Eulerian reference frames, rock trajectories will go *through* the grid.

As the description of a problem changes with the choice of reference frame, so does the relevant equation. Consider the example of Fig. 4.6a. In a Eulerian description, the thermal evolution of the geotherm during erosion may be described as advection of the material through the grid. The relevant equation is the diffusion - advection equation (e.g. eq. 3.42). In steady state problems (which Fig. 4.6a is not), the temperature at a given depth stays constant, which makes it easy to illustrate, but rocks track through the coordinate system, which makes it difficult to track them. Within a Lagrangian description, the same problem may be described without advection. The process of erosion is described by moving the boundary condition through the coordinate

system. This is mathematically much more difficult to handle, but has the great advantage that a fixed coordinate describes the thermal evolution of a given rock. We will encounter the description of a very similar problem in two different reference frames in the context of Fig. 4.37.

Figure 4.6b illustrates another example of Lagrangian and Eulerian descriptions of the same problem, using a two-dimensional deformation model. The gray area represents a body of rock and the grid is a two-dimensional coordinate grid. The rock is deformed towards the walls on the right and at the top by forces acting from the left. Within a Eulerian reference frame, this deformation appears like material is being transported through the coordinate system. The coordinates of individual rocks change with time. For a Lagrangian observer the deformation appears to be caused by an approaching side wall and a retreating upper wall. Individual rocks remain at the same coordinates and the grid is distorted together with the rock (Fig. 4.6). Again, the principle advantage of the Lagrangian description is that a given coordinate remains fixed to a rock.

As a final example, consider the flow of a river across a waterfall from two different observers. The Eulerian observer (e.g. a fisherman on the river bank) will not see that the river accelerates. If he fishes above the waterfall he will observe that the water is slow and flows with constant velocity. If he looks towards the fall he will see that the water flows fast and also at constant velocity. The Eulerian observer (e.g. a fish floating with the stream) will feel the acceleration across the edge, but he has no sense of being elsewhere as the water around him does not move.



**Figure 4.6.** Differences between reference frames according to Lagrange and to Euler. **a** One-dimensional model of the evolution of geotherms during erosion. **b** Two-dimensional deformation of a square block. In both cases, the Eulerian description is within a coordinate system that is fixed externally (material is transported relative to the coordinate system) and the Lagrangian description is in a coordinate system that moves with the material

### 4.2.3 The Undeformed Reference Lithosphere

For most geological problems of the field geologist most reference levels discussed above are not very instructive. In fact, only if we deal with gravity data or global seismology, reference frames like the geoid, or the spheroid become important. For the large majority of the problems of the field geologist, the most useful reference frame is the *undeformed reference lithosphere* (Le Pichon et al. 1982). The undeformed reference lithosphere is a hypothetical lithospheric column which remains unchanged relative to the orogen under observation. If the undeformed state of the lithosphere under observation is unknown, then it is most simple to assume that the surface of the undeformed reference lithosphere corresponds to the mean surface of the continents. This is 840 m above sea level and the mean depth of the oceans is 3 700 m. However, about 80 % of the land surface is only between 100 and 200 m above sea level. All heights and elevations that are given in the following sections *without* a detailed mention of the reference level, are understood to be relative to the surface of the reference lithosphere. That is, they are relative to the elevation of the respective lithospheric level (e. g. the surface, the Moho etc.) *prior* to onset of thickening or thinning.

**The  $f_c$ - $f_l$ -Plane.** This plane is a diagram that allows to explore a range of parameters in comparison with an undeformed reference lithosphere. The axes of the plane show:

- 1. the thickness of the crust,
- 2. the thickness of the mantle part of the lithosphere,

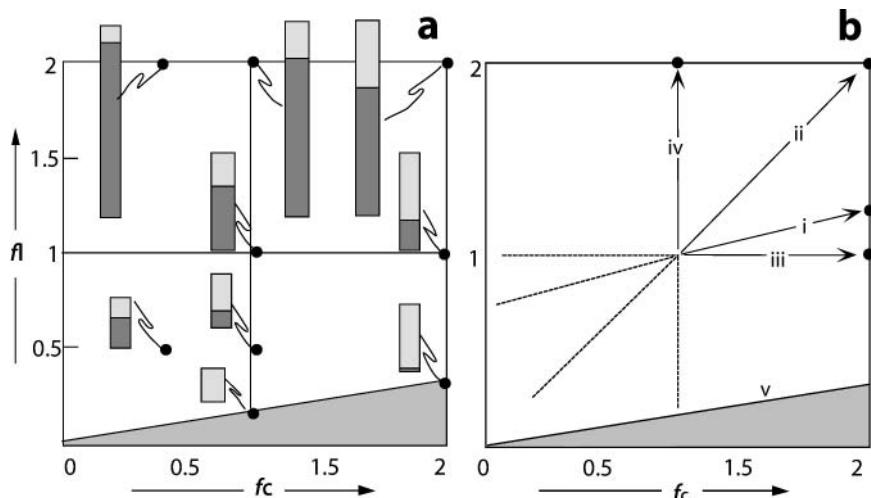
both relative to the reference lithosphere (s. Fig. 2.18). During orogenesis, these two parts of the lithosphere may change their thickness at different rates and by different amounts. It is therefore instructive to explore the influence of thickening of the two parts of the lithosphere explicitly but simultaneously. Strictly speaking,  $f_c$  and  $f_l$  are nothing but the vertical strains of the crust and lithosphere, but instead of the typical abbreviation fro strain,  $\epsilon$ , we follow the original authors and use  $f_c$  and  $f_l$ . They are defined as:

$$f_c = \frac{z_{\text{defc}}}{z_c} \quad \text{and} \quad f_l = \frac{z_{\text{defl}}}{z_l} \quad , \quad (4.9)$$

where  $z_{\text{defc}}$  and  $z_{\text{defl}}$  are the thicknesses of crust and lithosphere at a given time during orogenesis and  $z_c$  and  $z_l$  are the thicknesses of undeformed reference crust and lithosphere, respectively. Orogenic thickness evolutions may be plotted in the  $f_c$ - $f_l$ -plane as paths. Such paths are parametric in time ( $P - T$  paths in  $P - T$  space are also parametric int ime s. p. 338). This diagram was given to us by Sandiford and Powell (1990; 1991) and has since been used by a number of authors (e.g. Zhou and Stüwe 1994, Hawkesworth et al. 1995, Turner et al. 1995). Note that, according to the original definition of Sandiford and Powell (1990),  $f_l$  is the thickening strain of the *entire* lithosphere and not only that of the mantle lithosphere. While other definitions

are of course possible, we retain this original nomenclature here. Note also that  $f_1$  is the inverse of the stretching factor  $\beta$  ( $f_1 = 1/\beta$ ) that is commonly used in the description of continental extension (sect. 6.1.4).

Fig. 4.7a shows the  $f_c$ - $f_1$ -plane with some schematic lithospheric profiles of different deformation geometries of crust and mantle part of the lithosphere. As it is impossible that the whole lithosphere is *thinner* than the crust, the  $f_c$ - $f_1$ -plane is not defined in the region  $f_1 z_l < f_c z_c$  (shaded region in Fig. 4.7). The slope of the limiting line of this region is given by the initial thickness ratio of crust and whole lithosphere in the undeformed reference state:  $\phi = z_c/z_l$ . Fig. 4.7b shows a range of deformation paths that are end members of different orogenic evolutions.



**Figure 4.7.** The  $f_c$ - $f_1$ -plane. **a** The little schematic lithosphere columns indicate the thickness relationships of crust and mantle lithosphere in different parts of the diagram. Light shaded part of the columns is the crust, dark shaded part is the mantle part of the lithosphere. The point  $f_c = f_1 = 1$  is the reference lithosphere. **b** Some important lines in the  $f_c$ - $f_1$ -plane. Note that  $f_1$  describes the thickening strain of the entire lithosphere (that is: the mantle part of the lithosphere *plus* the crust). *i* is the line of constant thickness of the mantle lithosphere. *ii* is the line of homogeneous thickening (thinning) of the entire lithosphere. *iii* is the line of variable crustal thickness at constant lithospheric thickness. It implies that the mantle part of the lithosphere must thin along this line as the crust thickens and vice versa. *iv* illustrates thickening of the mantle part of the lithosphere without crustal thickening. *v* limits the allowed space in the  $f_c$ - $f_1$ -plane. Along this line, the crustal thickness and the lithospheric thickness are the same. That is, there is no mantle part of the lithosphere. Clearly, the lithosphere can not be thinner than the crust (not allowed dark shaded region). All lines are dashed in the region of thinning and continuous in the region of thickening relative to the undeformed reference lithosphere

The principal value of the  $f_c$ - $f_l$ -plane is that it may be contoured for a range of important geodynamic parameters, for example surface elevation, Moho temperature (Fig. 3.22), potential energy (Fig. 5.32), strain rate and others. The influence of the deformation geometry of the lithosphere on these parameters may then be explored in this diagram. As such, the  $f_c$ - $f_l$ -plane is in contrast to Fig. 3.20, where the thickness of the mantle part of the lithosphere is only *implicitly* determined by thermal considerations (s. Fig. 6.20). One of the disadvantages of the  $f_c$ - $f_l$ -plane is that both axes of the horizontal plane are used to describe the thickening strains. Thus, the use of the  $f_c$ - $f_l$ -plane and the two dimensions of this page limit us to the illustration of a single variable for which the plane may be contoured. For example on Fig. 3.22 the  $f_c$ - $f_l$ -plane is contoured for Moho-temperature as a characteristic temperature for the entire lithosphere.

## 4.3 Uplift and Exhumation

In the past decade, much progress in our understanding of mountain building processes has been made with studies in the interdisciplinary field between geomorphology and tectonics. In such studies, the interaction of two different processes that occur with respect to two different reference frames are studied (Fig. 4.3): 1. The evolution of surface elevation and landforms at the surface. 2. The evolution of depth and distance of rocks in the crust. The two different reference levels are:

- The distance to the surface of an undeformed reference lithosphere. This is the reference level labeled with  $H$  on Fig. 4.3.
- The distance to the surface of the lithosphere under consideration. This is the reference level we have used for most considerations in chap. 3.

The latter tells us about depth and pressure of rocks in the crust, the former about the geomorphic evolution of the surface. In the following section we discuss vertical motions of rocks relative to both of these reference levels. In this context, it is pointed out that the vertical axis  $z$  is sometimes measured positively upwards and sometimes positively downwards, depending on the question that is being asked (s. sect. 3.7.3). As it is very easy to get confused with these different reference frames, we begin with a careful definition of our terminology.

### 4.3.1 Definition of Uplift and Exhumation

Motion of rocks relative to the surface of the lithosphere under consideration is called *exhumation* or *burial*, depending on the motion being *towards* the surface or away from it (Tab. 4.1). In contrast to the geomorphological use of the word “exhumation” (where it is only used to describe the surfacing of

**Table 4.1.** Definition and methods of interpreting uplift

<b>uplift:</b>	vertical motion of the earth's surface relative to a reference level
direction of motion upwards:	called: uplift
direction of motion downwards:	called: subsidence
may be directly interpreted from:	palaeobotany, palaeoclimatology,
may be indirectly interpreted from:	sediments in the surrounding basins

**Table 4.2.** Definition and methods of interpreting exhumation

<b>exhumation:</b>	vertical motion of rocks, relative to the surface
direction of motion towards the surface:	called: exhumation
direction of motion away from the surface:	called: burial
may be directly interpreted from:	geobarometry geothermometry (via assumption of a geotherm)
may be indirectly interpreted from:	geochronology (via assumption of a geotherm)

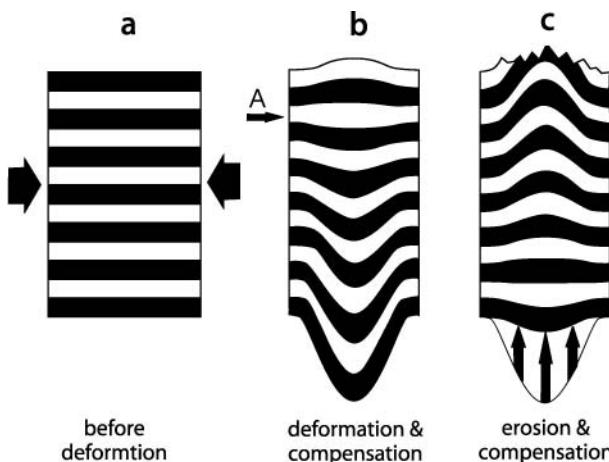
rocks that were previously at the surface, for example exhumation of a fossil or the exhumation of a river delta), tectonicists use the word “exhumation” also when describing upwards motion that has not brought rocks all the way to the surface (e.g. partial exhumation), or upwards motion of rocks that have never been on the surface previously (e.g. exhumation of a core complex) (Stüwe and Barr 1998). Exhumation is often used interchangeably with the word *denudation*.

Motion of rocks relative to the surface of an undeformed reference lithosphere is called *uplift* or *subsidence* depending on the motion being upwards or downwards with respect to an externally defined reference frame. In general, the words *uplift* and *subsidence* are only used to describe the vertical motion of the surface itself. Thus, when using these terms for another level in the crust, for example the uplift of a rock relative to another rock e.g. on the other side of a fault, then the term “uplift” should only be used together with a specification of the reference level – in this case: “the other side of a fault” (England and Molnar 1990). The terminology of *uplift* and *exhumation* is not very consistently used in the literature. England and Molnar (1990) have defined these terms precisely. In the following we will follow their definition and the expansion of their logic by Stüwe and Barr (1998). Table 4.1 and 4.2 summarize these definitions and their use.

Uplift and exhumation are both measured in units of distance. Uplift and exhumation *rates* are measured in velocities: in  $\text{m s}^{-1}$ . For the following discussion we call the uplift rate  $v_{\text{up}}$  and define it positively upwards. The exhumation rate is abbreviated with  $v_{\text{ex}}$ . The rate of *uplift of rocks*  $v_{\text{ro}}$  (relative to a fixed reference frame) is given by the sum of the uplift and the exhumation rate:

$$v_{\text{ro}} = v_{\text{ex}} + v_{\text{up}} . \quad (4.10)$$

The variable  $v_{\text{ro}}$  describes the vertical motion of rocks relative to a fixed reference level and is one of the most used and misused variables describing vertical motions. Ironically,  $v_{\text{ro}}$  is the only of the three vertical motions of eq. 4.10 which can not be determined directly from field or laboratory observations. It can *only* be determined from eq. 4.10. The importance of this equation can not be overemphasized. It is also emphasized that geobarometric or thermochronological data may *only* be used to infer exhumation and can not be used to interpret uplift. If there is no exhumation, then  $v_{\text{ro}} = v_{\text{up}}$ . In the next section we illustrate with some examples, how important the discrimination of different reference levels can be for the interpretation of tectonic features.



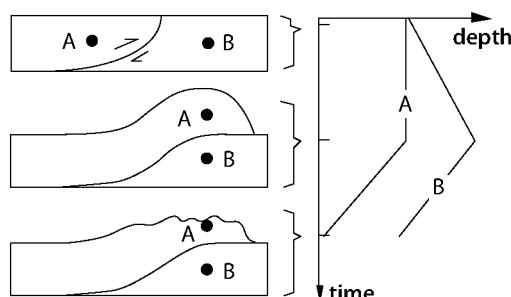
**Figure 4.8.** Schematic sketch of the vertical kinematics in isostatically compensated mountain belts. **a** shows a crustal column before deformation. **b** shows the crustal column after shortening. Note that all rocks below depth  $A$  are displaced downwards relative to **a**. **c** shows the crustal column after exhumation by erosion. Note that the dome structure that is now observed on the surface was formed by the exhumation processes only and *not* by the shortening

### 4.3.2 Vertical Motion in Active Orogens

In many mountain belts the highest grade metamorphic rocks crop out near the region of the highest topography. Well documented examples for the simultaneous occurrence of vertical motions are found in just about all convergent orogens, for example by the fact that the highest grade metamorphic rocks often crop out near the axis of highest surface elevation. In the eastern European Alps,  $P$ - $T$ - $t$ -paths have been documented that show that exhumation of rocks occurred synchronously with shortening deformation phases (Cliff et al. 1985). Equivalent observations have even been made in the granulite facies roots of ancient mountain belts (Carson et al. 1997). This suggests that the regions of the largest exhumation are also typically the region of largest uplift. This may not appear unusual, but the coincidence of the two has been the seed for much of the confusion about the different vertical motions. In fact, there are many examples where uplift and exhumation do not correlate. For example, there are small sedimentary basins on the Tibetan Plateau which imply that the rocks at the base of these basins were buried at a time of rapid surface uplift. Correspondingly, there are many sedimentary basins in which the surface subsides (negative uplift) during burial of rocks.

**Vertical Motion in Convergent Tectonics.** During convergent tectonics all rocks will be buried, i.e. increase their distance relative to the surface. However, Fig. 4.8b illustrated that – in isostatic equilibrium – rocks above depth  $A$  will be uplifted relative to an externally fixed reference frame during the process, while most rocks in the crustal column will not only increase in distance from the surface, but also move downwards with respect to an external reference frame. The upwards motion of rocks below  $A$  can only be caused by exhumation processes as shown in Fig. 4.8c. In the literature upwards and downwards motions are often confused, for example by drawing the vertical arrows like those on Fig. 4.8c on the same diagram as the shortening arrows like those on Fig. 4.8a.

**Figure 4.9.** Schematic illustration of the vertical motion of the hanging wall ( $A$ ) and the foot wall ( $B$ ) of a ramp anticline, relative to the surface. Note that the thrusting process itself does *not* cause any exhumation and that it is only the subsequent erosion that brings the rocks nearer the surface



If the shortening is not homogeneous as shown on Fig. 4.8b but discontinuous, for example during nappe stacking, then rocks will still only be buried

by the process. “Thrusting to the surface” is a commonly used term when talking of exhumation of rocks in association with a thrust, but the term is misleading. Fig. 4.9 shows that exhumation is *not* caused by the thrusting, but because the thrusting process places rocks in a position where they can easily be removed. If there is no erosion or other process removing material from the surface, then thrusting can not bring rocks closer to the surface! During thrusting, rocks in the upper plate will remain at constant depth below the surface, even though they are uplifted with respect to a reference frame fixed to an undeformed reference lithosphere (Fig. 4.9). Rocks in the foot wall on the other hand can only be buried.

### 4.3.3 Exhumation Processes

We discern a range of fundamentally different exhumation processes. Many of the will be encountered again in later sections of this book when we talk about dynamic processes (sect. 6.3.4). Here we restrict our discussion to the exhumation “end members” and how they differ with respect to the vertical motions of a crustal column. We discern:

- Exhumation by erosion,
- Exhumation by extension,
- Exhumation by compressive deformation.

**Exhumation by Erosion.** This is one of the earliest models proposed for exhumation, largely because of the recognition that mountain belts exposing high grade metamorphic rocks are flanked by sedimentary basins. Several studies have correlated volumes derived from sedimentary basins with those estimated from geobarometry and shown that there is a good correspondence (England 1981; Kuhlemann et al. 2001). Although we restrict our discussion here to erosion at the surface, it should be noted that erosion also occurs at depth, for example by scraping off material by an overriding plate. This process is often referred to as *subduction erosion* and has been described in Costa Rica, Mexico, Alaska and in connection with several other subduction zones. However, subduction erosion does not cause exhumation and will not be discussed further here.

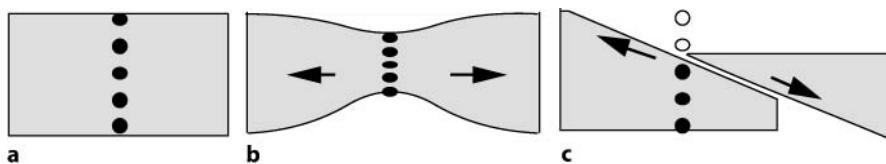
Exhumation by erosion involves always a vertical translation of the column upwards, or – more generally – in direction normal to the erosion surface (Fig. 4.11). As such, the column below the eroding surface remains intact, which has implications for the interpretation of metamorphic field gradients recording burial depth of rocks.

**Exhumation by Extension.** This is by far the most efficient exhumation mechanism for rocks from large depths. It is also referred to *tectonic denudation* and is a much discussed process in connection with the interpretation of high pressure metamorphic rocks (Avigard 1992; Platt 1993b) (sect. 6.3.4). Platt (1993b) subdivided exhumation processes that occur due to extension

in convergent orogens into those that occur in connection with underplating in orogenic wedges and those that occur in collisional orogens (sect. 6.2.2). This distinction is important when interpreting high pressure metamorphism. However, for the purpose of understanding vertical motions, it may be simpler to discern between

- Exhumation by normal faulting,
- Exhumation by ductile thinning.

Similar to erosion, normal faults and crustal scale detachments can bring large areas of high pressure metamorphic rocks practically undeformed to the surface and translate the vertical column upwards. In contrast, *pure shear* thinning or *simple shear* deformation can pervasively deform enormous rock packages causing telescoping or thinning of metamorphic isograds. Also note that ideal pure shear can never lead to complete exhumation as the complete thinned column of rock will always remain preserved (Fig. 4.10b). During extension by simple shear processes it is easy to exhume rocks as part of the column is physically removed from the top (Fig. 4.10c).



**Figure 4.10.** Exhumation during pure and simple shear of the lithosphere. **a** Starting situation with five marked depths. **b** In pure shear thinning rocks get partially exhumed, but can never exhume completely to the surface, even at very large extensional strains. The entire column above the rocks remains preserved in thinned form. **c** In simple shear it is easy to exhume rocks to the surface

**Exhumation by Compressive Deformation.** Although we have shown in section 4.3.2 that convergent deformation leads only to burial, it must be said that this is only true when this is considered one-dimensionally. In two or three dimensions exhumation during convergence can occur by buoyant up-raise, by channel flow, wedge forcing and a range of other processes that will be discussed in some more detail when we talk about orogenic wedges and exhumation of high pressure metamorphic rocks (sect. 6.3.4). Most exhumation mechanisms that occur during compressive deformation are characterized by a very heterogeneous distribution of vertical motions. In the field this may be reflected by the close proximity of high and low pressure metamorphic rocks.

#### 4.3.4 Modeling Vertical Kinematics

In most convergent orogens the various vertical motions discussed above occur simultaneously. Thus, it is often not clear at what stages of the evolution

rocks should exhume or get buried and if the surface uplifts or subsides at the time. In the following, we want to introduce an extremely simplistic one dimensional kinematic model that can be used to illustrate typical relationships between different vertical motions during orogenesis. The model is based on eq. 4.10 and describes a convergent orogen subject to homogeneous crustal thickening and erosion at the surface (Stüwe and Barr 1998). It bears many characteristic features of real orogens.

**Vertical Coordinates for the Model.** For the model we define  $z$  as the vertical distance of a rock from the surface of an undeformed reference lithosphere and  $z'$  as the vertical distance of the same rock from the surface above it, i.e. the depth of burial (Fig. 4.11). With these assumptions it is possible to describe the vertical motions of rocks with:

$$v_z = v_{\text{ro}} - \dot{\epsilon}(z + H) \quad . \quad (4.11)$$

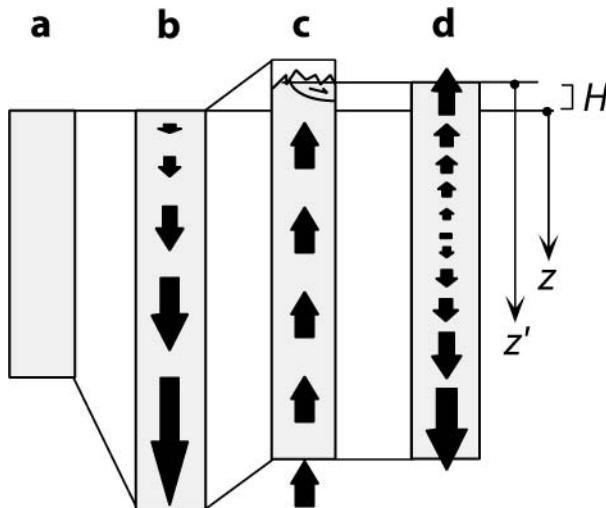
There,  $v_z$  is the *rate* with which rocks move relative to the reference column (measured positive if they move upwards),  $\dot{\epsilon}$  is the vertical strain rate (i.e. thickening rate) and  $H$  is the surface elevation of the lithosphere under consideration, again measured relative to the surface of an undeformed reference lithosphere. Eq. 4.11 is fundamental for the understanding of the following considerations so read it carefully before going on. If  $v_{\text{ro}} = 0$  and  $z = -H$  (this means: rocks at the surface), then:  $v_z = 0$ . Relative to the reference frame of the deforming lithosphere under consideration we can also write:

$$v_{z'} = v_{\text{ex}} = v_{\text{er}} - \dot{\epsilon}z' \quad . \quad (4.12)$$

There,  $v_{\text{er}}$  is the rate with which material is removed from the surface, e.g. the erosion rate. Note that  $v_{\text{er}}$  is different from  $v_{\text{ex}}$ . The rate of exhumation depends on the difference between  $v_{\text{er}}$  (exhuming the rocks) and the deforming strain rate  $\dot{\epsilon}$  (burying the rocks at a rate that is proportional to their depth  $z'$ ). However, it should be clear that  $v_{z'} = v_{\text{ex}}$ . The relationship between the different reference levels may also be written as:

$$z' = z + H \quad . \quad (4.13)$$

The first term on the right hand side of eq. 4.11 describes the vertical motion of the lithosphere as a consequence of removal of material at the surface *and* isostatic compensation thereof. It is positive, because  $v_z$  is defined positively upwards. The second term of the equation describes the vertical thickening of the lithosphere during thickening. It is negative because thickening leads to burial of rocks. The sum of both motions is  $v_z$  or  $v_{z'}$ , depending on the chosen reference frame (eq. 4.11). From eq. 4.11 we can read the following: At high crustal levels, where  $z$  is small, the second term of the equation is also small. The contribution of the first term is relatively large, so that  $v_z$  is likely to be positive. Rocks move upwards in the crust during simultaneous thickening and erosion at the surface (Fig. 4.11). At deep crustal levels the



**Figure 4.11.** Schematic illustration of the vertical motions of rocks in the crust relative to the surface of the undeformed reference lithosphere. *a* reference lithosphere. *b* vertical motion as the consequence of thickening. *c* vertical motion as the consequence of erosion at the surface and isostatic compensation. *d* vertical motion during simultaneous thickening and erosion (after Zhou and Stüwe 1994; Stüwe and Barr 1998)

second term is larger than the first term and  $v_z$  is likely to be negative. Rocks move downwards. It may be concluded that - even during homogeneous thickening - vertical motions in the crust may be very heterogeneous if there is simultaneous erosion.

**Modeling Simple Vertical Motions.** If we want to analyze these heterogeneous vertical motions in a bit more detail, we must pre empty some information from sect. 4.4, because we need eq. 4.35. This equation may be used to calculate the surface elevation of an isostatically compensated mountain range as a function of the thickness of crust and mantle lithosphere (expressed by  $f_c \times z_c$  and  $f_l \times z_l$ ) using some very simple assumptions on the density structure (expressed by the density terms  $\delta$  and  $\xi$ ). If we can find the time derivative of this equation we have a description for the *change* in elevation as a function of time: the uplift rate. The time derivative of eq. 4.35 is:

$$\frac{dH}{dt} = \delta z_c \left( \frac{df_c}{dt} \right) - \xi z_l \left( \frac{df_l}{dt} \right) . \quad (4.14)$$

We can find the time derivatives of  $f_c$  and  $f_l$  in eq. 4.14 if we assume that the crustal and lithospheric thickening rates (the change of  $f_c$  and  $f_l$  with time) are described by the difference between *thickening* due to deformation and *thinning* due to erosion at the surface:

$$\frac{df_c}{dt} = f_c \dot{\epsilon} - \frac{v_{er}}{z_c} \quad \text{and :} \quad \frac{df_l}{dt} = f_l \dot{\epsilon} - \frac{v_{er}}{z_l} .$$

Inserting these into eq. 4.14 we get the surface elevation as a function of time (the uplift rate). After some rearrangement of this equation, which you should be able to follow on paper, we find this to be:

$$\frac{dH}{dt} = v_{up} = v_{er}b - \dot{\epsilon}(H + a) . \quad (4.15)$$

In this equation,  $a$  and  $b$  summarize the constants from eq. 4.35 that are explained in eq. 4.32. They are:  $a = (\delta z_c - \xi z_l)$  and:  $b = (\delta - \xi)$ . Eq. 4.15 describes the evolution of surface elevation of a mountain range in which there is simultaneous vertical thickening (for example due to lateral shortening) and material removal from the surface (for example due to erosion).

Eq. 4.15 may be solved, if the erosion rate  $v_{er}$  is known. So let us assume a simple erosion model for a mountain range in which the erosion rate is proportional to the elevation of the range:

$$v_{er} = \frac{H}{t_E} . \quad (4.16)$$

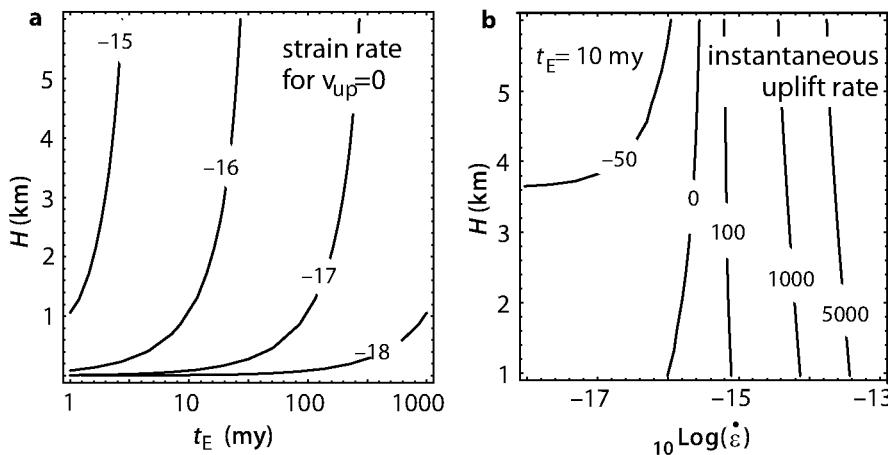
In sect. 4.5.1 we will discuss this equation in a bit more detail and show that this model may be quite a realistic description for many mountain belts. In eq. 4.16  $t_E$  is an erosional time constant that indicates the time scale over which the elevation  $H$  is removed by erosion. According to eq. 4.16 erosion is more rapid, if  $t_E$  is small. If we insert eq. 4.16 into eq. 4.15 it is possible to calculate the incremental uplift rates of a mountain range subject to the simple model boundary conditions assumed here.

**Geomorphic Steady State.** Eq. 4.16 states that the erosion rate gets higher as the elevation increases. The consequence of this model is that a mountain range will reach a geomorphic steady state when the elevation – and therefore the erosion rate – becomes high enough to balance further thickening. This is a realistic scenario observed in many mountain belts and can be described with the model of eq. 4.15. In this steady state,  $v_{ro}$  and  $v_{ex}$  at the surface have the same absolute value and it is true that:  $dH/dt = 0$ . From eq. 4.15 we get (s. Zhou and Stüwe 1994):

$$\dot{\epsilon} = \frac{bH}{t_E(H + a)} \quad \text{or} \quad H = \frac{t_E \dot{\epsilon} a}{b - \dot{\epsilon} t_E} . \quad (4.17)$$

Eq. 4.17 can be used to explore the geomorphic steady state constraints on mountain belts as a function of thickening strain rate and erosion rate as characterized by  $t_E$  (Fig. 4.12a). It may be seen that - for the surface to remain constant - the erosion rate must be larger (i.e.  $t_E$  must be smaller) the *higher* the mountain range and the *faster* the thickening rate. This result is actually quite intuitive.

Let us now consider the case of a mountain range *not* in the steady state and explore the uplift or subsidence rates it will undergo. However, we need

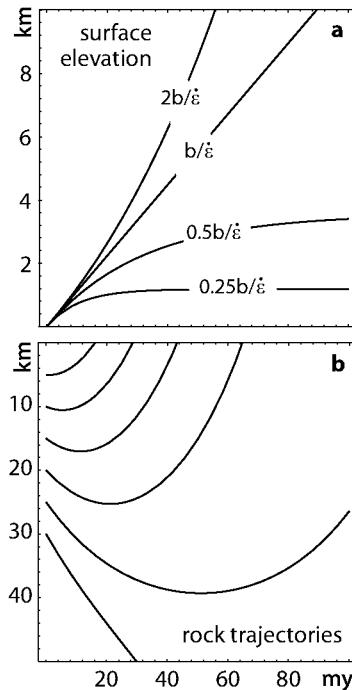


**Figure 4.12.** **a** Strain rates (contoured in  $10 \log(\dot{\epsilon})$ ) necessary to maintain a geomorphic steady state (where the uplift rate is zero:  $v_{up} = 0$ ) as a function of surface elevation  $H$  and erosion rate (characterized by the erosion parameter  $t_E$ ). The diagram was calculated with eq. 4.17. **b** The instantaneous uplift rate (contoured in meters per million years) of an isostatically compensated mountain range as a function of surface elevation and thickening strain rate. Calculated with eq. 4.15 using eq. 4.16 to describe the erosion rate. For uplift the values are positive, for subsidence negative. Assumptions for the physical parameters are:  $\rho_m = 3200 \text{ kg m}^{-3}$ ,  $\rho_c = 2700 \text{ kg m}^{-3}$ ,  $z_c = 35 \text{ km}$ ,  $z_l = 100 \text{ km}$ ,  $T_l = 1280^\circ\text{C}$ ,  $\alpha = 3 \cdot 10^{-5} \text{ K}^{-1}$ . With these parameters the constants in the equations are:  $a \approx 3545$  and  $b \approx 0.14$  (s. eq. 4.32)

to keep in mind that we can only explore the *instantaneous* uplift rates as they will immediately change as a different elevation is reached. This may be recognized as both  $(dH/dt)$  and  $H$  occur in eq. 4.15 and we have not expressed  $H$  as a function of time. Fig. 4.12b shows the instantaneous uplift rates as a function of thickening rate and surface elevation as calculated with eq. 4.15 and eq. 4.16. The figure still shows some interesting results. For example, it shows that for thickening strain rates below about  $\dot{\epsilon} = 10^{-15}$ ) the uplift rate is negative (subsidence occurs) and the subsidence rate increases with elevation, while this relationship is reversed for other thickening rates. Above around  $\dot{\epsilon} = 10^{-15}$  there is uplift with the uplift rate increasing for higher mountain belts. Let us remember that such counter intuitive results are useful to teach us about the processes controlling the evolution of surface elevation in mountain belts, but that the model presented here is far too simplistic for any direct application.

**Evolution of Surface Elevation.** We can also use the model above to explore the *evolution* of surface elevation with time. For this it is necessary to integrate eq. 4.15. Using the erosion model of eq. 4.16 integration of eq. 4.15 (using the method of Appendix B.5.1) gives:

**Figure 4.13.** The evolution of surface elevation and depth of rocks in the crust in a mountain belt during simultaneous thickening and erosion. **a** Surface elevation for four different erosion parameters  $t_E$  (eq. 4.18). **b** Rock trajectories for the assumption of:  $t_E = 0.5b/\dot{\epsilon}$  in **a**, as calculated with eq. 4.19. Note that all rocks with an initial depth of less than 30 km will be exhumed under these assumptions. All other assumptions are the same as in Fig. 4.12



$$H(t) = \frac{\dot{\epsilon}a}{\dot{\epsilon} - b/t_E} \left( e^{t(\dot{\epsilon} - b/t_E)} - 1 \right) . \quad (4.18)$$

Fig. 4.13a shows four examples for the evolution of surface elevation with time as described by eq. 4.18. It may be seen that surface elevation converges only to a steady state if  $t_E < b/\dot{\epsilon}$  (s. Problem 4.4).

So far we have only considered the vertical motion of the surface. However, we can use our model also to describe the vertical motions of rocks in the crust. From Fig. 4.11 we know that rock trajectories in the crust may be divided into two groups: paths that track upwards and those that track downwards in the crust. The two groups are separated by the point where  $v_z = 0$  or  $v_{ro} = 0$ . This point is of great geological importance, as rocks can only get exhumed by material removal from the surface if their depth  $z$  is  $z < z_{v_z=0}$ . We can actually calculate this point and all rock trajectories in the crust within our simple model. We can do that because we know that the vertical motion of rocks relative to an externally fixed reference frame is described by  $v_{ro} = (dz/dt)$ . Rock trajectories showing the evolution of depth through time may therefore be calculated by integrating eq. 4.10. After inserting eq. 4.15 and eq. 4.13 in eq. 4.10 and integrating (which we do not derive in detail here) we get:

$$z(t) = z_i e^{\dot{\epsilon}t} + \frac{a}{b} (1 - e^{\dot{\epsilon}t}) + H \left( \frac{1 - b}{b} \right) . \quad (4.19)$$

There,  $z_i$  is the initial depth of rocks in the crust. Fig. 4.13b shows rock trajectories calculated with eq. 4.19. Within our simple model, the depth  $z_{v_z=0}$  or  $z_{v_{ro}=0}$  may now be calculated reasonable easy. It is given by inserting eq. 4.15 and eq. 4.13 in eq. 4.10 assuming  $v_{ro} = 0$ . Solving for  $z$  gives:

$$z_{(v_z=0)} = a + \frac{H}{\dot{\epsilon}_E} (1 - b) . \quad (4.20)$$

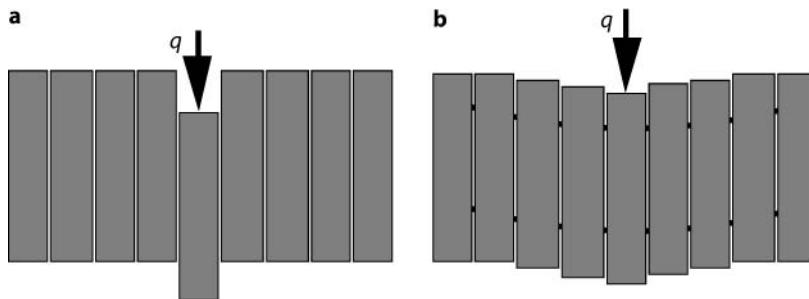
Inserting a range of realistic numerical values for the physical parameters into eq. 4.20 shows that this transition point lies at roughly 30–40 km depth. This means, that rocks can only be exhumed in orogens subject to simultaneous thickening and erosion at the surface if they lie at depths shallower than 30–40 km. This is in good correspondence with our observation that greenschist and amphibolite facies rocks with metamorphic pressures up to roughly 10 kbar (see eq. 7.1) are common in convergent orogens, while eclogites and other high pressure rocks are rare and usually confined to structures of the orogen that can only be described with two-dimensional models (e.g. subduction zones, lithosphere scale thrusts etc.). Our model also shows that the exhumation of high pressure rocks requires consideration of other exhumation mechanisms (s. sect. 6.3.4; s. Platt 1993b).

## 4.4 Isostasy

Isostasy is a stress balance. However, since it is used to describe the elevation of the surface the concept of isostasy is discussed in this chapter. Isostasy relates the vertical distribution of mass to elevation in a state of equilibrium in which the lithosphere is considered to be floating on the underlying relatively weak asthenosphere. Isostasy does a good job of explaining the first-order variation of elevation over most of the earth's surface. In general, isostasy is concerned with the comparison of the surface elevation in two different places. For example, we might want to interpret the elevation difference between a mountain range and its foreland (assuming *isostatic* equilibrium) in terms of its implications for their different thicknesses. When we consider isostatic equilibrium it is useful to discern:

- hydrostatic isostasy and
- flexural isostasy.

*Hydrostatic isostasy* is a stress balance in the vertical direction only (s. sect. 5.1.1). Thus, hydrostatic isostasy is a model that should really only be applied to regions that are large compared to the elastic thickness of the lithosphere. In other words, to geological features that are of at least several hundreds of kilometers in extent, i.e. areas like the Tibetan Plateau or the Canadian Shield. *Flexural isostasy* describes a stress balance in two or even three dimensions (s. Fig. 4.14). As a consequence, flexural isostatic considerations can be used to interpret the shape of much smaller scale features, for example foreland basins or subduction zones.



**Figure 4.14.** Illustration of the difference between **a** hydrostatic isostasy and **b** flexural isostasy. In **a** all vertical columns are considered independently of each other. In **b** the shear stresses between vertical columns are also considered.  $q$  is the load

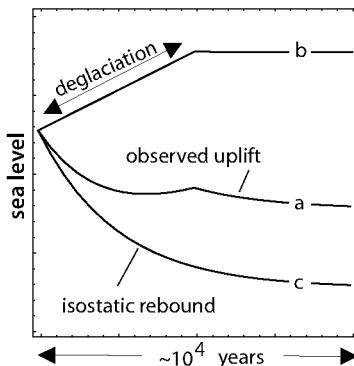
**Isostatic Equilibration Rates.** Isostasy describes an equilibrium state and is therefore independent of time. Nevertheless, many geologists misinterpret the temporal evolution of isostatic rebound as a feature inherent to isostasy. For example, we observe that isostatic equilibrium of Scandinavia in response to its deglaciation in the last ice age, occurs on a time scale of  $10^4$  years (e.g. Sabodini et al. 1991). Such isostatic compensation rates can be measured, for example by dating raised beaches (Fig. 4.15). However, this observation does *not* tell us that isostasy itself is time-dependent. Isostasy is a stress balance and as such *independent* of time. If a plate tries to rise or sink to reach its isostatic equilibrium state in response to a changed load, it has to displace the underlying asthenosphere. Thus, the rate of isostatic compensation can be used to estimate the viscosity of the asthenosphere (e.g. Lambeck 1993).

#### 4.4.1 Hydrostatic Isostasy

The hydrostatic isostatic model is based on the assumption that all vertical profiles through the lithosphere may be considered independently of each other. That is, shear stresses on vertical planes are neglected (Fig. 4.14a). Then, there will be a depth at which the vertical stresses of all vertical profiles are equal. This depth is called the *isostatic compensation depth*. At this depth, the weight of all columns are equal. If you dive underneath a boat you dive beneath this isostatic compensation depth: Regardless if the boat is above you or not, the water pressure is the same. If we consider two profiles A and B, the *isostasy condition* may be formulated in terms of an equation (s. Fig. 4.16):

$$\sigma_{zz}^A|_{z=z_K} = \sigma_{zz}^B|_{z=z_K} . \quad (4.21)$$

In this equation  $\sigma_{zz}^A$  and  $\sigma_{zz}^B$  are the vertical normal stresses of the two columns A and B and the depth  $z_K$  is the isostatic compensation depth. The vertical dash stands for “at the location”. For most geological purposes we want to compare the elevation of two neighboring lithospheric columns



**Figure 4.15.** Observed and interpreted sea level changes. *a* Typical evolution of surface uplift relative to sea level in regions of recent deglaciation, for example Scandinavia (e.g. Lambeck 1991). Such curves typically contain two distinct parts and may be interpreted as the sum of sea level *rise* due to deglaciation (because of increased water mass in the oceans) and sea level *drop* because of isostatic rebound. A simple example for the former is shown in curve *b* consisting of a linear sea level rise until deglaciation ceases. Isostatic rebound (curve *c*) decreases exponentially as isostatic equilibrium is approached and therefore outlasts the deglaciation. From such curves, mantle viscosities of the order of  $10^{20}$  Poise have been calculated

in isostatic equilibrium. For this, it is useful to assume as isostatic compensation depth the shallowest possible depth below which there is no density differences between two neighboring columns. For most examples this can be assumed to be the base of the lithosphere of the column which reaches deepest into the asthenosphere.

The downward force that is exerted by one cubic meter of rock is given by the product of its mass  $\times$  gravitational acceleration. The downward force that is exerted by an entire vertical column per square meter (the vertical normal stress) is thus the product of density and acceleration integrated over the thickness of the column:

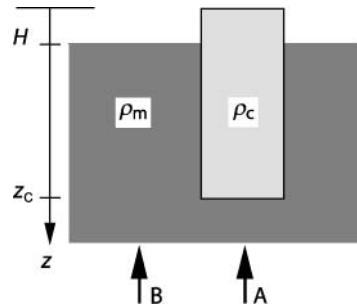
$$\sigma_{zz} \Big|_{z=z_K} = \int_0^{z_K} \rho g dz . \quad (4.22)$$

Inserting eq. 4.22 into eq. 4.21 gives:

$$\int_0^{z_K} \rho_A(z) g dz = \int_0^{z_K} \rho_B(z) g dz , \quad (4.23)$$

where  $\rho_A(z)$  and  $\rho_B(z)$  are the densities of the two columns that are to be compared, both as a function of depth,  $z$ . Within the coordinate system shown in Fig. 4.16, the lower limit of integration 0 corresponds to the upper surface of the higher of two columns that are to be compared. The upper limit of integration is the isostatic compensation depth  $z_K$ .  $g$  is the gravitational acceleration. Eq. 4.23 is the basis of all calculations of isostasy.

**Figure 4.16.** Illustration of isostatic equilibrium. Note that the  $z$ -axis is defined positively downwards and has its origin at the surface of the light shaded block (e.g. an iceberg or the lithosphere) that is assumed to float in a dark shaded region of higher density (e.g. water or the asthenosphere)



When considering the isostatically supported elevation of a mountain belt, it is useful to divide the density variations in the lithosphere into two parts:

- density variations that are due to material differences and
- density variations that are caused by thermal expansion.

The fact that both types of density variations may form significant contributions to the density structure of the lithosphere is familiar to us from Fig. 2.18. We begin by considering material differences.

**Isostasy due to Material Differences.** We begin by solving the balance written out in eq. 4.23 for the elevation of a single lithospheric column above the asthenosphere. For simplicity we forget in the first instance about the mantle part of the lithosphere and consider the crust only so that  $z_K = z_c$  (Fig. 4.16; s. also sect. 5.1.1). The block in Fig. 4.16 has a constant density  $\rho_c$  (e.g. density of the crust) and floats in a denser medium of the constant density  $\rho_m$  (e.g. density of the mantle). We call its elevation above the surface of the denser medium  $H_{\text{mat}}$ , although it is just labeled as  $H$  in Fig. 4.16. We use the subscript <sub>mat</sub> to emphasize that – for now – we consider only the *material* contribution to density differences between the profiles A and B. The densities and the acceleration are independent of  $z$ . Thus, they can be drawn out of the integrals on both sides of eq. 4.23 and integration is easy. By integrating the left half of the equation and splitting up the right half of eq. 4.23 we get according to Fig. 4.16:

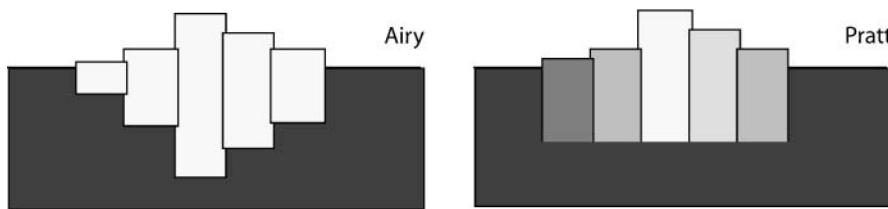
$$\rho_c g z \Big|_0^{z_c} = g \int_0^{H_{\text{mat}}} \rho_{\text{air}} dz + g \int_{H_{\text{mat}}}^{z_c} \rho_m dz . \quad (4.24)$$

The density of air is negligible in comparison with  $\rho_m$  or  $\rho_c$ . Thus, the first integral on the right hand side of eq. 4.24 is also negligible. After finishing the integration, canceling out  $g$  and inserting the integration limits we get:

$$\rho_c z_c = \rho_m z_c - \rho_m H_{\text{mat}} . \quad (4.25)$$

Solving for elevation  $H$  gives:

$$H = H_{\text{mat}} = z_c \left( \frac{\rho_m - \rho_c}{\rho_m} \right) . \quad (4.26)$$



**Figure 4.17.** Comparison of the interpretations of the isostatic model according to Airy and Pratt. The shading indicates density. Darker shading means higher density

This relationship describes the hydrostatically balanced elevation of the surface of a floating body above the medium it floats in. Remember that  $H = H_{\text{mat}}$  emphasizes the fact that this elevation difference is only based on the material difference between the block and the liquid. We can control this equation for some end member scenarios: If  $\rho_c$  is zero, then this equation states that  $H = z_l$ : the entire column floats on top of the liquid. This is the scenario given by a kids balloon floating on a lake. Alternatively if the two densities approach each other ( $\rho_m = \rho_c$ ), then the entire body is submerged ( $H = 0$ ). This is the scenario we observe with water soaked logs that float almost completely submerged in water. We can conclude that our observations confirm the simple model.

**Isostasy According to Airy and Pratt.** Two centuries ago, different models were developed to explain elevation differences observed in the mountain belts of the world in terms of the isostasy model. The two most notable models are those of *Airy* and *Pratt* (Fig. 4.17). Both earth scientists recognized that mountain belts are likely to rest in isostatic equilibrium and that their elevation is proportional to the density contrast between crust and mantle, as expressed by eq. 4.26. Pratt observed that many low lying Proterozoic shields are made up of high grade metamorphic rocks of high density, while mountain belts are often made up of hydrated, low grade metasediments and carbonates. He concluded that most continental crusts extend to roughly similar depths and that the observed differences in surface elevation are the consequence of horizontal density variations in the crust.

In contrast, Airy estimated that the density of the crust is largely the same in all continental regions and therefore concluded that topographically higher regions, must be compensated by crustal roots at depth. The models of Airy and Pratt still bear their names. Seismic studies in many mountain belts show that most regions of high surface elevation are indeed compensated by significant roots at depth. On the other hand, it is true that there is a relationship between surface elevation and density of rocks. In short, the truth lies between the models of Airy and Pratt, although much success has been made by following Airy's model.

**Isostasy Due to Thermal Expansion.** In order to calculate the contribution of thermal expansion to surface elevation we need to introduce

$\alpha$ : the *coefficient of thermal expansion*, which we encountered already in eq. 3.29.  $\alpha$  has the units of strain per temperature increment, which is  $\text{K}^{-1}$  (s. sect. 5.1). For most rocks the coefficient of thermal expansion is of the order  $\alpha = 3 \cdot 10^{-5} \text{ K}^{-1}$ . Using  $\alpha$  and the density of the mantle  $\rho_m$  (at the temperature of the asthenosphere), the density of colder rocks of the same material as a function of temperature may be calculated with:

$$\rho(T) = \rho_m(1 + \alpha(T_1 - T)) . \quad (4.27)$$

There,  $T_1$  is the temperature at the base of the lithosphere at  $z = z_1$ . According to eq. 4.27:  $\rho = \rho_m$ , where  $T = T_1$ . At lower temperatures, the density increases linearly. At the surface, where we can assume that the temperature is  $T_s = 0^\circ\text{C}$ , eq. 4.27 becomes:

$$\rho(T=T_s) = \rho_0 = \rho_m(1 + \alpha T_1) . \quad (4.28)$$

If the density of the mantle is about  $\rho_m = 3200 \text{ kg m}^{-3}$  at  $T_1$ , then the density at the surface is:  $\rho_0 = 3300 \text{ kg m}^{-3}$ . Assuming a linear geotherm in the lithosphere, we can describe the mean density of the lithosphere with:

$$\bar{\rho} = \rho_m \left( 1 + \alpha \frac{T_1 + T_s}{2} \right) . \quad (4.29)$$

In order to estimate which proportion of the elevation of a mountain belt is due to thermal expansion ( $H_{\text{therm}}$ ), we insert eq. 4.29 into the left hand side of eq. 4.23. The following algebra remains the same as in eq. 4.24 and eq. 4.25 except that the upper limit of integration is not the base of the crust, but the base of the lithosphere, because thermal expansion and contraction affects the entire lithospheric column. After integration according to the same principles as we did before we get here:

$$H_{\text{therm}} = -z_1 \alpha (T_1 + T_s)/2 . \quad (4.30)$$

The negative sign arises because  $\bar{\rho}$  is larger than  $\rho_m$ .

**The Elevation of Mountain Belts.** First off a warning: Gravimetric data tell us that many active orogens are *not* in isostatic equilibrium, but that their topography is dynamically supported. This means the surface elevation is actively held up or pushed down and is *out of* isostatic equilibrium. Dynamically supported topography may generally be found on length scales that are comparable to the elastic thickness of the lithosphere and will be discussed there (e. g. Forsyth 1985; Lyon-Caen and Molnar 1983; Molnar and Lyon-Caen 1989) (sect. 4.4.2). It is therefore emphasized that the model of hydrostatic isostasy should only be used for topographic features that are at least some hundreds of kilometers in lateral extent. For example, the European Alps are barely 200 kilometers across and are only partly compensated isostatically (Karner and Watts 1983). This limitation of the hydrostatic model should be kept in mind when we interpret the simple considerations below.

Nevertheless, let us now consider the elevation of a lithosphere with the thickness  $z_l$  and a crustal thickness of  $z_c$  above its surroundings considering both the influence of the different materials and the influence of thermal expansion. The higher density of the cold lithosphere provides a *negative* contribution to the overall buoyancy (eq. 4.30). The *material* contribution of the crust to the elevation, on the other hand, is positive and was derived in eq. 4.26. Density variations within the mantle part of the lithosphere are neglected here (s. however p. 37). Then, the isostatically supported surface elevation relative to the surroundings is given be the sum of the thermal and the material contributions:

$$H = H_{\text{mat}} + H_{\text{therm}} = z_c \left( \frac{\rho_m - \rho_c}{\rho_m} \right) - z_l \alpha (T_l + T_s)/2 . \quad (4.31)$$

If we summarize all the material parameters into the constants:

$$\delta = (\rho_m - \rho_c)/\rho_m \quad \text{and :} \quad \xi = \alpha(T_l + T_s)/2 , \quad (4.32)$$

then this eq. 4.31 simplifies to:

$$H = \delta z_c - \xi z_l . \quad (4.33)$$

If we insert meaningful numbers into eq. 4.31 (e.g.  $\rho_m = 3200 \text{ kg m}^{-3}$ ,  $\rho_c = 2700 \text{ kg m}^{-3}$ ), we get:

$$\delta \approx 0.15 \quad \text{and :} \quad \xi \approx 0.018 . \quad (4.34)$$

This implies that the influence of material difference between crust and mantle, per meter of lithospheric column, is about ten times more important to the isostatically supported surface elevation than the influence of the thermal expansion. However, because the crust constitutes only about one third of the lithosphere, the crustal material contribution to the elevation is in total only about 3 times larger than the contribution of thermal contraction, which applies to the whole lithosphere. In total,  $H$  is about 3600 m.

This is the elevation of the upper surface (of a lithosphere with  $z_c$  and  $z_l$  as above) above the hypothetical surface of a liquid mantle, as we illustrated in Fig. 4.16. Mid-oceanic ridges are the only place on the globe where we can measure the depth of this reference level. It turns out that mid-oceanic ridges lie indeed about 3600 m below the average elevation of the continents and lie at a very constant depth below sea level (Turcotte et al. 1977; Cochran 1982).

In most geological problems it is much more interesting to know the elevation of a mountain belt above its surroundings, rather than above the mid-oceanic ridges. For this purpose, it is useful to reformulate eq. 4.31, so that the elevation is given as the elevation difference between a thickened (or thinned) lithosphere and an undeformed reference lithosphere:

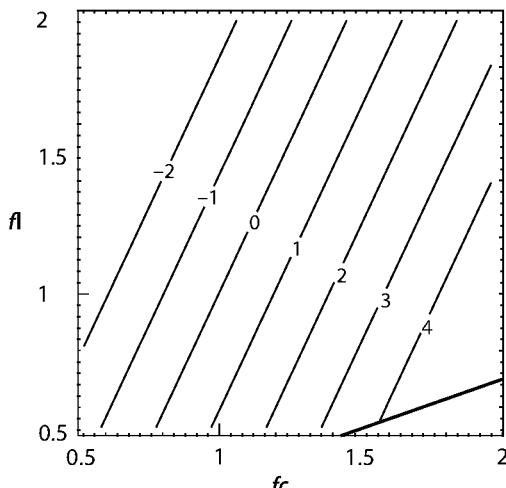
$$H = (\delta f_c z_c - \xi f_l z_l) - (\delta z_c - \xi z_l) = \delta z_c (f_c - 1) - \xi z_l (f_l - 1) . \quad (4.35)$$

The parameters  $f_c$  and  $f_l$  describe the thickening strains of the crust and the mantle lithosphere and were discussed in detail in sect. 3.4.4 and 4.2.3 (also: sect. 6.1.4, eq. 6.10). The elevation of isostatically supported mountain belts above the undeformed reference lithosphere is shown in Fig. 4.18 (for the concept of an undeformed reference lithosphere see: Le Pichon et al. 1982). More detailed assumptions about the thermal expansion have no influence on the surface elevation (e. g. Zhou and Sandiford 1992). Fig. 4.18 shows clearly that homogeneous thickening of the entire lithosphere (a diagonal line from bottom left to top right in this diagram) causes relatively small changes of the surface elevation, because the two contributions in eq. 4.33 and eq. 4.35 have opposite signs. Accordingly, the negative buoyancy caused by the thickening of the mantle part of the lithosphere is largely compensated by the positive buoyancy of the thickened crust. It may also be read from this figures, that doubling of the crust, without thickening of the lithosphere would imply an isostatic uplift of about 3–4 km (compare the paths in Fig. 4.7b, s. sect. 2.4.1).

**The Depth of the Oceans.** The water depth of the oceans is a direct function of the distance to the mid-oceanic ridges (Fig. 2.1). The functional relationship between water depth and distance from the mid-oceanic ridge was described with a fantastically simple model by Parsons and Sclater (1977). Their model is one of the largest successes of the theory of heat conduction (sect. 2.1, 3.5.1). It can be derived using the principles of hydrostatic isostasy.

Oceanic lithosphere consists (except for a thin 7 km thick crust) largely of asthenosphere material that has cooled to form lithospheric mantle. Because of the small and constant thickness of the crust, material contributions to density variations may be neglected and thermal expansion (contraction) is the governing factor for variations in the density structure. In order to use

**Figure 4.18.** Isostatically supported surface elevation of mountain belts in the  $f_c$ - $f_l$ -plane (after Sandiford and Powell 1990). Contours were calculated with eq. 4.35 and are labeled in km. Following assumptions were used:  $\rho_m = 3\,200$ ,  $\rho_c = 2\,750$ ,  $\alpha = 3 \cdot 10^{-5}$ ,  $z_c = 35$  km,  $z_l = 100$  km. Using these values, the two constants are:  $\delta \approx 0.14$  and  $\xi = 0.018$ . The thick line in the lower right hand corner of the diagram delineates the not allowed region of this parameter space (s. sect. 4.2.3)



to estimate the isostatically supported elevation of the ocean floor, we use the model sketched in Fig. 4.19. According to eq. 4.23 the vertical normal stresses of the columns A and B must be the same in the compensation depth  $z = z_1$ . For column A the vertical normal stress at depth  $z = z_1$  is given by:

$$\sigma_{zz}^A|_{z=z_1} = \rho_w gw + \int_0^{z_1} \rho_{(z)} gdz . \quad (4.36)$$

There,  $w$  is the water depth in column A,  $\rho_w$  is the water density,  $g$  is the gravitational acceleration and  $\rho_{(z)}$  is the density of the lithosphere as a function of depth which we shall discuss on the next page. For column B we can formulate:

$$\sigma_{zz}^B|_{z=z_1} = \rho_m gw + \rho_m g z_1 . \quad (4.37)$$

It should be possible to follow eqs. 4.36 and 4.37 by considering Fig. 4.19. After inserting eqs. 4.36 and 4.37 into eq. 4.23, the isostasy condition of gets the following form:

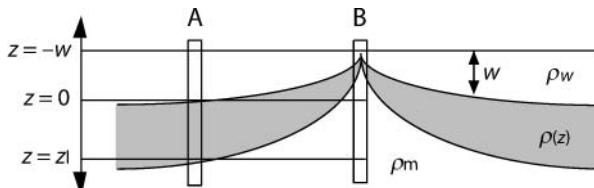
$$\rho_m z_1 + w(\rho_m - \rho_w) = \int_0^{z_1} \rho_{(z)} dz . \quad (4.38)$$

With foresight to the following steps, we bring the first term of this equation to the right hand side, find its derivative with respect to  $z$  and write it therefore into the integral. Eq. 4.38 gets the form:

$$w(\rho_m - \rho_w) = \int_0^{z_1} (\rho_{(z)} - \rho_m) dz . \quad (4.39)$$

This equations states that the water depth is dependent on the density structure as a function of depth  $\rho_{(z)}$ . In oceanic lithosphere this density function is a direct function of the temperature profile (s. sect. 3.5). Thus, if we know the temperature as a function of depth, then  $\rho_{(z)}$  in eq. 4.39 is known, because we know already the relationship between density and temperature from eq. 4.27. Thus we can begin by inserting eq. 4.27 into eq. 4.39:

$$w(\rho_m - \rho_w) = \int_0^{z_1} \rho_m \alpha (T_1 - T(z)) dz . \quad (4.40)$$



**Figure 4.19.** Schematic profile through a mid-oceanic ridge and the oceanic lithosphere as used for the calculation of water depth. The oceanic crust is neglected because it is everywhere of the same thickness

The variable  $T(z)$  is the only unknown in this equation, but we determined it in sect. 3.5.1. It is well-described by the half-space cooling model. Thus, the temperature profile of eq. 3.80 may be directly inserted into eq. 4.40:

$$w(\rho_m - \rho_w) = \int_0^{z_l} \rho_m \alpha \left( T_l - T_s - (T_l - T_s) \operatorname{erf} \left( \frac{z}{\sqrt{4\kappa t}} \right) \right) dz . \quad (4.41)$$

After simplification using eq. 3.16 we get:

$$w(\rho_m - \rho_w) = \int_0^{z_l} \rho_m \alpha (T_l - T_s) \operatorname{erfc} \left( \frac{z}{\sqrt{4\kappa t}} \right) dz \quad (4.42)$$

or, after taking the constants out of the integral and solving for  $w$ :

$$w = \frac{\rho_m \alpha (T_l - T_s)}{(\rho_m - \rho_w)} \int_0^{z_l} \operatorname{erfc} \left( \frac{z}{\sqrt{4\kappa t}} \right) dz . \quad (4.43)$$

If we introduce the variable  $n = z / \sqrt{4\kappa t}$ , we can take all the constants out of the integral (s. Appendix C) and get:

$$w = \sqrt{4\kappa t} \frac{\rho_m \alpha (T_l - T_s)}{(\rho_w - \rho_m)} \int_0^{z_l} \operatorname{erfc}(n) dn . \quad (4.44)$$

The definite integral of the error function is not known for integration limits of 0 and  $z_l$ , but it *is* known for integration with limits at infinity. It is:

$$\int_0^{\infty} \operatorname{erfc}(n) dn = \frac{1}{\sqrt{\pi}} .$$

This is a close enough approximation, in particular since  $\rho \rightarrow \rho_m$  at the base of the lithosphere. Thus, the integral of eq. 4.44 may be substituted by the integral from above. The water depth as a function of distance from the mid-oceanic ridge may thus be described with this model by:

$$w = \frac{2\rho_m \alpha (T_l - T_s)}{(\rho_m - \rho_w)} \sqrt{\frac{\kappa t}{\pi}} . \quad (4.45)$$

If we insert standard values for all the constants in this equation we get:

$$w \approx 5.91 \cdot 10^{-5} \sqrt{t} . \quad (4.46)$$

In words, the depth of the water is proportional to the square root of age of the oceanic lithosphere. Note that this water depth is only the *additional* water depth on top of the water depth *at* the mid-oceanic ridge (Fig. 4.19). We can convert this into water depth as a function of distance from the mid-oceanic ridge if we substitute age by the ratio of distance to rifting rate:  $x/u$ , (which is also age). Fig. 4.20 shows some water depth profiles calculated with this equation. The fantastic coincidence of these curves with bathymetric measurements in the oceans of the world to at least an age of the oceanic lithosphere of about 50 my, confirm the model (see also its correspondence with heat flow data in sect. 3.5.1).

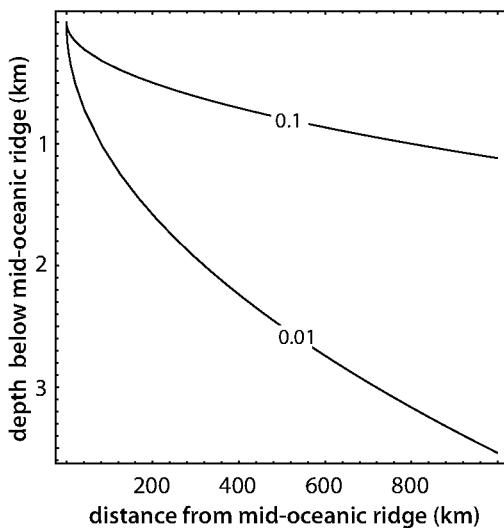
**Figure 4.20.** Profiles of water depth as a function of distance from the mid-oceanic ridge as calculated with eq. 4.45. The curves are shown for different rifting rates in  $\text{m y}^{-1}$ . Following constants were used:

$$\rho_m = 3200 \text{ kg m}^{-3},$$

$$\rho_w = 1000 \text{ kg m}^{-3},$$

$$\alpha = 3 \cdot 10^{-5} \text{ K}^{-1}, T_1 = 1280^\circ\text{C},$$

$$T_s = 0^\circ\text{C} \text{ and } \kappa = 10^{-6} \text{ m}^2 \text{s}^{-1}$$



#### 4.4.2 Flexural Isostasy

Most topographic features of our planet that are less than many hundreds of kilometers across are *not* completely in hydrostatic isostatic equilibrium. This includes whole mountain ranges like the European Alps (Karner and Watts 1983; Lyon-Caen and Molnar 1989) and can be measured gravimetrically: Gravimetry measures mass and in isostatic disequilibrium the total mass above the isostatic compensation depth is not everywhere the same. Thus, gravity anomalies may be interpreted in terms of the degree of isostatic disequilibrium. Isostatic disequilibria may form in response to a large range of processes. For example, a continental plate may be *actively* pushed downwards by the load of another plate, or it may be *actively* held up by mantle convection exerting an upwards force to the bottom of a plate. Topographic features that are created by non-isostatic processes are called: *dynamically supported*.

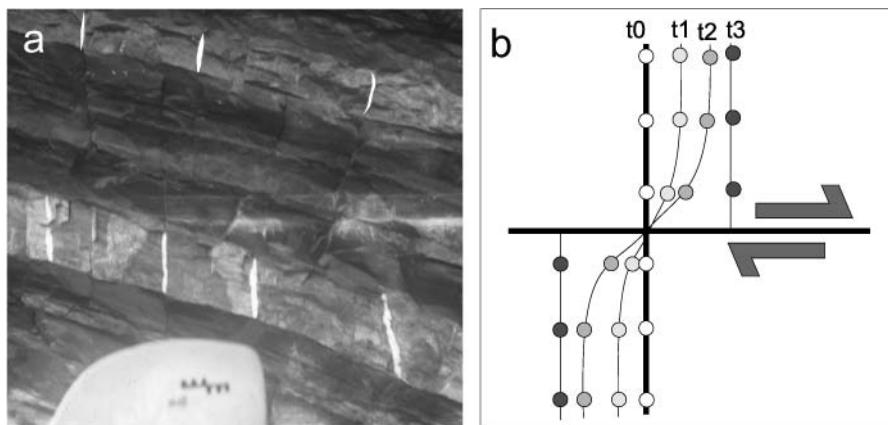
Flexural isostasy is a stress balance that also considers *horizontal* elastic stresses (Fig. 4.14b). Flexural isostasy is therefore at least a *two-dimensional* stress balance. It may be used to interpret surface topography in terms of both, hydrostatic balance and elastic flexure. In flexural isostasy, lithospheric plates are viewed as elastic plates that are bent by vertical loads. Interestingly, this model describes a large number of observations extremely well, although it is not at all trivial that the lithosphere should behave elastically at all. For example, we will show in sect. 5.2, that deformation of the lithosphere on geological time scales may be best described by brittle and ductile deformation mechanisms (sect. 5.1.2). Nevertheless, we observe a number of

large scale features that are well-described by elastic lithospheric models. In the following we summarize a few of these observations.

**Examples of Elastic Deformation.** Although it may not be intuitive that rocks can be elastic, there are quite a few observations that show us that they are! For example, regular spacing between joints and other cracks is a function of the elastic behavior of rocks and the continuous versus discontinuous displacement across seismically active structures is an elastic deformation that can be measured even with GPS measurements (Fig. 4.21). Elastic strains are of the order of about one per mil at the most.

- *Examples in oceanic lithosphere.* Oceanic lithosphere is rheologically stronger than continental lithosphere and is therefore little internally deformed. It has a very uniform thickness and a largely flat surface. As a consequence, plate scale elastic features that develop in response to vertical loads may spectacularly be seen without much disturbance by features created by other deformation mechanisms. The best known example for elastic deformation of the oceanic lithosphere are the valleys around *sea mounts*, for example around the Hawaii-Emperor chain. Sea mounts are volcanoes that have formed far from mid-oceanic ridges. They were created by *hot spots* that have their origin deep inside the mantle (Fig. 4.22). Thus, hot spot volcanoes that have formed on the surface of the oceanic lithosphere have no compensating root at the base of the plate. The volcano may be considered as an external load to a plate of more or less constant thickness that bends it downwards.

Another example of elastic deformation of oceanic lithosphere is the bending of the plates at subduction zones. The shape of trenches and the fore



**Figure 4.21.** Examples of elastic deformation of rocks. **a** Tension veins in competent layers between incompetent shale from Prince William Sound, Alaska (paddle for scale). The spacing is determined by the length scale that can support elastic deformation. Note that the spacing of veins is larger in the narrower layer at the top than in the wider layers at the bottom. **b** Schematic illustration of the elastic deformation around a fault. The marker lines are before deformation (t0), during elastic deformation (t1 and t2) and after rupture (t3).

bulge on the seaward side of the trench are also the consequence of elastic bending of the plate.

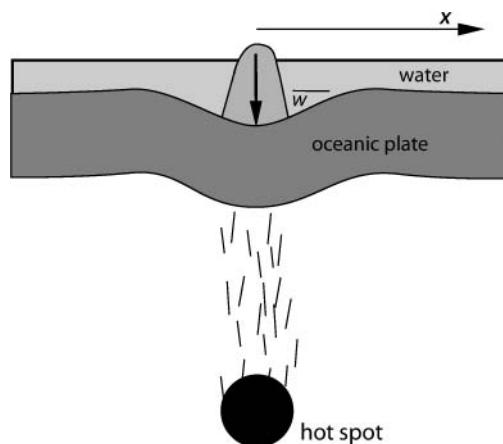
- *Examples in continental lithosphere.* The elastic bending of continental plates may be observed in the foreland of many collisional orogens, where molasse basins form as the consequence of the elastic deflection of the plate in response to the load of the mountain belt. One of the best known examples is the northern molasse of the European Alps. There, the European Plate is bent downwards under the load of the alpine mountain chain. The deepest point of the deflection is the valley of the river Donau. However, in collisional orogens the *external* load applied by the weight of the mountain belt is partly compensated by an *internal* load: the root of the mountain belt. Thus, the distribution of loads is not as clear or as easy to interpret as the examples of oceanic plates.

*Passive continental margins* also show often evidence for elastic bending of continental lithosphere (Fig. 4.26). The best known examples for this are the great escarpments along the coasts of southern Africa and Australia (Tucker and Slingerland 1994; Kooi and Beaumont 1994). There, the unloading of the plate that is caused by the asymmetric erosion of the continental margin is compensated by elastic updoming of the coastal foreland. The Australian Great Barrier Reef, for example, may be interpreted as an elastic fore bulge similar to those observed in the vicinity of subduction zones (Stüwe 1991).

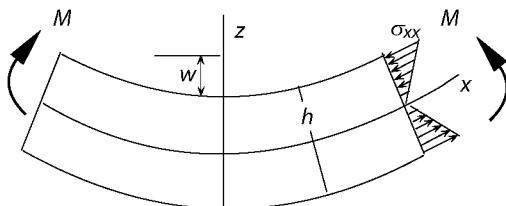
Interestingly, many geological structures that may be explained with the theory of elasticity – as for example the shape of the Australian escarpment – are preserved for many tens of millions of years. In sect. 5.3 we will show that on such a large time scale most stresses should be compensated by viscous deformation.

**The Flexure Equation.** In order to describe flexural isostasy quantitatively, we need to preempt some information on elastic deformation from

**Figure 4.22.** Flexure of oceanic lithosphere due to the loading of a sea mount



**Figure 4.23.** Bending of an ideal elastic plate in a simplified model view which is useful for the description of bending lithospheric plates



sect. 5.1.2. Elastic deformation describes an empirically derived *constitutive relationship* in which stress and strain are proportional to each other (eq. 5.24). The proportionality constant between stress and strain is called the modulus of elasticity or *Young's modulus*  $E$ . How much a plate bends under an applied stress depends on  $E$  and its compressibility, which is described by the *Poisson ratio*  $\nu$ .

Let us now consider the bending of a simple, ideal elastic plate like the one sketched in Fig. 4.23. We also neglect buoyancy forces for now. When integrating the horizontal normal stresses  $\sigma_{xx}$ , over the thickness of the elastic plate  $h$ , then it may be shown (or even intuitively seen) that the bending moment  $M$  is proportional to the curvature of the plate (s. Fig. 4.23):

$$M = -D \frac{d^2w}{dx^2} . \quad (4.47)$$

In this equation,  $w$  is the vertical deflection of the plate and the constant of proportionality  $D$  is called the *flexural rigidity* of the plate. The bending moment  $M$  is the integrated torques on both sides of the load. The derivation of eq. 4.47 and the following steps are explained in some detail by Turcotte (1979) and Turcotte and Schubert (1982) and will not be repeated here.

Eq. 4.47 is in itself not yet very useful. However, it may be coupled with a force balance equation that relates bending moments, the vertical load  $q$ , any applied horizontal forces  $F$  and the shear forces (s. Fig. 4.23) to each other (s. Turcotte and Schubert 1982; Ranalli 1987). When coupled with eq. 4.47, one arrives at the one-dimensional flexure equation:

$$D \frac{d^4w}{dx^4} = q_x - F \frac{d^2w}{dx^2} . \quad (4.48)$$

There,  $q_x$  is the vertical load as a function of horizontal distance  $x$  and has the units of force per area: stress. Thus, if the distribution of loads is known, this equation may be solved for either the deflection of the plate  $w$  or for its flexural rigidity  $D$  (in  $N \times m$ ). Usually, the deflection is well known from bathymetric or topographic observation and eq. 4.48 is used to derive the rigidity or “stiffness” of the plate. This flexural rigidity is a direct function of the elastic material properties of an ideal elastic plate of thickness  $h$  and is related to these by:

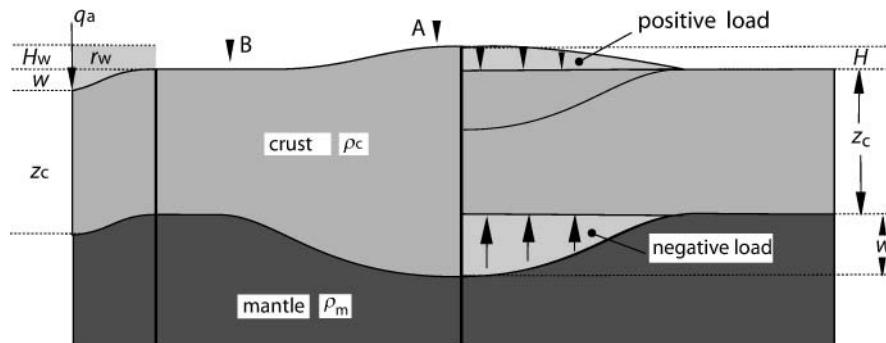
$$D = \frac{Eh^3}{12(1-\nu^2)} . \quad (4.49)$$

Thus, if the material constants  $E$  and  $\nu$  are known and the flexural rigidity of a plate was derived from modeling its shape using eq. 4.48, then this may be converted directly into an elastic thickness of the lithosphere using eq. 4.49. All descriptions of the bending of elastic plates are based on the integration of eq. 4.48, or its two-dimensional equivalent.

**Application to the Lithosphere.** Eq. 4.48 may be directly applied to describe flexural isostatic equilibrium, i.e. the elastic bending of lithospheric plates under external and internal loads. When we do this, we need to be aware of at least three important points:

1. We have to keep in mind that the flexure equation is based on completely different deformation mechanisms from those that we will discuss in sect. 5.2, where we consider the rheology of the lithosphere. In other words: eq. 4.48 is only a model that describes some field observations very well, but may be quite useless for the description of many other observations.

2. The flexural rigidity  $D$  must be interpreted correctly. Field observations tell us that the rigidity of lithospheric plates is of the order of  $D \approx 10^{23}$  Nm ( $\pm$  about one order of magnitude) and laboratory experiments show that the material constants are about  $E \approx 10^{11}$  Pa and  $\nu \approx 0.25$ . According to eq. 4.49 these parameters imply that the elastic thickness of the lithosphere  $h$  is only some tens of kilometers. Thus, the elastic thickness of the lithosphere is much thinner than the lithosphere according to thermal or mechanical definitions. The elastic thickness must be considered as the theoretical thickness of a plate with homogeneous elastic properties. Considering that the brittle strength of the upper crust as well as the ductile strength of the lower most lithosphere are likely to be very small, it is only the central part of the lithosphere that



**Figure 4.24.** Distribution of loads during the elastic bending of lithospheric plates. The shown example illustrates the displacement of the mantle by crustal material. At the left margin of the diagram the case of water covered oceanic lithosphere is illustrated (Problem 4.10). The central part of the figure illustrates the case of a continental plate. The right hand part of the figure shows how this plate can be split up into parts in order to consider the different forces acting on the plate.  $w$  is the plate deflection in both the left and the right hand example. Points  $A$  and  $B$  are marked for Problem 4.10

is dominated by elastic behavior on time scales that are short compared to the viscous response. However, it is important to note that the concept of an elastic thickness remains a theoretical one. Ranalli (1994) showed that the elastic thickness of the lithosphere is largely dependent on the depth of the 900 °C-isotherm (s. sect. 5.2.1, Fig. 5.16). As the rigidity is proportional to  $h^3$  (eq. 4.49), it might therefore be expected that  $D$  is indirectly proportional to the cube-root of the geothermal gradient (Molnar and Lyon-Caen 1989). This, however, is not the case. For example, the rigidity of the Indian and Adriatic plates varies by about *three* orders of magnitude but the difference in geothermal gradients is less than *one* order of magnitude. Thus, the elastic thickness is likely to be not only dependent on the 900 °C-isotherm.

3. The distribution of loads on the plate must be thought through. The load as a function of distance  $q_x$  as used in eq. 4.48 is the sum of a series of *internal* and *external* loads that act upwards and downwards onto a plate.

In order to clarify *which* different forces act on the plate, it is useful to divide the plate under consideration according to the scheme illustrated in the right hand part of Fig. 4.24. There, it may be seen that the downward force exerted by the mountain range on the plate is given by the vertical normal stress  $q_{\text{ext}} = \rho_c g H$ . This is the *external* or the *positive* load. This load is opposed by a buoyancy force in the region of the displaced mantle. This is the *internal* or *negative* load shown on Fig. 4.24 with the upwards arrows. This internal load has the magnitude  $q_{\text{int}} = (\rho_m - \rho_c) g w$ , where  $w$  is the deflection of the plate. The net load that is applied to the plate is therefore:

$$q(x) = q_{\text{ext}} - q_{\text{int}} = \rho_c g H(x) - (\rho_m - \rho_c) g w . \quad (4.50)$$

Note that the load is here already expressed as a function of horizontal distance  $x$ . If eq. 4.50 is inserted into eq. 4.48, this may be solved for  $w$  numerically or – for some simple boundary conditions – also analytically. When considering multi-layered plates, the same principle may be followed in deriving the net load on the plate (s. Problem 4.10).

**Applications to the Oceanic Lithosphere.** A series of elastic bending problem in the oceanic lithosphere may be well-described with eq. 4.48 if two simplifying assumptions are made:

- 1. We assume that there are no horizontal forces applied to the plate. Then, the entire last term of eq. 4.48 is zero. This is justified as many oceanic plates are not pushed from the side, but have open ends. Moreover, the forces that *do* act from the sides cause only a second order effect in plate deflection.
- 2. We assume that the vertical load is only applied at a single location at the end of the plate; i.e. there is no dependence of the load on  $x$ . When we consider subduction zones or island chains, this is also justified as both are largely linear features loading the plate.

Based on the second assumption, and assuming that the downwards deflected region is filled with water, eq. 4.50 simplifies to:

$$q = q_a - (\rho_m - \rho_w)gw \quad . \quad (4.51)$$

as illustrated on the very left hand edge of Fig. 4.24 ( $\rho_w$  is the water density). Eq. 4.48 simplifies to:

$$D \frac{d^4 w}{dx^4} = -(\rho_m - \rho_w)gw \quad . \quad (4.52)$$

Eq. 4.52 describes a range of geological features surprisingly well and has the great advantage that it may be integrated analytically for a range of geologically relevant boundary conditions. After integration, the constants  $D$ ,  $g$ ,  $\rho_m$  and  $\rho_c$  often occur in the following relationship:

$$\alpha = \left( \frac{4D}{g(\rho_m - \rho_w)} \right)^{1/4} \quad . \quad (4.53)$$

$\alpha$  is called the *flexure parameter* of the lithosphere (s. Problem 4.11).

- *Seamount chains.* The first example we want to discuss is that of a line-shaped load of islands on a continuous plate of constant thickness. For appropriately formulated boundary and initial conditions (e.g. the load applies only at  $x = 0$ , symmetry of the deflection so that  $dw/dx=0$  at  $x = 0$  and others) a solution of eq. 4.52 is:

$$w = w_0 e^{-x/\alpha} (\cos(x/\alpha) + \sin(x/\alpha)) \quad . \quad (4.54)$$

There,  $w_0$  is the maximum deflection of the plate directly underneath the load and  $w$  is normalized to this value (we can see from eq. 4.54 that  $w \rightarrow w_0$  for  $x \rightarrow 0$ ). Interestingly, the maximum deflection  $w_0$  is given by:

$$w_0 = \frac{q\alpha^3}{8D} \quad . \quad (4.55)$$

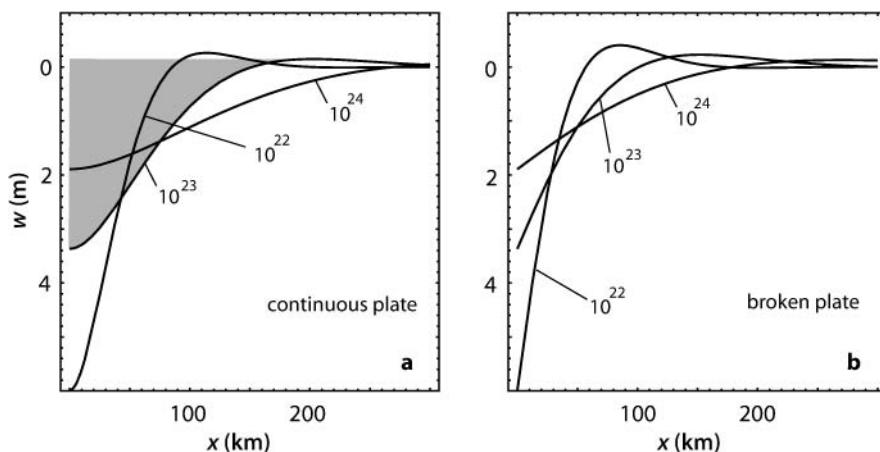
Eq. 4.54 is a good approximation for the description of the water depth around the Hawaii and Emperor Island chains (s. Fig. 4.25a). The equation is also historically important, as is was one of the first models used to estimate the elastic thickness of the lithosphere using the bathymetric surveys around Hawaii.

- *Trench morphology.* The second example that may be described with the approximation of eq. 4.52 is the shape of oceanic lithosphere near trenches. There, the loading of the subducting oceanic plate may be viewed as a line-loading by the margin of the upper plate. For this case, boundary conditions must be assumed that describe a broken half plate which is subjected to a load at its end. For appropriately formulated boundary conditions a solution of eq. 4.52 is:

$$w = w_0 e^{-x/\alpha} (\cos(x/\alpha)) \quad . \quad (4.56)$$

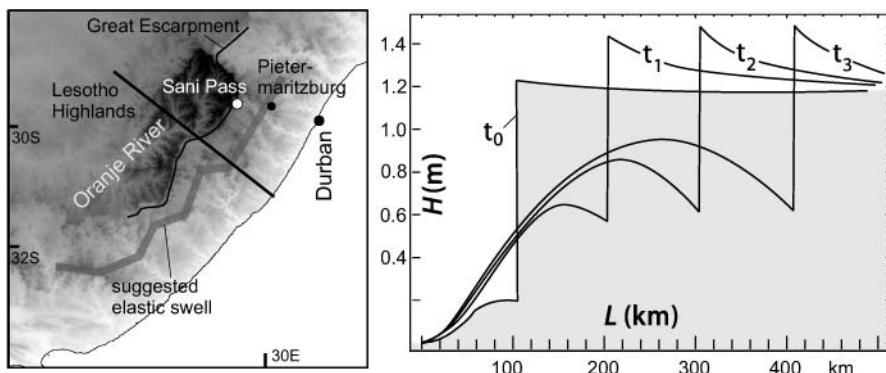
The shape of plates as described by eq. 4.56 is illustrated in Fig. 4.25b. Note how similar this solution is to eq. 3.100: They both describe sine-functions that decay exponentially with distance. A comparison of the curves shown on Fig. 4.25b with bathymetric measurements shows that most subduction zones are steeper near the trench than what is described by the curves at the left margin of Fig. 4.25b. It is interpreted that this indicates that subducted plates are not only loaded by the upper plate but that convection in the mantle wedge and other forces exert a additional torques on subducting plates. For other reasons of strange subduction angles see p. 22.

**Applications to the Continental Lithosphere.** Continental lithosphere deforms internally much easier than oceanic lithosphere by pervasive ductile mechanisms. Thus, elastic features are often not so clearly exposed and loads of mountain ranges and the like are distributed over large parts of the plates. As a consequence, continental lithosphere does not lend itself so easily to description with analytical solutions of eq. 4.48. Nevertheless, it should be said that the load of long mountain chains on homogeneous continental plates is analogous to the problem of long island chains on oceanic lithosphere. Thus, eq. 4.54 can – in principle – also be used to describe foreland basins, but care must be taken by accounting for sedimentary fill of foreland basins, compensating crustal roots etc (s. Turcotte and Schubert 1982). For example,  $\rho_w$  must be replaced by  $\rho_c$  in the formulation of the flexural parameter (s. eq. 4.50). However, much more progress has recently been made



**Figure 4.25.** Shape of elastically bent plates. **a** Continuous plate loaded only at  $x = 0$ : the left margin of the diagram (eq. 4.54). Only half of the plate is shown. **b** Broken plate, also loaded only at  $x = 0$  (eq. 4.56). The curves are labeled with the flexural rigidity of the plates in Nm. In **a**, the volume of the formed basin is shown by the shaded region for onw curve. However, the basin volume is for all rigidities the same! Note the elastic bulges that occur outwards of the basins before the plates return to their normal position in the far field

in foreland basin modeling by using numerical approximations, where more realistic scenarios may be described (Garcia-Castellanos et al. 1997).



**Figure 4.26.** Elastic flexure at passive continental margins. The map shows the Great Escarpment of southern Africa. The straight line shows the profile which is schematically modeled at right. The model at right is a cross section through an idealized passive margin showing surface elevation  $H$  as a function of distance from the continental shelf  $L$  at four different time steps  $t_0$  to  $t_3$  during the successive erosional retreat of a 1 000 m high escarpment. The land region of the cross section is shaded for time step  $t_0$ . Topographic profiles similar those at  $t_1$  to  $t_3$  can be found in southern Africa and eastern Australia (after Stüwe 1991)

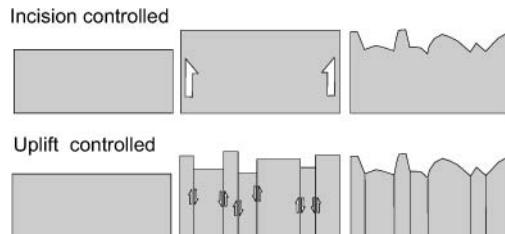
The topography of *passive continental margins* is probably the example in the continental lithosphere that is most obviously described by elastic flexure (Fig. 4.26). While the geometry of passive margins may be characterized with a very simple geometry, there is no analytical solution of the flexure equation to describe it and models in the literature rely on numerical approximations. Passive margins feature a range of interesting morphological relationships between the position of the drainage divide, the shape and direction of the drainage networks, the position of a characteristic great escarpment and the slope of the range (s. p. 184). Many of these relationships indicate that the relief of the plate margin is characterized by elastic bulges in the foreland of the escarpment that are interpreted to have formed in response to the unloading of the plate margin by erosional retreat of the escarpment. The Great Barrier Reef in Australia has been explained as such a forebulge (e. g. Stüwe 1991; Gilchrist et al. 1994; Tucker and Slingerland 1994). Other curious features of passive continental margins, for example the inland drainage of rivers, are consistent with such interpretations.

## 4.5 Geomorphology

The interpretation of geomorphological features in terms of an integrated geodynamic understanding of plate tectonic processes has become an impor-

tant topic in the earth sciences. This new direction has become known by the name “tectonic geomorphology” or “neotectonics” and the appearance of a range of recent excellent textbooks testifies of the topical nature of this new field (e.g. Burbank and Anderson 2001; Summerfield 1991; Keller and Pinter 1996; Schumm et al. 2000; Kirkby 1994; Julien 1998).

**Figure 4.27.** Schematic cartoon illustrating the development of uplift versus incision controlled topography. From left to right the three little blocks in each row show different time steps during the landscape evolution



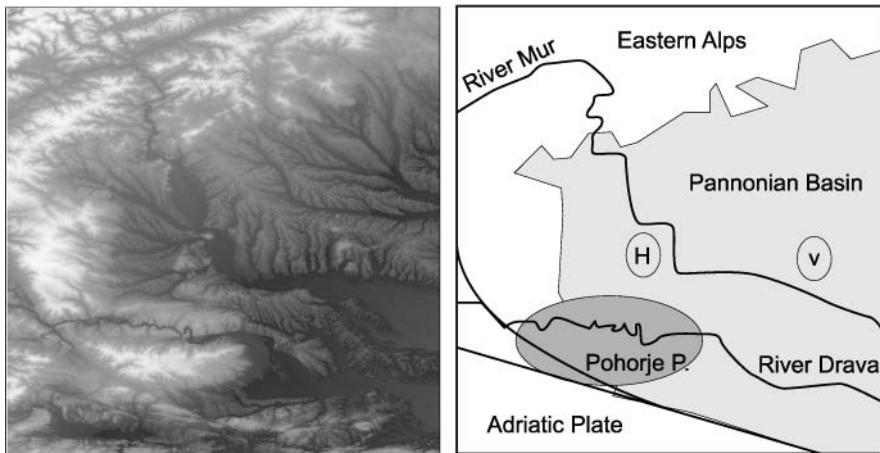
For the description of geomorphic processes it is often useful to discriminate between:

- geomorphic shaping by tectonic (“endogenic”) processes, and
- geomorphic shaping by erosion and sedimentation (“exogenic”) processes.

We could define these two basic end members as “uplift controlled” and as “incision controlled” topography, respectively (Fig. 4.27). In many tectonically active regions both processes interact (Fig. 4.28) and an understanding of the relative contributions of these two processes is crucial for an integrated understanding of the evolution of such regions. Most of both process groups may be described using some very basic principles of which some are introduced on the following pages. From the viewpoint of a Eulerian observer, both process groups may be described as a material transport into or out of the system. *Tectonic* (“endogenic”) processes move material by faulting, by uplift or by subsidence. *Erosion* and *sedimentation* (“exogenic”) processes move material by removing or depositing material. In many cases the material transport may be described with the same principles as the transport of heat. That is, by:

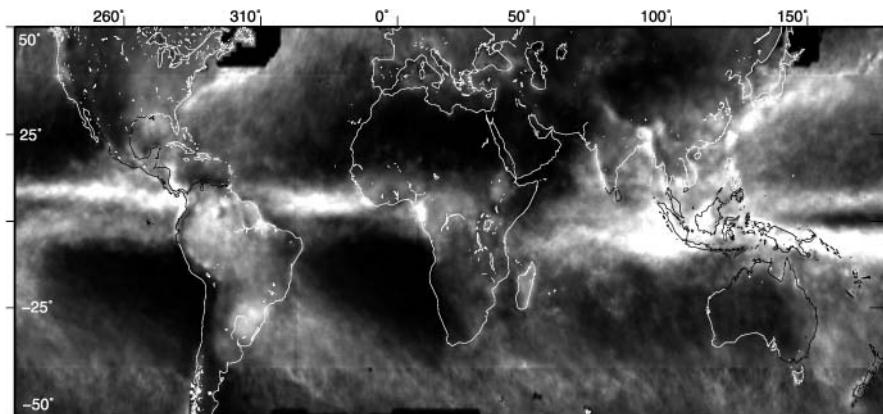
- diffusion,
- advection or
- production

of material in the system. Accordingly, many geomorphic processes may be described with the same principles discussed in sect. 3.1, 3.2 and 3.3, by considering the transport of *mass* instead of *energy*. However, there are also processes that are unique to landscape formation, for example, the hydrological processes in drainage networks or the threshold mechanisms governing landslides. Before we discuss the description of individual processes in some detail, the following paragraphs illustrate the difference between *endogenic* and *exogenic* landscape formation.



**Figure 4.28.** Digital elevation model and tectonic cartoon of a 120 km by 120 km section at the transition zone between the Alpine orogen and the Pannonian Basin (centered around 47°N and 15°E). The region shows a series of spectacular neotectonic features, for example the incision of the river Drava which is clearly antecedent with respect to the uplift of the dome made by the Pohorje Pluton. This pluton is the eastern most of a series of plutons within the suture zone between the Adriatic and European plates. Also note that some of the north south striking ridges within the basin show a clear asymmetry with steep west flanks and shallow east flanks indicating that the basin inversion was associated with morphological shaping by block tilting. The topographic highs at H and V are tectonic horsts and active volcanoes, respectively, both evidencing the high activity of the region

- *Geomorphic shaping by tectonic processes.* Relief development by tectonic processes is generally described with mechanical and kinematic models and – while undoubtedly an integral process of landscape formation – are usually not the subject of the modeling of geomorphic processes as such. Rather, tectonic processes form *boundary conditions* to geomorphic modeling, which generally concentrates on the description of the exogenic processes. For example, tectonic processes that are rapid compared to the subsequent geomorphic shaping can be assumed as starting condition, not unlike the assumption we discussed for the description of contact metamorphism on p. 81. Alternatively, slow tectonic uplift over long time scales can be modeled as an uplift function that interacts with the geomorphic shaping.
- *Geomorphic shaping by erosion and sedimentation.* Erosion and sedimentation processes are governed by a large range of different physical processes including soil creep, solution, rain splash, chemical- and aeolian weathering, down- and sideways cutting of drainages, debris flows, as well as discontinuous processes like landslides and many others (e. g. Carson and Kirkby 1972). In order to describe these processes with simple models it is useful to summarize them into three groups:



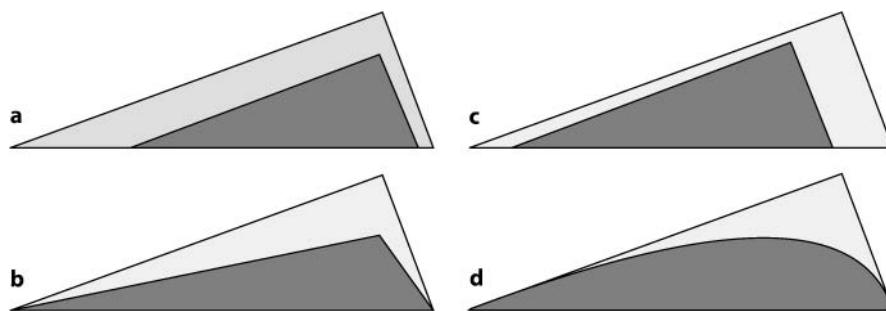
**Figure 4.29.** Map of global rain fall. The plotted range is from zero (black) to 3 m (white) annual mean precipitation. Data from <http://daac.gsfc.nasa.gov/data/datapool/TRMM>. Data processed and plotted by Robl, pers. comm. (2006)

- short range continuous transport,
- long range continuous transport,
- discontinuous processes.

*Short range transport* describes the local redistribution of mass on a hill slope scale and is discussed starting on p. 184. *Long range transport* describes the erosion and sedimentation processes in rivers (p. 193) and *discontinuous processes* are processes that are episodic in time, like landslides (p. 200). Fortunately, geological time scales are long enough so that many local *discontinuous* processes may be described by *continuous* models. However, as always, the chosen model description depends critically on the question being asked. For example, when modeling the first order morphological features of orogens as a whole, the subdivision discussed above is unnecessary. We begin therefore with a summary of geomorphic modeling on the largest scale.

#### 4.5.1 Erosion Models on Orogenic Scale

Erosion models are not only important for geomorphologists. Many questions typically asked by tectonicists and metamorphic petrologists also require the consideration of erosion in some simplified way. For example, when we discuss the exhumation mechanism of a metamorphic rock we may want to know if erosion or extension drove the exhumation. The studies of England (1981), Summerfield and Hutton (1994) or Harrison (1994) show that at least in some orogens, erosion is the principle exhumation mechanism. Thus, it is crucial to consider erosion processes as an integral part of geodynamic models (e. g. sect. 6.2.1). In many models where the description of the landscape evolution is not the primary focus of the study, the process of surface erosion



**Figure 4.30.** The influence of four different erosion models on the shaping of an asymmetric mountain belt. The light shaded area shows the mountain before the onset of erosion. The dark shaded region shows the shape of the mountain after some time. **a** constant erosion rate, **b** erosion rate proportional to elevation, **c** erosion rate proportional to slope, **d** erosion rate proportional to surface curvature. In all four erosion models, the erosion rate is the rate of removal of vertical section of the topography, measured in  $\text{m s}^{-1}$ . Note that only in models **a** and **b** the highest point of the topography remains laterally fixed during erosion

is simplified as much as possible. Then, one of the following erosion models is often used (Fig. 4.30):

- a** Erosion rate is constant through time and space,
- b** erosion rate is a function of elevation,
- c** erosion rate is a function of slope,

There are also more refined models in which erosion rate is assumed to be proportional to terrain curvature (diffusion) and those that model fluvial incision. These latter descriptions are important for landscape evolution models and are discussed in some detail in sect. 4.5.2 and sect. 4.5.3, respectively. Which of the three more simple models listed above should be used for a given problem to describe erosion on the most simple level is discussed in the three next paragraphs.

**Constant Erosion Rate.** Assuming that the erosion rate during a given orogenic process is constant through space and time is the most dramatic thinkable simplification of real erosion processes. However, let us recall that a good model must find the right balance between accurate description and simplicity (p. 4). For example, erosion models assuming constant erosion rate have been very successful to explain the causes of the clockwise shape of metamorphic *PT* paths during regional metamorphism (sect. 6.2.1). Thus, for many purposes and problems of metamorphic petrologists, this model is the best. The model can be formulated as:

$$v_{\text{er}} = -\frac{dH}{dt} = \text{constant} \quad . \quad (4.57)$$

where  $v_{\text{er}}$  is the erosion rate,  $H$  is elevation and  $t$  is time. Note that the sign must be negative for erosion to *decrease* surface elevation (this is different

from other vertical reference frames we have used elsewhere in this book). Remember that the erosion rate can only be set equal to the rate with which the surface goes downwards if the uplift rate of rocks is  $v_{ro} = 0$  (s. eq. 4.10).

**Erosion Rate Proportional to Elevation.** Assuming that the erosion rate of a mountain belt is proportional to its elevation is the next closer approximation of nature. In fact, this model describes many aspects of real erosion processes very well and we have used it extensively in the section around page 157. For example, we observe that high mountain ranges like the European Alps erode much quicker than low lying hills in the German or Italian foreland. However, there are also examples where the opposite is true: the Tibetan Plateau is 5 000 m high, but erosion rate is practically zero. Nevertheless, this erosion model has found much application in the literature. In the most simple case, the proportionality between elevation and erosion rate is linear and may be described by the relationship:

$$v_{er} = -\frac{dH}{dt} = -\frac{H}{t_E} . \quad (4.58)$$

Note again that the sign convention used here is consistent with the convention used in the last paragraph, but opposite to that used in eq. 4.16, where we used a different reference frame. The erosion parameter  $t_E$  (in units of time) describes how long it takes to erode a mountain of the elevation  $H$ . The model implies that the erosion rate drops immediately as the first increment of erosion has decreased the elevation of a mountain. If eq. 4.58 is integrated (e.g. using the principle explained on p. 416), the model describes an exponential decrease of the elevation through time:

$$H = H_0 \times e^{(-\frac{t}{t_E})} , \quad (4.59)$$

where  $H_0$  is the original elevation at the onset of erosion.

**Erosion Rate Proportional to Slope.** A proportionality between erosion rate and slope may be described by:

$$v_{er} = -\frac{dH}{dt} = -u \frac{dH}{dx} . \quad (4.60)$$

There,  $x$  is a horizontal spatial coordinate and  $dH/dx$  is the topographic gradient: the slope. The proportionality constant  $u$  is the horizontal rate of displacement of the slope. As with eq. 4.57, it is important to note that erosion rate only corresponds to the rate of elevation change, if all other uplift or subsidence processes are zero. Equation 4.60 should remind us of eq. 3.40. Both are one-dimensional transport equations, which may be solved with the methods discussed in sect. 3.3 (s. Fig. 3.13 or Fig. B.7). The model is a good description for the evolution of many landforms, for example the motion of sand dunes. The model has also been applied to describe the geomorphic evolution of passive continental margins.

- *Advection at passive margins.* Markedly asymmetric mountain ranges are developed at several locations around the globe along the passive margins of continents. The best developed examples are the *Great Dividing Range* along the east coast of Australia and the coastal ranges along much of the southern African continent (Fig. 4.26). In both these examples, the inland side of the mountain range has a very small slope angle, while the coastal side is characterized by much steeper slope, often referred to as *Great Escarpment* (Ollier 1985). Because of this marked asymmetry, King (1953) concluded that erosion occurs largely along the escarpment which causes that the shape of the range remains largely self similar through time and that the great escarpment retreats inland (s. Fig. 4.26). In southern Africa this process is aided by the fact that the escarpment is often made up by the very resistant Karoo-basalts. This erosion model and the implied morphological evolution of passive continental margins has found much interest in the past decade. The simple model of King (1953) is now largely superseded (e. g. Stüwe 1991; Kooi and Beaumont 1994; Tucker and Slingerland 1994), but it still must be acknowledged that it describes a number of features quite well.

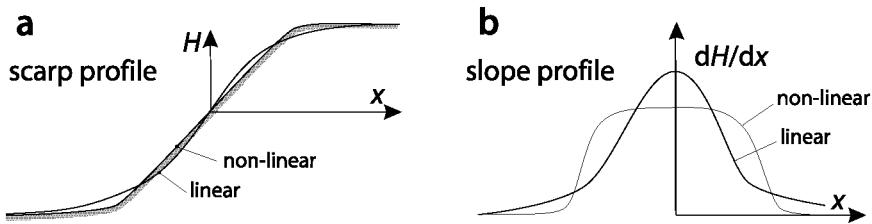
#### 4.5.2 Short Range Transport

Modeling erosion rate as a function of terrain curvature is the most refined description of erosion on short length scales, in particular on hill slopes. It is important enough so that we will spend some pages on it. This mass transfer process is also called diffusion and we have already met and discussed this process in much detail on p. 55. We will recall from there, that diffusion describes a proportionality between *temporal* change and *spatial* curvature (eqs. 3.1, 3.4). That this model can also be applied to describe the geomorphic shaping of landforms was established by Culling (1960), Ahnert (1970) or Andrews and Bucknam (1987). Many observations in nature lend themselves to description with this proportionality: We can observe that ragged, pointy mountains erode much quicker than flat plateaus (even if they lie very high); that sharp escarpments erode quicker than smoothly curved hills and many more. Among diffusion models we discern:

- linear diffusion,
- non-linear diffusion.

Before we describe and apply diffusion models on the next pages we want to recall that the diffusion model is only a simple description summarizing a range of physical processes (including even discontinuous processes like landslides if they may be temporally averaged).

- *Linear diffusion.* Linear diffusion of mass is completely analogous to the logic discussed on p. 52. It is based on the assumption that the rate of down slope transport of mass (described by mass flux  $q$ ) is proportional to the hill slope:



**Figure 4.31.** Comparison between linear and non-linear diffusion in the evolution of a degrading scarp. Note that in the non-linear case slope curvatures are more localized than for the linear case

$$q = -D \frac{dH}{dx} . \quad (4.61)$$

where  $dH/dx$  is the topographic gradient, (i.e. the slope) and the erosional diffusivity  $D$  corresponds to  $\kappa$  in the theory of heat conduction and has the units of  $\text{m}^2 \text{s}^{-1}$ . This equation is directly analogous to Fourier's first law (eq. 3.1). Note that the mass flux  $q$  has the units of  $\text{m}^2 \text{s}^{-1}$  which may be interpreted as the volumetric flow normalized to the width of a profile. In other parts of this chapter we will encounter fluxes that are normalized to the channel cross section and have therefore the units of  $\text{m s}^{-1}$  (e.g. p. 198). The diffusivity  $D$  may be interpreted as the product of horizontal rate of mass transport  $v$  and thickness of an erodable near surface layer  $h_s$  (Beaumont et al. 1992; Carson and Kirkby 1972):

$$D = vh_s . \quad (4.62)$$

From this equation, we can recognize the origin of the variability of  $D$ :  $h_s$  is dependent on rock type, but  $v$  is not. As in the theory of heat transfer, the flux equation eq. 4.61 may be combined with a one-dimensional mass balance of the form:

$$\frac{\partial H}{\partial t} = -\frac{\partial q}{\partial x} . \quad (4.63)$$

This equation is analogous to eq. 3.2 and is not derived in detail here. By inserting eq. 4.61 in eq. 4.63 we obtain the mass diffusion equation in one dimension:

$$\frac{\partial H}{\partial t} = D \frac{\partial^2 H}{\partial x^2} . \quad (4.64)$$

Eq. 4.64 is analogous to eq. 3.6. If we insert concentration instead of elevation into eq. 4.64 then this equation may also be used to model the distribution of elements in minerals (s. eq. 7.3). In mineralogy and geochemistry, the equations governing diffusion are known as *Fick's laws*.

- *General conservation of mass.* The diffusion equation describes conservation of mass on a local scale. However, just as we have combined conduction, advection and production of heat in a complete thermal energy balance in eq. 3.52, we can write a more general form of a mass conservation equation in two dimensions as:

$$\frac{\partial H}{\partial t} = D \left( \frac{\partial^2 H}{\partial x^2} + \frac{\partial^2 H}{\partial y^2} \right) + u_x \frac{\partial H}{\partial x} + u_y \frac{\partial H}{\partial y} + v . \quad (4.65)$$

This equation describes the change in elevation at a given point of the topography as the sum of diffusion (first term on right hand side), sideways advection of a topographic profile (e.g. dune motion or retreat of an erosional escarpment) at the horizontal rates  $u_x$  and  $u_y$  (second and third term) and material “production”, i.e. upwards advection at the rate  $v$ . This “upwards advection rate”  $v$  may be seen as the sum of uplift and sediment deposition and is particularly important for descriptions in Lagrangian reference frames.

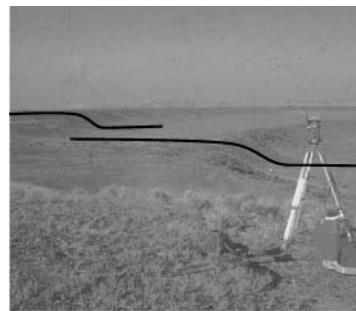
Note that this equation describes a two-dimensional model, although all three spatial dimensions  $x$ ,  $y$  and  $H$  occur in it and although the model can be represented as a three-dimensionally (e.g. as a plaster model). However,  $H$  is evaluated only on the basis of two model variables and landscape models are therefore in general two-dimensional models (s. p. 7).

- *Non-linear diffusion.* In linear diffusion it is assumed that the mass transport is *directly* proportional to slope. However, this need not be so. Non-linear diffusion describes in effect a down-slope change of the erosional diffusivity  $D$  (Pierce and Coleman 1986). It may be intuitively seen that this may be a good description of nature: On many slopes, the material is more and more loose the steeper the slope gets and erosion may therefore be more rapid there. A more general formulation of diffusion that accounts for this may therefore look like:

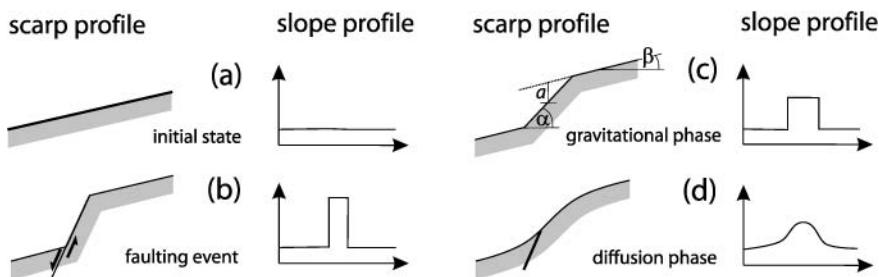
$$\frac{\partial H}{\partial t} = D \frac{\partial}{\partial x} \left( \frac{\partial H}{\partial x} \right)^n . \quad (4.66)$$

If  $n = 3$  this equation describes what is called *cubic diffusion*. (For an explanation of the definition of a general non-linear differential equation see p. 396). If the rate of down slope transport is related to some power of slope, then this will have the consequence that more even slope profiles develop with sharper edges as illustrated in Fig. 4.31. Non-linear diffusion was discussed by Newman (1983) and established for the modeling of scarps by Andrews and Bucknam (1987) and most recently discussed by Roering et al. (2001). Hanks and Andrews (1989) suggested a “linear plus cubic” diffusion model as a good fit to some field data. Modern discussion of non-linear versus linear diffusion in landform processes are found by Avouac (1993) or Avouac and Peltzer (1993).

**Figure 4.32.** Double terrace riser in the Tien Shan. The risers are very much degraded as they are high above the present river bed. More details on these particular risers are described by Avouac (1993) (photo by Avouac)



**Scarp Degradation.** As a first example for the application of the short range diffusion model, let us discuss the morphological dating of fault scarps. Scarps form by earthquakes as a practically instantaneously created steps in the surface (Fig. 4.33). Their subsequent morphological decay may be described as a mass diffusion process and lends itself to a description with similar solutions as we discussed for the diffusion of heat around step-shaped temperature perturbations in sect. 3.6 (Fig. 3.30, eq. 3.84). As we did there, boundary conditions can often be assumed at infinity because scarps are often small compared to the extend of the slope they occur on. Using  $H$  for elevation,  $x$  for a spatial coordinate system normal to the fault scarp and with its origin in the center of the scarp (as shown in Fig. 4.31), the boundary conditions may be formulated as  $H = a$  at  $x = \infty$  for  $t \geq 0$  and  $H = -a$  at  $x = -\infty$  for  $t \geq 0$ . These boundary conditions describe a scarp displacing a flat surface by the height  $2a$ . For these conditions, eq. 4.64 may be integrated to give:



**Figure 4.33.** Stages in the evolution of a fault scarp (after Avouac 1993). Note the similarity of the slope profile during the diffusion process from **c** to **d** to the cooling of intrusions shown in Fig. 3.32. Parameters labeled in **c** are those used in eq. 4.68

$$H = a \times \operatorname{erf} \left( \frac{x}{\sqrt{4Dt}} \right) , \quad (4.67)$$

which is completely analogous to eq. 3.84, except that we shifted the coordinate system to the middle of the scarp. However, field observations show that many fault scarps are not vertical, in part because they did not form on vertical faults and in part because they sag gravitationally after a faulting event and assume a stable angle of repose (angle  $\alpha$  on Fig. 4.33c) only some time thereafter. Fortunately, this “gravitational phase” is usually short compared to the subsequent degradation and the decay of the landform may still be described with the diffusion equation. However, an analytical solution of eq. 4.64 for the initial geometry shown in Fig. 4.33 is unlike more complicated than eq. 4.67. It is:

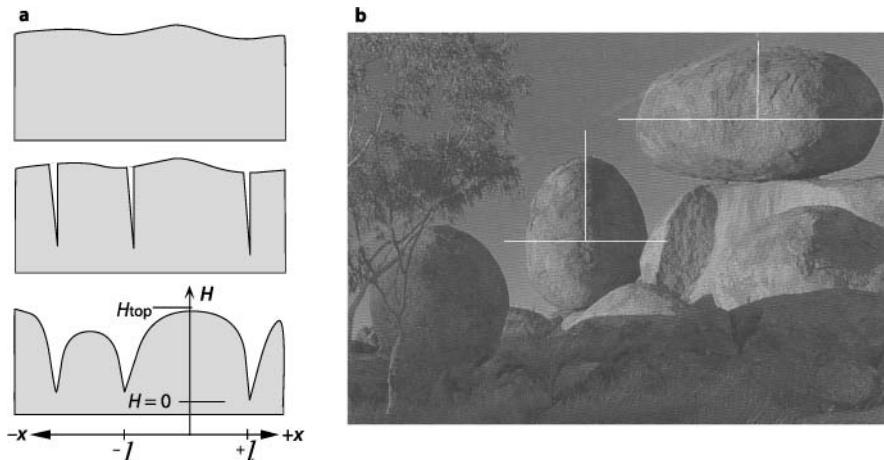
$$H = \left( \frac{c}{a} \sqrt{\frac{Dt}{\pi}} \right) \times \left( \exp \left( -\frac{(x+c)^2}{4Dt} \right) - \exp \left( -\frac{(x-c)^2}{4Dt} \right) \right) + \left( \frac{c}{2a} \right) \times \left( (x+c) \operatorname{erf} \left( \frac{x+c}{\sqrt{4Dt}} \right) - (x-c) \operatorname{erf} \left( \frac{x-c}{\sqrt{4Dt}} \right) \right) + bx , \quad (4.68)$$

where  $c = a/(\alpha - b)$ , the half height of the scarp is  $a$  and  $b = \tan(\beta)$ , as shown in Fig. 4.33c.

Instead of fitting this equation to measured scarps, it is often possible date the decay of scarps by using a single value: the maximum scarp slope  $\tan(\theta)$ . That may be found by taking the derivative of eq. 4.68 with respect to  $x$  and evaluating this at  $x = 0$ . This gives much more simple expressions which have been used to date scarps by a variety of authors (e.g. Avouac 1993; Avouac and Peltzer 1993). The slope distribution shown in Fig. 4.31 and 4.33 shows that diffusive decay of scarps is characterized by a Gaussian distribution of slope. Diffusion of landforms may therefore also be described by Gaussian smoothing (Avouac 1993).

**Mass Diffusion with Fixed Boundary Conditions.** While the degradation of scarps may be described with boundary conditions fixed at infinity, most geomorphological diffusion problems are characterized by spatially fixed boundary conditions. For example, the rounding of hill slopes is usually spatially confined to a hill between two drainages from where material is efficiently transported out of the system. Similarly, the weathering and erosion of granitic boulders occurs between joints from where material is removed out of the system. Below we discuss two examples for which there are analytical solutions of eq. 4.64.

- *Spheroidal weathering.* Mechanically isotropic rocks (for example granite) often weather in typical rounded shapes (Fig. 4.34). This *spheroidal weathering* occurs because individual blocks are separated relatively quickly by jointing, but the rounding off by erosion takes a much longer time scale. The weathering process is most effective on surfaces that have a high spatial curvature: it is a diffusion process and may be described with eq. 4.64. In order to formulate appropriate boundary conditions for the integration of eq. 4.64 it is useful to choose a coordinate system with an origin at the center of the



**Figure 4.34.** Illustration of spheroidal weathering. **a** Schematic illustration. The top diagram shows the starting geometry, the middle diagram shows how jointing of the rock surface occurs rapidly due to climatic influence and occurs rapidly in comparison with the subsequent erosion. The bottom diagram shows the typically rounded shapes that form during the subsequent weathering and the coordinate system is shown that is used for the formulation of eq. 4.69. **b** shows an example from the Devils Marbles, Central Australia. The coordinate systems on the photograph are drawn for Problem 4.14

block (Fig. 4.34). If the diameter of the block is  $2l$ , then the joints on either sides of the block lie at  $x = l$  and  $x = -l$ . Initial and boundary conditions may be formulated as:

- Initial condition:  $H = H_{\text{top}}$  in the region  $-l < x < l$  at time  $t = 0$ .
- Boundary conditions:  $H = 0$  at  $x = l$  and  $x = -l$  at  $t > 0$ .

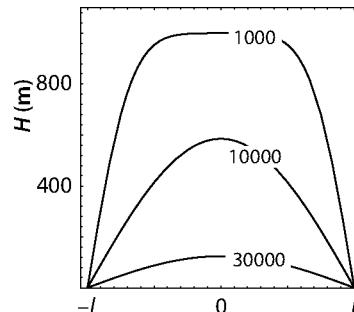
$H_{\text{top}}$  is the height of the block as shown in Fig. 4.34a. A solution of eq. 4.64 subject to these boundary conditions requires the use of Fourier series (for reasons explained on p. 413 and p. 120) and therefore contains infinite summations and trigonometric functions. It is:

$$H = \frac{4H_{\text{top}}}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp(-D(2n+1)^2 \pi^2 t / 4l^2) \times \cos\left(\frac{(2n+1)\pi x}{2l}\right) . \quad (4.69)$$

The shape of weathering profiles as a function of time calculated with this equation are shown in Fig. 4.35. The mismatch between the curves shown in this figure and the photograph in Fig. 4.34b is predominantly because eq. 4.69 is one-dimensionally, while the boulders in the photograph also round off in

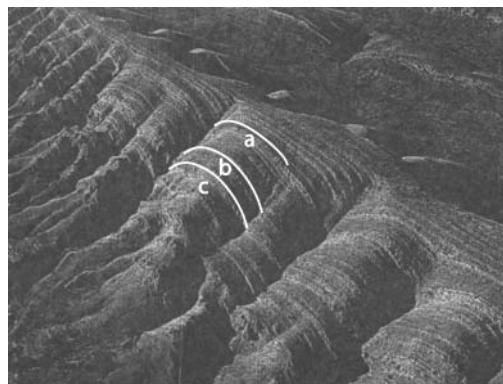
the plane normal to the photograph, plus from below. Fortunately, a two-dimensional equivalent to eq. 4.69 is easily formulated as the product of two solutions in orthogonal directions, as we discussed on p. 119.

**Figure 4.35.** The shape of an initially rectangular ridge that erodes by mass diffusion processes after 1 000, 10 000 and 30 000 years. Streams transporting eroded material out of the system must be imagined in the bottom left and right corners to run normal to this page. Calculated with eq. 4.69 using  $l = 1 \text{ km}$ ,  $H_{\text{top}} = 1 \text{ km}$  and  $D = 10^{-6} \text{ m}^2 \text{ s}^{-1}$



- *Hill slope profiles.* Eq. 4.69 and its graph on Fig. 4.35 may not only be used to date spheroidally weathered granite boulders. It can also be used to model the shape of ridge profiles between parallel drainages (Fig. 4.36). However, when modeling hill slopes between drainages we need to be careful, as the boundary conditions we have formulated in the last paragraph describe stationary (non-incising) river beds. Thus, the model applies only to rivers where the material transport in the stream equals the amount of material fed into them by the ridge. If the streams incise into the landscape simultaneously with the diffusion on the hill slope, we must formulate the problem with variable boundary conditions.

**Figure 4.36.** Example of hill slope formation due to mass diffusion with incising boundaries. Western Mac Donnell Range, Central Australia. The length of the visible part of the principle ridge is about 500 m. Different stages of hill slope development may be seen. Near the principle drainage divide, (where head waters in the gullies have incised last), profiles are similar to the youngest profile of Fig. 4.37 (profile *a*), while hill slopes have adapted a steady state by *b* which is maintained in *c* before the slope disintegrated into smaller landforms



**Mass Diffusion with Variable Boundary Conditions.** If the incision rate of drainages is comparable to the rate of diffusion, then the incision itself is part of the hill slope shaping process and the boundary conditions of the previous sections may not be used. The description of simultaneous diffusion and incision at the model boundaries depends on the reference frame used. Two possibilities offer themselves.

- *Eulerian description.* For an observer fixed to an external reference frame, the incision of drainages may be formulated as follows:

- Initial condition:  $H = 0$  in the region  $-l \leq x \leq l$  at time  $t = 0$ .
- Boundary conditions:  $H = v_{\text{fl}}t$  for  $x = l$  and  $x = -l$  at time  $t > 0$ .

In this formulation we used  $H = 0$  at the surface before the onset of erosion as the origin of a vertical axis going positively downwards and a horizontal coordinate system with the origin half way between two parallel streams that are a distance  $2l$  apart.  $v_{\text{fl}}$  is the vertical incision rate of the rivers. The elevation  $H$  increases linearly at the model boundaries with time. Note that we changed the direction of the vertical axis from the last section, for clearer illustration. Solving eq. 4.64 subject to these temporally varying boundary conditions is difficult and will not discuss here. However, an example of a result from a corresponding solution is shown in Fig. 4.37a.

- *Lagrangian description.* From the observation point in the river bed it appears as if the landscape uplifts between the river beds (Fig. 4.37b). We can describe this using the uplift rate  $v_{\text{ro}}$  for which it is true that:  $v_{\text{ro}} = -v_{\text{fl}}$ . The uplift is a kind of “material production” between the model boundaries and we can not use eq. 4.64 for a description in a Lagrangian reference frame. The equation we need to solve includes a term describing the rock uplift. It is:

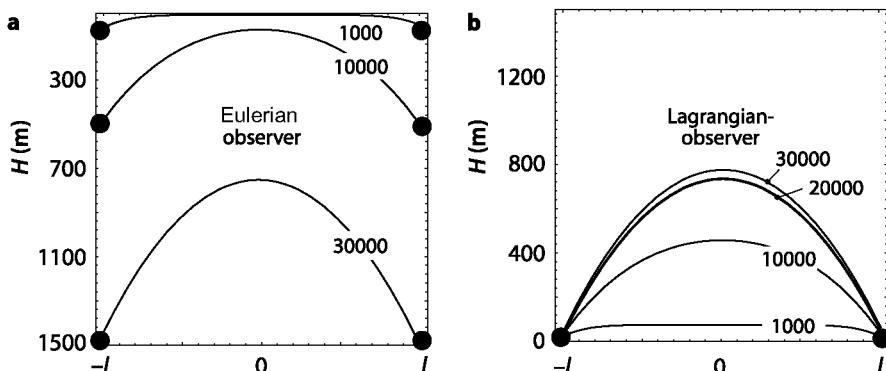
$$\frac{\partial H}{\partial t} = D \frac{\partial^2 H}{\partial x^2} + v_{\text{ro}} \quad . \quad (4.70)$$

which is equivalent to the heat conduction equation with a heat production term (see eq. 3.23, also see eq. 4.65). The initial and boundary conditions that apply for the solution of this equation are:

- Initial condition:  $H = 0$  in the region  $-l \leq x \leq l$  at the time  $t = 0$ .
- Boundary condition:  $H = 0$  at  $x = l$  and  $x = -l$  at all times  $t > 0$ .

For these boundary conditions a solution of eq. 4.70 is given by Crank (1975):

$$H = v_{\text{ro}} \left( \frac{l^2 - x^2}{2D} \right) - \frac{16l^2 v_{\text{ro}}}{D \pi^3} \times \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^3} \exp \left( -D(2n+1)^2 \pi^2 t / 4l^2 \right) \cos \left( \frac{(2n+1)\pi x}{2l} \right) \quad . \quad (4.71)$$



**Figure 4.37.** Landscape evolution of a ridge subject to mass diffusion processes between vertically incising drainages. **a** Within in a Eulerian reference frame (spatially fixed) and **b** in a Lagrangian reference frame (fixed to the incising drainages). The shape of the ridge is shown after 1 000, 10 000 and 30 000 years and was calculated using eq. 4.71,  $l = 1$  km,  $v_{ro} = 0.05$  m per year and  $D = 10^{-6}$  m $^2$  s $^{-1}$

A model example of a hill slope profile calculated with this solution is shown in Fig. 4.37b. Note that the landscape profiles in Fig. 4.37a and b have the same shape at the same times. They differ only in the reference frame and thus the profiles occur at different absolute elevations.

On Fig. 4.37b we can see that the landscape evolution described here approaches a steady state shape after about 20 000 years. This is what is called an *erosional steady state* (Ahnert 1984). In the steady state, the shape of the hill slope profile does not change anymore. This steady state is reached when the curvature of the landscape profile is exactly large enough, so that the diffusive mass transport balances the mass production described by the rock uplift  $v_{ro}$  (s. p. 66 to estimate when such steady states are reached). The attainment and preservation of an erosional steady state can be observed on Fig. 4.36. We have discussed an equivalent steady state in the theory of heat conduction when discussing the steady state temperatures in subduction zones (p. 107).

**Time Scales of Mass Diffusion.** For the construction of the curves in Fig. 4.35 or 4.37 we have assumed that  $D = 10^{-6}$  m $^2$  s $^{-1}$ . We have assumed this value in order to retain some analogy to chapter 3 where we have discussed that the thermal diffusivity is  $\kappa \approx 10^{-6}$  m $^2$  s $^{-1}$  and rarely varies by more than a factor of two from this value. However, the diffusivity of mass in erosion processes is much more variable. It depends on climate, material and many other parameters. Mass diffusivities have been reported to range between  $D = 1 \times 10^{-10}$  m $^2$  s $^{-1}$  or  $D = 1.7 \times 10^{-10}$  m $^2$  s $^{-1}$  for largely unconsolidated materials in different regions in China and the Tien Shan (Tapponier et al. 1990; Avouac and Peltzer 1993), to  $D = 5.3 \times 10^{-12}$  m $^2$  s $^{-1}$  and  $D = 1.8 \times 10^{-11}$  m $^2$  s $^{-1}$  as estimates for the in-strike and cross-strike

diffusivities for vertically bedded sandstone at Ayers Rock, Australia (Stüwe 1994). Mass diffusivities may vary by many orders of magnitude and unless we have measured  $D$  in our region of interest, we can usually not know if a given landform formed over a long time with a small diffusivity or a short time with a high diffusivity.

However, we can see that in eqs. 4.69 and 4.71, *time* and *diffusivity* occur always as a linear product and profiles are therefore identical if the product of time and diffusivity is a constant. Thus, it is often useful to define a “degradation coefficient”,  $t_a$ , as the product of time and diffusivity:

$$t_a = Dt \quad . \quad (4.72)$$

Degradation coefficients are also known by the name “diffusion age”, although we can see from eq. 4.72 that it has the units of  $\text{m}^2$ . Note however, that degradation coefficients are only defined by the simply linear product of eq. 4.72 if the diffusivity is constant in time. If the diffusivity changes with time (as often is the case if climate or the state of consolidation change during erosion), then the degradation coefficient is defined by a more complicated function discussed and explained in detail by Avouac (1993).

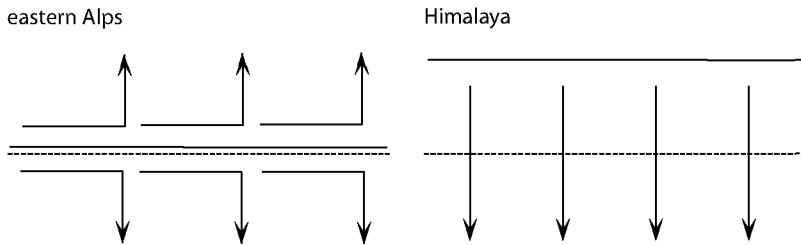
- *Diffusive time versus length scale.* The degradation coefficient corresponds to the product  $t_{eq} \times \kappa$  in eq. 3.17 and can therefore be used to convert between time and length scale in diffusion processes (s. also eq. 7.4). The diffusive time scale argument used on p. 62 can directly be applied to estimate the rough time scale of erosion processes and is therefore a useful tool for the field geologist. As the analogy between energy and mass diffusion is straightforward, the discussion from p. 62 is not repeated here.

### 4.5.3 Long Range Transport: Drainages

Fluvial erosion is one of the most important landscaping processes and the transport of material by rivers is an efficient mechanism for redistribution of mass on a large scale. Fluvial erosion is therefore often called *long-range* transport. In the field, this is documented by the enormous incision rates of some rivers and their extensive sedimentary fans elsewhere; for example the Indus or the Tsangpo in the Himalayas. In fluvial erosion processes it is often useful to discern between:

- supply limited erosion,
- transport limited erosion.

In supply-limited processes the transport of material out of the system is much more efficient than the sediment production rate, while transport limited processes are limited by the fluvial transport of rapidly supplied sediment (e.g. Tucker and Slingerland 1994). In the following we discuss some methods of characterizing fluvial erosion.



**Figure 4.38.** Schematic sketch of two different spatial relationships between drainage divide (continuous line), axis of the highest topography (dashed line) and direction of drainages (arrows). In the eastern European Alps the principle drainage divide corresponds to the region of the highest topography and the principle drainages are parallel to this axis. In the Himalayas, the principle drainage divide is some hundreds of kilometers north of the line of the highest topography and the drainages are perpendicular to this axis

**Networks.** The geometry of drainages – both with respect to a longitudinal profile along the drainage, as well as in plan view – usually have spatial patterns that are characteristic of tectonic and erosion processes (s. Summerfield 1991, Burbank and Anderson 2001), (Fig. 4.38). It is therefore useful to discuss some model tools that can be used to describe the spatial pattern of drainages. In *plan view*, the spatial characteristics of drainage networks may be described by their:

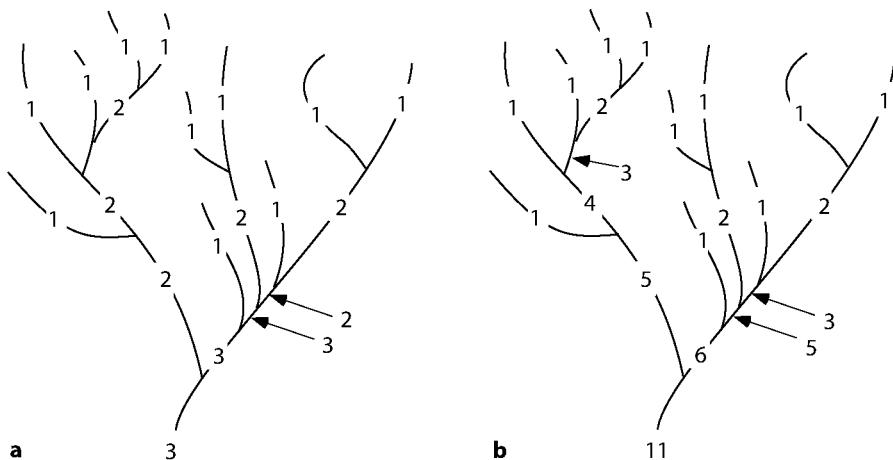
- topological properties,
- geometrical properties.

Both properties are often largely independent of scale. They are *self-similar* and lend themselves to a description as fractal shapes (s. sect. 4.5.6).

- *Network rules.* The topological properties of drainage networks may be characterized by allocating each stream section an order. There are different rules, how this may be done. According to Horton (1945), Strahler (1964) and Schumm (1956) the first stream after the spring has the order 1. When two streams of different order merge, the subsequent stream has the order of the higher order stream at the confluence. If both streams have the same order, then the order of the subsequent stream is larger by one (Fig. 4.39a). According to the scheme of Shreve (1967) the order of a stream is the number of contributing springs (Fig. 4.39b). If one uses the scheme of Horton (1945), then the topology of most natural networks appear to follow simple exponential laws. For example, the law of stream number states that the number of streams of order  $i$  may be described by:

$$N_i = a_1 e^{b_1 i} \quad . \quad (4.73)$$

There,  $N_i$  is the number of streams of the order  $i$  and  $a_1$  and  $b_1$  are constants. Similar laws apply to the length of streams of different order  $l_i$  and the size of the catchment of each stream  $A_i$ :



**Figure 4.39.** Stream order according to two different schemes: **a** according to Horton (1945) and Strahler (1964); **b** according to Shreve (1967)

$$l_i = a_2 e^{b_2 i} \quad \text{and} \quad A_i = a_3 e^{b_3 i} \quad . \quad (4.74)$$

There,  $a_2$ ,  $a_3$ ,  $b_2$  and  $b_3$  are constants. These network laws can be used to characterize the topology of an entire network with only a small portion of the network. However, Kirchner (1993) showed that practically all networks follow these laws and that it is therefore difficult to discern between artificial, random and natural networks using these rules (s. also Tarboton 1996).

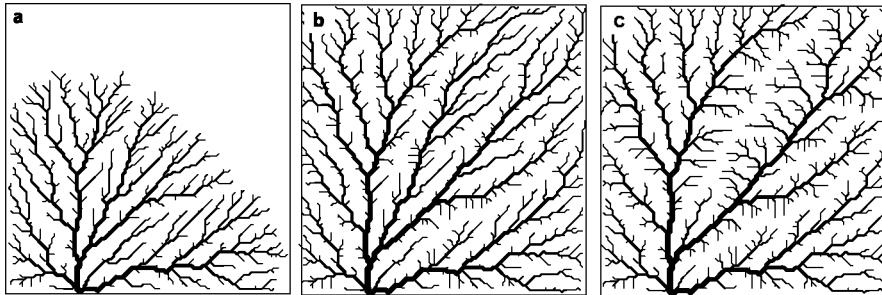
**Models Describing Fluvial Erosion.** In order to describe landscaping by rivers, a series of elegant models have been designed in the last few years (Ahnert 1976; Kooi and Beaumont 1994; Beaumont et al. 1992; Willgoose et al. 1991; Chase 1992; Tucker and Slingerland 1994, 1996). In many of these models two different types of fluvial channels are distinguished which are thought to form due to different physical processes:

- bedrock channels,
- alluvial channels

- **Bedrock channels.** For the description of the erosion of bedrock channels, it may be assumed that the rate of incision due to bedrock erosion is a function of discharge and channel gradient:

$$v_{\text{er}} = - \frac{\partial H}{\partial t} \propto q_r^\alpha \left( \frac{\partial H}{\partial l} \right)^\beta \quad . \quad (4.75)$$

There,  $t$  is time,  $H$  is the elevation of the channel,  $v_{\text{er}}$  is the erosion rate,  $l$  is the horizontal distance along the channel and the discharge  $q_r$  has the units of discharge volume per drainage width and time ( $\text{m s}^{-1}$ ). The negative sign arises because the elevation gets smaller for larger discharge and slope.  $\alpha$  and



**Figure 4.40.** Three different time steps of a landscape evolution model. The shown drainages incise exclusively according to the erosion law formulated in eq. 4.75 into a flat topography disturbed only by numerical noise. Despite this simple model assumption, interesting patterns develop. In particular note that the river developing in the central diagonal of the shown section disappears between **b** and **c**. The pattern shown in **c** approaches a steady state and further time steps do not show significant differences (after Hergarten 2002)

$\beta$  are exponents defining the non-linearity and are not to be confused with the “proportional to” sign  $\propto$ . This model is based on an empirical relationship between slope and drainage area found by Hack (1957). If it is assumed that  $q_r$  correlates with the size of the drainage basin above a given point, eq. 4.75 can be numerically solved for erosion rate at any given point of a digital elevation model to produce an elegant description of evolving landscapes (Fig. 4.40) (Beaumont et al. 1992; Hergarten 2002).

However, in many natural channels, the general law of eq. 4.75 can be simplified enough so that an analytical solution can be found that can then be compared with measured longitudinal river profiles (Fig. 4.42). For this, we assume that the erosion rate is *directly* proportional to the discharge (i.e.  $\alpha = 1$ ) and that the discharge is directly proportional to the area  $A$  of the drainage basin (which is a fair assumption, unless there are hydrological peculiarities like strong precipitation gradients or karst features like sinks). Then:

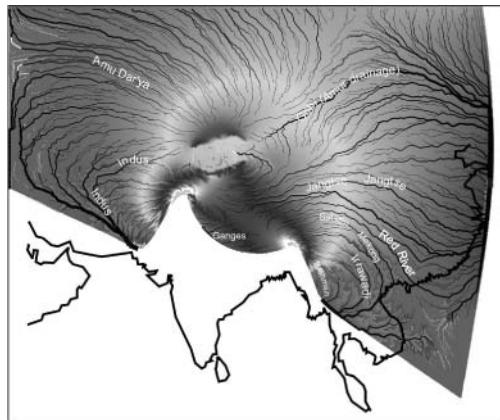
$$v_{er} = -\frac{\partial H}{\partial t} \propto A \left( \frac{\partial H}{\partial l} \right)^\beta . \quad (4.76)$$

The proportionality in eq. 4.76 could be replaced by an equality if a proportionality constant is added to the right hand side. However, in the following steps, several other proportionality constants must be inserted, so that we rather stick with eq. 4.76 and retain the “ $\propto$ ” sign. The area  $A$  is usually related to the channel length  $l$  by:

$$l \propto A^h . \quad (4.77)$$

The exponent would be  $h = 0.5$  if the basin area rose quadratically with river length. However, most basins are somewhat elongated and for many rivers this exponent appears to be around  $h = 0.56$ . The exact exponent can be

**Figure 4.41.** Example of an integrated landscape evolution model for the India- Asia collision zone where a simple erosion model according to eq. 4.75 (causing thinning of the crust) was coupled with force balanced finite deformation producing topography (and thickening of the crust). See also Fig. 1.6 which was performed with similar boundary conditions, but no erosion. Shading is surface elevation. Drainages are exclusively dynamically developed, white parts of drainages are where they sediment



derived from an area - distance plot as shown on Fig. 4.42. Using eq. 4.77, eq. 4.76 becomes:

$$v_{er} = -\frac{\partial H}{\partial t} \propto l^{(1/h)} \left( \frac{\partial H}{\partial l} \right)^\beta . \quad (4.78)$$

In an equilibrium channel, where the erosion rate is the same at every point, the erosion rate  $\partial H / \partial t$  is a constant and eq. 4.78 may be written as:

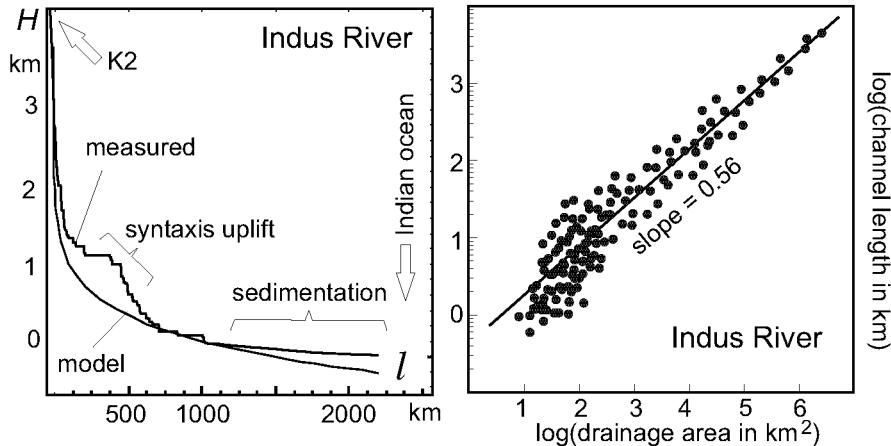
$$\frac{dH}{dl} \propto l^{(-1/\beta h)} \quad \text{or :} \quad dH \propto l^{(-1/\beta h)} dl . \quad (4.79)$$

Integration (according to the rules given in appendix B.5.2) gives:

$$H \propto \frac{-l^{1-1/(ah)}}{1 - 1/(\beta h)} + C . \quad (4.80)$$

The integration constant  $C$  determines the absolute elevation of the profile. The proportionality constant that must be inserted to make eq. 4.80 into a proper equation summarizes a series of erodibility terms that determine how rapid  $H$  falls with  $l$ . However, the *shape* of curves plotted with eq. 4.80 is independent of the constant. Eq. 4.80 provides a good description of eroding rivers in equilibrium and is a useful model that can be used to determine knick points by comparing equilibrium channels with channel profiles measured from digital elevation models (Fig. 4.42).

- *Alluvial channels.* Alluvial channels are those where the sediment transport of the channel is important to its erosion / deposition characteristics. For the description of those, it is often useful to consider the erosion/sedimentation processes in terms of the sediment carrying capacity of the stream



**Figure 4.42.** Longitudinal profile along the Indus River (between the mountain K2 in the Karakorum and the Indian ocean), in comparison with the analytical description of eq. 4.80. The deviation in the Himalayan syntaxis is clearly due to recent uplift causing the channel to be out of equilibrium. The deviation in the lower course is because there is no erosion but sedimentation in this part of the river. It can therefore not be described with the model discussed here. The right hand diagram is a simplified area - distance plot for the Indus River basin

$$q_{f(\text{eq})} = -K_f q_r^m \left( \frac{\partial H}{\partial l} \right)^n . \quad (4.81)$$

where  $q_{f(\text{eq})}$  is the equilibrium carrying capacity of a river (i.e. that a graded river bed carries),  $K_f$  is a constant of proportionality and  $m$  and  $n$  are exponents defining the non-linearities (Beaumont et al. 1992; Begin et al. 1981; Willgoose et al. 1991). The other parameters are as above. Most modern models assume simply that  $m = n = 1$  (see Kooi and Beaumont 1994; Beaumont et al. 1992). Erosion / sedimentation processes may then be modeled by comparing the equilibrium carrying capacity of the river  $q_{f(\text{eq})}$  given by eq. 4.81 with the actual sediment freight carried by a drainage  $q_f$ . Whether the river erodes- or sediments onto its bed, then depends on whether the sediment flow  $q_f$  is larger or smaller than the sediment equilibrium carrying capacity. The actual sediment freight  $q_f$  can be calculated using the following two assumptions:

- 1. The rate of change of the sediment content of a river is proportional to the magnitude of the disequilibrium  $dq_f/dt \propto (q_{f(\text{eq})} - q_f)$ . If  $q_f < q_{f(\text{eq})}$ , then the amount of sediment in the river increases and the river will erode its bed. If  $q_f > q_{f(\text{eq})}$ , then there is more sediment in the stream than it can hold and there is sedimentation in the river bed.
- 2. The change of the sediment content  $dq_f/dt$  is inversely proportional to a *length scale of reaction*,  $l_f$ , which may be interpreted as the distance that water needs to flow along a river bed in order to do work on the river bed.

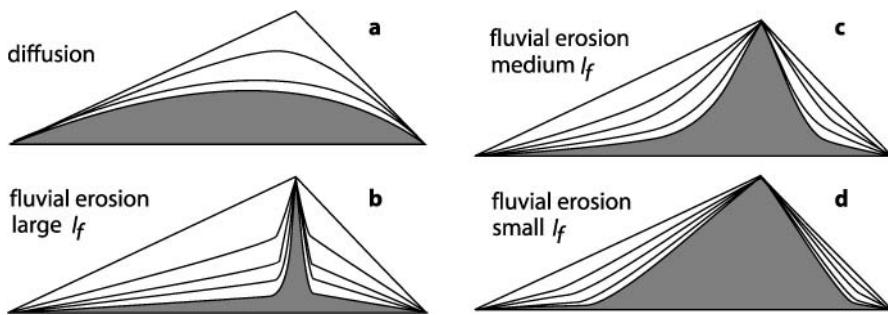
If the reaction length scale is large, then the change of sediment freight of the river occurs slowly and over larger length scales and vice versa.

Using these two assumptions  $q_f$  may be evaluated by numerical integration along the river bed. The rate of vertical lowering (by erosion) or rising (by sedimentation) of the river bed may be then described by:

$$\frac{dH}{dt} = -\frac{dq_f}{dl} = -\frac{1}{l_f} (q_{f(eq)} - q_f) . \quad (4.82)$$

Figs. 4.43b, c and d show schematic illustrations of the temporal evolution of drainages as calculated with eq. 4.82. All drainages may be divided into two sections: a lower part where the slope of the river *decreases* and an upper part where the slope of the river *increases*. In the lower section of the rivers a *graded river bed* develops. The section of the river bed which is *graded* grows with the evolution of the drainage upwards. The length of the section of the grade river bed and the rate with which it grows drainage upwards, depends on the reaction length scale.

- *Graded river beds.* If we plot the surface elevation of a river bed against distance from the spring (a cross section along the river) many drainages have a shape that reminds of an exponential function: The drainage bed is steep at first and then gets shallower with increasing distance from the spring. This natural development often leads to the development of *graded river beds* (Mackin 1948). In a graded river, the slope of the drainage and the flow of water are in equilibrium so that neither erosion nor sedimentation takes place. The flow rate is exactly large enough to transport the sediment from the drainage basin above. Because of this, a graded river can maintain its shape and is in geomorphic equilibrium.
- *Integrated landscape models.* Eq. 4.82 describes many field observations made about the geometry of river profiles very well and has therefore been used by a range of authors to describe the temporal evolution of drainages and drainage networks. However, it fails to describe processes at the drainage divide for which different models must be invoked. Within the model of eq. 4.82 the parameters  $q_r$  and therefore  $q_f^{eq}$  and  $q_f$  are zero at the watershed and drainage divides will therefore remain preserved as steep ridges. The steepness of the ridge is proportional to the reaction length scale. However, the decay of the drainage divide itself may be well-described by diffusive mass transport discussed in the previous sections. Beaumont et al. (1992), Tucker and Slingerland (1994) and others have integrated combinations of diffusive short range transport and long range transport to describe two-dimensional landscape evolutions. This model was transferred onto irregular grids by Braun and Sambridge (1997) which is today probably one of the most elegant model for the integrated description of landscape evolutions.



**Figure 4.43.** Erosion of a model landscape. **a** Relief formation by diffusion. If the starting profile is asymmetric, then diffusion will lead to a lateral shift of the drainage divide. **b, c** and **d** Relief formation by fluvial erosion (eq. 4.82). The model rives flow on both sides of the drainage divide along the surface of the gray shaded region outwards. The water shed (drainage divide) itself is not affected by this. The three examples differ in the magnitude of the reaction length scale  $l_f$ . In **b**,  $l_f$  is much smaller than the length of the river, in **c** it is comparable and in **d**  $l_f$  is much longer than the length of the river

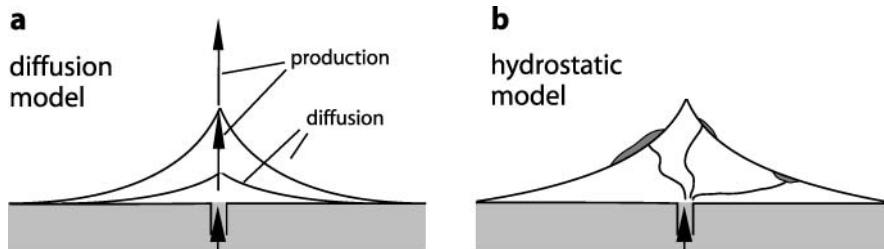
#### 4.5.4 Discontinuous Landscape Formation

Not all geomorphic processes may be described using continuous models like those discussed in the previous sections. For example, the occurrence of landslides is a classic discontinuous process that requires its own class of model description. How *continuous* geological processes may cause *discontinuous* processes will be discussed in some detail in sections 6.3.6 and 6.3.6. We will show there that discontinuous processes may be triggered by *threshold mechanisms*, or be the direct cause of *non-linear feedback* between different processes. The particular example of landslide occurrence is a beautiful example illustrating the meaning of self-organized criticality (Hergarten 2002).

Consider an incising valley. As a river incises into the landscape, the slopes will steepen until they reach a critical angle where a steady state geomorphic profile is reached (s. p. 192, Fig. 4.37b). However, in many real landforms, this steady state will not be maintained by continuous incremental transport of material into the drainage, but by discontinuous landslides that cause a temporal fluctuation of slope around the steady state angle. The size-frequency distribution of such land slides is fractal and the state of the slope is said to be in a state of self-organized criticality (Turcotte 1997).

#### 4.5.5 The Shape of Volcanoes

Many strato-volcanoes have an intriguingly similar shape and surface elevation suggesting that they formed by similar processes (Turcotte 1997). For example, Etna in Sicily, Mt. Fuji in Japan as well as many volcanoes in Indonesia and in Alaska are all about 3 500 m high and have conical profiles



**Figure 4.44.** Two models for the formation of strato volcanoes. **a** This model assumes that volcanoes get their shape from the interaction of mass *production* at a single point at the center of the volcano and mass *diffusion* distributing the material (s. also Fig. 3.9). **b** This model assumes that volcanoes describe surfaces of constant hydrostatic pressure. Magma does not necessarily extrudes at the tip of the volcano, but follows the way of least resistance giving the mountain its shape without erosive influence. While both models lead to similar shapes, detailed comparison between measured shapes of volcanoes with the two models may help to provide constraints on future eruption sites

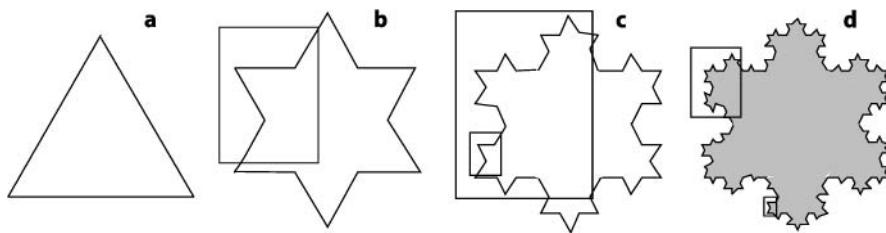
with concave flanks. Two very simple (but very different) models may be used to explain this shape. One relies on the principals of mass production and diffusion as we have discussed in several parts of this book; the other is a hydrostatic model.

- *Diffusion model.* The diffusion model is based on the assumption that magma erupts always from the same point on the surface and that the erupted material is distributed from there by mass diffusion (Fig. 4.44a). This model may be described with the equations we have used in sections 3.1 and 3.2 for the simultaneous production and diffusion of heat. If we assume that the erupted magma is distributed concentrically around the eruption point, then we can describe this problem in cylindrical coordinates using the eruption point as the origin and using an equivalent of eq. 3.23 in cylindrical coordinates (s. p. 59). Initial and boundary conditions for the problem may be formulated as:

- Initial condition:  $H = 0$  for all  $r$  at  $t = 0$  as well as  $v_{ro} =$  magma production rate at  $r = 0$  and  $v_{ro}=0$  at  $r > 0$ .
- Boundary condition:  $dH/dr = 0$  at  $r = 0$  and  $H = 0$  at  $r \rightarrow \infty$  for  $t > 0$ .

$v_{ro}$  is the rate of magma production at the coordinate origin and replaces the heat production rate in eq. 3.23. These conditions are equivalent to those we have used on p. 71 for the description of the influence of frictional heating around shear zones. The only difference lies in that we have used Cartesian coordinates there (as shear zones are generally planar) and will require cylindrical coordinates for the description of volcanoes. However, the curves in Fig. 3.9 give a qualitative indication of the shape obtained with this model.

- *Hydrostatic head model.* The second model that has been used to describe the shape of volcanoes relies on the assumption that the surface of a volcano



**Figure 4.45.** The first four stages of Koch's snowflake, a typical example for the development of a fractal geometry. The final form at **d** has a fractal geometry as it is self similar. The sections inside the box in **b** and **c** or in **c** and **d** are identical

corresponds to a surface of constant hydrostatic head over a point source where the magma erupts (Fig. 4.44b). Thus, it is assumed that magma will always erupt to the surface where it finds the least resistance, even if this is not the crater vertically above the point source. This second model is discussed in detail by Turcotte and Schubert (1982).

Both models discussed above lead to cone shapes with concave surfaces, just like we observe volcanoes to look like. However, there is sufficient difference in the details of the two models, so that it may be possible to use the observed shape of a volcano in comparison with both model to predict what processes govern the magma distribution in the chosen example. It may be possible to use the models described here to make predictions about the likely points of future eruptions.

#### 4.5.6 Fractals

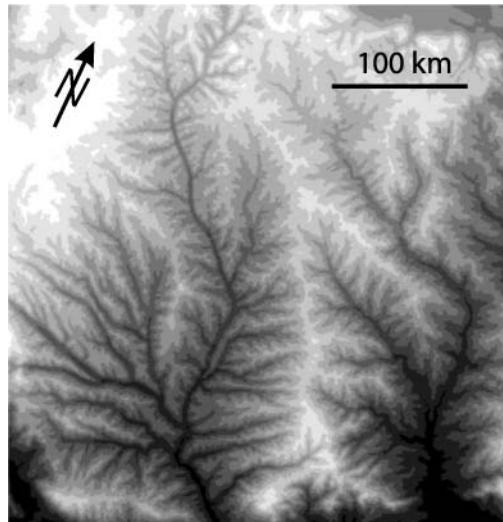
Many morphological forms on the earth's surface have a fractal shape, for example coast lines (Mandelbrot 1975) or the shape of the earth's surface itself (Chase 1992). Fractals are usually brought in connection with non-linear feed back and chaotic processes because many geometric representations of non-linear phenomena do indeed render fractal shapes (e.g. Turcotte 1997; s. sect. 6.3.6). However, fractals themselves have nothing to do with chaos or feedback. Rather, they are simply defined as a geometric object that has the following properties (Fig. 4.45):

- The shape of the object must be able to be characterized without having to give a scale. This property is called *self similarity* or *scale independence*.
- It must be possible to characterize the object with a *fractal dimension*.
- *Fractal dimension.* The fractal dimension of an object is defined as follows:

$$D = \frac{\log(m)}{\log(n)} . \quad (4.83)$$

There,  $m$  is the number of objects and  $n$  is a characteristic linear dimension. This may be illustrated with the fractal dimension of a simple geometric

**Figure 4.46.** Fractal drainage patterns in the Chinese Loess Plateau. Similar fractal patterns can be found in the shape of the Grand Canyon and countless other examples around the world. However, the Chinese Loess Plateau is probably the largest and most spectacular example in the world (because of the isotropy and high erodability of the material). This example was also chosen by Turcotte (1997) as a title image for his textbook on fractals



object (called a *geometrical fractal*). Consider a square that is subdivided into four sub-squares. If the edge of the square is subdivided into  $n$  pieces, then the total number of small squares is  $m = n^2$ . If one magnifies the length of the edge of each sub-square by the factor  $n$  we return to the original square. Using eq. 4.83 we arrive at  $D = \log(m)/\log(n) = 2$ . For a square (having obviously two dimensions) this is quite trivial. However, for the snowflake of Koch things are not so clear (Fig. 4.45). Each edge of a triangle with length  $a$  corresponds to four edge sections each of which has a length of  $a/3$  in each subsequent figure in Fig. 4.45. We can write:  $D = \log(4)/\log(3) = \log(16)/\log(9) \approx 1.262$ .

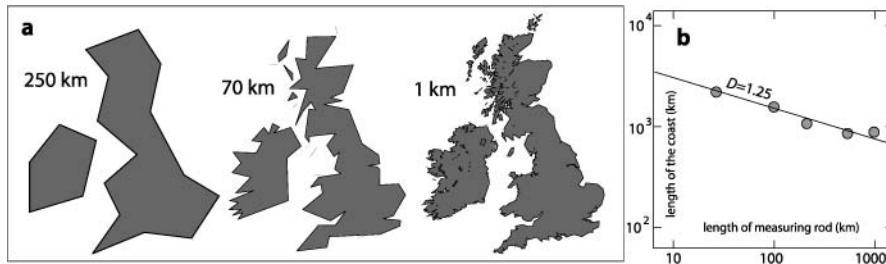
In natural landscapes there are two important fractal dimensions that may be defined and used to characterize the landscape:

- the fractal dimension of the drainage network,
- the fractal dimension of the topography itself.

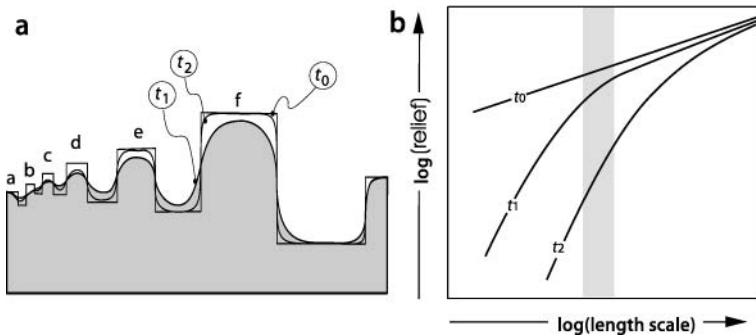
Both can be obtained *statistically* from natural landscapes and they are therefore called “statistical fractal” (in contrast to the geometrical fractal dimension of Koch’s snowflake). The fractal dimension of *drainage networks* is defined as the ratio of the logarithms of bifurcation ratio and length-order ratio:

$$D = \frac{\log(R_b)}{\log(R_r)} . \quad (4.84)$$

where  $R_b$  is the ratio of the number of streams of a given order, to the number of streams of the next higher order and  $R_r$  is defined as a length ratio in a corresponding way. Typical drainage networks (like that shown in Fig. 4.46) have fractal dimensions that are somewhat less than space filling with  $D \approx 1.8$ . The fractal dimension of *coast lines* is similar to that of drainage networks



**Figure 4.47.** a The coast of Britain in three different resolutions (labeled in km). The figure was constructed using the generic mapping tool GMT. The length of the coast line segments are determined by the resolution of the digital elevation models used and are therefore not exactly of the same length. Nevertheless, the figure illustrates the relationship between segment length and coast line length used to determine the fractal dimension in b (after Mandelbrot 1967)



**Figure 4.48.** a Schematic cartoon of an idealized landscape made up of horsts and grabens which were subsequently eroded and sedimented upon by mass diffusion processes. Curve  $t_0$  describes the landscape at the end of the tectonically induced horst and graben formation. The other two curves show the landscape at two subsequent times  $t_1$  and  $t_2$  b shows the fractal dimension of the landscape at times  $t_0$ ,  $t_1$  and  $t_2$  in a plot of relief against length scale  $l$ . There, relief is the elevation difference between two points separated by a distance  $l$  (after Braun and Sambridge 1997)

inasmuch as it is smaller than 2. The fractal dimension of coast lines may be measured by taking  $m$  as the length of a measured coast line and  $n$  as the length of the measuring rod (Figs. 4.47). For the classic example of the west coast of Britain it is  $D = 1.25$  (Mandelbrot 1967).

Corresponding to the measurement of the fractal dimension of coast lines, the fractal dimension of entire landscapes may be defined as the ratio of the logarithm of the relief and the logarithm of length scale. Natural landscapes have a fractal dimension around  $D \approx 2.1\text{--}2.7$  (Mandelbrot 1982).

**Using Fractal Dimension.** Consider a landscape that was created by tectonic processes, for example the idealized horst and graben structure shown

for time step  $t_0$  in Fig. 4.48a. If the tectonic processes created a more or less random topography, then such a landscape may be characterized by a single fractal dimension as shown by the straight line for time  $t_0$  in Fig. 4.48b. Subsequent erosion and sedimentation processes will destroy landforms at a rate that is proportional to their length scale. Small landforms will be quickly destroyed by mass diffusion processes, while intermediate landforms are only rounded and the largest landforms will preserve their original shape the longest. The length scale of the landform that separates *smaller* landforms (that are now characterized by “erodic shapes”) from *larger* landforms (that still retain the shape created by the initial tectonic event) may be used to derive a characteristic diffusive length scale of the landscape. At time  $t_1$  this is roughly the length scale of the landform  $e$  for the example in Fig. 4.48a. At time  $t_2$  erosion has proceeded further and only the valley located to the right of landform  $f$  is large enough so that its original shape is still dominant. The landscape becomes multi fractal as indicated by the curves for  $t_1$  and  $t_2$  in Fig. 4.48b. As we have discussed on p. 193, the diffusive length scale may now be converted into a time scale and therefore can be used to date the time of the last landscape forming tectonic event. Such qualitative considerations can be considerably improved by using statistic evaluation of length scales over an entire landscape (Chase 1992). Braun and Sambridge (1997) used this approach to date the age of the eastern Australian highlands.

## 4.6 Problems

### **Problem 4.1. Vertical reference frames (p. 143):**

The summit of Chimborazzo in Ecuador is 6 180 m above sea level. The summit of Mt Everest is 8 848 m above sea level. Use eq. 4.8 to calculate the elevation difference between the two summits measured from the center of earth rather than from sea level. Chimborazzo lies practically at the equator, Mt Everest at roughly  $\lambda = 28^\circ\text{N}$ .

### **Problem 4.2. Uplift versus exhumation (p. 150):**

- How much was a rock uplifted relative to a fixed external reference level if it was exhumed from 10 km depth to 5 km depth by erosion and the eroding mountain range decreases in this time from 8 000 m to 3 000 m elevation.
- What is the amount of exhumation a rock experienced if the surface uplift was 5 km and the rock was uplifted 4 km relative to a fixed reference level.

### **Problem 4.3. Vertical kinematics (p. 159):**

Use eq. 4.15, eq. 4.17 and eq. 4.18 to calculate the uplift and the exhumation history of a simple model mountain range that may be described with these equations. Assume that the range is in geomorphic steady state, so that its surface elevation remains constant because  $v_{\text{er}} = v_{\text{ro}}$ .

- What is the magnitude of the erosion parameter  $t_E$  of this range? Use the following constants:

$a = 2500$  m and  $b = 0.13$  and  $\dot{\epsilon} = 10^{-15}$  s $^{-1}$ . b) What was the depth of rocks that reach the surface after 40 my at the start of their evolution? Use eq. 4.19. c) How long is the hiatus between the onset of surface uplift and the onset of exhumation of these rocks? Use eq. 4.19.

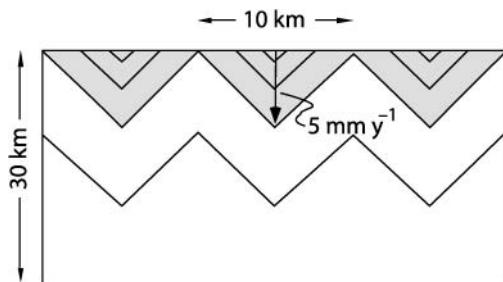
**Problem 4.4.** *The principle of isostasy (p. 163):*

We are familiar with the fact that 90% of icebergs are under water and 10% above the surface. Use Fig. 4.16 and eq. 4.23 to determine the density of ice from this observation. The density of water is  $\rho_w = 1000$  kg m $^{-3}$ .

**Problem 4.5.** *Relationship between relief and uplift (p. 163):*

Parallel rivers incise an isostatically compensated plateau with a vertical incision rate of:  $v = 5$  mm y $^{-1}$  (Fig. 4.49). The slopes on both sides of the rivers are 45° steep. All rivers have a distance of  $l = 10$  km from each other. Draw the surface elevation of the ridges and the river beds as a function of time assuming the plateau remains isostatically compensated during the incision process. Assume the following numerical values for the necessary physical parameters:  $\rho_m = 3200$  kg m $^{-3}$  and  $\rho_c = 2700$  kg m $^{-3}$ . (An interesting discussion of this problem may be found by Montgomery 1994.)

**Figure 4.49.** Illustration of Problem 4.5



**Problem 4.6.** *Controls on surface elevation (p. 163):*

Calculate the surface elevation changes of an isostatically compensated mountain belt that occur in response to the following processes: a) Continental crust of thickness  $z_c = 30$  km ( $\rho_c = 2700$  kg m $^{-3}$ ) is being underplated by a 10 km thick basaltic layer of density  $\rho_u = 2900$  kg m $^{-3}$ . The density of the underlying mantle is  $\rho_m = 3200$  kg m $^{-3}$ . b) Continental crust is being thinned to half its thickness ( $f_c = 0.5$ ) and underplated by a 5 km thick layer of basaltic underplate of density  $\rho_u = 2950$  kg m $^{-3}$ . c) What is the difference in surface elevation between a) and b)?

**Problem 4.7.** *Age depth relationship of oceanic lithosphere (p. 167):*

a) Is the rifting rate at the Mid-Atlantic Ridge 1.5 cm y $^{-1}$ , 2.5 cm y $^{-1}$  or 3.5 cm y $^{-1}$ ? Calculate the answer assuming that the bathymetry of the Atlantic may be described with the model of eq. 4.45 (Fig. 4.20), and

use the following bathymetric data from the north Atlantic Ocean. The data are given in pairs for water depth/distance from the Mid-Atlantic Ridge: 300 m/50 km; 500 m/175 km; 800 m/250 km; 1 300 m/500 km; 1 500 m/700 km; 1 800 m/825 km; 2 300 m/1 300 km; 2 600 m/1 575 km; 2 800 m/1 775 km; 2 900 m/1 950 km; 3 200 m/2 500 km; 3 200 m/3 125 km; 3 300 m/3 375 km; 3 200 m/3 625 km.  $\rho_m = 3200 \text{ kg m}^{-3}$ ;  $\rho_w = 1000 \text{ kg m}^{-3}$ ;  $\alpha = 3 \cdot 10^{-5} \text{ }^{\circ}\text{C}^{-1}$ ;  $T_l = 1200 \text{ }^{\circ}\text{C}$ ;  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ . b) From a certain distance from the ridge (age) onwards, these data do not correspond very well with the model, even for the correct answer of a). What is this distance (age) and why does the model not work for older oceanic lithosphere?

**Problem 4.8. Understanding the flexure equation (p. 173):**

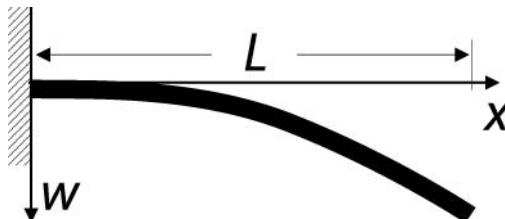
Perform a dimensional analysis of eq. 4.49 and eq. 4.48. a) What are the units of  $D$ ,  $F$  and  $q$ ? b) Typical continental lithosphere may be described with a Poisson's ratio around  $\nu = 0.25$  and a Young's modulus of  $E = 10^{11} \text{ Pa}$ . What is the flexural rigidity of a continent that may be described with an elastic thickness of  $h = 10 \text{ km}$  to  $h = 70 \text{ km}$ ?

**Problem 4.9. Understanding the flexure equation (p. 173):**

(a) Integrate the flexure equation (eq. 4.48) to describe the shape of a fishing rod of length  $L$  held horizontally at its end and bent downwards under its own weight. Use the geometry shown in Fig. 4.50. (Equivalent curves are described by a piece of paper hanging off the edge of a table, or a bridge under construction extending half over a valley).

Some hints: 1.) Since there is no horizontal force applied to the rod, the last term of eq. 4.48 is zero. 2.) Since the load per unit length is everywhere the same, the load is independent of  $x$ . Eq. 4.48 simplifies to  $d^4w/dx^4 = q/D$ . 3.) Since there is no torque or change in torque applied to the free-hanging end of the rod, it is true that  $d^2w/dx^2 = 0$  at  $x = L$  as well as that:  $d^3w/dx^3 = 0$  at  $x = L$ . 4.) Since the rod is held out horizontally, it is also true that  $w = 0$  at  $x = 0$  and that:  $dw/dx = 0$  at  $x = 0$ . (b) Use your result to estimate the flexural rigidity of a piece of paper hanging  $L = 10 \text{ cm}$  off the edge of a table with the lowest point of the paper tip hanging  $w_{x=L} = 3 \text{ cm}$  down. The weight of the paper corresponds to a load of  $q = 2 \text{ N m}^{-2}$ .

**Figure 4.50.** Illustration of Problem 4.9



**Problem 4.10.** *Internal and external loads (p. 174):*

a) Repeat the derivation of eq. 4.50 using the principles of eq. 4.21. In other words, derive eq. 4.50 by comparing the weights of the vertical columns in Fig. 4.24 at the points A and B. b) Derive an equation for the load on oceanic plates that corresponds to eq. 4.50. This case is illustrated in Fig. 4.24 on the very left. The variable  $q_a$  is a line load. That is, it has to be imagined to extend infinitely in the direction normal to this page.

**Problem 4.11.** *Understanding the flexural parameter  $\alpha$  (p. 176):*

What are the units of the flexure parameter  $\alpha$  in eq. 4.53? How large is  $\alpha$  roughly for the continental lithosphere? Use  $D \approx 10^{23}$  Nm;  $\rho_m = 3200 \text{ kg m}^{-3}$  and  $\rho_w = 1000 \text{ kg m}^{-3}$ .

**Problem 4.12.** *Plate flexure near Hawaii (p. 176):*

The water depth of the Pacific in the vicinity of the Hawaii- and Emperor Island chains is somewhat deeper than in the remainder of the abyssal planes of the Pacific. Also, there is a high point in the ocean floor, about 250 km off the coast of Hawaii. (Fig. 4.22). This water depth profile corresponds well with the shape of an elastically flexed plate loaded by the weight of the island chain with the high point in the ocean floor being the flexural bulge. What is the thickness of the elastic lithosphere in the Pacific? Use eq. 4.54, 4.53, 4.49 and the following values for the physical parameters:  $E = 70 \text{ GPa}$ ,  $\rho_m = 3200 \text{ kg m}^{-3}$ ;  $\rho_w = 1000 \text{ kg m}^{-3}$ ;  $g = 10 \text{ m s}^{-1}$  and  $\nu = 0.25$ . For the most elegant solution of this problem we need not even know the details of the shape of the flexed plate or the maximum water depth at the point of maximum flexure.

**Problem 4.13.** *Fault scarp degradation (p. 187):*

Avouac and Peltzer (1993) report a fault scarp from China with the following profile: (27,1; 51,1.1; 68,1.2; 77,1.8; 90,1.8; 110,1.9; 111,2.0; 120,2.2; 122,2.45; 125,2.6; 128,3.1; 130,4; 133,4.5; 135,5; 140,6.6; 143,7.5; 147,8.8; 150,10; 155,11; 160,12; 170,12.8; 180,13.5; 190,13.5; 220,13.7). This list shows the scarp coordinates  $(x_1, H_1; x_2, H_2 \dots)$ , where  $x$  horizontal distance from a fix point in the far field and  $H$  is the surface elevation above the base, both in meters. Assuming that the surface is flat in the far field and that the scarp was initially vertical, use eq. 4.67 to date the earth quake that formed it. Assume that the mass diffusivity is  $5.5 \text{ m}^2$  per 1000 years.

**Problem 4.14.** *Spheroidal weathering (p. 189):*

Measure the topographic profiles off the two boulders with the marked coordinate systems on Fig. 4.34 and use eq. 4.69 to estimate whether the two boulders commenced weathering at the same time, or if the smaller one has been subjected to weathering longer.

## 5. Mechanics: Force and Rheology

In this chapter we discuss the forces involved in geodynamic processes. Knowing a bit about forces is a great tool for the field geologist to test field observations. Let us consider an example: A field geologist finds folds and thrusts in a Precambrian terrain that he or she interprets to have formed as the consequence of crustal shortening. Strain analysis shows that 80% shortening occurred and the geometry of shortening indicates that this resulted in four-fold thickening of the crust. He therefore further infers (using the principle of isostasy) that - at the time - a mountain range of some 15 km elevation existed above the metamorphic terrain. This interpretation is consistent with the field observations, but it has no independent test. In this particular example we could argue that we have no knowledge of any present day mountain range that is this high and that, therefore, this interpretation is unlikely. However, in many less obvious examples there are no direct analogies and the resulting model is - albeit perfectly imaginable and fully consistent with field observations - wrong. One way to provide an independent test of such models is to make a rough estimate of the forces involved. In the next chapters we want to perform such estimates. In order to do so, it is necessary to commence with a brief repetition of the basics of stress and strain. For more details on the basics the interested reader is referred to a range of excellent text books on the subject. For example:

- Pollard and Fletcher (2006) Fundamentals of Structural Geology
- Twiss and Moores (1992) Structural geology.
- Ramsay and Huber (1983) Modern Structural Geology. Volume 1: Strain Analysis
- Ramsay and Huber (1987) Modern Structural Geology. Volume 2: Folds and Fractures
- Ramsay and Lisle (2000) Modern Structural Geology. Volume 3: Applications of Continuum Mechanics in Structural Geology
- Weijermars (1997) Principles of Rock Mechanics.

## 5.1 Stress - The Basics

Many field geologists and metamorphic petrologists are only ever confronted with a single variable that has the units of force per area: pressure. Only with in situ stress measurements and some palaeopiezometric methods is it possible to resolve the stress state in more detail than this single value. However, as soon as we try to describe the forces that act on a structural feature in some more detail we will discover that the words “pressure” or “stress” are simply not precise enough to describe what we see. In fact, we will soon discover that the state of stress of a rock can only be described by nine numbers that form a tensor. Because many field geologists remain uncertain about the relationship between the pressure they measure and the complete stress tensor, we begin in sect. 5.1.1 with a very brief summary of the stress tensor and its most important derived quantities. Much of the logic discussed below also applies to the strain tensor discussed in sect. 4.1.

### 5.1.1 The Stress Tensor

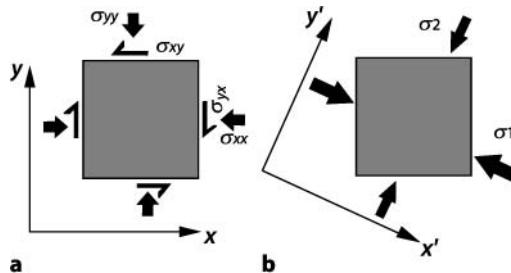
There are many excellent descriptions of stress in an abundance of good text books (e.g. Means 1976; Suppe 1985; Engelder 1993; Pollard and Fletcher 2006). Here we only summarize the definitions of a few terms related to stress (Engelder 1994).

**Force.** Force is a vector and - like all vectors - is described by a *magnitude* and a *direction*. It has the units of mass  $\times$  acceleration:  $1 \text{ N} = 1 \text{ kg m s}^{-2}$ . A related vector quantity is *traction*. Traction is a force (with magnitude and direction) *per area*, where the orientation of this area is not defined. Traction may be subdivided into normal and parallel components called *normal traction* and *shear traction*. It is important to note that tractions are vectors, although they have the same units as stress. In contrast, *stress* is a tensorial quantity described by all the tractions acting on a unit cube (s. sect. B.3). We will now discuss the stress tensor in a bit more detail.

**The Stress Tensor.** In three-dimensional space, the state of stress of a single point inside a rock (i.e. a unit cube) is given by nine numbers, all of which have the units of force per area. These nine numbers are all tractions acting on different planes and different directions that need to be defined by using subscripts. These nine numbers define the stress tensor which is typically written as:

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \quad (5.1)$$

(s. sect. B.3). The two different notations of subscript used in eq. 5.1 are both common in the literature. The first of the two specifying spatial indices  $x, y$



**Figure 5.1.** a The state of stress of a unit square within the two dimensions of this book page. In two dimensions, the stress tensor has only four independent components illustrated here by the four labeled arrows. In b the coordinate system  $x'$ ,  $y'$ ,  $z'$  was chosen in such an orientation so that the shear stresses become zero. The normal stresses become therefore *principal stresses*. The state of stress of the square in a and b is identical

and  $z$  (or 1, 2 and 3) indicates the direction in which the stress component acts. The second index indicates the normal to the plane on which this stress component acts. We can see that the three tensor components in the diagonal of this matrix have two identical indices. They are called *normal stresses* because the surface onto which the stresses act are *normal* to the direction in which they act, (i.e. the indices for “direction in which it acts” and “plane onto which it acts” are the same). In the following we abbreviate normal stresses with  $\sigma_n$ . The remaining six components of stress in eq. 5.1 are *shear stresses*. In these, the stress components they describe act *parallel* to the plane onto which they are exerted. Shear stresses are abbreviated in this book with  $\sigma_s$

The stress components in a given column of eq. 5.1 act on the same plane, but in different directions. The rows contain stress components oriented in the same direction, but acting on different planes.

In the literature, shear stresses are often abbreviated with  $\tau$  and normal stresses with  $\sigma$ . However, this notation is somewhat confusing as all components of the tensor have the same units and should be abbreviated with the same symbol. We therefore stick to the notation of eq. 5.1 in particular to the use of  $x$  and  $y$  rather than 1 and 2 as subscripts and describe shear stresses with  $\sigma_{i \neq j}$  or  $\sigma_s$  and normal stresses with  $\sigma_{i=j}$  or  $\sigma_n$ . We reserve  $\tau$  as a symbol for deviatoric stress.

The stress tensor is symmetrical, that is, each component above the diagonal has an equivalent component of equal magnitude below it:  $\sigma_{yx} = \sigma_{xy}$ ,  $\sigma_{zx} = \sigma_{xz}$ ,  $\sigma_{yz} = \sigma_{zy}$ . Thus, the stress tensor consists of only six independent numbers: three normal stresses (written in the diagonal) and three shear stresses (the off diagonal terms).

The state of stress described by eq. 5.1 can be expressed a bit more simply in a differently oriented coordinate system. Because of the symmetry of the stress tensor, it is always possible to assume a coordinate system with the coordinates  $x'$ ,  $y'$  and  $z'$ , in which all shear stresses (all off diagonal terms

in eq. 5.1) become zero. The diagonal components in this new coordinate system are called *principal stresses*. Principal stresses are denoted with a single subscript as  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$ . In the earth sciences it is common to use the subscript “1” for the largest principal stress and “3” for the smallest. Thus, the state of stress at a point may always be characterized by only three principal stresses (Fig. 5.1):

$$\sigma' = \begin{pmatrix} \sigma'_{xx} & 0 & 0 \\ 0 & \sigma'_{yy} & 0 \\ 0 & 0 & \sigma'_{zz} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} . \quad (5.2)$$

The order in which  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$  appear in eq. 5.2 implies that the new coordinate system was chosen here so that the  $x'$ -axis is parallel to the largest of the three principal stresses. Of course this need not be the case. Also note that the numbers denoting the three principal stresses have nothing to do with the spatial subscripts we briefly used in eq. 5.1. They simply refer to the largest, the intermediate and the smallest of the three principal stresses.

In the stress diagrams of this book the coordinates are generally drawn parallel to the principal stress directions so that we can write:  $\sigma = \sigma'$ . Fortunately, in the earth's crust the principal stresses are often oriented roughly parallel to the vertical and the horizontal directions, because the shear stresses at the earth's surface (e.g. by wind) and at the base of the lithosphere (e.g. by mantle convection) are both negligible.

The stress tensor has a series of invariants that are independent of the choice of the coordinate system. For example the trace (the sum of the diagonal terms) or the determinant of the stress tensor is constant regardless how the coordinate system is oriented. There are three important invariant quantities. The first is given by the sum of the three normal stresses:

$$I_1 = \sigma_{xx} + \sigma_{yy} + \sigma_{zz} = \sigma_1 + \sigma_2 + \sigma_3 . \quad (5.3)$$

Note that pressure is directly related to the first invariant (see eq. 5.5). The second invariant of the stress tensor is given (in terms of the principal stresses) by:

$$I_2 = -(\sigma_2\sigma_3 + \sigma_3\sigma_1 + \sigma_1\sigma_2) ,$$

and the third invariant is given by the determinant of the tensor, which – again in terms of the principal stresses – is:

$$I_3 = \sigma_1\sigma_2\sigma_3 . \quad (5.4)$$

The 2nd and 3rd invariants are relevant for the understanding of some fundamental geodynamic processes (e. g. sect. 6.2.2).

**Mean Stress.** The *mean stress*  $\sigma_m$  is given by the mean of the three principal stresses. It is therefore independent of the coordinate system:

$$\sigma_m = P = \frac{\sigma'_{xx} + \sigma'_{yy} + \sigma'_{zz}}{3} = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3} . \quad (5.5)$$

The mean stress is also called pressure  $P$ . Strictly speaking, the mean stress is the mechanical definition of pressure, while a chemist or thermodynamicist would say that work is the product of pressure and volume change and that, therefore, pressure has the units of energy per volume ( $1 \text{ Pa} = 1 \text{ J m}^{-3}$ ). The most common place where geologists encounter these non-intuitive units for pressure is when looking up the molar volumes of mineral phases. These are generally quoted in the units Joule per bar. As the volume change may be highly anisotropic in an anisotropic stress state, chemically defined pressure may be determined by integrating the volume change over the surface of a unit volume. *Chemical* and *mechanical* pressure only correspond in an isotropic state of stress. Earth scientists measure pressure using geobarometers (sect. 7.2.1). Many geobarometers rely on the pressure sensitivity of chemical equilibria. It is therefore not clear if we measure mechanical or chemical pressure with them. This is some of the reason why the relationship between geobarometrically measured pressure and depth is often discussed in the literature (e.g. Harker 1939; Wintsch and Andrews 1988). However, in the following we assume that the differences between chemical and mechanical pressure are so small (if any) that we need not be concerned with them.

**Differential Stress.** Differential stress is a scalar value defined as the difference between the largest and the smallest principal stress:

$$\sigma_d = \sigma_1 - \sigma_3 . \quad (5.6)$$

It is a measure of how far the stress state deviates from the isotropic state. As such, differential stress relates directly to deviatoric stress  $\tau$ . In fact, we will see that  $\tau_1 = \sigma_d/2$  and that  $\tau_3 = -\sigma_d/2$ . During viscous (ductile) deformation, the application of any differential stress will cause permanent deformation. Of course, differential stresses also arise when the deformation is not permanent, for example during elastic deformation. In the brittle regime, deforming rocks will only record the differential stress that causes failure (i.e. if it is large enough to touch the failure envelope on the Mohr circle).

**Deviatoric Stress.** Unlike mean stress, pressure or differential stress, *deviatoric stress* is not a single number, but a tensor, denoted commonly with  $\tau$ . This tensor is defined by the deviation of the stress tensor in a general coordinate system (i.e. eq. 5.1) from the mean stress (i.e. pressure):

$$\tau = \begin{pmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} - P & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} - P & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} - P \end{pmatrix} . \quad (5.7)$$

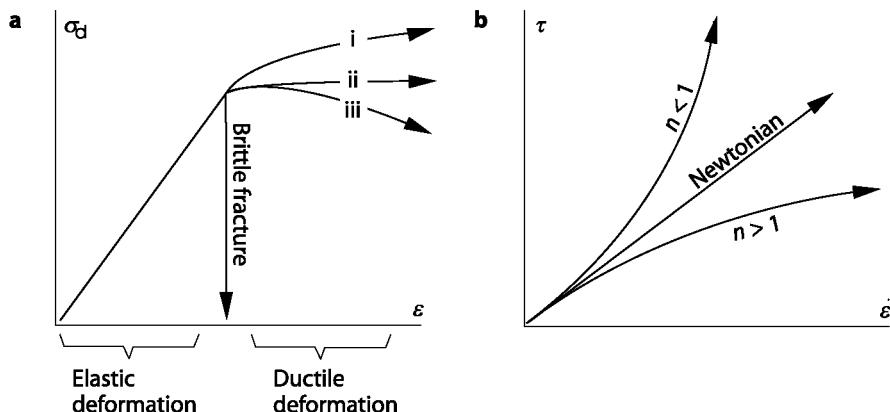
It can be seen that the total stress tensor is the sum of the isotropic stress tensor plus the deviatoric stress tensor:

$$\begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} = \begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{pmatrix} + \begin{pmatrix} \sigma_{xx} - P & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} - P & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} - P \end{pmatrix} . \quad (5.8)$$

This equation may also be written in short as:

$$\sigma = P\mathbf{1} + \tau . \quad (5.9)$$

where  $\mathbf{1}$  is the identity matrix. In the viscous regime, only deviatoric stresses (right hand term in eq. 5.8) cause deformation. In contrast, elastic deformation occurs in response to the total stress state as described by the left hand side of eq. 5.8 (Fig. 5.2).



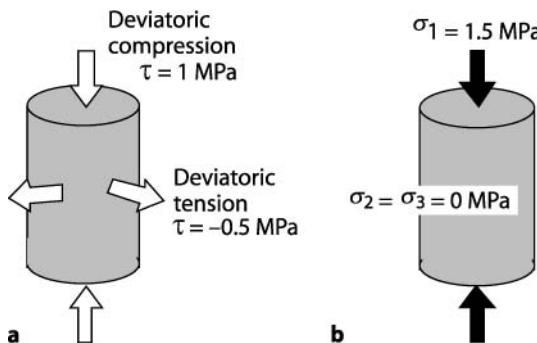
**Figure 5.2.** **a** Relationship between stress  $\sigma$  and strain  $\epsilon$  for elastic, plastic and brittle deformation. Curve *ii* is for the ideal case of plastic (ductile) deformation; *i* is with strain hardening; *iii* with strain softening. **b** Relationship between stress and strain rate  $\dot{\epsilon}$  for three different viscous materials. For a Newtonian fluid this relationship is linear. The number  $n$  is the power law exponent which we will discuss in some detail in sect. 5.1.2

For a coordinate system parallel to the principal stress directions the deviatoric stress tensor may simply be written as (in analogy to eq. 5.2):

$$\tau' = \begin{pmatrix} \tau'_{xx} & 0 & 0 \\ 0 & \tau'_{yy} & 0 \\ 0 & 0 & \tau'_{zz} \end{pmatrix} = \begin{pmatrix} \sigma_1 - \sigma_m & 0 & 0 \\ 0 & \sigma_2 - \sigma_m & 0 \\ 0 & 0 & \sigma_3 - \sigma_m \end{pmatrix} . \quad (5.10)$$

In this book, just about all discussed examples are for orientations of the stress field where  $\tau = \tau'$ . Also, we usually use many terms related to the stress tensor as if they were a scalar, i.e. we use them in equations that involve no tensor calculus. This is because most problems discussed in this book are reduced – for simplicity – to mere one-dimensional scenarios where there is only a single direction in which stresses are non-zero.

- *Simple examples for the use of deviatoric stress.* The deviatoric stress tensor is important as its components cause viscous deformation. The absolute magnitude of the deviatoric stress tensor components indicates how rapidly a rock will strain (deform). A rock will *extend* in the direction in which the deviatoric stress components are negative (negative is tensional in the earth science convention), even if all the principal stresses indicate compression (s. Fig. 5.3). Thus, when making cartoons of a field terrain it is always most instructive to sketch arrows for the principal components of the deviatoric stress tensor onto them, as their magnitude and direction corresponds to what is observed kinematically in the field (s. Fig. 6.26). Two rocks from different crustal levels may suffer the same deviatoric stresses and therefore deform similarly, but they may be in completely different states of total stress (Problem 5.4).



**Figure 5.3.** Cartoon illustrating a typical uniaxial deformation experiment. The state of deviatoric stress of the cylinder in **a** is identical to that of the cylinder in **b**. However, in **a** is not specified what the confining pressure is. This may be an arbitrary number added to all deviatoric stress values. Nevertheless, if there is no additional confining pressure, then **a** is consistent with the pressure inside both cylinders being 0.5 MPa, although both  $\sigma_d$  and  $\sigma_1$  are 1.5 MPa. In **a** the state of stress is illustrated in terms of the components of the deviatoric stress tensor. In **b** in terms of the uniaxially applied stress. Because the experiment is uniaxial, all illustrated stress components are also principal stress components both for the deviatoric stresses in **a** and the applied stress in **b** (after Engelder 1994)

Let us consider a simple two dimensional example (there are only two principal stresses  $\sigma_1$  and  $\sigma_2$ ) where both  $\sigma_1$  and  $\sigma_2$  are positive (the rock is under compression) and the mean stress is  $\sigma_m = (\sigma_1 + \sigma_2)/2$ . Then, according to eq. 5.8, the principle components of deviatoric stress are:  $\tau_1 = \sigma_1 - \sigma_m =$

$(\sigma_1 - \sigma_2)/2$  and  $\tau_2 = \sigma_2 - \sigma_m = -(\sigma_1 - \sigma_2)/2$ . We can see that the largest principal component of the deviatoric stress tensor is  $\tau_1 = \sigma_d/2$  and the smallest is  $\tau_2 = -\sigma_d/2$ . We can summarize this information with a more applied example of a continent that is under horizontal compression where the principle components of stress  $\sigma_{xx}$  and  $\sigma_{zz}$  are parallel to the horizontal ( $x$ ) and vertical ( $z$ ) directions. There, the principle horizontal deviatoric stress is:

$$\tau_{xx} = \sigma_{xx} - \sigma_m = \frac{\sigma_{xx} - \sigma_{zz}}{2} = \frac{\sigma_d}{2} . \quad (5.11)$$

Fig. 5.3 illustrates the state of stress of a rock cylinder as expressed in terms of the deviatoric stresses (left) and total stresses (right).

**Strength.** *Strength*, *failure strength* or *shear strength* are terms used to describe the critical value which the differential stress must reach to cause permanent deformation. As such it is a material property. In the *elastic* regime, the term “strength” does not really have a meaning and it is better to refer to rigidity or other terms explained in some more detail in section 4.4.2 and 5.1.2. In the *brittle* regime, the strength depends directly on the magnitude of the principal stresses and is given by the stress where the curve on Fig. 5.2a deviates from its linear course. In the *viscous* regime, all differential stresses will lead to permanent deformation (Fig. 5.2). and “strength” is dependent on strain rate. Strictly, therefore, strength has no meaning in viscous deformation either and it is better to describe viscous stresses via viscosity. Nevertheless, in this book we use the terms “strength”, “viscous strength” and “differential stress” in a similar meaning and note that they relate by a factor two:

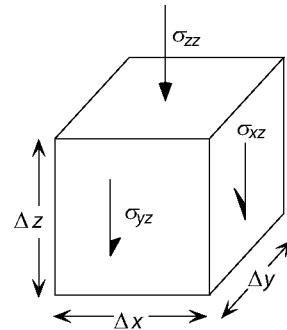
$$\text{strength} = \frac{\sigma_d}{2} . \quad (5.12)$$

**Stress Balance.** The equations describing the balance of stresses are the basics for all mechanical descriptions of deformation. A stress balance is a generalized form of Newton’s second law:

$$\text{force} = \text{mass} \times \text{acceleration} . \quad (5.13)$$

This equation is applied to a small volume of rock which may be subjected to surface forces (applied to the surfaces of the small volume, e.g. by pushing it) and body forces (applied to the small volume itself, e.g. by gravity). Eq. 5.13 has its only complication in that it is a vector equation, because force is a vector. That is, it consists of three equations each of which describe a force balance in one of the three spatial directions. Also, within each of these equations, several *surface* and *body forces* must be summed up and set equal to the product of mass times acceleration on the right hand side of the equations. Also note that the equations of force balance are generally considered per unity volume so that eq. 5.13 is usually written in terms of force/volume =

**Figure 5.4.** Different surface forces acting on a unity cube in the  $z$  direction. If  $z$  is the vertical, then there is also a body force due to gravity of the magnitude  $\rho g$ . At rest, each of the three labeled forces is compensated by a force of equal magnitude but opposite direction. Other forces in the  $y$ - and  $x$ -directions are not labeled for clarity



density  $\times$  acceleration. The different forces acting in the  $z$  direction can be summed up from Fig. 5.4. The sum is:

$$\begin{aligned}
 & \left( \sigma_{zz} + \frac{\partial \sigma_{zz}}{\partial z} \Delta z \right) \Delta x \Delta y - \sigma_{zz} \Delta x \Delta y + \left( \sigma_{zx} + \frac{\partial \sigma_{zx}}{\partial x} \Delta x \right) \Delta y \Delta z \\
 & - \sigma_{zx} \Delta y \Delta z + \left( \sigma_{zy} + \frac{\partial \sigma_{zy}}{\partial y} \Delta y \right) \Delta x \Delta z - \sigma_{zy} \Delta x \Delta z - \rho g \Delta x \Delta y \Delta z \\
 & = \rho a_z \Delta x \Delta y \Delta z , \tag{5.14}
 \end{aligned}$$

where  $x$ ,  $y$  and  $z$  are the three spatial directions,  $\rho$  is density,  $g$  is acceleration due to gravity and  $a_z$  is the acceleration of the body in the  $z$  direction. Even if this equation appears enormously complicated, it should be easy to follow it using Fig. 5.4. It simply states that the difference in forces between any two sides of a unity cube result in acceleration. We can see that the equation has six similar looking terms on the left hand side. Every group of two terms describes a difference between the force on one side of the cube (e.g. the 2nd term in eq. 5.14:  $\sigma_{zz} \Delta x \Delta y$ ) and the force on the opposite side of the cube (e.g. the 1st term in eq. 5.14:  $(\sigma_{zz} + (\partial \sigma_{zz} / \partial z) \Delta z) \Delta x \Delta y$ ). If these are equally large, then the body does not accelerate. If this difference is finite, then the body accelerates with the rate written on the right hand side of the equation:  $\rho \frac{\partial u_z}{\partial t} \Delta x \Delta y \Delta z$ . Writing eq. 5.14 a bit shorter we can write:

$$\frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} - \rho g = \rho a_z . \tag{5.15}$$

- *Navier-Stokes equation.* Eq. 5.15 (plus its corresponding versions for the  $x$  and  $y$  directions) can be reformulated into the famous Navier-Stokes equation if it is coupled with a viscous flow law. We will encounter this in eq. 5.42, but use it already here briefly to introduce this important equation. We also need the definition of strain rate from eq. 4.6 and the relationship between the stress and deviatoric stress tensors shown in eq. 5.9. Then, we can insert the definition of strain rate into eq. 5.42 and that into eq. 5.9. The resulting description of stress is then differentiated according to eq. 5.15 and we arrive

at the Navier-Stokes equation for an incompressible medium with constant viscosity:

$$-\nabla P + \eta \nabla^2 u = \rho a + \rho g , \quad (5.16)$$

In this form, we have placed both acceleration terms (that due to surface forces and that due to body forces) on the right hand side of the equation. In particular the surface force related acceleration term in eq. 5.16 (which is  $a$  instead of the earlier  $a_z$  meaning that acceleration may be in all three directions) can also be expressed in terms of velocity changes, but we need not do this here, as we will show below that is is negligible for most geological problems.

• *Force balance equations.* In most geodynamic problems, acceleration is negligible. Then, in the horizontal directions the term  $\rho \frac{\partial u_x}{\partial t} = \rho \frac{\partial u_y}{\partial t} \rightarrow 0$  and in the vertical direction  $\rho \frac{\partial u_z}{\partial t} \rightarrow \rho g$ , as gravitational force is still felt as a body force. Thus, eq. 5.14 simplifies to the following:

$$\frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial y} - \rho g = 0 . \quad (5.17)$$

Eq. 5.15 describes the equilibrium of stresses in the vertical direction and is generally applicable in the earth sciences. The first three terms of this equation are the sum of the *surface forces* acting in the  $z$ -direction, the fourth term is the *volume* or *body force* downwards. In analogous equations for the  $x$ - and  $y$ -direction this fourth term does not appear. The relationships for the  $x$ - and  $y$ -directions are:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} = 0 \quad (5.18)$$

and:

$$\frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial z} = 0 . \quad (5.19)$$

Eqs. 5.15, 5.18 and 5.19 are the basics of all mechanical equilibria discussed in this book (e. g. sect. 4.4.1; 6.2.2, s. Problem 5.5).

**Tectonic Relevance of Momentum.** In this section we discuss the nature and relevance of momentum in tectonic processes to show that momentum is practically negligible to most geological problems. The momentum of a body  $I$  is given by the product of its mass  $m$  and its velocity  $v$  (sect. 2.2.4):

$$I = mv . \quad (5.20)$$

While the velocities of plate tectonic motions are very small, the mass of plates is very large and it is therefore not immediately obvious if momentum plays a role in the tectonic force balance.

Momentum is a physical quantity that is preserved: During the collision of two plates the momentum of the entire system remains constant. However, the momentum of one of the plates may be transferred to the other. This transfer of momentum occurs by a force. The magnitude of this force is given by the change of momentum  $\Delta I$  per time  $\Delta t$  that occurs during the slowing of plate motion due to collision.

$$F = \frac{\Delta I}{\Delta t} . \quad (5.21)$$

If a plate is slowed down due to collision very abruptly, then the force is large. If it slows over a large time period, the force is small. The slowing of a plate has also the consequence that its kinetic energy  $E_k$  decreases. Kinetic energy is given by the integrated momentum integrated over the change in velocity:

$$E_k = \frac{mv^2}{2} . \quad (5.22)$$

Let us now check if momentum, kinetic energy and the forces that arise from them could be responsible for the acceleration or slowing of plates. For this, let us hypothesize that the slowing of a continental collision is caused by a waning momentum. Lets use the India-Asia collision and make some very simple assumptions. The area of the Indo-Australian Plate is roughly  $A = 5 \cdot 10^6 \text{ km}^2$ . If the mean plate thickness is  $z_1 = 100 \text{ km}$  and the mean density is  $\rho = 3000 \text{ kg m}^{-3}$  then its mass is:  $m = Az_1\rho = 1.5 \cdot 10^{21} \text{ kg}$ . If the relative plate velocity between India and Asia is  $v = 0.1 \text{ m y}^{-1} \approx 3.2 \cdot 10^{-9} \text{ m s}^{-1}$ , then, according to eq. 5.20, the momentum of the collision is:  $I = mv = 4.7 \cdot 10^{12} \text{ kg m}^{-1} \text{ s}^{-1}$  and the kinetic energy of the Indian Plate is:  $E_k = 7.6 \cdot 10^3 \text{ J}$ . If we now assume that the Indian Plate will be brought to a complete halt within one million year of the collision, then:  $F = 4.7 \cdot 10^{12} / 3.15 \cdot 10^{13} \approx 0.15 \text{ N}$ . Distributed over almost 5000 km of collision length this leaves only about:  $3 \cdot 10^{-8} \text{ N m}^{-1}$ . We can infer that plates would have to be brought to a halt within fractions of a second of a collision in order for momentum to have any influence on the orogenic force balance. In short, momentum is negligible in plate tectonics.

Indeed, we know that the relative rate of motion between the Indian and the Asian plates has hardly changed since the Tertiary. We can conclude that any changes in velocity are due to changes in the boundary conditions and have nothing to do with momentum. It is for this reason that force balance equations in the earth science (e.g. eq. 6.25) need not consider acceleration terms.

**The Difference Between Lithostatic and Non-Lithostatic.** The pressure measured by petrologists with geobarometers in metamorphic rocks is generally interpreted as the “burial pressure”, that is, the pressure is directly correlated with the depth of the rocks at the time of metamorphism according to eq. 7.1 (s. sect. 7.1.1). This interpretation is based on the assumption

that rocks have negligible strength, i.e. they cannot support any differential stress. This state of stress is called *lithostatic*. In this state, the *lithostatic pressure* is of the same magnitude as each of the principal stresses (eq. 5.5).

The state of stress is isotropic. An every day example is the state of stress inside fluids which have negligible strength, for example a glass of beer. The force exerted by the beer onto the outside glass is exactly as large as the weight of the vertical section of beer lying above this point (s. Problem 5.6). However, if we consider a more general state of stress (i.e. a material that *can* support differential stresses), then we can see that only part of the pressure is caused by depth. The magnitude of the difference  $\sigma_1 - \sigma_3$  also contributes to pressure and the orientation of  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$  determines how pressure relates to depth. The discussion on the magnitude of the differential stress  $\sigma_1 - \sigma_3$  in rocks can be summarized under the term “tectonic overpressure” (Rutland 1965; Ernst 1971) and will be discussed in a bit more detail in s. sect. 6.3.5.

For some special orientations of a general stress field it is possible to divide pressure into a lithostatic and a non lithostatic component. Such a division helps to illustrate the different contributions to pressure and allows us to estimate the magnitude of differential stresses under different boundary conditions. In a stress field where  $\sigma_1$  and  $\sigma_3$  are the maximum and minimum principle stresses and are oriented horizontally and vertically, respectively we can write:

$$P = \frac{\sigma_1 + \sigma_3}{2} = \sigma_3 + \frac{\sigma_1 - \sigma_3}{2} = \sigma_{\text{lith}} + \frac{\sigma_d}{2} = \rho g z + \tau_1 . \quad (5.23)$$

There,  $\sigma_{\text{lith}}$  is the component of pressure caused by the weight of the overlying rock column, and the non-lithostatic component is given by the largest principal component of the deviatoric stress tensor. Of course, eq. 5.23 is only valid if  $\sigma_2 = (\sigma_1 + \sigma_3)/2$ . For other values of  $\sigma_2$ , or for differently oriented stress fields, this simple subdivision in lithostatic and non-lithostatic terms of pressure is not possible and non-lithostatic components of pressure can only be calculated from the complete tensor. In sect. 6.3.5 we discuss how large such tectonic contributions to pressure might be.

### 5.1.2 Deformation Laws

For a mechanical description of deformation we need a mathematical rule that relates stress (or force) to strain (or strain rate). Such a relationship is called a *flow law*, *deformation law* or *constitutive relationship*. If we know the flow law for a given rock, then we can use the relationships of mechanical equilibrium (eqs. 5.14 to 5.19) to describe the deformation of rocks in response to an applied force.

On a microscopic scale, structural geologists discriminate between a large number of deformation laws. On geological time scales and lithospheric length scales, most geologists use only one of five terms to describe deformation mechanisms (Table 5.1):

- brittle deformation
- plastic deformation
- ductile deformation
- elastic deformation
- viscous deformation.

These five terms have very different meanings and some of them are very rigorously defined and others are not. We will deal with these five terms at length over the next sections (s. also: Weijermars 1997; Twiss and Moores 1992; Jaeger and Cook 1979). Here we summarize some of the most important information on these five terms and how they relate stress to strain in general terms.

- *Brittle deformation* is no really a deformation mechanism at all. We will see below that the laws of brittle deformation only describe a stress state and not a relationship between stress and strain. The deformation law usually used to describe rocks in a brittle fashion is plastic deformation.

- *Plastic deformation* law states that a constant stress is required to deform the rock. Irregardless how much or how fast we deform, the required force is all the same. Deforming sand is a good example.

- *Ductile deformation* is a wonderful term that has no specified meaning other than that we want to say that the deformation is not elastic and not brittle. It is therefore an extremely useful term for the field geologist who does not want to specify himself by using words like “viscous”, “plastic” or “dislocation creep” - all of which have very rigorously defined meanings.

- *Elastic deformation* is the law that states that the *strain* of a rock is proportional to the applied *total stress*. As such, it is the only deformation mechanism which is not permanent: As soon as the stress is released, the strain is gone as well.

- *Viscous deformation* is the law that is most commonly used to describe ductile deformation on the crustal scale. Viscous means that the strain *rate* of a rock is proportional to the applied *deviatoric stress*.

We all have every day encounters with these two deformation mechanisms, namely with rubber bands (elastic) and mixing cake dough (viscous). When stretching a rubber band, the amount of stretch depends on how hard we pull. The more pull, the more stretch. The applied stress and the resulting strain are proportional. On the other hand it does not matter at all whether we pull fast or slow. The stretch is always the same for the same applied force, independent on the speed (strain rate) with which we do the experiment. With mixing dough its exactly opposite: It does not matter at all whether we mix it only a bit or very thoroughly (little or much strain in the dough), the needed force is always the same. However, how much force we need depends very strongly on the mixing rate. If we mix it rapidly, we need much more force than if we mix slowly.

**Elastic Deformation.** Elastic deformation is characterized by a proportionality between stress and strain (Fig. 5.2). Both these parameters are described

**Table 5.1.** Deformation laws, for more information see also Tabel 5.2

brittle		no deformation law but a stress state; usually described with plastic law	
plastic	(ductile)	constant stress; example: sand	
viscous	(ductile)	stress and strain rate are proportional	- linear (Newtonian) - non linear (power law)
elastic		stress and strain are proportional	

by tensors which each include 6 independent values. However, if this proportionality is ideally linear, and we only consider uniaxial (one-dimensional) loading then this relationship is called *Hooke's law* and may be simply written as:

$$\sigma_{xx} = E\epsilon_{xx} \quad . \quad (5.24)$$

where the subscripts emphasize that this equation is meant to be one dimensional (and not a full tensor equation). When the deformation is uni-axial, then  $\epsilon$  is the (dimensionless) longitudinal normal strain and is defined as the change in length during deformation relative to the original length ( $s$  in eq. 4.1). The proportionality constant  $E$  is called the *Young's modulus* and has the units of stress ( $\text{N m}^{-2}$ ). The Young modulus corresponds to the slope of the elastic section of the curve in Fig. 5.2. For rocks, the Young's modulus is of the order of  $10^{10}$  to  $10^{11}$  Pa. Young's modulus is a kind of a summary of the Lame elastic constants which are the elastic coefficients that occur when eq. 5.24 is written as a full tensor equation.

If more than one of the three principal stresses is larger than zero, then it is also important to consider that rocks are compressible. This is described by the Poisson constant  $\nu$ . For the largest principal stress we can write:

$$\sigma_1 = \epsilon_1 E + \nu\sigma_2 + \nu\sigma_3 \quad (5.25)$$

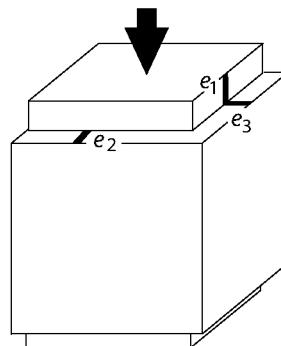
or, if strain is written as a function of stress:

$$\epsilon_1 = \frac{1}{E}\sigma_1 - \frac{\nu}{E}\sigma_2 - \frac{\nu}{E}\sigma_3 \quad . \quad (5.26)$$

For the other two spatial directions equivalent equations may be formulated. The Poisson constant is given by the ratio of two stretches, namely the infinitesimal strain normal to the applied stress and the stretch *in* direction of the applied stress (Fig. 5.5).

During compressive deformation, a rock will shorten in the direction of the applied stress. Thus, the incremental stretch  $e_1$  in Fig. 5.5 is negative. If the rock is isotropic, then this shortening is distributed evenly between expansion in the other two spatial directions. Thus,  $\nu = +0.5$  for incompressible

**Figure 5.5.** Stretch of a cube as the consequence of compression in the vertical direction. The Poisson constant is defined as  $\nu = -e_3/e_1$



materials. For example, rubber is almost incompressible and has a Poisson constant of almost  $\nu = +0.5$ . In contrast, the Poisson constant of rocks is of the order of 0.1–0.3. We can see that rocks are quite compressible in the elastic regime. However, the total strains of rocks in the elastic regime are quite small, because the Young's modulus of rocks is very large. Thus it is no obvious to us that rocks are actually quite compressible.

While the Poisson constant is directly related to how compressible a rock is, it should not be confused with the *compressibility*  $\beta$ . Under isotropic stress (with pressure  $P$ ) a rock will compress isotropically (i.e.  $\epsilon_1 = \epsilon_2 = \epsilon_3$ ) and

$$P = \frac{1}{\beta} (\epsilon_1 = \epsilon_2 = \epsilon_3) . \quad (5.27)$$

$\beta$  is related to Young's modulus and the Poisson constant by:

$$\beta = \frac{3 - 6\nu}{E} . \quad (5.28)$$

We can see from this equation that, for incompressible materials where  $\nu = +0.5$ , the compressibility becomes:  $\beta = 0$ . The inverse of the compressibility is called the *bulk modulus*:  $K = 1/\beta$ . For a more rigorous treatment of the theory of elasticity the reader is referred to the literature.

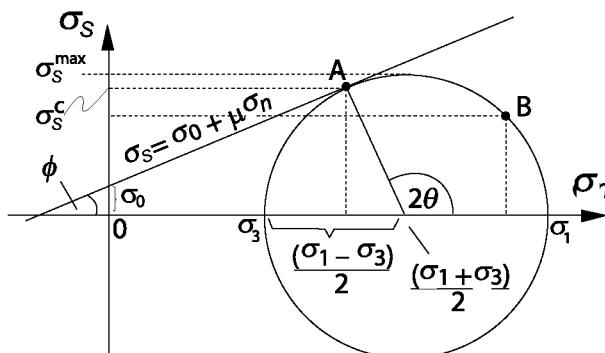
**Brittle Fracture.** When the stresses applied to rocks cannot be compensated elastically, permanent deformation will occur. This may occur by ductile or brittle processes. Among brittle processes, two different modes of brittle deformation may be discerned: rocks deform either by creating new cracks, or by friction along existing fractures. In both cases the friction along the failure planes plays a critical role. Brittle failure is commonly described with the *Mohr-Coulomb-criterion*. However, it should be said here that – strictly speaking – the Mohr-Coulomb-criterion describes only a state of stress, namely the critical state at which failure occurs. It does not place stress and strain in a relationship to each other and is therefore not a constitutive relationship or flow law.

- *Mohr-Coulomb-criterion.* Coulomb (1773) was the first to recognize that the brittle strength of materials is largely a linear function of the applied normal stress  $\sigma_n$  and that it depends only to the second order on a material constant called cohesion  $\sigma_0$ . At geological stresses cohesion is largely negligible. According to the Coulomb criterion, failure occurs when the shear stress on a given plane reaches a critical value  $\sigma_s^c$  that is a function of the normal stress acting on that plane  $\sigma_n$ , as:

$$\sigma_s^c = \sigma_0 + \mu \sigma_n . \quad (5.29)$$

The coefficient  $\mu$  that relates shear stress and normal stress on a failure plane is called the *internal coefficient of friction*. This coefficient is dimensionless. In the literature, the critical failure stress is often abbreviated with  $\tau$  (including in the first edition of this book). However, it is probably clearer if we reserve  $\tau$  exclusively for deviatoric stresses and we therefore choose a different notation in this edition: We use the subscript  $s$  for all shear stresses (for both stress:  $\sigma_s$  and deviatoric shear stress:  $\tau_s$ ) and apply additional superscripts if necessary. According to the Coulomb criterion (eq. 5.29) brittle deformation is a nearly linear function of total stress. It is independent of temperature or strain rate  $\dot{\epsilon}$  and almost independent of the material as the cohesion is almost negligible and the internal coefficients of friction are very similar for most rocks (Byerlee's law).

Mohr (1900) then discovered that the failure criterion of Coulomb may be elegantly portrayed graphically. His graphical analysis is called the Mohr diagram. In the Mohr diagram shear stresses are plotted against normal stresses



**Figure 5.6.** The relationship between normal stress (horizontal axis) and shear stress (vertical axis) in the Mohr circle. The normal stresses are compressive (positive) to the right of the origin and tensional (negative) to the left. Eq. 5.29 describes the tangent to the Mohr circles drawn around centers at  $(\sigma_1 + \sigma_3)/2$  with the radius  $(\sigma_1 - \sigma_3)/2$ . In the ductile regime shear stresses do not increase linearly with normal stresses anymore. However, for many rocks the curve is even in the brittle regime not completely linear, but slightly concave against the normal stress axis. For failure planes in a rock it is true that:  $\sigma_s^c = \sin 2\theta (\sigma_1 - \sigma_3)/2$  and  $\sigma_n = (\sigma_1 + \sigma_3)/2 - \cos(2\theta)(\sigma_1 - \sigma_3)/2$

and the stress state in a rock is plotted as a circle. Fig. 5.6 shows (for example for point *B*) that the shear stresses in a rock  $\sigma_s$  are a function of the angle  $\theta$  between the considered plane and the directions of the principal stresses. This may be formulated in terms of the equation:

$$\sin(2\theta) = \frac{2\sigma_s}{(\sigma_1 - \sigma_3)} , \quad (5.30)$$

where the angle  $\theta$  is the angle between any considered plane in a rock and the principal stress directions. From eq. 5.30 we can see that  $\sigma_s$  is the largest on planes that lie at an angle of  $45^\circ$  to the principal stress direction (i.e.  $2\theta = 90^\circ$  and  $\sin(2\theta) = 1$ ):

$$\sigma_s^{\max} = \frac{\sigma_1 - \sigma_3}{2} . \quad (5.31)$$

Thus, the maximum shear stress a rock can support is half as large as the applied differential stress. However, it is important to note that the *largest* shear stress is *not* where failure occurs. From Fig. 5.6 we can see that the normal stress at  $\sigma_s^{\max}$  is just a little bit larger (it is:  $\sigma_n = -(\sigma_1 + \sigma_3)/2$  than the normal stress at the critical failure stress  $\sigma_s^c$  (at point *A*, for which the normal stress is explained in Figure caption Fig. 5.6).

The slope of the tangent to the Mohr circles in Fig. 5.6 is given by the internal *angle of friction*. This angle of friction  $\phi$  and coefficient of friction  $\mu$  are related by:

$$\tan\phi = \mu . \quad (5.32)$$

For most rocks this angle is about  $30\text{--}40^\circ$ , which is equivalent to an internal coefficient of friction between roughly 0.6 and 0.85. This relationship is called Byerlee's law as he was the first to measure  $\mu$  on a crustal scale and derived  $\phi$  from  $\mu$ .

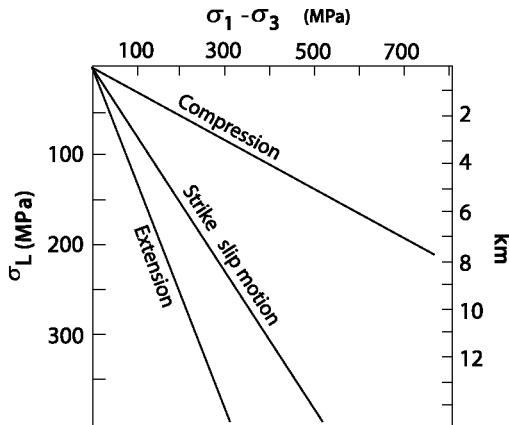
If fluid pressure plays a role, then the Mohr-Coulomb criterion is often formulated as:

$$\sigma_s^c = \sigma_0 + \mu(\sigma_n - P_f) , \quad (5.33)$$

where  $P_f$  is the pore fluid pressure. This can be approximated as  $\sigma_s^c = \sigma_0 + \mu\sigma_n(1 - \lambda)$ , where  $\lambda = P_f/\sigma_L$  is the ratio of pore fluid pressure to lithostatic stress,  $\sigma_L$ , if  $\sigma_n \approx \sigma_L$ . If both are of the same magnitude, then  $\lambda = 1$  and the shear stress necessary for failure is only a function of the cohesion. If there is no fluid, then eq. 5.33 reduces to eq. 5.29.

- *Byerlee's and Amonton's laws.* If preexisting cracks occur in a rock then there is no cohesion. To be more precise: the remaining cohesion is negligible compared to the cohesion of an intact rock. The shear stresses needed to deform a rock only need to overcome the coefficient of friction and the normal stresses applied to the rock. Eq. 5.29 simplifies to:

**Figure 5.7.** Brittle failure as a function of depth and normal stress or “lithostatic pressure”, as calculated with 5.38, 5.39 and 5.40



$$\sigma_s^c = \mu \sigma_n . \quad (5.34)$$

This equation is usually called *Amonton's law*. Byerlee (1968; 1970) showed empirically that, at pressures below 200 MPa, (roughly less than 8 km) the crust may be characterized by an internal coefficient of friction around 0.85:

$$\sigma_s^c = 0.85 \sigma_n . \quad (5.35)$$

At larger depths, but above the brittle ductile transition brittle failure in the crust appears to be best described by:

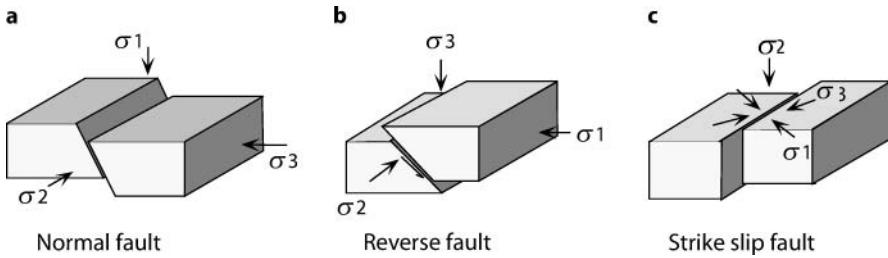
$$\sigma_s^c = 60 \text{ MPa} + 0.6 \sigma_n . \quad (5.36)$$

These empirical relationships are called *Byerlee's laws* (Fig. 5.7). Because of the fact that  $\mu \approx 0.85$ , most faults occur at 30 degrees angle to the maximum principle stress. Byerlee's laws state that rocks at 5 km depth will fail at roughly 110 MPa, in 10 km depth at roughly 230 MPa and in 15 km depth at about 300 MPa. If we want to consider fluid pressure as well we can reformulate Byerlee's laws to:

$$\sigma_s^c = 0.85 \sigma_n (1 - \lambda) \quad \text{and} \quad \tau_c = 60 \text{ MPa} + 0.6 \sigma_n (1 - \lambda) . \quad (5.37)$$

- *Anderson's theory.* Byerlee's laws describe the relationship between shear stresses (in particular  $\sigma_s^c$ ) and normal stresses (which relate to depth in the crust) in general, but they do not explain the spatial orientation of failure planes relative to the *principal stresses*.

Anderson (1951) reformulated the Mohr-Coulomb law in terms of differential stress ( $\sigma_1 - \sigma_3$ ) and lithostatic stress ( $\sigma_L = \rho g z$ ) instead of shear stress and normal stress. He took three cases to represent reverse faults, normal faults and strike slip faults and assumed that  $\sigma_L = \sigma_3$ ,  $\sigma_L = \sigma_1$  and  $\sigma_L = \sigma_2 = 0.5(\sigma_1 + \sigma_3)$  for each of these cases, respectively (Fig. 5.8). Then, the differential stress as a function of depth (or  $\sigma_L$ ) may be written as follows:



**Figure 5.8.** Geometry of a brittle fault during: **a** extension, **b** compression and **c** strike slip motion

$$\text{reverse faults : } \sigma_1 - \sigma_3 = \frac{2(\sigma_0 + \mu\sigma_L(1 - \lambda))}{\sqrt{\mu^2 + 1} - \mu} , \quad (5.38)$$

$$\text{normal faults : } \sigma_1 - \sigma_3 = \frac{-2(\sigma_0 - \mu\sigma_L(1 - \lambda))}{\sqrt{\mu^2 + 1} + \mu} , \quad (5.39)$$

$$\text{strike slip faults : } \sigma_1 - \sigma_3 = \frac{2(\sigma_0 + \mu\sigma_L(1 - \lambda))}{\sqrt{\mu^2 + 1}} . \quad (5.40)$$

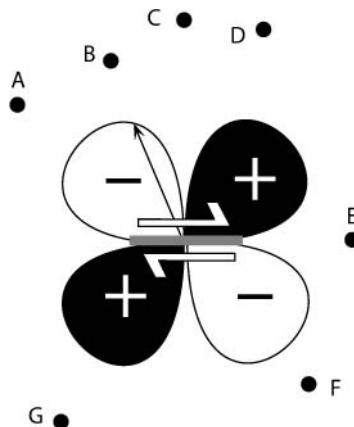
The relationship for each case can be directly derived from Fig. 5.6 using trigonometric relationships (e. g. Weijermars 1997). Fig. 5.7 shows these three linear relationships for  $\mu = 0.85$  and  $\sigma_0 = 0$ .

• *Interpretation of fault orientations.* The spatial orientation of faults in the earth's crust is one of the most important pieces of evidence for the interpretation of the magnitude and geometry of the stress field in the lithosphere (s. sect. 6.2; Zoback 1992). If the orientation of faults can be measured directly in outcrop, then there is a series of statistical methods that may be used to derive the palaeostress field at the time of fault formation (Angelier 1984; 1994). For the interpretation of the *present day* stress field of the lithosphere the most important methods are:

- Interpretation of seismic data,
- Interpretation of bore hole break outs (e. g. Bell and Gough 1979; Mastin 1988; Wilde and Stock 1997),
- direct in-situ- measurement of stresses (Zoback and Haimson 1983),
- Interpretation of GPS data (Global Positioning System; e. g. Argus and Heflin 1995),
- interferometric methods (e. g. Molnar and Gibson 1996).

The application of interferometric methods has only been possible since they are accurate enough to measure plate motions directly. However, strictly speaking, these methods (and those using GPS data) measure motions rather than stresses. Data from all these methods are collected in the world stress data base <http://www-gpi-physik.uni-karlsruhe.de/wsm/> ( see also Fig. 2.2).

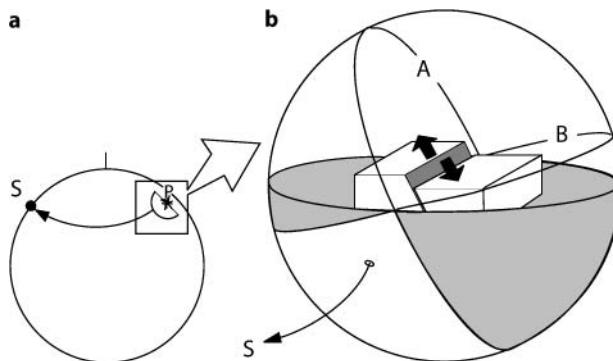
**Figure 5.9.** Polarity and intensity of P-waves in two dimensions. The two planes along which there are no P-waves are called nodal lines or nodal planes



The most reliable data source for the interpretation of the intra plate stress field remains the interpretation of seismic data, in particular because it allows to characterize processes that occur deep inside the lithosphere (Fig. 2.2). Thus, they can not only be used to determine the stress field in two dimensions, but also its depth dependence on a plate scale. Seismic data are commonly interpreted with the aid of *fault plane solutions* (s. e. g. Michael 1987; McKenzie 1969a).

- *Fault plane solutions.* Fault plane solutions illustrate the qualitative direction of motion (*polarity*) of seismic P-waves (primary longitudinal waves) on a Schmidt net (s. Figs. 2.11, 5.11). In order to be able to do this, a seismic event must be recorded by a large number of seismic stations around the globe. Fig. 5.9 illustrates how such data are then used to plot a fault plane solution using a two-dimensional example. The gray shaded bar and the arrows in the middle of this figure symbolize a dextral fault, and we assume that the center of the seismic event lies in the middle of the figure. Points A to G are seismic stations in the region that have recorded the event. The very first motion along the fault as measured by the seismic stations is either tensional or compressional. According to their position relative to the epicenter, the stations A, B and F will register a *tensional* first motion, while stations G and D will register a *compressional* first motion. Stations C and E will record no P-waves at all. These results are illustrated by the clove leaf around the center of the diagram. Tensional regions are colored white and characterized by a “-”, compressional regions are drawn black and characterized by a “+”. The distance of the clove leaf circumference from the middle indicates the intensity of the measured P-waves. The arrow from the origin towards station B is a vector indicating direction and intensity of the P-waves measured at B.

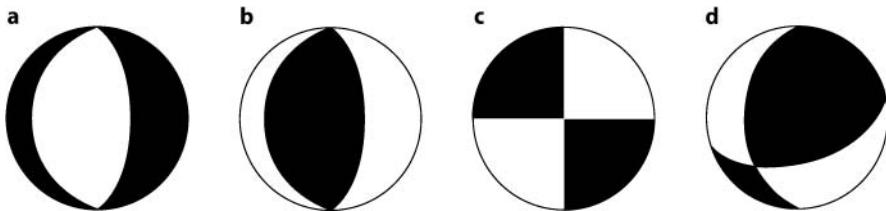
For the interpretation of a fault in three-dimensional space, we cannot restrict ourselves to the two-dimensional cartoon of Fig. 5.9 and it is necessary to use data from seismic stations around the globe. Fig. 5.10a shows



**Figure 5.10.** a Schematic cross section through the earth. Point  $P$  marks the occurrence of a seismic event. Point  $S$  marks a seismic station that has recorded the event. b Enlarged section from a. The polarity of all P-waves emanating downwards from point  $P$  is plotted in a Schmidt net surrounding the fault. The fault itself intersects the Schmidt net along line  $A$ . Plane  $B$  is the normal to this plane and to the movement direction.  $A$  and  $B$  are called nodal planes

a schematic cross section through earth. Now consider a seismic event that occurs at point  $P$ . From there, P-waves will begin to propagate in all directions. Because of the refraction of waves through the interior of the globe, the wave measured at point  $S$  will have departed from point  $P$  almost straight downwards. In order to know *exactly* which direction this wave has left from point  $P$ , we need to know the curvature of the wave through earth. However, this is known and may be looked up in seismic tables. It is therefore possible to plot the relative position of all seismic stations that have recorded the event into a Schmidt net. All stations in the vicinity of the epicenter (i.e. on a tangential plane to earth at the epicenter) will plot near the circumference of the net. For these stations we could use the two-dimensional illustration of Fig. 5.9. Fig. 5.10b is an enlarged section from Fig. 5.10a. The polarity of the P-waves on the projection hemisphere is indicated by the light and dark shading.

Fig. 5.11a shows the fault plane solution for the seismic event in the three-dimensional cartoon in Fig. 5.10b. Note that the solution has two alternative interpretations. Firstly, Fig. 5.11a may be interpreted as a normal fault with a steep eastwards dip. Then, the boundary between the right hand black colored region and the central white region is the fault plane (plane  $A$  in Fig. 5.10b) and the boundary between the left hand black region and the white region is the normal to this plane (*auxiliary plane*; plane  $B$  in Fig. 5.10b). However, the figure may alternatively be interpreted as a normal fault with a shallow westward directed dip. All fault plane solutions have two alternative interpretations for the orientation of the fault they represent. Fig. 5.11b is the fault plane solution of the reverse fault of the same orientation as in a and Fig. 5.11c is the fault plane solution of a strike slip fault. Fig. 5.11d is the fault plane solution for a fault containing reverse thrust as well as strike slip



**Figure 5.11.** Some examples of fault plane solutions. **a** Fault plane solution for a normal fault that dips steeper than  $45^\circ$  towards the right (east) *or* shallower than  $45^\circ$  towards the left (west). **b** Fault plane solution for a reverse fault of the same orientations as in **a**. **c** Fault plane solution of a north-south striking dextral- *or* an east-west striking sinistral strike slip fault. **d** Fault plane solution for a reverse fault that is inclined with roughly  $45^\circ$  towards the left (west) and has a dextral strike slip component in its motion. Alternatively, this fault plane solution could be for a reverse fault that is inclined towards the south east and contains a sinistral component. Note that the polarity of the center of the fault plane solution always indicates the overall kinematics of the structure (i.e. extensional vs. compressional)

components in its motion. Fault plane solutions as illustrated in Fig. 5.11 are a common means of interpretation of geodynamic processes in active orogens (Molnar and Lyon-Caen 1989) and plate boundary processes (s. e. g. Frohlich et al. 1997).

**Viscous Deformation.** On the scale of a thin section, rocks behave not viscously but according to a large range of deformation mechanisms (e.g. grain boundary migration, diffusion creep and many others). The dependence of deformation mechanism on the physical conditions may be portrayed in *deformation mechanism maps* (Frost and Ashby 1982). In such maps parameters like temperature, grain size, stress and viscosity are plotted against each other and the diagrams are divided into different fields where different mechanisms apply. However, on a larger scale, it is useful to average different deformation mechanisms and assume that rocks behave like viscous fluids. Viscous deformation of ideal fluids is described by a proportionality between deviatoric stress and strain rate. If we make a simple shear experiment in which we compare the scalar quantities of the shear strain rate  $\dot{\gamma}$  and the shear force (per area) we require to shear it  $\tau_s$ , then we can write:

$$\tau_s = \eta \dot{\gamma} \quad . \quad (5.41)$$

If we use the full deviatoric stress tensor  $\tau$  and the full strain rate tensor  $\dot{\epsilon}$ , then this is given by:

$$\tau = 2\eta \dot{\epsilon} \quad . \quad (5.42)$$

where the factor 2 arises from the definition of strain rate (eq. 4.6). Both equations are used in the literature (Ranalli 1987). In both the proportionality constant  $\eta$  is called the *dynamic viscosity*. There is also a parameter called the *kinematic viscosity* which is the ratio of dynamic viscosity and density

and has the units of diffusivity, namely:  $\text{m}^2 \text{s}^{-1}$ . The dynamic viscosity has the units of Pascal times second (Pas) or  $\text{kg m}^{-1} \text{s}^{-1}$ . For air it is roughly  $10^{-5}$  Pas, the viscosity of water is roughly  $10^{-3}$  Pas, the viscosity of ice roughly  $10^{10}$  Pas, of salt  $10^{17}$  Pas and of granite it is roughly  $10^{20}$  Pas. If  $\eta$  is constant with respect to strain rate then eq. 5.42 is linear. A fluid that behaves according to such a linear relationship is called a *Newtonian fluid*. Eq. 5.42 states that the larger the deviatoric stress that is applied, the faster the rock will deform. Note that, in the orientation of the *maximum* shear strain rate, the stress in eq. 5.42 will be  $\tau_{\max}$ , which is equivalent to half of the differential stress  $\sigma_d/2 = (\sigma_1 - \sigma_3)/2$ , (s. eq. 5.31). There are two reasons why rocks typically don't deform according to the simple form of eq. 5.41 with a constant viscosity:

- 1. *The Arrhenius relationship.* Viscosity is extremely strongly temperature dependent. This temperature dependence is described by the *Arrhenius relationship*:

$$\eta = A_0 e^{Q/RT} . \quad (5.43)$$

In this relationship the constants  $A_0$  and  $Q$  are material-specific constants called the *pre exponent constant* and the *activation energy* (in  $\text{J mol}^{-1}$ ), respectively. The parameter  $R$  is the universal gas constant and  $T$  is the absolute temperature. If we try to read eq. 5.43 we can see that it states that the viscosity of any material will trend towards infinity at absolute zero and will decrease exponentially from there to approach the value  $A_0$  asymptotically at high temperatures. The Arrhenius relationship will be discussed in some more detail in sect. 7.2.2 in connection with eq. 7.5. Do not forget to always use absolute temperature when performing calculations with eq. 5.43.

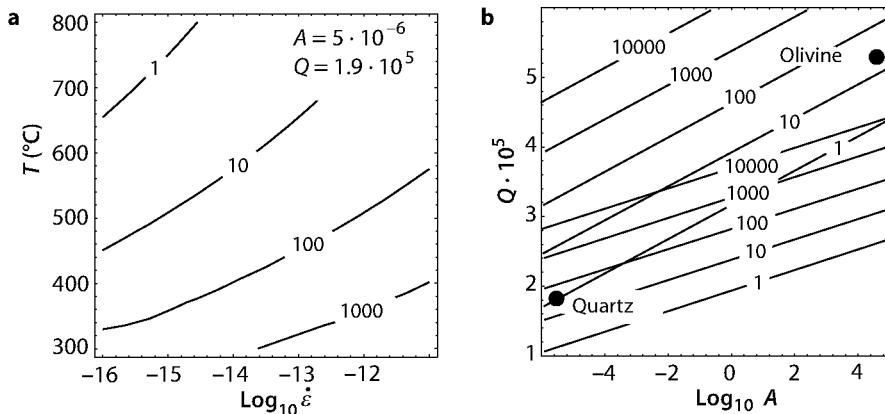
- 2. *Non linearity.* Rocks rarely deform as a Newtonian fluid (i. e. there is rarely a linear relationship between the applied deviatoric stress and strain rate). In fact, many rocks deform roughly 8 times as rapid if the applied stress is doubled. More generally, this may be written in terms of a power law relationship:

$$\tau_s^n = A_{\text{eff}} \dot{\gamma} . \quad (5.44)$$

There, the exponent  $n$  is called the *power law exponent*. It is a material constant and is between 2 and 4 for many rock types. The parameter  $A_{\text{eff}}$  is a material constant. It is analogous to  $\eta$  in eq. 5.42, but does not have the units of viscosity and we therefore use a different symbol.  $A_{\text{eff}}$  has the units  $\text{Pa}^n \text{s}$ . However, in analogy to a Newtonian fluid, it is possible to derive an *effective viscosity* from eq. 5.44 which is given by the ratio of deviatoric stress and strain rate. This is:

$$\eta_{\text{eff}} = \frac{\tau_s}{\dot{\gamma}} = A_{\text{eff}}^{1/n} \times \dot{\gamma}^{(1/n)-1} . \quad (5.45)$$

We will meet this relationship again in eq. 6.29.



**Figure 5.12.** Differential stress (in MPa) during viscous deformation as a function of a range of parameters as calculated with eq. 5.46. A power law exponent of  $n = 3$  was assumed. **a** Differential stress as a function of temperature and strain rate for the material constants of quartz. **b** Differential stress as a function of activation energy  $Q$  (in  $\text{J mol}^{-1}$ ) and pre exponential constant  $A$  (in  $\text{MPa}^{-3} \text{s}^{-1}$ ). Continuous lines are for 500 °C, dashed lines are for 1 000 °C. The assumed strain rate is  $\dot{\epsilon} = 10^{-13} \text{ s}^{-1}$ . The rheological data for quartz and olivine from sect. 5.2.1 are plotted

- *General viscous flow law for the lithosphere.* If we want to apply a non-linear viscous relationship like eq. 5.44 to rocks, it is useful to couple it with the Arrhenius relationship. However, because of the parameter  $A_{\text{eff}}$  does not really have the units of viscosity, and because experiments are typically not performed as shear experiments (where  $\dot{\gamma}$  is measured), viscous flow laws are usually formulated somewhat differently. Typically, they are formulated as a relationship between differential stress ( $\sigma_1 - \sigma_3$ ) and longitudinal strain rate  $\dot{\epsilon}_l$ , as it is measured in uniaxial shortening experiments. Such an empirical relationship is called *Dorn's law* and is typically written as:

$$(\sigma_1 - \sigma_3) = \left( \frac{\dot{\epsilon}_l}{A} \right)^{(1/n)} e^{(\frac{Q}{nRT})} \quad \text{or :} \quad \dot{\epsilon}_l = \sigma_d^n A e^{-\frac{Q}{nRT}}, \quad (5.46)$$

(e.g. Sandiford et al. 1991; Sonder and England 1989 or Houseman and England 1986). The three material constants  $A$ ,  $Q$  and  $n$  are constrained by series of experiments performed at constant strain rate and temperature (e.g. Gleason and Tullis 1995). For exponents larger than 1, Dorn's law is also called simply *power law*. Note that the constant  $A$  has (in contrast to  $A_{\text{eff}}$ ), the units of  $\text{Pa}^{-n} \text{s}^{-1}$  and incorporates the factor 2 we encountered in eq. 5.42. Also note again that – if you want to use eq. 5.46 to estimate the lithospheric strength or non lithostatic contributions to pressure – only half of the differential stress contributes to pressure (e.g. eq. 5.12; eq. 5.23 or: eq. 6.33).

For mechanical models in which temperature is not considered explicitly, it is useful to summarize the temperature dependent terms of eq. 5.46. We

**Table 5.2.** Dependence of brittle and viscous deformation on some physical parameters

dependent on	brittle	viscous
total pressure (depth)	yes (linear)	no
material	no	yes ( $\approx$ power of 3)
strain rate	no	yes
temperature	no	yes (exponential)

will discuss this in some more detail in sect. 6.2.2 on p. 304. Fig. 5.12a shows the differential stress as a function of temperature and strain rate for the material constants of quartz. Fig. 5.12b shows differential stress as a function of activation energy and pre exponent constant at fixed temperatures and strain rates.

Dorn's law is an empirical deformation law and in some cases it is necessary to modify it empirically. One example where this is necessary is the deformational behavior of olivine. Fig. 5.12b shows that olivine deforms at 500 °C (which may be a realistic assumption for the Moho-temperature) only at unrealistically high stresses around 10<sup>7</sup> MPa if it were described with eq. 5.46. Thus, Goetze (1978) and Goetze and Evans (1979) suggested that a better description of the behavior of olivine above 200 MPa is given by the relationship:

$$(\sigma_1 - \sigma_3) = \sigma_D \left( 1 - \sqrt{\frac{RT}{Q_D} \ln \left( \frac{\dot{\epsilon}_D}{\dot{\epsilon}_i} \right)} \right) . \quad (5.47)$$

There,  $Q_D$  is again an activation energy,  $\sigma_D$  a critical stress that must be exceeded and  $\dot{\epsilon}_D$  is the critical strain rate (s. Table 5.3). Comparing eq. 5.47 with eq. 5.46 shows that this law is by far not as temperature dependent as the power law. Combinations of eqs. 5.46, 5.47 and Byerlee's laws form the basics of many simple quantitative models describing the rheology of the lithosphere as a whole (Brace and Kohlstedt 1980).

Eq. 5.46 shows that the stresses during viscous deformation are strongly dependent on *temperature*, *strain rate* and *material constants*, but are independent of the confining pressure. Thus, ductile deformation is subject to completely different laws than brittle deformation (Table 5.2).

## 5.2 Rheology of the Lithosphere

Rheology is the science of the *flow characteristics of materials*. In a more general sense, *rheology* is often used as a term to describe the deformational behavior of material (in the case of geologists: rocks), independent of whether

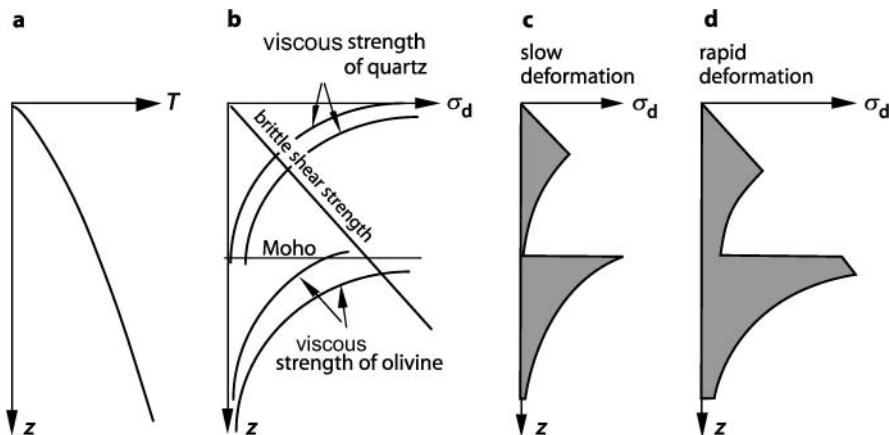
the deformation is actually by *flow* or rather by brittle fracture or other deformation mechanisms. Rheology describes the relationships between forces and motions, and between stress and strain. As such, constitutive relationships form the basics for the all rheological questions (s. p. 220).

In the previous section we showed that two of the most important deformation mechanisms in the lithosphere, namely brittle and viscous deformation, depend on very different physical parameters (Table 5.2). This is the reason for strong rheological heterogeneities in the lithosphere and also for the different rheological behavior of continental versus oceanic lithosphere. This will be the subject of the next sections.

### 5.2.1 Rheology of the Continental Lithosphere

In the late seventies of last century Brace, Goetze and others summarized much of the information from the previous sections to formulate a simple rheological model for the lithosphere. This rheological model is sketched in Fig. 5.13 and the following figures and will be the basis for our discussion in this book. Note that in these strength profiles (e.g. Figs. 5.13, 5.14 and 5.15) differential stress rather than deviatoric stress is usually plotted on the horizontal axis, because  $\sigma_d$  is a single scalar value that may be used to characterize the stress state. It is important to note throughout this chapter that  $\sigma_d/2$  corresponds to the maximum deviatoric stress (s. discussion around eqs. 5.31 and 5.42).

The strength profiles in Fig. 5.13 consist of two different types of curves. The straight lines are for brittle fracture. They show increasing rock strength with increasing depth as the normal stresses in the crust increase with depth as shown in eq. 5.29 (Fig. 5.13b, s. sect. 5.1.2). The curved lines describe viscous deformation. The strength they describe decreases exponentially downwards, because temperature *increases* with depth roughly *linearly* (s. sect. 5.1.2) and viscosity for a given mineral decreases exponentially with temperature. Each curve is for a given strain rate that is assumed to be constant over the entire lithosphere. A higher strain rate will yield a curve that has a higher strength at a given depth. Fig. 5.13b shows that, for a given strain rate, two different failure strengths may be associated with each depth. A rock at a given depth will always deform according to the deformation mechanism that requires *less* stress. Using this logic, we can draw strength profiles like those illustrated in Fig. 5.13c and 5.13d. We can see that the model predicts that, at shallow levels, rocks deform in a brittle fashion, as is familiar to us from the fact that rocks at near the surface break in brittle fashion or break (instead of flow) when we hit them with a hammer. At large depths and high temperatures, viscous deformation prevails as is familiar to us from the fact that high grade metamorphic rocks are often characterized by folding and other features indicating viscous (ductile) deformation. The depth at which brittle strength and viscous strength have the same magnitude is called the *brittle-ductile transition*. Note that the depth of this transition is strain rate dependent in this model.



**Figure 5.13.** Schematic illustration of a *Brace-Goetze lithosphere*. **a** Temperature  $T$  as a function of depth  $z$ . This curve has the shape of a typical continental geotherm. **b**, **c** and **d** show shear strength as a function of depth. **b** Shear strength due to brittle failure (straight line) and viscous deformation (curved lines) for two different strain rates and the material constants for quartz and olivine. At any given depth, the curve with the higher strength corresponds to the higher strain rates. **c** and **d** Strength profiles constructed from **b** for low and high strain rates. Integrating the shaded area yields the vertically integrated strength. The units of this integrated strength are  $\text{N m}^{-1}$ . This integrated strength may be interpreted as the force per meter length of orogen applied to the orogen in direction normal to the orogen (assuming the orogen is everywhere deforming). Note that the cartoon indicates that - at high strain rates (profile **d**) - the upper mantle will deform in a brittle fashion just below the Moho

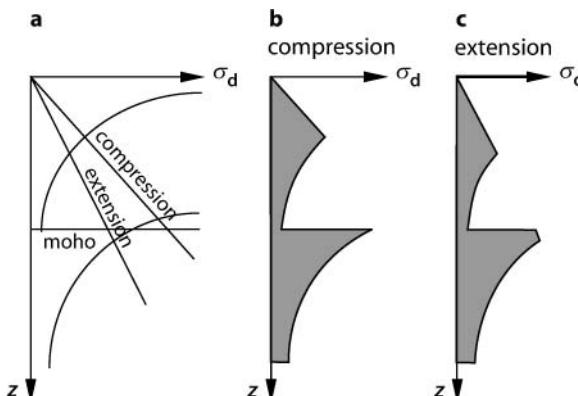
In first approximation it is fair to assume that a rock will begin to deform when the rheologically weakest phase fails. As quartz is one of the softer minerals and most crustal rocks contain quartz, the ductile deformation of the crust may well be described with the rheological data for quartz. Rocks in the mantle part of the lithosphere are quartz absent and dominated by olivine. Therefore, Fig. 5.13 shows two pairs of curves for power law creep; one pair for the creep stresses of quartz at low and high strain rates, the other for the creep behavior of olivine at low and high strain rates. Together, all these curves result in a strength profile for the continental lithosphere that contains two strength maxima, one at mid crustal levels, the other in the uppermost portions of the mantle part of the lithosphere. This extremely simple model for the rheological stratification of the lithosphere is called a *Brace-Goetze lithosphere* (after a suggestion by Molnar 1992).

**Qualitative Features of the Brace-Goetze Lithosphere.** The model of the Brace-Goetze lithosphere has a large number of features that are in phenomenal correspondence with observations in nature. Some of these are discussed here. Whilst we discuss these features, we want to keep in mind that

the Brace-Goetze lithosphere is only a model (also referred to as the “jelly sandwich model” and that there are also arguments against it (Jackson 2002).

- *Brittle failure in the mantle.* A comparison of Figs. 5.13c and d shows an interesting qualitative difference between the two strength profiles. At *low* strain rates the entire lithosphere below the brittle ductile transition deforms viscously. However, at *large* strain rates, the viscous strength of the upper mantle is larger than its brittle strength and the uppermost mantle will fracture. Of course, the occurrence of brittle fracture in the upper mantle depends on a large number of other factors as well. However, we want to note that the brittle strength of the upper mantle is comparable to its viscous strength at geologically realistic strain rates. Should it be true that the upper most mantle deforms brittle under some circumstances, then this process might have important consequences for the accumulation of mafic material (underplating) at the Moho (Huppert and Sparks 1988).

The transition from viscous flow to brittle failure in the upper most mantle may not only occur due to a change in strain rate (*increase* of viscous strength), but may also occur due to a *decrease* of the brittle strength. This may occur if there is a transition from compression to extension (sect. 5.1.2; Fig. 5.14, Sawyer 1985). Eq. 5.39 shows that the brittle strength of rocks is smaller in extension than it is in compression. It is therefore possible that brittle failure of the upper most mantle is caused by a qualitative change of the deformation regime.

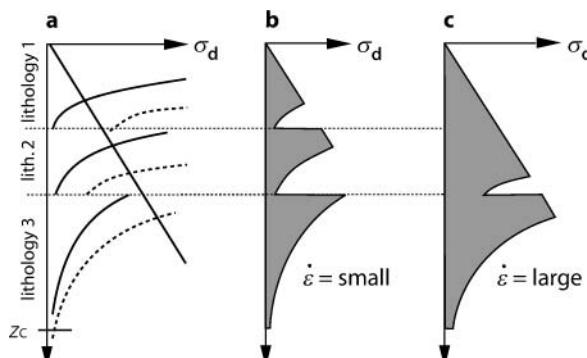


**Figure 5.14.** Schematic diagram showing the changes in mechanical strength of a Brace-Goetze lithosphere when changing the deformation regime qualitatively, i.e. from compression to extension. **a** The change from compression to extension decreases the brittle failure strength (see. eq. 5.38 and eq. 5.39; Fig. 5.7), while the viscous strength remains unaffected by this change, if the absolute value of the strain rate remains constant. Potentially, this may be reflected in brittle failure of the upper mantle. This is illustrated in the strength profiles in **b** and **c** that were constructed from **a**: In **b** the entire mantle part of the lithosphere deforms viscously, while in **c** the upper most mantle fails in brittle fashion

• *Changes in the rheological stratification.* Changes in the strain rate of an orogen can also change the rheological stratification of the lithosphere. This is illustrated in Fig. 5.15 using a simple model lithosphere made up of three lithological layers. The figure shows that a change in the strain rate may change the rheological layering. At low strain rates there are three strength maxima, while at high strain rate there are only two (Fig. 5.15b,c). Such weak points may be the nucleus for the formation of a tectonic nappe boundary. Thus, it is possible that the thickness of nappes in a lithologically stratified crust is a function of the strain rate (Kuznir and Park 1986).

• *Changes in the geotherm.* During viscous deformation it is not only changes in the strain rate that can change the strength of the lithosphere. Changing the geotherm may have the same influence. In the following we will encounter a series of examples where the shape of the geotherm is critical to the deformation mechanism. Other mechanisms that can cause changes in the strength of the lithosphere are, for example, strain hardening, or metamorphism.

• *Strength change due to metamorphism.* During metamorphism and deformation, both mineralogy and grain size change. It is therefore conceivable that a rock has a higher shear strength after metamorphism than before. For example, a garnet mica schist has a larger shear strength than its precursor: a clay. This is an interesting aspect which may be crucial in the consideration of postorogenic extension of mountain belts. In general it is thought that the stresses required for the late extension of an orogen are smaller than those required for its shortening. This is based on the fact that brittle deformation requires smaller stresses in tension than in compression. However, this is contrasted by the fact that the crust may have increased its strength by



**Figure 5.15.** Schematic diagram showing how changes of the deformation rate can cause changes in the rheological stratification of the lithosphere. In the shown example the crust is assumed to consist of three rock strata with different viscous rheologies (e.g. the upper, middle and lower crust). During deformation at low strain rate the lithosphere has three strength maxima. If the strain rate increases, the central strength maximum disappears. In a the continuous lines are for low strain rate and the dashed curves for large strain rate

metamorphism by the time it wants to extend. It is therefore possible that *larger* rather than *smaller* stresses are required to extend an orogen, than to build it.

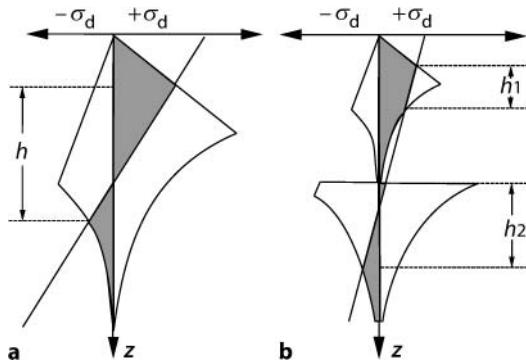
- *Rheology considering elastic criteria.* The Brace-Goetze lithosphere as we have discussed it in the last sections is based on the assumption that only viscous and brittle deformation mechanisms dominate its behavior. As such, it is in contrast to our assumption of the lithosphere as an elastic plate as we have done in sect. 4.4.2 when we have discussed flexural isostasy. Which model assumptions are made always depends on the question we are trying to answer (sect. 1.1).

Ranalli (1994) suggested to describe the rheology of the lithosphere using a coupled viscous, elastic and brittle approach. His model is illustrated in Fig. 5.16, but we need to be careful not to confuse it with the strength profiles in Figs. 5.13 or 5.14. Elastic deformation is instantaneous and does not reflect a strength envelop for a given strain rate as the viscous curves do. Thus, the model illustrated in Fig. 5.16 may be used to infer a stress state, but should not be interpreted as a failure envelope.

In a downward bent elastic lithosphere there is a stress neutral layer in the middle of the lithosphere. Above this point the lithosphere is under compression, below this point it is under extension (Fig. 4.23). These elastic stresses are shown in Fig. 5.16 by the straight line that goes from positive to negative stresses in the middle of the diagram. It may be seen that – in the upper most and lower most lithosphere – elastic stresses are extremely large but brittle and ductile stresses are small. However, in the central lithosphere, elastic stresses are smaller than brittle or viscous stresses and elastic stresses can therefore support internal and external loads. This model also shows that the elastic part of the lithosphere is significantly thinner than the thermally defined lithosphere (Fig. 5.16).

- *Quantitative description of a Brace-Goetze lithosphere.* In order to describe the Brace-Goetze lithosphere quantitatively we require quantitative information on 1. the depth dependence of temperature, i.e. a description of a geotherm; 2. the material constants (both 1 and 2 we need in order to calculate viscous stresses); 3. we need density and thickness of the crust and mantle part of the lithosphere in order to calculate vertical stresses and therefore the brittle strength. Table 5.3 lists typical numerical values for these parameters (Brace and Kohlstedt 1980). For the thermal structure of the lithosphere we will assume in the following that the radiogenic heat production decreases exponentially with depth according to eq. 3.67 with the characteristic drop off depth  $h_r = 10$  km. We also assume that the thermal conductivity is  $k = 2 \text{ J s}^{-1} \text{ m}^{-1} \text{ K}^{-1}$ , and that the temperature at the base of the lithosphere is  $T_1 = 1280^\circ\text{C}$ . Assumption on thickness and density are also listed in Table 5.4.

We also make the assumption that the viscous behavior of olivine is described by eq. 5.46 below stresses of 200 MPa and by eq. 5.47 for stresses above 200 MPa. Brittle failure is described with Byerlee's law. That is, below



**Figure 5.16.** Rheology of the lithosphere considering elastic, viscous and brittle constitutive relationships. Note that this model illustrates a *stress state* rather than a *failure envelope* and should therefore not be confused with illustrations like those in Figs. 5.13 or 5.14. Compressive stresses are plotted towards the right, tensional stresses towards the left. **a** Strength profile for the crust (viscous stresses are only plotted for a single mineral phase, e.g. quartz). **b** Strength profile through the lithosphere considering the rheology of quartz and olivine. In both strength profiles elastic stresses are also plotted as the straight line with positive stresses in the upper crust and negative stresses in the lower parts (s. Fig. 4.23). The deformation mechanism that dominates at a given depth is given by the lowest stresses at a given depth. The elastic section of the crust in **a** is therefore restricted to the region  $h$ . In **b** the elastic portion of the lithosphere is in two regions  $h_1$  and  $h_2$ , that are separated from each other (after Ranalli 1994)

500 MPa brittle failure is assumed to occur without cohesion and an internal coefficient of friction of 0.8 and above 500 MPa the cohesion is 60 MPa and the internal coefficient of friction is 0.6. Some strength profiles calculated with these assumptions are shown in Fig. 5.17.

**Strength of the Lithosphere.** When considering the distribution of stresses in the continental lithosphere, we have so far always only considered the stresses at a given depth. However, if we want to consider the deformation of entire continental plates, we need to know the mean stresses averaged over the entire lithosphere, or we need to know the total force that it needed to deform the entire lithosphere from top to base. Within the model of a Brace-Goetze lithosphere, this force is given by the vertically integrated stresses. This integrated strength is abbreviated with  $F_l$  and corresponds to the shaded region in Figs. 5.13, 5.14 and 5.15). If we make the thin sheet approximation, then the integrated strength of the lithosphere may be calculated as:

$$F_l = \int_0^{z_1} (\sigma_1 - \sigma_3) dz . \quad (5.48)$$

It has the units of force per meter or  $\text{Pa m} = \text{N m}^{-1}$ .  $F_l$  may be interpreted as the force acting in the direction normal to the orogen *per* meter length of orogen (i.e. in direction parallel to the orogen), that is required to deform the orogen with a given strain rate (Fig. 5.13, 5.14, 5.15). In the literature,

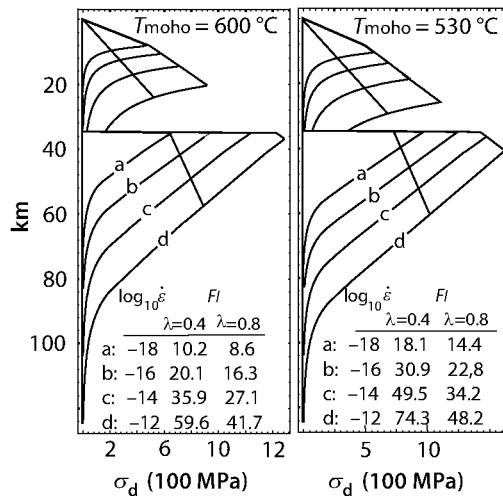
**Table 5.3.** Rheological parameters of the continental lithosphere relevant for its viscous behavior. These data are largely after Sonder and England (1986). However, the activation energy was changed from Sonder and England (1986) ( $Q_D = 5.4 \cdot 10^5$ ) to the value of  $Q_D = 5.454 \cdot 10^5$  in order to allow a smooth transition between the stresses at 200 MPa. Instead of changing  $Q_D$ , Zhou and Sandiford (1992) changed for the same reason  $\dot{\epsilon}_D$  from  $5.7 \cdot 10^{11}$  to  $3.05 \cdot 10^{11}$ . If all other parameters remain constant, then the two changes have the consequence that stresses are about 40 MPa larger than those of Sonder and England (1986)

parameter	value/unit	definition
<i>power law</i>	(eq. 5.46)	
$A_q$	$5 \cdot 10^{-6}$ MPa $^{-3}$ s $^{-1}$	pre exponent constant for quartz
$Q_q$	$1.9 \cdot 10^5$ J mol $^{-1}$	activation energy for quartz
$n_q$	3	power law exponent for quartz
$A_o$	$7 \cdot 10^4$ MPa $^{-3}$ s $^{-1}$	pre exponent constant for olivine
$Q_o$	$5.2 \cdot 10^5$ J mol $^{-1}$	activation energy for olivine creep
$n_o$	3	power law exponent for olivine
<i>Dorn's law</i>	(eq. 5.47)	
$Q_D$	$5.4 \cdot 10^5$ J mol $^{-1}$	activation energy for olivine creep
$\dot{\epsilon}_D$	$5.7 \cdot 10^{11}$ s $^{-1}$	strain rate
$\sigma_D$	8500 MPa	critical stress

**Table 5.4.** Rheological parameters of relevance for the brittle deformation of a Brace-Goetze lithosphere

parameter	value/unit	definition
$\mu(<500\text{ MPa})$	0.8	coefficient of friction in the crust
$\mu(>500\text{ MPa})$	0.6	coefficient of friction in the mantle
$\sigma_0(<500\text{ MPa})$	0	cohesion of the crust
$\sigma_0(>500\text{ MPa})$	60 MPa	cohesion of the mantle
$\lambda$	0.4 and 0.8	pore fluid/lithostatic pressure ratio
$z_c$	35 km	thickness of the crust
$z_l$	125 km	thickness of the lithosphere
$\rho_c$	$2750 \text{ kg m}^{-3}$	density of the crust
$\rho_m$	$3300 \text{ kg m}^{-3}$	density of the mantle lithosphere

the terms “strength” is often used very loosely. Strength (in Pa), *integrated strength* (in N m $^{-1}$ ), sometimes *stress* and occasionally even *force* are all often confused. We want to remember that *strength* has the units of stress (it is the stress that leads to brittle failure or viscous flow) and that *integrated strength* is a force per meter (which is equal to stress  $\times$  meter). In the viscous regime strength is only defined for a given strain rate. This should be clear from



**Figure 5.17.** Strength profiles for the continental lithosphere as calculated with the model for a Brace-Goetze lithosphere and the data from Table 5.3 and 5.4). *a, b, c* and *d* are profiles for four different geologically relevant strain rates. The two diagrams show the strength profiles for two different Moho-temperatures that result from assumptions for the radiogenic surface heat production of  $S_0 = 5 \cdot 10^{-6} \text{ W m}^{-3}$  and  $S_0 = 7 \cdot 10^{-6} \text{ W m}^{-3}$ . In each diagram two linear curves for brittle failure for  $\lambda = 0.4$  and  $\lambda = 0.8$  are plotted. The stress curve with the higher stresses is for the lower value of  $\lambda$ . It was assumed that  $\lambda_c = \lambda_l$ . The vertically integrated stresses  $F_1$  are given in  $10^{12} \text{ N m}^{-1}$

eq. 5.46, where it is shown that the viscous stresses (strength) are strongly dependent on strain rate.

- *Integrated strength as a function of Moho-temperature.* The strength of the lithosphere is very strongly dependent on the Moho-temperature. The details of the temperature distribution above and below the Moho are only a second order effect (Sonder and England 1986). Thus, for many mechanical questions on the scale of the lithosphere it is sufficient to characterize the geotherm by a single number: the Moho-temperature. Above and below the Moho it is sufficient to assume linear geotherms. However, in the following sections we continue to use curved geotherms characterized by exponentially decreasing radiogenic heat productions and so use the relationships derived in sect. 3.4.4. Thus, we determine the Moho-temperature indirectly by assuming the radiogenic surface heat production, the thermal conductivity and the surface heat flow. As a reminder to sect. 3.4.1, Fig. 5.18 shows the Moho-temperature and the surface heat flow as a function of thermal conductivity and heat production.

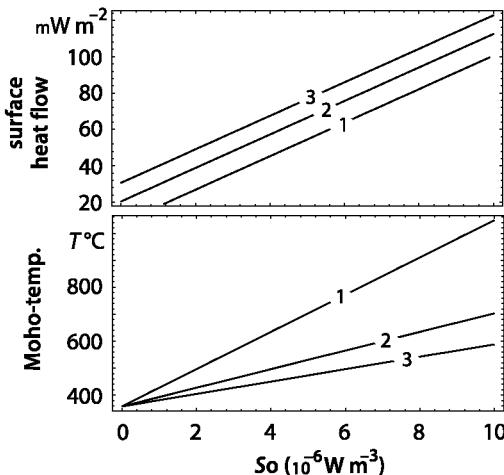
In order to quantify the integrated strength of the lithosphere we must integrate eq. 5.48. However, it is near impossible to integrate this equation analytically as the strength profile between  $z = 0$  and  $z = z_1$  is composed of several very different functions (Zhou and Sandiford 1992). The values for

integrated strength quoted in the following sections were derived by numerical integration of eq. 5.48. Fig. 5.19 illustrates that the integrated strength rises dramatically with increasing strain rate (which corresponds to eq. 5.46). At geologically realistic strain rates the integrated strength ranges between  $10^{12}$  and  $10^{14}$  N m $^{-1}$  (s. Fig. 5.17). These magnitudes correspond well with the magnitude of estimated plate tectonic driving forces (sect. 5.3).

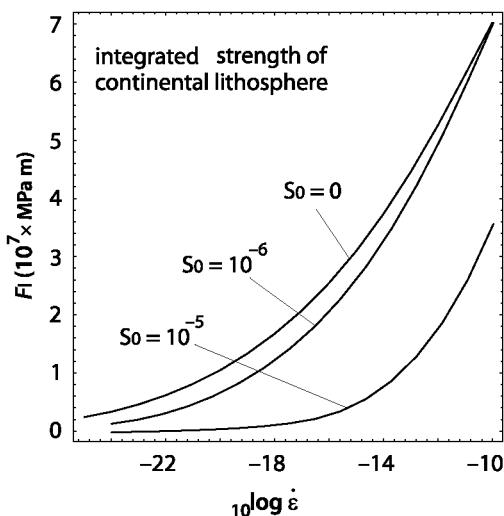
### 5.2.2 Rheology of the Oceanic Lithosphere

The fundamental assumptions which we have made for the calculation of stresses and strength profiles for the continental lithosphere are also valid for

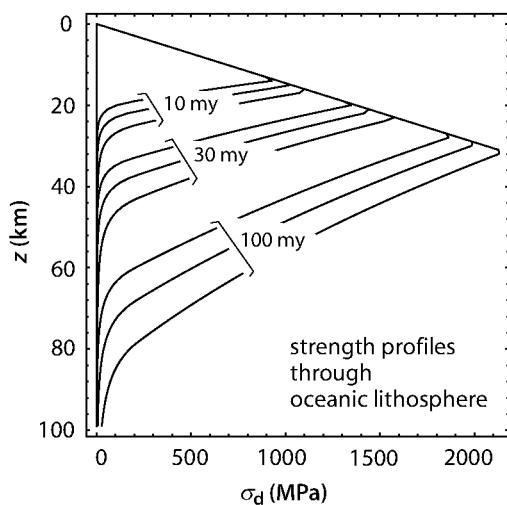
**Figure 5.18.** Surface heat flow and Moho-temperature as a function of radiogenic heat production at the surface  $S_0$ . It is assumed that this heat production rate decreases exponentially with depth. The characteristic skin depth for this exponential drop off is assumed to be 10 km (s. eq. 3.67). Curves are labeled for conductivities in  $\text{J s}^{-1} \text{m}^{-1} \text{K}^{-1}$  (calculated with eq. 3.76, eq. 3.77 and the data from Table 5.4)



**Figure 5.19.** Vertically integrated strength of the continental lithosphere  $F_1$  as a function of strain rate. The curves are shown for three different geotherms characterized by three different surface heat productions ( $S_0$  in  $\text{W m}^{-3}$ ). The Moho-temperatures that correspond to these assumptions may be read from Fig. 5.18. Similar diagrams were first discussed by Sonder and England (1986). The integrated strengths shown here are relatively small, because the assumed temperature profiles are relatively high



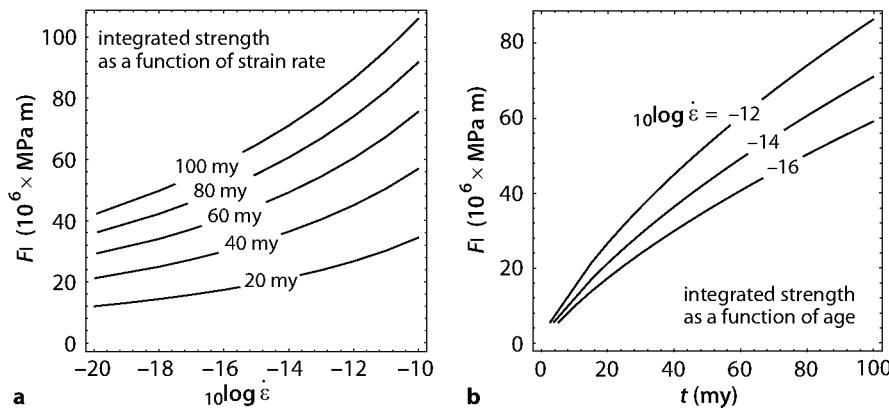
**Figure 5.20.** Strength profile through the oceanic lithosphere at ages of 10, 30 and 100 my. For each of these ages, stresses were calculated for three strain rates of  $\dot{\epsilon} = 10^{-16}$ ,  $10^{-14}$  and  $10^{-12} \text{ s}^{-1}$ . For each age, the curve for the highest strain rate has the largest strength. The temperature profiles needed to calculate the stresses were calculated using eq. 3.80; for the rheological data the values of Table 5.3 and 5.4 were used



the oceanic lithosphere. However, there are two important differences: 1. In contrast to the continental lithosphere, oceanic geotherms are time dependent and there is no radiogenic heat production in the oceanic lithosphere. As a consequence, different relationships must be used to calculate the temperature profile with depth and ultimately the rheology (sect. 3.5). 2. There is practically no quartz bearing crust in the oceanic lithosphere and the rheology of oceanic lithosphere is therefore largely governed by the rheology of olivine. As a consequence, there is only one maximum in the strength profile (Fig. 5.20). On the other hand, the strength profiles of oceanic lithosphere are highly dependent on its age (sect. 3.5.1). The depth dependent temperature profile of oceanic lithosphere may be calculated with eq. 3.80. A strength profile for the oceanic lithosphere may then be calculated using eqs. 5.38, 5.46 and 5.47 as well as the data from table 5.3 and 5.4. Fig. 5.20 shows some examples for such strength profiles through oceanic lithosphere.

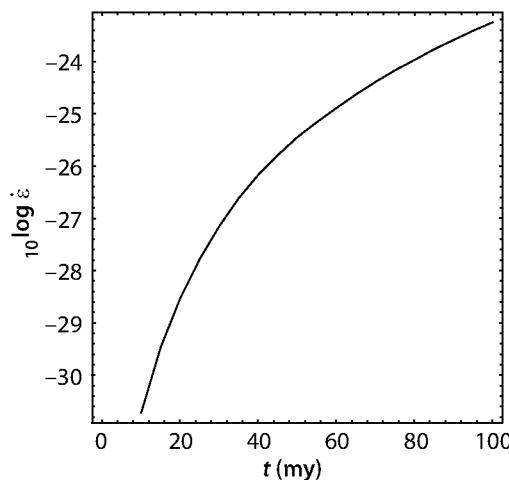
**Strength of the Oceanic Lithosphere.** Eq. 5.48 may be used to calculate the integrated strength of the oceanic lithosphere just like we used it above to calculate the integrated strength for the continental lithosphere. In fact, it is possible to calculate the integrated strength of the oceanic lithosphere with much higher accuracy than of the continental lithosphere, because oceanic geotherms are much better known than continental geotherms. Fig. 5.21a shows the calculated integrated strength of oceanic lithosphere as a function of different strain rates and age as calculated numerically with eq. 5.48.

Fig. 5.21a may be compared directly with Fig. 5.19. In Fig. 5.21b the same information is shown in a diagram with age on the horizontal axis. A comparison with Fig. 5.19 shows that only very young oceanic plates have a smaller integrated strength than continental lithosphere. This result from Figs. 5.21 and 5.19 corresponds to our observations: We know that most intra plate



**Figure 5.21.** **a** Integrated strength of the oceanic lithosphere as a function of strain rate for 5 differently aged oceanic lithospheres. Curves were calculated with eq. 5.48 using the results of Fig. 5.20. **b** Integrated strength of oceanic lithosphere as a function of age for three different strain rates. **a** and **b** contain the identical information

**Figure 5.22.** The strain rate with which oceanic lithosphere will deform in response to ridge push. The strain rate is plotted against age of the oceanic lithosphere. Clearly, the older the oceanic lithosphere is, the larger the applied ridge push and the larger the strain rate. Note that all strain rates are several orders of magnitude *below* anything that might be geologically relevant. In other words, oceanic lithosphere does practically *not* deform under the force exerted by ridge push force



seismicity occurs in the continents and not in the oceans (Fig. 2.2). There is practically no deformation inside the oceanic plates of the earth. An oceanic plate acts - because of its high integrated strength - like a passive transmitter of stresses from the mid oceanic ridges to the continents (s. sect. 6).

- *Strength relationships between continental and oceanic lithosphere.* In the previous paragraph we have shown that oceanic lithosphere is significantly stronger than continental lithosphere, even though it is generally much thinner. We have come to this important conclusion by comparing the integrated strength of continents and oceans if they would deform under the same strain rate. However, in geodynamics it is often more meaningful to compare the de-

formation rate of continental and oceanic lithosphere under the same applied stresses, rather than the stresses under the same applied strain rates.

One of the most important plate tectonic driving forces is the potential energy of the mid oceanic ridges (sect. 5.3.2). The force that is exerted by these ridges onto the surrounding continents increases with the age of the oceanic lithosphere. We can now ask ourselves if it is the continents or the oceanic lithosphere that are deformed by this force. In order to answer this question, we have plotted the strain rate of oceanic lithosphere with which oceanic lithosphere would deform under its own ridge push against age of the oceanic lithosphere (s. sect. 6.2.2). Fig. 5.22 shows that these strain rates are geologically irrelevant. The vast majority of the plate divergence at the mid oceanic ridges is compensated by deformation inside the continents and not inside the oceanic lithosphere.

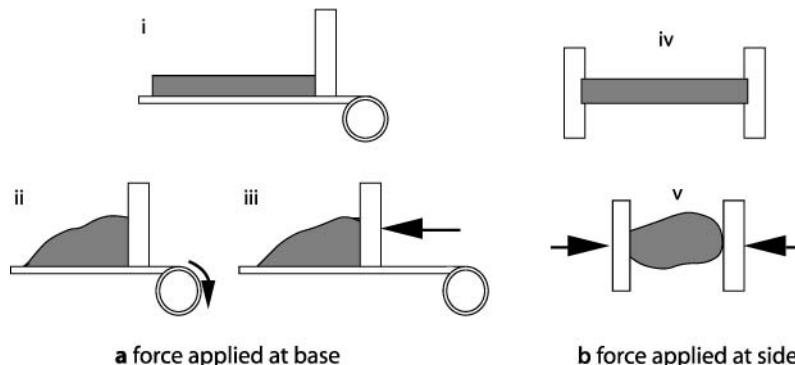
## 5.3 Forces Applied to Lithospheric Plates

### 5.3.1 Transmission Mechanisms

Plate tectonic driving forces may be divided into two fundamental groups according to the way they are transmitted:

- transmission by shear stresses,
- transmission by normal stresses.

Because plate tectonic driving forces act horizontally, *shear stresses* must be applied to horizontal surfaces and *normal stresses* to vertical surfaces. If the transmission occurs by shear stresses, this is often called *basal traction*. If the transmission occurs by normal stresses, we speak of *end loading* or *side*



**Figure 5.23.** Illustration of the two fundamental mechanisms for the transmission of plate tectonic driving forces. **a** illustrates transmission by *basal friction*. In *ii* basal friction is shown in a *Eulerian* reference frame (“conveyer belt” model). In *iii* basal friction is shown in a *Lagrangian* reference frame (“bulldozer” model). **b** illustrates transmission by lateral normal stresses (“side forcing”)

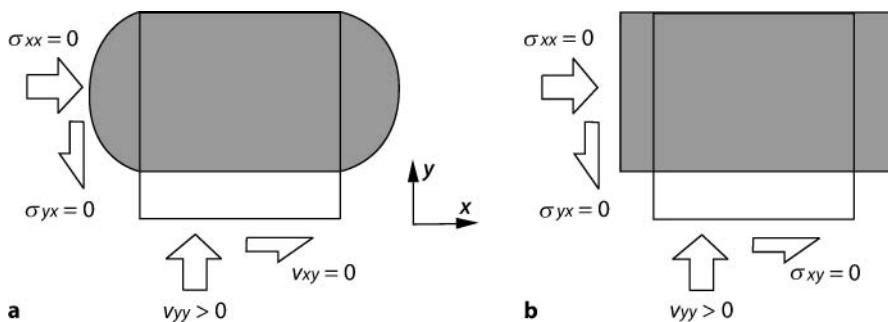
*forcing*. Fig. 5.23 illustrates that different deformation geometries may arise, depending on the transmission mechanism. However, in reality it is often difficult to interpret which of the two mechanisms is responsible for a given plate motion (Wilson 1993). In fact, the scientific community remains divided between those who believe that plate tectonics is driven by *basal traction* and those who consider *lateral normal stresses* as the principal driving force.

**Transmission of Stresses by Shear- or Normal Stresses.** One model for the explanation of plate motions is that the friction between the base of the lithosphere and the convective motion in the asthenosphere is the principal driving mechanism (Ziegler 1992; 1993). The most important argument *for* this model comes from the reconstruction of past plate motions. These do not correspond very well with the global geometry of mid oceanic ridges and subduction zones. Thus, it is thought that these plate motions reflect the geometry of convection cells in the mantle instead. The most important argument *against* this model is implicit in Fig. 5.13. This figure shows that differential stresses at the base of the lithosphere are much too small to be able to transmit forces from the mantle into the lithosphere. It is therefore hard to imagine that this softest part of the lithosphere can transmit stresses large enough to build the mountain ranges of our planet (s. mechanical definition of the lithosphere in sect. 2.4.1). The tractions at the base of the lithosphere are not likely to be larger than  $10^{-2}$  MPa (Richardson 1992).

The other - and by far more accepted - model for the explanation of plate motions is that *plate boundary* forces drive plate tectonics by lateral normal stresses (e.g. Forsyth and Uyenda 1975). These forces are predominantly caused by *potential energy variations*. Such variations occur inside the continents and along the boundaries of oceanic lithosphere and will be discussed on the following pages.

Despite the two different models for the origin of plate tectonic driving forces we should not forget that, ultimately, *all* plate tectonic forces find their origin in the thermal energy of earth. On the scale of an individual orogen it is often easier to determine the way forces are transmitted into the orogen. For example, the geometry of accretionary wedges shows clearly that they form by basal traction with the subducting plate below (sect. 6.2.3). An example for deformation in response to lateral normal stresses is the deformation of intra continental mountain ranges like the Tien-Shan in central Asia. The Tien Shan developed at the same time as the Himalaya. The forces involved in the India - Asia collision were transmitted horizontally across the Tarim Basin into the Tien Shan.

**Boundary Conditions of Deformation.** In the last section we have discussed normal and shear *stresses* that cause the deformation of plates. However, it is not clear that it is stresses that are the appropriate boundary conditions for plate deformation. For example, the India-Asia collision keeps going at constant convergence rate irregardless the stresses that arise as a consequence of the build up of the Himalaya. Thus, this may be an example



**Figure 5.24.** Illustration showing how apparently small differences in the boundary conditions can cause very different deformation geometries. In both examples a square body is subjected to two-dimensional plane strain deformation. In both **a** and **b** the tangential and normal components of the boundary conditions on the left and right boundaries are given by stresses, which are assumed to be zero ( $\sigma_{xx} = \sigma_{yx} = 0$ ). Also, in both examples the normal component at the bottom boundary is given by a velocity with which the bottom boundary moves towards the top boundary ( $v_{yy} > 0$ ). **a** and **b** differ only in the tangential component of the boundary condition along the bottom boundary. In **a** this is given by a velocity ( $v_{xy} = 0$ ), in **b** it is given by a stress ( $\sigma_{xy} = 0$  = free slip)

where *velocities* rather than stresses form an appropriate boundary condition. We therefore discriminate between:

- Orogenic boundary conditions given by velocities,
- Orogenic boundary conditions given by stresses.

Both types of boundary conditions may have a *normal* and a *tangential* component. Thus, for a two-dimensional mechanical model with the two spatial coordinates  $x$  and  $y$ , we require a tangential and a normal boundary condition on each boundary. A total of four variables must be defined by the boundary conditions. Fig. 5.24 illustrates that the difference between velocity and stress boundary conditions may have a very profound influence on the deformation geometry. A boundary condition is called *free slip* if the shear stresses along this boundary are considered to be zero.

The reason why we require *four* boundary conditions for the description of two-dimensional deformation may be also seen from the stress balance equations (eq. 5.15, 6.25 and 6.26). If we integrate these equations with two variables, then there are *four* constants of integration. In order to determine these constants we need *four* independent pieces of information: the four boundary conditions.

If a medium is not everywhere a continuum, for example because it contains a brittle fracture, then all mechanical properties may have discontinuities. Such problems may not be solved using a single set of boundary conditions as in Fig. 5.24 and we must use internal boundary conditions or other special tricks to be able to solve such problems. For example, the medium can be subdivided into several continuous regions that are described separately.

**Potential Energy.** Practically all important plate tectonic driving forces find their origin in differences of the potential energy of different parts of the earth (Turcotte 1983). In this section we explain what we understand with the term *potential energy* in a plate tectonic context. We will return to this concept again in the sections 5.3.2, 5.3.3 and 6.2.2.

In sect. 4.4.1 we have shown that the vertical normal stress at a given depth in the crust  $z$  is given by the product of density, gravitational acceleration and the height, or thickness of the vertical rock column above it. This vertical normal stress is the vertically acting *force per area*. It may be calculated by integrating  $\rho g$  between 0 and  $z$ , as we did in eq. 4.22. If the density over the thickness  $z$  remains constant, then this is simply  $\rho g z$ . This term has the units of Pa or  $\text{kg s}^{-2} \text{m}^{-1}$  or  $\text{J m}^{-3}$ . We can see that *stress* has the same units as *energy per volume*.

This quantity can also be interpreted as the potential energy of a cubic meter of rock at depth  $z$ . If we want to know the potential energy not of a single cubic meter, but that of a whole body, for example that of a mountain range, then we need to integrate this *potential energy per cubic meter* over the lateral and vertical extent of the range. Fortunately, it is usually sufficient to know the potential energy *per area*, i.e. that of a complete vertical column, but only for one square meter of area. Using this *potential energy per area* we can compare different regions on the globe, for example two neighboring lithospheric columns of different thickness and density distribution. In the following we will represent the potential energy *per area* with  $E_p$ . In order to determine  $E_p$  at depth  $z$  we simply need to sum up (i.e. integrate) the vertical stresses in the lithospheric column of interest between the surface (which usually is  $z=0$  in the reference frame we use) and the depth of interest  $z$ :

$$E_p = \int_0^z \sigma_{zz} dz = \int_0^z \int_0^z \rho(z) g dz dz . \quad (5.49)$$

Very often the “depth of interest” is the isostatic compensation depth. If the density is independent of depth, then eq. 5.49 may be simplified to give:

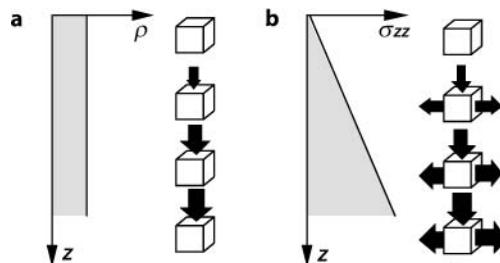
$$E_p = \int_0^z \sigma_{zz} dz = \int_0^z \rho g z dz = \frac{\rho g z^2}{2} . \quad (5.50)$$

This integral corresponds to the gray shaded region in Fig. 5.25b. We want to remember that  $E_p$  has the units of energy *per area* and is, therefore, strictly speaking, no energy as such.

**Horizontal Forces Arising from Potential Energy Variations.** In a static, non-deforming lithosphere the horizontal and vertical normal stresses have the same magnitude (see Fig. 5.25). It is true that:

$$\sigma_{zz} = \sigma_{xx} = \sigma_{yy} . \quad (5.51)$$

This is also stated in eq. 5.42, which says that there is no deviatoric stress if the strain rate is zero. The sum of all vertical stresses integrated over the



**Figure 5.25.** Density  $\rho$  and vertical normal stress  $\sigma_{zz}$  as a function of depth  $z$ . The value  $\sigma_{zz}$  is the vertically integrated density, times acceleration. Thus, the curve in **b** corresponds to the gray shaded region in **a**. The row of little unity cubes next to **a** illustrates how the vertical stress increases with depth. The column of cubes next to **b** illustrates that the horizontal force exerted by the column on its surroundings is given by the sum of all vertical stresses. This corresponds to the gray shaded area in **b**

thickness of a plate is the potential energy of the plate per area. Since horizontal and vertical stresses are the same, this potential energy per area is equivalent to the force exerted by the lithosphere onto its surroundings, per meter length of orogen. If two neighboring vertical lithospheric columns have the same potential energy per unit area, then they also exert equally large horizontal forces onto each other and there is no “net force” between them. However, if they have different potential energies per area, then this potential energy difference between the two plates may be interpreted as the net force  $F_b$  that is exerted by one column onto the other in the horizontal direction and per meter length of orogen. This net force arising from potential energy differences is also called *horizontal buoyancy force* (somewhat cumbersome) or *gravitational stress* and it is important to remember that it has the units of force per meter length of orogen. This potential energy difference may be written as (s. Fig. 5.26):

$$\Delta E_p = F_b = \int_0^{z_K} \int_0^{z_K} \rho^A(z) g dz dz - \int_0^{z_K} \int_0^{z_K} \rho^B(z) g dz dz . \quad (5.52)$$

There,  $z_K$  could be any depth, but for many purposes it is useful to assume that it is the same isostatic compensation depth we used on p. 161. Below this depth there is no density difference between the vertical columns  $A$  and  $B$  (s. eq. 4.21).  $\rho^A(z)$  is the density of profile  $A$  as a function of depth  $z$ .

If density is a continuous function of depth, then eq. 5.52 may be usually integrated without too much trouble. However, in the lithosphere, the density distribution has (at the least) a discontinuity at the Moho so that it may be necessary to split the integral in eq. 5.52, even for very simple assumptions on the density distribution in the lithosphere. For example, the potential energy of the two-layered column  $A$  in Fig. 5.26 for column  $A$  may be described by:

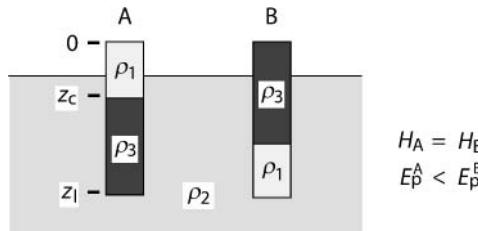
$$\begin{aligned}
 \int_0^{z_1} \int_0^{z_1} \rho^A(z) g dz dz &= \int_0^{z_c} \sigma_{zz} dz + \int_{z_c}^{z_1} \sigma_{zz} dz , \\
 &= \int_0^{z_c} \rho_1 g z dz + \int_{z_c}^{z_1} \rho_1 g z_c + \rho_3 g(z - z_c) dz , \\
 &= \frac{\rho_1 g z^2}{2} \bigg|_0^{z_c} + \rho_1 g z_c z \bigg|_{z_c}^{z_1} \\
 &\quad + \frac{\rho_3 g z^2}{2} \bigg|_{z_c}^{z_1} - \rho_3 g z_c z \bigg|_{z_c}^{z_1} . \tag{5.53}
 \end{aligned}$$

It should be easy to understand this equation graphically by plotting density and vertical stress as a function of depth as we did in Fig. 5.25, 5.27, 5.31 and 5.33.

The importance of the density *distribution* in the lithosphere for the potential energy may be illustrated nicely with an interesting example. Fig. 5.26 shows two columns in isostatic equilibrium. The two columns have the same isostatically supported surface elevation, because they are made up of sections of the same densities and thicknesses. However, they have different potential energies because in column *B* the dense part lies up high. Potential energy does not only depend on thickness and density, but also on the *distribution* of density with depth. Thus, there is a net buoyancy force between the two columns shown in Fig. 5.26. This net force is exerted by column *B* towards column *A*.

We can conclude that it is dangerous to infer lateral forces from topography on the surface of earth (England and Molnar 1991). Surface elevation is a linear function of thickness (eq. 4.23 and 4.24), while potential energy per unit area is a quadratic one (eq. 5.49 and 5.52). In fact, it is even possible, that topographically *lower* regions exert a gravitational stress on topographically *higher regions*, averaged over the thickness of the lithosphere (Stüwe and Barr 2000). Geoid anomalies on the other hand can be used to estimate the density distribution within the lithosphere. Coblenz et al. (1994) have used a combination of information on surface elevation and on geoid anomalies to estimate the potential energy of the lithosphere.

**Force Balance Between Mountains and Foreland.** In this section we estimate the forces exerted by a mountain range onto its foreland (Fig. 5.27). For this, we will follow the logic of Molnar and Lyon-Caen (1988) and also use their choice for the vertical axis of the cross section. We assume an origin at the Moho and measure the vertical direction positively upwards as illustrated in Fig. 5.27b. This choice for the vertical axis helps the intuitive understanding if the integration of eq. 5.52, as one of the integration limits is always zero (s. Molnar and Lyon-Caen 1988). However, note that the results are independent of the chosen reference frame as we do not calculate *absolute* potential energies, but only potential energy differences between two neighboring columns. Thus, as long as we choose the same coordinate system for



**Figure 5.26.** Schematic cartoon showing two columns in isostatic equilibrium ( $\rho_1 < \rho_2 < \rho_3$ ). The surface of both columns has the same elevation above the liquid of density  $\rho_2$ , because both bodies consist of equally thick sections of the densities  $\rho_1$  and  $\rho_3$ , i.e. they have the same weight. However, column *B* has a much higher potential energy per unit area than column *A*, because the distribution of density is different. In column *B* the high density part of the section lies higher. As a consequence, *B* exerts a net force towards *A*. The cartoon obviously represents no geologically realistic scenario for the lithosphere. However, it is useful to illustrate why mountain ranges need not exert a net force onto their lower lying surroundings even if they have a higher surface elevation. Note however, that this logic only applies if forces may be averaged over the thickness of the lithosphere. Problem 5.16 is related to this figure

the two columns that are to be compared, it does not matter which reference frame we pick.

We begin by calculating the potential energy per unit area of the foreland following the logic of Molnar and Lyon-Caen (1988). We can find this by integrating eq. 5.49. For the undeformed lithosphere in the foreland the potential energy above the Moho is simply:

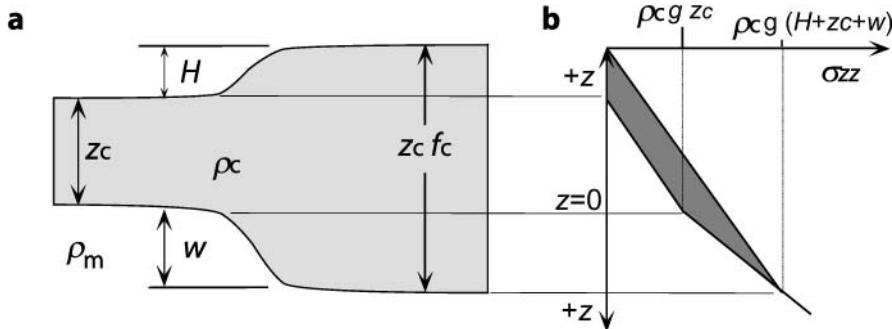
$$E_p^{\text{foreland}} = \rho_c g z_c^2 / 2 . \quad (5.54)$$

Correspondingly, the potential energy of the thickened crust relative to the Moho is:

$$E_p^{\text{range}} = \rho_c g (H + z_c)^2 / 2 + \Delta \rho g w^2 / 2 . \quad (5.55)$$

where  $\Delta \rho = (\rho_m - \rho_c)$  and the thicknesses  $H$ ,  $Z_c$  and  $w$  are as labeled on Fig. 5.27. The first term in the equation above is simply the potential energy of the thickened crust above the chosen origin at the Moho of the undeformed lithosphere. the second term is in the negative  $z$  direction, but the density contrast is also negative (as it acts as a buoyant force) providing in total a positive contribution to the potential energy. Note that these relationships were derived according to the same principal as eq. 5.53. The potential energy difference per unit area is given by the difference of eq. 5.54 and eq. 5.55 (s. eq. 5.52). It is:

$$\begin{aligned} \Delta E_p &= F_b = E_p^{\text{range}} - E_p^{\text{foreland}} \\ &= \rho_c g H^2 / 2 + \rho_c g H z_c + \Delta \rho g w^2 / 2 . \end{aligned} \quad (5.56)$$



**Figure 5.27.** Cartoon contrasting the distribution of vertical stresses in mountain ranges relative to their foreland. **a** The thickness of the crustal root is  $w$ , the surface elevation relative to the reference lithosphere in the foreland  $H$ . In isostatic equilibrium it is true that:  $H\rho_c = w(\rho_m - \rho_c) = w\Delta\rho$ . In **b** the vertical stresses are drawn for the mountain range and the foreland. The dark shaded area between the two stress curves has the units of stress  $\times$  meters or force per meter length of orogen exerted by the range onto the foreland. It corresponds to the potential energy difference between the mountain range and the foreland

Eq. 5.56 may be simplified because we assume that both, mountain range and foreland are in isostatic equilibrium. The isostasy condition states that:  $\Delta\rho w = H\rho_c$ . Using this we can simplify eq. 5.56 to:

$$\Delta E_p = F_b = \rho_c g H (H/2 + z_c + w/2) . \quad (5.57)$$

The force  $F_b$  corresponds to the dark shaded region in Fig. 5.27b. It is the difference between the vertically integrated vertical stresses  $\sigma_{zz}$  of two vertical columns in the mountain range and in the foreland, respectively (Tapponier and Molnar 1976). For a 3 km high mountain range with a 30 km root, eq. 5.57 gives a force  $F_b$  of the order of  $3\text{--}4 \cdot 10^{12} \text{ N m}^{-1}$ . We will see that this number is comparable with the forces applied to and exerted by mid ocean ridges.

Despite its simplicity, eq. 5.56 may be used to draw some very fundamental conclusions. For one, we can see that the third term is significantly larger than the first term. Thus, the potential energy difference between two mountain ranges of the same elevation becomes larger if the compensating root is thicker. For example, a 100 km thick root of a mountain range made up of low density mantle material contributes significantly more to the potential energy of a range than a 60 km thick root of crustal material. We can also see from eq. 5.56 that the potential energy of a mountain range grows with the square of both the surface elevation *and* the thickness of its root. The work that must be done to increase the surface elevation of a mountain range by one meter increases therefore as the mountain range gets higher (Molnar and Tapponier 1978; s. sect. 6.2.2). This is the reason why mountain ranges do not grow infinitely on this planet and have a limiting elevation. As poten-

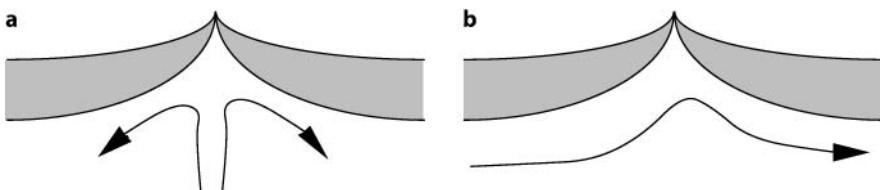
tial energy variations are some of the most important driving forces in the lithosphere we will continue with more details in the following sections 5.3.2 and 5.3.3.

### 5.3.2 Forces in Oceanic Lithosphere

The forces exerted by oceanic lithosphere onto the continents around them are considered to be the fundamental driving mechanism for plate tectonic motion (McKenzie 1969b). There are two important driving forces in oceanic lithosphere: 1. the potential energy of the mid-oceanic ridges and 2. the forces that occur in subduction zones. The former is called *ridge push*, the latter are called *slab pull* and *trench suction*. In the following we discuss the nature of both types of forces and discuss their magnitude.

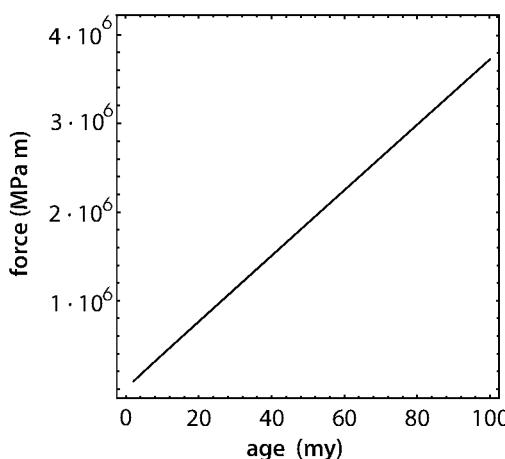
**Ridge Push.** Mid-oceanic ridges have a high topography and a high potential energy relative to the average oceanic lithosphere. This potential energy is one of the more important (and certainly best known) plate tectonic driving forces. While strictly speaking the mid-oceanic ridge applies a torque to the plate (s. p. 25), we will neglect here the curvature of the earth and continue using the term “ridge push”. It is important to understand that ridge push finds its origin in the high potential energy of the ridge, rather than in the frictional stresses between an outward welling mantle plume and the oceanic plate as drawn in Fig. 5.28a.

The ridge push force per meter length of ridge (equivalent to the potential energy of the ridge per unit area) may be calculated with eq. 5.52, using similar assumptions to those we have made when designing a model to explain the water depth of the oceans (s. Fig. 4.19). The density of oceanic lithosphere must be expressed in terms of temperature (eq. 4.27) and temperature as a function of depth (eq. 3.80; s. Turcotte and Schubert 1982; Parsons and Richter 1980). Then - using the half space cooling model - it may be shown that the ridge push force is a function of the thermal profile through the



**Figure 5.28.** Cartoon showing two possible motions of the asthenosphere below mid oceanic ridges. **a** Asthenospheric material wells up below the mid ocean ridge in form of a *mantle plume*. During this process, adiabatic decompression of asthenosphere material will cause massive partial melting. It is thought that this situation pertains to regions where these melts are now present as large igneous provinces like the Karoo Basalts in southern Africa or the Deccan Traps in India and may be Iceland (Fig. 6.34). **b** shows the mantle motion that is thought to be representative for most mid oceanic ridges

**Figure 5.29.** The force exerted by mid-oceanic ridges onto the surrounding plate per meter length of ridge, shown as a function of age of the oceanic lithosphere. Calculated with eq. 5.58



oceanic lithosphere and therefore of age. Without reiteration the derivation of the ridge push force here, we simply state that it is given within this model by the equation:

$$F_b = g\rho_m\alpha T_l \kappa t \left( 1 + \left( \frac{\rho_m}{\rho_m - \rho_w} \right) \frac{2\alpha T_l}{\pi} \right) \approx 1.19 \cdot 10^{-3} t \quad . \quad (5.58)$$

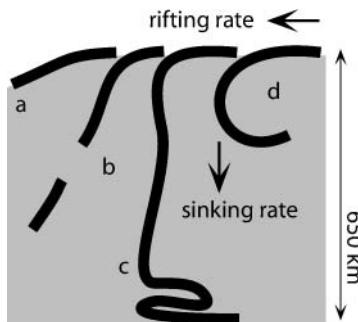
All parameters of this equation are explained in sect. 4.4.1. From eq. 5.58 we can see that the ridge push force is a *linear* function of age of the oceanic lithosphere (Fig. 5.29). As such it is different from water depth which - within this model is described by a square root function of age (Fig. 4.20). The numerical value of the proportionality constant between age and force in eq. 5.58 ( $1.19 \cdot 10^{-3}$ ) is derived using the following constants:  $T_l = 1200^\circ\text{C}$ ;  $\rho_m = 3200 \text{ kg m}^{-3}$ ;  $\rho_w = 1000 \text{ kg m}^{-3}$ ;  $\alpha = 3 \cdot 10^{-5} \text{ K}^{-1}$  and  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ . Fig. 5.29 shows that ridge push is about an order of magnitude *smaller* than the integrated strength of continents at normal orogenic strain rates (s. Fig. 5.19). Thus, we may conclude that ridge push alone is insufficient as the principal plate tectonic driving force.

- *Asthenospheric flow at mid-oceanic ridges.* In the past, ridge push has been interpreted to be related to frictional stresses of upwelling asthenosphere that “pushes” the ridge apart as illustrated in Fig. 5.28a. However, several observations speak against this model. For example, if upwelling material causes ridges, then it would be expected that different ridges have different elevations above the abyssal planes - dependent on the force exerted by the upwelling materials. In contrast, practically all mid ocean ridges lie at a constant water depth. Today we know that there are only very few places where mid-oceanic ridges coincide with diapirically upwelling mantle material. Rather, the asthenospheric flow at most mid-oceanic ridges is of the geometry shown in Fig. 5.28b. Among other arguments, this was recognized by McKenzie and Bickle (1988) using on geochemical arguments. These authors showed that partial melting that would occur due to adiabatic decompression of upwelling

melt in a mantle plume would be enough to form a 15 km thick oceanic crust. In contrast, normal oceanic crust is measured to be only about 5–7 km thick. This thickness can be produced by adiabatic melting of only the upper most asthenospheric regions. Asthenospheric flow as sketched in Fig. 5.28b is sufficient to produce a 5–7 km thick oceanic crust. Thus, it is thought that the flow directions of asthenospheric convection have little to do with the position of the mid-oceanic ridges. There are only very few places where mid-oceanic ridges coincide with diapirically upwelling mantle material. One of these places is Iceland (Fig. 6.34). There, not only is the oceanic crust significantly thicker than 7 km, but also the mid oceanic ridge is uplifted by the stresses exerted to its base by the upwelling mantle material not unlike a jet of water shot from below onto a rubber sheet. Other regions where mantle plumes are thought to coincide with rifting margins are those of flood basalts (p. 315) (White and McKenzie 1989).

**Slab Pull and Trench Suction.** Old oceanic lithosphere is denser than the underlying asthenosphere and it has therefore a negative buoyancy and it wants to sink. However, because oceanic lithosphere is very strong and stiff, it cannot immediately do this as soon as it reaches this critical age where its density becomes large compared to that of the underlying asthenosphere. Rather, the oceanic plate “glides” along the surface of the asthenosphere until this gravitationally unstable configuration is brought out of balance and a subduction zone forms. Once the edge of such an old oceanic plate has begun to subduct, it drags the remainder of the plate behind it. This is what is called *slab pull*. Such subduction processes may cause, or may be caused by, small scale convection in the upper mantle. This convection occurs predominantly in the wedge shaped region between the subducting and the upper plate. Once such a convection system is set up, it may actually drag both the upper plate and the subducting plate into the subduction zone (s. Fig. 3.28). This is what is called *trench suction*.

Slab pull is gravitationally induced, simply because the dense oceanic lithosphere wants to sink into the less dense upper mantle. In fact, the slab pull force is reinforced by the fact that the density of the down-pulling slab increases significantly once it has passed the olivine-spinel-transition at roughly 400 km depth. The magnitude of slab pull is roughly  $10^{13} \text{ N m}^{-1}$  (s. Turcotte and Schubert 1982). Thus, slab pull is about an order of magnitude larger than ridge push. However, it is likely that slab pull is being counteracted by frictional stresses of about the same magnitude between the sinking plate and the surrounding asthenospheric mantle. Thus, the net force exerted by subduction zones onto the foreland need not be very large. Estimates by Bott (1993) and Bott et al. (1989) suggest that both slab pull and trench suction may be of the magnitude of roughly  $4 \cdot 10^{12} \text{ N m}^{-1}$ . In general it may be said that the force balance in subduction zones is much less well understood than that around mid-oceanic ridges. Nevertheless, most authors agree that forces



**Figure 5.30.** Four possibilities for the shape of subducted slabs at trenches. *a* Shallow subduction as probably occurs if subduction occurs in the same direction as the convective flow of the upper mantle (Doglioni 1993). *b* Steep subduction as it is often observed around the globe with slab break off at depth (Blanckenburg and Davies 1995). *c* Vertically hanging plate with folding at the 650-km-discontinuity (Frottier et al. 1995; Houseman and Gubbins 1997). *d* Hypothetical (in nature not observed) shape of a subducted plate that would occur if *no* deformation of the plate would occur subsequently to subduction in the trench

in and around subduction zones may be much larger than those exerted by the mid-oceanic ridges.

**Roll Back of Subduction Zones.** *Slab pull* and *trench suction* are predominantly forces acting downwards, while *ridge push* acts mainly in the horizontal direction. Slab pull and trench suction are not related to potential energy variations, but to *gravitational instabilities* similar to those that are responsible for convective motions. Because slab pull and trench suction act *downwards* it is possible that the kink in the subducting plate near the trench shifts laterally. If the rifiting rate at the mid-oceanic ridge exceeds the sinking rate of the subducting slab, then the trench will move *towards* the upper plate (Fig. 5.30). However, if the sinking rate is larger than the rifiting rate, then the trench moves *away* from the upper plate towards the mid-oceanic ridge (e.g. Dewey 1988). This process is called *roll back* of a subduction zone. The most famous example for roll back is the Scotia arc west of South Georgia. The formation of extensional basins in front of subduction zones, in particular the formation of Fore-arc- and Back-arc-basins, is thought to be related to roll back (Royden 1993a). Roll back does not only depend on the relative rates of rifiting and sinking of the oceanic plate, but also on how easily the asthenosphere may be displaced underneath the subducting plate.

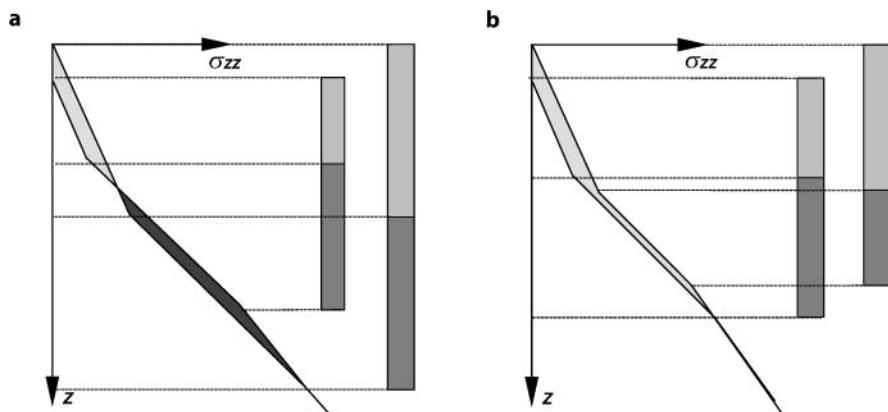
**Deformation of Subducting Plates.** The deformation of subducted plates in the upper mantle is not very well-understood. We know from bathymetric data that oceanic lithosphere is kinked at the trenches (Fig. 2.23). However, tomographic and seismic imaging indicates that subducted slabs are mainly planar slabs below the trench. This implies that the kinks of the subducted plates are unbent again at depth (Houseman and Gubbins 1997). Blanck-

enburg and Davies (1995) showed that slabs may also break off at depth (Fig. 5.30).

The long term evolution of subducting plates depends on the processes at the upper-lower mantle transition in about 650 km depth. Creager and Jordan (1984) showed that the processes there are of large importance to the possibility of recycling lithospheric material. In general it is thought that subducting plates cannot perforate this 650 km transition, predominantly because the density of the lower mantle is higher than that of the subducting slabs (e. g. Christensen and Yuen 1984) (Fig. 5.30c). This model has since been confirmed by tomographic imaging (e. g. van den Hilst et al. 1991; Gudmundsson and Sambridge 1998). It appears that the 650 km discontinuity is a *graveyard* for subducted slabs. There is a range of recent analogue and numerical experiments that test details of the geometry of deformation of subducted slabs at this discontinuity (Frottier et al. 1995).

### 5.3.3 Forces in Continental Plates

Inside the continents, plate tectonic driving forces arise predominantly from lateral variations in the density structure, which cause lateral variations in *potential energy*. When we discussed Fig. 5.27 we have already estimated the magnitude of these forces for a plate of constant density but variable thick-



**Figure 5.31.** Illustration of vertical stresses and potential energy differences between two neighboring lithospheric columns. Vertical normal stress is plotted as a function of depth. The shaded region between the two curves is the potential energy difference per area between the two adjacent columns. In **a** this difference is positive in the upper part of the lithosphere (light shading) but negative in the lower part (dark shading). This means, that there is a net force acting from the right hand column towards the left hand column, while this net force is directed towards the right in the lower part. Because both shaded regions are roughly of the same area, there is practically no net force between the two columns, averaged over the thickness of the lithosphere. In **b** the entire right hand lithospheric column exerts a net force onto the left hand column

ness (eq. 5.56). In this section we want to refine our estimates by considering a more realistic lithosphere consisting of a dense mantle lithosphere and a much less dense crust (s. Fig. 2.18). We also want to remember that momentum (often considered by students as an important contributor to mountain building processes) is completely irrelevant in plate tectonics (s. p. 218).

Fig. 5.31 illustrates two examples of potential energy differences between two lithospheric columns. Similar to Fig. 5.27 this potential energy difference is given by the shaded region between the two curves for vertical normal stress as a function of depth. This area corresponds to  $F_b$  in eq. 5.52 and may be interpreted as the net force exerted by one column onto the other per meter length of orogen and averaged over the thickness of the lithosphere (horizontal buoyancy force). In Fig. 5.31b the vertical normal stresses in the right hand lithosphere column is larger than that of the left hand column at all depths. Thus, there is a net force from the right towards the left column at all depths. However, for the two columns shown in Fig. 5.31a the situation is different. In the upper part of the profile the vertical stresses in the column of lower surface elevation are smaller. Interestingly, the vertical normal stresses are *smaller* for the column of *higher* surface topography in the lower part of the profile. This means that there is a net force exerted from the right hand column towards the left hand column in the *upper* part (light shaded region in the  $\sigma_{zz}$ -z-diagram), but that this force is directed in the opposite direction in the lower part (dark shaded region; see Problem 5.19).

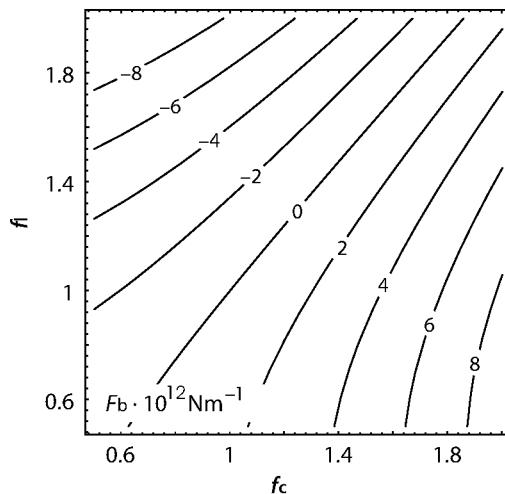
The qualitative considerations of Fig. 5.31 may be quantified by integrating eq. 5.52 and using simple descriptions for density as a function of depth. If we assume a simple lithosphere of two layers (a crust and a mantle lithosphere) and assume a linear thermal profile in the lithosphere so that the density due to thermal expansion may be described with eq. 4.29, then the lateral buoyancy force is described by:

$$\frac{F_b}{\rho_m g z_c^2} = \frac{\delta(1-\delta)}{2} (f_c^2 - 1) - \frac{\alpha T_l}{6(z_c/z_l)^2} (f_l^2 - 1 - 3\delta(f_c f_l - 1)) + \frac{\alpha^2 T_l^2}{8(z_c/z_l)^2} (1 - f_l^2) \quad (5.59)$$

(Turcotte 1983; Sandiford and Powell 1990). All parameters in this equation are the same as those we used in eq. 4.35 to calculate the elevation of mountain belts in isostatic equilibrium but the definition of  $\delta$  differs from that of Sandiford and Powell (1990) and the way eq. 5.59 is written here differs therefore slightly from theirs as well. Here (and elsewhere in this book)  $\delta$  is the density ratio of crust and mantle lithosphere  $\delta = (\rho_m - \rho_c)/\rho_m$ ,  $g$  is the gravitational acceleration,  $T_l$  the temperature at the base of the lithosphere and  $\alpha$  is the coefficient of thermal expansion. Lateral forces calculated with eq. 5.59 are shown in Fig. 5.32. The shape of these curves hardly changes for more refined assumptions on the thermal structure of the lithosphere (Zhou and Sandiford 1992). Fig. 5.32 shows that the absolute values of the net lateral forces exerted by very *thin* or very *thick* continental lithosphere on its

surroundings are of the order of  $10^{12}$ – $10^{13}$  N m $^{-1}$ . Thus, they are comparable to the magnitude of forces arising from ridge push or slab pull. This is really no surprise, as the thickness variations within the continental lithosphere are themselves only caused by plate driving forces in the oceans. Fig. 5.32 can also be used to explain why there is no place on the earth where the crustal thickness is significantly thicker than double of normal ( $f_c \gg 2$ ). Such regions can only form by forces in excess of  $F_b > 10^{14}$  N m $^{-1}$ , which is greater than any known plate tectonic driving force.

- *Potential energy excess created by external forces.* The potential energy of plates may not only be increased by internal deformation of the plates, but also by passively uplifting the entire plate, for example by the vertical stresses exerted from upwelling mantle plumes to the base of the lithosphere. McKenzie et al. (1974), McKenzie (1977a;b) as well as Houseman and England (1986b) showed that these forces are large enough to lift lithospheric plates by several hundreds of meters. On the abyssal planes of the oceans, such upwelling convection streams may even cause topography of the order of 1 km (Crough 1983; Watts 1976). Such topography has a higher potential



**Figure 5.32.** Diagram of lithospheric thickening strain  $f_l$  plotted against crustal thickening strain  $f_c$  and contoured for potential energy difference per area (equivalent to: “horizontal buoyancy force per meter” or: “lateral force”). The  $f_c$ - $f_l$  plane was explained in detail in Fig. 4.7. The potential energy difference is always that between any point in  $f_c$ - $f_l$  space and the reference lithosphere at  $f_c=f_l=1$ . The diagram was calculated with eq. 5.59 and is contoured for  $F_b$  in  $10^{12}$  N m $^{-1}$ . Other assumptions are:  $\rho_m = 3\,200$  kg m $^{-3}$ ;  $\rho_c = 2\,750$  kg m $^{-3}$ ;  $\alpha = 3 \cdot 10^{-5}$ ;  $z_c = 35\,000$  m;  $z_l = 125\,000$  m;  $T_l = 1\,200$  °C. The curvature of the contours arises because of the quadratic dependence of potential energy on thickness. As such, these contours for lateral force between two columns are fundamentally different from those for surface elevation (s. Fig. 4.18) (Sandiford and Powell 1990; Stüwe and Barr 2000)

energy than its surroundings and will - just like a mountain range in a region of thickened crust - exert a horizontal buoyancy force onto its surroundings. The geometry of this phenomenon is schematically drawn in Fig. 5.33. The horizontal force exerted by plate  $A$  onto plate  $B$  may also be calculated by integrating eq. 5.52, just as we did when calculating the potential energy difference between mountains and their foreland (Fig. 5.27). The magnitude of this force corresponds to the shaded region between the two curves for vertical stress as a function of depth in Fig. 5.33. The principal difference between this and the example discussed in Figs. 5.27 and 5.31 is that the vertical stresses of the two profiles do not converge, because the two columns in Fig. 5.33 are not in isostatic equilibrium. Integration according to the same principles we used in eq. 5.54 to eq. 5.57 gives:

$$F_b = \Delta E_p = \frac{\rho_m g H^2}{2} + \rho_c g H z_c . \quad (5.60)$$

Using the same numerical values for the physical parameters as we did in Fig. 5.32, we get for an uplifted elevation of  $H = 1$  km a horizontal force of the order of  $F_b \approx 9 \cdot 10^{11} \text{ N m}^{-1}$ . This results shows that convectional stresses in the mantle may have a significant influence onto the stress regime and therefore on the deformation of continents (s. sect. 6.1.4).

## 5.4 Problems

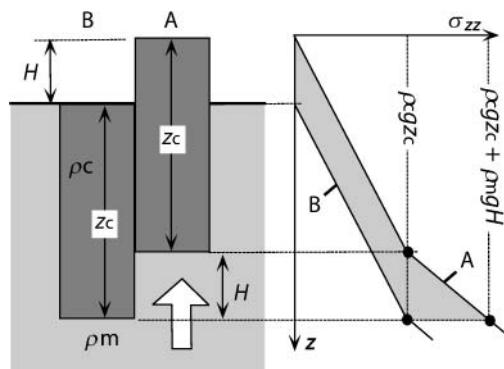
### Problem 5.1. Units of strain (p. 140):

During orogenesis a continental crust has thickened from 30 km to 60 km. What is the stretch, the elongation and the vertical strain it has experienced? Use eq. 4.1.

### Problem 5.2. Difference between weight and mass (p. 210):

What is the weight of 1 kg of rock at the surface? Give the result in SI units.

**Figure 5.33.** Illustration for the calculation of the potential energy change that is caused by actively lifting plate  $A$  with the thickness  $z_c$  and the density  $\rho_c$  by the amount  $H$ . Note that the two plates are *not* in isostatic equilibrium and that, therefore, the two curves for  $\sigma_{zz}$  do not meet at the depth  $z_c + H$ . The shaded area corresponds to the net horizontal force exerted by plate  $A$  onto plate  $B$



**Problem 5.3.** *Conversion of different energy forms (p. 213):*

A continent collides with another with a force of  $10^{13}$  N and rams 100 km into the other continent. How much mechanical energy is released in the process? Discuss where this energy goes, i.e. into which other forms of energy it may be transformed.

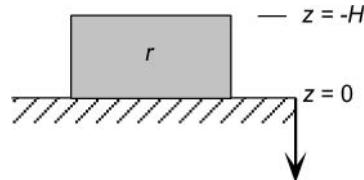
**Problem 5.4.** *Formulations of stress state (p. 215):*

A continent is under extension because an attached subducting plate pulls it apart. The tensional stress has the absolute value  $A$ . Another continent is under extension because it collapses under the weight of a mountain range on its surface. Let us assume that the vertical normal stress exerted by the mountain range onto the plate has also the magnitude  $A$ . Are the states of stress of the two continents the same? For simplicity, consider the two continental plates to be represented by little cubes that have no weight of their own: one that is pulled on its side and the other that is loaded from above. Consider the problem only in two dimensions.

**Problem 5.5.** *Stress balance, pressure and deviatoric stress (p. 210– 220):*

Fig. 5.34 shows a rock that lies atop the plane  $z = 0$ . a) How large are the vertical and horizontal normal stresses inside the rock at a given depth  $z$ ? b) How large is the pressure at this depth? c) How large are the components of the deviatoric stress tensor? For your answers, use the stress balance equations (eq. 5.15 and corresponding relationships in the other spatial directions; s. also eq. 6.25 and 6.26) and ignore elastic effects. Note that the coordinate system for this problem (Fig. 5.34) implies that the upper surface of the rock is at negative  $z$ . This is different from many other examples in this book where the origin of the vertical axis is often located at the highest point of the surface (see Fig. 4.3).

**Figure 5.34.** Illustration to Problem 5.5

**Problem 5.6.** *Stress balance, pressure and deviatoric stress (p. 210– 220):*

The rock from Problem 5.5 has melted. The shape it has in Fig. 5.34 is only maintained because we have put a box around it. a) How large are the vertical and horizontal normal stresses inside the rock in a given depth  $z$ ? b) How large is the pressure at this depth? c) How large are the components of the deviatoric stress tensor? For your answers, use the stress balance equations (eq. 5.15 and corresponding relationships in the other spatial directions; s. also eq. 6.25 and 6.26).

**Problem 5.7. Stress balance and deviatoric stress (p. 210– 220):**

Assume that the rock from Problem 5.6 cannot flow freely apart but pushes onto a fixed side wall (e.g. an indenter). How large is the force per meter that the rock exerts onto the side wall?

**Figure 5.35.** Illustration for Problem 5.8

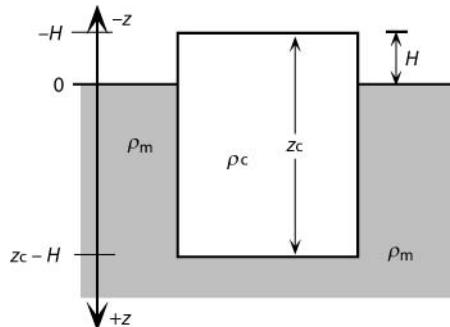
**Problem 5.8. Stress balance and deviatoric stress (p. 210– 220):**

Fig. 5.35 shows a rock of density  $\rho_c$ , floating in a fluid with the density  $\rho_m$ . Write relationships describing the principal stresses, the pressure and the principal components of the deviatoric stress tensor in the rock and in the fluid, both as a function of depth.

**Problem 5.9. Understanding elastic deformation (p. 221):**

Granite has a Young's modulus of roughly 50 GPa. What is the elastic strain of a granite to which a uniaxial stress of 50 MPa is applied? Use eq. 5.24 and eq. 5.25.

**Problem 5.10. Understanding elastic deformation (p. 221):**

What is the elastic change in thickness of the lithosphere (assuming that it is perfectly elastic) that arises solely as a function of its own weight? Assume the pre-elastic thickness  $z_l = 100$  km as well as  $E = 60$  GPa, and  $\rho = 3\,000 \text{ kg m}^{-3}$ .

**Problem 5.11. Fault plane solutions (p. 227):**

Draw qualitative fault plane solutions for the following faults: a) a north - south striking vertical dip-slip fault where the eastern block is displaced downwards; b) a horizontal fault where the top wall is displaced to the west; c) a north - south striking vertical fault at which the eastern side was displaced with roughly  $45^\circ$  towards the north and down; d) a thrust that dips a bit steeper than  $45^\circ$  to the east along which the upper plate was thrust obliquely to the north.

**Problem 5.12. Power law rheology (p. 232):**

Use the material constants for quartz and olivine from Table 5.3 (the activation energies  $Q$ , the pre exponent constants  $A$  and power law exponent

$n$ ) and calculate the shear stresses supported at strain rates of  $\dot{\epsilon} = 10^{-14}\text{s}^{-1}$  and a temperature of  $500^\circ\text{C}$ . Investigate how the stresses change if the power law exponents were actually 2 or 4 instead of 3. Compare your result with Fig. 5.12.

**Problem 5.13. Power law rheology (p. 232):**

The data from experiments that are used to derive the rheological material constants for viscous deformation (those required in eq. 5.46 are usually represented as lines in diagrams where  $\log(\sigma_d)$  is plotted against  $\log(\dot{\epsilon})$ , or in diagrams where  $\log(\dot{\epsilon})$  is plotted against  $1/T$ . Why ?

**Problem 5.14. Lithospheric strength (p. 244):**

Fig. 2.5 shows that seismicity along continental plate margins is distributed over much larger regions than seismicity along oceanic plate margins. This indicates that continental lithosphere is softer than oceanic lithosphere, although it has a larger thickness. Why is this so?

**Problem 5.15. Understanding potential energy (p. 248):**

Calculate the potential energy per unit area of the rock shown in Fig. 5.34.

**Problem 5.16. Understanding potential energy (p. 248):**

Estimate the potential energy (per unit area) of the two columns in Fig. 5.26 above the base of the lithosphere algebraically and graphically. For your graphic estimate, use the scheme illustrated in Fig. 5.31. For the calculation use the scheme explained in eq. 5.53 and  $z_c = 30\text{ km}$ ;  $z_l = 100\text{ km}$ ;  $\rho_1 = 2700\text{ kg m}^{-3}$ ,  $\rho_3 = 3200\text{ kg m}^{-3}$  and  $g = 10\text{ m s}^{-2}$ .

**Problem 5.17. Understanding gravitational stress (p. 252):**

What is the lateral buoyancy force exerted by a mountain range onto its foreland if the mountain range is characterized by a crust that is twice the thickness from the foreland ( $f_c = 2$ ), but the same thickness of the entire lithosphere ( $f_l = 1$ ). Use eq. 5.57 and  $z_c = 30\text{ km}$ ,  $\rho_c = 2700\text{ kg m}^{-3}$ ,  $\rho_m = 3300\text{ kg m}^{-3}$  and  $g = 10\text{ m s}^{-2}$ . In order to use eq. 5.57 we also need to know the surface elevation of the range  $H$  and the thickness of the root  $w$ , both at  $f_c = 2$  and  $f_l = 1$ . Recalling eq. 4.35 the elevation is  $H = 5454\text{ m}$ . Consequently, the thickness of the root is  $w = 24546\text{ m}$ . (See also Problem 6.9.)

**Problem 5.18. Understanding gravitational stress (p. 258):**

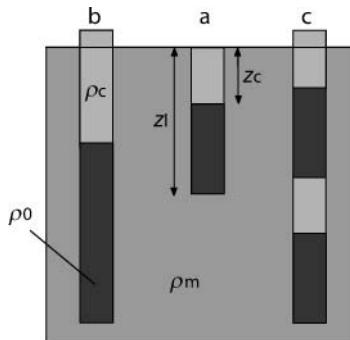
Refine your estimate from Problem 5.17 using eq. 5.59 with  $T_l = 1200^\circ\text{C}$ ,  $z_l = 100\text{ km}$  and  $\alpha = 3 \times 10^{-5}\text{ }^\circ\text{C}^{-1}$ . (Note that the surface elevation need not be known when using eq. 5.59). Compare your result with the result from Problem 5.18 and the graphically presented result in Fig. 5.32.

**Problem 5.19. Understanding gravitational stresses (p. 257):**

Fig. 5.36 shows a schematic lithosphere: a) Normal thick lithosphere of thickness  $z_l$  with crust of thickness  $z_c$ ; b) after homogeneous thickening to double

thickness ( $2z_l$ ,  $2z_c$ ); c) after doubling the entire lithosphere by over thrusting. How large are the net horizontal forces that the columns exert on each other? Estimate the result graphically using the scheme used in Fig. 5.31. Calculate the result using eq. 5.49 to eq. 5.53) and use the parameter values given in the caption of Fig. 5.36.

**Figure 5.36.** Illustration to Problem 5.19.  $\rho_c = 2700 \text{ kg m}^{-3}$ ,  $\rho_0 = 3300 \text{ kg m}^{-3}$ ,  $\rho_m = 3200 \text{ kg m}^{-3}$ ,  $g = 10 \text{ m s}^{-2}$ . Note that the surface elevation of the columns in b and c is the same as both columns have the same weight



**Problem 5.20. Understanding momentum (p. 218):**

A 100 km thick continental plate that has a mean density of  $3000 \text{ kg m}^{-3}$  and is  $1000 \times 1000 \text{ km}$  large collides with a much larger continent at a velocity of  $0.03 \text{ m y}^{-1}$ . The collision stops the plate. How large is the kinetic energy of the plate? How high can a mountain belt be built if all kinetic energy of the collision process is transformed into potential energy? Assume that all other forces that may apply to the plates may be neglected.

# 6. Dynamic Processes

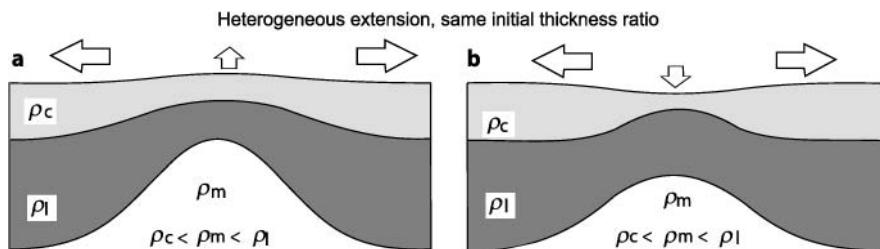
This chapter is the first of two chapters in which we integrate the information of the previous three chapters into “real” tectonic models. The first two thirds of this chapter are dedicated to the description of continents in extension and in collision. In the last third of this chapter we touch upon a range of interesting and currently very topical geodynamic problems.

## 6.1 Continents in Extension

Under certain stress states continents may extend. In the process, they usually decrease in their thickness. In general, extensional processes are divided into *active* and *passive* processes (e.g. Allan and Allen 1990; Ruppel 1995). An extensional process is considered to be *active* if the extension occurs as the consequence of forces inherent to the extending area, for example gravitational collapse of a region of high potential energy. Examples are mid ocean ridges, high continental plateaus or regions actively uplifted by mantle plumes (Keen 1980). Active rifting is therefore also called “mantle generated” and we have estimated the forces that cause active rifting in eq. 5.60. We describe an extension process to be *passive* if the forces causing extension are applied *outside* the extending area, for example the force of a subducting plate that pulls at the passive margin of an adjacent continent (e. g. Le Pichon 1983). Passive rifting is also called “lithosphere generated” rifting. However, we should remember that the stress state of an extending plate is identical for both active and or passive extension (s. Fig. 5.3 and Problem 5.4).

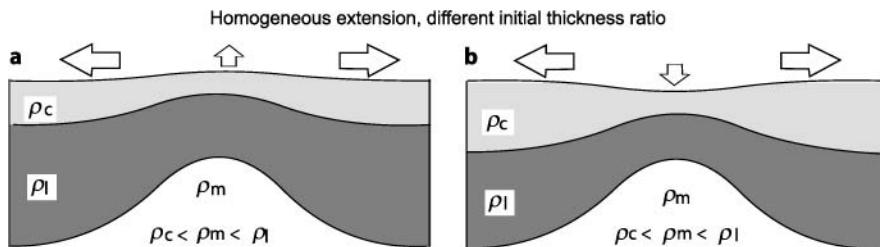
- *Uplift or subsidence.* It is not trivial that extension of the lithosphere must lead to subsidence of the surface. The density of continental crust is *lower* than that of the underlying asthenosphere and the density of the mantle lithosphere is *higher* (s. Fig. 2.18). Whether or not extension leads to subsidence depends therefore on the partitioning of the extensional strain between the crust and the mantle part of the lithosphere (Fig. 6.1). It also depends on the initial thickness ratio of the two prior to onset of extension (Fig. 6.2 s. also: Fig. 4.18). McKenzie (1978) showed that surface subsidence during extension will only occur during homogeneous lithospheric extension if the crustal thickness at the onset was more than 14 km, as part of a thermally

equilibrated continental lithosphere. If the crust is thinner than this, then lithospheric extension will lead to surface uplift.



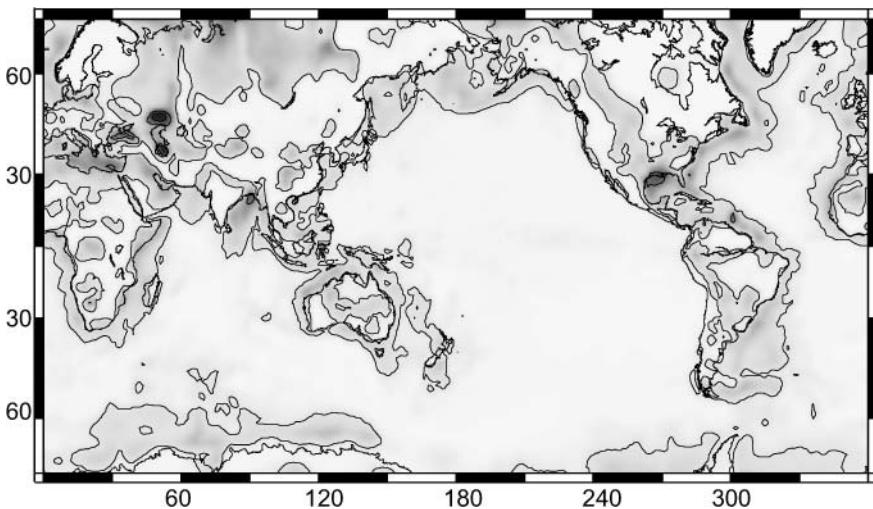
**Figure 6.1.** The influence of the heterogeneous stretching on surface uplift/subsidence during lithospheric extension. The variables  $\rho_c$ ,  $\rho_l$  and  $\rho_m$  are the density of the crust, mantle lithosphere and asthenosphere, respectively. In a extension causes surface uplift, because the mantle lithosphere is stretched more than the crust. In b extension causes subsidence because the crust is stretched more than the mantle part of the lithosphere

Subsidence and the development of sedimentary basins must not have been caused by extension: Several types of sedimentary basins form in collisional environments during crustal thickening. However, the processes of *continental extension* and *sedimentary basin formation* are so intimately related, that this first section of this chapter is concerned with the origin, nature and processes of the formation of all types of sedimentary basin. The processes that form basins in collisional environments are also discussed here.



**Figure 6.2.** The influence of the initial thickness ratio of crust (light shaded area) and mantle part of the lithosphere (dark shaded areas) on the nature of the vertical motion of the surface during homogeneous lithospheric extension. The variables  $\rho_c$ ,  $\rho_l$  and  $\rho_m$  are the density of the crust, mantle lithosphere and asthenosphere, respectively. In a extension causes surface uplift, because the mantle lithosphere constitutes a large proportion of the lithosphere by thickness. In b extension causes subsidence because the crust forms a larger proportion of the lithosphere

The temporal evolution of the subsidence of sedimentary basins is one of the most important sources of information for the geodynamic interpretation of continental extension processes. In many ways, the processes and



**Figure 6.3.** Global sediment thickness map. Contours are shown for 1, 10 and 15 km sediment thickness. Greatest sediment thicknesses are in the oceans, in the Bengal Fan and the Gulf of Mexico, greatest thicknesses on land are north of the Caspian Sea (data from Laske and Masters 1997; see also Divins 2006; data set at: <http://www.ngdc.noaa.gov/mgg/sedthick/sedthick.html> )

methods of interpretation of the subsidence of basins is analogous to the processes and interpretation of surface uplift during collisional processes (e.g. in sect. 4.3.1, 4.4.1). However, they differ fundamentally in one respect: the thickness of a basin fill is usually substantially larger than the amount of subsidence that could be accredited to tectonic processes alone. Earth scientists first explained this phenomenon at a time when the COST drill holes on the continental shelf of the eastern USA were drilled by the ocean drilling program: the density of sediments is substantially larger than that of water or air. Thus, in isostatically compensated basins, sedimentary loading of tectonically formed basins will cause additional subsidence which in turn makes room for additional sediment loads.

We therefore discriminate between *tectonic* subsidence and *total* subsidence. *Total* subsidence is the total amount of vertical change of the former surface. The rate of downward motion of the former surface is called the subsidence rate. As such this definition is not quite analogous to the definition of uplift rate (s. sect. 4.3.4). The *tectonic* subsidence is only the component of total subsidence that is caused by tectonic mechanisms. In order to interpret the tectonic processes that lead to sedimentary basin formation, knowledge of the total subsidence (i.e. the sediment thickness in the basin) is only the starting point. It is then necessary to subtract the influence of sedimentary loading from the total subsidence to determine the tectonic contribution to

subsidence (Sclater and Christie 1980; Steckler and Watts 1978; 1981). This exercise is referred to as *back stripping* and will be discussed in sect. 6.1.3.

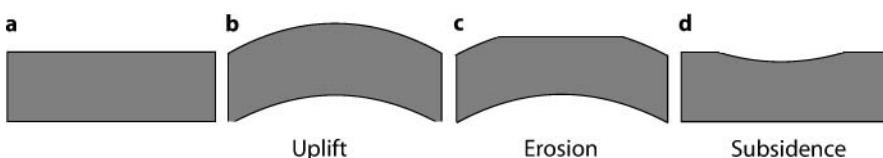
### 6.1.1 Basin Subsidence Mechanisms

The subsidence of just about all sedimentary basins is caused by one or more of the following three processes. Each of these processes has been discussed in previous sections of this book:

- isostatic subsidence,
- flexural subsidence,
- thermal subsidence.

The three mechanisms may be intimately related. In the following we discuss briefly the nature of these three mechanisms and then go on to explain different types of sedimentary basins using these terms. Good summaries of models for the development of sedimentary basins are published by Angevine et al. (1990) and Allen and Allen (1990) and we rely here – in part – on their work.

- *Isostatic subsidence* is caused by physical changes in the thickness of the lithosphere. For example, if physical stretching of the lithosphere causes thinning, then isostatic compensation will generally lead to subsidence (sect. 4.4.1; Fig. 6.4).
- *Flexural subsidence* relies on elastic bending of the lithosphere (s. sect. 4.4.2). If the lithosphere is loaded, it bends and a basin forms near the load (Fig. 4.22). For very strong plates, such basins are wide and shallow, while for less competent plates such basins are narrow and deep. However, the basin volume is independent of the rigidity of the plate (s. Fig. 4.25).
- *Thermal subsidence* occurs if the density structure of the lithosphere is thermally changed by cooling (sect. 4.4.1). Thus, thermal subsidence is also a type of isostatic subsidence, except that the thickness change is caused thermally and not mechanically. As cooling of the lithosphere occurs only in thermally destabilized lithosphere, thermal subsidence can only occur in



**Figure 6.4.** Sketch illustrating one of the first models for the development of sedimentary basins (Sleep 1971). A continent is uplifted by external processes, for example by upwelling asthenospheric material of a mantle plume. Erosion thins the uplifted region. When the uplifting process terminates, the continent returns to its original position and a shallow sedimentary basin is formed

lithosphere that has previously been heated. Everything else being equal, the amount of thermal subsidence during cooling is exactly as large as the amount of thermal uplift during the heating phase. Thus, no sedimentary basin can form as the consequence of thermal processes on their own. Erosion or extension must have separated the thermal uplift and thermal subsidence in order to form a sedimentary basin (Fig. 6.4; sect. 6.1.4; McKenzie 1978).

### 6.1.2 Basin Types

Different types of sedimentary basins were classified by Dickinson (1976) according to the three subsidence mechanisms discussed above (s. also Buck 1991).

- *Passive margins and rift basins.* Rift basins form as the consequence of continental extension and ultimately rifting (Fig. 2.27a). The extensional process during the formation of rift basins may be symmetrical (Keen et al. 1989) or asymmetrical (Wernicke 1985; Lister et al. 1986; 1991) about the rift axis. The subsidence associated with the isostatic compensation of the rifting is usually followed by a later phase of thermal subsidence during which the mechanically rifted mantle lithosphere thickens by cooling. Thus, the subsidence of rift basins may usually be divided into a *rift phase* and a *sag phase*, both of which are characterized by specific sedimentary environments. During the rift phase, sedimentation is rapid, highly energetic and associated with the development of half grabens and other tectonic structures. During the sag phase, sedimentation is slow and static. Both phases are best developed if the rifting has not gone to completion. Examples for rift basins that *have* gone to completion in their successive stages of development are to be found in the Rhein Graben, the East African Rift system, the Red Sea and the Atlantic coast.
- *Transform basins.* Transform- or pull-apart basins also form due to continental extension. The most important difference between these and proper rift basins is that they are smaller because their extensional phase terminated much earlier. Transform basins never get to a rifting stage. They are bound on at least two sides by strike slip faults and they are usually rectangular or diamond shaped. Because of their limited size, heat conduction processes do not only occur in the vertical direction, but also in the lateral direction. As a consequence, thermal thinning of the mantle lithosphere is limited. Therefore, transform basins usually lack the sag phase that is so typical for rift basins (Pitman and Andrews 1985). Subsidence of transform basins is usually short-lived and is largely a linear function of time. This is especially because the deformation history of the brittle structures that control their shape is very short-lived. Other structures in the same orogen will soon take over and accommodate the strain. Examples of transform basins are the Death Valley in California, the Vienna Basin in Austria as well as a large number of small intramontane basins within the European Alps.

- *Foreland basins.* Foreland basins form during the collision of two continental plates and are the continental analogue to fore-arc and back-arc basins (see next paragraph). The principle subsidence mechanism is elastic flexure of the plate in response to the loading by external and internal loads (Beaumont 1981; Jordan 1981b; Karner and Watts 1983). According to their location relative to the lower plate, foreland basins may be divided into two groups (Fig. 2.23). *Peripheral* foreland basins form near subduction zones in collisional environments as a consequence of loading of the lower plate by the upper plate. *Retroarc* foreland basins form on the upper plate in the hinterland of a subduction zone. Good examples of peripheral foreland basins are the molasse basins near the Alps or the Himalayas. Examples of retroarc foreland basins are those that form east of the Andes and also the molasse basins of the Alps. The subsidence rate in peripheral foreland basins may be used to determine the rate of loading of the plate and therefore ultimately the collision rate of two plates.
- *Fore-arc and Back-arc basins.* Fore-arc-basins have their name because they form in front of an island arc. There is a range of models that have been used to explain their origin, but none of them is really well-constrained or completely satisfactory. Some of the models include: 1. Subduction of an oceanic plate underneath another leads to a doubling of the plate thickness beneath the accretionary wedge. Since the density of oceanic lithosphere may be higher than that of the underlying asthenosphere, doubling the plate thickness leads to subsidence and formation of a fore-arc basin. 2. Subduction of a cold plate underneath a hot plate may cause cooling of the upper plate and thus lead to thermal subsidence and basin formation. 3. Loading of the plate from above by an island arc and loading from below by the buoyancy of the accretionary wedge may lead to elastic back-bending of the plate. Basins on oceanic lithosphere that form behind a subduction zone are called *back-arc* basins. Their formation is usually interpreted as the consequence of upwelling asthenospheric material in the mantle wedge and they have a characteristic basaltic signature (Taylor and Martinez 2003). However, they also have been thought to be connected with potential energy differences (Stüwe and Barr 2000). Examples of Fore-arc basins are those developed along the south margin of Sumatra and Java, along the east margin of Japan and a famous ancient example is the Californian Great Valley. Examples of back-arc basins are the Mariana back-arc basin west of the Mariana Ridge, or the North Fiji Basin.
- *Intracontinental basins.* Some large sedimentary basins form intracontinentally, for example the Michigan Basin in the USA. The amount of tectonic subsidence in these basins is rarely more than 2 km. Their round shape and slow subsidence rates indicate thermal subsidence as a subsidence mechanism. However, the origin of these basins remains largely unconstrained.

### 6.1.3 Subsidence Analysis

The large variety of tectonic processes that form sedimentary basins discussed above illustrates the necessity to find a data set that can be collected in the field and that can be used to constrain the nature of the basin forming process. Such a data set exists in the subsidence history of a basin as recorded by the sedimentary basin fill. In order to constrain the evolution of subsidence from the stratigraphic record, three steps are necessary:

- Documentation of the stratigraphic section,
- Consideration of compaction of sediments,
- Consideration of the water depth.

If we are also interested in the tectonic component of the subsidence (as we usually are), then a fourth step is necessary:

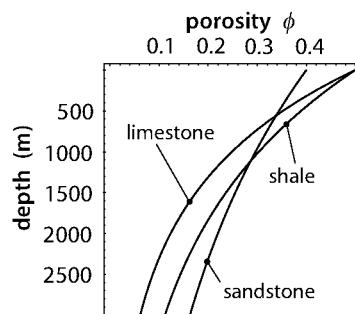
- Consideration of the sedimentary loading: back stripping.

When mapping the stratigraphy of the basin fill with the intention of using it for subsidence analysis, the following data must be collected or assumed for each layer: 1. thickness, 2. lithology, 3. age and 4. water depth at deposition. Porosity of the sediments and information on the thermal evolution are additional data that can be extremely helpful. On the following pages we show how the subsidence history may be extracted from this data.

**Compaction.** Because of their porosity, sedimentary strata are compacted by overlying layers after their deposition. Thus, the thickness of each layer in a sedimentary sequence was larger at the time of its deposition than it is when measured in the field. In order to consider the influence of sediment compaction on the thickness and density of the stratigraphic column, the porosity must be known. Empirical studies show that the porosity of rocks decreases exponentially with depth. In general we can describe this with the relationship:

$$\phi = \phi_0 e^{-cz} \quad . \quad (6.1)$$

**Figure 6.5.** The decrease of porosity of a range of rock types with depth. Calculated with eq. 6.1 and using the following material constants. Sandstone:  $\phi_0 = 0.4$ ,  $c = 3 \cdot 10^{-4} \text{ m}^{-1}$ ; limestone:  $\phi_0 = 0.5$ ,  $c = 7 \cdot 10^{-4} \text{ m}^{-1}$ ; slate:  $\phi_0 = 0.5$ ,  $c = 5 \cdot 10^{-4} \text{ m}^{-1}$ . The grain density  $\rho_g$  of these three rock types is: sandstone:  $\rho_g = 2650 \text{ kg m}^{-3}$ ; shale:  $\rho_g = 2720 \text{ kg m}^{-3}$ ; limestone:  $\rho_g = 2710 \text{ kg m}^{-3}$  (data from Sclater and Christie 1980; s. also Bond et al. 1983)



There,  $\phi$  is the porosity of the rock at depth  $z$ ,  $\phi_0$  is the porosity at the surface and  $c$  is a rock specific compaction constant (Fig. 6.5). If we now want to derive the original thickness of a layer  $L_0$  from that measured for this layer in the field,  $L$ , we must solve an integral. Assuming that the thickness was only changed by changing the porosity (and not by metamorphism, diagenesis, cementation or dissolution), it must be true that:

$$\int_0^{L_0} (1 - \phi) dz = \int_{z_L}^{z_L + L} (1 - \phi) dz . \quad (6.2)$$

This equation states that the rock volume without the pore space ( $1 - \phi$ ) remains a constant, regardless of the fact whether the upper surface of the layer  $L$  is at depth  $z = z_L$  or at  $z = 0$ . Eq. 6.2 is a one-dimensional volumetric balance. This equation forms the basis of our following considerations. Clearly it is important to check the applicability of this equation before embarking on a compaction analysis. If the sediment is cemented or partially dissolved, then it is easily possible that the processes did *not* occur at constant volume and the compaction analysis becomes more complicated. It may then be necessary to study the cementation material petrographically to see if it was derived internally or externally, and so on. It comes back to the universal fact that field and laboratory data determine how simple a model is allowed or how complicated it must be designed.

By substituting eq. 6.1 into eq. 6.2 and integrating we get:

$$L_0 + \frac{\phi_0}{c} e^{(-cz_0)} (e^{(-cL_0)} - 1) = L + \frac{\phi_0}{c} e^{(-cz_L)} (e^{(-cL)} - 1) . \quad (6.3)$$

Sadly, it is impossible to solve eq. 6.3 for  $L_0$  - the original depositional thickness of a layer of which we measured the thickness  $L$  at depth  $z$ . When using eq. 6.3,  $L_0$  can only be determined numerically by iteration (sect. B.5.2). However, for most cases this is not necessary. It is usually sufficient to use the following approximation:

$$L_0 = L \frac{(1 - \phi)}{1 - \phi_0} . \quad (6.4)$$

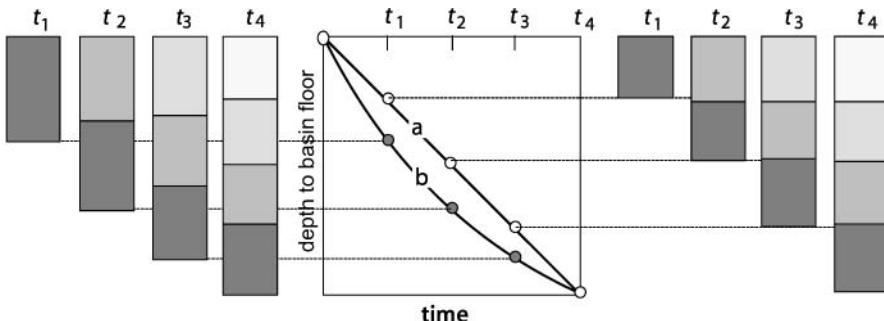
Using eq. 6.4 it is now possible to calculate the original thickness of a layer at the end of its sedimentation  $L_0$  from field data on the porosity  $\phi$  and the thickness  $L$  as measured from a drill hole. For this, the porosity  $\phi_0$  is calculated using eq. 6.1. The original thickness of the layer is needed for the further steps in the subsidence analysis. Obviously, eq. 6.4 can also be used to determine the thickness of a layer at any other stage of the decompaction process  $L^*$  and not only the fully decompacted thickness  $L_0$ . We just need to use the porosity (and depth) at the right stage of the analysis  $\phi^*$  instead of the original porosity  $\phi_0$ .  $\phi^*$  can also be calculated with eq. 6.1 (Fig. 6.5).

If we have additional information on the porosity or thermal evolution of our rocks, it is possible to refine the compaction analysis. However, we do not want to discuss this here. We only want to remember that - using eq. 6.1

and eq. 6.4 – it is possible to calculate subsidence curves for the basin floor. The method is also illustrated on Fig. 6.6.

The subsidence evolutions that may be calculated with the method described above give us the depth evolution of the basin floor underneath the surface of the basin fill as a function of time. However, it is important to note that the surface of the basin fill need not remain at a constant depth below (or above) sea level. If, for example the water depth in a sedimentary basin changes over time, then the water depth must be added to the subsidence curve to obtain the subsidence evolution relative to a fixed reference level. In marine basins, the water depth at the time of deposition can usually be constrained by lithology, sedimentary structures and fossil record. If the sea level itself changes during deposition, the interpretation becomes more difficult. Sometimes it is possible to document sea level changes by comparing synchronous stages in the sedimentation record of two independent basins subjected to the same sea level change. Basins that develop in terrestrial environments are much harder to interpret, as it is difficult to document the changes of surface elevation through time and the water depth is not so easy to determine.

**Backstripping.** In the previous sections we have shown how to determine the evolution of total subsidence (i. e. the depth evolution of the basin floor) as a function of time. This subsidence history is the sum of *tectonic subsidence* and subsidence caused by the *sedimentary loading*. The process of determin-



**Figure 6.6.** Illustration showing the influence of compaction on a sedimentary section. The columns on the left and right of the central diagram show cartoons of stratigraphic columns as they evolve through times  $t_1$ ,  $t_2$ ,  $t_3$  and  $t_4$ . On the left these columns are for sediments that compact during successive deposition, on the right these columns are for sediments that do not compact. For both sets of columns the oldest layer is shaded the darkest, the youngest the lightest. At present ( $t_4$ ) the stratigraphy of both profiles is identical. However, these identical columns at time  $t_4$  were arrived at by different sedimentation and subsidence histories. On the right sedimentation rates were constant through time and none of the units were compacted (curve *a* on the central diagram). On the left, sedimentation rates were rapid at time  $t_1$  and then decreased through time (curve *b*). The similar thickness of all units at time  $t_4$  arises here only because compaction has balanced the variations in sedimentation rate for the columns on the left

ing the *tectonic* contribution to subsidence from the *total* subsidence is called *back stripping* (Watts and Ryan 1976; Steckler and Watts 1978; Sclater and Christie 1980). The inverse of back stripping – *back stacking* – is also possible, using thermochronological data from sedimentary basins (Brown 1991).

Back stripping is a mental exercise in which layers are successively removed from the sedimentary column of a basin. During each step of removal, the hypothetical depth of the basin floor *without* being loaded is calculated. In other words, we determine the isostatic loading of the column during successive steps of sedimentation. Depending how careful we need to do this (as required by the large scale geological situation, for example the size of the basin) we discern between:

- Back stripping assuming hydrostatic isostatic compensation,
- Back stripping assuming flexural isostatic compensation.

Back stripping assuming hydrostatic isostasy is straight forward as all we need to do is apply eq. 4.23. We will illustrate this initially with an example of a basin filled by a single layer.

• *Back stripping of a single layer.* For our Example we assume a marine (i. e. water covered) basin that was created by a single tectonic process. We further assume that the tectonic process started when the surface was at sea level and caused a tectonic subsidence of the amount  $z_T$ . Today, the basin is filled by water of the depth  $w$  and a single sedimentary layer of thickness  $L$  and density  $\rho_L$ . On Fig. 6.7 we can see that the tectonic subsidence ( $z_T$ ) can be written as the sum of the water depth at present  $w$  and the basin depth change due to sedimentary loading:  $z_T = w + z_s$ . We can use this relationship to determine the tectonic subsidence. So we first need to calculate  $z_s$ : the depth of the basin floor prior to the sedimentary fill below sea level. The calculation of  $z_s$  follows the same principle as the relationships we have used for the calculation of isostatic equilibrium (sect. 4.4.1; Problem 6.4), (Steckler and Watts 1978):

$$z_s = L \left( \frac{\rho_m - \rho_L}{\rho_m - \rho_w} \right) , \quad (6.5)$$

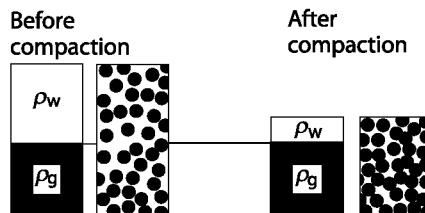
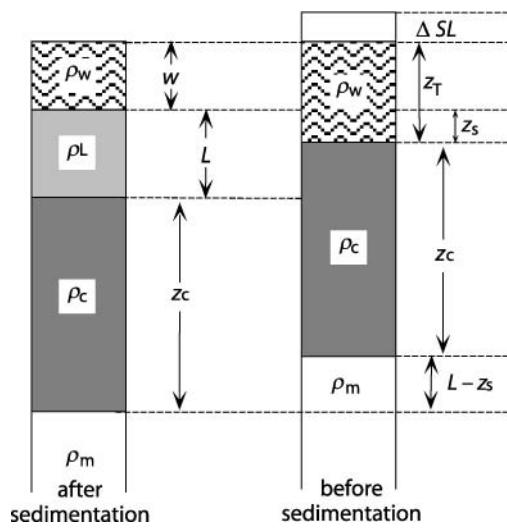
if we disregard any change in sea level during the period of loading. Somewhat more refined we can also write:

$$z_T = L \left( \frac{\rho_m - \rho_L}{\rho_m - \rho_w} \right) + w - \Delta SL \frac{\rho_m}{\rho_m - \rho_w} , \quad (6.6)$$

if we formulate the relationship in terms of  $z_T$  and also consider a sea level change that may have occurred during sedimentation. The densities  $\rho_m$ ,  $\rho_L$  and  $\rho_w$  are those of the asthenospheric mantle, that of the sediment layer  $L$  and that of water,  $w$  is the water depth and  $\Delta SL$  is the change in sea level during the deposition of the unit of thickness  $L$  (Fig. 6.7).

The thickness of the crust or even that of the lithosphere are not needed for eq. 6.5 or eq. 6.6, as we assume that no thickness change of the crust

**Figure 6.7.** Cartoon illustrating the origin of eq. 6.5 (s. also Problem 6.4). The right hand column shows a crustal profile in isostatic equilibrium *after* tectonic subsidence but *before* fill of the depression by sediments. The left hand profile shows the total subsidence of the column after the sedimentation of a layer with the thickness  $L$ .  $\rho_c$ ,  $\rho_m$  and  $\rho_w$  are the densities of crust, asthenospheric mantle and water and  $z_c$  is the thickness of the crust.  $z_s$  is the change of surface after removing the sedimentary basin fill and compensating isostatically. As  $\rho_c$  and  $z_c$  are the same in both columns, they cancel out in eq. 6.5

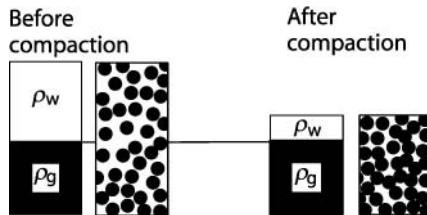


**Figure 6.8.** Cartoon illustrating eq. 6.7. Compaction of sediments decreases the water filled pore volume, but not the grain volume (unless diagenesis accompanies compaction). This is schematically shown by two profiles of water volume and grain volume for both the situation before compaction and after compaction. The variable  $\rho_w$  is the density of the pore fluid and  $\rho_g$  is the grain density

occurred during the sedimentation, i.e. for the purpose of this example we have assumed that the tectonic amount of subsidence occurred entirely before the onset of sedimentation. If the basin was not filled by water at any stage of the evolution, then  $\rho_w$  must be substituted by  $\rho_{\text{air}} = 0$  for the corresponding period. Eq. 6.5 is analogous to eq. 4.26, but it is solved for another variable. The density of the porous sediment layer  $\rho_L$ , that occurs in eq. 6.5 may be determined from the grain density  $\rho_g$  and the pore fluid density  $\rho_w$  (which - in most cases - is the density of water:  $\rho_w = 1\,000 \text{ kg m}^{-3}$ ) from the relationship:

$$\rho_L = \phi \rho_w + (1 - \phi) \rho_g , \quad (6.7)$$

if the porosity  $\phi$  is known. eq. 6.7 is illustrated in Fig. 6.9. The amount of tectonic subsidence of basins that are filled by more than one sedimentary



**Figure 6.9.** Cartoon illustrating eq. 6.7. Compaction of sediments decreases the water filled pore volume, but not the grain volume (unless diagenesis accompanies compaction). This is schematically shown by two profiles of water volume and grain volume for both the situation before compaction and after compaction. The variable  $\rho_w$  is the density of the pore fluid and  $\rho_g$  is the grain density

layer may also be determined with eq. 6.5. All we need to do is use the mean values for  $L$  and  $\rho_L$  from all layers in the sequence.

- *Back stripping for multi layered sedimentary basins.* Using eq. 6.5 we can determine the complete evolution of tectonic subsidence by stepwise removal of the top layer at any one stage during the analysis and performing the analysis as above. For the remaining column mean densities and thicknesses must be used at each time step. The value  $z_T$  is then the tectonic amount of subsidence during sedimentation of the top most layer only and  $L^*$  and  $\rho_{L^*}$  are the thickness and density of the entire remaining sedimentary column after removal of the top layer  $i$  (i.e. the decompacted thickness). The thickness of a sediment pile with  $i$  layers is then:

$$L^* = \sum_{j=1}^i L_j . \quad (6.8)$$

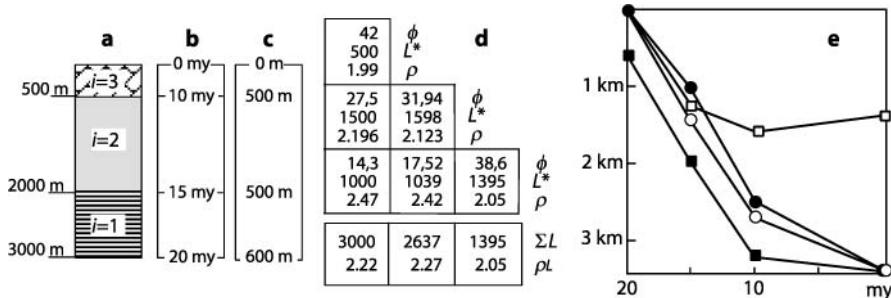
The density of the sedimentary column underneath layer  $i$  is given by the mean density of all remaining layers. This is simply the sum of all densities, multiplied by the respective thickness and divided by  $L^*$ :

$$\rho_{L^*} = \frac{\sum_{j=1}^i L_j (\phi_j \rho_w + (1 - \phi_j) \rho_g)}{L^*} . \quad (6.9)$$

Now we can use eqs. 6.8 and 6.9 to determine the tectonic subsidence history of a basin by stepwise reconstruction. If we do this by hand, then we need to iteratively apply eq. 6.1, eq. 6.5 and eq. 6.7 (using  $L^*$  and  $\rho_{L^*}$  instead of  $L$  and  $\rho_L$ ). An example is shown in Fig. 6.10 (s. also Problem 6.5). The analysis can be refined by considering flexural isostasy instead of hydrostatic isostasy as we did here. Computer programs that do sedimentary basin analysis follow exactly this procedure.

#### 6.1.4 Models of Continental Extension

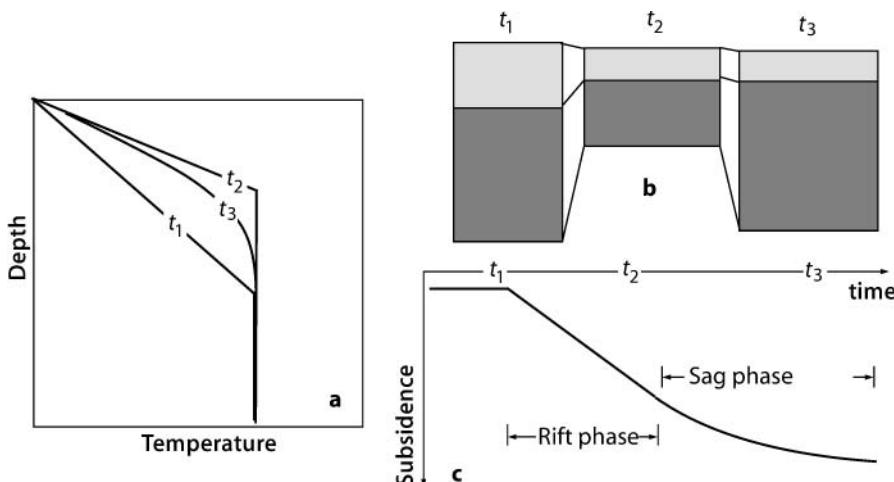
We have shown in the last section how a careful analysis of the sedimentary basin fill may be used to constrain the tectonic subsidence history of a



**Figure 6.10.** Example for a subsidence analysis with consideration of compaction. **a, b** and **c** show the field data of a simple basin filled by three sedimentary strata. **a** Lithology and thickness of strata (slate: horizontal lines; sandstone: shaded; limestone: brick signature); **b** age of the unconformities; **c** water depth as derived from fossil record. No sea level changes occurred during the sedimentation process. Strata are numbered from base to top. The center of each layer is the reference point used for the following calculations. For the values of porosity and density we use the data from Fig. 6.5. Porosity and density of the top layer (with a mean depth of 250 m) may be derived from eq. 6.1 and eq. 6.7. These data are written in the 1st column of Table **d**. The second column lists the porosity of the second layer (with a mean depth of 750 m = mean depth of second layer minus the thickness of the first layer) and so on. The thickness of the second layer is given by eq. 6.4 using the porosities of the 1. and 2. column as well as the thickness from the first column. The sum of the thicknesses and densities in the bottom two rows were calculated with eq. 6.8 and eq. 6.9 respectively. Their mean values are written in the third column. In **e** the result of the subsidence analysis are illustrated graphically. *Black dots* show the measured field data, *white dots* the calculated decompacted thicknesses and *black squares* show the decompacted thicknesses plus water depth. The curve given by the white squares was calculated with eq. 6.5 and shows the tectonic component of subsidence

basin. Based on analyses of this kind it was recognized that the evolution of subsidence of many sedimentary basins follows very process-specific patterns. Some of these characteristic patterns and simple models that have been used to explain them are discussed on the next pages.

**The McKenzie- and its Follow up Models.** Subsidence analysis has shown that the tectonic subsidence rate of many sedimentary basins is rapid at first and then decreases abruptly to continue at a much slower rate for a much longer time. The model of McKenzie (1978) was one of the first and certainly most famous model that has successfully explained this pattern. Despite its simplicity it remains the basis of a large range of more refined models. Like other models of its time (e.g. Le Pichon et al. 1982) the model is one-dimensional and describes the subsidence of the surface as a function of lithospheric extension. Within the model, typical subsidence histories are divided into two phases: a *rift phase* and a *sag phase*. The model holds well in examples where the phase of physical extension of the lithosphere (rift phase) is short compared to the duration of the subsequent thermal equilibration (sag



**Figure 6.11.** Subsidence of sedimentary basins according to the model of McKenzie (1978) and that of Jarvis and McKenzie (1980). **a** Geotherms at three different times: prior to onset of extension ( $t_1$ ), at the end of a rapid stretching phase ( $t_2$ ) and during subsequent thermal equilibration of the lithosphere ( $t_3$ ). **b** Thickness of the crust (light shaded region) and that of the mantle part of the lithosphere (dark shaded region) at these three different time steps. **c** Schematic subsidence curve corresponding to the cartoons in **a** and **b**

phase). In fact, within the original McKenzie (1978) model, the stretching phase (rift phase) is assumed to have occurred instantaneously (Fig. 6.11). The amount of subsidence during this instantaneous rifting phase  $H_{\text{rift}}$  may be calculated with the relationship:

$$H_{\text{rift}} = z_c \left( \frac{\rho_0 - \rho_c}{\rho_0 - \rho_w} \right) \left( 1 - \frac{1}{\delta} \right) - z_l \left( \frac{\rho_0 \alpha T_l}{2(\rho_0 - \rho_w)} \right) \left( 1 - \frac{1}{\beta} \right) . \quad (6.10)$$

Equation 6.10 is largely analogous to eq. 4.35, which we have used previously to describe the changes in surface elevation as a function of changed thicknesses of crust and mantle part of the lithosphere. Therefore, we do not derive the details of eq. 6.10 here and the reader is referred to sect. 4.4.1. In eq. 6.10,  $z_c$  and  $z_l$  are the thickness of crust and lithosphere.  $\rho_c$ ,  $\rho_w$  and  $\rho_0$  are the densities of crust, water and mantle at 0°C.  $\alpha$  is the coefficient of thermal expansion,  $T_l$  is the temperature at the base of the lithosphere,  $\delta$  the stretching parameter of the crust and  $\beta$  that of the mantle part of the lithosphere. The stretching parameters are defined as the ratio of starting thickness to stretched thickness and the stretching parameter for the crust  $\delta$  is therefore the inverse of the thickening strain  $f_c$  which we have used in several other sections of this book:  $\delta = 1/f_c$  (s. sect. 4.2). However,  $\beta$  does not correspond to  $1/f_l$ , but rather to  $1/f_{ml}$ . Because of this eq. 6.10 looks somewhat different from eq. 4.35, although both may be used to answer similar questions (Fig. 4.18).

Eq. 6.10 is already a refined McKenzie model: it allows to explore the consequence of different amounts of stretching in the crust and in the mantle lithosphere (s. e. g. Royden and Keen 1980). In the original McKenzie model  $\delta = \beta$ . In other words, the original McKenzie model assumed homogeneous stretching of the entire lithosphere.

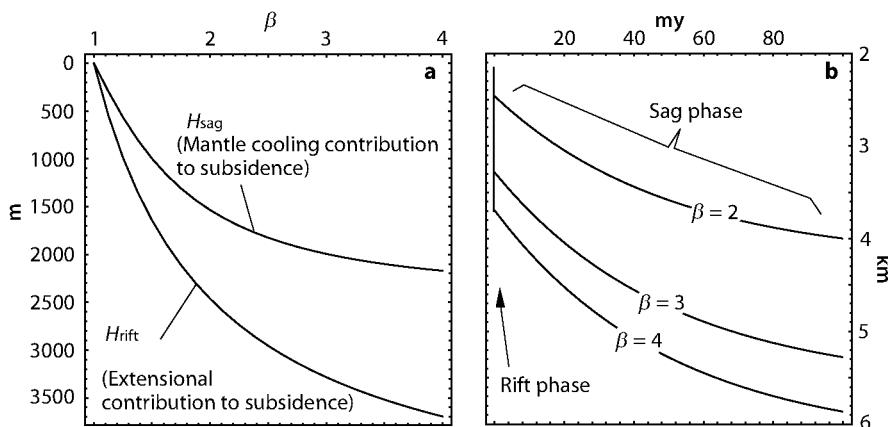
When we discussed eq. 4.35 we saw that the initial ratio of crust to mantle lithosphere determined if homogeneous thickening of the whole lithosphere leads to surface uplift or subsidence. The same is true for eq. 6.10 (s. Fig. 6.2, Problem 6.6).

The subsidence described by eq. 6.10 (Fig. 6.12a) is followed by a phase of thermal equilibration: the *sag phase*. During this subsequent phase, active extension has stopped and thickening of the mantle part of the lithosphere occurs due to cooling. The surface subsides as a consequence of this cooling (Fig. 6.11). This is largely analogous to the surface subsidence associated with cooling of the oceanic lithosphere (Fig. 3.24). Thus, we can describe this subsidence with a similar model to the one we have discussed in sect. 3.5.1. The principle difference to the cooling of oceanic lithosphere arises because McKenzie assumed the bottom boundary not at infinity, but at depth  $z_1$ . Because of this boundary condition at *finite* depth, the heat conduction equation may only be solved with the aid of Fourier series (s. sect. B.4). As a consequence, the solution presented below contains trigonometric functions instead of an error function (s. sect. B.4). The solution is:

$$H_{\text{sag}} = \left( \frac{4\rho_0\alpha T_1 z_1}{\pi^2(\rho_0 - \rho_w)} \right) \left( \frac{\beta}{\pi} \sin(\pi/\beta) \right) \left( 1 - e^{-t/t_{\text{eq}}} \right) . \quad (6.11)$$

This relationship describes the subsidence of rift basins during the sag phase as a function of time  $t$ . The parameter  $t_{\text{eq}}$  is the time scale of thermal equilibration and is given by:  $t_{\text{eq}} = z_1^2/(\pi^2\kappa)$  (which differs from the time constant discussed on p. 63 by the factor  $\pi^2$ , but is in principle equivalent). While we will not derive the origin of this equation in any detail, we note that it should be possible, at least in principle, to follow it from the information given in previous chapters of this book (s. e. g. eq. 4.69). In contrast to eq. 4.69, eq. 6.11 contains no infinite summations, because it is an approximation in which all terms for  $n > 1$  were omitted. Aside from this, eq. 6.11 (and also eq. 4.69) describe a similar model to that of eq. 4.45, which we have used to calculate the water depth in the oceans. Examples of subsidence during the sag phase are shown in Fig. 6.12b. For very large times, the last term of eq. 6.11 becomes 1 and the simplified equation may be used to estimate the total amount of subsidence during the entire sag phase.

- *Finite extension.* In contrast to the assumptions of the McKenzie model, many sedimentary basins show evidence that the duration of physical stretching or rifting of a continent was not instantaneous, but was comparable to the time scale of diffusive thermal equilibration of the lithosphere, i. e. rifting is slow and rift and sag phase are, in fact, of similar duration. In order to account for finite duration of the rifting phase Jarvis and McKenzie (1980) and



**Figure 6.12.** a Subsidence of rift basins at the end of the rift phase ( $H_{rift}$ ) and at the end of the sag phase ( $H_{sag}$ ) both as a function of the amount of initial stretching as given by  $\beta$ . It is assumed that the lithosphere stretched homogeneously, i. e.  $\delta = \beta$  (s. eq. 6.10). The total amount of subsidence is given by the sum of the two curves. For  $\beta < 1$  the curve go to negative values, i. e. thickening leads to uplift. b Subsidence during the rift and subsequent sag phase as a function of time for two different values of  $\beta$  (describing different amounts of rifting at the onset). All curves were calculated using eq. 6.10, eq. 6.11 and the following values for the variables:  $z_1 = 125$  km;  $z_c = 35$  km;  $T_1 = 1\,280$  °C;  $\rho_0 = 3\,300$  kg m $^{-3}$ ;  $\rho_c = 2\,750$  kg m $^{-3}$ ;  $\rho_w = 1\,000$  kg m $^{-3}$ ;  $\alpha = 3 \cdot 10^{-5}$  °C $^{-1}$  and  $\kappa = 10^{-6}$  m $^2$  s $^{-1}$

Cochran (1983) expanded the McKenzie model. Jarvis and McKenzie (1980) suggested as a rule of thumb that the duration of stretching,  $t$ , must only be considered in their model for basin development if the following relationship holds:

$$t < \frac{60\text{my}}{\beta^2} \quad , \text{if } \beta < 2 \quad \text{or:} \quad t < 60\text{my} \left(1 - \frac{1}{\beta}\right)^2 \quad , \text{if } \beta > 2 \quad . \quad (6.12)$$

If the duration is shorter than time  $t$ , then it is sufficient to assume that stretching was instantaneous and occurred prior to any thermal equilibration.

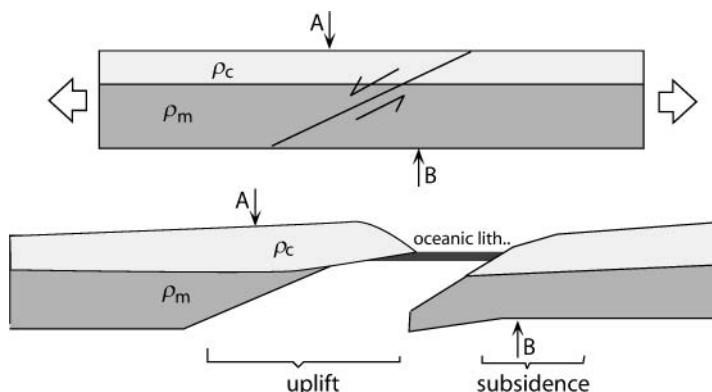
- *Two-dimensional models for continental extension.* Most rift basins are one-dimensional in their geometry, i. e. they are long compared to their width. If however, lithospheric stretching occurs two-dimensionally so that basins extend into two directions, then the modeling of subsidence during the sag phase must account for two-dimensional heat conduction. Some of the first two-dimensional models describing this were designed by Buck et al. (1988), Issler et al. (1989) and Wees et al. (1992). Those models assume that extension occurs symmetrically about the rifting axis.

- *Heterogeneous extension.* One of the first models describing asymmetric extension of the lithosphere was designed by Oxburgh (1982). Oxburgh assumed that both the crust and the mantle part of the lithosphere extend

homogeneously, but that the location of maximum crustal extension is laterally displaced from the location of maximum extension of the mantle part of the lithosphere. Wernicke (1985) and Lister et al. (1986) were among the first who explored the consequences of simple shear geometries on lithospheric extension processes by assuming low-angle normal faults that transect the entire lithosphere. These models were based on observations in the Basin and Range province of the western US. Lister and Etheridge (1989) and Lister et al. (1991) applied this model to the east coast of Australia in order to explain the simultaneous uplift of the Australian Great Dividing Range and the subsidence of continental lithosphere at the west coast of New Zealand (Fig. 6.13).

An elegant analytical solution describing some aspects of the models of Wernicke and Lister was published by Voorhoeve and Houseman (1988). Some geometric considerations of extension models based on heterogeneous stretching geometries may be found in Buck et al. (1988).

**Dynamic Extension Models.** In all previous paragraphs we have considered the evolution of extension and sedimentary basin formation purely on the basis of kinematic and thermal assumptions. We will now consider *dynamic* models, that is, models for continental extension that are based on mechanical assumptions. One of the many questions that can only be explained with dynamical models relates to the causes for terminating extension: we observe that some rift basins extend until rifting occurs and a passive continental



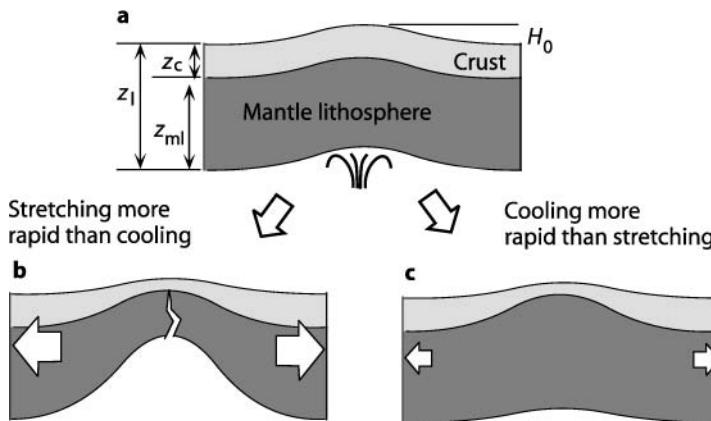
**Figure 6.13.** Extension of the lithosphere by normal faulting along a low angle normal fault that transects the entire lithosphere. The upper cartoon shows the situation at the onset of extension, the lower cartoon after full separation of the two plates and the development of passive continental margins. Note that at location *A*, extension only decreased the thickness of the mantle part of the lithosphere, while at location *B*, extension only decreased the thickness of the crust. Lister et al. (1986) and Wernicke (1985) interpreted that the uplift of the left hand plate at location *A* (e. g. eastern Australia) and the subsidence of the right hand plate at location *B* (i. e. western New Zealand) are only caused by the changed thickness ratio of crust and mantle lithosphere at these two locations

margin forms (sect. 2.4.4), while in others the extension was limited. Examples for the former are the central African rift or the Red Sea, examples for the latter include the Michigan Basin in the US, the Cooper Basin in Australia or the Pannonian Basin in Europe. There are two fundamentally different processes that may be responsible for the termination of extension:

1. *External reasons*: In this case the termination of extension has nothing to do with the plate itself and is only a function of processes in the surrounding plates. Extension is controlled by the boundary conditions. The forces or velocities with which a plate is pulled apart and a basin is formed may simply decrease.
2. *Internal reasons*: Extension may terminate because the rheology of the plate changes to become stronger. In this case no changes in the boundary conditions need to occur; the plate boundary forces may remain constant. The extension process itself is self limiting (s. sect. 6.2.2).

Analogous to the terms *active* and *passive extension*, these two processes might be called *active* and *passive termination* to extension. Which of these two mechanisms caused the termination of a particular basin is a question which often can only be answered using dynamical models. Some of the first dynamic models for continental extension were those of Bassi (1991), Bassi et al. (1993) and Cloetingh et al. (1995). These models showed that the extension geometry is strongly dependent on the tectonic processes in a plate that occurred *prior* to the onset of extension. For example, they showed that continental extension onsets most easily in regions of thickened continental crust, as this is the region where the plate is the weakest (s. sect. 5.2.1, 6.3.6; Houseman and England 1986b). As another example, it was shown by Buck (1991) that the width of continental rift basins depends on the geothermal gradient and ultimately on the rheology. One of the more elegant models explaining the temporal limitation of extension processes is that of Houseman and England (1986b). In the following section we discuss this model in some more detail.

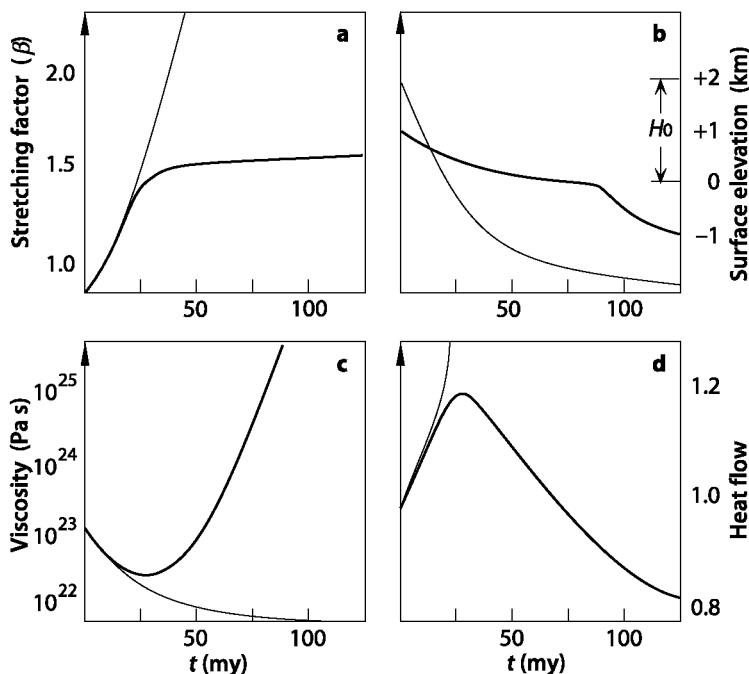
- *The model of Houseman and England*. The one-dimensional model for continental extension by Houseman and England (1986b) is an elegantly simple model that can be used to illustrate why extension can be self limiting. The model is a coupled thermal and mechanical model in which extension is a function of both thermal and rheological development. For this, the authors assumed a model geotherm and a model rheology for the lithosphere. Active extension is assumed to be driven by upwelling asthenosphere of a mantle plume which initially lifted the entire continental plate by the amount  $H_0$  (Fig. 6.14a). Note that the extension is called “active” because it is triggered by potential energy of the lithosphere that was created by an external process. Fig. 6.15 illustrates the evolution of extension in response to this uplift (s. also Fig. 5.33). According to the calculations of Houseman and England (1986b) this extension can ultimately lead to three different scenarios:



**Figure 6.14.** Possible evolution of an active extension process initiated by uplifting of the entire lithosphere for the amount  $H_0$  by upwelling asthenospheric material over mantle plumes. **a** Starting geometry at the onset of extension. **b** If the extensional forces are large enough, then the lithosphere is thinned and thermally weakened until a rift zone develops. **c** If the extensional forces are small, thickening (and strengthening) of the mantle lithosphere by cooling is more rapid than physical stretching. The extension process is self limiting. Note that the mantle lithosphere in **c** is thickest in the region of maximum crustal thinning

1. If  $H_0$  is smaller than about 100 m, then the extensional forces caused by the excess potential energy are too small to be reflected in any appreciable strain rate.
2. If  $H_0$  is of the order of several hundreds of meters, then the extension is self limiting. The extensional forces are large enough to cause extension at an appreciable rate, but this rate is still slower than the rate of thickening of the mantle lithosphere by cooling (as in Fig. 6.11 between  $t_2$  and  $t_3$ ). Thus, the mantle part of the lithosphere thickens, while the crust thins. This leads to a successive strengthening of the lithosphere (because the rheology of the mantle lithosphere is much stronger than that of the crust; s. sect. 5.2.2) and ultimately to a termination of the extension process (thick lines in Fig. 6.15; Fig. 6.14c).
3. If  $H_0$  is of the order of one or even several kilometers, then the extension rates are rapid enough so that physical extension outweighs thermal thickening (as in Fig. 6.11 between  $t_1$  and  $t_2$ ). The geotherm steepens, the lithosphere is thermally weakened and the excess potential energy leads to an acceleration of the extension process. Extension leads to rifting and ultimately to the development of a new passive margin (thin lines in Fig. 6.15; Fig. 6.14b).

The basic principles of the force balance underlying the scenarios discussed above will be explained in sect. 5.3.1 and 6.2.2. Models equivalent to the model of Houseman and England (1986b) but with application to continental



**Figure 6.15.** Two model evolutions of continental extension according to the model of Houseman and England (1986b). Thick lines show a self limiting extension process. Thin lines are for conditions that ultimately lead to rifting. **a** the evolution of the stretching factor; **b** evolution of surface elevation; **c** evolution of viscosity; **d** evolution of surface heat flow (normalized to that at the onset of extension). The scaling of the vertical axes of all four plots depends strongly on the model assumptions. The shoulder in the evolution of surface elevation as seen for the thick line in **b** corresponds to the transition from rift phase to sag phase

*collision* will be discussed in the context of Fig. 6.23 (Sonder and England 1986; Molnar and Lyon-Caen 1989).

## 6.2 Continents in Collision

When two plates of continental lithosphere move toward each other, they will eventually collide. In contrast to the collision between two oceanic plates (during which one of the plate generally dives below the other and *no* intense deformation occurs in either of the plates), continental collision leads to intense deformation and interfingering of both plates (Fig. 2.22). This fundamental difference between the collision of continental and oceanic plates occurs because of three reasons:

- because of their different thickness,

- because of their different strength,
- because of their different densities.

Continental lithosphere is much *thicker*, much *weaker* much *less dense* than oceanic lithosphere. The lower density and larger thickness of continental plates make it more difficult to subduct them underneath each other. However, importantly, the weakness of continental lithosphere allows internal pervasive deformation including nappe stacking, folding and more, while the much larger strength of oceanic lithosphere does not allow such processes. This internal deformation leads to mountain building, metamorphism and a series of other tectonic events which we can observe today in active orogens and which are preserved for us in the metamorphic rocks of ancient orogens.

Although only a relatively small number of all plate boundaries around the globe are formed by collisional orogens, such orogens are among the best studied tectonic features of our planet. This is certainly in part because continental collision processes form some of the most eye catching features on the planet: the spectacular mountain ranges of the Himalayas or the European Alps. In this section we deal with aspects of the thermal and dynamical evolution of such orogens.

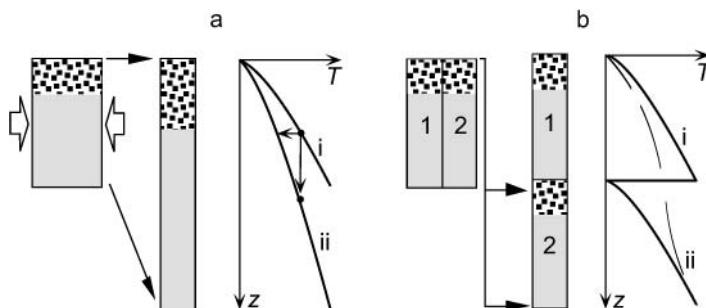
### 6.2.1 Thermal Evolution of Collisional Orogenes

It is widely observed that collision of two continental plates leads to heating of rocks at depth. In present day orogens this is evidenced by increased heat flow. In ancient orogens it is documented by the preservation of metamorphic rocks from a well known range of metamorphic facies. However, heating and convergent deformation need not occur simultaneously. In fact, in many convergent orogens it is observed that peak metamorphism occurred later than the deformation phases that were responsible for the crustal thickening.<sup>359</sup> Many aspects of this typical relationship between deformation and metamorphism may be explained by a simple comparison of the duration of three processes: 1. The duration of crustal thickening processes; 2. The duration of thermal equilibration of the crust and 3. the time scale of exhumation processes. A comparison of these three time scales was the basis of the model by England and Richardson (1977) which is one of the first elegant models unifying deformation and metamorphic processes in collisional orogens into a single model. This model will form the basis of the following section. Precursors of the model were already published by Oxburgh and Turcotte (1974) as well as Bickle et al. (1975) and the model was quantified later by England and Thompson (1984), but the model is often referred to as the “England and Richardson” model and we use this name here. Today, their concept is still the standard model to explain “regional” or: “Barrovian” metamorphism (s. Fig. 7.1).

**Fundamentals of the Thermal Evolution.** England and Richardson (1977) recognized that the following relationships are fundamental to the thermal characteristics of collisional orogens (s. also Ridley 1989):

- Thickening of the crust is substantially more rapid than thermal equilibration on the scale of the crust. The former generally occurs at rates of the order of  $\dot{\epsilon} \approx 10^{-14} \text{ s}^{-1}$  (i. e. doubling the crustal thickness in less than 10 my), the latter takes of the order of several tens of my (as given by eq. 3.17).
- The time scale of thermal equilibration is comparable to that of many exhumation processes, in particular to erosion: both often take several tens of my.

The time scale of deformation and the time scale of exhumation are functions of the geological boundary conditions of the orogen in question and may vary. Their absolute and relative magnitudes are known to us from field observations. In contrast, the time scale of thermal equilibration is independent of the geological processes. It is given directly by the laws of heat conduction and depends largely on the length scale of the equilibrating body (s. sect. 3.1.4). These relationships give rise to the following stages in the tectonic evolution of collisional orogens:



**Figure 6.16.** Cartoon illustrating the thermal changes that occur in the crust as the consequence of rapid crustal thickening. **a** and **b** show two end member scenarios of crustal thickening geometries: homogeneous crustal thickening in **a** and overthrusting of the whole crust (block 1 over block 2) in **b**. In each **a** and **b** a cartoon of a crustal column is shown on the left (the radioactive heat producing part of the crust is shown with the coarse dots) and a corresponding temperature - depth ( $T$ - $z$ ) diagram is shown on the right. The  $T$ - $z$  diagrams show geotherms before thickening (labeled *i*) and after thickening but before thermal equilibration (labeled *ii*). The dashed line in **b** is the geotherm during early subsequent thermal equilibration. It shows that the “saw tooth geotherm” has rapidly equilibrated to a  $T$ - $z$  profile not unlike *i* in **a** (rapidly in comparison to the overall duration of subsequent thermal equilibration). Thus, the overall thermal evolution of thickened crust is robust towards the initial thickening geometry. The vertical arrow between *ii* and *i* in **a** indicates the  $T$ - $z$  path of a rock during thickening. The horizontal arrow shows that thickening leads to cooling at a constant depth

• *Thermal evolution during thickening.* The fact that continental deformation is typically much more rapid than thermal equilibration on the length scale of the crust has two immediate consequences: 1. Thickening leads to burial of rocks without appreciable heating. This is indicated in Fig. 6.16a by the small vertical arrow in the  $T$ - $z$ -diagram. 2. From the point of view of an Eulerian observer the crust cools. This is indicated in Fig. 6.16a by the small horizontal arrow in the  $T$ - $z$ -diagram. Most of the interplay of heating and cooling mechanisms that governs the thermal evolution of collisional orogens occurs *after* the thickening. In order to understand this interplay it is useful to discuss heating and cooling mechanisms separately.

• *Heating mechanisms following thickening.* There are two mechanisms that lead to heating:

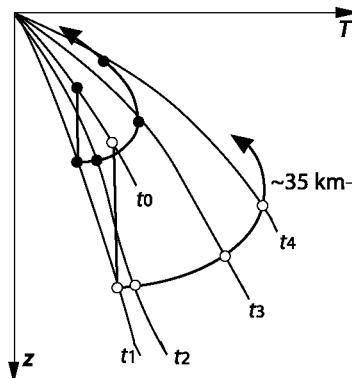
1. Prior to thickening the stable gotherm formed an equilibrium between surface heat flow *out* of the crust, the heat flow from the mantle *into* the crust and the radiogenic heat production *in* the crust. This equilibrium was perturbed by the thickening process. The geotherm was cooled. Given that the mantle heat flow does not change, heating will occur to reestablish the equilibrium geotherm.
2. Thickening of the crust increases the total amount of radiogenic elements in each vertical column the crust (coarse dotted part in Fig. 6.16). This increased radiogenic heat production has the consequence that the new equilibrium geotherm will have a steeper gradient than before. Heating will therefore occur in order to attain an equilibrium beyond the initial stage.

• *Cooling mechanisms following thickening.* Crustal thickening leads not only to heating, but also to mountain building. This is followed by erosion and extension processes thinning the crust and leading to exhumation of metamorphic rocks. In analogy to the last paragraph, there are two cooling processes that interact with the heating processes:

1. Denudation of the upper crust removes the heat producing elements from the upper crust. This has the consequence that the thermal equilibrium heat flow decreases in the whole crust.
2. All rocks must cool to the surface temperature by the time they reach the surface. The closer rocks get exhumed to the surface, the stronger the cooling influence of this surface becomes.

The thermal evolution of rocks in a given orogen depends on the interplay and competition of the heating and cooling mechanisms listed in the last paragraphs. Usually, heating mechanisms outweigh cooling mechanisms in the early phase following the crustal thickening. During the later evolution, in particular after the onset of denudation at the surface, the heating mechanisms wane and the influence of cooling mechanisms increases. Eventually, cooling processes win over the heating processes and rocks begin to cool

**Figure 6.17.** Geotherms (thin lines) and  $T$ - $z$ -paths (thick lines) of rocks in typical collisional orogens as predicted by the model of England and Richardson (1977).  $t_0$  is the starting geotherm at the onset of collision,  $t_1$  the geotherm immediately after thickening,  $t_2$ ,  $t_3$  and  $t_4$  geotherms at later time steps. The Moho heat flow (given by the slope of the lower tip of all geotherms) is assumed to remain constant throughout the evolution. The black and white dots that lie on each geotherm label the evolution of rocks at two different depths. Note that the rocks of the shallower  $T$ - $z$ -path (black dots) experience their thermal peak at time  $t_3$  and cool thereafter, while the rocks at larger depths (white dots) reach their metamorphic thermal peak only at time  $t_4$ .



(s. Fig. 6.17). The competition between these various heating and cooling mechanisms has the consequence that the  $T$ - $z$ -paths of rocks in collisional orogens have smooth curvature in a  $T$ - $z$  diagram. They follow this curvature in a clockwise sense if the depth axis is drawn upwards and the temperature axis is drawn to the right (Fig. 7.7). Depending on the relative importance of heating and cooling mechanism, such paths are more or less tight in a  $T$ - $z$  diagram (Fig. 6.17). The slope and curvature of  $T$ - $z$  paths is therefore an important indicator for tectonic processes (Fig. 7.5). It is the art of the metamorphic petrologist to extract such information from metamorphic mineral equilibria.

**The Model of England and Thompson.** England and Thompson (1984) designed a quantitative numerical model based on the “England and Richardson” model discussed above. Their model is one of the most simple and intuitive models that can be used to explain the shape and nature of metamorphic  $P$ - $T$  paths in collisional orogens. Thus, we will discuss the assumptions and results of their model in some more detail here.

- *Thermal assumptions.* The thermal assumptions made by England and Thompson (1984) to describe the evolution of collisional orogens are those necessary to calculate an initial geotherm with eq. 3.23 (s. sect. 3.4.3. Following boundary conditions are assumed:

- The temperature at the surface of earth  $T_s$  is constant.
- The mantle heat flow at the Moho  $q_m$  is also constant.

These boundary conditions are assumed to remain the same throughout the thermal evolution. It is also assumed that the radiogenic heat production is of the constant value  $S_{\text{rad}}$  down to the depth  $z_{\text{rad}} = 15$  km and is zero below that (sect. 3.4.3). With those assumptions, integration of eq. 3.23 gives a description of a stable geotherm as given by eq. 3.65.

The numerical values needed in this equation are those for  $q_m$  and  $S_{\text{rad}}$ . England and Thompson chose them so that the surface heat flow  $q_s$  (which is one of the few thermal parameters that may be measured directly) is between 0.045 and 0.075  $\text{W m}^{-2}$ , which is reasonable for continental shield regions (Table 6.1).

- *Heat flow relationships.* The relationship between surface heat flow, mantle heat flow and radioactive heat production can be illustrated clearly by interpreting the surface heat flow  $q_s$  as the sum of the mantle heat flow  $q_m$  and the heat flow caused by radiogenic heat production  $q_{\text{rad}}$ :

$$q_s = q_m + q_{\text{rad}} . \quad (6.13)$$

In this equation, the radiogenic heat flow is given by:  $q_{\text{rad}} = S_{\text{rad}} z_{\text{rad}}$ , as we explained when we discussed Fig. 3.17 (see also eq. 3.61). England and Thompson (1984) assumed that the radiogenically caused heat flow is comparable to the mantle heat flow ( $q_{\text{rad}} \approx q_m$ ; Table 6.1) and that the mantle heat flow remains unchanged, regardless of the thickness of the crust (the problems and implications of this are discussed on p. 293).

Thickening of the crust without thickening of the mantle part of the lithosphere doubles the radiogenic heat flow (because  $z_{\text{rad}}$  is doubled) but does nothing to the mantle heat flow. We can write:

$$q_s = q_m + 2q_{\text{rad}} . \quad (6.14)$$

Thus, the surface heat flow in thermal equilibrium after thickening is expected to be of the order of 1.5 times as high as before if  $q_{\text{rad}} = q_m$  (Eq. 6.13).

However, if the mantle part of the lithosphere thickens together with the crust (homogeneous lithospheric thickening), then this halves the heat flow through the Moho (as the mantle lithosphere is thermally defined). We can then write:

**Table 6.1.** Three different simple but geologically realistic assumptions about the distribution of heat sources in the crust (after England and Thompson 1984). The mantle heat flow  $q_m$  and the radiogenic heat production  $q_{\text{rad}}$  are chosen in a way so that they give three different values for surface heat flow covering a realistic range of measured surface heat flows. The thickness of the heat producing layer in the crust is always assumed to be  $z_{\text{rad}} = 15$  km. England and Thompson have coupled these three heat source distributions with three different thermal conductivities (1.5, 2.25 and 3.0  $\text{W m}^{-1} \text{K}^{-1}$ ) to obtain a total of nine groups of  $T$ - $z$ -paths

$q_s (\text{W m}^{-2})$	$S_{\text{rad}} (\cdot 10^{-6} \text{ W m}^{-3})$	$q_{\text{rad}} (\text{W m}^{-2})$	$q_m (\text{W m}^{-2})$
0.045	1.666	0.025	0.020
0.060	2.000	0.030	0.030
0.075	2.333	0.035	0.040

$$q_s = \frac{q_m}{2} + 2q_{\text{rad}} \quad . \quad (6.15)$$

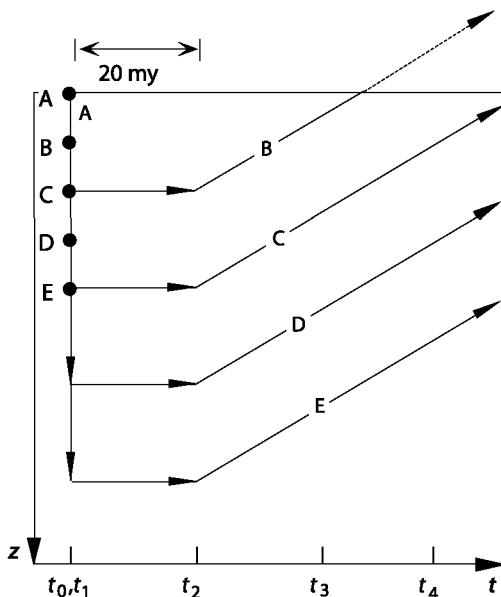
Thus, if  $q_{\text{rad}} = q_m$  and the entire lithosphere thickens to double thickness, the surface heat flow in thermal equilibrium and after thickening would be only 1.25 times as large as the value given by Eq. 6.13. If  $q_{\text{rad}} = q_m/2$ , then thickening or thinning of the lithosphere as a whole does not change the surface heat flow at all.

- *Kinematic assumptions.* England and Thompson (1984) made the following simple assumptions about the kinematic evolution of collisional orogens (s. Fig. 6.18):

- Thickening of the crust occurs homogeneously and instantaneously at the start of an orogenic cycle (there is no separation between  $t_0$  and  $t_1$  in Fig. 6.18; s. p. 286). This assumption was made because many observations of plate tectonic velocities show that the rate of continental deformation is indeed about an one or two orders of magnitude more rapid than the lifetime of an orogen (at rates of centimeters per year, crustal thickening events last only few millions of years, while orogenic cycles last of the order of tens to one hundred million years).
- Following initial thickening, there is no vertical motion in the crust for about 20 my (Fig. 6.18). This assumption was made based on the observation that erosion and extension (causing the most important vertical motions in orogens) do not commence with the onset of thickening, but only once substantial topography is developed and potential energy is created.
- After 20 my erosion sets in. Denudation of the thickened crust is linear in time and lasts for several tens of my. While this assumption may be also an extreme simplification, it is the most simple assumption that can be made for the orogen to return to its normal thickens after some tens of my.

These three kinematic assumptions are illustrated in Fig. 6.18. The  $z$ - $t$ -paths in this figure correspond to homogeneous thickening. Fig. 6.16 shows the thickening geometry and geotherms during the initial thickening process for two end member scenarios of thickening explored by England and Thompson (1984): homogeneous thickening and stacking of the entire crust. The dashed curve in Fig. 6.16b is drawn for a short time (some few my) after thickening by thrusting. It shows that the subsequent  $T$ - $z$ -paths of rocks are quite robust towards the geometry of initial thickening. This justifies the assumption of the most simple of all thickening geometries.

- *Model results and application.* Fig. 6.17 illustrates the thermal evolution of a rock in a collisional orogen according to the model of England and Richardson (1977). According to this figure and the model, the involved processes occur in the following order:



**Figure 6.18.** Kinematic assumptions of the continental collision model of England and Thompson (1984). The depth-time- $(z-t)$ -paths A–E correspond to 5 different rocks that originally (before thickening) were located at the depths indicated by the black dots. The labeled times  $t_0$  to  $t_4$  correspond to those of Fig. 6.17

1. At the onset of the evolution (time  $t_1$ ) crustal thickening causes burial of rocks to great depths. In metamorphic rocks this is likely to be documented by an early high pressure metamorphic event.
2. Some tens of my later rocks reach their thermal metamorphic peak at mid-crustal depths. In Fig. 6.17 it may be seen that, depending on crustal level of the rocks, this occurs at times  $t_3$  or  $t_4$ . In metamorphic rocks this stage may be documented by the metamorphic peak paragenesis.
3. Following the thermal peak,  $T$ - $z$ -paths are first characterized by isothermal exhumation and later by cooling and finally exhumation to the surface (s. a. sect. 7.3).

This evolution is shifted in time for different crustal levels (s. sect. 7.4.1).

In Fig. 6.17 cooling of the upper crust commences at time  $t_3$ , while the lower crust heats at least until time  $t_4$ . Overall, there is a positive correlation between metamorphic grade and the time of metamorphism: The higher the metamorphic grade of a rock, the later its peak metamorphism occurred. For contact metamorphic rocks this relationship is exactly the opposite (sect. 3.6.2). We can conclude that the relationship between metamorphic grade and time of peak metamorphism is an important tool for the interpretation of heat sources of metamorphism (s. sect. 7.4.2).

The space, time and grade relationships predicted by the model discussed above (and shown in Fig. 6.17) are documented in the eastern Alps. There, Cretaceous high pressure parageneses are overprinted by mid Tertiary amphibolite facies parageneses. This observation was made by Oxburgh and his

students in the early seventies of last century and formed the basis of the model discussed here (e. g. Oxburgh and Turcotte 1974).

- *An example of some calculated  $T$ - $z$ -paths.* Fig. 6.19a,b shows depth-time and temperature-depth paths calculated with the model discussed above. We can see that these paths pass through conditions typical of Barrovian metamorphism (s. Fig. 7.1). Metamorphic peak temperatures between 600 °C and 800 °C are measured at metamorphic peak depths of 20–40 km in many regional metamorphic terrains. However, the temperatures that the model predicts for the lower crust are unrealistically high. This is indeed a major problem of the “England and Richardson” and “England and Thompson” models. It illustrates that – although the model is fantastic to teach us the fundamental relationships governing the thermal evolution of convergent orogens – it required refinement for more realistic modeling.

**More Refined Thermal Modeling.** The model discussed above has two important limitations that can be improved upon, even if we want to retain a one-dimensional model geometry:

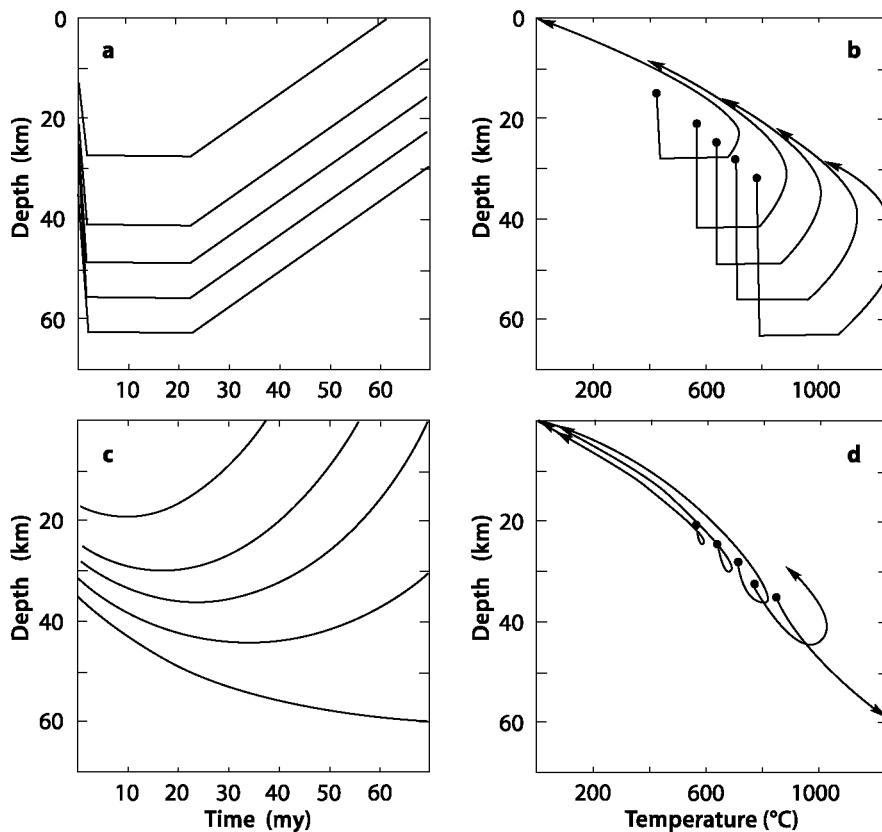
- The model does not allow for simultaneous deformation and metamorphism as observed in many orogens.
- By assuming a constant mantle heat flow at the Moho, the model does not allow for any processes in the mantle part of the lithosphere.

The influence of both is discussed here.

- *Simultaneous deformation and metamorphism.* Fig. 6.19c,d show  $T$ - $t$ - and  $T$ - $z$ -paths calculated assuming that deformation occurs simultaneously with heating, in particular with the kinematic assumptions discussed in sect. 4.3.4. We can see that there are important differences between the  $T$ - $z$ -paths shown in b and d:

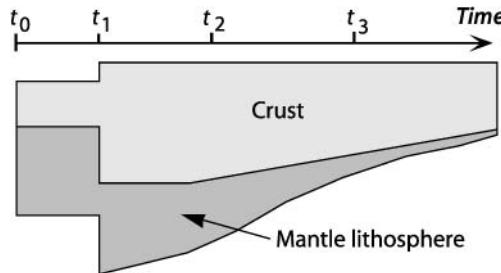
- If deformation, exhumation and thermal development occur on similar time scales, then metamorphic rocks reach their maximum depth and peak metamorphic temperature roughly at the same time (Fig. 6.19d). In contrast, if deformation of the orogen is short compared to the thermal development, then metamorphic depth and temperature peak are separated in time (Fig. 6.19b).
- The pro- and retrograde sections of the  $T$ - $z$ -paths are distinctly different if deformation precedes thermal development (Fig. 6.19b), but are much more similar if deformation and metamorphism occur simultaneously (Fig. 6.19d).

These differences may form useful criteria for the tectonic interpretation of a metamorphic terrain for which the  $T$ - $z$ - and  $T$ - $t$ -path is roughly known, but any other geodynamic information is unknown.



**Figure 6.19.** Kinematic and thermal evolution of rocks in collisional orogens: **a** and **b** show depth-time ( $z$ - $t$ ) and depth-temperature ( $z$ - $T$ ) curves using the kinematic assumptions of England and Thompson (1984); **c** and **d** show corresponding paths assuming deformation, exhumation and thermal evolution occur simultaneously (as described in terms of a simple model in sect. 4.3.4 (eq. 4.19) (Stüwe and Barr 1998)

- *The influence of the mantle lithosphere.* Within the “England and Richardson” the mantle heat flow through the Moho is assumed to remain constant during orogenesis. Considering the thermal definition of the mantle part of the lithosphere, this implicitly defines the thickness of the mantle part of the lithosphere (Fig. 6.20; s. also Fig. 3.20). Within the model heating of the crust causes thinning of the mantle part of the lithosphere. This in turn leads ultimately to unrealistically high temperatures in the lower crust (Fig. 6.19) and has some quite peculiar implications for the evolution of surface elevation and the dynamic state of orogens as a whole (Fig. 3.20 and Fig. 4.18). For example, England (1987) and Ridley (1989) showed that orogens with geotherms as predicted by the England and Thompson (1984) model, would



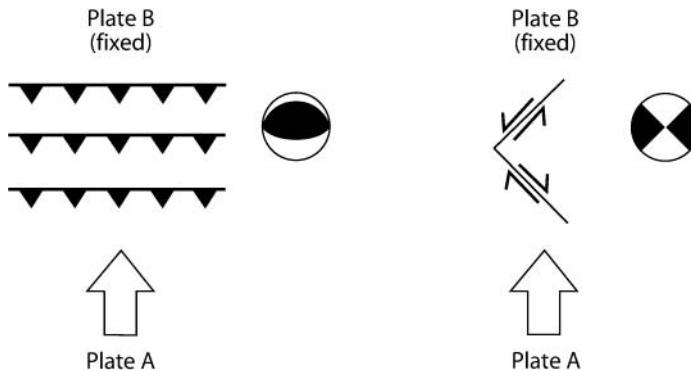
**Figure 6.20.** The thickness evolution of the lithosphere according to the models of England and Richardson (1977) and England and Thompson (1984). The thickness evolution of the *crust* is explicitly prescribed by the model. The thickness evolution of the *mantle part* of the lithosphere is determined implicitly by the lower boundary condition of constant heat flow at the Moho. The labeled times  $t_0$  to  $t_3$  correspond to the time steps labeled in Fig. 6.17 and 6.18

be mechanically extremely unstable (s. sect. 6.2.2). Interestingly, the implied evolution of the thickness of crust and mantle part of the lithosphere has a lot of similarities with an orogenic evolution characterized by early homogeneous thickening of the whole lithosphere and later delamination of the mantle part of the lithosphere (s. also sect. 6.3.2). Whatever the case may be, we may conclude that more refined models for the thermal evolution of the crust must consider the thickness of the mantle part of the lithosphere, even in one dimensional models.

### 6.2.2 Mechanical Description of Colliding Continents

So far we have considered the collision of continents only in terms of their *thermal* (sect. 6.2.1) and *kinematic* (sect. 4.3.4) evolution. If we want to consider their *mechanical* evolution, we need to consider the force balance equations we discussed on page 218 and probably use a constitutive relationship to relate forces to motions (or stresses to strain rates) (sect. 5.3.1). In *thermo mechanical* descriptions, thermal processes and mechanical evolution are considered simultaneously. If the mechanical problem we are interested in is a three-dimensional one, then all three equations of mechanical equilibrium must be integrated. Because this is often very difficult and even more difficult to interpret, most models try to reduce the explain the mechanics of orogens in two dimensional plan view or in cross section. We dicuss such two dimensional mechanical models and their implicit assumptions later in this chapter. However, in the first part of this section we make it even simpler and make a series of one dimensional considerations about mechanical equilibria in orogens.

- *Boundary conditions to orogens.* Orogenes may be subject to either “stress” or “velocity” type boundary conditions that drive their deformation and it



**Figure 6.21.** Illustration of the fact that the *same* boundary conditions may be reflected in two completely different stress states and strain fields (after Jackson and McKenzie 1988; England and Jackson 1989); (for an interpretation of the fault plane solutions see p. 227). The processes in the left diagram may best be illustrated on a vertical section. The processes in the right hand diagram are best described with a plane strain model

is important to discriminate between the two (s. Fig. 5.24). The India-Asia collision is an example for an orogenesis subject to a “constant velocity” type boundary condition: The convergence rate between the two plates has been largely constant over the last 30 million years, even though the potential energy of the Himalayas has increased dramatically during this time and is now opposing the driving force of the Indian Plate.

- *Plan view or cross section?* It is also important to realize that the *same* boundary conditions may lead to very different strain distributions. Fig. 6.21 illustrates an example. In both cartoons of this figure plate A moves towards plate B. However, in the example on the left, the convergence is compensated by overthrusting. In the example on the right, the convergence leads to orogen parallel extension. Although the overall velocity across the boundaries is the same, the strains are different. Clearly, the two geometries are different, because the boundary conditions in the *vertical* direction are different.

**Force Balances in Orogens.** Proper force balance equations were discussed in other sections of this book (e.g. eq. 5.18). Here, we want to introduce a more intuitive way to formulate a force balance for orgens that we can use without too much algebra. For this we divide (very loosely and not very precisely) the forces that keep orgens in mechanical equilibrium into three groups:

1. *Driving forces:* Driving forces are forces applied from the outside to an orogen, for example ridge push or slab pull. In the following we abbreviate these forces with  $F_d$ . Some of these forces were already discussed in sect. 5.3.2.
2. *Internal forces:* These are the forces internal to the lithosphere which resist the driving forces and are limited by the inherent strength of the rocks in

the lithosphere. These are the forces discussed in detail in sect. 5.2.1. As we are often interested in the whole lithosphere we often use the vertically *integrated* strength of the lithosphere, which has the units of force/meter and was explained on p. 239. We represent this in the following with  $F_l$ .

3. *Potential energy*: Forces resulting from the potential energy difference of an orogen relative to its surroundings are also called *gravitational stresses* or: *horizontal buoyancy forces*. We denote those in the following with  $F_b$ .

This division is not completely sound, as many of the plate tectonic driving forces themselves are also caused by potential energy differences and many of the other forces are also coupled. However, it helps us to understand the balance of forces in orogens which we can write as:

$$F_l = F_d - F_b \quad . \quad (6.16)$$

Basically this equation states that the strength of the lithosphere balances the effective force applied to the orogen, with the “effective force” being the difference between the external driving force causing convergence and the buoyancy force causing extension. We will discuss this equation in some detail in a few pages. However, first we want to discuss the process of building up potential energy in an orogen in some more detail. Note also that all orogenic forces are usually not given in the units of force (N), but that they are discussed in terms of force *per* meter ( $\text{Nm}^{-1}$ ) and that the unit of “force per meter” is equivalent to the units of “potential energy per area” or the units of “stress  $\times$  distance”.

- *Building up potential energy*. In sect. 5.3.1 we showed that the potential energy of orogens grows with the square of the surface elevation *and* with the square of the thickness of the orogenic root. Thus, it takes significantly more energy to increase the surface elevation of a high mountain range by one meter than it takes to increase the elevation of a low range by the same amount (Molnar and Tapponier 1978). As a consequence, the height of a mountain range and the thickness of an orogenic root are limited, if the driving force is a constant. This limiting elevation is reached when the potential energy of the range per square meter area is exactly as large as the tectonic driving force per meter length of orogen. Then, a steady state equilibrium of the forces is reached.

In order to understand how this equilibrium is reached, consider Fig. 6.22a, which illustrates a very simple model orogen. The left of this diagram shows normal thick crust of the thickness  $z_c$  and the density  $\rho_c$ . On the right, this diagram shows an elevated mountain range in isostatic equilibrium of the elevation  $H$ . The diagram is equivalent to Fig. 5.27. The difference in potential energy between the two mountain range and the foreland per square meter of area is given by eq. 5.56 and 5.57. Let us also recall that  $\Delta E_p$  is a potential energy *per* area and has the units of  $\text{J m}^{-2}$  and may also be interpreted as the mean net horizontal force exerted by the mountain range onto the foreland per meter length of orogen.

By analogy, the potential energy *per meter length of orogen* may also be interpreted as the product of the potential energy per area times the width of the mountain range  $l$ . From eq. 5.57 we can derive directly that:

$$\Delta E_{p,m-1} = \rho_{cg} H l (H/2 + z_c + w/2) . \quad (6.17)$$

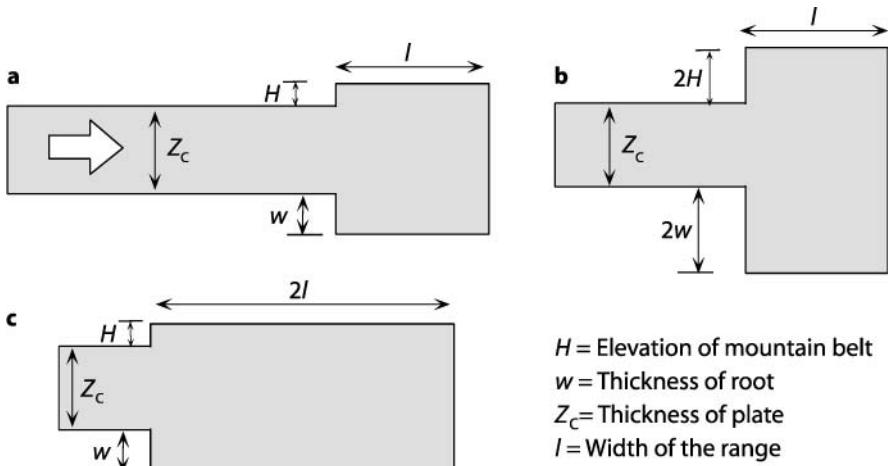
The subscripts are used to emphasize that we are dealing with the units of potential energy difference *per meter*, while the  $\Delta E_p$  that we used in eq. 5.56 and eq. 5.57 has the units of potential energy difference *per area*. Further growth of the mountain range may now proceed either in the vertical direction (Fig. 6.22b) *or* in the horizontal direction (Fig. 6.22c). If the crust inside the orogen is doubled in *thickness*, then the potential energy of the range per meter grows to the following value:

$$\Delta E_{p,m-1}^{\text{high}} = 2\rho_{cg} H l (H + z_c + w) . \quad (6.18)$$

If the growth of the mountain range is by doubling its width (at constant thickness, as shown in Fig. 6.22c), then the potential energy per meter growth to the following value:

$$\Delta E_{p,m-1}^{\text{wide}} = 2\rho_{cg} H l (H/2 + z_c + w/2) . \quad (6.19)$$

The difference of the potential energy increases between the two deformation styles is given by the difference between eq. 6.18 and eq. 6.19:



**Figure 6.22.** a Cartoon of a collisional orogen showing crust of normal thickness on the left and a mountain range on the right. Further displacement of the crust from left to right is compensated in b by further thickening and in c by lateral growth of the range. The difference in deformation style between b and c causes a significant difference of the potential energy of the mountain range (see eqs. 6.17 to 6.20) (s. also Fig. 5.27; after Molnar and Lyon-Caen 1988)

$$\Delta E_{p,m-1}^{\text{high}} - \Delta E_{p,m-1}^{\text{wide}} = \rho_c g H l (H + w) = \left( \frac{\rho_c \rho_m}{\rho_m - \rho_c} \right) g l H^2 . \quad (6.20)$$

The last simplification in the equation above was performed using the isostasy condition  $\Delta \rho w = H \rho_c$  that we also used in eq. 5.57. Eq. 6.20 shows us that it takes significantly less energy to thicken the crust in the foreland of a mountain belt (i.e. to widen the range) than it takes to increase the thickness of the crust in the mountain range itself (i.e. to increase the elevation of the range). Because of this, it is not necessary that convergence between two plates will stop when the gravitational extensional force  $F_b$  has reached the same magnitude as the tectonic driving force  $F_d$  acting towards the orogen. It is just that the convergence cannot be compensated anymore by *vertical* growth of the range, but will be compensated by *lateral* growth of the range towards the fore- or hinterland. Thus, active deformation in the range itself will come to a halt, the zone of active deformation propagates into the fore- and hinterland. A plateau will form in the center. In the process, the transition zone between the region where the largest principle stress is oriented horizontally and the region where it is oriented vertically will shift also towards the foreland.

Note that - despite these dramatic changes of the deformation and stress fields in the orogen - nothing has changed in the overall kinematics or stresses of the collision zone as a whole (Molnar and Lyon-Caen 1988) (s. p. 301): The driving forces have remained constant and so has the convergence between the two colliding plates. Understanding these relationships should therefore serve as a warning to structural geologists who are tempted to infer the overall kinematics of an ancient orogen from field observations on the kinematics of a few rocks.

- *Evolution of orogens in the equilibrium of forces.* The force balance we have discussed in the last paragraphs may be summarized in the following equation:

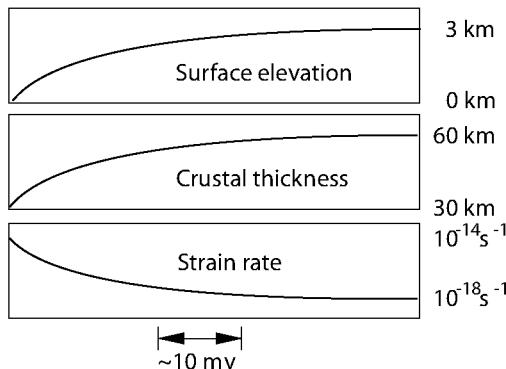
$$F_{\text{eff}} = F_d - F_b . \quad (6.21)$$

which we already introduced in eq. 6.16. There,  $F_d$  is the tectonic driving force *per* meter length of orogen,  $F_b$  is the gravitational stress *times* the thickness of the lithosphere.  $F_b$  is also called *horizontal buoyancy force*, or: *extensional force* or: *potential energy per area*. The difference between the driving force and the horizontal buoyancy force is the effective driving force applied to a continent  $F_{\text{eff}}$ . Equation 6.21 is often referred to as the “*ogenetic force balance*”. Note that – although this equation is called a “*force balance*” – it *really* balances parameters that have the units of force per meter or stress  $\times$  meter. Eq. 6.21 is often also written as:

$$F_{\text{eff}} = F_d - F_b = F_l . \quad (6.22)$$

There,  $F_l$  is the vertically integrated strength of the lithosphere in  $\text{Nm}^{-1}$  and corresponds to the area under the failure envelope discussed in Figs. 5.13,

**Figure 6.23.** Schematic illustration of the evolution of a collisional orogen subject to the force balance of eq. 6.22. Surface elevation and crustal thickness converge to a steady state when the magnitude of the horizontal buoyancy force approaches the tectonic driving force. Then, the convergent strain rate goes towards zero



5.14 and several others. Note that  $F_l$  can only equal the left hand side of the equation if the orogen is deforming (i.e. at the point of failure). When  $F_{\text{eff}} < F_l$ , there is no deformation. However, we assume that active orogens are always on the point of failure so that  $F_{\text{eff}} = F_l$  (sect. 5.2.1, eq. 5.48). The bulk of the lithosphere is dominated by viscous deformation mechanisms where deviatoric stress and strain rate are proportional. Thus, an orogen will always deform with a strain rate that is just large enough so that the vertically integrated flow stresses balance exactly the effective driving force (per meter). If the strain rate would be *lower* than this, the integrated strength of the lithosphere would be smaller than the effective driving force (per meter) and the deformation rate would increase. Conversely, if the strain rate would be *larger* than the effective driving force, then the strength would be too large for any deformation to occur. Note also that, within eq. 6.22, the integrated failure strength of the lithosphere is zero when the effective driving force is zero.

Because of the balance described by eq. 6.21 it is possible to solve this equation for strain rate of an orogen, if a relationship is assumed that relates stress to strain rate (e.g. a viscous flow law). Such an analysis has been done by a number of authors and provides insights into the basic principles of the mechanical evolution of collisional orogens (e.g. Sonder and England 1986; Sandiford et al. 1991; Stüwe et al. 1993a). If the tectonic driving force is assumed to be constant, then such orogenic evolutions track towards an equilibrium where  $F_b = F_d$  and  $F_{\text{eff}} = F_l = 0$  (Fig. 6.23). Thus, collisional orogens are self limiting. As such, collisional orogens are fundamentally different from extensional orogens, which are not necessarily self limiting (sect. 6.1.4).

The limiting values for a range of geodynamical parameters (e.g. surface elevation, crustal thickness) depend not only in the driving force  $F_d$ , but also on the integrated strength of the lithosphere. This may be illustrated if we reformulate eq. 6.22 to:

$$F_d = F_l + F_b \quad . \quad (6.23)$$

Remember that  $F_b$  is a direct function of the surface elevation and the thickness of the lithosphere. Thus, if the strength ( $F_l$ ) of the lithosphere becomes smaller, for example because the lithosphere becomes warmer or the material changes due to metamorphism, then it is possible that the potential energy of the range per unit area ( $F_b$ ) increases, even at constant driving force - the mountain belt grows (Houseman and England 1986a). Because of this process, sudden changes of the thermal structure of the lithosphere may cause sudden deformation events without changes in the driving force.

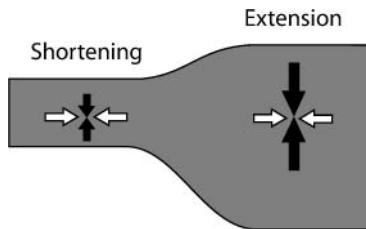
- *The mean strength of the lithosphere.* Differences in surface elevation of the continental lithosphere can only be created if the lithosphere has a finite strength (Fig. 2.2). That is: if the horizontal and vertical principle stresses are of different magnitude (Artyushkov 1973; McKenzie 1972; Molnar and Lyon-Caen 1988). If there were no stress differences, then the surface of a plate subjected to lateral forces from the outside would lift everywhere by the same amount; like water between two converging sides of an aquarium. There would be no mountain ranges and the surface of the continents would look rather boring. Conversely, it is possible to use the thickness and surface elevation of a mountain belt to estimate the mean strength of the lithosphere (Molnar and Lyon-Caen 1989).

Consider a mountain range which collapses under its own weight and to which there is no forces applied externally. We could then reformulate eq. 6.23 to:

$$F_b = -F_l \quad . \quad (6.24)$$

The left hand side of eq. 6.24 is the potential energy difference between mountains and foreland per unit area and was evaluated in eq. 5.57 or, somewhat more precisely, with eq. 5.59 (s. also Fig. 5.27). The right hand side of eq. 6.24 is the integrated strength of the lithosphere (s. eq. 5.48, Fig. 5.19 and Fig. 5.21). It is the product of the mean differential stress of the extending mountain range and its thickness. Thus, the elevation contrast between mountain belts and their foreland may directly be used to provide an upper bound on the mean strength of the lithosphere.

According to the estimates of Molnar and Lyon-Caen (1988), the surface elevation contrast between the Tibetan Plateau and the Indian foreland indicates a mean strength of the Asian lithosphere of  $\sigma_d = 69$  MPa. For the Altiplano in the Andes similar estimates indicate a mean strength of  $\sigma_d = 52$  MPa. This mean strength is estimated purely on the basis of topography differences and is therefore quite a sound estimate. If we acknowledge that some parts of the lithosphere will be significantly “softer” than this value (e.g. the uppermost and lowermost parts of the crust as shown in Fig. 5.13), then there *must* be other parts of the lithosphere that are significantly “stronger” than this value to maintain the mean value given by these estimates. These considerations provide a strong argument for the existence of a significant shear strength of parts of the lithosphere (s. sect. 6.3.5).



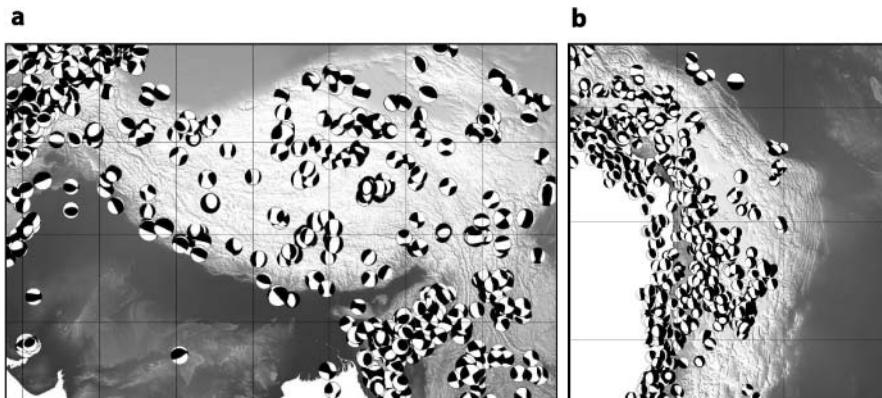
**Figure 6.24.** Distribution of horizontal and vertical stresses in a simple collisional orogen. If the topographic gradients at the surface and the base of the lithosphere are small, then the horizontal and vertical stresses  $\sigma_{xx}$  and  $\sigma_{zz}$  are parallel to the principal stresses. The horizontal stresses are constant across the orogen. However, the vertical stresses at a constant crustal level are higher in the orogen and smaller in the foreland. Thus, the largest principal stress in the foreland is given by  $\sigma_{xx}$ , while it is given by  $\sigma_{zz}$  in the orogen

**Mechanics on Vertical Sections.** Many continental orogens are long compared to their width. In such orogens many parameters do not change very much in the direction parallel to the orogen and it is often possible to neglect this direction altogether when describing the orogen: We can characterize them with a description on a vertical cross section and the equations that must be solved to describe this are eq. 5.15 and eq. 5.18, but omitting all terms that contain  $y$ . However, in this section we refrain from integrating these equations and simply expand on the discussion of the last section. More elegant two-dimensional models on cross sections are then discussed in sect. 6.2.3.

- *Changes in the stress field in collisional orogens.* In the discussion of eq. 6.20 we have shown that the stress field in an orogen may change over time, even if the far field plate boundary stresses remain constant. Here we illustrate this in some more detail by looking at the changes of the stress state across a mountain belt. In this discussion we follow the logic of Dalmayrac and Molnar (1981) as well as Molnar and Lyon-Caen (1988).

If the shear stresses at the base of the lithosphere are negligible, then the *horizontal* forces in a simple orogen (simplified as shown in Fig. 6.24) are constant, regardless of thickness of the plate or surface elevation (Artyushkov 1973; Dalmayrac and Molnar 1981) (s. Problem 5.7). In other words, the product of the mean horizontal stress  $\sigma_{xx}$  and the thickness of the plate remains a constant. Thus, if the stresses are a similar function of depth in different parts of the orogen, then the horizontal stress  $\sigma_{xx}$  is constant at any one depth across the orogen. This also implies that mountain ranges and plateaus transmit horizontal forces from the foreland to the hinterland of the orogen without changing their magnitude. On Fig. 6.24 this is indicated by the horizontal white arrows that are of the same size everywhere across the orogen.

This logic does *not* apply to the vertical stresses. Vertical stresses are the largest in regions where the overlying rock column is the thickest and the



**Figure 6.25.** Fault plane solutions for the two great plateaus on this planet. **a** The Tibetan Plateau as the consequence of the India – Asia collision and: **b** the Altiplano as the consequence of the collision between the Pacific and the South American plates. Note that the majority of the fault plane solutions at low elevation regions indicate compression, while those on top of the plateau indicate largely extension

smallest where it is the thinnest (s. Fig. 5.27). As a consequence, the stress distribution in an orogen may be like that shown in Fig. 6.24. In the foreland (on the left in this figure) the *vertical* stress is *smaller* than the horizontal stress. The region is thickening, for example by thrusting. In the mountain belt (strictly: in the region of high potential energy, s. sect. 5.3.3), the largest principle stress is the vertical stress. The region is extending. In short: although the horizontal stress on Fig. 6.24 is everywhere the same, there is thickening in parts of the Figure and extension in others. On earth, there are two orogens that have reached mechanical equilibrium and have formed plateaus. The Altiplano and the Tibetan Plateau. On both the transition from compression (in the foreland) to extension (on the plateau) can be observed (Fig. 6.25).

The lateral qualitative change in the deformation regime is *not* caused by changes in the horizontal- but changes in the vertical stress. This also explains why the observation of extension in mountainous regions must not occur because the surrounding plates are moving apart. The Tibetan Plateau is an example for such a situation: although the plateau is extending laterally, there is thrust tectonics in the surrounding regions (s. p. 303 and p. 306).

- *Changes in the stress field during the aging of plates.* If the driving force in an orogen remains constant, the qualitative nature of the orientation of the principal stresses may be changed during orogenesis because of: 1. Increasing the potential energy of the mountain range or 2. *decreasing* the mean potential energy of the entire plate (Coblentz and Sandiford 1994; Sandiford and Coblentz 1994). The former was discussed in the last section, the latter is the subject of this paragraph.

The geoid anomalies at continental margins tell us that the potential energy of old oceanic lithosphere is very low, while that of the mid oceanic ridges and that of large parts of the continental lithosphere is relatively high. Thus, the mean potential energy of a lithospheric plate is largely governed by the relative proportions of oceanic and continental lithosphere within this plate. If a plate is largely surrounded by mid ocean ridges (as both the African and the Antarctic Plate are) then the proportion of oceanic lithosphere in the plate will grow over time and the mean potential energy of the plate will sink. Correspondingly, the part of the *continental* lithosphere that has a higher potential energy than the mean value, will also grow over time. Thus, some areas that have long been under compression because of the high potential energy of the surrounding mid ocean ridges, may go into extension only because the plate is aging (the proportion of oceanic to continental lithosphere rises as the area of oceanic lithosphere rises). Sandiford and Coblenz (1994) have suggested that all continental plates will ultimately go into extension as a function of their age and suggested that features like the central African rift system may be largely caused by the aging of the African Plate.

**Mechanics in the Plane.** Many collisional orogens have features that may only be described by considering deformation in plan view, for example processes like lateral extrusion. In order to explain such observations (and avoid a full three dimensional description) two dimensional models in plan view are used. Because the vertical direction can then not be considered, vertically averaged assumption for lithospheric rheology and thermal structure have to be made. There are two common ways how to do this. These are the *plane strain* assumption and the *thin sheet* (plane stress) assumption (s. p. 10; Fig. 6.21). Which of the two assumptions is more appropriate for the description of orogens has been a subject of debate between the schools of Tapponier on the one side (e.g. Molnar and Tapponier 1975; 1978) and that of England, Houseman and McKenzie on the other side (e.g. England and McKenzie 1982; Houseman and England 1986a; England and Houseman 1986; 1988; Molnar and Lyon-Caen 1988). Within both the plane strain and the thin sheet model we need to consider the force balance equations:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} = 0 \quad \text{and :} \quad \frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} = 0 \quad . \quad (6.25)$$

which are the two dimensional versions of the force balance introduced on p. 218 for the  $x$  and  $y$  direction, respectively. Within the plane strain model these equations are solved subject to the appropriate boundary conditions. In contrast, the thin sheet model (e.g. England and Jackson 1989; England and McKenzie 1982; England et al. 1985; Houseman and England 1996) also considers a reduced vertical force balance with the body force term:

$$\frac{\partial \sigma_{zz}}{\partial z} = -\rho g \quad (6.26)$$

Note that – as long as there are no vertical strain rate gradients considered – such a body force term can simply be considered as a variably on a two dimensional model region described by eqs. 6.25 and the model is still two dimensional. Substituting these equations into eq. 5.8, the force balance may be written in terms of the depth averaged components of deviatoric stress and a horizontal buoyancy force arising from horizontal gradients in the vertical stress. For the  $x$  and  $y$  directions these equations are:

$$\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} - \frac{\partial \tau_{zz}}{\partial x} = -\frac{\partial \sigma_{zz}}{\partial x} \quad \text{and :} \quad \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{yy}}{\partial x} - \frac{\partial \tau_{zz}}{\partial y} = -\frac{\partial \sigma_{zz}}{\partial y} \quad .(6.27)$$

If the horizontal gradients of the vertical stress can be obtained independently, i.e. without solving a three dimensional set of equations, then eq. 6.27 remains a two dimensional description called the thin sheet approximation. It may be solved for evolving orogens if a flow law is assumed. Usually this flow law is assumed to be a non-linear viscous flow according to the relationship:

$$(\sigma_1 - \sigma_3) = B \dot{\epsilon}_l^{(1/n)} \quad \text{or :} \quad \dot{\epsilon}_l = B^{-n} (\sigma_1 - \sigma_3)^n \quad (6.28)$$

in which we can recognize a simplification of the viscous relationships explained in eq. 5.42 and eq. 5.45. The constant  $B$  summarizes all temperature dependent terms of the power law (eq. 5.46) and represents a depth averaged value when doing thin sheet calculations. A comparison of eq. 5.46 with eq. 6.28 shows that:  $B = A^{(-1/n)} e^{Q/nRT}$ . In short,  $B$  depends on strongly on temperature, but it can be shown that it is largely independent of the *distribution* of temperature within the lithosphere. Using the simplification of eq. 6.28 the lithosphere may be considered as a simple medium deforming according to a power law relationship between stress and strain rate (described by eq. 4.6). However, eq. 6.28 is not really a tensor equation (s. discussion of eq. 5.46) and it needs to be generalised to be coupled with the force balance above. Usually eq. 6.28 is generalised in the form:

$$\tau = B \dot{E}^{(1/n-1)} \dot{\epsilon} \quad (6.29)$$

where  $\dot{E}$  is the 2nd invariant of the strain rate tensor ( $\dot{E} = \sqrt{\dot{\epsilon}_{ij} \dot{\epsilon}_{ij}}$ ) and  $\tau$  and  $\dot{\epsilon}$  are the deviatoric stress and strain rate tensors as elsewhere in this book. Eq. 6.29 is a proper vector equation (like eq. 5.42) and the flow law can now be coupled with the force balance above. The non linearity between deviatoric stress and strain rate is taken care of by introducing the deformation dependent term  $\dot{E}$ . In this form, eq. 6.29 is the basis of many dynamic models for the description of continental deformation, for example those of England and McKenzie (1982) or Villette et al. (1982). In these models, the nature of deformation is often characterized by a single value: the Argand number.

- *The Argand number.* The Argand number  $Ar$  is a measure for the ease with which the lithosphere deforms in response to gravitational stresses. It tells us if an orogen is likely to flow apart at the same rate it is being built, or if significant amounts of potential energy may be stored within it before it would

Length scale	$L$	=	$U_0/\dot{\epsilon}_0$
Velocity scale	$U_0$	=	$L \times \dot{\epsilon}_0$
Strain rate scale	$\dot{\epsilon}_0$	=	$U_0/L$
Viscosity scale	$B$	=	$\tau_0/\dot{\epsilon}_0$
Stress scale	$\tau_0$	=	$B \times \dot{\epsilon}_0$
Argand number	$Ar$	$\propto$	$L/\tau_0$

**Table 6.2.** Scaling parameters in non dimensional viscous calculations. For simplicity, the stress exponent is assumed to be  $n = 1$ . Time scale  $t$  is not specifically listed here as it is simply  $t = \dot{\epsilon}_0^{-1}$ . In calculations where there is no Argand number, two of the top three variables and one of either viscosity- or stress scale are independent. Specifying  $Ar$  relates the top three to the next two variables, so that only two other scaling parameters must be chosen to define the system

collapse slowly under the influence of gravitational stress. It stems from the need to obtain horizontal gradients in the vertical stress when solving eq. 6.27. England and McKenzie (1982) showed that these gradients in vertical stress are proportional to the square of the crustal thickness  $S$ :

$$\frac{\partial \sigma_{zz}}{\partial x} \propto \frac{\partial S^2}{\partial x} \quad \text{or :} \quad \frac{\partial \sigma_{zz}}{\partial x} = \left( -\frac{g\rho_c}{2L} (1 - \rho_c/\rho_m) \right) \frac{\partial S^2}{\partial x} \quad (6.30)$$

for reasons discussed on p. 248. In eq. 6.30 the densities are those commonly used in this book and  $L$  is a thickness of the considered layer. If this is substituted into eq. 6.27 and the equation is brought into a non-dimensional form by normalizing to the length scale  $L$ , stress scale  $\tau_0$  and velocity scale  $U_0$  (the collision velocity), then the Argand number is:

$$Ar = \frac{\rho_c g L (1 - \rho_c/\rho_m)}{B(U_0/L)^{1/n}} = \frac{P_{(L)}}{\tau_0} . \quad (6.31)$$

and may be interpreted as dimensionless ratio of the additional pressure  $P_{(L)}$ , that arises because of the thickness difference  $L$  between two plates and the stress  $\tau_0$ , that is necessary to deform a plate with a significant rate  $\dot{\epsilon}_0 = U_0/L$  (England and McKenzie 1982).

In this form,  $Ar$  may be used as an input parameter for mechanical modelling of orogens without having to explicitly consider the rheology, the material constants or the temperature profile of the lithosphere (Table 6.2). The additional pressure rises linearly with the thickness of the orogen and the stress  $\tau_0$  increases with the effective viscosity of the plate.

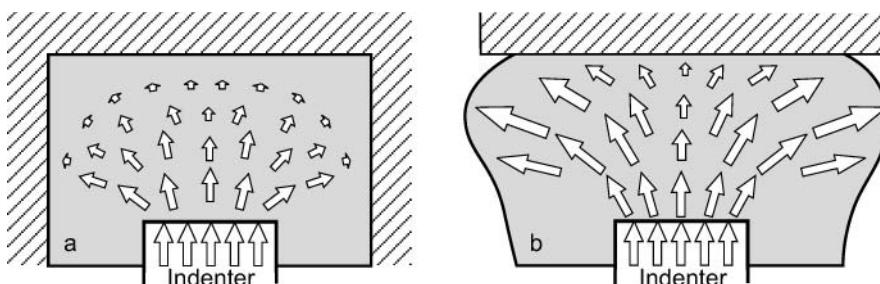
We can see that – if the effective viscosity of a plate is large, then the Argand number is small. Then, the flow properties of a mountain belt will depend largely on the orogenic boundary conditions. The belt will only begin to extend once its potential energy is very large. In contrast, if the Argand number is large (say between 10 and 20), then the effective viscosity of the range is very small and the forces caused by potential energy differences are

large. The crust will quickly flow in response to applied forces. No significant thickness variations between foreland and orogen will ever develop during orogenesis. England and McKenzie (1982) have shown that orogens characterized by an Argand number of 30 show practically no thickness variation and their deformation is already nearly plane strain.

- *The Deborah number.* A number similar to the Argand number is the Deborah number. In contrast to the Argand number (which is defined in terms of stresses), it is defined as the ratio of two time scales, namely the time scale of viscoelastic stress relaxation and the characteristic time scale of deformation (Reiner 1964; 1969). Akin to the Argand number it can be used as a measure of the fluid-like behavior of continents. When applied to the evolution of continents, the Deborah number may be interpreted as the duration for which an orogenic driving force is applied to a plate, relative to the duration an orogen takes to flow apart (England 1996). The Deborah number has also been employed on a much smaller scale to characterize outcrop scale folding (e.g. Mancktelow 1999; Schmalholz and Podladchikov 1999).

- *Orogen parallel displacement.* Collision of continents causes displacement of rocks in all three spatial directions. The *vertical* displacement results in thickening, the horizontal displacement in direction normal to the indentation direction is called loosely *lateral extrusion* and may occur in either a compressional or extensional regime.

Fig. 6.26a shows the collision of a plate with a rigid indenter that deforms the plate in front of it. The rocks in front of the indenter are displaced both in the direction of indentation and perpendicular to that direction. In Fig. 6.26a the amount of displacement decreases with distance from the indenter as the deformation there dissipates more. However, despite the orogen parallel *displacement* of rocks, all points of the indented plate are under *compression*. There is no lateral extension. This conclusion from Fig. 6.26a is in contrast to

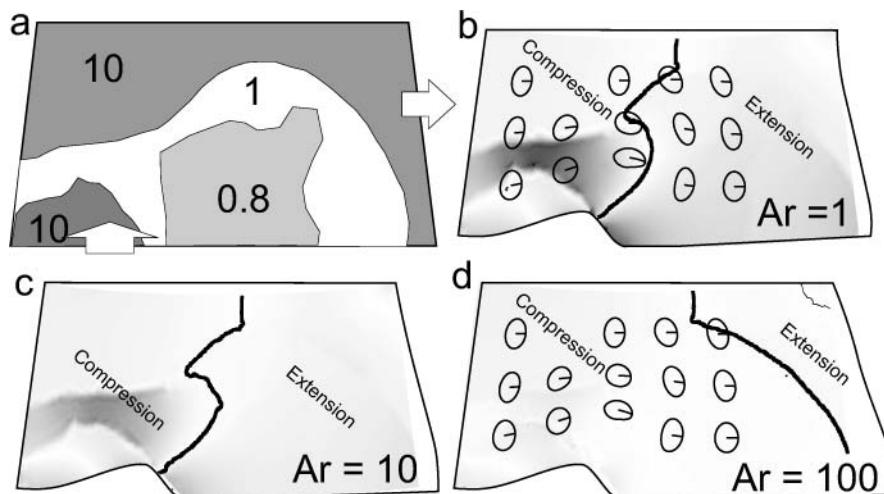


**Figure 6.26.** Different deformation regimes that occur during collision of an indenter with a much larger continental plate (gray shaded region). The arrows are velocity vectors. In **a** the plate is infinite or bound at all sides. In **b** the side boundaries of the plate are free and lateral extension occurs. The absence of confined model boundaries is one of four mechanisms that can account for orogen parallel extension during convergent plate motion

many observations in active collisional orogens where lateral extension *does* occur. According to England and Houseman (1989) orogen parallel extension in convergent orogens must find its nature in one of the following four processes:

1. unconstrained boundaries,
2. decrease in the convergence rate between two plates,
3. changes in the rheology of the plate,
4. external addition of potential energy to the plate.

The first of these four processes is illustrated in Fig. 6.26b. There – in contrast to Fig. 6.26a – the gray shaded region is not bound on the sides. The other three processes may be illustrated with an analysis of eq. 6.23. A decrease in the convergence rate is reflected in this equation by a decrease in  $F_d$ . If  $F_l$  remains constant, the horizontal buoyancy force must decrease and extension sets in. This process is generally known as “post orogenic collapse”. Changing the rheology of the plate (e.g. by heating, recrystallization, metamorphism etc.) is reflected in eq. 6.23 by changes in  $F_l$ . In order to maintain the force balance, strengthening of the plate must be accompanied by a decrease in the deformation rate or a decrease in the horizontal buoy-



**Figure 6.27.** Model for the lateral extrusion of the Eastern Alps illustrating the influence of Argand number  $Ar$  on the deformation regime during continental indentation (after Robl and Stüwe 2005). **a** The 4 considered regions and their relative viscosity contrasts: European foreland (10), Adriatic indenter (10), Pannonian Basin (0.8) and Eastern Alps (1). For  $Ar = 1$  and even for  $Ar = 10$ , significant topography (gray shading) is built in the Alps and lateral extrusion is exclusively due to tectonic forcing. For  $Ar = 100$ , lateral extrusion is increased (a little bit) by an additional contribution of extensional collapse. However – although barely any topography is supported – the extensional collapse is still in an overall compressive regime (because indentation goes on).

**Table 6.3.** Definition of terms related to lateral displacement of rocks during indenter tectonics on an orogenic scale

term (key reference)	meaning
lateral extrusion (Ratschbacher et al. 1991)	kinematic description for displacement of rocks normal to indentation direction (stress regime: undefined)
gravitational collapse (Dewey 1988)	extension under its own potential energy (stress regime: extensional)
tectonic escape (Tapponier et al. 1982)	active lateral forcing along strike slip faults (stress regime: compressional)

ancy force  $F_b$ . Extension occurs as a consequence. The external addition of potential energy, for example by delamination of the mantle part of the lithosphere, has a similar influence on eq. 6.23. It may also cause the transition from compression to extension.

- *Lateral extrusion.* Lateral *extrusion* of orogens is a great term that can be used when referring to lateral displacement of rocks without wanting to specify if its under a compressional or extensional regime. According to Ratschbacher et al. (1991), lateral extrusion describes lateral motion of rocks as the consequence of a combination of tectonic escape (in a compressional regime) and gravitational collapse (in an extensional regime). While the term lateral extrusion is mechanically not specified and is a largely a kinematic description, the terms *tectonic escape* and *gravitational collapse* have well defined mechanical implications that were described by Tapponier et al. (1982) (Jones 1997) and Dewey (1988), respectively. Both processes are observed in the eastern European Alps, where extensional tectonics is observed along the east and west margins of the Tauern window ) and simultaneous tectonic forcing along major strike slip lineaments has occurred (Selverstone 1988; Ratschbacher et al. 1991; Robl and Stüwe 2005).

### 6.2.3 Orogenic Wedges

All considerations of the last paragraphs are based on the thin sheet model assumption. That is, we have assumed thickening processes to be homogeneous. Using this simplification we were able to understand some important aspects of the mechanics of continental orogens, but we have limited ourselves to the understanding of symmetrical orogens and plateaus. However, most orogens have a fundamentally asymmetric geometry in cross section

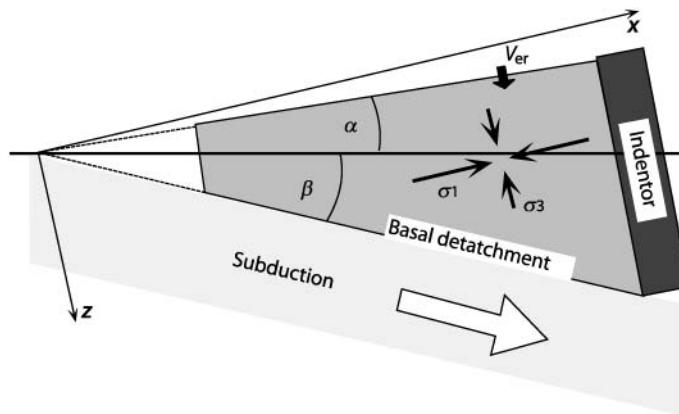
(e. g. Pfiffner et al. 1997). In fact, on a sub-orogen scale, wedge shaped packages of rock are very common in the foreland of orogens. According to the appropriate assumptions used to describe them, wedge shaped orogens may be very loosely divided into:

- Mohr-Coulomb wedges (fold and thrust belts, accretionary wedges)
- S-point wedges (used to describe orogens as a whole)

Both have in common that they have characteristic geometries that are a function of the fact that stresses are imposed at the base and that the upper surface is stress free. As such they differ significantly from another use of the term “orogenic wedge”: The term is also used for wedge shaped packages of rocks that are extruded in the horizontal or vertical direction between two fault zones, for example the Central Himalayan Crystalline complex (that extrudes between the Main Central Thrust and the South Tibetan Detachment), The Chugach Metamorphic complex of Southern Alaska or the Eastern European Alps. Extruding wedges of this type were discussed in the last section and we spend some time on them on page p. 323,

**Mohr-Coulomb Wedges.** Wedge shaped orogenic bodies that are characterized by a fixed angle between the surface of the earth and a detachment surface at the base of the plate form in many orogens both on land and under water. The mechanics and geometry of such wedges is well described with the Mohr-Coulomb criterion (Dahlen 1984; sect. 5.1.2) and we call them therefore here “Mohr-Coulomb wedges” although we also discuss thermal models in this section. Mohr-Coulomb wedges that form in connection with the subduction of oceanic lithosphere are usually below sea level and are called *accretionary wedges*. In continental orogens, such wedges form above sea level and are called *fold and thrust belts* (McClay 1992). One of the fundamental characteristics of both fold and thrust belts and accretionary wedges is their constant overall shape during growth and in different examples throughout the world. Most accretionary wedges have a surface slope which dips at about 1°. In fold and thrust belts, this angle is typically of the order of 3°. In the following section we use the term “wedge” or “orogenic wedge” for both accretionary wedges and fold and thrust belts.

The wedge shape of all orogenic wedges stems from the fact that an inclined plane moves material towards a fixed back stop. This inclined plane is usually referred to as the basal detachment and the transmission of forces into the wedge is by friction along this detachment (s. sect. 6.2.2). In accretionary wedges marine sediments that lie on top of the subducting oceanic lithosphere are moved towards the upper plate (s. Fig. 2.23; 5.23). The upper plate is the back stop or indenter, depending on whether the process is seen in a Lagrangian or Eulerian reference frame (Fig. 6.29). A good example of such a wedge is the accretionary wedge that forms between the Pacific Plate and the North American continent in Alaska. The best known example of a fold and thrust belt is Taiwan, which formed as a consequence of the subduction of



**Figure 6.28.** Geometrical parameters of accretionary wedges and fold and thrust belts. The angles  $\alpha$  and  $\beta$  are the inclination angle of the surface and that of the basal detachment both measured relative to the horizontal. In models that are based on Mohr-Coulomb criteria, the coordinate system is often chosen so that the axes and the indenter are parallel to the direction of the principle stresses as shown here

the Eurasian continental margin underneath the island arc of Luzon (Suppe 1981; 1987). Using Taiwan as an example, a number of models have been developed in the past 20 years that may be used to explain orogenic wedges around the world (Davis et al. 1983; Dahlen et al. 1984; Dahlen 1984; Barr and Dahlen 1989; Dahlen and Barr 1989; Platt 1990; 1993a).

Modeling orogenic wedges is a typical two-dimensional problem. The most important parameters necessary for the geometrical description of the wedge are the angle of inclination of the basal detachment  $\beta$  and that of the surface  $\alpha$  (both relative to the horizontal Fig. 6.28). Models describing accretionary wedges may be divided into:

1. Models, describing the geometry and the state of stress.
2. Models, describing the kinematics and thermal development.

Both types of models will be discussed briefly in the next sections. In the context of doing so we will show that indenters are not necessary for the formation of an orogenic wedge.

- 1. *Geometry and State of Stress.* The origin of the typical shape of orogenic wedges may be understood by comparing orogenic wedges with the familiar analog of a snow plow. If a snow plow starts moving through a fresh layer of snow, we know that there are two possibilities of how the snow deforms: 1. If the internal strength of the snow is higher than the friction with the road (e.g. if the snow is icy and the road is warm), then the snow will be moved as an undeformed slab in front of the plow. 2. If the internal strength is smaller than the friction with the road (as usually is the case) then the snow deforms internally and its surface will become inclined. According to the Mohr-Coulomb criterion, the strength of the snow will increase as the

thickness of the snow wedge increases (and the normal stresses within it do, s. eq. 5.29). This will continue until a critical taper between surface and basal detachment (in our example: the road) is reached where the strength of the snow is exactly as large as the basal traction (on Fig. 6.28 this is the angle  $\alpha + \beta$ ) (Davis et al. 1983). In this state, the wedge can now move along the basal detachment and during progressive deformation the wedge stays constant in shape. However, while constant in *shape*, it still may grow in size as more snow is scraped up always retaining a balance between (increasing) internal strength and (increasing) basal friction. The wedge has a self similar geometry.

Dahlen (1984) showed that many orogenic wedges may be described very well with the assumption that they consist of cohesion-free material deforming according to Mohr-Coulomb criterion (sect. 5.1.2). In this case the orientation of the principal stresses are constant everywhere within the wedge (Fig. 6.28). Summarizing this we can write:

$$\alpha + \beta = \text{constant} \quad . \quad (6.32)$$

The constant in this equation depends on two parameters: the strength of the wedge material and the strength of the basal detachment. Both parameters are functions of the coefficient of internal friction  $\mu$  and the fluid pressure (eq. 5.33). High stresses on the basal detachment *increase* the critical taper, high internal strength decreases it. Similarly, high fluid pressure inside the wedge decreases the strength of the wedge and increases the critical taper, while high fluid pressure along the basal detachment decreases the friction along this surface and thus decreases the critical taper.

The model discussed in the last paragraph was a great advance in our understanding of orogenic wedges (Dahlen 1984). Before the consideration of Mohr-Coulomb rheologies, wedges were generally described with models that do not consider the depth dependence of stresses (Chapple 1978; Stockmal 1983).

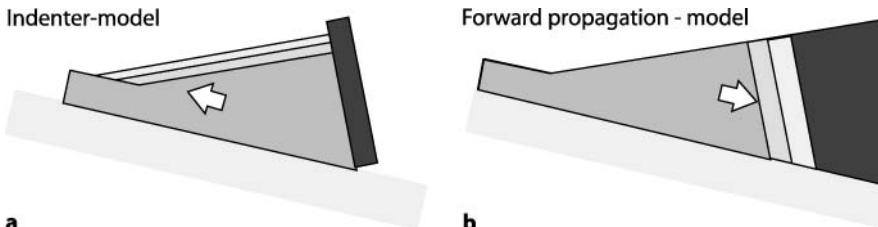
- *Differences between limited and unlimited wedge size.* In the last paragraph we have shown that an orogenic wedge will grow infinitely while maintaining a self similar shape. Depending on the position of the observer this growth may be interpreted in two different ways. An observer on the indenter (e.g. a snow plow driver) will see that the wedge grows at the toe and increases in thickness. This is illustrated in Figs. 6.29a and Fig. 6.30a and this situation corresponds to most accretionary wedges. However, if seen from the point of view of the wedge toe on the subducting plate, the wedge grows at its wide end and the region of most intense deformation moves progressively towards the foreland (Fig. 6.29b and Fig. 6.30b). This is observed in Fold-and-thrust belts where the deforming area progressively propagates forward into the orogen. There are two possibilities to limit not only the *shape* but also the *size* of a wedge:

- In the snow plow model, the size of the wedge is limited by the height of the plow. During progressive deformation, the wedge will eventually become as

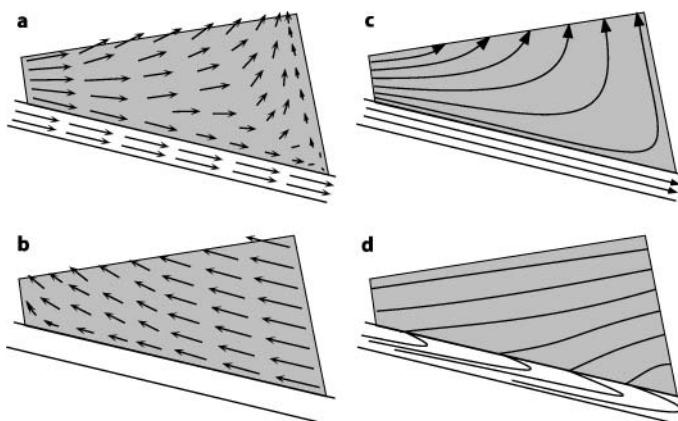
high as the plow itself and the snow will be removed from the system over the top of the plow blade. Koons (1990) showed that exactly this process may be relevant to the geometry of deformation in New Zealand.

- The surface of the wedge may erode rapidly enough so that the erosive removal of material from the system balances the input of material at the wedge toe.

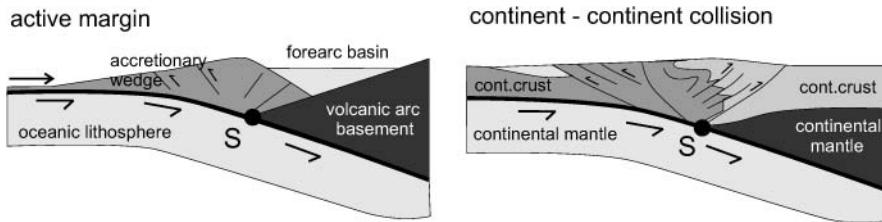
The second possibility appears to be realized in many accretionary wedges around the world. It forms the basis of the kinematic models we discuss in the next section.



**Figure 6.29.** Cartoon illustrating the growth of accretionary wedges and fold and thrust belts. The arrows and light shaded wedge sections indicate the growth direction. The black regions show undeformed parts of the foreland. **a** From the point of view of the indenter, the wedge grows at its toe. **b** From the point of view of the basal detachment, the wedge growth appears to be caused by forward propagation of the deformation into the foreland. The indenter model is more appropriate to accretionary wedges. The forward propagation model more appropriate to fold-and-thrust belts



**Figure 6.30.** Schematic illustration of the kinematic and thermal structure in accretionary wedges (after Dahlen and Barr 1989; Barr and Dahlen 1989). The velocity vectors penetrate the surface because it is assumed that there is a steady state equilibrium between material input at the wedge toe (left edges) and material output by erosion at the surface. **a** Velocity field relative to the indenter (upper plate); **b** Velocity field relative to the subducting plate; **c** Rock trajectories in the accretionary wedge; **d** Isotherms



**Figure 6.31.** Two different plate margin settings showing that they have similar geometries. In both, point **S** is a discontinuity in the basal boundary condition

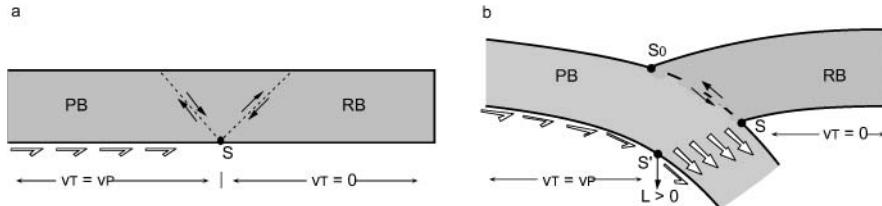
- *2. Kinematics and Thermal Structure of Orogenic Wedges.* If there is no erosion and the height of the indenter is infinitely high, then wedges would grow infinitely. The velocity field of rocks in the wedge relative to the indenter would look like a radial vector field starting at the toe of the wedge (coordinate origin in Fig. 6.28). There is no exhumation. However, if the wedge formation is accompanied by erosion at the surface, then the rock trajectories follow curved paths. If the total volume of the eroded material is exactly as large as the amount of material that is put into the wedge at its toe by the subducting plate, then the wedge does not grow and the rock trajectories are in a steady state (Fig. 6.30c). For such wedges with a constant shape and size, there is a range of models that describe their internal kinematics.

The most simple models are based on methods used in fluid dynamics for the description of the flow behavior of fluids in corner (*corner flow models*; Cowan and Silling 1978; Cloos 1984; Cloos and Shreve 1988). In further developments, Barr and Dahlen (1989) as well as Dahlen and Barr (1989) succeeded in developing analytical models that may be used to describe the internal kinematics of wedges that are subject to the Mohr-Coulomb-criterion. For this, they chose a coordinate system that is parallel to the principal stresses in the wedge (Fig. 6.28).

The thermal structure of wedges has also been described in models designed by Royden (1993b); Platt (1993a) as well as Barr and Dahlen (1989); (Fig. 6.30d). Similar models have been designed by Bird and Piper (1980), Beaumont et al. (1992) and Willett (1992).

**S-point Wedges.** “S-point wedges” bear their name because Willett et al. (1993) and Beaumont et al. (1996) recognized that a series of plate margin settings can be described by a simple discontinuity at the basal boundary and that a back stop – like in the bulldozer model discussed on the last pages – is not explicitly required to describe wedge shaped orogens or parts thereof (Fig. 6.31). The model of S-point wedges is also more appropriate to doubly vergent orogens and in fact, effective back stops develop dynamically during model evolutions. As such, the S-point wedge model is more simple, more elegant and it now largely supersedes the geometry described for Mohr Coulomb wedges in the last section in most model studies describing orogenic wedges.

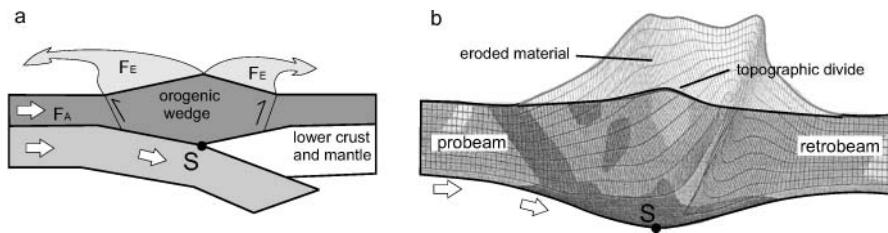
The S-point wedge model of Willett et al. (1993) is described by a rectangle that is subjected to zero stress boundary conditions at the sides and at the top (Fig. 6.32a). The lower boundary is subjected to kinematic boundary conditions with a discontinuity at point  $S$  in the center. To the left of point  $S$  there is a constant tangential velocity imposed at the base ( $v_T = v_P$ ). To the right of point  $S$  the velocity is zero. The lower boundary conditions is assumed to represent the Moho and implies that the mantle lithosphere of the plate  $PB$  (ProBeam) subducts at point  $S$  towards the right underneath plate  $RB$  (RetroBeam).



**Figure 6.32.** Boundary conditions of the two-dimensional model for S-point wedges. **a** The original model of Willett et al. (1993). The white arrows show the velocity boundary condition along the lower margin of the plate. On the other three boundaries it is assumed that both normal and shear stresses are zero. The dashed lines with the shear sense indicators are not part of the model assumptions. They illustrate regions of high shear as they develop dynamically during model simulations with these boundary conditions (Fig. 6.33b). **b** The general subduction-collision model of Beaumont et al. (1996). The geometry of the initial condition shown here is the consequence of the magnitude of the downwards acting load  $L$  applied at point  $S'$ .  $PB$  and  $RB$  are the lower and upper plate ( $PR$  for pro beam and  $RB$  for retro beam). The shear zone between  $S$  and  $S_0$  is no model assumption but develops dynamically during model runs with these boundary conditions. **b** If the load  $L$  is assumed to be zero, the model recovers the geometry shown in **a**

Beaumont et al. (1996) expanded this model to describe subduction of the entire lower plate within the same model (Fig. 6.32b). For this modification they assumed a vertical load  $L$  (simulating a downwards pulling plate) to apply at point  $S'$ . To the left of point  $S'$  and to the right of point  $S$ , the same boundary conditions as in Fig. 6.32a apply. Between  $S'$  and  $S$  the velocities are assumed to be the same as at point  $S'$ . These boundary conditions imply that the entire crust is subducted between these two points. As the load  $L$  in Fig. 6.32a becomes zero, the model becomes that of Fig. 6.32b.

Model runs performed with these boundary conditions reproduce a series of structures commonly observed in collisional orogens, in particular the conjugate shear zones that are schematically indicated in Fig. 6.32. It is important to note that these shear zones are not predefined, but develop dynamically during the model runs. The crustal scale shear zones give the orogen quite a symmetric appearance, despite the fundamental asymmetry of the basal boundary conditions. This apparent symmetry warns us to be careful with the interpretation of the direction of subduction of plates in an orogen, if this



**Figure 6.33.** Results of model evolutions using the boundary condition geometry shown in Fig. 6.32. **a** Cartoon of a typical model geometry Convergence between probeam and retrobeam has produced a doubly vergent orogen.  $F_E$  and  $F_A$  are the mass outflux and influx showing that there may be a balance between erosion and convergence in steady state orogens (Willett and Brandon 2002). **b** Lagrangian tracking grid of a typical model run used to study morphological evolutions by considering erosion explicitly (Willett 1999). Gray shading is strain rate. Note that two conjugate directions concentrate the strain

is solely interpreted on the basis of observations of structural trends at the surface

The model shown in Fig. 6.32 has been used extensively in the past years (e.g. Fig. 6.33, for example to illustrate the effects of erosion on the dynamics of orogens (Willett 1999), to explore the balance between mass in flux by collision and mass outflux by erosion (Willett and Brandon 2002), and even effects of asymmetric precipitation or climate change (Willett 1999; Willett et al. 2006).

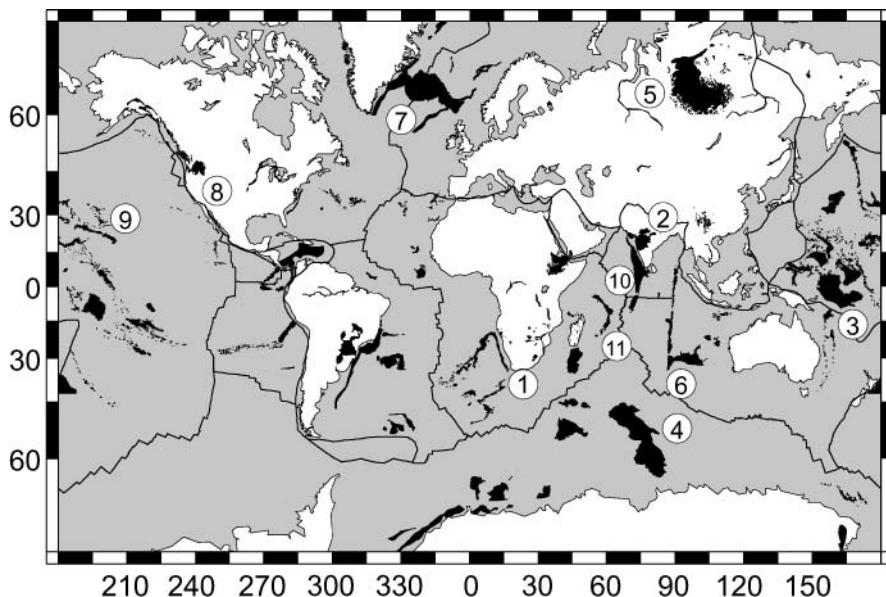
## 6.3 Selected Geodynamic Processes

In this section we discuss, very superficially, a random selection of some important and currently actively debated geodynamic problems. There is no direct connection between the individual sections.

### 6.3.1 Flood Basalts and Mantle Plumes

Around the globe there are a large number of regions where enormous quantities of basalts have erupted (Fig. 6.34). These regions are known by the name of “*large igneous provinces*” or just “LIP”. Aside from the rocks formed at the mid oceanic ridges, they are the second largest accumulations of mostly mafic igneous rocks on earth. LIPs are a typical feature of Phanerozoic geology and are largely unknown from the Proterozoic or Archaean. LIPs include three types of basaltic provinces:

- continental flood basalts,
- rifted continental margin volcanic sequences,
- oceanic plateaus.

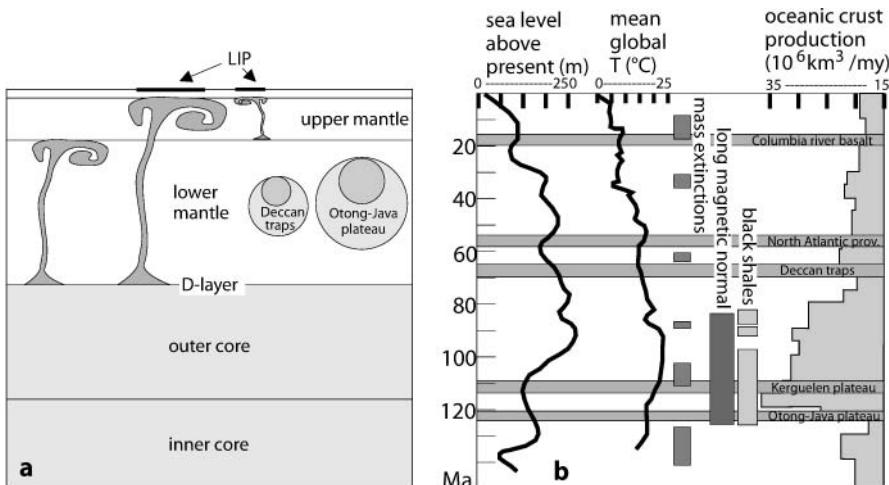


**Figure 6.34.** Large igneous provinces on earth. The most important labeled provinces are: 1=Karoo Basalts; 2=Deccan Traps; 3=Otong-Java Plateau (largest province with  $2 \times 10^6 \text{ km}^2$ ); 4=Kerguelen Plateau; 5=Siberian Traps; 6=Broken Ridge Plateau; 7=Iceland; 8=Columbia River basalts; 9=Hawaii; 10=Chagos Kac-cadive Ridge; 11=Maskarene Plateau. Data are from Coffin and Eldholm (1993a,b)

The formation of LIPs has been discussed by a large number of authors (see recent summary edited by: Mahoney and Coffin 1997). In general, it is accepted that most of such provinces are caused by plume activity in the mantle. Initially mantle plumes were considered to be part of the general convection system of earth (Morgan 1971). However, within our modern understanding of plumes, they are known to be secondary features unrelated to the plate scale convection in the mantle (Sleep 1992). Today, authors discern between two types of plumes:

- Plumes that initiate from the core - mantle boundary that have narrow stems and large heads (Fig. 6.35). These plumes will be associated with “active” rifting that follows active uplift of the overlying lithosphere by the plume (s. p. 253).
- Plumes that form due to adiabatic upwelling of the asthenosphere in response to passive lithospheric extension (White and McKenzie 1989). Such plumes are themselves the consequence of the rifting but there may be positive feedbacks that cause an accelerating extension process.

(For a discussion of active versus passive rifting see also p. 265 and 253). The island chain of Hawaii was the first place where mantle plume activity was



**Figure 6.35.** **a** Scaled cross section through earth showing the size and geometry of mantle plumes. The circles for the Otong-Java Plateau and the Deccan Traps show minimum and maximum estimates for the volumes of melt extracted from the mantle during their activity (after Coffin and Eldholm 1994). See also Fig. 4.22. **b** The occurrence of the last 5 large igneous provinces in the last 150 my in relationship to other episodic events during this time. Shown are the time of the long magnetic normal in the Cretaceous, some important mass extinctions, important black shale deposition events as well as sea level, global temperature and oceanic crust production curves (after Larson 1991, as well as Coffin and Eldholm 1994)

suggested to be responsible for the appearance of the volcanic chain (Wilson 1963), although Hawaii is now known to be one of the smaller basaltic provinces on the planet (Fig. 6.34). Some of the largest LIPs form in places where rifted continental margins coincide with plumes (see also discussion of Fig. 5.28 and White and McKenzie 1989). There, LIPs may occur on the surface or below sea level, depending on the relationship between rifting rate and magma supply (Saunders et al. 1996). If the spreading rate is low in comparison to the rate of magma supply, then the plateaus become subaerial as in Iceland. If the spreading rate is low the igneous province remain submarine.

**Geometry.** In the nineties people have begun to explore the geometry and volumes of LIPs mainly using the facilities of the international ocean drilling program ODP. For example, it is now known that plumes account at present for about 5-10% of the mass and energy flux from the mantle to the crust and that this value may have been much larger in the past (Coffin and Eldholm 1994). Thicknesses of LIPs are between 20 and 40 km and appear to have formed in relatively short-lived pulses of increased global production. Many oceanic plateaus have refractory depleted keels that are more buoyant than normal oceanic crust. As a consequence, oceanic plateaus may be preserved much longer in the plate tectonic cycle than normal oceanic crust. When they

collide with active margins, they choke the subduction zone and may cause back-stepping of the zone.

Thicknesses and aerial extent are well enough known to estimate magma production rates. For example, it is now known that the magma production rate for the formation of the Otong-Java Plateau may have exceeded the entire production rate of mid ocean ridge around the world at present (Coffin and Eldholm 1994).

**Plumes or no Plumes.** While mantle plumes have been generally accepted as the cause for the formation of most LIPs, it is worth mentioning that they may not be the only cause (Coffin 1997). Several authors have discussed alternative mechanisms for the formation of large igneous provinces (King 1996; Seth 1999). Three alternative ideas have been discussed:

- Linear chains of volcanoes may indicate propagating rifts where the volcanic chain delineates the stress field, rather than the displacement field. This idea has been discussed by Turcotte and Oxburgh (1973) and Jackson and Shaw (1975) and the origin of the Deccan Traps as well as the Laccadives-Reunion hot spot has been suggested as an example by Seth (1999). Within this model volcanic chains are piezometers rather than speedometers.
- LIPs may be leaky transform faults (Smoot 1997).
- Linear chains of volcanoes are produced by magma surge channels (Meyerhoff 1995).

However, it should be said that none of these models have found wide acceptance and plumes are still the most widely accepted model for the formation of LIPs.

**Large Igneous Provinces and Mass Extinctions.** Large igneous provinces have formed episodically during the entire Phanerozoic. This episodicity has been brought in connection with a range of other episodically occurring events (Fig. 6.35b). For example, super plume activity and oceanic crust production have been brought in connection with cessation of the magnetic field reversal in the mid Cretaceous and in the Permian (Larson 1991) (s. also Yale and Carpenter 1998). Correlations with mass extinction events have been attempted by a number of authors. However, in a recent summary Wignall (2001) considers most of these correlations as unduly optimistic. On the other hand, Wignall (2001) does recognize correlations with some mass extinction events, namely with the Karoo basalts, the Siberian traps, the central Atlantic volcanism and the Emeishan flood basalts.

### 6.3.2 Delamination of the Mantle Lithosphere

The mantle part of the lithosphere is colder and thus may be denser than the underlying asthenosphere (Fig. 2.18, eq. 4.30). Thus, the mantle part of

the lithosphere can have a *negative* buoyancy (in contrast to the crust) and it is conceivable that it falls down into the asthenosphere. If this happens, the overlying orogen experiences dramatic changes to its potential energy, surface elevation and thermal structure. There are several mechanisms why and how in detail such sinking of the mantle lithosphere into the asthenosphere may occur. In particular, there are two models regarding how this may happen:

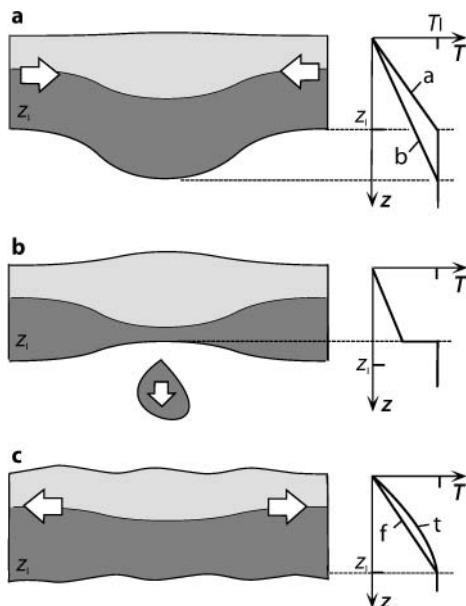
1. Delamination of the entire mantle part of the lithosphere from the crust along the Moho.
2. Convective removal of only the mechanically unstable thick root of the mantle lithosphere which, during progressive thickening, successively protrudes into the asthenosphere (Fig. 6.36).

- *Delamination.* The first mechanism was initially suggested by Bird (1979) to explain the uplift of the Colorado Plateau in the western US. However, the model implies that the asthenosphere comes in direct contact with the crust and it is therefore expected that abundant crustal melts occur in connection with the surface uplift. This is not observed in Colorado. However, mantle xenolith studies in several regions around the world show that the uppermost mantle lithosphere is significantly younger than the overlying crust, indicating that this process does happen under some conditions. Jull and Kelemen (2001) have even suggested that the lowermost (eclogitized) crust may delaminate together with the mantle lithosphere.

- *Convective removal.* The second mechanism was suggested by Houseman et al. (1981) and has since been confirmed by many field observations (e.g. Platt and England 1994). The model relies on the following argument: The uppermost part of the mantle lithosphere is so viscous that its sinking rate is geologically irrelevant, despite its high density. The viscosity of the lowest part of the mantle lithosphere, on the other hand, approximates that of the asthenosphere. This part, where heat is still being transported mainly by conduction (and therefore part of the thermally defined lithosphere), but which has a negative buoyancy and a viscosity comparable to that of the asthenosphere, is also called the *thermal boundary layer* (Parsons and McKenzie 1978) (Fig. 2.19). This part of the mantle lithosphere may be removed from the rest of the mantle lithosphere by convective processes in the surrounding asthenosphere and it may ultimately sink (Houseman et al. 1981; Fleitout and Froidevaux 1982; Houseman and Molnar 1997; Molnar et al. 1998).

**Temporal Evolution of Convective Removal.** The evolution of the convective removal of the mantle part of the lithosphere may be divided into three temporal stages:

- The first stage is the development of a lithospheric root during collision. It is necessary to develop a substantial root of the mantle lithosphere in order for the negative buoyancy forces to be large enough to cause strain rates that overcome geologically relevant values. Only then convective removal can set in.



**Figure 6.36.** Schematic illustration of the Houseman et al. (1981) model for the convective removal of the mantle lithosphere in thickening orogens. **a** Thickening of the lithosphere during collisional orogenesis (crust is shaded light, mantle part of the lithosphere dark). The arrows indicate convergent motion. The labels on the schematic geotherms in the little  $T$ - $z$  diagrams on the right are:  $b$  = before thickening,  $a$  = after thickening.  $T_1$  is the temperature at the base of the lithosphere and  $z_1$  is the thickness of the lithosphere prior to deformation. **b** Delamination of the thermal boundary layer from the mantle lithosphere. Note the uplifted surface and the dramatically changed thermal profile. **c** Subsequent thermal equilibration will cause the mantle lithosphere to thicken again. This process counteracts the simultaneously occurring rapid extension that occurs in response to the increased potential energy of the orogen. The two geotherms in the  $T$ - $z$  diagram are:  $t$  = during thermal equilibration,  $f = a$  = final stable stage

- The second stage is the *beginning* of the removal process. This process is initially very slow and may take of the order of 1–10 my.
- The third stage is the *completion* of the removal. This occurs very rapidly once the sinking velocity has reached its maximum.

At the end of the detachment of the thickened root, the mantle part of the lithosphere is of similar thickness, or even thinner than at the start of the orogenic process (Fig. 6.36b).

The numerical experiments of Houseman et al. (1981) showed that convective removal may occur much more rapidly than orogenic evolution as a whole. Removal may be completed within a total of 1–10 my, while orogenic cycles last on the order of several tens of my. Thus, convective removal of the mantle part of the lithosphere may even be a cyclic process that occurs

more than once during the evolution of a collisional orogen. Recent studies of Houseman and Molnar (1997), Molnar et al. (1998), Conrad and Molnar (1999) and others have shown that the non-linear viscous behavior of the lithosphere may cause asymmetries between the location of maximum thickening and the location of convective removal.

**Mechanical Consequences of Convective Removal.** The removal of the dense root of the mantle lithosphere of an orogen causes rapid isostatic uplift of the overlying orogen. The amount of surface uplift depends on the thickness of the removed thermal boundary layer. England and Houseman (1989) estimated about 1–3 km surface uplift, which in turn causes an increase of the potential energy of the orogen by about  $2\text{--}10 \cdot 10^{12} \text{ N m}^{-1}$ . This increase is comparable to the magnitude of several plate tectonic driving forces and will therefore significantly influence the evolution of deformation events in the orogen. In particular, it is very likely that the removal of the lithospheric root may trigger the onset of extension in convergent orogens (sect. 6.1.4).

**Thermal Consequences.** The rapid removal of a lithospheric root has the consequence that hot asthenospheric material is brought much closer to the Moho than before (Fig. 6.36b). This causes increased heat flow through the Moho and ultimately partial melting in the lower crust. The amount and chemistry of partial melts that may form due to decompression in the mantle and due to partial melting of the lower crust is discussed by McKenzie and Bickle (1988) as well as White and McKenzie (1989).

However, it is not trivial that the increased heat flow at the Moho can also lead to high temperature metamorphism in the middle crust. Because of the slow rates of heat conduction on crustal length scales it may take up to tens of millions of years until the middle crust “feels” the thermal effects at the base of the crust. In this time span extension may also have caused an increase of the geothermal gradient. Thus, a metamorphic event that occurred synchronously with removal of a lithospheric root may occur due to rapid extensional processes rather than heat conduction. Platt and England (1994) showed that if the extensional processes are short-lived, then metamorphism caused by heat flow changes at the Moho may be characterized by isobaric heating and cooling.

### 6.3.3 Low Pressure - High Temperature Metamorphism

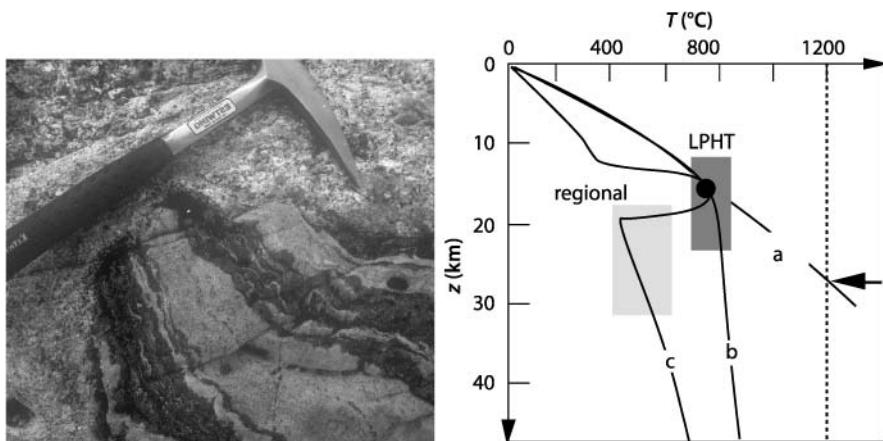
In many regions of this planet, in particular on the Precambrian shields, we can find metamorphic terrains that experienced peak metamorphism at unusually high temperatures, if compared with the depth of metamorphism (Fig. 7.1). In other words, the ratio of peak pressure to peak temperature in these terrains is much higher than that corresponding to a “normal” geothermal gradient or that predicted by models for regional (Barrovian) metamorphism (e. g. sect. 6.2.1). Such terrains are generally called “low-pressure-high-temperature”, or short LPHT- terrains and the metamorphism is often referred to as “Buchan style”. LPHT terrains occur at all grades, ranging from

greenschist facies metamorphism at less than a kilobar peak pressure (e.g. Xu et al. 1994) to granulite facies metamorphism at less than 3 or 4 kilobars (Greenfield et al. 1998). The heat sources of metamorphism in these terrains are intensely debated. In principal there are two fundamentally different heat sources that might be considered: “external” and “internal” heating.

**External Heat Sources.** One school of thought argues that the  $T - P$  ratio of peak metamorphic conditions in LPHT terrains is much too high to be possibly attainable by a conductive geotherm. Thus, so it is argued, the heat sources must originate from “outside” the terrain under consideration (the heat sources are: “external heat”) (e.g. Bohlen 1987; Lux et al. 1986; Sandiford et al. 1991). Examples of external heat sources would be heat sources that are advected from larger depths into the terrain, for example magma or fluids. This process can be considered as “contact metamorphism” in the widest sense. The most important arguments in support of this external heating model are:

- If the terrain was heated by conductive response of the lithosphere to a changed thickness geometry of crust and mantle lithosphere, then this implies that the measured  $PT$  ratio corresponds more or less to a geothermal gradient (curve *a* in Fig. 6.37). Typical  $PT$  ratios of LPHT terrains imply that a geotherm would reach the base of the lithosphere ( $\approx 1200^\circ\text{C}$ ) at a depth of about 30 km. Today, we observe such small lithospheric thicknesses only in regions of active extension and intra continental rift zones. In contrast, LPHT terrains are usually characterized by convergent structures and evidence for a plate margin setting are usually absent. Thus, alternative (external) heat sources must be made responsible.
- In many LPHT terrains metamorphism occurred synchronously with deformation. This observation is easily explained if external heat sources are responsible for metamorphism (s. p. 360). However, it is in contrast with the models that explain regional metamorphism as a function of conductive processes (s. sect. 6.2.1) (s. Problem 3.5).
- Many LPHT terrains are characterized by isobaric cooling curves. This observation indicates that the rate of cooling was substantially larger than the rate of burial or exhumation (s. sect. 7.3.1). As the duration of conductive cooling of a terrain is proportional to the square of the size of the cooling region, the rates of exhumation or burial may be used to constrain the length scale of the heated terrain. Assuming normal rates of vertical motion of rocks, such estimates indicate that only a region substantially smaller than the entire lithosphere could have been affected by the LPHT event.

**Internal Heat Sources.** In contrast to the arguments presented above, another school of thought argues that neither enough magmatic bodies nor sufficient evidence for fluid infiltration is found in LPHT terrains to justify external heat transfer into the terrains (s. sect. 3.6.4, Problem 3.14). Thus,



**Figure 6.37.** Left: Typical field appearance of LPHT terrains with multiple generations of partial melts (some syndeformational, some post deformational) and garnet – cordierite bearing melanosomes. Right: Three different models for the interpretation of geotherms in LPHT terrains. *a* is a monotonously rising geotherm. Such a geotherm implies that the base of the thermally defined lithosphere at 1200 °C is located at a depth of only 30 km (arrow). *b* and *c* show two other possibilities for geotherms that are characterized by LPHT metamorphism, but allow normal lithospheric thicknesses. However, geotherms can usually not get such a shape by conductive processes only

so it is argued, LPHT metamorphism must have similar causes as regional Barrovian type metamorphism (s. Harley 1989). In order to explain the exceptional peak metamorphic *PT* ratios a series of models have been invoked that all are based on extremely unusual thickness geometries of crust and mantle lithosphere. For example, extreme thinning of the crust *and* the mantle lithosphere may cause conditions appropriate to LPHT metamorphism. Another possibility that has recently received some attention is unusually high radioactive heat production within the crust (Chamberlain and Sonder 1990; Sandiford and Hand 1998a;b). This might lead to a geotherm of the shape of curve *b* on Fig. 6.37. Spear and Peacock (1989) discuss models of internal and external heating of metamorphic terrains in some detail. Their contribution also includes a series of computer programs that can be used to model these heating mechanisms.

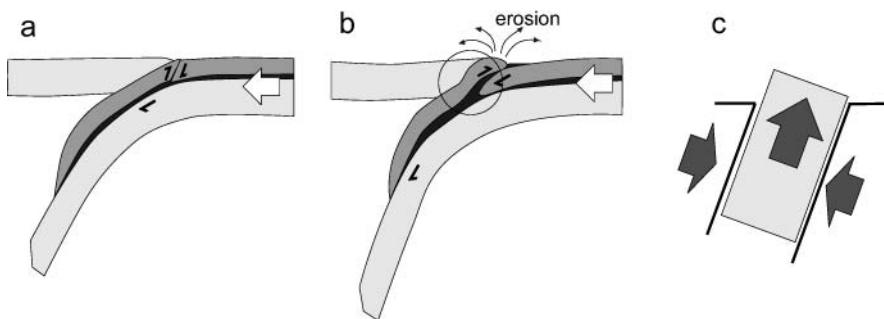
### 6.3.4 High Pressure Metamorphism

Metamorphic rocks that were buried to depths above 60 km are a common minor constituent of the lithological assembly of many orogens (Ring et al. 1999). In sect. 4.3.4 (and several others of this book) we have shown that it is difficult to exhume rocks in convergent orogens from more than 30 km depth both via extension or erosion, when considering only a simple one-dimensional

model. Therefore, the terms “high pressure metamorphism” (and: “ultrahigh pressure metamorphism” above the quartz - coesite transition, s. Fig. 7.1) have found their own meaning in the literature. The exhumation of such rocks must be considered as a two dimensional problem and a series of models have been suggested that can be used to explain exhumation from much larger depths (s. sect. 4.3.3). These models can be grouped according to the nature of the exhuming forces into (s. Platt 1993b):

1. Models that rely on forces that are applied externally to the metamorphic terrain (extrusion).
2. Models that rely on buoyancy forces caused by density differences.
3. Models that rely on extensional processes caused by gravitational stresses.

Compare this division also to the subdivision for exhumation processes made on p. 153. The third of these three mechanisms was already discussed on p. 153 and cannot be responsible for exhumation of more than from about 30 km depth. The first two both play a role in extruding high  $P$  terrains (Fig. 6.39). The currently most popular model quoted for the exhumation of high pressure rocks is the model of Chemenda (Chemenda et al. 1996; Shemenda 1994), which relies on a complicated interaction of processes from all three model groups in subduction zones. Before we discuss these model groups in some more detail, we note that the observation of high pressure metamorphic parageneses must not necessarily be interpreted as an indicator of large burial depths (e. g. Ernst 1971; Mancktelow 1993; 1995; s. sect. 6.3.5).

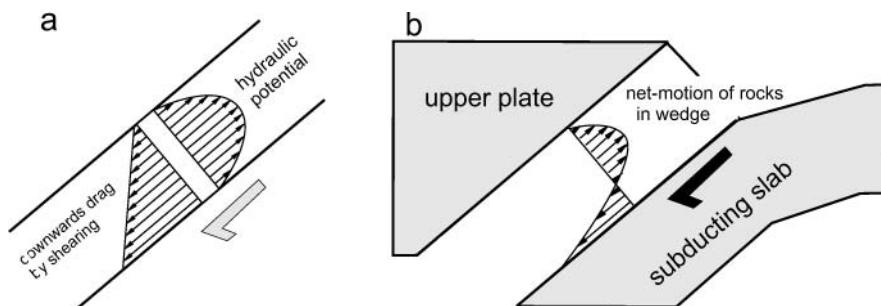


**Figure 6.38.** Exhumation of rocks according to the Chemenda (1996) model. **a** and **b** show two time steps during subduction and extrusion of the accretionary wedge; **c** shows schematically the extrusion process in the circled region in **b**

**1. Exhumation by External Forces.** Exhumation mechanisms of the first group may be characterized by the key words *Extrusion*, *strike slip faulting* and *corner flow* (s. p. 308). The process of vertical extrusion means that material is squeezed out between two hard blocks of rock, for example in a flower structure. The process is quite similar to what we discussed as “lateral

extrusion" on p. 308, but in the vertical direction. Vertical extrusion does not help to exhume the rocks by itself, but it brings rocks rapidly into a position where they may be exhumed by erosion. The amount of vertical extrusion as a function of the forces between the two hard blocks may be calculated from the relationships we discussed in sect. 5.1.1, 5.3.1 and 6.2.2 (e.g. eq. 6.21, 5.57).

The *corner flow* model is different. During the continuous deformation of accretionary wedges it is possible that rocks will be exhumed without the removal of a corresponding amount of material from the surface (s. sect. 6.2.3). However, this process can only be invoked as an explanation of exhumation if the viscosity of the rocks is very low or the exhumed rocks occur as isolated blocks in a soft melange (Cloos 1982).



**Figure 6.39.** Competing flow during shearing and buoyancy in a subduction channel (England and Holland 1979; Mancktelow 1995). The net flow shown in **b** corresponds well with the distribution of metamorphic rocks in parts of the Himalayan crystalline complex that extrudes between the Main Central Thrust and the South Tibetan Detachment zone

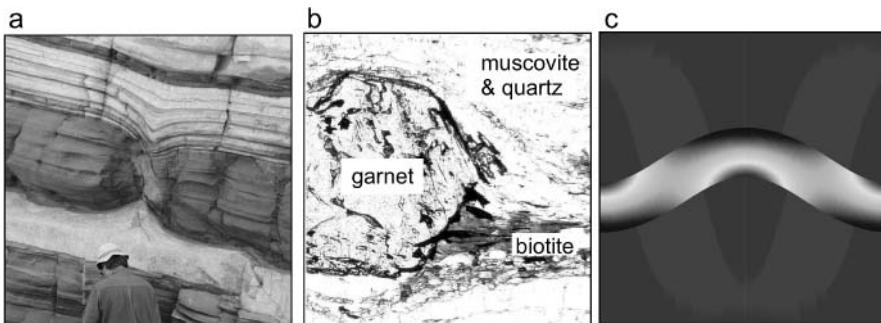
**2. Exhumation by Buoyancy Forces.** If high pressure metamorphic rocks have a lower density than their surroundings, then it is conceivable that they rise through the crust only due to their positive buoyancy - a bit like plutons in the solid state (Fig. 6.39). One example where this may occur is if crustal material is brought down into the mantle (Wheeler 1991; Chemenda et al. 1996). However, one of the most common high pressure metamorphic rocks is *eclogite* which is generally embedded into rocks of much lower density and grade. This observation cannot be explained by buoyancy forces. England and Holland (1979) observed that eclogites from the eastern Alps are often embedded in carbonates. They infer that the buoyancy of eclogites *plus* carbonates may be low compared to the density on a regional scale, so that exhumation by buoyancy forces is possible.

### 6.3.5 Tectonic Overpressure

The term “tectonic overpressure” is very loosely used for all non-lithostatic components of pressure (s. sect. 5.1.1; eq. 5.23). In other words, the component of pressure that deviates from eq. 7.1. Ernst (1971) and Rutland (1965) suggested that the non lithostatic stresses may form a significant contribution to the pressure measured with geobarometers in metamorphic parageneses. However, today most petrologists assume that this component of pressure is so small that it may be neglected completely when interpreting results from geobarometers applied to metamorphic parageneses. Whether this assumption is always justified is not trivial and needs testing. In very general terms, we can say that there are two mechanisms to create tectonic over- or underpressure:

1. Deformation (via shear stresses that contribute to pressure)
2. Static confinement (via increased principal stresses)

Deformation causes overpressure because the inherent strength of rocks is responsible for a deviation from the isotropic stress state. Then, shear stresses contribute to the mean stress. Deformation is only possible if rocks have a finite strength so that the mean stress (pressure) will *always* be different from the lithostatic stress. Static confinement may be important when rocks are squeezed into an area where all the principal stresses are increased.



**Figure 6.40.** Examples of local pressure variation in outcrop and model. **a** Partial melt in a boudin neck from Namibia. Did it form there (?) or migrate there (?) – do to local pressure gradients during boudinage? (photo by Grasemann, hat by Mancktelow). **b** Biotite growth in the pressure shadow of a garnet crystal from the Plattengneiss shear zone in the Eastern Alps. The rock contains only muscovite and quartz in the matrix and biotite forms exclusively in the pressure shadows – probably due to local decompression. **c** Finite element model of pressure variation during viscous folding (Mancktelow pers. comm. 2006). Competence contrast between central layer and matrix is 20:1. Highest pressure in the core of folds is about twice the normal flow stress

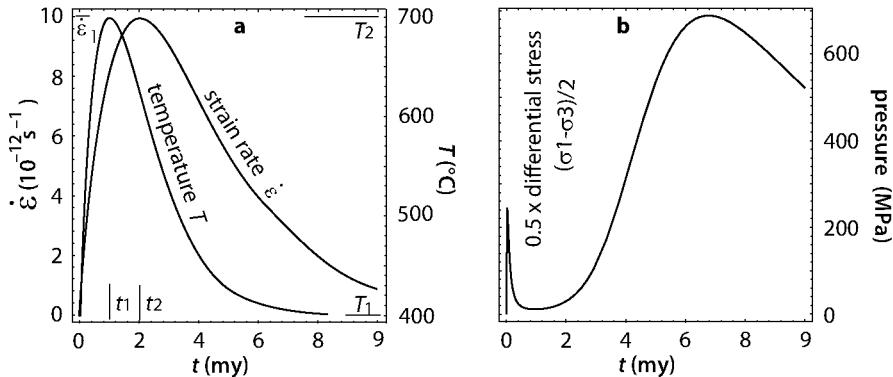
**Some Field Observations.** The fact that non lithostatic stresses have an influence on the pressure distribution in rocks is quite obvious on a large range of scales from the hand specimen scale (Fig. 6.40) to the intra plate stress field on the largest scale (Fig. 2.2). On the hand specimen scale we know of the existence of pressure shadows behind minerals, pressure solution, boudinage and folding (Fig. 6.40). In fact, any fabric formation in rocks can all only occur if there is deviations from the isotropic stress state of rocks. For the mechanical analysis of such deviations, numerical methods, in particular the finite element method are now of common use. The interested reader is referred to the classics of Strömgard (1973) and Stephansson (1974), more modern studies of Barr and Houseman (1996), Bons et al. (1997), Tenczer et al. (2001) or Schmid (2005) and – of course – to the wonderful text book of Ramsay and Lisle (2000). On a larger scale, large isothermal sections of the crust (as they are common in many areas around the world, s. Fig. 7.14) may indicate overpressure on a crustal scale.

**Estimating its Magnitude and Interpretation.** To illustrate overpressure, let us consider the different parameters that contribute to pressure during viscous deformation. For this we consider a simple convergent deformation geometry in which the largest principal stress is oriented horizontally and the smallest vertically and where we can write:  $\sigma_2 = (\sigma_1 + \sigma_3)/2$ . In this case pressure is the mean of the largest and smallest principal stress and may be described as follows: (s. sect. 5.1.1, eq. 5.23 and eq. 5.46):

$$P = \frac{\sigma_1 + \sigma_3}{2} = \sigma_3 + \frac{\sigma_1 - \sigma_3}{2} = \sigma_3 + \frac{\sigma_d}{2} = \rho g z + \frac{1}{2} \left( \frac{\dot{\epsilon}_l}{A} \right)^{(1/n)} e^{\frac{Q}{nRT}} \quad (6.33)$$

where  $\dot{\epsilon}_l$  is longitudinal strain rate, as we described it in eq. 5.46. The second term in this equation may be called the “non lithostatic component of pressure”. It depends on the material constants  $Q$ ,  $A$  and  $n$  as well as temperature  $T$  and strain rate  $\dot{\epsilon}$ . As it is very difficult to determine the material constants experimentally, geologists have tried to constrain their magnitude using field observations (e.g. Molnar and England 1990a; England and Molnar 1991; Mancktelow 1993, 1995; Stüwe 1998a).

- *Constraints from the elevation of mountain ranges.* One of the most important field observations used to constrain the magnitude of differential stress (strength) of the crust is the elevation of mountain belts: The mean elevation of a mountain ranges implies that the mean differential stress in the lithosphere must be of the order of 50 MPa (s. p. 300). This implies that the mean contribution of non-lithostatic components to pressure is about 0.25 kbar. However, considering the rheological anisotropy of the lithosphere, it is very likely that some parts of the lithosphere are significantly weaker than this and others therefore stronger. For example, the rheological model of the Brace-Goetze-lithosphere implies that differential stresses near the brittle ductile transition are of the order of several hundreds of MPa (s. Fig. 5.17 and Fig. 5.20). Thus, the Brace-Goetze-model is in contrast to the wide spread



**Figure 6.41.** This diagram illustrates that the temporal evolution of differential stress need not correlate in any intuitive way with either the evolution of temperature or the evolution of strain rate. This is because differential stress in the ductile regime has such a complicated exponential and power law dependence on temperature and strain rate (eq. 6.33). **a** Schematic illustration of temperature and strain rate in a metamorphic terrain. These curves were calculated with the purely artificial assumptions that temperature may be described by:  $T = T_1 + (T_2 - T_1)(t/t_1)e^{(1-t/t_1)}$  and strain rate by:  $\dot{\varepsilon} = \dot{\varepsilon}_1(t/t_2)e^{(1-t/t_2)}$ . **b** The differential stress contribution to pressure with the assumed temperature and strain rate evolutions from **a**. Calculated with eq. 6.33 using  $A = 2 \cdot 10^{-4} \text{ MPa}^{-3} \text{ s}^{-1}$ ;  $Q = 2 \cdot 10^5 \text{ J mol}^{-1}$ ;  $n = 3$

opinion that differential stresses are of negligible influence to the interpretation of geobarometric data.

- *Constraints from metamorphic PT paths.* Metamorphic PT paths – in particular those from low- $P$  high- $T$  terrains – show a large variability including decompression, compression and isobaric cooling. Many of these paths are constrained only by directions of PT changes rather than by absolute pressure changes and it has therefore been argued that they may reflect counter intuitive variations of non-lithostatic stresses, rather than depth changes. From eq. 6.33 we can see that temperature and strain rate are related to differential stress by complicated exponential and power law functions. Thus, it is not trivial to see if differential stresses (and therefore the tectonic overpressure) will rise or fall when the temperature or the strain rate in a metamorphic terrain change. In low- $P$  high- $T$  terrains deformation and temperature change are often intimately related, as shown, for example in Fig. 6.41a. Fig. 6.41b shows the differential stress that results from these strain rate and temperature evolutions if inserted into eq. 6.33. Despite the simple and intuitive shape of the functions in Fig. 6.41a, the resulting evolution of stress as a function of time is completely counter intuitive. Fig. 6.41b should serve us as a warning to attribute too much significance to some complications observed in P-T paths of metamorphic rocks (s. Stüwe and Sandiford 1994).

### 6.3.6 Feedback and Episodicity

Feedback between different geological processes is a common phenomenon in the earth science. In general, we discriminate between two types of feedback processes:

- positive feedback,
- negative feedback.

In both cases one process has an effect on another, the changes of which influence in turn the first process. Feedback processes are called *positive* if two or several processes “accelerate” each other. Feedback processes are called *negative*, if two or several processes hinder each other. We have encountered positive feedback processes between lithospheric extension and thermal weakening in Fig. 6.15: There, the onset of lithospheric extension causes steeping of the geothermal gradient (thermal weakening), which accelerates the extensional strain rates, which in turn results in accelerated thermal weakening. The process begins to “run away” until rifting occurs. Another currently very topical positive feedback processes occur between global glaciation and CO<sub>2</sub> content of the atmosphere possibly resulting in very rapid formation of a snowball earth (North et al. 1981; Hoffman et al. 1998). Negative feedback is more intuitive, for example the feedback between potential energy and effective driving force during collisional orogenesis (Fig. 6.23). There, collision causes an increase of the potential energy of a new orogen which opposes the driving force and therefore causes the collisional strain rate to decrease, which in turn leads to a slowed increase of the potential energy until orogenesis comes to a rapid halt. Because of feedback processes the question on the *cause* or *consequence* of one or the other geological process is often difficult to answer. Most feedback processes in the earth sciences are non-linear, that is, there is not a direct linear relationship between cause and consequence.

**Non-linear Feedback.** Consider a very simple theoretical example. If we iterate the non-linear function  $x = x^2$  for many iterations, then the result will depend on our assumption for the starting value of  $x$ . For all positive starting values of  $x$  that are smaller than 1, this function will converge toward zero. For all starting values of  $x$  that are larger than 1 this function will diverge towards  $x \rightarrow \infty$ . Because of the non-linearity, it will do so at an increasingly large rate, the more iterations we perform. Only a single starting value, namely  $x = 1$ , separates the two different trends. (If the starting value is not a real number but a complex number, then this critical starting guess separating different evolutions will become a series of starting values forming a line in complex plane. This line often has a complicated fractal shape and is known by the name *Julia set*).

Many geological processes behave that way due to non-linear feedbacks: depending on the starting situation, they either “run away” or converge to a steady state. For instance, in the example we discussed in Fig. 6.15, we

noted that there is single set of starting conditions separating two completely different orogenic evolutions. These are evolutions terminating in self-limiting extension and those terminating in “run away” rifting processes. Even a very small difference in starting conditions is sufficient to result in these very different evolutions, if the two starting conditions lie on either side of this critical set of starting conditions. The example is therefore equivalent to the simple illustrative iteration of  $x = x^2$ .

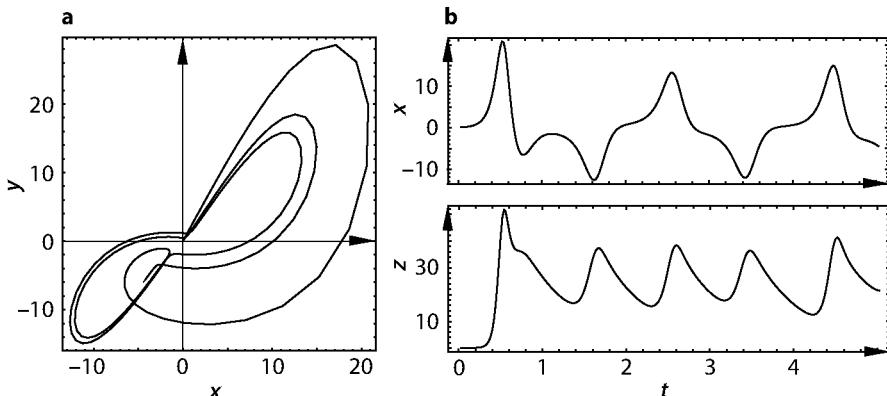
For some functions, convergent and diverging evolutions are *not* separated by a single starting value, but by a whole region of starting values. Within this critical region the evolution may behave *oscillatory* or *chaotic*. In this context *chaotic* means that the evolution of the function may not be predicted directly from the choice of starting parameters.

- *Lorentz Attractor*. Although it has no direct geodynamic applications, the so called Lorentz attractor provides a nice simple illustration how non linear feedback processes may cause chaotic and oscillatory behavior of processes. The Lorentz attractor describes the feedback between three independent differential equations:

$$\begin{aligned}\frac{dx}{dt} &= -10x + 10y , \\ \frac{dy}{dt} &= -xz + 28x - y , \\ \frac{dz}{dt} &= xy - \frac{8}{5}z .\end{aligned}\tag{6.34}$$

All three equations are functions of the same parameters:  $x$ ,  $y$  and  $z$ . Also, all three equations are non-linear (the fact that variables occur as themselves *and* as their own derivative within the same equation indicates that they are exponential functions; s. p. 416). Numerical simultaneous solution of these three equations using the method of finite differences is straight forward and results in the irregular temporal evolution of the three variables  $x$ ,  $y$  and  $z$  that is illustrated in Fig. 6.42b. We can see that  $x$ ,  $y$  and  $z$  have unpredictable ups and downs as a function of time, although each of the three functions in eq.6.34 are continuous functions. It is therefore very tempting to interpret the  $x$  and  $z$  as shown in Fig. 6.42b as – for example – the evolution of metamorphism and deformation of a polymetamorphic terrain and suggest that the repeated occurrence of deformation or metamorphic events in orogens may have nothing to do with separate events and is solely a function of non-linear feedback processes.

**Geological Examples of Non-linear Feedback.** Non-linear feedback processes are very familiar to geophysicists who model dynamic processes, but their potential influence on shaping the observations that field geologists make in the outcrop is not appreciated. In this section we therefore illustrate feedback processes with three examples with direct relevance to field observations.



**Figure 6.42.** a The Lorenz-attractor in the  $x$ - $y$ -plane, calculated with the eqs. 6.34. b chaotic evolution of the variables  $x$  and  $z$  plotted as a function of time  $t$

• **1. Feedback between erosion and deformation.** Feedback processes between surface erosion and the tectonic and thermal evolution of the crust are well recognized (England and Richardson 1977; Willett 1999). Even feedbacks between climate changes and mantle processes have been suggested (e. g. Molnar et al. 1993; Koons 1990; Harrison et al. 1992; Pinter and Brandon 1997). Many of these models are based on the fact that the rate of erosion and that of the thermal equilibration on the scale of the crust are comparable. As a consequence, erosion may or may not advect heat towards the surface, depending on the relative rates. This in turn may cause thermal weakening, which may be reflected in deformation events (Hoffman and Grotzinger 1993). We have discussed this interaction in several sections of this book and it is well documented in New Zealand, where the difference in surface heat flow between the west and the east coast of the South Island may be correlated with the difference in rain fall on the two sides of the island (Koons 1990). Similarly, Molnar and England (1990b) have speculated that the Indian Monsoon is closely linked to the uplift of the Tibetan Plateau and Zhou and Stüwe (1994) showed that this connection of events is only possible for exceptional rates during orogenesis.

• **2. Deformation and metamorphism – cause or consequence.** Observations from many metamorphic terrains show that there are characteristic temporal relationships between metamorphism and deformation indicating that feedback processes between deformation and metamorphism are likely. In principle, there may be three different timing relationships: the deformation of a terrain may occur pre-, syn- or postdate to metamorphism (s. sect. 7.4.1). Clearly, it is useful to know if one caused the other, or if there are feedback processes between metamorphism and deformation.

Models that are used to explain regional Barrovian type conductive metamorphism (where deformation precedes metamorphism) consider metamorphism to be the *consequence* of deformation (e.g. the model of England and Richardson, s. sect. 6.2.1): the deformation causes thickening which in turn causes a thermal disequilibrium. The subsequent thermal equilibration that causes metamorphism is therefore the result of deformation (s. sect. 7.4.1; Fig. 7.13). Conversely, metamorphism can cause a deformation event. For example, if a strong lithospheric plate is thermally-weakened, then it might deform in response to the thermal event, while the far field stress field remains constant (s. eq. 6.23, sect. 5.3.1). In this model, deformation is the result of metamorphism and deformation is likely to be *syn-* metamorphic.

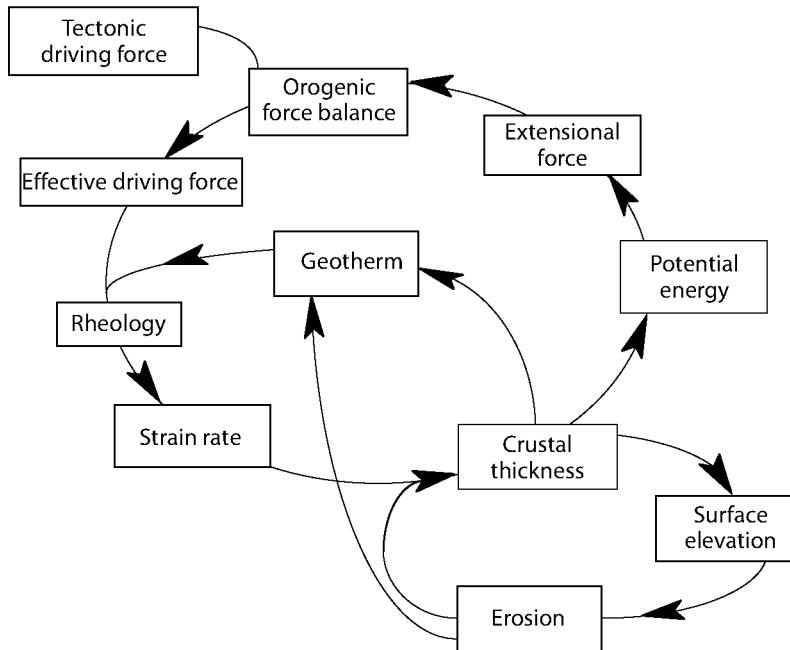
- *3. Orogeny as a non-linear feedback between many processes.* Mountain building processes are characterized by a series of non-linear relationships between physical parameters (e.g. Malanson et al. 1992; Hodges 1996; Stüwe et al. 1993). Fig. 6.43 gives an overview over some non-linear processes and their interaction during orogenesis. Many of the processes illustrated there may be described by mathematical relationships not unlike those of Eq. 6.34. A correlation between the evolutions of  $x$ ,  $y$  and  $z$  on Fig. 6.42b and strain rate of thermal events may therefore be not completely wrong.

**Episodicity in Geological Events.** Most geological events occur episodically: Field geologists map out *discrete* events of repeated deformation and metamorphism (s. Fig. 7.11); volcanoes erupt cyclically; earthquakes, mass extinctions, glaciations, magmatic activity and orogenesis all occur over time spans that are short compared to the periods of quiescence before they occur again (e.g. Waschbush and Royden 1992; Malniverno and Pockalny 1990). In a very general way, the episodicity of a process may be explained in three different ways:

- Episodicity at the boundaries,
- threshold mechanisms,
- non-linear feedback.

The last of these three mechanisms was dealt with in the last section. The first mechanisms (“episodicity at the boundaries”) simply refers to an explanation of the cyclicity of events by cyclicity of events elsewhere. For example, the cyclic nature of tectonic events in the western US has been interpreted as the consequence of the episodic subduction of the Pacific Plate under the North American continent (Elison 1991). This interpretation explains the cyclic occurrence of events in North America, but it does not explain episodicity as such. Rather, it defers the causes of episodicity to an area outside the region of consideration - in this case outside North America.

- *Threshold mechanisms.* The most common way how continuous processes are broken up into discrete episodically occurring events is by *threshold mechanisms*. In principle this means that a certain value of a parameter – the threshold – needs to be exceeded by a continuous evolution before a process can occur (Tong 1983; Hergarten 2002). Two examples:



**Figure 6.43.** Interaction and feedback between some selected processes during orogenesis. Note that the functional relationship between several of the processes is non-linear. For example, the strain rate is an *exponential* function of temperature, a *cubic* function of stress (eq. 5.46) and potential energy is a quadratic function of thickness (eq. 5.49). It is therefore possible that interaction between the different processes may lead to a chaotic evolution of processes during orogenesis

- Earthquakes occur if a certain *threshold stress* is exceeded. During deformation elastic stresses are cumulatively built up until the failure stress of a rock is exceeded. An earthquake dissipates these stresses then by deformation. This is followed by a period of quiescence during which stresses build up again, until the threshold stress is exceeded again.
- Partially molten rocks often intrude into higher crustal levels. For a melted rock to leave its source rocks and form an intrusion a *threshold melt volume* must be exceeded in the source rock (Wickham 1987).

Both the above examples are characterized by a regular, but episodic evolution, even though the applied boundary conditions remain constant. However, in reality, earth quakes and magmatic activity occur much more randomly than the above models and thresholds suggest. This is because the thresholds in question are coupled with a spatial length scale. For example, an earth quake will not occur if the necessary threshold stresses are only exceeded along a single micro fracture. Instead, it will be necessary that the mean stress over a certain area exceeds the threshold stress. Similarly, magma

segregation will not occur if a critical melt volume is exceeded on the scale of a thin section. Instead, it will be necessary that melt can only segregate once the critical melt volume is exceeded over a certain length scale.

Thus, in order to be able to predict the occurrence of earth quakes or magmatic intrusion (or any other process relying on threshold mechanisms), it is necessary to know the relationship between the magnitude of the threshold and the length scale over which this threshold needs to be exceeded. Sadly, this relationship is a function of so many variables that is practically impossible to know it. The problem is akin to the problem of meteorologists, who would need to know the motion of every gas molecule in the atmosphere in order to be able to predict weather patterns with any accuracy. The fact that infinitely small variations of a single small parameter are sufficient to bring the whole system out of predictability, this called the *butterfly effect*, reminding us of the possibility that the air current caused by the flight of a butterfly may cause complete weather patterns to change.

## 6.4 Problems

**Problem 6.1.** *Extension: uplift or subsidence? (p. 266):*

Fig. 6.2 illustrates that continental extension can lead to surface uplift or to surface subsidence, depending on the initial thickness ratio of crust and mantle lithosphere and their densities. Use eq. 4.35 and Fig. 4.18 to understand which parameters control whether uplift or subsidence will occur (s. also Problem 6.6).

**Problem 6.2.** *Porosity estimates (p. 271):*

Use eq. 6.1 to estimate the depth at which the porosity of sandstone is the same as that of shale. Use the parameter values given in the caption for Fig. 6.5.

**Problem 6.3.** *Compaction of sediments (p. 272):*

Use eq. 6.4 to calculate the thickness of a 100 m thick sandstone unit at 2 000 m depth at the time of its deposition. For the porosity data of sandstone use those in Fig. 6.5.

**Problem 6.4.** *Isostasy of sedimentary basins (p. 275):*

Derive eq. 6.5 using Fig. 6.7 and eq. 4.23.

**Problem 6.5.** *Subsidence analysis (p. 277):*

Describe the subsidence evolution of the sedimentary basin characterized below. What tectonic process might have formed the basin and discuss field observations you would look for to test your model? What subsidence mechanism might have been responsible for the evolution of subsidence you have derived? Note that your subsidence analysis must include back stripping to answer these questions.

The basin we consider is filled by a 6 km thick pile of sediments. A drill hole has shown that the pile is made up of five strata which we number from the base to the top with  $i = 1$  to  $i = 5$ . Fossils were found in each layer and allowed to determine a detailed record of the sedimentary succession: The 1. layer is a 3 km thick sandstone layer that was deposited during an increasing water depth from 500 m to 800 m between 55 my and 60 my. The second layer is a 500 m thick shale layer that was deposited in a water depth of 800 m between 40 my and 55 my. The third layer is a 1 km thick sandstone unit, deposited during decreasing water depth from 800 m to 300 m in the time between 40 my and 30 my. The 4. layer is a 500 m thick shale unit that was deposited during a decrease of water depth from 300 m to 0 m in the time between 30 my and 15 my. The 5. layer was deposited between 15 my and present in the tidal environment.

For your analysis use Eq. 6.1, 6.4 and 6.5 (note that you will also need Eq. 6.7 for Eq. 6.5). Use the values for densities and porosities from Fig. 6.5. The density of the asthenosphere is:  $\rho_a = 3\,200 \text{ kg m}^{-3}$ .

**Problem 6.6. Extension: uplift or subsidence? (p. 278):**

According to McKenzie (1978) homogeneous stretching of the lithosphere will only lead to subsidence if the initial ratio of crustal thickness to lithospheric thickness exceeds a certain ratio. Derive how large this ratio is using Eq. 6.10 and making the following assumption for the physical parameters.  $\rho_0 = 3\,300 \text{ kg m}^{-3}$ ;  $\rho_c = 2\,750 \text{ kg m}^{-3}$ ;  $\rho_w = 1\,000 \text{ kg m}^{-3}$ ;  $\alpha = 3 \cdot 10^{-5} \text{ }^\circ\text{C}^{-1}$ ;  $T_1 = 1\,280 \text{ }^\circ\text{C}$  and  $z_1 = 130 \text{ km}$ . (See also Problem 6.1.)

**Problem 6.7. Estimating thermal sag (p. 279):**

Estimate roughly the total duration of the thermal sag phase that may be expected in a hypothetical basin forming in Europe, where the lithosphere is roughly 100 km thick and another one in southern Africa, where the lithosphere is almost 200 km thick. Use eq. 6.11.

**Problem 6.8. Radioactive contribution to heat flow (p. 290):**

Table 6.1 lists a range of geologically realistic radiogenically caused heat flows  $q_{\text{rad}}$  and mantle heat flows  $q_m$  for the continental lithosphere. Calculate the expected equilibrium surface heat flow in lithosphere that is (a) doubled in thickness and (b) halved in thickness using the logic outlined in eq. 6.13 to eq. 6.15

**Problem 6.9. Understanding orogenic force balance (p. 295 – p. 301):**

A mountain belt has formed in response to the collision of two plates. How high can the (isostatically supported) mountain range grow if the driving force for collision is  $F_d = 5 \cdot 10^{12} \text{ N m}^{-1}$ ? Use eq. 6.21. Tackle the problem in 3 different ways and compare the answers: (a) Solve the problem graphically by using the logic of eq. 6.21 and comparing Fig. 4.18 and Fig. 5.32 (assuming that only the crust thickens but that the complete lithospheric thickness remains unchanged, i.e.  $f_l = 1$  at all times). (b) Calculate the answer using eq. 5.57

and  $z_c = 30$  km,  $\rho_c = 2700 \text{ kg m}^{-3}$ ,  $\rho_m = 3300 \text{ kg m}^{-3}$  and  $g = 10 \text{ ms}^{-2}$ . (c) Refine that answer of (b) by using eq. 5.59 (using  $\alpha = 3 \times 10^{-5} \text{ }^\circ\text{C}^{-1}$ ,  $T_l = 1200 \text{ }^\circ\text{C}$  and  $z_l = 100$  km as additional variables) to calculate the buoyancy force and eq. 4.31 to calculate the isostatically supported surface elevation (assuming again that only the crust thickens but that the complete lithospheric thickness remains unchanged), (s. also Problem 5.17.)

**Problem 6.10. Estimating the Argand number (p. 305):**

Use eq. 6.31 to make a rough estimate how an orogen will look like in which the thickness increase is  $L = 50$  km,  $\rho_c = 2700 \text{ kg m}^{-3}$ ,  $\rho_m = 3300 \text{ kg m}^{-3}$  and the collision rate is 5 cm per year. For the parameter  $B$  use the relationship given in the text following eq. 6.29 (see eq. 5.46) and the values for the material constants of olivine given in Table 5.3. Calculate the Argand number for  $400 \text{ }^\circ\text{C}$  and  $500 \text{ }^\circ\text{C}$  and discuss what difference this makes.

**Problem 6.11. Estimating tectonic overpressure (p. 327):**

How large is the tectonic overpressure in a quartz dominated rock in 10 or 15 km depth if the thermal gradient from the surface to that depth is  $30 \text{ }^\circ\text{C km}^{-1}$  and the strain rate is large enough to double the thickness of the crust in 5 my? Use eq. 6.33 and the rheological data for quartz given in Table 5.3. Assumed eq. 5.46 to be the governing constitutive relationship.

## 7. Metamorphic Processes

One of the basic data sets used by geologists for the geodynamic interpretation of a metamorphic terrain is the spatial and temporal evolution of pressure  $P$ , temperature  $T$  and deformation  $D$  that the rocks experienced: the metamorphic evolution of the rocks. Data on the metamorphic evolution are particularly important when interpreting ancient orogens where it is impossible to measure many other parameters directly (e.g. surface elevation, surface heat flow, gravity etc.). The relative evolution of pressure, temperature and deformation may well be illustrated as curves in  $P$ - $T$ -space. Such curves are called  $P$ - $T$ -paths or  $P$ - $T$ - $t$ - $D$ -paths, if the path is also labeled for deformation events and time. As the interpretation of metamorphic rocks is so crucial to any geodynamic interpretation performed by a field geologist, we dedicated it here its own chapter. For detailed treatment of thermodynamics underlying all petrological studies of  $P$ - $T$ -paths we recommend: Anderson and Crerar (1993) or: Atkins (1994). For more petrologically oriented texts with geodynamic applications we recommend: Spear (1993) or: Spear and Peacock (1989).

### 7.1 Introduction

Most tectonic processes are characterized by process-specific spatial and/or temporal relationships between the evolutions of pressure, temperature and deformation. Thus, when inferring tectonic processes from such relationships, it is important to subdivide these three parameters in as much detail as possible: The thermal evolution may be subdivided into a heating phase and a cooling phase, the baric evolution may be subdivided into a phase of increasing pressure and into one of decreasing pressure and the evolution of deformation may be subdivided into phases of increasing and phases of decreasing strain rate. When collecting field observations it is important to discriminate between two different sets of data pertaining to two different questions:

- What is the temperature, pressure and deformation evolution of a single rock? (*temporal evolution*)

- How does this temporal evolution change in space across the terrain in question; for example as a function of a regionally decreasing or increasing metamorphic or strain gradient? (*spatial structure*)

The more detailed these questions can be answered in the field or in the laboratory, the easier it is to constrain the tectonic processes that formed the terrain in question. However, answering these questions is not always easy. An observed regional strain gradient could be the consequence of both a regional change in the *duration* of deformation or a regional change in *strain rate*. In other words, it remains unclear if the observation should be interpreted in terms of a *spatial* or *temporal* change. Corresponding problems may arise during the interpretation of pressure or temperature changes: A rise of temperature need not imply heat conduction and may be related to other thermal processes like those discussed in sect. 3.1, 3.2 and 3.3. A change in pressure need not correlate with a change in depth. Finally, the interpretation of these questions is complicated by the fact that the evolution of temperature, pressure and deformation may not be independent. Thermal expansion may have an influence on the pressure field, pressure changes may be followed by adiabatic heating or friction heat released during deformation may influence the thermal evolution.

Fortunately, such couplings do not necessarily hinder the geodynamic interpretation. To the contrary, this interaction of processes is often the very reason for very process specific spatial or temporal relationships between metamorphic rocks. They provide an exciting playground for the geodynamically oriented field geologist in which we can probe our art of interpretation of metamorphic *P-T*-paths.

### 7.1.1 What Exactly are *P-T*- and *P-T-t-D*-Paths?

*P-T*-paths are curves on a diagram in which pressure and temperature form the axes. Thus, a *P-T*-path illustrates the relative change of pressure and temperature in a rock, but it cannot show the temporal evolution of either of them. *P-T*-diagrams are *parametric* diagrams because they show the relationship between two independent functions of the same variable (in this case: time). In rocks it is much easier to document the *relative* evolution of pressure and temperature, than the absolute temporal changes of these two parameters. Thus, *P-T*-diagrams have become one of the standard tools for the interpretation of metamorphic rocks. If absolute timing information is available, or the timing of deformation relative to pressure and temperature is known, then we speak of *P-T-t*- or even *P-T-t-D*-paths.

**Relationship Between Pressure and Depth.** The interpretation of *P-T*-paths is often made with the basic assumption that pressure is only caused by the weight of the overlying rock column, i.e. it is *hydrostatic* or, as geologists say: *lithostatic*. Then the principal stresses are of the same magnitude in all

spatial directions:  $\sigma_1 = \sigma_2 = \sigma_3$ . Moreover, if this stress field is oriented parallel and normal to the earth surface, then  $\sigma_1 = \sigma_2 = \sigma_3 = \sigma_{xx} = \sigma_{yy} = \sigma_{zz}$  and pressure correlates directly with depth, more specifically: to the vertical normal stress. At depth  $z$  this may be calculated by integrating the density over depth (s. also sect. 4.4.1):

$$P = \sigma_{zz} = \int_0^z \rho(z) g dz . \quad (7.1)$$

If the density is independent of depth, then this integral is easily solved. It is:  $P = \rho g z$ . Checking the units confirms that this equation is dimensionally consistent: stress = force per area, force = mass times acceleration and mass = density times volume. For a 10 km thick column of rock with the density  $\rho = 2700 \text{ kg m}^{-3}$ , we get a mass per area of:  $2700 \times 10000 \text{ kg m}^{-2}$ . Multiplied with the gravitational acceleration ( $g \approx 10 \text{ m s}^{-2}$ ), this gives a pressure of  $2.7 \cdot 10^8 \text{ kg m}^{-1} \text{ s}^{-2} = 2.7 \cdot 10^8 \text{ Pa} = 270 \text{ MPa} = 2.7 \text{ kbar}$ . In short:

$$P(\text{kbar}) \approx 0.27 \times z(\text{km}) \quad \text{or:} \quad z(\text{km}) \approx 3.7 \times P(\text{kbar}) \quad (7.2)$$

In words: the depth in kilometers times  $\approx 0.27$  gives us the pressure in kilobar. Correspondingly, the pressure in kilobar times  $\approx 3.7$  gives us the depth in kilometers. The errors that arise in this estimate from rounding of the gravitational acceleration or the density changes with depth are significantly smaller than the accuracy of geobarometers and may therefore be neglected. If the three principal stresses are *not* the same (as they cannot be in deforming rocks), the deviatoric stresses contribute to the overall pressure and it has been of continuous debate if and how this “tectonic overpressure” is part of the pressure measured by petrologists when applying geobarometers to rocks (see sect. 6.3.5).

Note that pressure is generally given in kilobars in the geological literature, in energy per volume by the thermodynamics community and in Pascal by the geophysicists. The conversions between these units are given by:  $1 \text{ kbar} = 100 \text{ MPa} = 10^8 \text{ Pa} = 10^8 \text{ J m}^{-3}$ .

### 7.1.2 The Metamorphic Facies

The  $P$ - $T$  space relevant to crustal rocks may be divided into *metamorphic facies*. The metamorphic facies concept was introduced by Eskola (1915) and is widely used by metamorphic geologists (e.g. Yardley 1989). Nevertheless, placing exact boundaries on facies fields is difficult. This is largely due to the strong dependence of parageneses on bulk composition and the fact that parageneses, not individual minerals are characteristic of metamorphic grade. Modern petrologists may argue that the facies concept is obsolete in view of the availability of pseudosections. However, the idea is still well entrenched in the literature and certainly justifies a small excursion here. In our division (Fig. 7.1a) we follow Hacker et al. (2003) and Brown (2006) and use phase

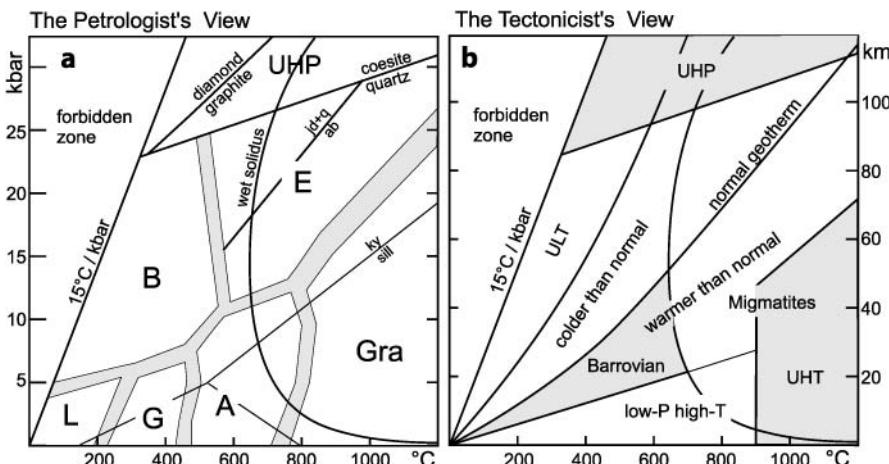
diagrams of Dale et al. (2005) and White et al. (2001), but we also introduce a new facies concept that appears useful to distinguish between metamorphic facies fields used by petrologists and facies fields as used by tectonicists (Fig. 7.1). Petrologists use the facies concept to group rocks with similar characteristic parageneses. Tectonicists use the concept to discern between  $P$ - $T$  fields reached by different tectonic processes. For both, metamorphic facies space is limited by a region which is simply too cold to be reached by metamorphic rocks. Brown (2006) suggests that this limit is given by a  $T/P$  ratio lower than  $15^{\circ}\text{C} / \text{kbar}$ . Although there is – in principle – no limit, no metamorphic rocks have been found with lower  $T/P$  ratios. Above this ratio there are a series of facies fields:

**Metamorphic geologists** discern the following main fields (Fig. 7.1a): *Greenschist facies* is defined in mafic rocks by the presence of actinolite above about  $250^{\circ}\text{C}$  to  $300^{\circ}\text{C}$  (instead of glaucophane at blueschist and hornblende at amphibolite facies). Below this there is a series of low grade fields including prehnite, pumpellyite and zeolite facies that will not be discussed here. In pelitic rocks, it starts with the coarse grained appearance of white micas.

*Amphibolite facies* is defined by the first appearance of staurolite in metapelites. For subaluminous pelites, staurolite appears at the KFMASH univariant breakdown of garnet + chlorite to staurolite + biotite (the “staurolite isograd”) above a line from about  $570^{\circ}\text{C}$  at 4 kbar to  $630^{\circ}\text{C}$  at 9 kbar. For aluminous pelites, staurolite may appear in continuous equilibria at somewhat lower  $T$ . In mafic rocks, the amphibolite facies begins with the appearance of hornblende in continuous reaction at similar  $P$ - $T$  to the staurolite isograd. Within the amphibolite facies, amphibolites begin to bear garnet at high  $P$  and epidote at low  $T$ , giving rise to a series of possible subdivisions. The wet solidus of pelitic rock is also within the amphibolite facies field and even the muscovite breakdown (the “effective” solidus) is generally considered to be part of this facies.

The start of *granulite facies* is defined by the first appearance of orthopyroxene in mafic rocks. In fact, a “granulite” in the narrow sense of the word is an orthopyroxene bearing gneiss. In pelitic rocks, granulite facies is generally considered to begin above the KFMASH univariant breakdown of biotite + sillimanite to garnet + cordierite, a reaction with positive slope from about  $700^{\circ}\text{C}$  at 2 kbar to  $800^{\circ}\text{C}$  at 8 kbar. These values are similar to those of the first appearance of orthopyroxene in mafic rocks. In pelitic rocks, the sillimanite - kyanite transition marks the boundary between low- $P$  and high- $P$  granulites. In mafic rocks the continuous appearance of jadeite-rich pyroxenes and the disappearance of plagioclase marks a broad transition to eclogites at roughly 15 - 20 kbar.

*Eclogite and blueschist facies* cover the high- $P$  region. Their boundary is very loosely defined by the disappearance of garnet to lower  $T$ . The limit to the lower  $P$  fields is defined by the disappearance of lawsonite and glaucophane from the blueschist field and the disappearance of zoisite from the

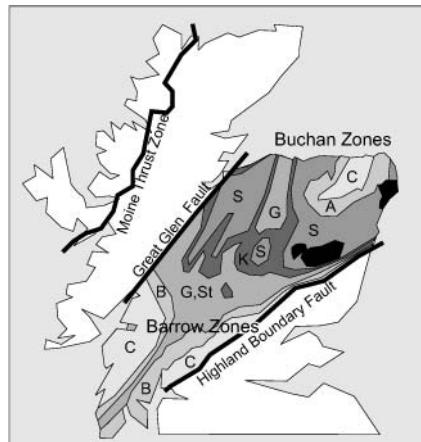


**Figure 7.1.** Metamorphic facies according to the interests of the metamorphic petrologist (left) and geodynamicist (right). Shown is the  $P$ - $T$  region of relevance to the lithosphere and fields are inserted semi-quantitatively. The axes are the same in both diagrams but labeled in kbar in **a** and in km (assuming a density of  $2800 \text{ kg m}^{-3}$ ) in **b**. In the “forbidden zone” phase equilibria are not drawn. Bands delineating fields in **a** indicate that boundaries vary according to bulk composition. The normal geotherm in **b** is drawn as to pass about  $500^\circ\text{C}$  at 35 km depth and reach  $1200^\circ\text{C}$  in 120 km depth for reasons discussed in chapter 3. Abbreviations are: L = low grade; G = greenschist; A = amphibolite facies; Gra = granulite facies; B = blueschist facies; E = eclogite facies; UHP = ultrahigh pressure facies; UHT = ultrahigh temperature facies. Minerals are: sill = sillimanite; ky = kyanite; jd = jadeite; ab = albite; q = quartz. Field boundaries are discussed in the text

eclogite field. Again, these phases disappear in continuous reaction so that these boundaries are wide bands. Only the boundary to *ultrahigh pressure facies* (UHP metamorphism) is sharply defined by the quartz - coesite transition.

**Tectonicists** divide  $P$ - $T$  space into conditions that are warmer and those that are colder than a normal continental geotherm (Fig. 7.1b). On the *warm side* a line with  $\approx 100^\circ\text{C}$  per kilobar limits conditions that can possibly be reached conductively (Barrovian style metamorphism) to the low- $P$  high- $T$  facies. This slope corresponds to a situation where the base of the lithosphere (i.e.  $\approx 1200^\circ\text{C}$ ) is located at the Moho at normal depth ( $\approx 35$  km), a situation that would lead to wholesale melting and rift formation and is rarely realized. It is suggested here that this line provides a useful boundary between a “conductive” or “Barrovian” part of  $P$ - $T$  space and the *low-P high-T facies* (Buchan style metamorphism) where metamorphism must be related to contact effects in the widest sense. On the *cold side* of a normal geotherm we suggest to discern a “*colder than normal*” and an “*ultra low temperature*” (ULT) facies field. The “*colder than normal*” field encompasses all conditions on the “*cold*” side of a normal geotherm that can be reached

**Figure 7.2.** Two classic metamorphic locations: The Barrow and Buchan zones in Scotland. Note that the Buchan zones form a concentric pattern, while the Barrow zones increase in metamorphic grade with distance from the fundamental structures of the Highland Boundary and the Great Glen Faults. Zones are labeled as follows: C = chlorite zone; B = biotite zone; G = garnet zone; St = staurolite zone; A = andalusite zone; S = sillimanite zone; K = kyanite. Granites are shown in black.



by one-dimensional instantaneous thickening of the crust up to a maximum thickness where it is in potential energy equilibrium with standard plate tectonic driving forces, i.e. about double normal thickness or half the geothermal gradient. Beyond that is the ULT field that lies at conditions less than half of a normal geothermal gradient. This region can only be reached by processes other than normal crustal thickening, for example in subduction zone settings. Finally, ultrahigh pressure metamorphism (UHP) is as described for Fig. 7.1a and the ultrahigh temperature metamorphic facies field (UHT) is defined somewhat arbitrarily to be above 900°C, a temperature that requires quite unusual magmas as heat sources.

**Barrovian and Buchan style metamorphism.** Metamorphic facies fields reached by normal collisional processes causing regional (conductive) metamorphism as we discussed it in sect. 6.2.1 are also referred to as Barrovian metamorphism. This refers to the Scottish Barrow zones where there is the type sequence of regional metamorphic facies grades (Fig. 7.2). The type location for the other end member of metamorphism – contact metamorphism – lies also in Scotland: the Buchan zones. The Buchan areas exposes a sequence of zones that include all three aluminosilicate polymorphs and is interpreted to be related to a contact metamorphism.

## 7.2 Basic Principles of Petrology

Much of petrology is concerned with the physical and chemical conditions that prevailed at the time of formation of a rock. As such, petrology is inseparably connected with structural geology. Textural interpretation of microstructures and reaction textures in thin section is one of the basic tools to infer both the structural and the metamorphic evolution of a rock. Any reader thinking of him- or herself as either a petrologist or a structural geologist only, is warned when embarking on field or laboratory work with the

aim of explaining the tectonic evolution of a region. Because petrology and structural geology are inseparable field tools for a geologist trying to understand the tectonic evolution of a terrain, a section on basic petrology finds its place in a text book on geodynamics.

Classic phase petrology is based on the principle of thermodynamic equilibrium. In contrast, most of the geodynamic processes discussed in this book are inherently out of equilibrium, for example the thermal evolution of orogens. The large success that the equilibrium thermodynamics approach has had in determining metamorphic conditions largely relies on the Arrhenius relationship which we will discuss in some detail on p. 344. However, here we want to begin by stating that petrology may be seen to involve:

- equilibrium considerations,
- non-equilibrium considerations.

We will deal with these two fundamentally different approaches in sect. 7.2.1 and sect. 7.2.2, respectively. Throughout this chapter we remind the reader that all topics are dealt with in a very sweeping way and refer to the wide range of excellent teaching texts in petrology (e. g. Winter 2001; Spear 1993; Anderson and Crerar 1993; Yardley 1989; Powell 1978).

**Chemical Reaction.** Microstructural and chemical changes in rocks depend on two factors:

1. The rate of diffusion with which atoms are brought from one point of the rock to another.
2. The rate of reaction or nucleation which actually binds the atoms at the new location structurally into the crystal lattice.

These two processes depend on very different parameters. The rate of *diffusion* depends very strongly on temperature (in a manner described by eq. 7.5). The rate of *nucleation* and *reaction* is independent of temperature, but depends on how far the reaction is overstepped. The slower of the two processes will be the bottleneck for the overall process and will determine the process rate (Fischer 1973; Joesten 1977; Putnis and McConnell 1980). Petrologists refer to *diffusion controlled* processes and *reaction controlled* processes. At geologically relevant temperatures the rate determining factor is usually diffusion and we discuss chemical diffusion therefore in the next paragraph.

- *Chemical diffusion.* In the most simple case, one-dimensional diffusion of a single element through a crystal lattice may be described by:

$$\frac{\partial C}{\partial t} = D_{(T)} \frac{\partial^2 C}{\partial x^2} . \quad (7.3)$$

There,  $C$  is the concentration of a given element in a given mineral,  $t$  is time and  $x$  is a spatial coordinate, for example the distance from the center of a garnet crystal to its surface.  $D_{(T)}$  is the cation diffusivity and we note already

here that it is *not* a constant (as it was in eq. 3.6), but a strong function of temperature. Other than that, element diffusion in minerals is completely analogous to the diffusion of heat (sect. 3.1, eq. 3.6), or the diffusion of mass on a larger scale (e.g. sect. 4.5.2) (Stiwe 1998b). As for those, analytical solutions of eq. 7.3 may be found for a large range of initial and boundary conditions in Carslaw and Jaeger (1959). The specifics of element diffusion is described in detail by the excellent text book of Crank (1975).

The application of eq. 7.3 is not only hindered by the temperature dependence of the diffusivity. It is also complicated by that fact that diffusion rates depend on the diffusion path. For example, diffusion of atoms along grain boundaries is orders of magnitude faster than volume diffusion through the crystal lattice. Thus, volume diffusion is the process rate limiting factor. The rate of volume diffusion of chemical elements through a crystal lattice itself depends on many factors and is – akin to the determination of creep constants for power law creep – difficult to determine experimentally. Also, if several elements diffuse at the same time in the same crystal, then they may influence each others diffusion rates (Onsager 1931).

Nevertheless, eq. 7.3 may be used to make some important estimates on the degree to which chemical equilibration of a rock has gone to completion, for example by using the same argument we have used on p. 62 to estimate the time scale for the diffusion of heat (sect. 3.1.3 and 3.1.4, eq. 3.17). In analogy to there we can write here:

$$t_{eq} \approx \frac{l^2}{D} . \quad (7.4)$$

There,  $l$  is the length scale of diffusion, which a measure of the spatial distance over which elements diffuse in time  $t_{eq}$ . Eq. 7.4 may be used to estimate the time it takes for a garnet crystal of a given size to equilibrate at a given temperature, or vice versa. For example, we will show below that the diffusivities of iron through a garnet lattice at 400 °C is about:  $D_{400} \approx 2.7 \cdot 10^{-27} \text{ m}^2 \text{ s}^{-1}$  and at 800 °C it is about  $D_{800} \approx 2.3 \cdot 10^{-20} \text{ m}^2 \text{ s}^{-1}$ . Using eq. 7.4 we can estimate that – if a metamorphic event of 800 °C temperature lasted of the order of 10 my – only garnets smaller than 2.7 mm can be equilibrate by volume diffusion in this time. Correspondingly, we can judge that garnet crystals of 2 mm diameter must be heated for at least 5 my to 800 °C or at least for  $4.7 \cdot 10^7$  my to 400 °C, so that they can equilibrate. If we want to learn about mean diffusion times over a period of temperature change, then we need to use mean diffusivities that we can estimate using eq. 7.9.

**Why does Equilibrium Work?** The fantastic success of equilibrium thermodynamic considerations in petrology may be largely accredited to the Arrhenius relationship. The Arrhenius relationship states that diffusion processes are an exponential function of temperature:

$$D_{(T)} = D_{0e} e^{-\frac{Q+VP}{RT}} . \quad (7.5)$$

In this equation,  $D_{(T)}$  is the diffusivity of elements as a function of temperature (in  $\text{m}^2 \text{ s}^{-1}$ ),  $D_0$  is a pre exponential factor,  $Q$  is the activation energy,  $R$  is the gas constant and  $T$  is absolute temperature. The numerator of the exponent also includes the product of the activation volume  $V$  and pressure  $P$ . However, the activation volumes of petrological processes are generally so small, so that the pressure dependence of element diffusion is generally neglected. According to eq. 7.5 the diffusivity goes towards zero when temperature goes towards zero. In other words, all equilibria that rely on diffusion freeze at low temperatures. At higher temperatures, the diffusivity rises and goes asymptotically towards  $D_0$ . Thus, with higher temperatures, it becomes increasingly likely that different minerals in a rock can communicate with each other and are in chemical or structural equilibrium. The Arrhenius relationship applies not only to the diffusion of ions on a crystal lattice but also to the diffusion of radioactive isotopes or lattice dislocations. The former is highly relevant to geochronologists, the latter is relevant when discussing flow processes and microstructures. In this context we have met the Arrhenius equation already when we discussed deformation mechanisms (sect. 5.1.2).

**The Irreversibility Principle.** Because of the temperature dependence of diffusion processes (eq. 7.5), only small parts of a metamorphic  $P$ - $T$  path are actually preserved in a rock. According to which parts of a  $P$ - $T$  path are preserved in a rock and which parts are not, the metamorphic evolution of rocks may be divided into 5 different sections (Fig. 7.3). In the 1st section there is no chemical reaction or mineral growth. Temperatures and diffusion rates are too slow. In the second part chemical equilibration does occur. However, because temperature and diffusion rates continue to increase, each equilibrium will be erased and superseded by that of the next higher temperature. The 3rd section of the  $P$ - $T$ -path is reached at the metamorphic temperature peak where - by definition - the rate of temperature change is small. There, chemical and textural equilibrium is reached to the best degree. The 4th part of the  $P$ - $T$  path occurs during the initial cooling history. Diffusion rates decrease and the volume of the rock which can be in equilibrium at any given temperature decreases. Thus, larger and larger parts of the rock will cease to equilibrate with their surroundings and preserve larger and larger parts of the rocks. Partial retrogression will occur in this 4th section. In the 5th section of the  $P$ - $T$  path all equilibria are frozen and reaction has terminated.

Fig. 7.3a summarizes this information. It shows that only the thermal peak and a small section of the path thereafter are likely to be preserved by rocks, while the heating phases is likely to be only preserved in relics. Because temperature is so much more important to the equilibration process than pressure (see discussion of eq. 7.5) the *thermal* maximum (rather than the *baric* maximum) is generally referred to as the metamorphic peak. However, it is emphasized that pressure and temperature peak of a  $P$ - $T$  path must not coincide (Fig. 7.3b). Indeed, valuable tectonic information may be extracted

from the temporal relationship between pressure and temperature maximum on a metamorphic  $P$ - $T$  path.

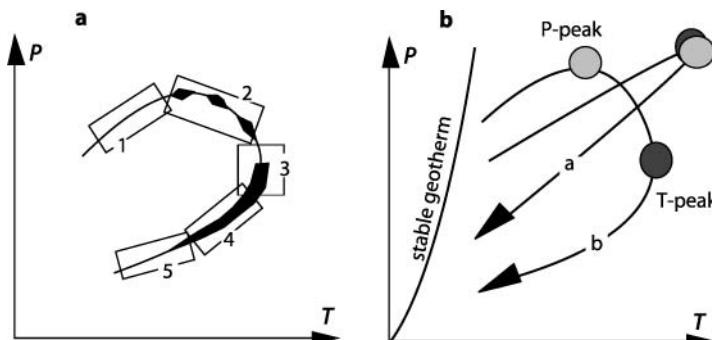
### 7.2.1 Equilibrium Information: Thermobarometry

If we assume that rocks do, in fact, record some sort of chemical equilibrium from some stage of their metamorphic evolution it is possible to use the principles of equilibrium thermodynamics to infer their formation conditions. Very often this concerns the derivation of metamorphic pressures and temperatures using geothermometers and geobarometers. In general, the term *geothermobarometry* summarizes the methods used. It is useful to discriminate between:

- petrological thermobarometry,
- mineralogical - crystallographic thermobarometry and:
- structural thermobarometry.

An example for *structural barometers* are the palaeopiezometric methods (e. g. Christie and Ord 1980; Dunning et al. 1982) and an example of *structural thermometry* is textural analysis like lattice preferred orientations of quartz (e. g. Jessel and Lister 1990). *Mineralogical thermobarometry* uses the pressure and temperature dependence of parameters that may be measured in a single mineral, for example the lattice parameters, the concentration of a given element or the composition of fluid inclusions.

*Petrological thermobarometry* is based on the fact that the distribution of elements between minerals is a function of pressure and temperature. For energetic reasons, this distribution always aims to be in chemical equilibrium.



**Figure 7.3.** **a** Schematic  $P$ - $T$ -path that is divided into five different sections that are recorded in different detail in a rock. The different sections are discussed in the text. The sections of a  $P$ - $T$ -path that are well recorded by a rock are shown thick **b**  $P$ - $T$ -path in which pressure and temperature peak coincide (*a*); and  $P$ - $T$ -path in which pressure and temperature peak do *not* coincide (*b*). A schematic stable geotherm is shown for reference

In particular with high grade metamorphic rocks, petrological thermobarometry has become the standard method to determine the physical conditions of metamorphism. However, for a meaningful application it is crucial that chemical equilibrium was reached and that reaction textures between different minerals are interpreted correctly (Cooke et al. 2000). About the methods and problems of textural interpretation of mineral parageneses there is abundant literature (e. g. Spear and Florence 1992; Robinson 1990).

**The Chemical Equilibrium Condition.** Most of the *petrological thermobarometry* is based on the assumption of chemical equilibrium in rocks. Then, the thermodynamic equations describing equilibrium can be used to predict the distribution of a chemical element between the different minerals in a rock. The *thermodynamic equilibrium condition* may be formulated as:

$$0 = \Delta G^0 + RT \ln K \quad . \quad (7.6)$$

In this equation,  $R$  is the gas constant,  $\Delta G^0$  is the change in Gibb's free energy between all phases involved in a reaction in their standard states and  $K$  is the equilibrium coefficient. This coefficient may be derived from the activities of different mineral end members in a mineral, which is related to their composition.  $T$  is the absolute temperature. For a simple exchange reaction of the element C between the phases A and B,  $K$  is given by the relationship:

$$K = \left( \frac{1 - X_C^A}{X_C^A} \right) \left( \frac{X_C^B}{1 - X_C^B} \right) \quad . \quad (7.7)$$

There,  $X_C^A$  is the mol fraction of the element C in mineral A and  $X_C^B$  is the corresponding mol fraction in mineral phase B. Thus,  $K$  may be measured directly from the compositions of different minerals in a rock, for example by electron microprobe. The expression  $\Delta G^0$  is the difference between the free enthalpies of the reactants and the products (in this case the pure phases A and B) in their standard state. These free enthalpies are functions of material constants like heat capacity, formation energy, pressure and temperature. Thus, eq. 7.6 may be solved for pressure as a function of temperature or vice versa, if all the other constants are known. If eq. 7.6 describes a curve with a shallow slope in a pressure-temperature diagram, then this reaction is called a *barometer*. Correspondingly, if the slope of this curve is steep, then it is called a *thermometer*. Clearly, the information given in this short paragraph is by no means sufficient to learn about the thermodynamics of thermobarometry. It is just meant to indicate that an entry into this field is not all this hard. For details see an abundance of literature on the thermodynamics of rocks, for example Holland and Powell (1990), Atkins (1994), Powell (1978), Spear (1993) or Will (1998).

**The Phase Rule.** The concept of the phase rule is a much undervalued tool that can also be used by the field geologist. This section shows how. The phase rule relates the number of degrees of thermodynamic freedom to the number of chemical components and the number of mineral phases (Zen 1966). It may be formulated as:

$$\text{Phases} = \text{Components} - \text{Degrees of Freedom} + 2 \quad . \quad (7.8)$$

For example, a reaction involving *one* chemical component and *two* mineral phases will have *one* degree of thermodynamic freedom. We may freely choose the pressure where this reaction is supposed to occur, but then the temperature and compositions of all phases are predetermined by this assumption - they are fixed. The reaction between ice and water is a good example of a univariant (one degree of freedom) reaction between 2 phases in the 1 component system  $\text{H}_2\text{O}$ . Similarly, the reaction between periclase and quartz to form enstatite is a univariant reaction between 3 phases in a 2 component system  $\text{MgO} - \text{SiO}_2$ . We can also see that - in a 6 component system - it is impossible for more than 8 phases to exist in stable equilibrium. In fact, it is extremely unlikely that rocks equilibrate on invariant points or even along univariant reactions. This information can help us enormously:

- ... *by considering observed minerals as phases.* Then, the phase rule can be used to infer two overprinting events. For example, pelitic rocks typically contain 6 chemical components:  $\text{K}_2\text{O}$ ,  $\text{FeO}$ ,  $\text{MgO}$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{SiO}_2$  and  $\text{H}_2\text{O}$ . If a pelitic rock contains 8 minerals, for example garnet, staurolite, biotite, muscovite, kyanite, chlorite, quartz and water, then we must conclude that it is invariant. As this unlikely (or impossible if there is even more minerals), we can conclude that one or several of the observed phases grew later or earlier than the rest and we are observing at least two different equilibrium assemblages. As such, the phase rule can aid us to infer overprinting events when doing petrography. Alternatively, it may tell us that there are more chemical components in the system, for example,  $\text{CaO}$  or  $\text{MnO}$ .

- ... *by considering end members as phases.* Then, the phase rule may be used to estimate the number of possible thermobarometers that we can use to constrain our paragenesis. If the parageneses above contains only 6 components and is in fact invariant then there is at least 10 end members with appreciable activities in it: There are pyrope and almandine in the garnet, phlogopite and annite in biotite and so on. Since we need 7 end members to formulate a univariant reaction for a thermobarometer equation (eq. 7.6), 10 end members allows us to formulate at least 4 independent thermobarometers. In short, the phase rule can help us to estimate how easy it will be to constrain our formation conditions. Using several thermobarometers simultaneously is the principle underlying the idea of formulating *internally consistent thermodynamic data sets* that can then be used to do average calculations between many thermobarometers (e.g. Holland and Powell 1990;

Berman 1988). Because of these reasons it is generally useful to collect in the field rocks that contain as many phases as possible for as few minerals as possible.

### 7.2.2 Non-equilibrium Information: Kinetics

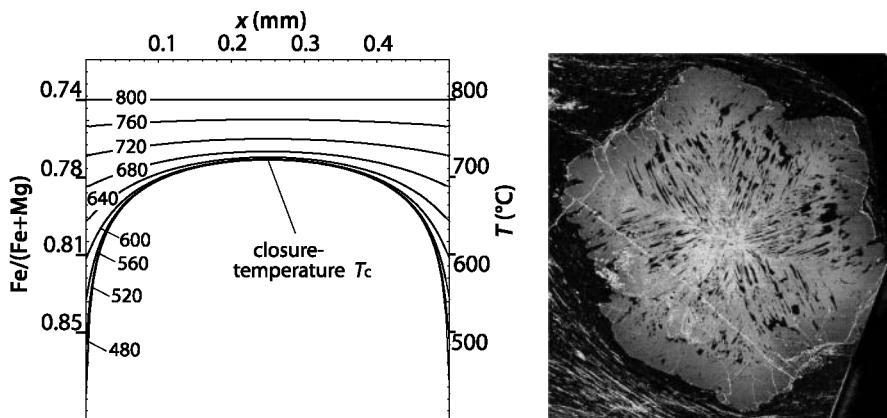
While equilibrium thermodynamic methods are the most widely used tool by petrologists, much additional information can be gained by using the non-equilibrium information recorded by rocks. Very loosely, it may be said that this is done using two different approaches:

- using *non-equilibrium* thermodynamics,
- using *equilibrium* thermodynamics and “being careful”.

The use of non-equilibrium thermodynamics in petrology was pioneered in the seventies (e.g. Fischer 1973; Joesten 1977), but – while undoubtedly a rigorous treatment of the subject – it has not found wide application by the field geology community. In contrast, the application of equilibrium considerations to the interpretation of metamorphic disequilibria (e.g. partial reaction textures) by “being careful” is widely applied by geologists to the interpretation of metamorphic  $P$ - $T$  paths. In particular the consideration of a time scale in diffusion processes has led to a lot of progress which may be summarized under the term “geospeedometry” (Lasaga 1983). The following section presents a few highlights from this field.

**The Concept of Closure Temperature.** Because of the strong temperature dependence of diffusivities (eq. 7.5) diffusing systems are subject to a closure temperature. We illustrate this with the rate of volume diffusion of cations in garnet crystals. The magnitude of the diffusion rate of magnesium in garnet is roughly known. It is described by the material constants  $Q = 239\,000 \text{ J mol}^{-1}$  and  $D_0 = 9.81 \cdot 10^{-9} \text{ m}^2 \text{ s}^{-1}$  (Cygan and Lasaga 1985). Using eq. 7.5 we can derive a diffusivity of  $D_{400} \approx 2.7 \cdot 10^{-27} \text{ m}^2 \text{ s}^{-1}$  or  $D_{800} \approx 2.3 \cdot 10^{-20} \text{ m}^2 \text{ s}^{-1}$  at temperatures of  $400^\circ\text{C}$  or  $800^\circ\text{C}$ , respectively. We can see that the diffusivity changes over many orders of magnitude over this geologically relevant temperature interval.

From the above and eq. 7.4 we can see that a crystal of a given size can *only* equilibrate in a given time if its  $P$ - $T$ -path exceeds a certain temperature over a certain time. If the temperature decreases below this temperature, then full equilibration is not any more possible and the crystal will only equilibrate partially. The diffusive length scale (given by  $l$  in eq. 7.4) gets smaller than the radius of a crystal and at least part of the crystal center will not be able to keep up with the processes on the crystal surface. It will cease to react with the surroundings and its composition will freeze. This phenomenon is called *closure* and the temperature at which this occurs is called the *closure temperature*. The successive closure of the chemical composition of a crystal leads to the development of a zoning profile and can be observed in the minerals of



**Figure 7.4.** Modeled and observed concentric zoning pattern in garnet. The modeled profile illustrates the definition of closure temperature  $T_c$ . The horizontal axis is a profile through a garnet crystal that is 0.5 mm in diameter. The grain center is located at  $x = 0.25$  mm. The left hand vertical axis is labeled with the Fe concentration in the crystal (given as mol fraction  $\text{Fe} / (\text{Fe} + \text{Mg})$ ). In equilibrium this concentrations may be directly correlated with temperature (right hand vertical axis). The surface of the crystal is assumed to be in chemical equilibrium with a much larger biotite crystal characterized by an Fe concentration of  $\text{Fe} / (\text{Fe} + \text{Mg}) = 0.5$  at all times. A cooling rate of  $100^\circ\text{C my}^{-1}$  was assumed. Calculated by numerical coupling of eqs. 7.3, 7.5 and 7.7 and the garnet - biotite thermometer equation of Holland and Powell (1990). The shown garnet crystal is from the Mt Isa Inlier of Central Australia (Image by M. Hand)

many rocks (Fig. 7.4). The chemical non-equilibrium that is frozen into such zoning profiles is an important piece of evidence for the interpretation of the cooling history of a rock.

**Mean Diffusivities.** Even if a mineral grain is open to diffusive equilibration (i.e. it is above its closure temperature) diffusivities are very strong functions of temperature so that it is difficult to use eq. 7.4 to estimate diffusive length scales without having found an appropriate mean diffusivity that is representative for the diffusion rate over a given interval of temperature change. As the Arrhenius relationship is strongly non linear it is not a trivial exercise to estimate such a mean diffusivity. However, if the cooling rate within the temperature interval of consideration is constant, then the mean diffusion rate may be estimated with the following relationship (Itaya and Stüwe 1974):

$$\overline{D} \approx \frac{D_A}{\left(\frac{Q}{RT_E}\right)(1 - T_E/T_A)} . \quad (7.9)$$

There,  $\overline{D}$  is the mean diffusion rate averaged between the starting temperature  $T_A$  and the final temperature  $T_E$  (e.g. the metamorphic peak temperature and the temperature of the earth's surface).  $D_A$  is the diffusion rate at the starting temperature and may be calculated from eq. 7.5.

**Determining Cooling Rates.** From the above we can summarize, that the following parameters have a close linking in diffusive processes:

- Diffusive length scale (grain size),
- cooling rate (temperature, time),
- diffusivity (material constants).

For simple exchange systems and simple geometries of the crystal eq. 7.3 may be solved analytically to describe this relationship:

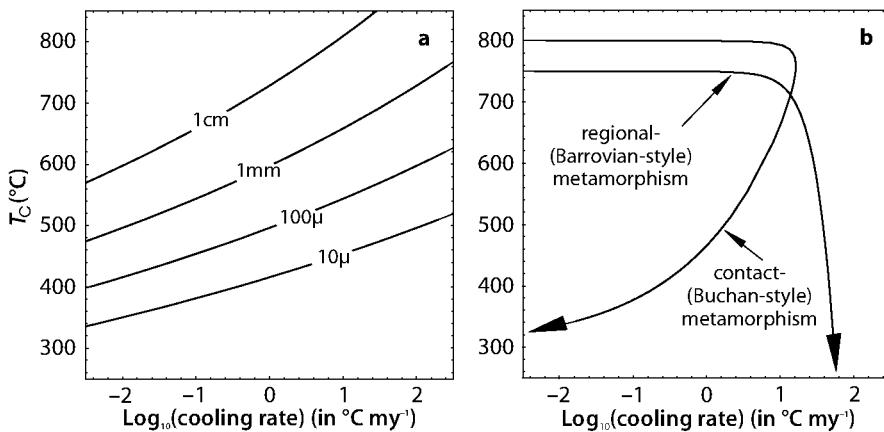
$$s = D_0 e^{-\frac{Q}{RT_c} + G} \left( \frac{RT_c^2}{(l/2)^2 Q} \right) . \quad (7.10)$$

There,  $s$  is the cooling rate,  $T_c$  is the closure temperature, (defined in Fig. 7.4) and  $l$  is the radius of a crystal. The constant  $G$  is a parameter that describes the geometry of the crystal in which diffusion occurs. For spherical crystals (as is well-approximated by garnet) this constant is:  $G = 1.96$ . The variables  $Q$ ,  $D_0$  and  $R$  were defined in eq. 7.5. Eq. 7.10 was derived by Dodson (1973) for application in geochronological systems but it is equally applicable to cation diffusion in minerals where it has become an important tool to estimate cooling rates of rocks using data that are easily obtained using optical microscopy (to determine  $l$ ) and electron microprobe (to determine  $T_c$ ) (Ehlers and Powell 1994; Ehlers et al. 1994a). The relationship between grain size, closure temperature and cooling rate defined by eq. 7.10 is illustrated in Fig. 7.5a.

- *Tectonic use of cooling rate information.* The slope and curvature of cooling curves is characteristic of the underlying cooling process. Thus, determination of cooling curves is an important tool for the interpretation of geodynamic processes. For example, a comparison of Fig. 3.15 with 3.33 (as contrasted in Fig. 7.5b) shows, that cooling curves terminating a regional metamorphic cycle are fundamentally different from those that characterize cooling after contact metamorphism (Harrison and Clark 1979). If cooling occurs because the rocks are exhumed, then the cooling rate *increases* with decreasing temperature. In contrast, the cooling rate *decreases* with decreasing temperature towards the end of contact metamorphism (see Fig. 7.5b).

Fig. 7.5a illustrates the different relationships between closure temperatures and grain size that are expected in two rocks that cooled according to the two cooling curves shown in Fig. 7.5b. Overlaying Fig. 7.5a onto Fig. 7.5b shows that these relationships are different! Ehlers et al. (1994b) have documented closure temperatures between 400 °C and 700 °C in a single thin section and were able to correlate these closure temperatures with grain size. This range of closure temperatures is sufficient to document crucial parts of a cooling curve.

**Geochronology.** Many geochronological methods also rely on the concept of closure temperature. At high temperatures, diffusivities of radiogenically produced isotopes are large enough so that they do not accumulate in a



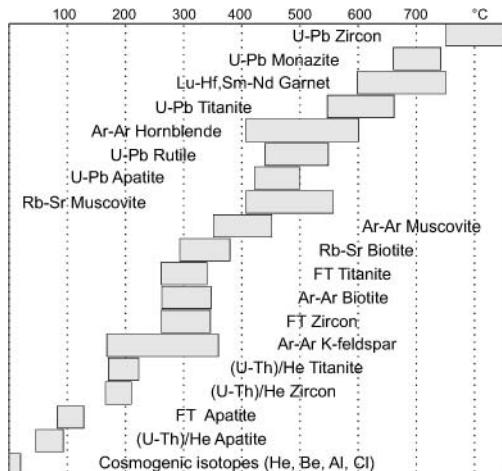
**Figure 7.5.** Relationships between cooling rate and temperature. **a** Cooling rate plotted against closure temperature  $T_c$  of crystals with different diameters. The curves were calculated with eq. 7.10 assuming a spherical grain geometry. The diffusion data of Cygan and Lasaga (1985) for garnet were used. **b** Cooling rate against temperature for two different cooling processes. The curve for contact metamorphism was calculated with eq. 3.88 using  $T_i = 800\text{ }^{\circ}\text{C}$ ,  $T_b = 300\text{ }^{\circ}\text{C}$ ,  $l = 60\text{ km}$ ,  $\kappa = 10^{-6}\text{ m}^2\text{ s}^{-1}$  and  $z = 0$ . The curve for regional metamorphism was calculated with eq. 3.48 and eq. 3.49 using  $g = 30\text{ }^{\circ}\text{C km}^{-1}$ ,  $z_i = 25\text{ km}$  and an erosion rate of  $u = -1\text{ km my}^{-1}$ . While the assumptions underlying these equations are very much simplified from real settings, they do show some typical characteristics. For example, note that during cooling from regional metamorphism the cooling rate *increases* with decreasing temperature, while the cooling rate *decreases* with decreasing temperature following contact metamorphism. Such differences may be extracted, at least in principle, from zoned crystals

crystal lattice. Below the closure temperature the diffusivities are low and the isotopes accumulate. Thus, the content of daughter isotopes in a crystal is proportional to the time it has spent below the closure temperature and the content of the parent isotope. If the latter is known, the former can be inferred (Fig. 7.6). However, as we have learned above, closure temperatures are dependent on cooling rates and many other factors, so that dating with isotopes has a large range of problems that geochronologists battle with.

A common method to document cooling curves uses isotopic dating of a series of isotopic systems in minerals with different closure temperatures. Closure temperature of the various isotopic systems is then plotted against the radiometrically determined age and a cooling curve is constructed. However, in the previous section we have shown that the closure temperature depends on cooling rate. Thus there is an obvious circular argument: Cooling rates obtained by this approach can – strictly taken – only be estimated once a cooling rates is assumed. Only then the closure temperature of the system is known.

The successful application of geochronological methods to the determination of cooling curves indicates that the variability of closure temperatures of isotopic systems is small – even for a large range of cooling rates – in comparison to many other factors influencing closure of isotopic systems during cooling. However, it is important to be aware of the implicit circularity in the uncritical acceptance of a certain value for closure temperature of an isotopic system.

**Figure 7.6.** Isotope systems commonly used for absolute geochronology and the temperatures they date. For most systems these temperatures are “closure temperatures”, where the diffusion of the radioactive element of interest ceases. For some, for example fission track (FT) analysis, it is a “track retention temperature”. For cosmogenic isotopes it is simply an accumulation time. (After Fitzgerald and Baldwin pers comm 2006).



## 7.3 Classification of *P-T*-Paths

*P-T*-paths may be divided into two groups, according to their slope following the metamorphic temperature peak:

- 1. *P-T*-paths which are characterized by a decrease in pressure before cooling commences.
- 2. *P-T*-paths which are characterized by a decrease in temperature before decompression commences and *P-T*-paths where pressure increase accompanies cooling (Fig. 7.7).

In a diagram in which the positive temperature axis is drawn to the right and the positive pressure axis is drawn upwards, these two paths follow a clockwise and an anticlockwise curvature, respectively (Fig. 7.7a). Thus, the two path groups 1. and 2. are referred to as clockwise *P-T*-paths and anticlockwise *P-T*-paths, respectively.

However, in the geological literature it is common to plot the pressure axis downwards to assist the intuitive understanding that pressure increases with depth in the crust (Fig. 7.7b). Obviously, the direction of a *P-T*-path in such a diagram is reversed from that in the conventional representation. Thus, it

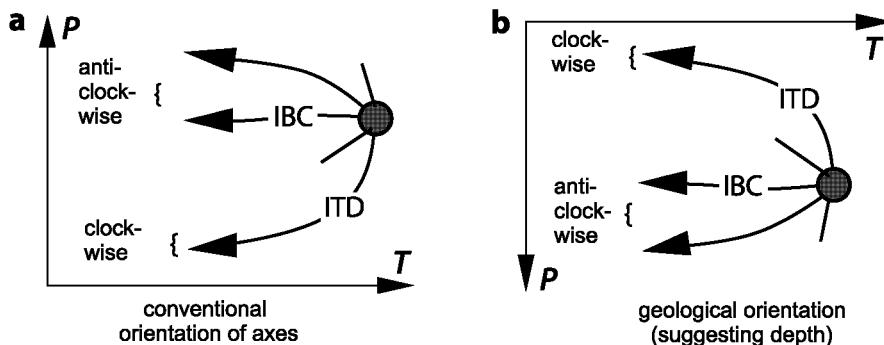
may be better to use the less ambiguous description of ITD- (IsoThermal Decompression) or IBC-paths (**IsoBaric Cooling**) instead of clockwise and anticlockwise.

This qualitative division of  $P$ - $T$ -paths may be formulated a bit more rigorously. At the thermal peak, the rate of temperature change is, by definition, zero:  $dT/dt=0$ . Thus, differently shaped  $P$ - $T$  paths may be defined on the basis of the qualitative nature of the pressure change at the temperature peak. We can write:

- $\frac{dP}{dt} \Big|_{(dT/dt=0)} = \text{negative:} \quad \text{clockwise (ITD)}$
- $\frac{dP}{dt} \Big|_{(dT/dt=0)} = 0: \quad \text{anticlockwise (IBC)}$
- $\frac{dP}{dt} \Big|_{(dT/dt=0)} = \text{positive:} \quad \text{anticlockwise}$

While this is quite a rigorous classification, this definition does not state how long this condition is held upright. For example, it is possible that the  $P$ - $T$ -path is characterized by pressure increase only at the very temperature peak, but that most of the time the path occurred during decompression (Fig. 7.8a). In short, it remains difficult to define a rigorous quantitative classification of  $P$ - $T$ -paths.

- *Continuous or discontinuous P-T-paths?*. The most common problem with interpreting the shape of P-T-paths is illustrated in Fig. 7.8b. There, two black dots indicate the P-T-conditions for the formation of two overprinting metamorphic parageneses in a single rock. Usually it is very difficult to see in a rock if the two parageneses formed in a continuous P-T-evolution during one



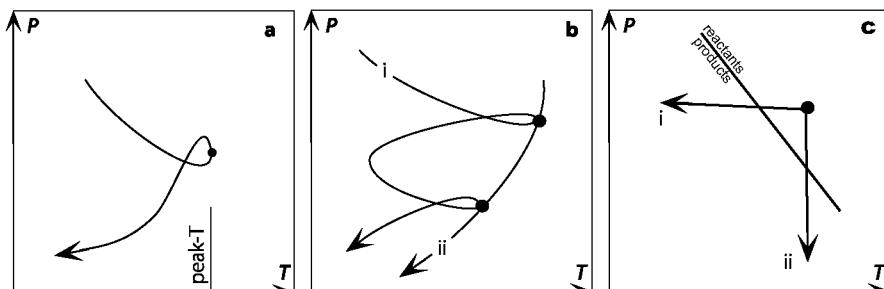
**Figure 7.7.** Directions of  $P$ - $T$ -paths in different presentations. The thick dot is the metamorphic peak. **a** In the conventional representation, isothermal decompression paths follow a clockwise path and isobaric cooling paths have an anticlockwise shape. **b** In the geologically most commonly used representation pressure is usually plotted to increase downwards to indicate that it increases with depth in the crust. Then, “anticlockwise paths” actually follow a clockwise path and vice versa.

orogenic event (ii), or as the consequence of two independent metamorphic events that may be hundreds of million years apart (i) (Fig. 7.8b).

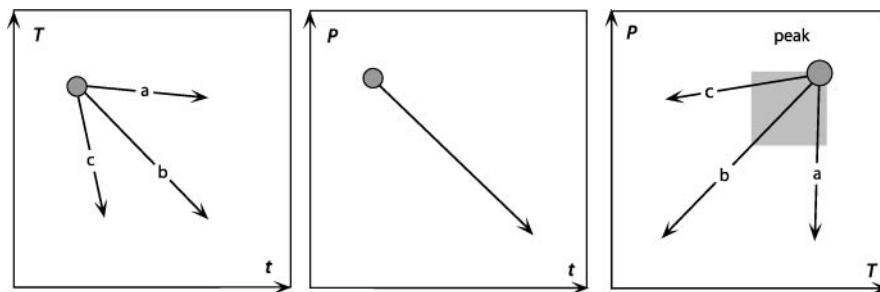
### 7.3.1 Slope and Curvature of $P$ - $T$ -Paths

The slope and curvature of metamorphic  $P$ - $T$ -paths contains important information on the relative and absolute rate of different geological processes. The two diagrams on the left in Fig. 7.9 show the temporal evolution of pressure a rock experienced (bottom diagram) and three different possibilities of the temporal evolution of temperature in the same rock. The diagram on the right shows three  $P$ - $T$ -paths that correspond to the temporal evolutions shown at left. The temporal evolution of neither pressure nor temperature may be inferred from the  $P$ - $T$ -path, but their *relative* temporal evolution is. This may be inferred from the slope of the three different  $P$ - $T$ -paths. From the IBC-path *c* we can infer that the rate of cooling was rapid compared to the rate of pressure change. Correspondingly, we can conclude from the ITD-path *a* that the rate of decompression was rapid compared to the rate of cooling. Note that we extracted this information solely on the basis of its  $P$ - $T$ -path.

If pressure and temperature are not linear functions of time, then the relationships discussed above need not apply. For example, if the temporal evolution of pressure and that of temperature are the same, then the  $P$ - $T$ -path will be linear, irregardless of *what* temporal evolution pressure and temperature undertake. This is shown on Fig. 7.10 where both the combination of temper-



**Figure 7.8.** Examples of  $P$ - $T$ -paths that are easily misinterpreted. **a**  $P$ - $T$ -path which – according to conventional classification – would be interpreted to be anti-clockwise, but which has the overall form of a ITD-path. **b** Example of a fundamental problem of the interpretation of  $P$ - $T$ -paths. If two metamorphic parageneses are observed in a rock that indicate the two  $P$ - $T$  conditions shown by the two black dots, then it is usually very difficult to discern if the path between the two was characterized by two different events with IBC-paths (shown by path *i*) or a single ITD path (shown by path *ii*). **c** Example of another problem of interpreting  $P$ - $T$ -paths: The observation of two overprinting parageneses and the knowledge of the reaction that produced them is often insufficient to discern between IBC and ITD paths.

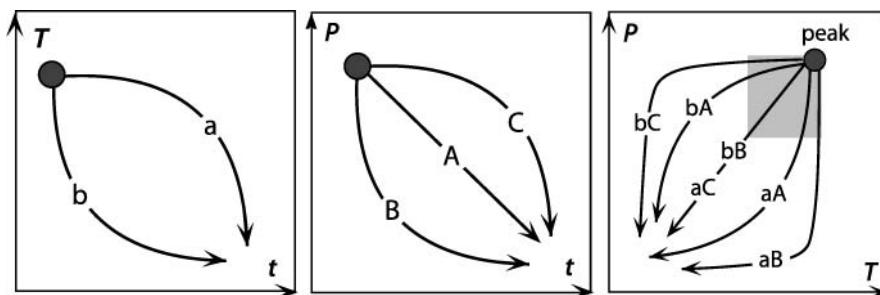


**Figure 7.9.** Different shapes of  $P$ - $T$ -paths (diagram on right), as they arise from different cooling rates of rocks (at left) even though all paths experienced the same decompression rate (at center). The  $P$ - $T$ -paths  $a$ ,  $b$  and  $c$  in the diagram on the right correspond to the three different cooling curves as labeled on the left. The shaded rectangle indicates schematically the  $P$ - $T$ -region which is likely to be preserved in a rock

ature evolution  $a$  with pressure evolution  $C$  and the temperature evolution  $b$  with pressure evolution  $B$  produce the same linear  $P$ - $T$ -path.

**Interpreting the Slope of a  $P$ - $T$ -Path.** Let us illustrate the enormous geodynamical relevance of the considerations of the last sections with an example. In many high grade metamorphic terrains in the Precambrian shields IBC-curves have been documented (e. g. Harley 1989). We can conclude that the cooling rate of all these terrains must have been much more rapid than their rate of exhumation or burial during the thermal evolution (assuming that all pressure change is related to depth change). As most of these terrains are characterized by intense synmetamorphic deformation (both extensional and compressional) it is likely that the terrains did indeed exhumed or get buried during their thermal history. Thus, if the rates of exhumation and burial during erosion or continental deformation were known, the we can use the fact that a terrain is characterized by an IBC paths to constrain minimum values for the rate of cooling. During normal continental deformation rates of  $\dot{\epsilon} = 10^{-13} - 10^{-14} \text{ s}^{-1}$  the rate of exhumation or burial is of the order of some millimeters per year. At this rate of depth change an IBC-path can only occur if the cooling phase of the terrain lasted a mere few millions of years. For example, if we just take an arbitrary number of 3 millimeters per year of depth change and assume that a  $P$ - $T$ -path will look like an IBC path if it cools more rapidly than  $300^\circ\text{C}$  per kilobar of pressure change, then this implies a cooling rate of at least  $243^\circ\text{C my}^{-1}$ .

If we now recall sect. 3.1.4 where we have shown that the duration of conductive processes is proportional to the square of the length scale of the conducting body, then we can use the estimate for the total duration of cooling to infer the spatial size of the terrain in question, at least to the order of magnitude. It turns out that for the total duration of the cooling phase as estimated above, the size of the heated region is a mere few kilometers. We can conclude from our derived IBC-path that it is unlikely that the metamor-



**Figure 7.10.** Possibilities for  $P$ - $T$ -paths as they may arise as a consequence of various non-linear evolutions of pressure and temperature. The figure is analogous to Fig. 7.9

phic event was caused by a process that affected the entire lithosphere. It is more likely that the heating mechanism was localized. Indeed, it is this very argument that has often been used to interpret that IBC-paths are atypical for regional metamorphic events and are more likely to be associated with contact metamorphism around magma bodies (e. g. Lux et al. 1986; DeYoreo et al. 1991; see. sect. 6.3.3). You may argue that all the numbers used here are wrong or not well constrained, but the thought process carried out here is independent of the numbers and in an example where they *are* known, it may extremely useful.

## 7.4 Interpretation of $P$ - $T$ - $t$ - $D$ -Relationships in Orogens

We will now go into some detail how to interpret spatial and temporal field relationships and  $P$ - $T$ - $t$ - $D$ -paths and refer the reader also to Passchier et al. (1990). Conceptual interpretations like those presented below are particularly important in ancient or badly exposed metamorphic terrains, where little is known about the geometry, style or tectonic setting of the underlying orogenic event. However, even in well known orogens like the European Alps, such conceptual interpretations may help geologists to free the eye from a huge abundance of detailed regional knowledge and help to understand some fundamental background of orogenic processes. In general, it is important to discriminate between two different types of relationships:

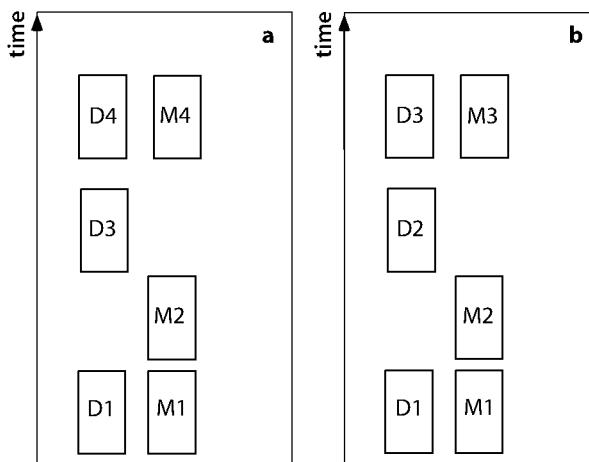
- 1. The *temporal* relationship between deformation and metamorphism.
- 2. The *spatial* change of the relationship between metamorphism and deformation across the terrain in question.

Such relationships, both in space and time, may be well illustrated on an event diagram (Fig. 7.11). When mapping such changes in the field it is important to be aware of the fact that different mapped parameters may have completely different strike directions in the field. For example, the lithological

boundaries may strike completely different from the direction of the metamorphic isograds or from lines of constant strain or those of constant age.

**Terminology Definition.** When discussing spatial and temporal relationships between metamorphic pressure, temperature and time across a terrain, or within a single rock, it is easy to confuse a number of quite different issues each of which contains important interpretative information. We therefore define the following terms:

- *Metamorphic field gradient.* The term “metamorphic field gradient” is used strictly to describe the change of metamorphic grade with distance as observed in the field (measured normal to the metamorphic isograds). As such, metamorphic field gradients have the units of  $dT/dx$  or  $dP/dx$  depending whether the gradient describes the change of metamorphic temperature or pressure with lateral distance  $x$ . The ratio of the metamorphic temperature gradient to the metamorphic pressure gradient does *not* need to document a metamorphic geotherm (although it also has the units of  $dT/dP$  and may be easily converted into  $dT/dz$  assuming the conversion discussed in eq. 7.1). In contrast, the  $T/P$  ratio of a single rock will *always* record one point on a metamorphic geotherm (by definition - see below).



**Figure 7.11.** “Event diagrams” are a useful way to record field observations. **a** and **b** show two different nomenclatures used for event diagrams for the same set of field observations. Both diagrams are drawn for a terrain in which three deformation phases ( $D$ ) and three metamorphic events ( $M$ ) have been mapped in the field. In **a** the different events were numbered according to their temporal sequence, but without consideration to whether the event is a deformation event or a metamorphic event. In **b** deformation events and metamorphic events are numbered separately. Both according to their respective temporal sequence, but with no consideration of the temporal relationship *between* deformation and metamorphic events. Which representation is more useful depends on the question that is to be answered with the illustration

- *Metamorphic geotherm.* The term “metamorphic geotherm” is used to describe the relationship between temperature and depth at a chosen point in time during metamorphism. As such, a metamorphic geotherm is a transient feature and contrasts the term *stable geotherm*. Very generally, metamorphic geotherms change at rates of  $\gg 10^\circ\text{C my}^{-1}$ , whereas stable geotherms change at rates of  $\ll 10^\circ\text{C my}^{-1}$ . During regional (Barrovian) metamorphism, metamorphic geotherms are generally characterized by a monotonous increase of temperature with depth (England and Thompson 1984). However, during contact metamorphism or in staked nappe piles, metamorphic geotherms may transiently be characterized by some crustal sections in which temperature decreases with increasing depth.
- *Piezothermal array.* The term “piezothermal array” was clearly defined by Richardson and England (1979) as the line that connects the metamorphic temperature peaks of all rocks in a vertical crustal section through depth and time. Piezothermal arrays are diachronous in time and they cross metamorphic geotherms. In general, it is a piezothermal array that will be preserved by an obliquely exposed crustal section, but piezotherms may coincide with metamorphic geotherms for some tectonic settings, for example if exhumation is practically instantaneous. Piezotherms are discussed in some more detail further below.

#### 7.4.1 Interpreting Temporal Relationships

The relative temporal relationship between deformation and metamorphism is usually well-preserved in metamorphic rocks (Fig. 7.11). There are three different possible relationships between one phase of deformation and one phase of metamorphism:

1. Deformation of rocks precedes metamorphism,
2. deformation and metamorphism occurred contemporaneously,
3. deformation occurred *after* metamorphism.

Clearly, it is always possible that the causes for deformation are unrelated to the causes of metamorphism. However, the timing relationship may also be the consequence of a causal coupling between the two. In the following, we discuss possible interpretations of these timing relationships in terms of a single underlying cause.

- *Deformation precedes metamorphism.* This timing relationship is characteristic for regional metamorphism, where metamorphism was caused by heat conduction and radiogenic heat production in overthickened continental crust (sect. 6.2.1). This is because continental deformation rates are typically about one order of magnitude more rapid than thermal equilibration over the length scale of the crust (s. sect. 3.1.4, 6.3.6). Thus, thickening of the crust at “normal” collision rates is usually completed before conductive equilibration may catch up. As a consequence, it may be expected that metamorphism is separated from peak metamorphism by up to several tens of millions of years.

• *Deformation occurred contemporaneously with metamorphism.* This relationship is typical of many low-pressure high-temperature metamorphic terrains. Three possible relationships may be thought of (s. also p. 331):

1. Deformation is caused directly by the heating. This is possible if the terrain in question is subjected to a constant plate boundary driving force. Then, heating of the terrain may weaken the rocks sufficiently so that strain rate rises rapidly in response to the far field stress (Sandiford et al. 1991).
2. Heating is caused directly and instantaneously by deformation. This is only possible if the rocks are strong enough so that shear heating causes the metamorphism. This is widely-observed on a local scale, for example around pseudotachylites, but the possibility for regional shear heating is subject to debate.
3. Both heating and deformation may have been caused by advection of hot material towards shallower depths. Such processes are certainly important when stacking nappes or when exhuming core complexes, but even whole sale convection has been suggested to operate in the crust (Weinberg 1997).

• *Deformation occurs after metamorphism.* Field examples where deformation of a terrain occurred after its metamorphic peak, are typically interpreted in terms of two independent events. That is, deformation and metamorphism are unrelated. Unlike the other two timing relationships discussed above, there is currently no elegant models which allow us to interpret this timing relationship in terms of a single tectonic process.

#### 7.4.2 Interpreting Spatial Relationships

In many metamorphic terrains the metamorphic grade changes across the terrain. The change in grade is often accompanied by a change of the timing of peak metamorphism. For example, high grade parts may experience their metamorphic peak earlier or later than lower grade parts, either relative to an absolute marker (e.g. a dike swarm), or as determined by absolute geochronology. Such spatial changes may be interpreted in three different ways:

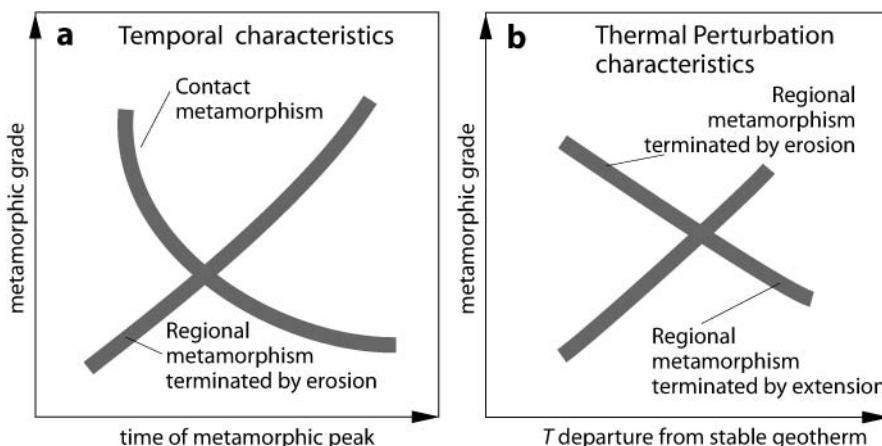
1. Independent heating mechanisms heated different parts of the terrain in question at different times. In this case, it is expected that there is a discontinuity in metamorphic grade or metamorphic age somewhere in the terrain; for example, because the rocks of different metamorphic age were juxtaposed much later.
2. Lateral variations in the physical parameters, for example, thermal conductivity or heat production rate may cause different parts of a terrain to heat at different rates and to different degrees (Sonder and Chamberlain 1992).

3. The changes in the timing of metamorphism across the terrain are inherent to the nature of a single heating process. A typical example for this is heat conduction: because of the time it takes for a terrain to equilibrate conductively, it is possible that some parts of a terrain experienced metamorphism earlier than others (e. g. s. sect. 3.1.4). In this case, changes in metamorphic timing and grade are expected to be continuous across a terrain.

The curve that must be interpreted to understand some of these relationships is the *piezothermal array*.

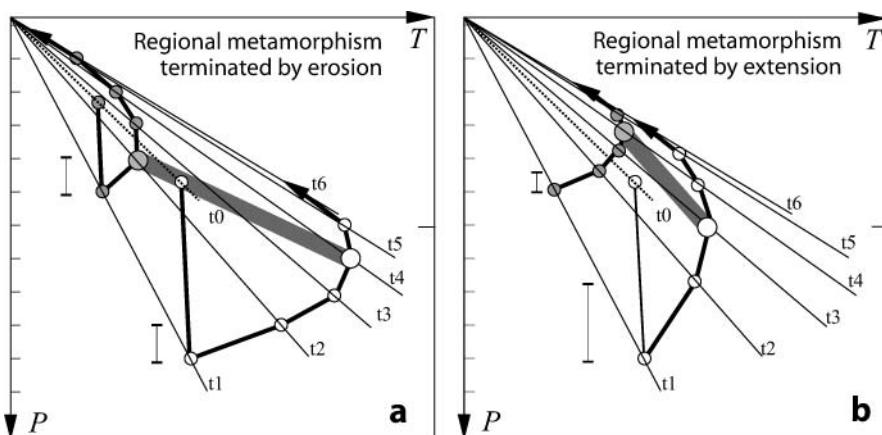
**Piezothermal Arrays.** Piezothermal arrays are curves that connect the metamorphic temperature peaks of rocks from all crustal depths through grade and time (England and Thompson 1984; Fig. 7.12). Accordingly, piezotherms may be plotted in pressure-temperature, in temperature-time, or in pressure-time diagrams. As the slopes of the piezotherms in two of these different diagrams are independent, it is useful to discriminate between two different features of a piezotherm. These are: 1. their temporal characteristics and 2. their thermal perturbation characteristics.

- *Time characteristics of piezotherms.* The temporal characteristics of a piezotherm describe if the metamorphic peak of high grade rocks occurs earlier or later than that of their lower grade equivalents (Fig. 7.12a). These timing relationships may be characterized by the slope of the piezotherm in a temperature-time diagram (e. g. Stüwe et al. 1993; Stüwe 1998c). Three qualitatively different relationships may be discerned:



**Figure 7.12.** Schematic cartoons illustrating **a** the temporal characteristics and **b** the thermal perturbation characteristics of piezothermal arrays for some different tectonic settings and metamorphic heat sources. As the slope of the piezotherms in these diagrams are mapable in the field, these functions are extremely important to consider when interpreting metamorphic terrains

1. If the metamorphic peak in high grade rocks occurred later than in low grade rocks, then the slope of the piezothermal array in a temperature-time diagram is positive. Piezothermal arrays with a positive slope are typical of regional metamorphism (sect. 6.2.1, Fig. 7.12a). Rocks from shallow crustal levels will not heat significantly because of their proximity to the surface and will be the first to cool when exhumation commences (Fig. 6.17). Rocks from deeper crustal levels have more time to equilibrate and will reach their metamorphic peak therefore later (Fig. 7.13a).
2. If the metamorphic peak in high grade rocks occurred earlier than in low grade rocks, then the slope of the piezothermal array in a temperature-time diagram is negative. Piezothermal arrays with a negative slope are typical of contact metamorphism (Fig. 3.33, 7.12). Near the contact of a heat source, rocks are heated very rapidly to high temperatures. With increasing distance from the heat source, the thermal effects of contact metamorphism are felt later and the metamorphic grade is lower (for detail s. sect. 3.6.2).
3. The situation that the time of metamorphism is independent of metamorphic grade is unusual. More likely, the temporal variation is too small to be discernable in the field. Nevertheless, one mechanism that *would* allow the

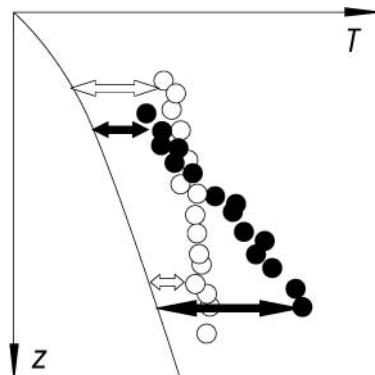


**Figure 7.13.** Cartoons illustrating the thermal evolution of thickened continental crust during: **a** exhumation by erosion (all rocks in a vertical column move upwards by the same amount during each time step as indicated by the vertical bars) and **b** exhumation by pure shear extension (the exhumation rate is depth dependent). Shown are: stable geotherms prior to thickening (dotted lines), *metamorphic geotherms* (thin lines labeled from time  $t_1$  immediately after thickening to  $t_6$  near the final stages of exhumation), *piezothermal arrays* (shaded bars) and *PTt* paths (thick lines). Note that in **a** the slope of the piezothermal array has a higher  $T/P$  gradient than any one geotherm but in **b** it has a lower  $T/P$  gradient. Moreover, the temporal relationships between high and low grade rocks are reversed between the two exhumation mechanisms

metamorphic peak to occur at the same time in rocks of all metamorphic grades, is if metamorphism is caused by mechanical heat production in a region with a strong strain gradient.

While the slope of the piezothermal array contains a lot of information on the underlying thermal event, the age difference between metamorphism of the highest grade rocks and that of the lowest grade rocks is often too small to be resolved with geochronological methods. Thus, we are often limited to interpret timing relationships from detailed observations in the field using structural observations.

**Figure 7.14.**  $T$ - $z$  diagram showing two hypothetical data sets collected from metamorphic terrains with a field gradient (large dots). The data set shown by the black dots has positive thermal perturbation characteristics: low grade rocks are less thermally perturbed from stable metamorphic conditions than high grade rocks. For the data set shown by the white dots this relationship is reversed: high grade rocks are less thermally perturbed than low grade rocks. This latter relationship is actually very common in many terrains around the world (e.g. Tenczer and Stüwe 2003). The information of this figure is equivalent to that shown in Fig. 7.12b



- *Thermal perturbation characteristics of piezotherms.* The thermal perturbation characteristics of a piezotherm describe if high grade rocks are more or less thermally perturbed from stable geothermal conditions than their lower grade equivalents. This may be characterized by the slope of the piezotherm in a  $PT$  diagram. For example, during regional metamorphism according to the classic description of England and Thompson (1984), high grade rocks may experience their peak at large depths and substantially derived from stable geothermal conditions (Figs. 7.12, 7.13a, black dots on Fig. 7.14). In contrast, many metamorphic terrains show reversed relationships between grade and magnitude of thermal perturbation (white dots on Fig. 7.14). For example, in the eclogite type locality in the eastern Alps, eclogites equilibrate at  $650^{\circ}\text{C}$  and almost 20 kbar, (which is not very perturbed as it lies near a “normal” geothermal gradient around  $9^{\circ}\text{C km}^{-1}$ ), but amphibolite facies rocks further north equilibrated at  $650^{\circ}\text{C}$  and 6 kbar (which is very warm for that pressure, s. Fig. 7.1). Thus, the amphibolite facies rocks are substantially more thermally perturbed than their eclogite facies equivalents (Tenczer and Stüwe 2003; Stüwe 1998a). One possible interpretation of this relationship is presented in Fig. 7.13b.

## 7.5 Problems

**Problem 7.1.** *Converting pressure and volume (p. 338):*

a) What is the mean density of the atmosphere assuming that it is 10 km thick and that the pressure on the surface is 1 atmosphere? b) What is the barometric pressure in kilobar on the surface of the earth?

**Problem 7.2.** *Converting pressure and volume (p. 338):*

The molar volume of almandine-garnet is about  $11.5 \text{ J bar}^{-1}$ . Convert these units of volume (volume = energy/pressure) into cubic centimeters. Check your result by comparing the weight of one formula unit of almandine with its density. (The garnet almandine has the formula  $\text{Fe}_3\text{Al}_2\text{Si}_3\text{O}_{12}$  and a density of about  $4 \text{ g cm}^{-3}$ .

**Problem 7.3.** *Estimating cation diffusivities (p. 343):*

Use the Arrhenius relationship (eq. 7.5) to estimate the diffusivity of cations through a garnet lattice at  $300^\circ\text{C}$ ,  $500^\circ\text{C}$  and  $1000^\circ\text{C}$ . For the material constants use:  $Q = 239\,000 \text{ J mol}^{-1}$ ,  $D_0 = 9.81 \cdot 10^{-9} \text{ m}^2 \text{ s}^{-1}$  and assume the activation volume is negligible (Cygani and Lasaga 1985). The value of the gas constant may be found in Table D.4. Discuss the significance of the enormous difference between these cation diffusivities and the thermal diffusivity  $\kappa \approx 10^{-6} \text{ m}^2 \text{ s}^{-1}$ .

**Problem 7.4.** *Time scale of chemical equilibration (p. 344):*

How long does it roughly take for a 1 mm large garnet crystal to equilibrate chemically at  $600^\circ\text{C}$ ? Use eq. 7.4 and the data from Problem 7.3.

**Problem 7.5.** *Estimating mean diffusivity (p. 350):*

The Arrhenius relationship and the results of Problem 7.3 show a strong exponential dependence of cation diffusivity on temperature. In metamorphic processes cations in minerals diffuse typically during temperature change. It is therefore useful to know some sort of mean diffusivity that characterizes the diffusion rate during the entire thermal evolution. Use eq. 7.9 to estimate this mean diffusivity of cations in garnet between the starting temperature  $T_A = 700^\circ\text{C}$  and  $T_E = 400^\circ\text{C}$ . Use the data from Problem 7.3 and compare your result.

**Problem 7.6.** *The meaning of “metamorphic peak” (p. 346):*

Figure 7.3 shows in a a  $P$ - $T$ -path in which the thermal peak coincides with the pressure peak and in b a  $P$ - $T$ -path in which pressure and temperature peak occur at different times. The case of b is more common. In the literature we often read simply about the “metamorphic peak” of a terrain. a) Does this mean the temperature peak or the pressure peak? b) Why is this imprecise usage of the term “metamorphic peak” so common? c) Draw a  $P$ - $T$ -path in which the temperature peak is reached before the pressure peak. d) What tectonic process can you think of that causes such a path?

**Problem 7.7. Estimating cooling rates (p. 351):**

Assume you have found a garnet that is 2 mm in diameter (1 mm radius) that has a retrograde zoning profile similar to that shown on Fig. 7.4. You have analyzed it and found it to have a closure temperature of 600 °C. Estimate the cooling rate of the metamorphic event that caused the zoning profile using eq. 7.10.

**Problem 7.8. Using diffusivity estimates in the field (p. 343-344):**

(This is an integrated problem using knowledge on closure temperatures as well as the time scale of diffusion of both mass and heat. To tackle this problem you should solve problems 7.3, 7.4 and 7.5 first).

A 2 km thick mafic dike has intruded rocks that were cooling from a previous regional metamorphic event at the time of intrusion. The intrusion temperature of the mafic magma was  $T_i = 1200$  °C. The temperature of the cooling host rocks at the time of intrusion was  $T_b = 500$  °C. The dike has a contact metamorphic aureole that overprints the metamorphic parageneses from the previous event. In this contact metamorphic aureole, 5 mm large mica crystals were found at a distance of about 50 m from the dike contact. Microstructural observations show that these crystals grew prior to intrusion during the previous regional metamorphic event.

Question: Will the isotopic systems of the micas record the older regional metamorphic event or will they have been reset by the contact metamorphic event?

a) Find an answer by estimating the time scale of the contact metamorphic event (eq. 3.17) ( $\kappa = 10^{-6}$  m<sup>2</sup> s<sup>-1</sup>). Compare this with the time scale of diffusive equilibration of the micas at 1200 °C and 500 °C (from eq. 7.4 and using eq. 7.5). Use the following diffusion constants for mica:  $Q = 163\,000$  J mol<sup>-1</sup> and  $D_0 = 7.7 \cdot 10^{-9}$  m<sup>2</sup> s<sup>-1</sup> (Fortier and Giletti 1991).

b) The result from a) should have shown that, in principle, the micas *can* be used to date the intrusion event, but that the estimate we have made is very crude and needs refining. The dating may be improved by calculating the maximum temperature ( $T_{\max}$ ) the mica experienced during the contact metamorphic event (by substituting eq. 3.90 into eq. 3.88). Also, a better value for the diffusive time constant of the micas can be derived by calculating the mean diffusion rate between  $T_{\max}$  and  $T_b = 500$  °C using eq. 7.9.

c) Sketch out a flow chart for a computer program that could be used to find a detailed solution for the problem.

**Problem 7.9. Constructing P-T-paths (p. 355):**

Draw pressure-time and temperature-time paths that correspond to the  $P$ - $T$ - $t$ -paths shown in Fig. 7.15. Discuss some possible interpretations. The numbers shown along the  $P$ - $T$ -paths are geochronologically determined ages in my.

**Problem 7.10. Constructing P-T-paths (p. 355):**

Construct  $P$ - $T$ -paths from the pairs of pressure-time and temperature-time

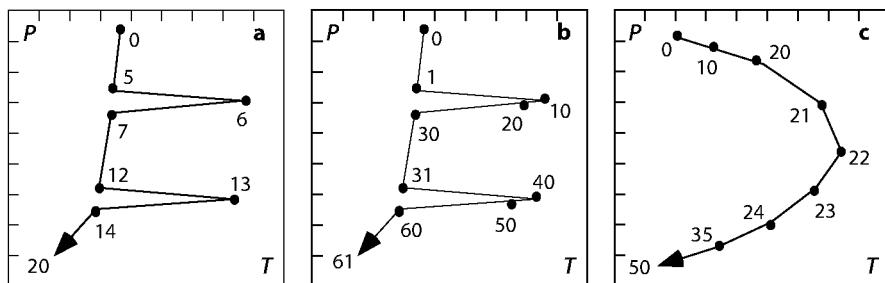


Figure 7.15. Illustration to Problem 7.9

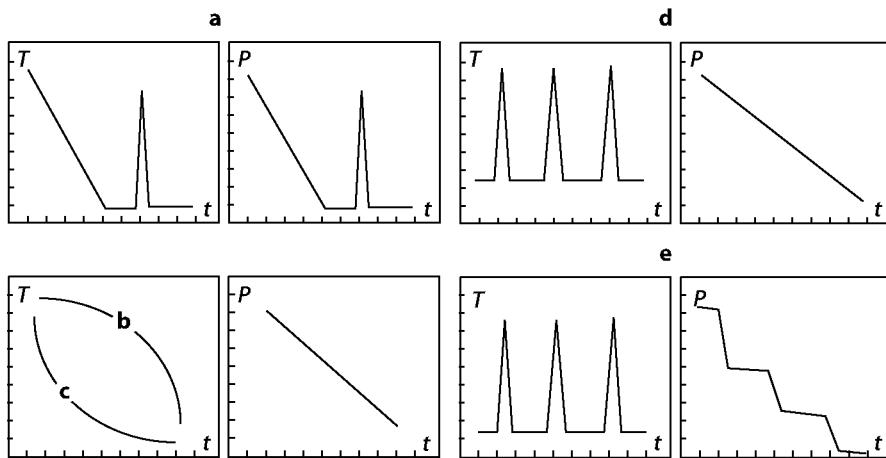


Figure 7.16. Illustration to Problem 7.10

curves shown in Fig. 7.16. The  $P$ - $T$  paths you have constructed contain no information on time. It is therefore possible to interpret them in terms of different temporal evolutions for  $P$  and  $T$  from those shown in Fig. 7.16 (if only the  $P$ - $T$  path is known). Discuss some.

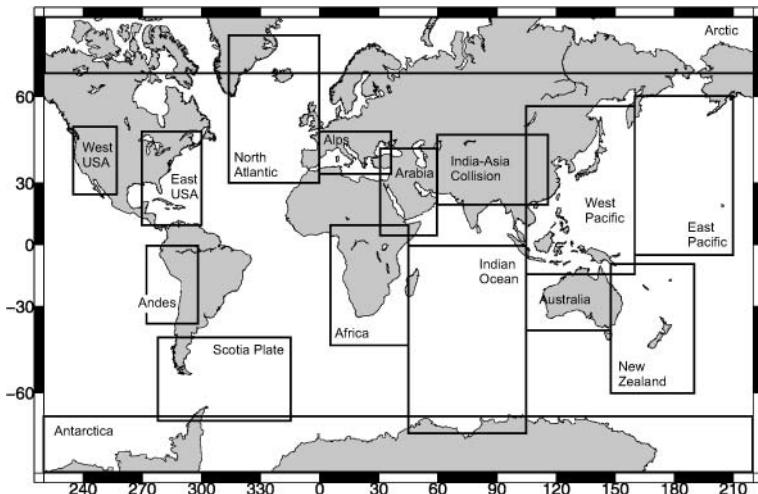
**Problem 7.11. Construction of  $P$ - $T$ -paths (p. 355):**

Draw the retrograde  $P$ - $T$ -path of a rock that experienced the following evolution following its metamorphic peak: Peak metamorphism occurred in 20 km depth at  $700^\circ\text{C}$ . Then, the rock was exhumed at a rate of  $1200 \text{ m my}^{-1}$  for 5 my. After that, the exhumation rate decreased to  $400 \text{ m my}^{-1}$ . The metamorphic temperature remained at  $700^\circ\text{C}$  for another 3 my after the metamorphic peak before cooling commenced. Then cooling commenced at a rate of  $100^\circ\text{C my}^{-1}$  and lasted for 3 my. After that, the cooling rate decreased to  $20^\circ\text{C my}^{-1}$  until a stable geotherm is reached. Final cooling occurred along the stable geotherm. Assume that the stable geotherm is characterized by a constant gradient of  $20^\circ\text{C km}^{-1}$ . Hint: It is easiest to first draw depth-time and temperature-time curves and construct a depth-temperature diagram

from those. The conversion from depth to pressure can be done as a final step in the calculation using  $\rho = 2\,700 \text{ kg m}^{-3}$  for density and  $g = 10 \text{ m s}^{-2}$  for the gravitational acceleration. Non lithostatic components of pressure are to be neglected.

## A. Topographic Maps

This section is a small Atlas of 16 of the most important areas in geodynamics around the world. All maps are shown in Mercator projection and were produced from the ETOPO2 global digital elevation model using the GMT software of Wessel and Smith (1995, 1998). Slab contours on some maps are after Gudmundsson and Sambridge (1998) (<http://wwwrses.anu.edu.au/seismology/projects/RUM/>). Plotted seismicity is from the IRIS seismic data base (<http://www.iris.edu>). All events above magnitude 5 since 1973 are plotted (except for Fig. A.3, Fig. A.10, Fig. A.14, where only magnitudes above 5.5, 6 and 5.5 are plotted, respectively). Depth of hypocenters is color coded: Red – yellow = 0 – 600 km. Size of dots refers to magnitude. Surface elevations are colored according to the scale bar shown on Fig. A.2. Major permanent rivers are shown and water bodies below 5000 km<sup>2</sup> are omitted. For scaling recall that one arc minute of latitude (not longitude !) is defined as one nautical mile or:  $10^\circ \approx 1080$  km, (10 degrees latitude are very roughly 1000 km) (s. p. 25). All maps (with and without labeling) and the GMT scripts used to produce them, are available from the author.



**Figure A.1.** World map showing the locations of the 16 maps presented here.

**Fig. A.2: India - Asia collision zone.** Topographic map of the most spectacular orogen on earth - the India Asia collision zone. The European Alps are plotted for scale the top right corner. The scale bar at top left is for all maps. 1 = Tibetan Plateau (covering the entire region between about 78°E and 95°E and 30°N to 37°N. 2 = The Shillong Plateau is the smallish elevated region north of the number. 3 = Tarim Basin. 4 = Altyntag Range. 5 = The deepest point in Central Asia is 154 m below sea level. 6 = Pamir. 7 = Kirgis Tien Shan. The low area northwest of the number is Lake Issyk Kul. 8 = Chinese Tien Shan. 9 = Western- or Namche Barwa syntaxis. The Brahmaputra River breaks through the chain here. 10 = Red River and Red River shear zone system. 11 = Chinese Loess Plateau. The region is shown in detail on Fig. 4.46. 12 = Mt. Everest. 13 = Mt Kailas. Both the Indus and the Tsangpo (later forming the Brahmaputra) come from here. 14 = Kazakhstan Platform. 15 = Ganges delta. 16 = Hindukush. 17 = Eastern syntaxis of the Himalaya. The Indus River breaks through the chain here. 18 = Gobi Desert. 19 = Dabie Shan - a famous ultra high pressure terrain. 20 = Arabian Basin and fan shaped cone of the Indus River. 21 = Sichuan Basin. 22 = Qilian Shan. 23 = Qaidam Basin. 24 = Long Men Shan. 25 = Kun Lun Range. 26 = Western Kun Lun Range. 27 = Western Ghats with the Deccan Traps flood basalts (mainly south of the map sheet). 28 = Helmand Basin and river. Kandahar is just northeast of the number. 29 = Kabul and the Afghan highlands to the west.

**Fig. A.3: Andes.** Topographic map of the Andes: the earth's most spectacular example of the collision between an oceanic plate and a continental plate. Labeled are: 1 = Altiplano. 2 = Lake Titicaca. 3 = Nasca Ridge. 4 = Juan Fernandez Islands. 5 = Carnegie Ridge with the Galapagos Islands at the west end (off the map). 6 = San Ambrosia Island. 7 = Peru-Chile Trench. Slab contours of Pacific Plate are drawn every 100 km depth. 8 = Sierras Pampeanas: a Panafrican granulite terrain. 9 = Amazonas Basin. 10 = Pre Cordillera. 11 = Frontal Cordillera. 12 = Cordillera Principal. 13 = Coastal Cordillera. 14 = Buenos Aires. 15 = Parana Basin and river. 16 = Rio de la Plata. 17 = Arequipa. 18 = Escondida. The world's largest copper mine.

**Fig. A.4: Europe.** Topographic map of southern Europe showing the Alpine orogen as well as the geometry of micro plates in the Mediterranean realm. Labeled are: 1 = Eastern Alps. The number is drawn into the Tauern window. The triangular feature just south of the number is the Peradriatic lineament marking the boundary between European and Adriatic plates. 2 = Western Alps. 3 = Pannonian Basin. 4 = Moesian Platform. The Iron Gate - where the Danube breaks through the Carpathian Arc is to be seen northwest of the number. 5 = Carpathian arc. 6 = Dinarides. 7 = The west - east striking structures transecting most of Turkey is the North Anatolian fault system. 8 = Hellenic Trench along the north directed subduction zone between the Anatolian Plate and the Adriatic Plate. 9 = Tyrrhenian sea. This

sea is interpreted as a backarc basin between Italy and the islands of Corsica and Sardinia. 10 = Ligurian sea. 11 = Ionian sea with the island of Malta just north of the number. 12 = Aegean sea. 13 = Rhine graben flanked by the Black Forest on the east and the Voges Range in the west. To the south is the Banana shaped Jura Mountains. 14 = Pyrenees. 15 = Subduction zone between the Adriatic Plate and the Apulian Plate. Mt Etna located northwest of the number along the east coast of Sicily is related to this subduction. 16 = Apuseni Mountains. 17 = Bohemian Massif. 18 = Massif Central. 19 = Rotation pole of the Adriatic Plate. 20 = Atlas Mountains. 21 = Vesuvio. 22 = Appenine Range. 23 = Rhodope Range.

**Fig. A.5: Arabian Region.** Topographic map of the Arabian region. 1 = Afar Depression. 2 = Egyptian Eastern Desert. Pan African core complexes are exhumed here. 3 = The Oman ophiolite complex. 4 = The deepest point on land in the world: the Dead Sea shore is 417 m below sea level. The Dead Sea lies along the Dead Sea Transform Fault that continues to form the boundary between the Arabian Plate and the Anatolian Plate (black line). 5 = Zagros Mountains. 6 = Luxor on the Nile river. 7 = Ethiopian Highlands. 8 = North Anatolian Fault System. 9 = Mt Elbrus (5633 m) in the Caucasus is the highest point in Europe. 10 = United Arab Emirates – where most of the oil is. 11 = Caspian Sea. 12 = Red Sea Rift. 13 = Persian Gulf. 14 = Gulf of Oman – separated from the Persian Gulf by the Straits of Hormuz. 15 = Indian Plate. 16 = Socotra Island at the entrance to the Gulf of Aden. 17 = Mt Ararat. 18 = Elburz Mountain range between Teheran and the Caspian Sea. 19 = Troodos ophiolite complex on Cyprus.

**Fig. A.6: Africa.** Topographic map of Africa. 1 = Lesotho highlands. East of the highlands is the city of Durban at the costs. The marked topographic jump from the highlands down to the coastal lands is the South African Great Escarpment that can be seen all around the southern edge of the continent. In Lesotho it is called the Drakensberg escarpment. Note that the drainage in Lesotho drains west into the Oranje River. 2 = Namaqualand and the mouth of the Oranje River. 3 = Kalahari Desert. 4 = The lowlands south of the Great Escarpment are made up of the Cape fold belt. Note large scale folding of strata. 5 = Limpopo River drainage basin. Note how the basin breaks through the Great Escarpment. 6 = Sambezi River. 7 = Lake Njassa. The southernmost of the large lakes along the East African rift system. 8 = Lake Tanganyika. 9 = Lake Victoria. 10 = Nairobi. The spots north and south of the number are Mt Kenia and Mt Kilimanjaro, respectively. The highlands just west of the rift valley are the Serengeti. 11 = Lake Rudolf. 12 = The lowlands at the north end of the map widen into the Afar Depression. 13 = Kongo Basin. 14 = Fernando Poo hot spot trace. Note how the chain of islands can also be seen onland in Kamerun. The line is also known as Guinea Ridge. 15 = Walvis Ridge. 16 = Agulhas Plateau. 17 = Mozambique Plateau. 18 = Malagasy fracture zone. This fracture zone is part of the ridge system separating the African Plate from the Antarctic Plate. 19 = Namibian Desert.

Brandberg is just north of the number. 20 = Angola Basin. 21 = In this region are the Witwatersrand gold reef and the Kimberley diamond mines. Pretoria and Johannesburg are near here. 22 = Aldabra and Comoro Islands. 23 = Etosha Pan. 24 = Olduvai Gorge - the site of the earliest human fossils. Also: Ngorongoro Crate and Oldolnya Lengai – the worlds only active carbonatite volcano. 25 = Cape Basin. 26 Agulhas Ridge.

**Fig. A.7: North Atlantic.** Topographic map of the North Atlantic. 1 = Mid-Atlantic Ridge. 2 = Azores. 3 = Madeira Islands. 4 = Biscay abyssal Plain. 5 = Flemish Cap. 6 = Faeroe Islands. 7 = Jan Mayen Island. South of the number extends the Jan Mayen micro-continent to Iceland. 8 = Reykjanes Ridge (active 55 Ma to now). Note the V-shaped arrangement of ridges and the continuous rise of the Mid-Atlantic Ridge towards Iceland for almost 1000 km. 9 = Rockall Island. 10 = Hatton Basin. 11 = Rockall Trough. 12 = Scoresby Sund. The worlds largest fjord with 475 km length. 13 = Gunbjorns Fjeld (3700 m). The highest mountain on this map. 14 = Greenland ice sheet is about 3000 m thick here. Bedrock is 300 m below sea level. 15 = Aegir Ridge was active from 55 - 33 Ma before it jumped across to: 16 = Kolbeinsey Ridge, active now. 17 = Present day position of Iceland Hot Spot. 18 = Liverpool Land: Remains of Caledonian crust with ultra high pressure eclogites. 19 = Huge sedimentary fans deposited from glacial erosion on Greenland. 20 = Mohns Ridge (active now).

**Fig. A.8: Eastern North America.** Topographic map of eastern North America. 1 = Bermuda Islands. 2 = Florida. 3 = Deepest point in the Atlantic Ocean at the west end of the Antilles Trench. 4 = Lesser Antilles. 5 = Caribbean Sea. 6 = Puerto Rico Trench. 7 = Bahamas. 8 = West Indies. 9 = Appalachian Mountains. 10 = Great Lakes. 11 = Blake Plateau. 12 = Gulf of Mexico. 13 = Bermuda Rise. The large fractures to the east belong to the Mid-Atlantic Ridge. 14 = The Mississippi Delta and New Orleans. Note the fan in the Gulf of Mexico. 15 = Outer ridge. 16 = Canadian Shield. 17 = New York and Long Island. 18 = Cape Cod and Boston to the west on the mainland. 19 = Nova Scotia. 20 = St Laurence seaway. 21 = Site of the Chicxulub crater which is often made responsible for the mass extinction at the Cretaceous - Tertiary boundary.

**Fig. A.9: Western North America.** Topographic map of western North America - a region which hosts a series of classic geodynamic settings. Labeled are: 1 = San Andreas Fault (from north west to south east). 2 = Lake Tahoe. 3 = Grand Canon and Colorado River. Hoover Dam with Lake Mead and Las Vegas are just south west of the number. 4 = Snake River Plain. 5 = Yellowstone Hot Spot. 6 = Columbia River. 7 = Sierra Nevada (outlined). 8 = Great Salt Lake. 9 = Great Valley. San Francisco Bay is west of the number. 10 = Colorado Plateau (outlined) and Upper Colorado River in Utah. The number is drawn over Monument Valley and west of the number is Lake Powell. 11 = Vancouver Island. 12 = Seattle. Just east of the number

is Mt Rainier as the north end of a linear chain of volcanoes including from north to south Mt St Helens, Mt Adams and Mt Hood as the highest peaks. 13 = Sierra Madre Occidental (outlined). 14 = Southern Basin and Range (outlined). 15 = Northern Basin and Range (outlined). 16 = Uinta Mountains. 17 = Klamath Mountains. 18 = Cascadia Mountains. 19 = Bighorn Mountains. 20 = Transverse Ranges. 21 = Rio Grande Rift. 22 = Southern Rocky Mountains. 23 = Black Hills. 24 = Bitterroot Mountains. 25 = Death Valley. Owens Valley extends further north west of the number. 26 = Meteor Crater. 27 = Yosemite Valley. 28 = Madre Archipelago. Outlines are after Thelin and Pike (1991). See also Ellis et al. (1999), Sonder and Jones (1999) or Leeman (1989).

**Fig. A.10: Western Pacific.** Topographic map of the western Pacific also known as the “ring of fire”. 1 = Kuril Trench. 2 = Japan Trench. 3 = Izu Bonin Trench. 4 = Mariana Trench. 5 = Ryukyu Trench. 6 = Challenger Deep - the deepest point known in the oceans. It is 10911 m deep and is located near 11°22'N and 142°36'E. It was discovered by the Royal Navy survey ship HMS Challenger in 1951 and the point was visited by J Piccard and D Walsh in the US Navy submersible Trieste in 1960. 7 = Java (Sunda) Trench. 8 = Bougainville Trench. 9 = Philippine Trench. 10 = Parece Vela Ridge. The number is central to the Philippine Plate which extends from the Mariana Trench to the Philippines. 11 = Sikhote Alin range. 12 = Lake Baikal and the Baikal rift. 13 = Shatsky Rise. 14 = Banda Sea. 15 = Taiwan. 16 = Luzon. 17 = Yellow Sea. 18 = Sea of Okhosk. 19 = South China Sea. 20 = New Britain Trench. 21 = Ottong Java Plateau - the worlds largest igneous province. 22 = Arafura Basin between Papua New Guinea and the Australian Gulf of Carpentaria. 23 = Sunda shelf. 24 = Wallace line separating different paleontological faunas between Borneo and Sulawesi (continues between Bali and Lombok to the south and just south of the Philippines to the north. Slab contours are drawn every 100 km depth for the Indonesian, the Luzon-, the Ryukyu, west Philippine- and east Philippine slabs. For slab contours of the Mariana and Japan slabs see Fig. 2.6.

**Fig. A.11: Eastern Pacific.** Topographic map of the eastern Pacific. 1 = Aleutian Trench. 2 = Hawaiian Ridge and Hawaiian islands. The largest island at the east end of the ridge is "Big Island". Note the diffuse light shading indicating shallow water several hundreds of kilometers north and south of the chain and the deeper water (dark color) in the immediate vicinity of the chain indicating lithospheric flexure. 3 = Emperor seamount chain. 4 = Midway Islands. 5 = Mid Pacific Mountain Range including Cape Johnson Guyot and Hess Guyot as the highest (under water) peaks. 6 = Bowers Bank. 7 = Kodiak Island in Alaska. 8 = North end of the Manihiki Plateau. A Cretaceous large igneous province. 9 = Gilbert Islands. 10 = Marshall Islands. 11 = Mendocino Fracture Zone. 12 = Murray Fracture Zone. 13 = Christmas Island Ridge. 14 = Shatsky Rise. 15 = Hess Rise. 16 = Kuril Trench. 17 =

Pribilof Islands. 18 = Emperor Trough. 19 = Chinook Trough. 20 = Musicians Seamounts. 21 = Necker Ridge. 22 = Hokkaido Trough.

**Fig. A.12: Indian Ocean.** Topographic map of the Indian Ocean - the birthplace of many plate tectonic concepts. 1 = The Amery ice shelf in the failed rift of the Lambert graben, Antarctica. Today, the world's largest glacier, the Lambert glacier, flows down this graben. 2 = Enderby land. 3 = These are not islands - they are icebergs. 4 = Kerguelen Plateau with the French Kerguelen Island to the northwest and the Australian Heard Island to the southeast of the number. 5 = Ninety East Ridge. 6 = Broken Ridge Plateau. 7 = Diamantina Fracture Zone. 8 = Madagaskar. 9 = Chagos Trench with the Chagos Lacadive Plateau to the west. 10 = Seychelles. 11 = Amirante Trench. 12 = La Reunion in the west and Mauritius in the east. 13 = Ile St Paul and Ile Amsterdam at the southwest end of the Amsterdam fracture zone. 14 = Crozet Islands. 15 = Cocos Islands. 16 = Triple junction between African Plate, Indo-Australian Plate and Antarctic Plate.

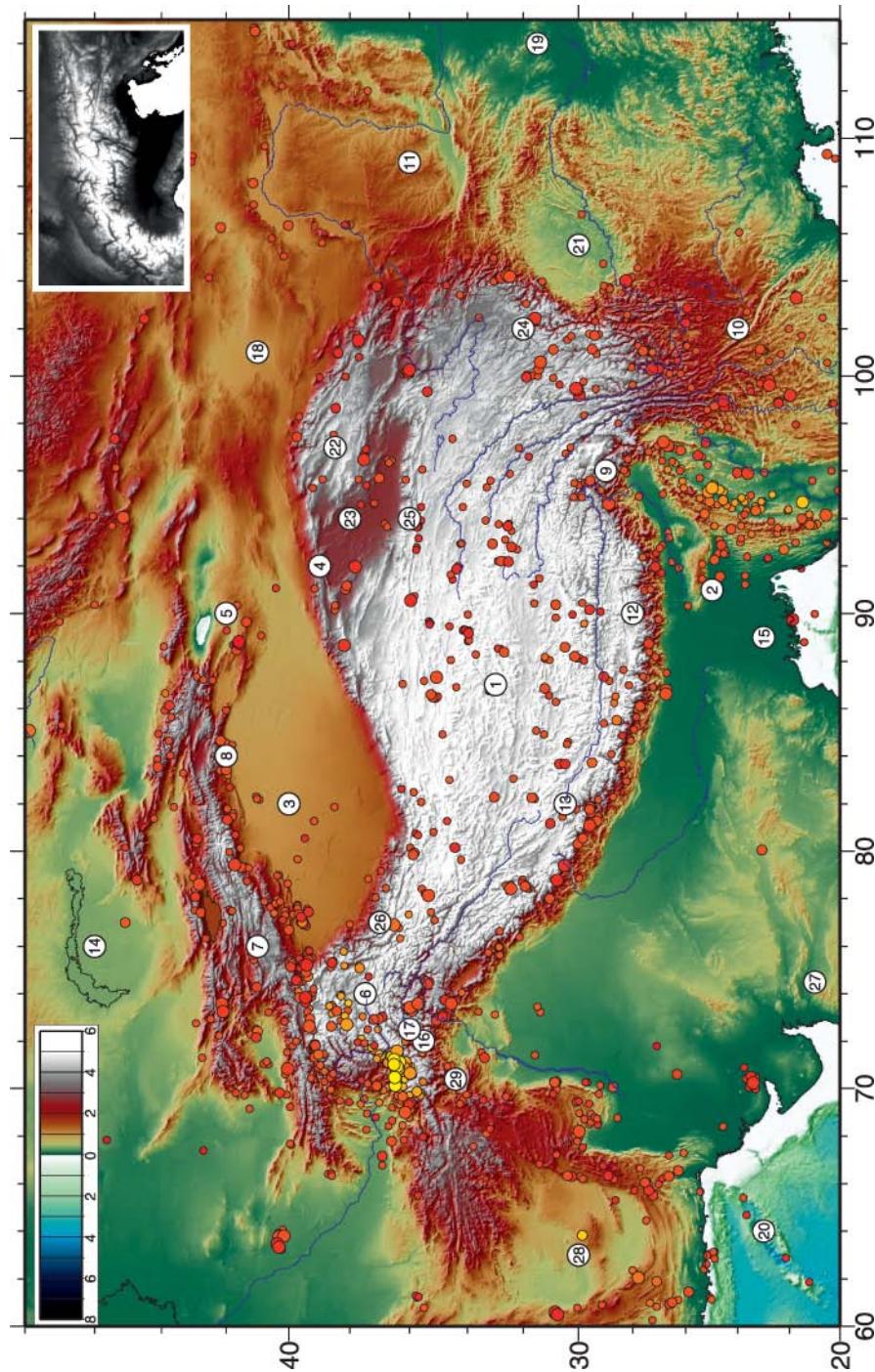
**Fig. A.13: Australia.** Topographic map of Australia. 1 = Lake Eyre and other salt lakes (just below sea level). 2 = The Proterozoic Broken Hill Block and its Pb - Zn deposit. 3 = Flinders Ranges. 4 = Murray Basin. 5 = Otway Basin (most of it is off shore). 6 = Cooper - Eromanga Basin hosting most of Australia's onshore gas deposits. 7 = Proterozoic Arunta Block. 8 = Musgrave Block. 9 = Mt Isa Block and Mt Isa deposit. 10 = Kimberley Ranges. 11 = Northwest Shelf hosting most of Australia's hydrocarbon deposits. 12 = Albany Fraser Province. 13 = Nullabor Plain and Great Australian Bight. 14 = Ayers Rock. 15 = The Archaean Pilbara Block. 16 = Diamantina Trench. 17 = Wallaby Plateau- 18 = Hammersley Basin and iron deposits. 19 = Jack Hills where the oldest zircons on Earth were found. 20 = Yilgarn Craton (going as far as the Jack Hills). 21 = Australian Dividing Range and Snowy Mountains. 22 = Gawler Craton.

**Fig. A.14: New Zealand.** Topographic map of New Zealand. The black line in New Zealand is the Alpine Fault. 1 = The Great escarpment of eastern Australia. 2 = Tonga Trench and Tonga Islands. Slab contours are drawn every 100 km. 3 = Kermadec Trench. 4 = Macquarie Trench and Macquarie Island just northeast of the number. 5 = Campbell Plateau. South of the number are the Auckland islands. 6 = Lord Howe rise and Lord Howe island west of the number. 7 = New Caledonia. 8 = New Hebrides Trench. Slab contours are drawn every 100 km depth. 9 = Fiji Plateau with the Fiji islands east of the number. 10 = Samoa Islands. 11 = Norfolk Island. 12 = Lau Ridge. 13 = Bounty Trough with the Catham Islands to the northeast. 14 = Queensland Plateau on the south side of the Solomon sea. 15 = Tasman Sea as part of the Indo-Australian Plate. 16 = New Hebrides. 17 = Louisville seamount chain. 18 = White Island and the Taupo volcanic zone on the New Zealand mainland. 19 = Colville Ridge.

**Fig. A.15: Scotia Plate.** Topographic map of the Scotia Plate at the south tip of south America: 1 = Falkland Plateau and Falkland Islands. 2 = South Sandwich Trench. Slab contours are drawn every 50 km depth. 3 = South Sandwich microplate with the South Sandwich Islands just north and south of the number. The microplate extends to the small ridge just west of the number. 4 = South Georgia. 5 = Scotia Ridge. 6 = Argentinian abyssal plain. 7 = Scotia Plate. 8 = South Orkney Islands. 9 = Antarctic Plate. 10 = Elephant Island. 11 = Antarctic Peninsula. 12 = Patagonia. 13 = South Shetland Islands. 14 = Shackleton Fracture Zone. 15 = Falkland Trough. 16 = Chile triple point. Currently the only place on earth where a mid oceanic ridge is being subducted.

**Fig. A.16: Arctic.** Topographic map of the Arctic: 1 = Franz Josef Land. 2 = Spitzbergen (Svalbard). 3 = Ellesmere Island. 4 = Baffin Island. 5 = Victoria Island. 6 = Barents abyssal plain. 7 = Wrangel Island. 8 = North magnetic pole. 9 = New Siberian Islands. 10 = Nansen Ridge. 11 = Lomonosov Ridge. 12 = Novaya Zemlya. 13 = Severnaya Zemlya. 14 = Jan Mayen Island is just west of the number. 15 = Greenland Fracture Zone. 16 = Canada Basin. 17 = Makarov Basin. 18 = The early Archaean Isua supracrustal belt, with 3.8 Ga believed to be the oldest rock on Earth.

**Fig. A.17: Antarctica.** Topographic map of Antarctica. The two percent of the continent that are ice free bear many unique geological features: 1 = Enderby Land – one of the first discoveries of ultra high temperature metamorphism. 2 = Lambert rift – a Mesozoic failed rift and home to the Lambert Glacier, the largest glacier in the world. 3 = Prince Charles Mountains. The last discovered major mountain range on the globe. 4 = Larsemann Hills and Vestfold Hills. Two classic low pressure - high temperature metamorphic areas. 5 = Mt Vinson. With 4892 m the highest peak of Antarctica. 6 = Antarctic Peninsula. 7 = Transantarctic Mountain Range. 8 = McMurdo Sound. 9 = Dry Valleys. 10 = Mt Erebus, Antarctica's only active volcano. 11 = Law Dome. A region of stationary ice where much of the worlds ice core drilling is done for palaeoclimatological research. 12 = Vostok Station. Russian permanent research station near the coldest and highest part of the continent. 13 = West Antarctica. 14 = Patriot Hills. 15 = The South Magnetic Pole (2005) is just north of here off the map. 16 = Ross Ice Shelf. 17 = Ronne Ice Shelf. 18 = Weddell Sea. 19 = Bunker Hills. 20 = Dronning Maud Land. 21 = Allan Hills in Victoria Land. Many meteorites have been found in this region.



**Figure A.2.** Topographic map of the most spectacular orogen on earth - the India-Asia collision zone.

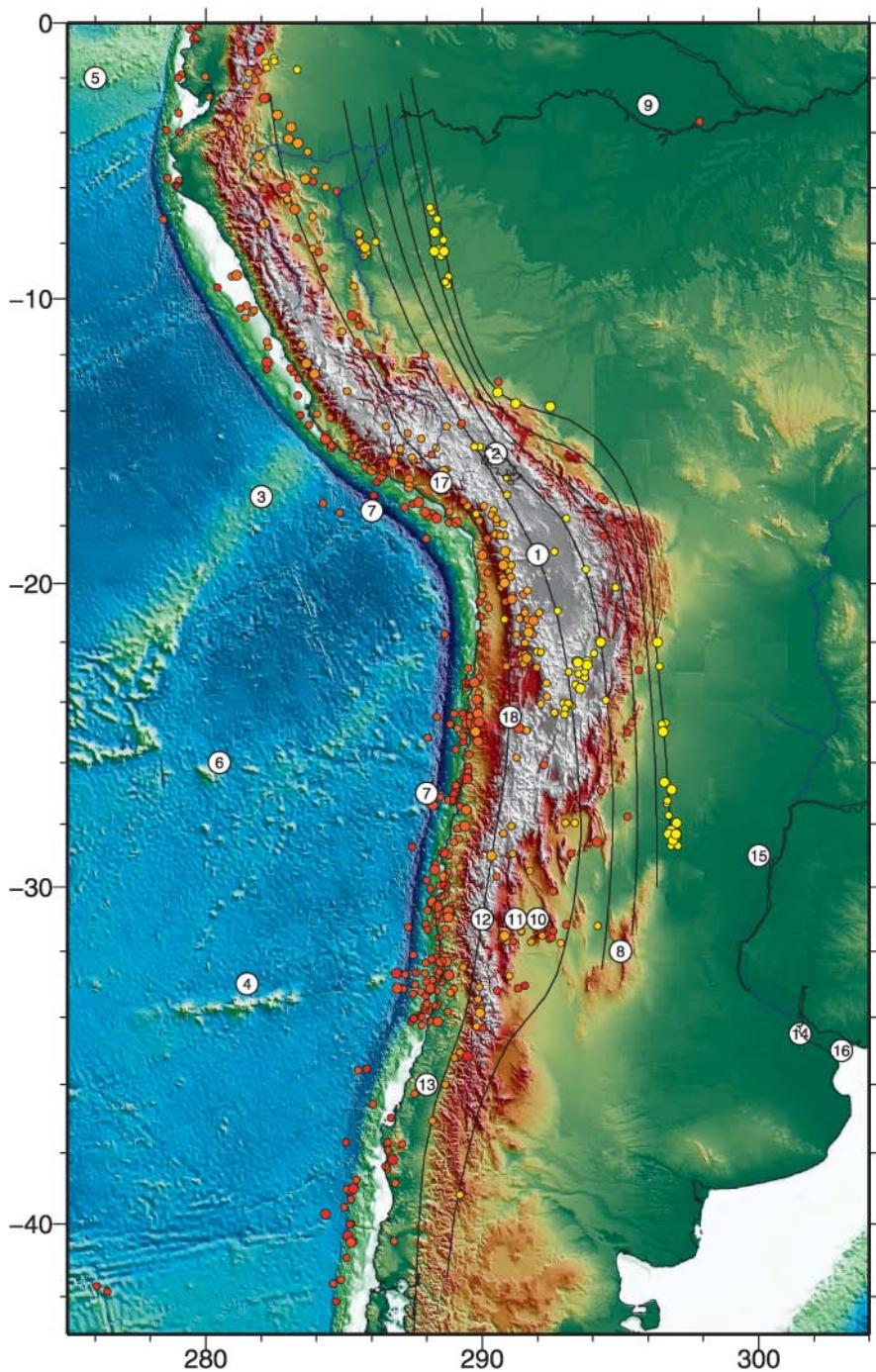


Figure A.3. Topographic map of the Andes.

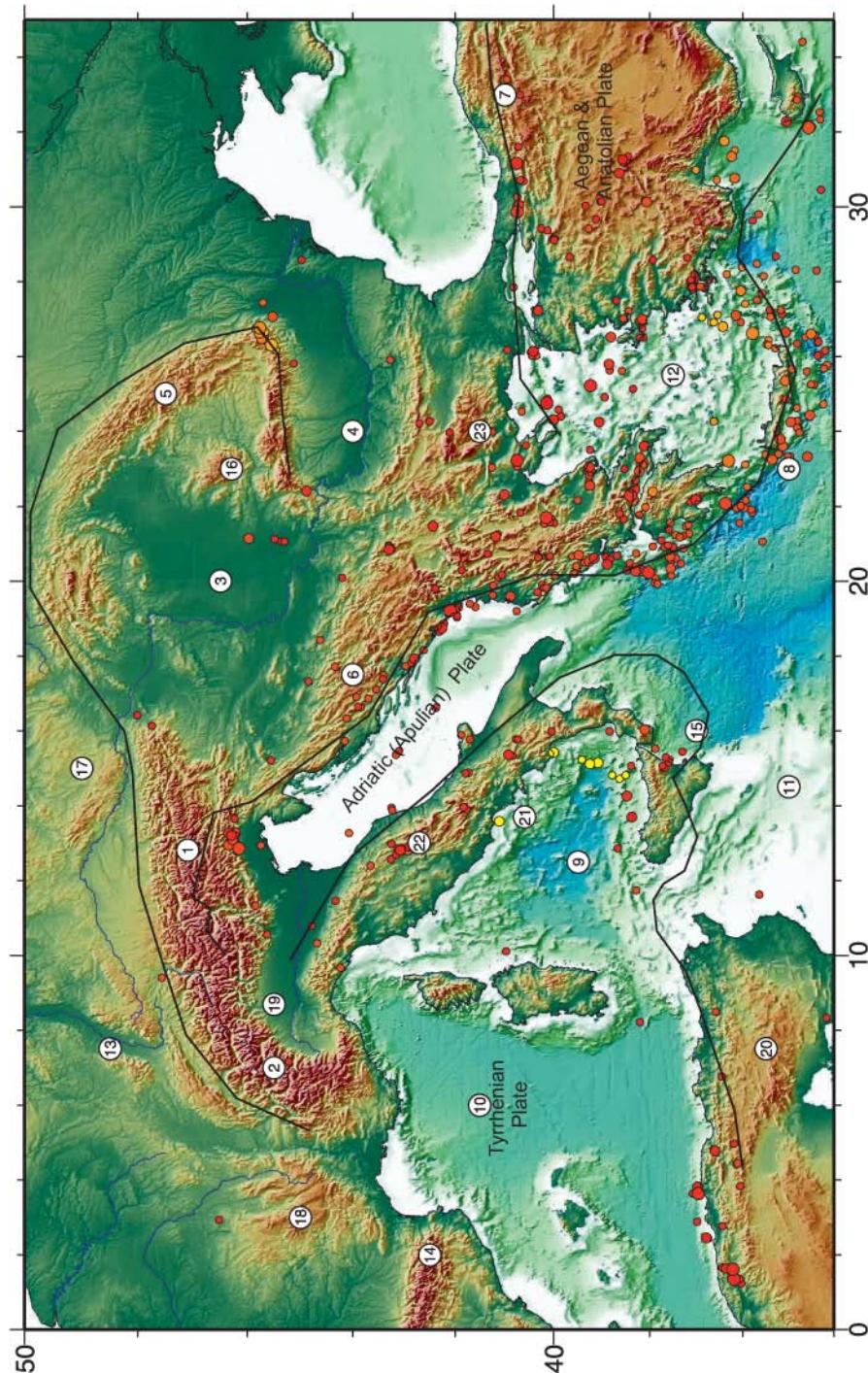
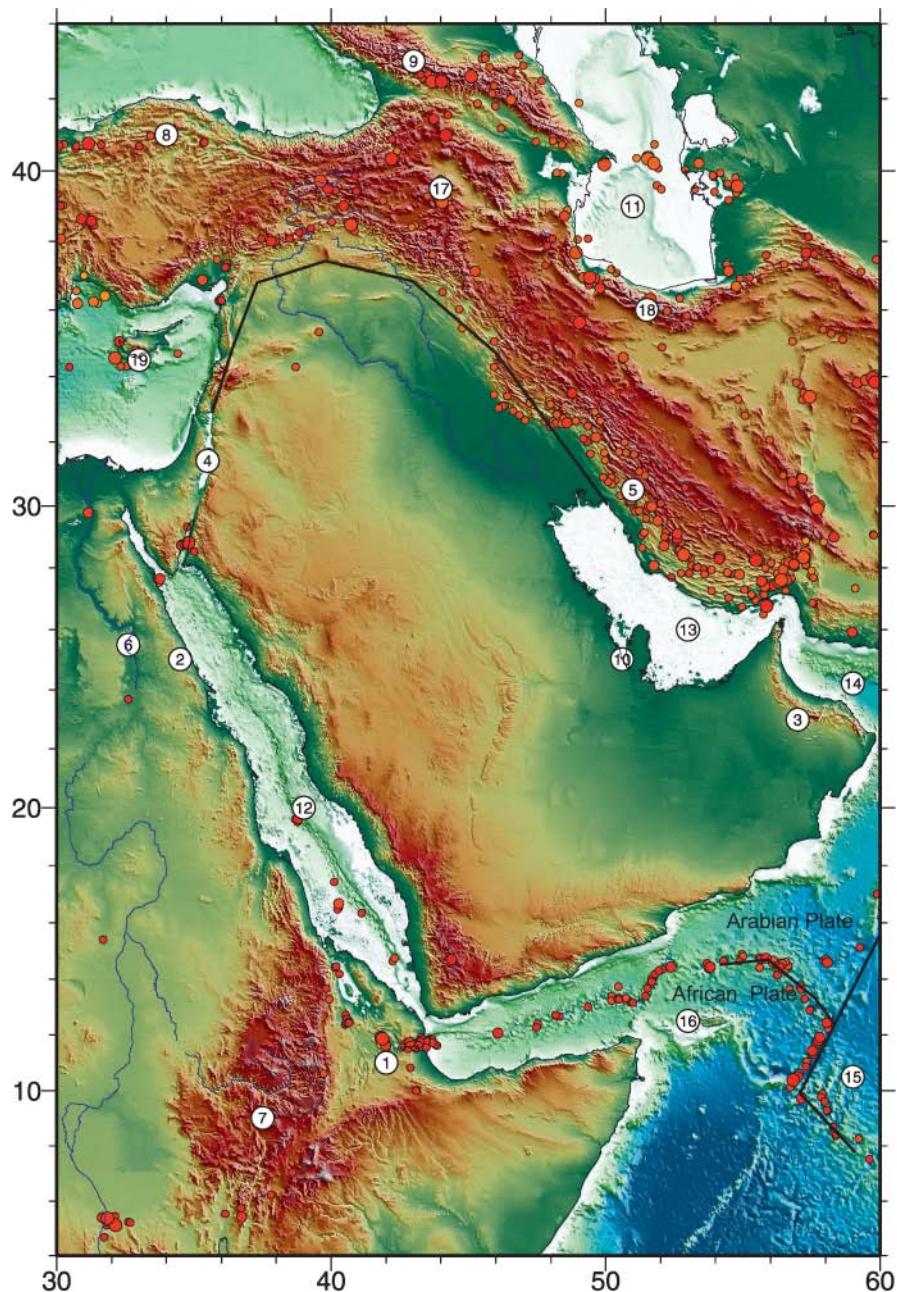


Figure A.4. Topographic map of southern Europe.



**Figure A.5.** Topographic map of the Arabian region.

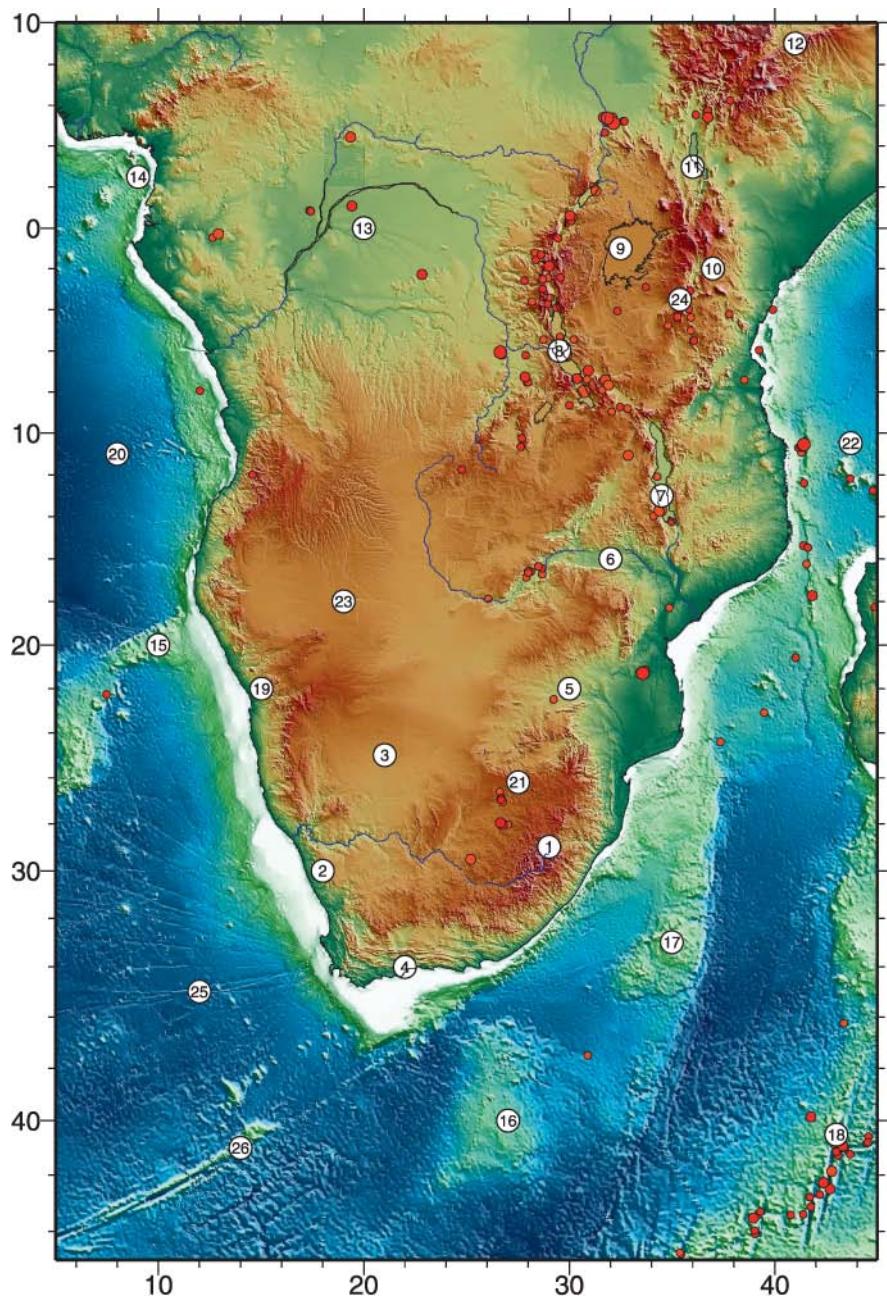


Figure A.6. Topographic map of Southern Africa.

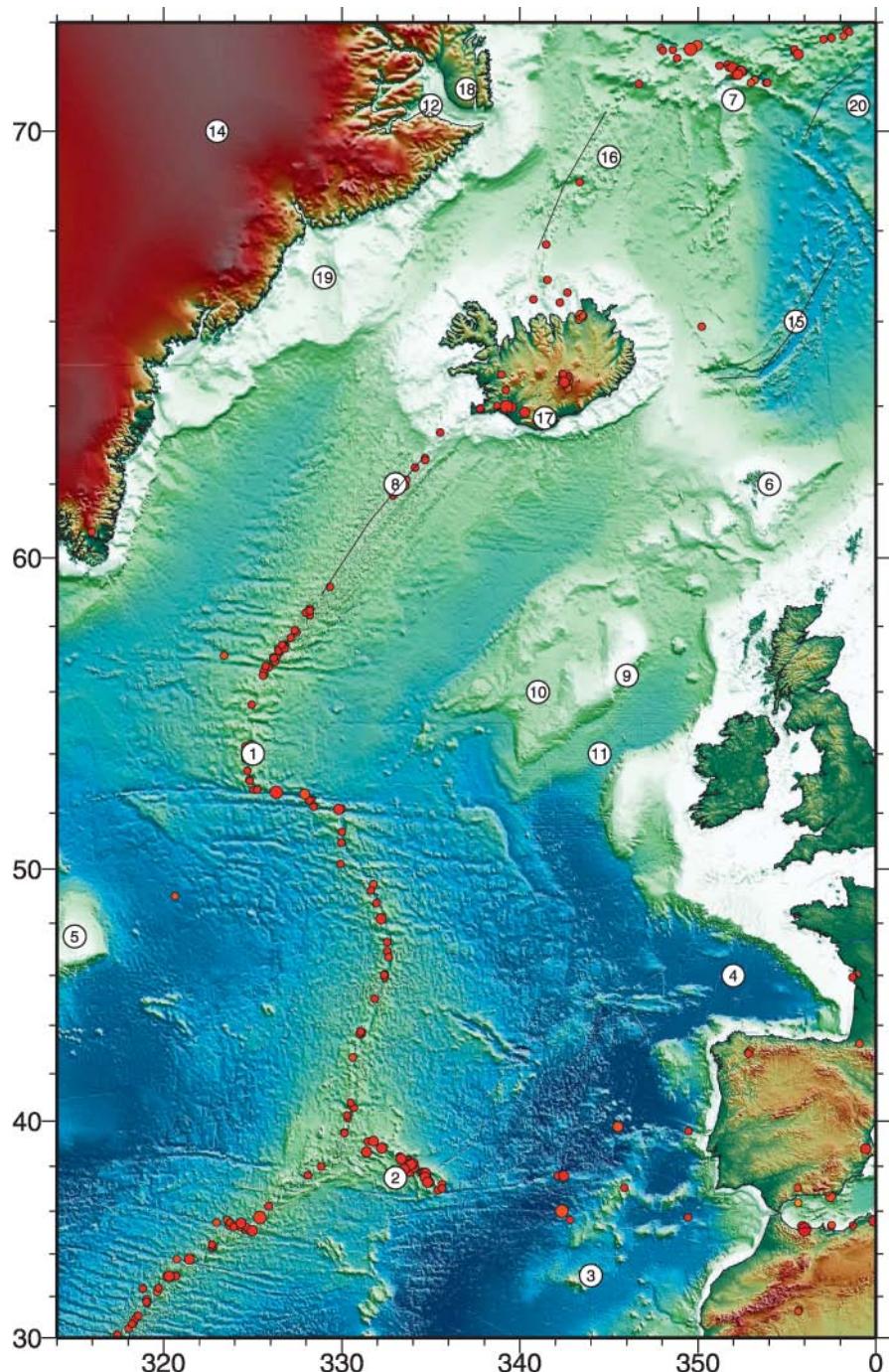


Figure A.7. Topographic map of the northern Atlantic Ocean.

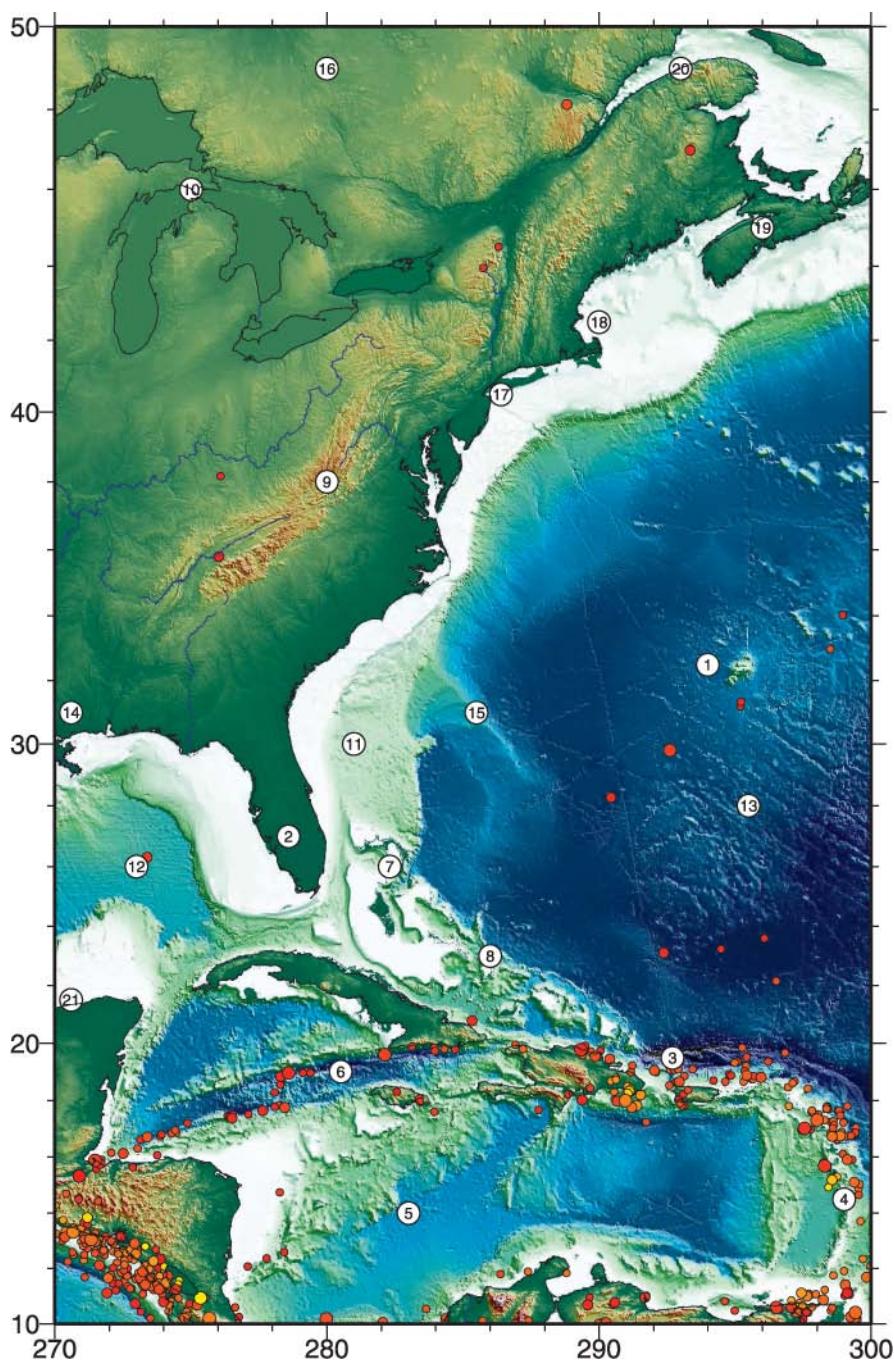
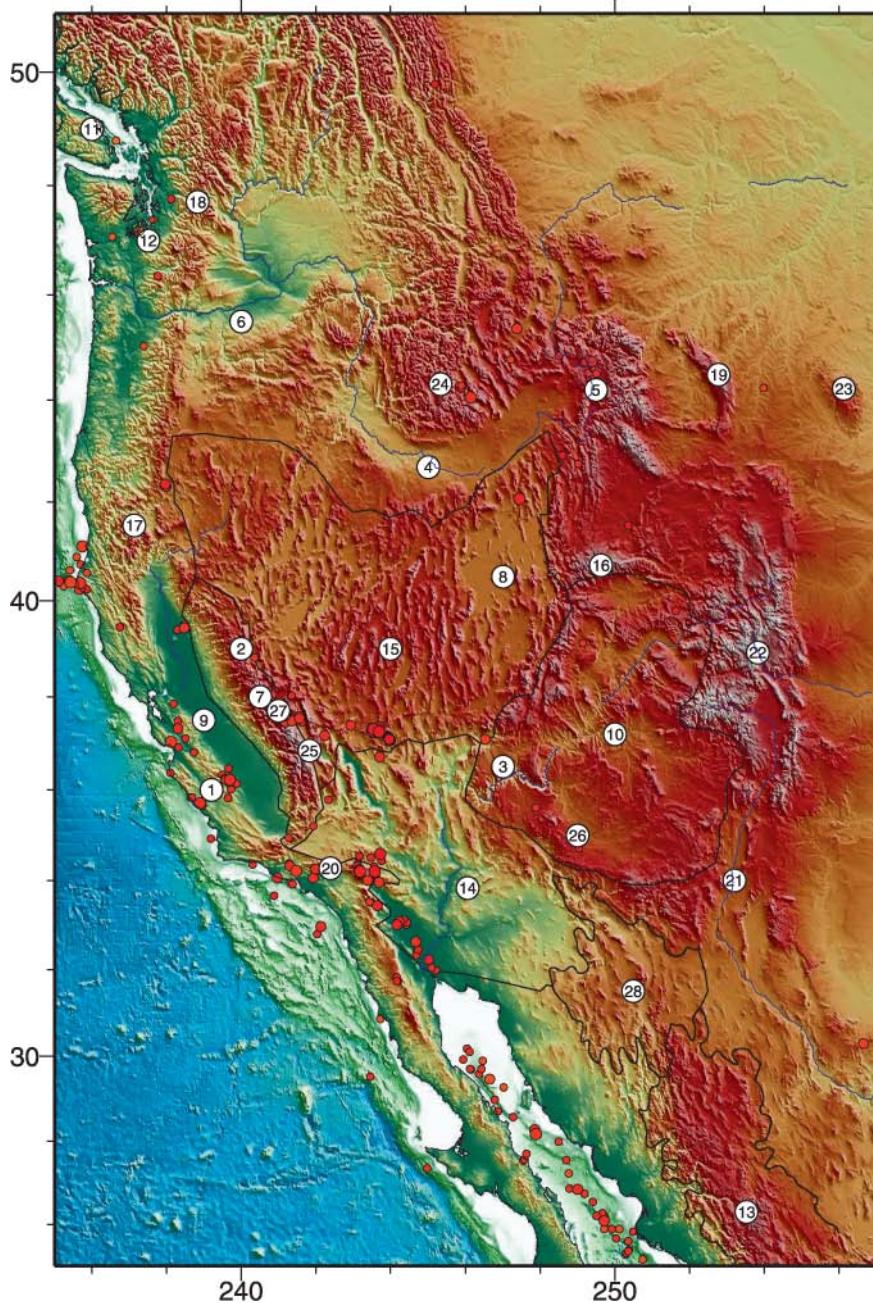
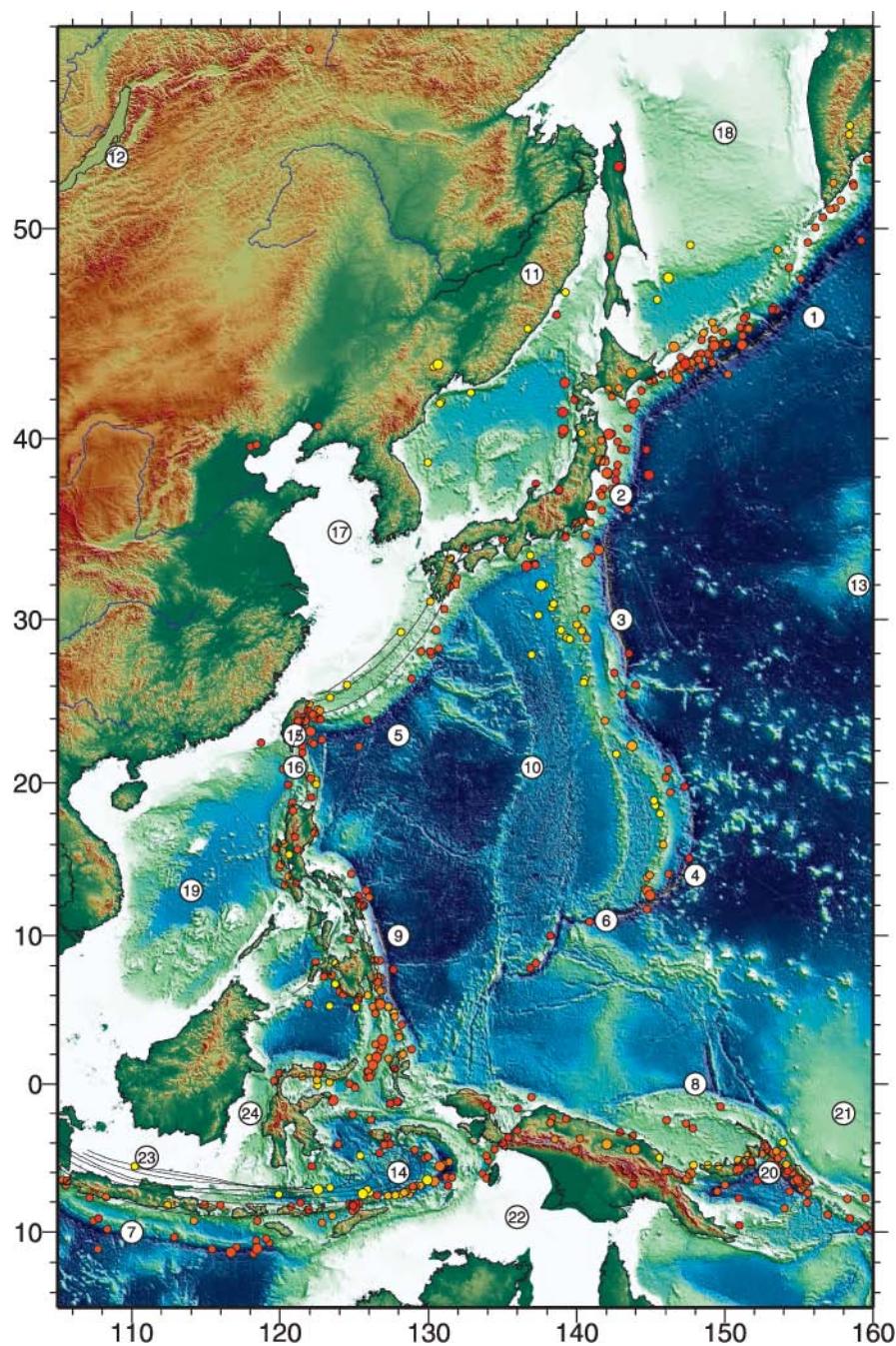


Figure A.8. Topographic map of eastern North America.



**Figure A.9.** Topographic map of western North America.



**Figure A.10.** Topographic map of the western Pacific also known as the "ring of fire".

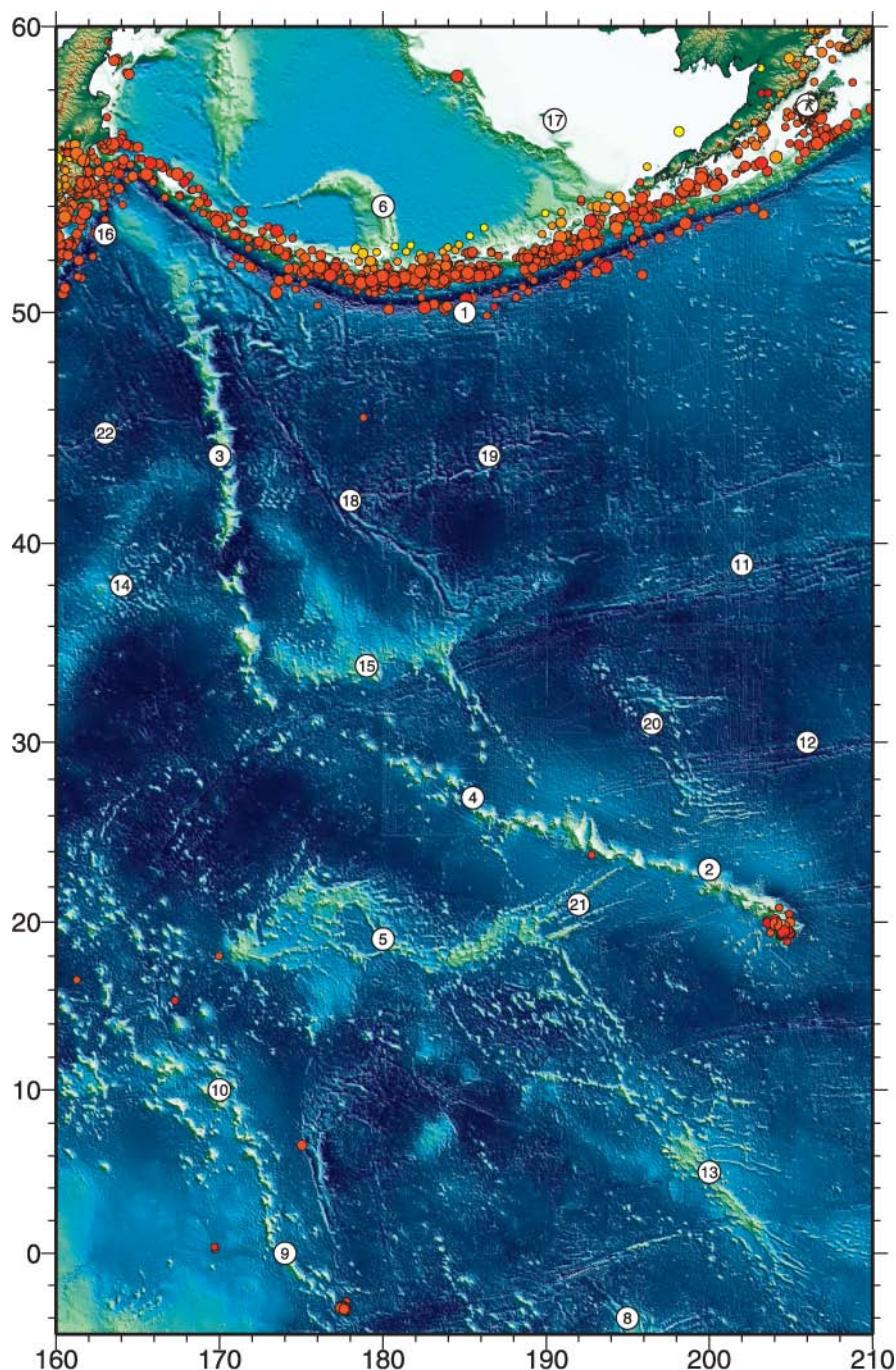
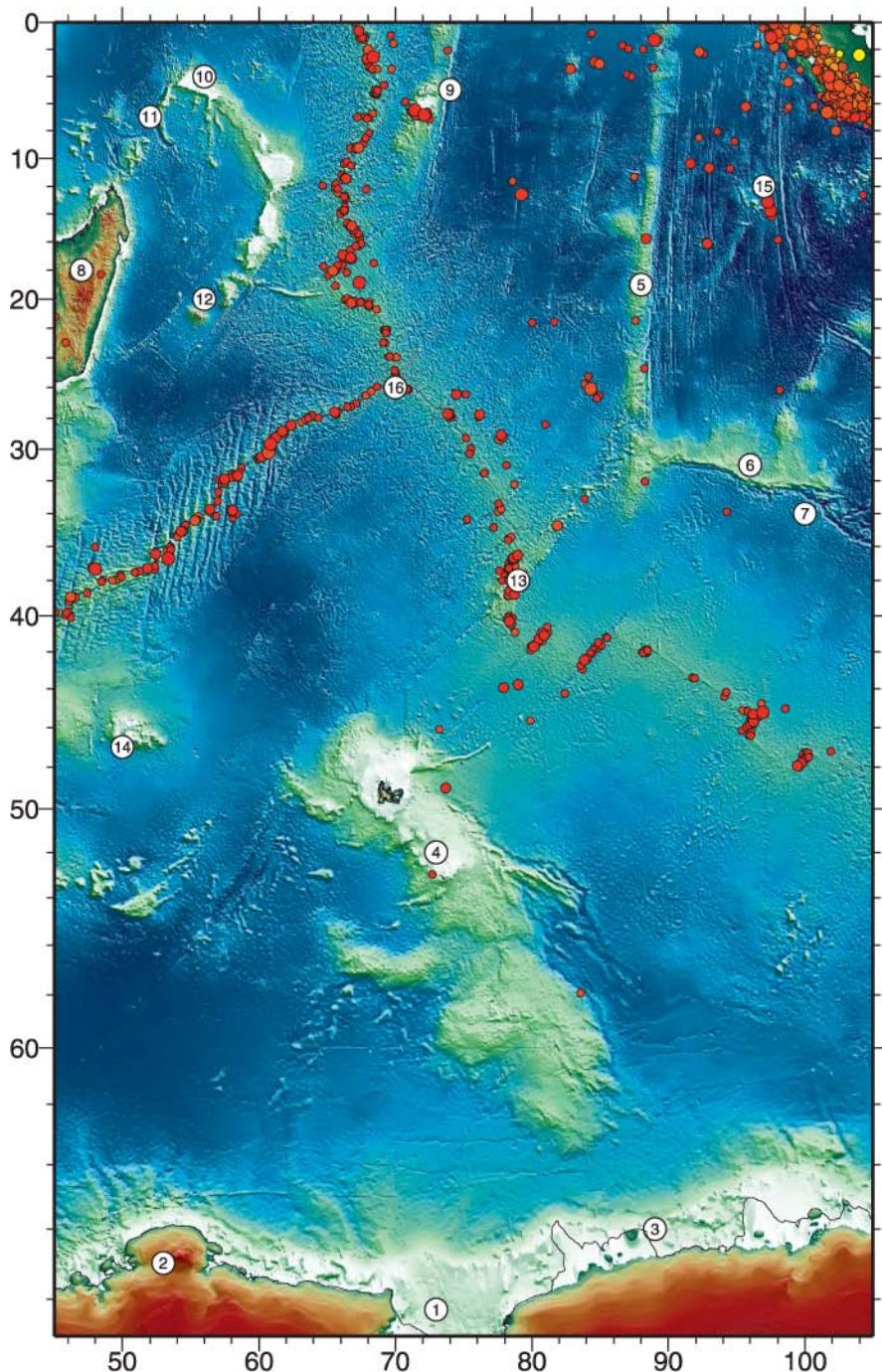


Figure A.11. Topographic map of the eastern Pacific.



**Figure A.12.** Topographic map of the Indian Ocean - the birthplace of many plate tectonic concepts.

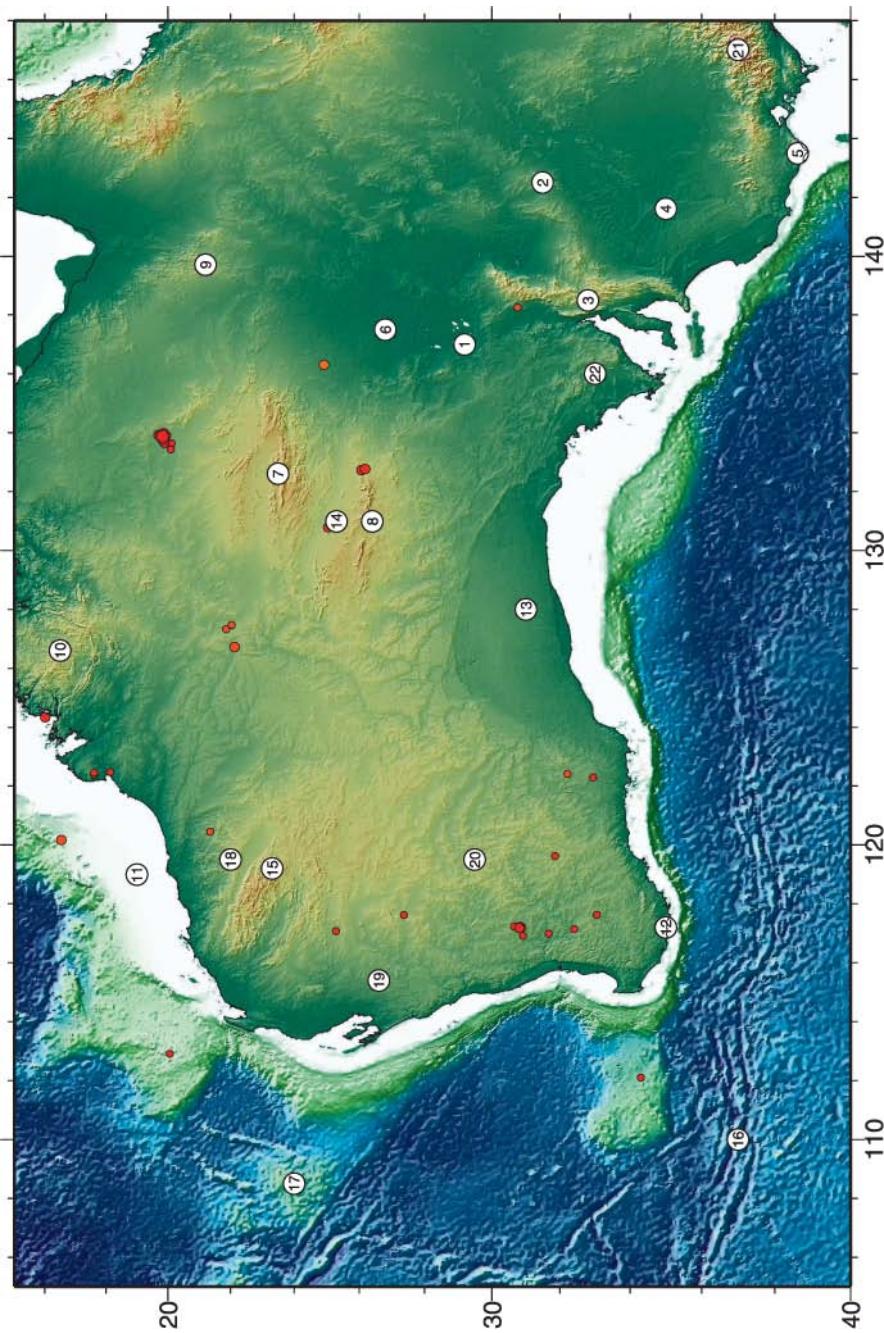


Figure A.13. Topographic map of Australia.

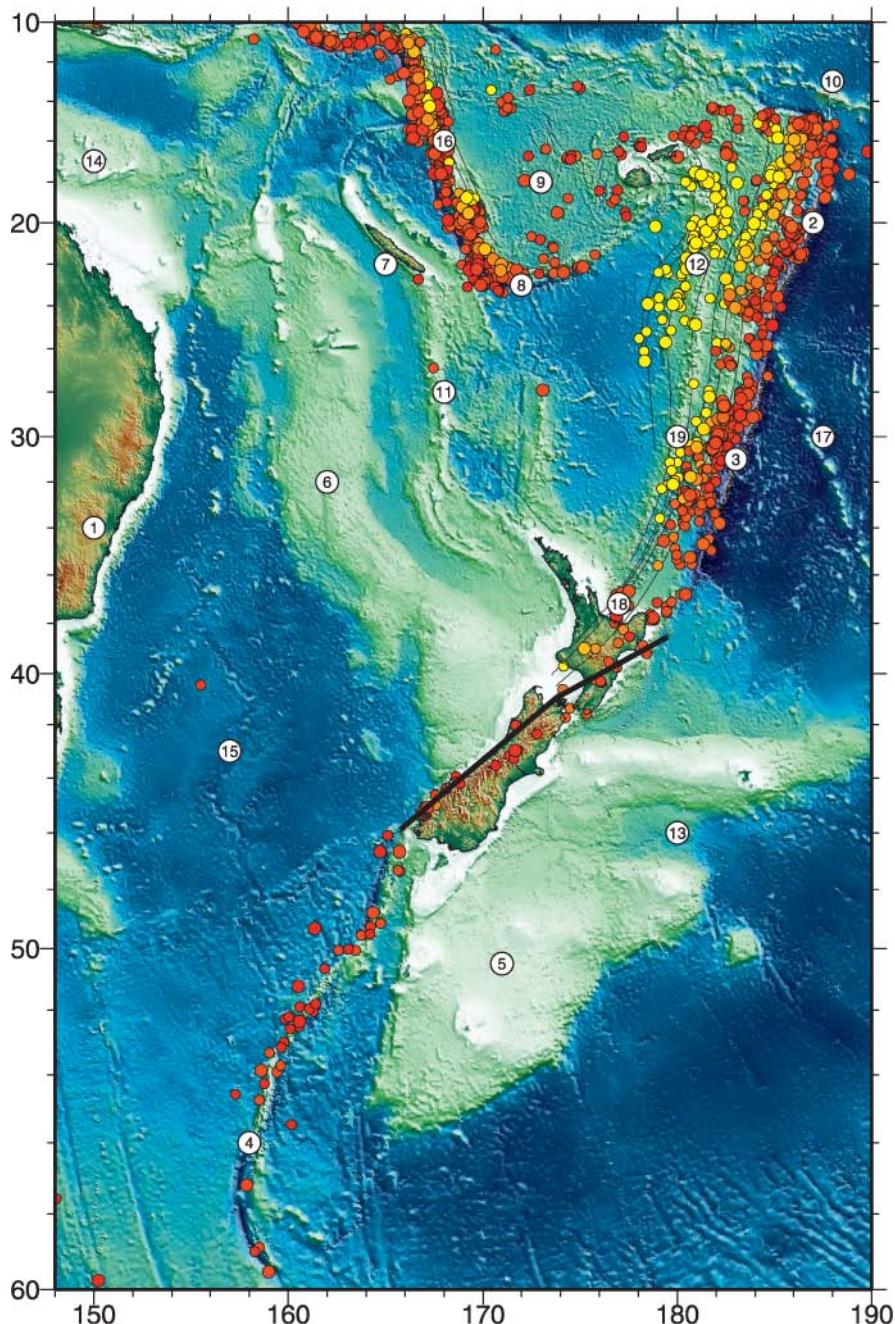
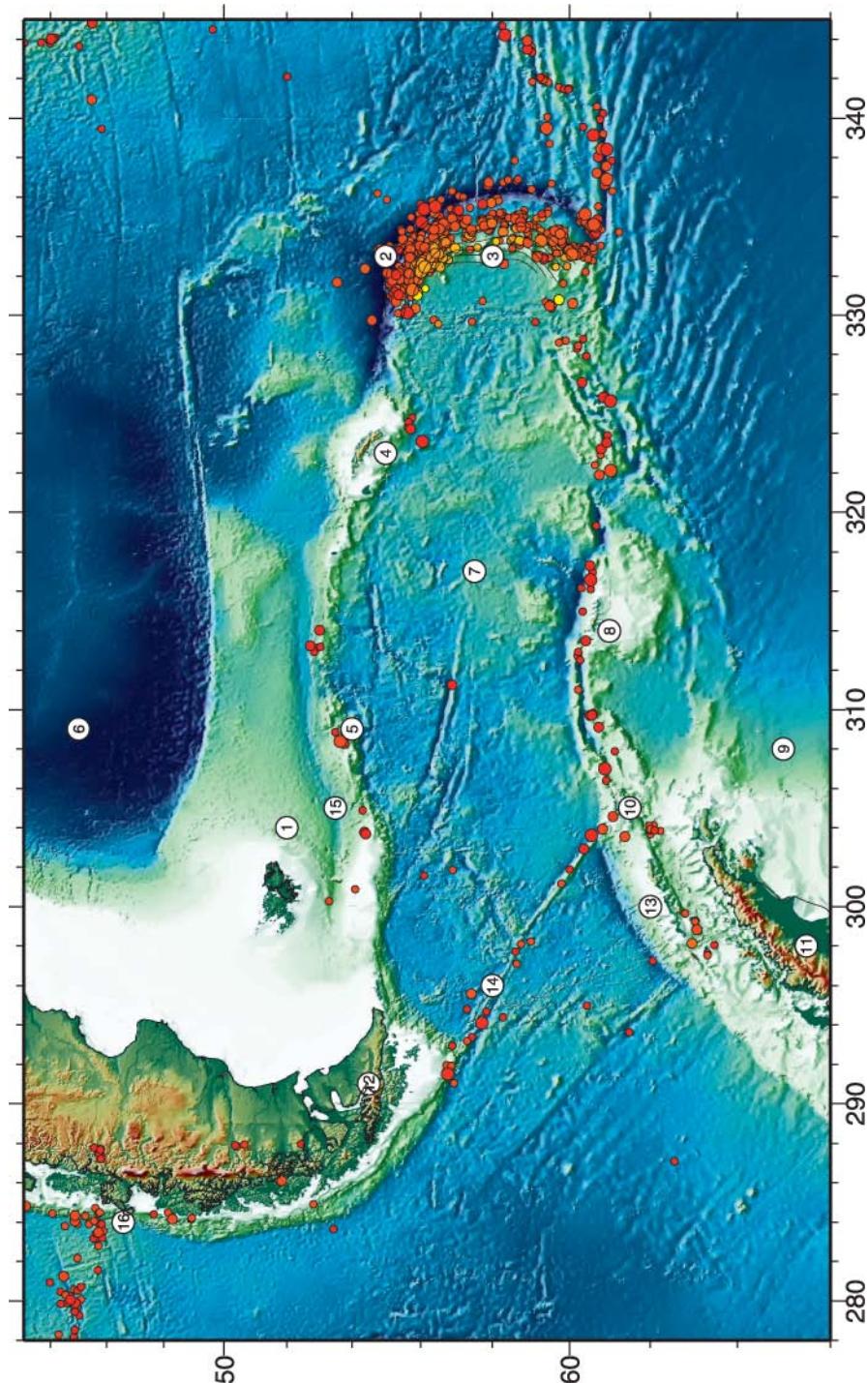
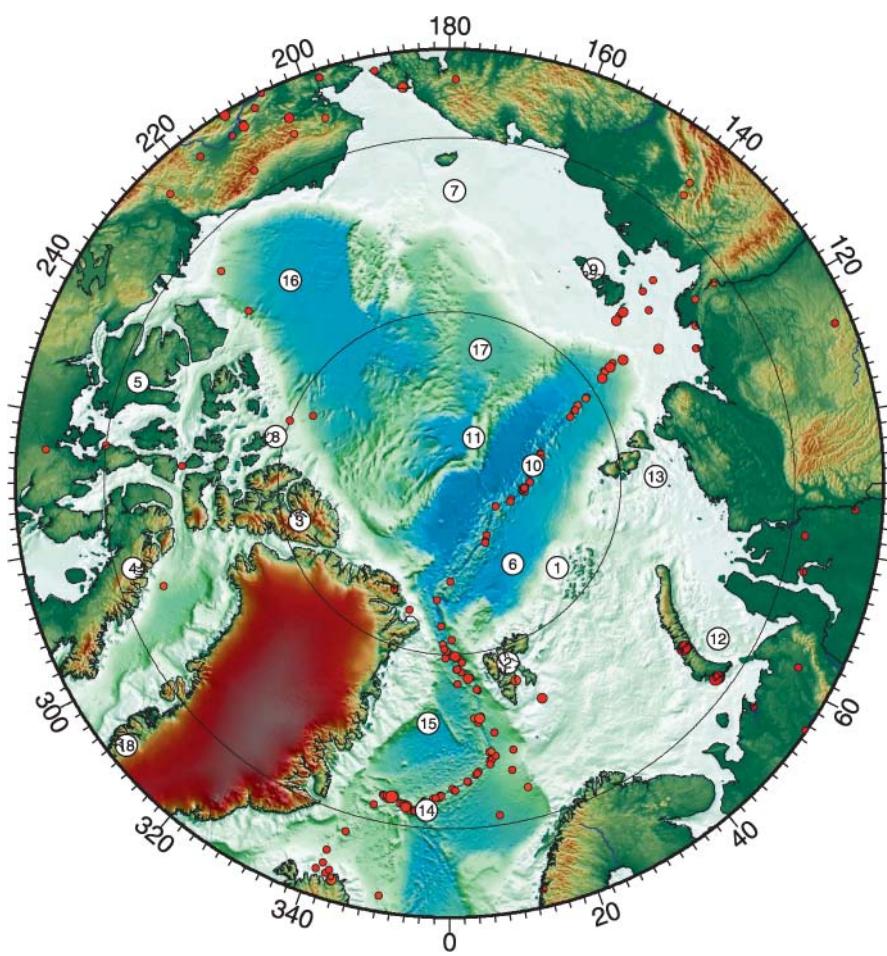


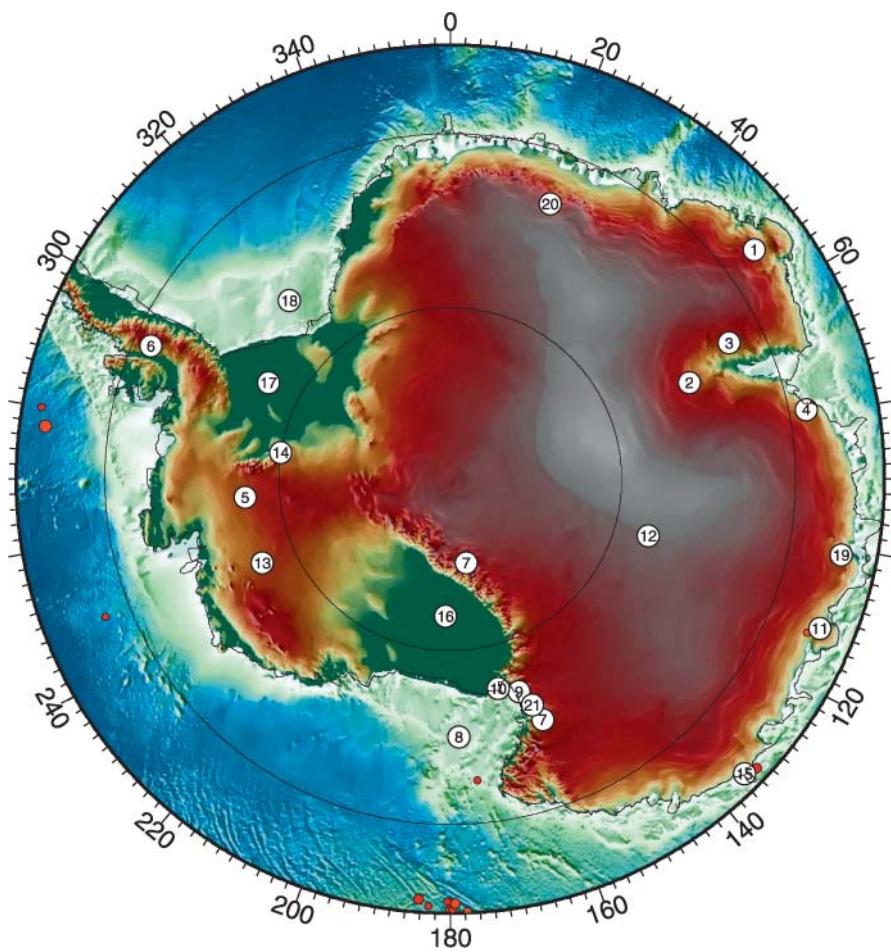
Figure A.14. Topographic map of eastern Australia and New Zealand.



**Figure A.15.** Topographic map of the Scotia Plate.



**Figure A.16.** Topographic map of the Arctic.



**Figure A.17.** Topographic map of Antarctica

## B. Mathematical Tools

Most geodynamic processes are processes that change in space and time. One of the most important tools to describe such changing processes are *differential equations*. This chapter is therefore mainly concerned with the use and interpretation of differential equations. A few selected other important numerical tricks and basic rules are summarized towards the end of this chapter.

### B.1 What is a Differential Equation?

The derivative (or: differential)  $dy/dx$  is a way to describe the *change* of  $y$  with respect to another variable  $x$ . It can be interpreted as the slope (or *gradient*) of the function  $y = f(x)$ . If the slope of this function is constant between two points along the  $x$  axis, for example between  $x$  and  $x + \Delta x$ , then we need no derivative and we can write:

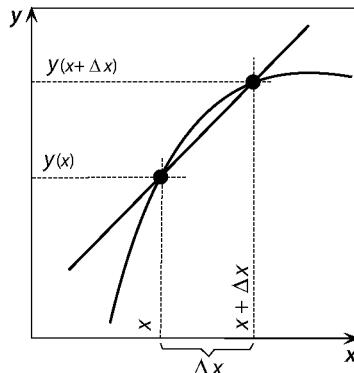
$$\text{gradient} = \frac{y(x + \Delta x) - y(x)}{\Delta x} . \quad (\text{B.1})$$

The numerator of the fraction on the right hand side of this equation is given by the difference between the  $y$  values of the function at the two points  $x$  and  $x + \Delta x$ . The denominator is given by the distance between the two points on the  $x$  axis between which the gradient is measured (Fig. B.1). Their ratio is the slope between the points  $x$  and  $x + \Delta x$ . If we consider a function where the slope is *not* constant between  $x$  and  $x + \Delta x$ , then eq. B.1 would give us only some mean of all the slopes of this function between the two points. However, the smaller we choose our  $\Delta x$ , the better will eq. B.1 describe the exact slope at point  $x$ . We can write:

$$y'(x) = \frac{dy}{dx} = \lim \Delta x \rightarrow 0 \left( \frac{y(x + \Delta x) - y(x)}{\Delta x} \right) . \quad (\text{B.2})$$

Eq. B.2 is the mathematical definition of a derivative or slope of a function. Note that we used a dash to indicate that  $y'$  is a derivative. This is a commonly used notation. The slope of a mountain road is a clear example to illustrate the meaning of *slope*. Assume that  $H$  describes the elevation of the

**Figure B.1.** Diagram illustrating the definition of the first derivative of a function (thick curve) (eq. B.1 and B.2). The slope of the thick drawn straight line may be accurately described by the ratio  $(y(x + \Delta x) - y(x))/\Delta x$ . The slope of the curved function at  $x$  is only moderately well approximated by this ratio. However, the smaller  $\Delta x$  becomes, the better this approximation will be to describe the slope at a single point



road surface as a function of distance from the valley  $x$ :  $H = f(x)$ . Then, the slope of the road is given by the first derivative of this function  $f'_x$ :

$$f'_x = \frac{dH}{dx} . \quad (\text{B.3})$$

The units of this differential are  $\text{m m}^{-1}$ . In short, it is dimensionless for this example. Familiar geological examples described by first derivatives (first differentials) of functions are the geothermal gradient, describing the change of temperature with depth (in  $^{\circ}\text{C m}^{-1}$ ), or cooling histories of metamorphic terrains that describe the change of temperature over time (in  $^{\circ}\text{C s}^{-1}$ ).

The *second derivative* (or second differential) of a function describes how the slope changes. The more familiar name of the second derivative is *curvature*. It is often abbreviated with  $f''_x$ . In our example of a mountain road, the vertical curvature of the road is:

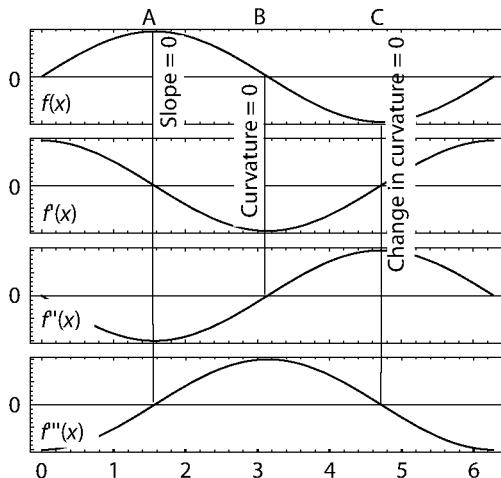
$$f''_x = \frac{d(\frac{dH}{dx})}{dx} = \frac{d^2H}{dx^2} \quad (\text{B.4})$$

and has the units of  $\text{m}/(\text{m m}^{-2})$ , which is:  $\text{m}^{-1}$ . It should be read as: “*d two H over dx square*”. The scheme we have followed to go from first to second derivative may be followed to describe the third, fourth or higher derivatives of functions. Corresponding to the first two derivatives, the third derivative of a function describes the *change* of the curvature of the function and the fourth the curvature of the curvature and so on (Fig. B.2):

$$f'''_x = \frac{d^3H}{dx^3} \quad \text{and :} \quad f''''_x = \frac{d^4H}{dx^4} . \quad (\text{B.5})$$

In our example of a mountain road, the units of the third and fourth derivatives are  $\text{m}^{-2}$  and  $\text{m}^{-3}$ , respectively. Be careful not to confuse these *linear* derivatives of the third and fourth order (in eq. B.5) with the *non-linear* first order derivatives  $(dH/dx)^3$  and  $(dH/dx)^4$ . Fortunately there are very few geological problems where more than 2nd order differentials are needed (the flexure equation (eq. 4.48) in elasticity theory is just about the only one).

**Figure B.2.** A sine function as an example for the function  $f_x$  and its first, second and third derivatives. At the maximum of the function (point A), the slope of the function is  $f'_x = 0$  and the curvature has a negative maximum:  $f''_x = -1$ . Conversely, at the inflection point of the curve (point B), the slope has a minimum the curvature is zero  $f''_x = 0$ . At the minimum of the function (point C), the slope is also zero and the curvature has a positive maximum



### B.1.1 Terminology Used in Differential Calculus

**Order.** The order of the highest derivative in a differential equation is called the *order of the equation*. For example, eq. 3.57 is a first order differential equation, eq. 3.6 or eq. 3.56 are second order differential equation and eq. 4.48 is a fourth order differential equation. Fig. B.2 illustrates the meaning of derivatives of higher order using the example of a simple sine function.

**Partial and Total Derivatives.** A function may have several variables. For example, the elevation of a point  $H$  on the surface of the earth can be described as a function of two spatial coordinates in the horizontal directions  $x$  and  $y$ , but it may also be a function of time or any other variables that we deem of importance, say vegetation or lithology. Let us consider the spatial dependence only. Then, using a simple hypothetical function to describe topography, elevation might be described by:

$$H = 3x^2 + 4y^3 \quad . \quad (\text{B.6})$$

Such analytic descriptions are great, because we can use it to determine parameters like slope directly from it. If we differentiate this function with respect to one of the variables only, then this is called a *partial derivative*. It describes the slope of the function in *one* spatial direction only. When forming a partial derivative with respect to one variable, then all other variables are kept constant during the process and are treated like any other constant of the equation. The symbol for the partial differential is  $\partial$  (say: “del”). However, “del” is no real Greek letter and should not be confused with  $\delta$ . The partial derivative of eq. B.6 after  $x$  is:

$$\frac{\partial H}{\partial x} \Big|_{y=const.} = 6x \quad . \quad (\text{B.7})$$

This partial derivative describes the slope of the function in direction  $x$ . The *total differential* is the sum of all partial derivatives. In our example it is:

$$dH = \left( \frac{\partial H}{\partial x} \right)_{y=const.} \times dx + \left( \frac{\partial H}{\partial y} \right)_{x=const.} \times dy . \quad (B.8)$$

This gives the total derivative of the function from eq. B.6 to be:

$$dH = 6x + 12y^2 . \quad (B.9)$$

This total derivative describes the tangential plane at the point  $x, y$ . In our example of surface elevation, eq. B.9 may also be interpreted as the slope of the surface in dip direction. Partial and total derivative are identical if the function contains only one variable. Partial derivatives after time are often called *rates* and symbolized by a little dot above the variable concerned (e. g. strain:  $\epsilon$  and strain rate:  $\partial\epsilon/\partial t = \dot{\epsilon}$ ). Differential equations that contain only total derivatives are called *ordinary* differential equations, in contrast to the *partial* differential equations, which contain partial derivatives. A good summary of these simple definitions may be found in Anderson and Crerar (1993) or Zill (1986), as well as many other mathematical hand books.

**Linear and Non-linear Differential Equations.** A differential equation is said to be linear if it is characterized by two properties: 1.) the dependent variable and all its derivatives are of the first degree, that is, the power of each term involving it is 1; and 2.) each coefficient depends only on the independent variable. In this context, the *dependent* variable is generally the one in the *numerator* of the derivative, while the variable in the *denominator* is called the *independent* variable. For example, in the differential equations in chapter 3, temperature,  $T$  or heat  $H$  are generally used as the dependent variables, while the spatial coordinates or time are generally the independent variables. Just about all differential equations in this book are *linear* differential equations (e.g. eqs. 3.6, 3.56 or 4.48). On the other hand, hypothetical examples of non-linear differential equations would be:

$$T \frac{\partial^2 T}{\partial x^2} - 5 \frac{\partial T}{\partial y} = x \quad \text{or:} \quad \frac{d^3 T}{dz^3} + T^2 = 0 . \quad (B.10)$$

These are examples of non-linear second and third order partial differential equations, respectively. The first example is non-linear because the coefficient of the first term depends on  $T$ , the second because  $T$  occurs in the power of 2. The only geologically relevant example of a non-linear differential equation discussed in this book occurs on p. 186.

**Analytical and Numerical Solutions.** In order to make use of a differential equation we must solve it. Only then, they can be used as a tool to extract numbers that describe some process. There are two fundamentally different ways to solve them.

- 1. *Analytical solutions.* Analytical or closed solutions of differential equations may be found by integrating them. Let us consider as an example the description of a geotherm by

$$\frac{dT}{dz} = \frac{1.5}{\sqrt{z}} . \quad (\text{B.11})$$

There,  $T$  is temperature in °C and  $z$  is depth. This differential equation can be integrated without difficulty:

$$T = 3\sqrt{z} + C . \quad (\text{B.12})$$

The integration constant  $C$  must be determined using boundary conditions. Eq. B.12 is said to be an “analytical solution of the differential equation eq. B.11”. If we assume (as our boundary condition) that the temperature at the earths surface is always zero and we assume a coordinate system where the surface is at  $z = 0$ , then this constant must be also zero:  $C = 0$ . Now eq. B.12 can be used to calculate temperatures at any depth of our choice by inserting numbers for  $z$ . For example, for  $z = 100\,000$  m eq. B.12 gives  $T = 949$  °C.

- 2. *Numerical solutions.* Numerical solutions of differential equations are used to extract numbers from differential equations *without* having to solve (integrate) them. With their aid we can arrive at the result that eq. B.11 describes a temperature of  $T = 949$  °C at 100 km depth without having to solve the differential equation, i.e. without having to go from eq. B.11 to eq. B.12. Great – results without having to solve the problem! However, there is nothing such as a free lunch: numerical solutions are not exact. Numerical approximations are always approximations and they are plagued by stability and accuracy problems (s. p. 401). The numerical solution of partial differential equations is a science on its own (sect. B.2, B.4). The two most important methods that are in use are *finite difference methods* and *finite element methods*.

The *finite element method* has the advantage that it is much more elegant to use it for the description of deformation on Lagrangian coordinates. The principal disadvantage of the finite element method is that it is not very intuitive and therefore requires quite an initial effort to learn it.

The *finite difference method* has the enormous advantage that it is quite intuitive, easy to implement on a computer (even by inexperienced mathematicians) and easily adaptable to many different problems. Its principal problems are those of instability, and that they are quite cumbersome when it comes to the treatment of discontinuous boundary conditions and deformed grids (sect. B.2).

- *Advantages and disadvantages.* Numerical and analytical solutions have both their advantages and disadvantages. The enormous advantage of numerical solutions is that they allow us to arrive at results without having

to know enough differential calculus to be able to integrate the equation in question. In fact many geological problems can be simplified enough to be able to formulate them into an equation, but are too complicated so that an analytical solution can be found. In such cases, numerical solutions are the only way to obtain results.

Analytical solutions have the advantage that they are much more useful to understand the nature of a geological process. For example, eq. B.12 may be used directly to infer that the temperature in the crust rises with the square root of depth. If this model corresponds well with our observations in nature, then we can continue to think about the significance of this quadratic relationship. Such considerations are difficult with numerical solutions as they only deliver numbers.

### Initial- and Boundary Conditions.

• *Boundary conditions.* When solving differential equations, *boundary conditions* are necessary in order to determine the integration constants. This is true for both numerical and analytical solutions. For differential equations of the first order we need one boundary condition, for those of the second order two and so on. The term *boundary condition* is exactly what it implies: it is a condition at the boundary of the model (s. sect. B.2.4). The most common types of boundary conditions are:

- A prescribed value of the function at the model boundary (e.g.  $T = 0$  at  $z = 0$ ; s. eq. B.12),
- Neumann boundary condition: A prescribed gradient of the function at the model boundary (e.g. the heat flow boundary condition we used in sect. 3.4.3),
- A functional relationship between *value* and *gradient* at the model boundary (e.g. the constant heat content boundary condition used on p. 123).

Boundary conditions given by higher derivatives of functions are also possible and play an important role when integrating differential equations of higher orders (s. sect. 4.4.2, eq. 4.48). In sect. B.2 we discuss how some of these boundary conditions may be implemented.

• *Initial conditions.* *Initial conditions* are necessary to determine the starting point of a model. For example, if we want to use the diffusion equation (eq. 3.6) to calculate the evolution of a diffusive zoning profile over time, then we must use a function  $T = f(z)$  at the time  $t = 0$  from which we can start calculating. The nature of this function  $T = f(z)$  must be determined by a known initial condition. Steady state problems usually do not require an initial condition.

## B.2 The Finite Difference Method

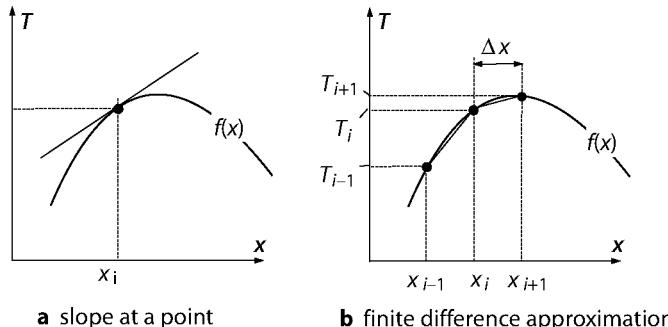
The finite difference method makes use of the discretization of the derivative from eq. B.2. Instead of describing the differential  $dy/dx$  by the limiting value  $\Delta x \rightarrow 0$ , a finite value of  $\Delta x$  is used (i.e. eq. B.1 is retained). For our explanation on the next pages we use Fig. B.3 showing the function  $T = f(x)$  and assume that this function is a temperature profile across a metamorphic terrain along the spatial axis  $x$ . Thus, we will use the variable  $T$  instead of the more abstract  $y$  that we have used up to now in this chapter. At the point  $x_i$  (labeled in Fig. B.3a by the dotted line) the function has the slope  $dT/dx$ . When using the method of finite differences, this slope is approximated by the discrete temperature difference at two different places with a *finite* distance to each other (a bit as we have already implicitly shown in Fig. B.1). There is many ways to formulate such a difference. From Fig. B.3b we can see that one way to formulate such a difference is:

$$\frac{dT}{dx} \approx \frac{T_{i+1} - T_i}{x_{i+1} - x_i} = \frac{T_{i+1} - T_i}{\Delta x} . \quad (\text{B.13})$$

The index  $i$  is just a description of the number of the grid point chosen here.  $T_i$  is the temperature at the “ $i^{\text{th}}$ ” point of a discrete grid of points.  $T_{i+1}$  is the temperature at the *next* point of the grid,  $T_{i-1}$  at the *previous* point. The finite difference method used in eq. B.13 is called *forward differencing method* as we have calculated the temperature gradient at  $x_i$  using the temperature at  $x_i$  as well as the temperature at the next *forward* point on the grid (Fig. B.4). Some other simple examples of differencing schemes have the form:

$$\frac{dT}{dx} \approx \frac{T_i - T_{i-1}}{\Delta x} \quad \text{or} \quad \frac{dT}{dx} \approx \frac{T_{i+1} - T_{i-1}}{2\Delta x} . \quad (\text{B.14})$$

For reasons that should now be obvious, these two methods are called *backward differencing* and *central differencing* schemes (Fig. B.4).



**Figure B.3.** Graphical illustration of the method of finite differences. In a the slope of the function  $f(x)$  at point  $x_i$  is accurately described by the touching tangent. Mathematically this slope is described by the differential  $dT/dx$ . In b the slope is approximated by the ratio of the differences of two temperature and two  $x$  values

- *Differentiating with respect to time.* All information we have discussed so far is generally applicable, regardless of what variable is described by  $x$ ,  $y$  or  $T$ . However, the use of  $x$  has suggested that we imply *spatial* differentials. In order to discriminate between the numbering of grid nodes of spatial and temporal grids, the symbols “+” and “-” are common to describe the *next time step* and the *previous time step* while  $i$  and  $i + 1$  is used for the spatial grid stepping. Thus we can write:

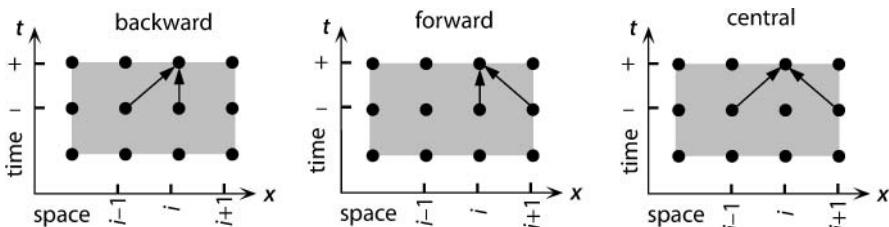
$$\frac{dT}{dt} \approx \frac{T^+ - T^-}{\Delta t} . \quad (\text{B.15})$$

In some books “ $j$ ” and “ $j + 1$ ” are used to denote time steps. However, this should not be confused with spatially two-dimensional problems in which “ $i$ ” subscripts are used for grid numbering in  $x$  direction and “ $j$ ” numbering of grid steps in  $y$  direction.

- *Approximations of derivatives of higher order.* For the approximation of derivatives of the second or higher order we can use the same scheme as that for the first derivative (eq. B.13, B.14). For the second derivative we must form the ratio of the difference in slope at two different grid points with the distance  $\Delta x$ :

$$\frac{d^2T}{dx^2} = \frac{d\left(\frac{dT}{dx}\right)}{dx} \approx \frac{\left(\frac{T_{i+1} - T_i}{\Delta x}\right) - \left(\frac{T_i - T_{i-1}}{\Delta x}\right)}{\Delta x} = \frac{T_{i+1} - 2T_i + T_{i-1}}{\Delta x^2} . \quad (\text{B.16})$$

From eq. B.16 we can see that, in order to formulate the difference between slopes at two point, the slope at point  $i$  was approximated once by forward differencing and once by backward differencing. This is necessary, as we want to calculate the curvature at point  $i$  from the differences between the slopes of the curve as near as possible to it (i.e. in front of it and behind it). We can see that the curvature is described by the difference of slopes, just like we describe the slope by the differences between two function values.



**Figure B.4.** Schematic illustration of three simple methods of discretization in the finite difference method. The  $x$  axis of each diagram shows four discrete points of a one-dimensional spatial grid. Each dot is a temperature value at this point in space. The  $y$ -axis shows three different time steps of the calculation. In the three diagrams, the temperature at the third grid point (labeled by subscript  $i$ ) is calculated by *backward*, *forward* and *central* differencing. In each diagram this calculation is for the values at the (as yet) unknown time step “+” from known information at time “-”

- *Solution of the diffusion equation using finite differences.* If differential equations contain more than one variable (e. g. the diffusion equation (eq. 3.6), in which both spatial and temporal derivatives occur) it is necessary to combine several indices with each other. This may lead to apparently quite complicated formulations. Here we will follow the use of temporal and spatial indices as discussed above, i.e. the temperature at the spatial grid point  $i$  at a future time step is called  $T_i^+$  and at a previous time step it is called  $T_i^-$  (Fig. B.4). Once we are familiar with this notation it should be straight forward to understand the following finite difference approximation of the diffusion equation (eq. 3.6) by using eqs. B.15 and B.16:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial x^2} \approx \frac{T_i^+ - T_i^-}{\Delta t} = \kappa \frac{T_{i+1}^- - 2T_i^- + T_{i-1}^-}{\Delta x^2} . \quad (\text{B.17})$$

Solved for  $T$  at the new time step of interest this gives:

$$T_i^+ = T_i^- + \left( \frac{\kappa \Delta t}{\Delta x^2} \right) (T_{i+1}^- - 2T_i^- + T_{i-1}^-) . \quad (\text{B.18})$$

Now we can insert into eq. B.18 known temperatures at known points in space from a previous time step (starting from the known temperature profile of our initial condition) to determine the temporal evolution of the temperature profile.

The last few pages have given you an overview over the basic principles of the finite difference method. All other finite difference approximations are refinements of the above aiming at higher accuracy, higher stability and higher speeds of calculation on computers.

Why better methods may be necessary can be illustrated with eq. B.18. There, the magnitude of the constant  $(\kappa \Delta t / \Delta x^2)$  is critical for the stability and accuracy of the approximation (sect. B.2.1). The magnitude of this term is so important so that this term has its own name. It is called the *Fourier cell number* and must be smaller than 0.25 so that the solution of eq. B.18 retains stability. Since the magnitude of  $\kappa$  depends on material constants, we cannot change it arbitrarily. Thus, in order to fulfill the stability criterion we must make a corresponding choice with the time and space stepping. With many simple finite difference approximations this leads to insurmountable problems: If a given problem requires high spatial resolution (small  $\Delta x$ ) it requires a correspondingly small choice of  $\Delta t$ . However, if this problem should now be solved over long geological times, then we may have to iterate through too many time steps for the problem to be solvable in realistic computer time. This is one of the reasons why large computers (and more refined finite difference methods) are required for many geological questions.

### B.2.1 Stability and Accuracy

Finite difference solutions of differential equations have two important disadvantages:

1. They are only approximations.
2. They are often unstable.

Criteria for accuracy and stability are extensively discussed in the literature (e. g. Smith 1985; Fletcher 1991; Anderson et al. 1984). However, both problems can be reduced to a minimum by some very simple checks:

- *Accuracy.* The accuracy of finite difference approximations can easily be checked by successively decreasing the time or spatial stepping (for a discussion of accuracy versus precision s. p. 5). If the result does not change, the exact solution has probably been approximated well enough. A second test can be performed by simplifying the initial and boundary conditions of a giving problem enough so that analytical integration of the descriptive equations is possible. Then the numerical solution may be compared directly with the analytical results. Time and space stepping can then be relaxed and finally the initial and boundary conditions readjusted to describe the problem in the required detail.
- *Stability.* A finite difference solution is called *stable* if it converges to the correct solution. Unstable solutions diverge with progressive calculation more and more. Most unstable solutions “explode” within a few time steps. Thus, stability problems are often relatively easy to recognize as all functional values trend towards infinity (s. Fig. B.7). Stability problems can often be brought under control by decreasing the discrete stepping in the approximation.

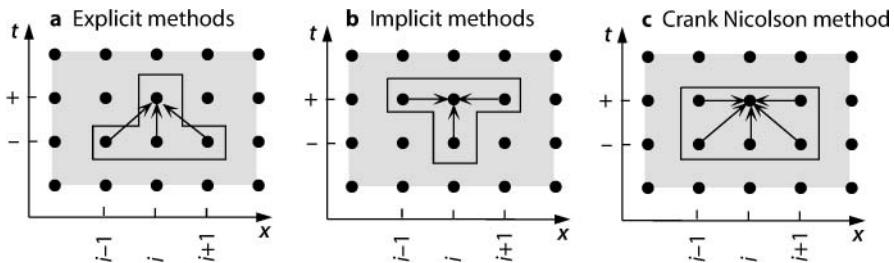
### B.2.2 Implicit and Explicit Finite Difference Methods

There are two fundamentally different types of finite difference methods that may be used to solve (approximate) differential equations:

1. explicit methods,
2. implicit methods.

There is also *mixed* methods that are partially implicit and partially explicit. Fig. B.5 illustrates what is meant with *implicit* and *explicit*. Both methods will be discussed briefly below using the example of temperature calculation with the diffusion equation. However, the principal difference between implicit and explicit solutions are the same regardless of the variables or the equations.

- *Explicit methods.* The idea behind explicit finite difference methods is illustrated in Fig. B.5a. This figure corresponds to the way the diffusion equation was solved in eq. B.18 (Fig. B.4). It may be seen that the temperature at point  $i$  at the new time step  $T_i^+$  is calculated from the *known* temperatures (those from the previous time step) at the points  $i - 1$ ,  $i$  and  $i + 1$ . As these temperatures ( $T_i^-, T_{i-1}^-$  and  $T_{i+1}^-$ ) are known, the application of eq. B.18 is no problem. All methods that use schemes where *new* information is calculated exclusively from known information are called *explicit* methods.



**Figure B.5.** Examples for implicit and explicit finite difference methods. The vertical axis is time, the horizontal space. Time step “+” is the time step to be calculated. Time step “-” denotes the time for which information is already available. The *Crank Nicolson* method is a mixed method. It consists of implicit and explicit parts

- *Implicit methods.* Implicit finite difference methods calculate the unknown temperatures  $T_i^-$  from other unknown temperatures at the same time step (Fig. B.5b). This sounds a bit counter intuitive if not impossible, but is possible if all temperatures are calculated *simultaneously*. Remember that we have boundary conditions that tell us the new temperatures at the two ends of the grid. Thus, in a grid with  $n$  points, there is only  $n - 2$  points where the temperature is unknown. It is therefore possible to formulate a set of  $n - 1$  equations with  $n - 2$  unknowns. This may be solved for all unknown variables. An example of an implicit approximation of eq. 3.6, (corresponding to Fig. B.5b) is:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial x^2} \approx \frac{T_i^+ - T_i^-}{\Delta t} = \kappa \frac{T_{i+1}^+ - 2T_i^+ + T_{i-1}^+}{\Delta x^2} . \quad (\text{B.19})$$

Solved for the temperature of interest this is:

$$T_i^+ = \frac{T_i^- + R(T_{i+1}^+ + T_{i-1}^+)}{1 + 2R} \quad \text{where :} \quad R = \frac{\kappa \Delta t}{\Delta x^2} . \quad (\text{B.20})$$

- *Mixed methods.* Mixed methods use explicit as well as implicit information to calculate the new data (Fig. B.5c). Mixed methods have the best accuracy and stability characteristics and are therefore commonly used. The most famous of all mixed methods is the *Crank-Nicolson-method* which is used to describe second order differentials, as they occur in the diffusion equation. The Crank Nicolson method describes this with:

$$\frac{T_i^+ - T_i^-}{\Delta t} = \frac{\kappa}{2} \left( \frac{T_{i+1}^+ - 2T_i^+ + T_{i-1}^+}{\Delta x^2} + \frac{T_{i+1}^- - 2T_i^- + T_{i-1}^-}{\Delta x^2} \right) . \quad (\text{B.21})$$

It may be seen that the expression inside the brackets is the sum of the right hand sides of eq. B.17 and eq. B.20 and that the mean of these expressions is formed. Ways to implement eq. B.21 are discussed in most books on numerical

mathematics. The *Thomas algorithm* is an elegant method that can be used to implement the simultaneous solution of equations as is necessary to solve eq. B.21.

**Two-dimensional Derivatives.** In the following we illustrate the approximation of two-dimensional differential equations using the two-dimensional form of the heat conduction equation as an example. This equation is:

$$\frac{\partial T}{\partial t} = \kappa \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) . \quad (\text{B.22})$$

The most commonly used finite difference method to approximate eq. B.22 is the *Alternating Direction Implicit Method*, in short: ADI method. In this method, each time step is divided into two. The first half time step is *explicit* in one spatial direction and *implicit* in the other, the second half time step is done in reverse. Thus, for each step in time, two steps of calculation are required. The ADI method has a lot of similarities with the Crank Nicolson method. A discrete version of eq. B.22 looks like this:

1. step:

$$\begin{aligned} \frac{T_{i,j}^{+/2} - T_{i,j}^-}{\Delta t/2} &= \frac{\kappa}{\Delta x^2} (T_{i+1,j}^{+/2} - 2T_{i,j}^{+/2} + T_{i-1,j}^{+/2}) \\ &\quad + \frac{\kappa}{\Delta y^2} (T_{i,j+1}^- - 2T_{i,j}^- + T_{i,j-1}^-) . \end{aligned}$$

2. step:

$$\begin{aligned} \frac{T_{i,j}^+ - T_{i,j}^{+/2}}{\Delta t/2} &= \frac{\kappa}{\Delta x^2} (T_{i+1,j}^{+/2} - 2T_{i,j}^{+/2} + T_{i-1,j}^{+/2}) \\ &\quad + \frac{\kappa}{\Delta y^2} (T_{i,j+1}^+ - 2T_{i,j}^+ + T_{i,j-1}^+) . \end{aligned} \quad (\text{B.23})$$

In this equation we have used the subscripts  $i$  and  $j$  to label the grid points in the  $x$ - and  $y$ -directions and we used “ $+/2$ ” as a description of half a time step.

### B.2.3 Approximation of the Transport Equation

In sect. 3.3 we introduced a simple equation that can be used to describe the advection of material, for example the advection of rocks to the surface by erosion or the advection of fluids through a marble (eq. 3.40). Although it looks much more simple than the diffusion equation, it is not. In fact, the equation requires the use of very different finite difference approximations from those discussed above for the diffusion equation. For example, *backward* and *central* differencing schemes are unstable when applied to approximate

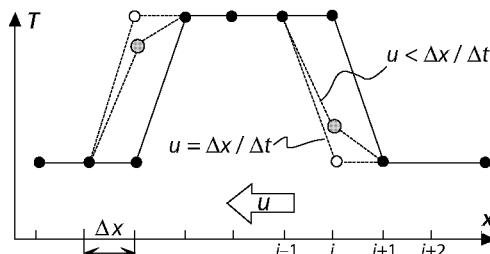
eq. 3.40. Thus *forward* differencing schemes must be used. A simple forward differencing scheme that can be used to approximate this transport equation is:

$$T^+ = T^- + u \frac{\Delta t}{\Delta x} (T_{i+1}^- - T_i^-) \quad . \quad (\text{B.24})$$

However, eq. B.24 is associated with *numerical diffusion*, a problem which makes the solution increasingly inaccurate if used over many time steps.

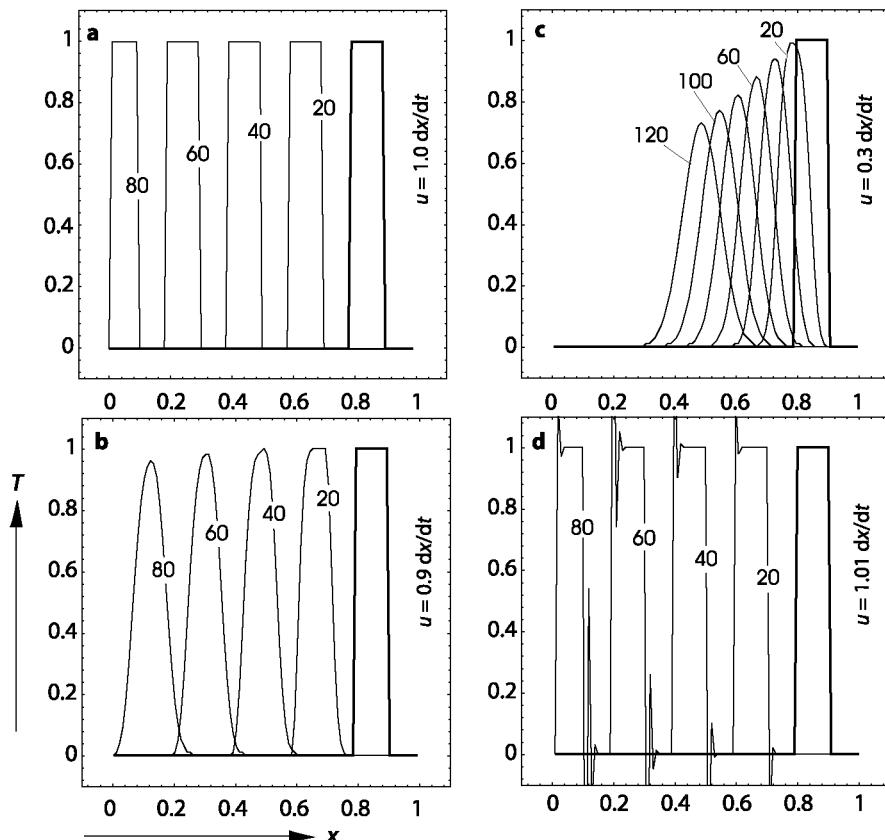
**The Problem of Numerical Diffusion.** The description of advection with forward finite differencing schemes of the transport equation causes a phenomenon known by the name *numerical diffusion*. During numerical diffusion progressive steps of the calculation cause a rounding of parts of the function that are strongly curved. This rounding is akin to the smoothing of curves by diffusion, but is only an artifact of the numerical approximation and has nothing to do with any real diffusion process. Fig. B.6 illustrates why this numerical diffusion occurs. There, the original temperature profile is shown by the continuous line. The black dots label discrete values of this function. Using eq. B.24 and positive values for  $u$  caused advection (motion) of the temperature profile towards the origin (against the  $x$  direction). If the transport velocity is exactly  $u = \Delta x / \Delta t$ , then the temperature profile will be moved by exactly one grid node with every time step (dashed line and white dots). If, however, the transport rate is  $u < \Delta x / \Delta t$  then the temperature profile is shifted accordingly, but also rounded off (dotted line and gray dots). If we look at eq. B.24 carefully we can see that the numerical diffusion occurs because of interpolations between grid points when the transport is not a full grid interval per time step.

Unfortunately, many geological transport rates are too small so that the number of time and space steps required to fulfill  $u = \Delta x / \Delta t$  is much too large to be sensibly used. It is often necessary to use a space stepping that is



**Figure B.6.** Illustration of the origin of numerical diffusion in the one-dimensional transport equation (eq. 3.40). The black dots mark discrete values of the temperature profile shown by the continuous line. The dashed line (and white dots) shows the temperature profile as calculated with eq. B.24 after one  $\Delta t$  time step if the transport velocity  $u$  is exactly  $u = \Delta x / \Delta t$ . It may be seen that the original temperature profile is advected towards the left without a change in shape. However, if  $u < \Delta x / \Delta t$ , then the temperature profile after one  $\Delta t$ , is not exactly like the original profile (dotted line, shaded dots). It has been victim to *numerical diffusion*.

$\Delta x > u\Delta t$ . As a consequence, we often have to deal with numerical diffusion when describing the transport of rocks using eq. 3.40. Fig. B.7b,c illustrates how much information of a step-shaped temperature profile is lost, if the transport rate is 90 % or 30 % of  $u = \Delta x/\Delta t$ . It may be seen that after about 50 time steps the shape of the initial and the final temperature profile have little in common. Fig. B.7d illustrates how eq. B.24 becomes unstable if the



**Figure B.7.** Graphical illustration of numerical diffusion and instabilities using the example of the advection of a step shaped temperature profile originally at the position indicated by the thick line. The advection of a step shaped temperature profile is shown as described with the finite difference approximation of eq. 3.40 given in eq. B.24. (However, note that we use  $x$  instead of  $z$  for the spatial coordinate here).  $u$  is positive so that advection is against the  $x$ -direction. The grid consists of 100 points between 0 and 1 in the  $x$ -direction. Labels are in numbers of time steps. **a**  $u = \Delta x/\Delta t$ ; **b**, **c**  $u < \Delta x/\Delta t$ ; **d**  $u > \Delta x/\Delta t$

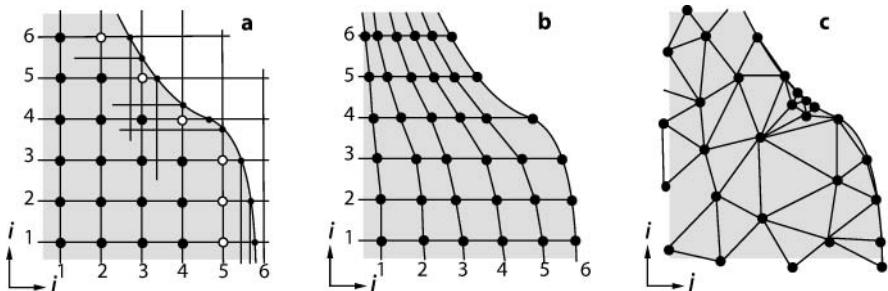
transport rate is  $u > \Delta x / \Delta t$ . The temperature profile gets jagged edges and quickly blows up meaninglessly.

### B.2.4 Grids and Boundary Conditions

If we want to use the finite difference method to solve a differential equation (e.g. eq. B.18), we need to build a discrete grid on which the function is evaluated (Fig. B.9). A regular grid with  $n$  grid points has  $n - 1$  grid spaces. If the total length of the grid is  $L$  and the grid spacing is regular, then the distance between any two grid points will be  $\Delta x = L/(n - 1)$  (Fig. B.9a). However, grid spacings need not be regular. For example, if a function is of particular interest in a special region it may be useful to make the grid especially fine in this region. On the other hand, it may not be wise to make the grid everywhere this fine as this may enlarge the time of calculation enormously. A spatially variable grid is the best solution for this. On such spatially variable grids we must substitute  $\Delta x$  by  $(x_{i+1} - x_i)$  (see: eq. B.13)

For most two dimensional models triangulated grids are used today. Such grids may be designed to be self adaptive, that is that they refine themselves as a function of the changes per time step (Fig. B.9c). Another way to deal with problems that require a high spatial resolution is the use of multigrid methods. Multigrid methods use a series of successively coarser grids for each time step.

- **Boundary conditions.** Closer consideration of eq. B.18 indicates that this equation may not be evaluated at the points  $i = 1$  and  $i = n$ , because no grid points “ $i - 1$ ” and “ $n + 1$ ” exist there for which we could insert the temperatures  $T_{i-1}$  and  $T_{i+1}$  into the right hand side of the equation. These two temperatures must be determined by the boundary conditions. These boundary conditions are equivalent to the integration limits of a definite integral that are required to determine the integration constants. Thus, it is



**Figure B.8.** Some examples for the discretization of a two-dimensional region with an irregular boundary. **a** The interpolation of the model boundary on an orthogonal grid is difficult and messy. **b** A non-orthogonal grid with a constant number of grid points in  $x$ - and  $y$ -direction. **c** The most elegant way to handle an irregular boundary is by using a triangulation of the area of interest

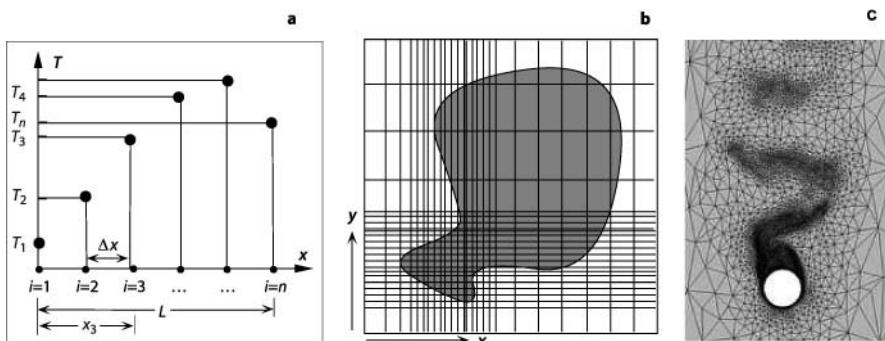
no coincidence that there is *two* grid nodes in finite difference approximations of *second* order derivatives, where the functional values can only be determined with the aid of boundary conditions.

### B.2.5 Dealing with Irregular Grid Boundaries

When dealing with two-dimensional geological problems, model boundaries are often not straight lines, but are curved. For example, if we want to use the Moho heat flow as a boundary condition and the Moho is curved underneath a mountain range. Another example was discussed in sect. 3.7.3 where the irregular boundary was given by the irregular shape of the earths surface. When dealing with such problems it is often difficult to find simple finite difference approximations of the descriptive differential equations. The type of problem that we may encounter is illustrated in Fig. B.8. This figure shows an irregular body (shaded region) for which a two-dimensional heat conduction problem is to be solved. If we want to discretize the region with a rectangular grid (as for example the ADI-method in eq. B.23 would require), then this problem can only be solved for the part of the grid highlighted by the thick black dots in Fig. B.8a. The marginal points (shown in white on Fig. B.8a) must be defined by the boundary conditions. Clearly this is highly unsatisfactory. If we want the model boundaries to correspond to the *real* boundaries of the problem, then we would need to introduce all kinds of new grid lines (thin lines on Fig. B.8a) and the grid would become irregular and the entire numerical approximation full of interpolations and, in fact, quite messy.

Such problems may be avoided with one of the two following alternatives:

1. A grid may be defined that has just as many points in the  $x$ - and  $y$  directions, but in which the grid lines are not on an orthogonal grid but



**Figure B.9.** Examples of discrete grids. **a** Discrete form of the one dimensional function from Fig. B.3 on a regular one-dimensional spatial grid. Different points are numbered from  $i = 1$  to  $i = n$ . **b** An irregular two-dimensional orthogonal grid. The grid serves the description of the dark shaded region. Thus, a finer grid spacing was used for the grid in the lower left hand portion of the grid. **c** Self adaptive grid for a model of flow past a heated cylinder (Pain pers. comm. 2005)

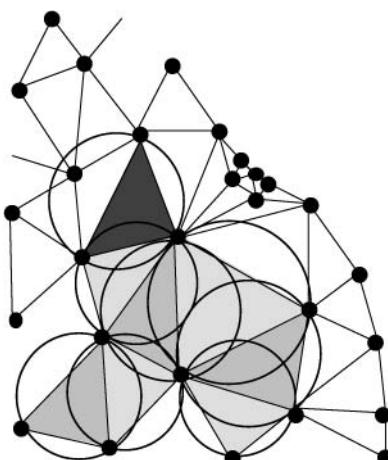
follow the problem boundaries (Fig. B.8b). Such curvilinear grids can still be dealt with using finite difference methods, but correction terms for the changes in direction at the grid nodes must be added. Calculations on such non-orthogonal grids become quite inaccurate if the direction changes at individual nodes become too large.

2. The most elegant (and common) method to deal with irregular boundaries is by using triangulations of the region of interest (Fig. B.8c). Triangles are the most simple of all geometric forms that can be used to subdivide two-dimensional regions and triangles are therefore with many respects superior to polygons with four corners. When triangulating a two-dimensional region, there are many ways to do this. For example, one could assume a certain allowed triangle size, a minimum triangle angle, or assume the number of triangles that are allowed to meet in one point. For many purposes the *Delaunay triangulation* is the best choice of *how* to triangulate a region (Fig. B.10) (e. g. Sambridge et al. 1995). The biggest disadvantage of triangulated grids is that finite difference approximations often become quite a lot more complicated than those for orthogonal grids. These days it is possible to download a large variety of mesh generators off the web that produce triangulated grids for any given geometric problem.

### B.2.6 Recommended Reading

- Anderson, Tannehill and Pletcher (1984) Computational Fluid Mechanics and Heat Transfer.
- Fletcher (1991) Computational Techniques for Fluid Dynamics Volume 1. Fundamental and General Techniques.
- Smith (1985) Numerical Solutions of Partial Differential Equations.
- Reece (1986) Microcomputer Modeling by Finite Differences.

**Figure B.10.** A section of the grid from Fig. B.8c for the illustration of the Delaunay triangulation. In a Delaunay triangulation every circle that contains all three corners of any given triangle does *not* include any other grid nodes. Thus, all shaded triangles, except the dark shaded one, are Delaunay triangles



- O’Rouke (1993) Computational geometry in C.
- Zill (1986) A First Course in Differential Equations with Applications.
- <http://mathworld.wolfram.com/FiniteDifferenceMethod.html>

## B.3 Scalars, Vectors and Tensors

- *Scalars*. Geological parameters that are described by their magnitude only are *scalar quantities*. The temperature at the Moho, the elevation of a mountain, density of a rock or pressure are examples. (According to Oertel (1996) pressure should be referred to as an *isotropic tensor of second rank* but for all intents and purposes of this book it is sufficient to treat it as a scalar). Variables that are scalar quantities are commonly denoted with italics, as most variables in this book.
- *Vectors*. Geological parameters that have both a *magnitude* and a *direction* are described by *vectors*. An example is the force with which India and Asia collide or the rate of intrusion of a magmatic body. The former is roughly  $10^{13} \text{ N m}^{-1}$  and is directed northwards; the latter might be some meters per year and directed upwards in the crust. Vectors are commonly represented by bold roman letters, although we refrain from this use in this book.
- *Tensors*. Parameters that are characterized by not only their magnitude and their direction, but also by a spatial dependence of this direction are described by *tensors*. The state of stress at a point or strain rate are two familiar examples of tensor quantities to a geologist. It is easy to see that magnitude and direction alone are insufficient to describe stress. For example, the tensor components  $\sigma_{xx}$  and  $\sigma_{yx}$  both act in the  $x$  direction and they also may both be of the same magnitude. However,  $\sigma_{xx}$  is a normal stress and  $\sigma_{xy}$  is a shear stress, i.e. they are exerted onto planes of different orientation. Tensors are represented as matrices and are commonly abbreviated with italics. As for vectors, we do not use this notation in the present book as the tensorial quantities occurring herein (e.g. strain rate or stress) are usually simplified enough so that they reduce to simple scalar quantities (e.g. by considering one-dimensional cases only).

Scalars, vectors and tensors are often called tensorial quantities of the 0., 1. and 2nd rank. In this book we treat many quantities that are actually described by vectors or tensors as if they were scalars (except in sect. 5.1.1). We have done so by making our problems so simple, so that they may be treated one-dimensionally. In fact, it is better to call them “pseudoscalars” because it is always implicit that their direction is known. Regardless, we introduce some of the basic principals of vector calculations on the following pages.

• *Common confusions.* Quantities described by scalars, vectors and tensors are often confused in the literature. Even in this book – while we try not to confuse them – we often treat tensorial quantities as if they were scalars. For a correct consideration of two and three-dimensional problems, the full tensor quantities must be considered. Products and sums of tensors are *not* described by the sum of the one-dimensional descriptions in several spatial directions alone (e.g. Oertel 1996; Strang 1988) and it is therefore often not trivial to understand the results of two- and three-dimensional models in comparison to their one-dimensional equivalents.

**Basics of Vector Algebra.** Vectors describe *direction* and *magnitude* of a parameter. Thus, in Cartesian coordinates, they are described by three components:

$$\mathbf{u} = u_x \mathbf{i} + u_y \mathbf{j} + u_z \mathbf{k} . \quad (\text{B.25})$$

$u_x$ ,  $u_y$  and  $u_z$  are called the vector components of the vector  $\mathbf{u}$  and  $\mathbf{i}$ ,  $\mathbf{j}$  and  $\mathbf{k}$  are called the unit vectors in the three orthogonal spatial directions. The unit vectors are often omitted and vectors are usually just written as a list of three scalar components. In the literature, these are variably named  $u_x$ ,  $u_y$ ,  $u_z$  or  $u$ ,  $v$ ,  $w$  or  $u_1$ ,  $u_2$ ,  $u_3$ . In the following we use the first of these three notation rules. Note that vectors are commonly represented with bold characters.

The sum of two vectors  $\mathbf{u}$  and  $\mathbf{v}$  is given by the sum of the vector components:

$$\mathbf{w} = \mathbf{u} + \mathbf{v} = (u_x + v_x) \mathbf{i} + (u_y + v_y) \mathbf{j} + (u_z + v_z) \mathbf{k} . \quad (\text{B.26})$$

This sum is often written as:

$$\mathbf{w} = (u_x + v_x, u_y + v_y, u_z + v_z) . \quad (\text{B.27})$$

The magnitude (or length) of a vector is given by:

$$|\mathbf{u}| = \sqrt{u_x^2 + u_y^2 + u_z^2} . \quad (\text{B.28})$$

Eqs B.25, B.26 and B.28 may be intuitively or graphically followed using the Pythagoras theorem.

The *scalar- or dot product* of two vectors is a scalar quantity which is defined as the sum of the products of two vector components:

$$\mathbf{u} \bullet \mathbf{v} = u_x v_x + u_y v_y + u_z v_z . \quad (\text{B.29})$$

This is equivalent to the product of the magnitudes of the two vectors and the cos of the angle  $\phi$  between them:

$$\mathbf{u} \bullet \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \cos\phi . \quad (\text{B.30})$$

The scalar product has its name because the result is a scalar quantity. A nice example for a scalar product is the work required to move a plate with the force  $\mathbf{F}$  (being a vector) for the distance  $\mathbf{l}$  (having a length and a direction).

The *cross product* or *vector product* is denoted with  $\times$  or  $\wedge$  and is important when calculating rotations, for example from the angular velocity vector  $\mathbf{w}$  and a position vector  $\mathbf{r}$ . The result is a vector. It is defined as follows:

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \mathbf{u} \wedge \mathbf{v} = (u_y v_z - u_z v_y, u_z v_x - u_x v_z, u_x v_y - u_y v_x) \quad . \quad (\text{B.31})$$

The three values on the right hand side of eq. B.31 have the form of the determinants of matrices.

- *Grad, Div and Curl.* The *gradient* of a scalar valued function (denoted with “Grad” or “Del” or:  $\nabla$ ; s. sect. 3.1.1) is a vector describing the spatial change of this function. It is defined as

$$\nabla \equiv \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \quad . \quad (\text{B.32})$$

Thus, the spatial change of temperature (as a function of  $x, y$  and  $z$ :  $T = T(x, y, z)$ ), may be described as:

$$\text{Grad } T \equiv \nabla T = \left( \frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right) \quad . \quad (\text{B.33})$$

The vector “Grad  $T$ ” is normal to surfaces of constant temperature just like the dip direction of a surface is always normal to the contour lines. “Grad” is a handy tool for the description of the topography of any potential surface.

The *divergence* of a vector field is a *scalar*. In the earth sciences it often describes the transfer rate of mass or energy. The divergence of a vector is defined as follows:

$$\text{Div } \mathbf{v} \equiv \nabla \bullet \mathbf{v} = \left( \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \right) \quad . \quad (\text{B.34})$$

Let us illustrate the *divergence* of a vector valued function dependent on the spatial coordinates  $x, y$  and  $z$  with an example. Assume that  $\mathbf{v}$  is the rate of mass or energy transfer. The flow of mass is  $q_f = \rho \mathbf{v}$  and the flow of energy is:  $q = H \mathbf{v}$ . There,  $\rho$  is density in  $\text{kg m}^{-3}$  and  $H$  is the volumetric energy content in  $\text{J m}^{-3}$ . Thus, flow has the units of  $\text{kg m}^{-2} \text{s}^{-1}$  or  $\text{W m}^{-2}$ , respectively. The divergence of these flows is the sum of the change in flow in the three spatial directions (eq. B.34). If the flow of energy or mass *into* a unity cube is just as large as the flow *out of it* (general criterion for the conservation of mass), then  $\text{Div } \mathbf{v} = 0$  (s. also sect. 3.1.1, 3.3).

The *Curl* or *Rot* of a vector field is a vector describing the rotation of a vector. A vector with  $\text{Curl } \mathbf{u} = 0$  is called non rotating. The Curl is defined by the relationship:

$$\begin{aligned} \text{Curl } \mathbf{v} &\equiv \text{Rot } \mathbf{v} \equiv \nabla \times \mathbf{v} \\ &= \left( \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}, \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}, \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \quad . \end{aligned} \quad (\text{B.35})$$

## B.4 Using Fourier Series

In sections 3.1.1, 3.6.1 and 3.4.1 we have been introduced to two different types of solutions of the diffusion equation (eq. 3.6). They are:

1. Solutions that may be found by integration. These include mainly problems for which the descriptive equations may be so much simplified so that it is straight forward to integrate them. Very often, these are *steady state* problems in which it is possible to assume  $dT/dt = 0$ .
2. Solutions containing an error function. These may be found for problems that have their boundary condition at infinity. For example, when describing the thermal evolution of intrusions that are much smaller than the thickness of the crust or their distance to the earths surface, it is possible to make this assumption (e.g. p. 111 or p. 187).

A third type of solution is necessary for time dependent problems with spatially fixed boundary conditions. We have encountered such examples when describing the erosion of mature landscapes between incising drainages with the diffusion equation, for example on p. 188. Such examples may be solved using *Fourier series*. As the diffusion equation is such a classic example where Fourier series find an important application, we will continue to use this equation as an example. The now well familiar equation that we want to use again (eq. 3.6) is:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial x^2}, \quad (\text{B.36})$$

with  $T$  being a function of both space  $x$  and time  $t$ :  $T = T(x, t)$ . Let us assume that this equation is subject to zero temperature boundary conditions at  $x = 0$  and  $x = l$  which may be formulated as:

- $T = 0$  at  $x = 0$  at time  $t \geq 0$ .
- $T = 0$  at  $x = l$  at time  $t \geq 0$ .

With these boundary conditions, this problem corresponds to that discussed on page 188. There,  $D$  and  $H$  correspond to what is here  $\kappa$  and  $T$  and the spatial extent of the problem was there measured between  $-l$  and  $l$ , while it is here only from 0 to  $l$ . On page 188 we just gave the solution of this problem in eq. 4.69 without detailing the methods of solution. In order to understand the process of solution here in some more detail, consider the following: Eq. B.36 is satisfied if we find a term for which the first time derivative is directly proportional to the second spatial derivative (s. also eq. 3.43 where we solved a very similar problem). The proportionality constant is  $\kappa$ .

It turns out that a general function that satisfies this condition and the boundary conditions has the form:

$$T = \sum_{n=0}^{\infty} a_n e^{b_n t} \sin\left(\frac{n\pi x}{l}\right) . \quad (\text{B.37})$$

There,  $a_n$  and  $b_n$  are constants. This is not very obvious at first and we will therefore now show why this solution satisfies eq. B.36 and how it may be derived:

1. We can see that the solution above contains an exponential function of time and a sine-function of  $x$ . This can be understood as follows: Differentiating an exponential function will always return an exponential function. Correspondingly, the second derivation of a sine-function is a negative sine function. This negative will result in the exponential function also being negative (as shown below), which gives a function that decays with time. Thus, the first derivative of eq. B.37 with respect to  $t$ , will always be proportional to its second derivative with respect to  $x$ . Thus, the condition of the diffusion equation is met, if the correct constants are found.
2. It may be seen that the boundary conditions at  $x = 0$  and  $x = l$  are always satisfied as the sine-function is always zero at these two values of  $x$ . Thus, temperature there is also always zero.
3. The fact that the solution contains an infinite sum is a generalization. If a single term of the infinite sum satisfies eq. B.36, so will the infinite sum of a series of terms.

Let us check if eq. B.37 actually satisfies eq. B.36. For clarity, we perform this check only for a single term of the infinite sum. For our check we differentiate this term with respect to time as well as space. The time derivative gives:

$$\frac{\partial T}{\partial t} = abe^{bt} \sin\left(\frac{n\pi x}{l}\right) . \quad (\text{B.38})$$

The spatial derivatives are:

$$\frac{\partial T}{\partial x} = \frac{n\pi a}{l} e^{bt} \cos\left(\frac{n\pi x}{l}\right) \quad (\text{B.39})$$

as well as:

$$\frac{\partial^2 T}{\partial x^2} = -\frac{n^2 \pi^2 a}{l^2} e^{bt} \sin\left(\frac{n\pi x}{l}\right) . \quad (\text{B.40})$$

Comparing eq. B.38 and eq. B.40 shows that eq. B.36 is satisfied if the constant  $b$  has the following value:

$$b = -\kappa \frac{n^2 \pi^2}{l^2} . \quad (\text{B.41})$$

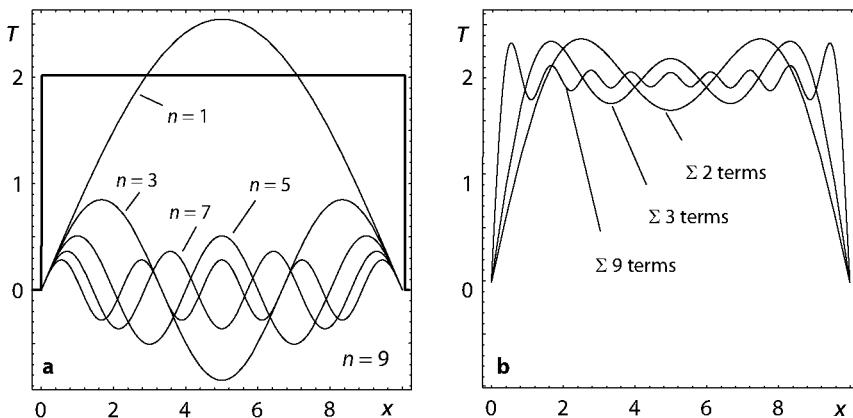
If we insert  $b$  from eq. B.41 in eq. B.37, we have an equation that satisfies all conditions of eq. B.36. The values for the constants  $a_n$  can be determined from the initial conditions. At time  $t = 0$ ,  $e^{bt} = 1$  and thus from eq. B.37 it is true that:

$$T(x, 0) = f(x) = \sum_{n=0}^{\infty} a_n \sin\left(\frac{n\pi x}{l}\right) . \quad (\text{B.42})$$

Eq. B.42 is an example of a Fourier series. The coefficients  $a_n$  can be determined from the integral:

$$a_n = \frac{2}{l} \int_0^l f(x) \sin\left(\frac{n\pi x}{l}\right) dx , \quad (\text{B.43})$$

the derivation of which does not follow directly from eq. B.42 and will not be discussed here. However, it may be found in any book on Fourier series. The coefficients may be evaluated from this integral if the initial condition  $T(x, 0) = f(x)$  is known. However, this integral is only easily evaluated for certain functions of  $f(x)$ . For more general functions, solutions to this integral may be obtained from either math tables or numerically.



**Figure B.11.** **a** The function  $f(x) = 2$  (thick line) and the first five terms of an infinite sum of sine functions from eq. B.42 at time  $t = 0$ . **b** The sum of the first two, three and nine terms of the function shown in **a**. It may be seen that the sum of only few terms is sufficient to approximate the thick drawn function in **a** quite good

- *Solving eq. B.36 for non-zero boundary conditions.* We can take this approach one step further to solve the diffusion equation eq. B.36 for non-zero boundary conditions:

- $T = T_1$  at  $x = 0$  at time  $t \geq 0$ .
- $T = T_2$  at  $x = l$  at time  $t \geq 0$ .

and initial conditions  $T = f(x)$  at  $t = 0$ . In this case, the temperature  $T$  should evolve with time to the steady-state solution that satisfies the boundary conditions. We will denote the steady-state solution as  $g(x)$ . It can be shown that  $g(x) = T_1 + (T_2 - T_1)x/l$ . The solution for  $T$  has the form:

$$T = g(x) + \sum_{n=0}^{\infty} a_n e^{b_n t} \sin\left(\frac{n\pi x}{l}\right) . \quad (\text{B.44})$$

where  $b_n$  is given by equation eq. B.41 and  $a_n$  has the form

$$a_n = \frac{2}{l} \int_0^l (f(x) - g(x)) \sin\left(\frac{n\pi x}{l}\right) dx . \quad (\text{B.45})$$

## B.5 Selected Numerical Tricks

### B.5.1 Integrating Differential Equations

For many differential equations there are mathematical reference books containing their solutions and it certainly goes beyond the scope of this book to go into details of complicated integration methods. However, one simple differential equation that describes an enormous number of natural processes is a rate law where the rate of change of a parameter is linearly related to this parameter. This differential equation has the form:

$$\frac{dy}{dx} = ay + b , \quad (\text{B.46})$$

where  $a$  and  $b$  are constants and we have seen examples of it in eq. 4.15 or eq. 3.44. The equation states that the differential of the variable  $y$  is proportional to  $y$ . This information is sufficient to be able to guess that the solution will contain an exponential function of the form  $e^{cx}$ , because exponential functions always remain exponential functions when they are differentiated (s. Table C.2). Thus, we may guess that the solution will have the form:

$$y = qe^{(cx)} + d \quad (\text{B.47})$$

and thus:

$$\frac{dy}{dx} = qce^{(cx)} . \quad (\text{B.48})$$

Inserting eq. B.47 and eq. B.48 in eq. B.46 shows that  $c = a$  and  $d = -b/a$ . It follows that:

$$y = qe^{(ax)} - \frac{b}{a} , \quad (\text{B.49})$$

for a fixed scalar  $q$ .

### B.5.2 Analytically Unsolvable Equations

Many equations cannot be solved analytically. However, they often may be evaluated numerically by separating them into two parts. We illustrate this using the transcendental eq. 6.3 as an example. This equation has the form

$$cx = ae^{-bx} + d . \quad (\text{B.50})$$

All other parameters that occur in eq. 6.3 are summarized in eq. B.50 into the constants  $a$ ,  $b$ ,  $c$  and  $d$ . Eq. B.50 cannot be solved for  $x$ . In order to solve it numerically it is useful to split the right hand side and the left hand side of the equation into two new equations. For the left side we write:

$$z = cx \quad (B.51)$$

and for the right side we write:

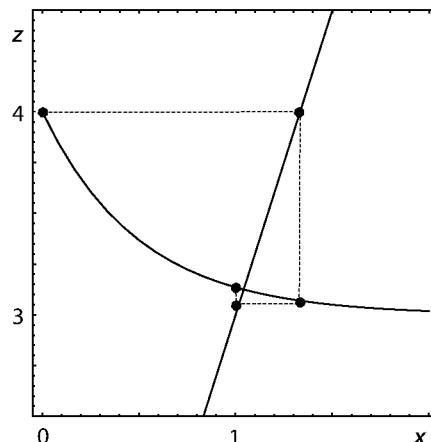
$$z = ae^{-bx} + d \quad \text{or :} \quad x = \frac{\ln\left(\frac{z-d}{a}\right)}{-b} \quad (B.52)$$

These two functions are plotted in Fig. B.12. With the constants  $a = 1$ ,  $b = 2$ ,  $c = 3$  and  $d = 3$ , the steep linear curve is eq. B.51 and the curve with a negative slope is eq. B.52. Their intersection is the solution of eq. B.50. This point may be found by alternating solution of eq. B.51 and B.52. For this we guess a value for  $x$ , insert this into eq. B.51 to calculate  $z$  and then insert this value for  $z$  into eq. B.52 to obtain a new  $x$ . For the example illustrated in Fig. B.24 an initial guess of  $x = 0$  leads to the series:  $z = 4$ ,  $x = 1.333$ ,  $z = 3.069$ ,  $x = 1.023$ ,  $z = 3.129$ ,  $x = 1.043$ ,  $z = 3.124$  and so forth. The result converges to a solution of approximately  $x \approx 1.04$  and  $z \approx 3.12$ . The exact solution may be approximated as closely as desired. While the method is very simple, it may also lead to wrong results, for example if one of the two functions has local minima or maxima.

### B.5.3 The Least Squares Method

A common problem in science occurs when a curve should be fitted to a number of data points and the fit of the data to this curve should be quantified. The most common method for this is to find the smallest sum of the squares of the deviations of the data from the curve, in short, the *least squares*. In the following section we explain how this is done with the example of a linear fit,

**Figure B.12.** Illustration of the numerical solution of eq. B.50. The constants are  $a = 1$ ,  $b = 2$ ,  $c = 3$  and  $d = 3$ . The straight line represents eq. B.51, the curved line is eq. B.52. The dashed line shows the iterative approximation of their intersection



i.e. the curve that is fitted to the data is a straight line. However, for more complicated choices of functions, the same rules apply. We assume that the data consist of  $n$  values for  $y$  and just as many for  $x$ . We label the data with  $y_i$  and  $x_i$  from  $i = 1$  to  $n$ . The straight line we will to fit through the data cloud has the form  $y = ax + b$  where  $a$  and  $b$  are unknown. If we insert our data pair for  $x$  and  $y$  into this linear equation, we obtain:

$$y_i = ax_i + b - e \quad , \quad (\text{B.53})$$

where  $e$  is the deviation of a given data point from the fitted line. In order to minimize the sum of the squares of this deviation, the sum of all  $e$  must be minimized. This means that:

$$\sum_{i=1}^n e^2 = \sum_{i=1}^n (ax_i + b - y_i)^2 \quad (\text{B.54})$$

must be minimized. In order to do so, eq. B.54 may be partially differentiated with respect to  $a$  and  $b$ , set to zero and solved for  $a$  or  $b$ . Using the simple differentiation rules from Table C.1 the derivative with respect to  $a$  may be found to be:

$$0 = \frac{\partial (\sum_{i=1}^n (ax_i + b - y_i)^2)}{\partial a} = \sum_{i=1}^n 2x_i(ax_i + b - y_i)^2 \quad . \quad (\text{B.55})$$

The derivative with respect to  $b$  is:

$$0 = \frac{\partial (\sum_{i=1}^n (ax_i + b - y_i)^2)}{\partial b} = \sum_{i=1}^n 2(ax_i + b - y_i)^2 \quad . \quad (\text{B.56})$$

Eqs. B.55 and B.56 may be simplified and solved simultaneously for  $a$  and  $b$ . We get:

$$a = \frac{n(\sum_{i=1}^n x_i y_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2} \quad , \quad (\text{B.57})$$

$$b = \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2} \quad . \quad (\text{B.58})$$

These are the coefficients of the *best fitting straight line* of eq. B.53. These two equations (eq. B.57 and eq. B.58) are straight forward to implement on a computer.

The errors of these values, as given by their standard deviations are simply given by:

$$(\partial a)^2 = \frac{n \sum_{i=1}^n e_i^2}{(n-2) \left( n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2 \right)} \quad (\text{B.59})$$

and:

$$(\partial b)^2 = \frac{\left(\sum_{i=1}^n x_i^2\right) \left(\sum_{i=1}^n e_i^2\right)}{(n-2) \left(n \left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2\right)} . \quad (\text{B.60})$$

### B.5.4 Basic Statistical Parameters

A fundamental task of geologists is the characterization of the location and the variability of a data set, for example  $P$ - $T$  data determined with thermobarometry, geochronological data, numbers from a whole rock analysis, digital data during image processing or dip and strike data measured in the field. Because such measurements may be very precise, but they are never perfectly accurate it is necessary to evaluate them statistically (s. p. 5). The most important parameters for such an evaluation are summarized here.

- *Normal distribution.* The statistical interpretation of many geological data is based on the assumption that the data have a normal (also called Gaussian) distribution around the exact value. A distribution is said to be normal if its probability density function is given by:

$$f(x) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right) e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} . \quad (\text{B.61})$$

This function is characterized by two parameters called the mean,  $\mu$ , and the standard deviation  $\sigma$ . Eq. B.61 is plotted in Fig. B.13 and may be interpreted as the enveloping curve of a histogram. If the data are centered around a mean of  $\mu = 0$  and the standard deviation is  $\sigma = 1$ , then eq. B.61 simplifies to:  $f(x) = e^{-(x^2/2)}/\sqrt{2\pi}$ . The distribution is said to be a *standard normal distribution*.

- *Mean.* The mean of a data set indicates the most probable location of the exact value. For a data set  $S$  containing  $n$  data  $S_i$  it is defined as:

$$\mu = \frac{1}{n} \sum_{i=1}^n S_i . \quad (\text{B.62})$$

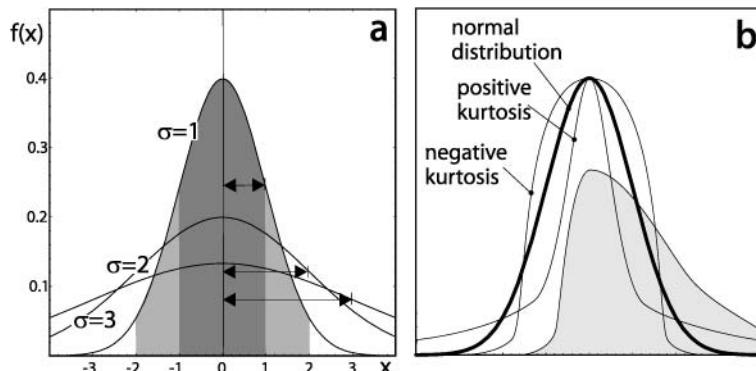
the mean gives the most probable location of the exact value.

- *Variance and standard deviation.* The variance is a measure of the variability of the data around the mean. It is defined as

$$Var = \frac{1}{n} \sum_{i=1}^n (S_i - \mu)^2 . \quad (\text{B.63})$$

If the data set is incomplete (univariat data set), where only a finite number of data points were collected, an unbiased estimate of the population variance is given by:

$$Var' = \frac{1}{n-1} \sum_{i=1}^n (S_i - \mu)^2 . \quad (\text{B.64})$$



**Figure B.13.** a Normal distribution of data points calculated with eq. B.61. The curve for  $\sigma = 1$  is a standard normal distribution. The area underneath all three curves is 1. The  $1\sigma$  confidence interval is shaded dark, the  $2\sigma$  confidence interval is shaded light. These two areas contain 68.26% and 95.45% of the data, respectively. b shows data distributions with high and low kurtosis as well as one with a significant skewness (shaded), relative to a normal distribution

The *standard deviation* is defined as the square root of the variance. If the data distribution is normal (Gaussian), then the standard deviation is abbreviated with  $\sigma = \sqrt{Var}$ . From eq. B.61 it may be calculated that 68.26% of all data of a normal distribution will fall within  $\pm 1\sigma$  of the mean and 95.45% of the data within  $\pm 2\sigma$ . The latter is therefore often loosely referred to as a “95% confidence interval”.

- *Skewness and kurtosis.* Skewness is a measure of the lack of symmetry of data in a histogram. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution (Fig. B.13b). Data with a high kurtosis have a distinct peak near the mean, decline rather rapidly and have heavy tails. Data with a low kurtosis have a flat top. A uniform distribution is an extreme case of kurtosis. Skewness (for a univariate data set s. eq. B.64) is defined as:

$$skew = \frac{1}{\sigma^3(n-1)} \sum_{i=1}^n (S_i - \mu)^3 \quad . \quad (B.65)$$

Kurtosis is defined as:

$$kurt = \left( \frac{1}{\sigma^4(n-1)} \sum_{i=1}^n (S_i - \mu)^4 \right) - 3 \quad . \quad (B.66)$$

Many classical statistical tests depend on the assumption that the data have a normal distribution. Significant skewness and kurtosis indicate that the data are not normal. In fission track analysis the skewness and kurtosis of track length distributions bear significant information on the cooling history of the rocks.

## B.6 Problems

**Problem B.1.** *Finite difference approximations (p. 399):*

Eq. 3.42 describes simultaneous diffusion and one-dimensional mass transport. Write an explicit finite difference approximation for this equation. Use a forward differencing scheme to approximate the transport term. Follow the schemes introduced in eq. B.14 and eq. B.15

**Problem B.2.** *Finite difference approximations (p. 400):*

Eq. 4.52 describes the elastic bending of oceanic lithosphere under applied loads. There,  $w$  is the variable for which we want to solve the equation. It is the vertical deflection of the bent plate as a function of distance  $x$ . Write an explicit finite difference approximation for this equation. Hint: it is easiest to just expand the scheme we have followed in eq. B.16

**Problem B.3.** *Finite difference approximations (p. 405):*

Redraw Fig. B.6 carefully to convince yourself why *backward* finite differencing schemes will be unstable when describing the transport equation. For your considerations look at the scheme of eq. B.24, and a corresponding backward differencing scheme for  $u = \Delta x / \Delta t$ .

**Problem B.4.** *Mean and standard deviation (p. 405):*

Determine the mean and the standard deviation of the following two data sets  $S1 = \{10, 10, 10, 10, 6, 6, 6, 6\}$  and  $S2 = \{1, 5, 10, 20, 15, 9, 4, 0\}$ , assuming that both data sets are samples from a normal distribution.

## C. Maths Refresher

### Recommended Reading:

- Abramowitz and Stegun (1972) Handbook of Mathematical Functions.
- Press, Flannery, Teukolsky and Vetterling (1989) Numerical Recipes.
- Strang (1988) Linear Algebra and Applications.
- <http://mathworld.wolfram.com>

**Table C.1.** General rules of differential calculus using the example of the function  $y = f(x)$ .  $u$  and  $v$  are also functions of  $x$ .  $a$  is a constant.  $f'(x)$  or  $y'$  is the first derivative of  $y$  with respect to  $x$

$f(x)$	$f'(x)$
$y = au$	$y' = a(\mathrm{d}u/\mathrm{d}x)$
$y = x^a$	$y' = ax^{a-1}$
$y = a^x$	$y' = a^x \ln(a)$
$y = x^x$	$y' = (1 + \ln(x))x^x$
$y = u + v$	$y' = (\mathrm{d}u/\mathrm{d}x) + (\mathrm{d}v/\mathrm{d}x)$
$y = uv$	$y' = u(\mathrm{d}v/\mathrm{d}x) + v(\mathrm{d}u/\mathrm{d}x)$
$y = \frac{u}{v}$	$y' = (v\mathrm{d}u/\mathrm{d}x - u\mathrm{d}v/\mathrm{d}x)/v^2$
$y = u^v$	$y' = u^v ((v/u)(\mathrm{d}u/\mathrm{d}x) + \ln u(\mathrm{d}v/\mathrm{d}x))$

**Table C.2.** Special derivatives of the function  $y = f(x)$ .  $f'(x)$  or  $y'$  is the first derivative of  $y$  with respect to  $x$ . The table may be read both ways:  $f(x)$  may also be interpreted as the integral of the function  $f'(x)$

$f(x)$	$f'(x)$
$y = \frac{1}{x}$	$= x^{-1}$
$y = \ln(x)$	$y' = \frac{1}{x \ln(10)}$
$y = \log(x)$	$= \frac{\ln(x)}{\ln(10)}$
$y = e^x$	$y' = e^x$
$y = e^{ax}$	$y' = ae^x$
$y = x \ln(x) - x$	$y' = \ln(x)$
$y = \sin(x)$	$y' = \cos(x)$
$y = \cos(x)$	$y' = -\sin(x)$
$y = \tan(x)$	$y' = \sec(x)$

**Table C.3.** Conversions between different logarithms. In this book we use “ln” for the natural logarithm (to the base of e) and “log” for the decimal logarithm (to the base of 10)

$\ln(xy)$	$= \ln(x) + \ln(y)$
$\ln(x/y)$	$= \ln(x) - \ln(y)$
$\ln(x^y)$	$= y \ln(x)$
$\ln(e)$	$= 1$
$\ln(1)$	$= 0$
$\ln(0)$	$= -\infty$
$\log(x)$	$= \ln(x)/\ln(10) = \log(e)\ln(x)$
$\ln(10)\log(e)$	$= 1$

**Table C.4.** Trigonometry of a triangle with the side lengths  $a$ ,  $b$  and  $c$ . The angles  $\alpha$ ,  $\beta$  and  $\gamma$  are those opposite to the sides  $a$ ,  $b$  and  $c$ , respectively

*plane triangles with a right angle ( $\gamma = 90^\circ$ )*

$$a^2 + b^2 = c^2 \quad (\text{Pythagoras})$$

$$\sin(\alpha) = a/c$$

$$\cos(\alpha) = b/c$$

$$\tan(\alpha) = a/b$$

$$\cot(\alpha) = b/a$$

$$\text{area} = (a \times b)/2$$

*general plane triangles*

$$a/\sin(\alpha) = b/\sin(\beta) = c/\sin(\gamma)$$

$$a^2 = b^2 + c^2 - 2bc \times \cos(\alpha)$$

$$\text{area} = (b c \sin(\alpha))/2 = [s(s - a)(s - b)(s - c)]^{1/2}$$

$$\dots \text{where: } s = 1/2(a + b + c)$$

*spherical triangles*

$$\sin(a)/\sin(\alpha) = \sin(b)/\sin(\beta) = \sin(c)/\sin(\gamma)$$

$$\cos(a) = \cos(b)\cos(c) + \sin(b)\sin(c)\cos(\alpha)$$

*angular separation  $\alpha$  of 2 points on a sphere (lats.  $\lambda_1, \lambda_2$  and longs.  $\phi_1, \phi_2$ )*

$$\cos(\alpha) = \sin(\lambda_1)\sin(\lambda_2) + \cos(\lambda_1)\cos(\lambda_2)\cos(\phi_2 - \phi_1)$$

...for any  $\alpha$

$$\sin(\alpha/2) = \left( \sin^2((\lambda_2 - \lambda_1)/2) + \cos(\lambda_2)\cos(\lambda_1)\sin^2((\phi_2 - \phi_1)/2) \right)^{0.5}$$

...for small  $\alpha$

*transformation between spherical and Cartesian coordinates (R is sphere radius)*

$$x = R\cos(\phi)\cos(\lambda)$$

$$y = R\sin(\phi)\cos(\lambda)$$

$$z = R\sin(\lambda)$$

*spatial distance  $a$  between two points  $x_1, y_1, z_1$  and  $x_2, y_2, z_2$*

$$a^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

**Table C.5.** Volume, surface and other important data of some geometrical bodies*cube with side length a*

surface	$= 6a^2$
volume	$= a^3$
length of side diagonal	$= a\sqrt{2}$
length of space diagonal	$= a\sqrt{3}$

*tetrahedron with side length a*

surface	$= a^2\sqrt{3}$
volume	$= \frac{1}{12}a^3\sqrt{2}$

*sphere with radius r*

surface	$= 4\pi r^2$
volume	$= \frac{4}{3}\pi r^3$

*cone with radius r and height H*

surface of the mantle	$= r\pi\sqrt{r^2 + H^2}$
volume	$= \frac{1}{3}\pi r^2 H$

*cylinder with radius r and height H*

surface of the mantle	$= 2r\pi H$
volume	$= r^2\pi H$

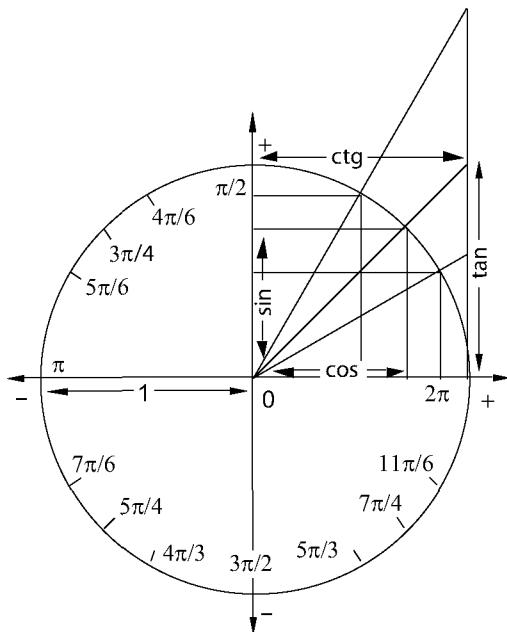
**Table C.6.** Definitions and conversions between trigonometric functions

$\csc(x) = 1/\sin(x)$
$\sec(x) = 1/\cos(x)$
$\operatorname{tg}(x) = \sin(x)/\cos(x)$
$\operatorname{ctg}(x) = \cos(x)/\sin(x) = 1/\tan(x)$
$\sin^2(x) + \cos^2(x) = 1$

**Table C.7.** Special values of trigonometric functions (Fig. C.1)

Angle rad	$\alpha = 0^\circ$	$\alpha = 30^\circ$	$\alpha = 45^\circ$	$\alpha = 60^\circ$	$\alpha = 90^\circ$
$\sin(\alpha)$	0	0.5	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\cos(\alpha)$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	0.5	0
$\operatorname{tg}(\alpha)$	0	$1/\sqrt{3}$	1	$\sqrt{3}$	$\infty$
$\operatorname{ctg}(\alpha)$	$\infty$	$\sqrt{3}$	$1/\sqrt{3}$	0	0

**Figure C.1.** The unity circle and definitions of the trigonometric functions. The four trigonometric functions are labeled for an angle of  $45^\circ$



**Table C.8.** Solutions of quadratic equations of the form  $ax^2 + bx + c = 0$

$$x_{1,2} = -\frac{b}{2a} \pm \frac{1}{2a} \sqrt{b^2 - 4ac}$$

**Table C.9.** Approximation of the error function using the constants  $a = 0.3480242$ ;  $b = -0.0958798$  and  $c = 0.7478556$  and  $y = 1/(1 + 0.47047x)$

$$\text{erf}(x) \approx 1 - (ay + by^2 + cy^3) e^{-x^2}$$

## **D. Symbols and Units**

**Table D.1.** Symbols and units of the variables used in this book. Physical constants of the earth are listed in Table D.3 and D.4. Variables abbreviated with Greek letters are explained in Table D.2. Variables used in the text are often specified in more detail by adding a subscript. The most commonly used subscripts in this book are  $l$  for lithosphere;  $c$  for crust;  $m$  for mantle;  $i$ ,  $j$  and  $n$  for numbering;  $x$ ,  $y$  and  $z$  for spatial directions and 0 for initial values

Symbol	Variable	Unit	1. Occurrence
$Ar$	Argand number	—	Eq. 6.31
$A$	area	$\text{m}^2$	Eq. 3.2
$A$	pre exponential constant	$\text{MPa}^{-n}\text{s}^{-1}$	Eq. 5.43
$A_q$	... of quartz creep	$\text{MPa}^{-n}\text{s}^{-1}$	Table 5.3
$A_o$	... of olivine creep	$\text{MPa}^{-n}\text{s}^{-1}$	Table 5.3
$a$	acceleration	$\text{ms}^{-2}$	Eq. 5.14
$a$	general constant	variable	Eq. 3.19
$b$	general constant	variable	Eq. 3.19
$c_p$	heat capacity (of rocks)	$\approx 1\,000 \text{ J kg}^{-1} \text{ K}^{-1}$	Eq. 3.4
$c_{pf}$	... of fluids	$\text{J kg}^{-1} \text{ K}^{-1}$	Eq. 3.50
$C$	concentration	—	Eq. 7.3
$C$	constant of integration	variable	Eq. 3.57
$D$	angular momentum	$\text{kg m}^2 \text{ s}^{-1}$	Sect. 2.2.4
$D$	rigidity of elastic plates	$\text{Nm}$	Eq. 4.47
$D$	displacement gradient	tensor	Eq. 4.3
$D$	diffusivity of mass	$\text{m}^2 \text{ s}^{-1}$	Eq. 4.61
$D_0$	pre exponential diffusivity	$\text{m}^2 \text{ s}^{-1}$	Eq. 7.5
$e$	elongation	—	Eq. 4.1
$e$	error	—	Eq. B.53
$E$	Young's Modul	$\text{Pa}$	Eq. 4.49
$E$	energy	$\text{J}$	—
$E_p$	potential energy (per area)	$\text{J m}^{-2}$	Eq. 5.49
$E_k$	kinetic energy	$\text{J}$	Eq. 5.22
$f$	frequency	$\text{s}^{-1}$	Eq. 3.100
$f$	ellipticity	—	Eq. 4.7
$f_c$	vertical strain of crust	—	Eq. 4.9
$f_l$	vertical strain of lithosphere	—	Eq. 4.9
$F$	force	$\text{N}$	Sect. 2.2.4
$F_b$	buoyancy force (per length)	$\text{N m}^{-1}$	Eq. 5.49
$F_d$	tectonic driving force (per length)	$\text{N m}^{-1}$	Eq. 6.21
$F_{\text{eff}}$	effective driving force (per length)	$\text{N m}^{-1}$	Eq. 6.21
$F_I$	integrated strength	$\text{N m}^{-1}$	Eq. 5.48
$G$	gravitation constant	—	Eq. 1.1
$G$	geometrical factor for $T_c$	—	Eq. 7.10
$G$	Gibb's energy	$\text{J}$	Eq. 7.6
$g$	geothermal gradient	$^{\circ}\text{C m}^{-1}$	Eq. 3.48
$g$	acceleration (gravitational)	$\text{m s}^{-2}$	Eq. 4.22

**Table C.1.** ... continuation

Symbol	Variable	Unit	1. Occurrence
$h$	elastic thickness	m	Eq. 4.49
$h$	dimensionless height	—	Eq. 1.2
$h_r$	exponential drop off	m	Eq. 3.67
$h_s$	erodable thickness	m	Eq. 4.62
$H$	heat content (volumetric)	$\text{J m}^{-3}$	Fig. 3.12; Eq. 3.3
$H$	elevation, height	m	Eq. 1.2
$I$	momentum	$\text{kg m s}^{-1}$	Sect. 2.2.4
$I$	tensor invariant	—	Eq. 5.3
$J$	moment of inertia	$\text{kg m}^2$	Sect. 2.2.4
$k$	thermal conductivity	$\text{J s}^{-1} \text{m}^{-1} \text{K}^{-1}$	Eq. 3.1
$K$	bulk modulus	Pa	Sect. 5.1.2
$K$	equilibrium coefficient	—	Eq. 7.6
$l$	thickness, length	m	Eq. 3.17
$L$	latent heat of fusion	$\text{J kg}^{-1}$	Eq. 3.31
$L$	sediment thickness	m	Eq. 6.2
$L_0$	... of decompacted layer	m	Eq. 6.2
$L^*$	... of decompacted pile	m	Eq. 6.2
$m$	mass	kg	Eq. 1.1
$M$	bending moment	N	Eq. 4.47
$N$	number, counter	—	Eq. 3.91
$n$	power law-exponent	—	Eq. 5.44
$n$	general counter	—	Eq. 3.15
$Pe$	Peclet number	—	Eq. 3.51
$P$	pressure	Pa	Eq. 5.5
$Q$	activation energy (diffusion)	$\text{J mol}^{-1}$	Eq. 7.5
$Q$	activation energy (creep)	$\text{J mol}^{-1}$	Eq. 5.43
$Q_q$	... of quartz creep	$\text{J mol}^{-1}$	Table 5.3
$Q_o$	... of olivine creep	$\text{J mol}^{-1}$	Table 5.3
$Q_D$	... of Dorn law creep	$\text{J mol}^{-1}$	Eq. 5.47
$q$	load on a plate	Pa	Eq. 4.48
$q$	heat flow	$\text{W m}^{-2}$	Eq. 3.1
$q_s$	... at the surface	$\text{W m}^{-2}$	Eq. 3.61
$q_m$	... at the Moho	$\text{W m}^{-2}$	Eq. 3.61
$q_{\text{rad}}$	... caused by radioactivity	$\text{W m}^{-2}$	Eq. 3.61; Eq. 6.13
$q_r$	water flux	$\text{m s}^{-1}$	Eq. 4.81
$q_f$	sediment flux in rivers	$\text{m s}^{-1}$	Eq. 4.81
$r$	radius	m	Eq. 1.1, 3.11
$r$	distance in polar coordinates	m	Eq. 1.1, 3.11
$R$	radius of earth	m	Eq. 2.1
$R_A$	... at the equator	6 378.139 km	Eq. 4.7
$R_P$	... at the pole	6 356.75 km	Eq. 4.7

Table C.1. *continuation*

Symbol	Variable	Unit	1. Occurrence
$s$	stretch	—	Eq. 4.1
$s$	cooling rate	$^{\circ}\text{Cs}^{-1}$	Eq. 3.89
$S$	rate of heat production	$\text{Js}^{-1} \text{m}^{-3}$	Eq. 3.23
$S_{\text{chem}}$	... chemical	$\text{Js}^{-1} \text{m}^{-3}$	Eq. 3.22
$S_{\text{mec}}$	... mechanical	$\text{Js}^{-1} \text{m}^{-3}$	Eq. 3.22
$S_{\text{rad}}$	... radioactive	$\text{Js}^{-1} \text{m}^{-3}$	Eq. 3.22
$S_0$	... radioactive at surface	$\text{Js}^{-1} \text{m}^{-3}$	Eq. 3.67
$S$	entropy	$\text{J K}^{-1}$	Eq. 3.30
$SL$	sea level	m	Eq. 6.5
$t$	time	s	Eq. 3.4
$t_a$	degradation coefficient	$\text{m}^2$	Eq. 4.72
$t_{\text{eq}}$	time scale of diffusion	s	Eq. 3.17
$t_E$	erosional time constant	s	Eq. 4.16
$T$	temperature	$^{\circ}\text{C}; \text{K}$	Eq. 3.1
$T_A$	... at the start	$^{\circ}\text{C}; \text{K}$	Eq. 7.9
$T_b$	... of the host rock	$^{\circ}\text{C}; \text{K}$	Eq. 3.84
$T_c$	... closure temperature	$^{\circ}\text{C}; \text{K}$	Eq. 7.10
$T_E$	... at the end	$^{\circ}\text{C}; \text{K}$	Eq. 7.9
$T_i$	... intrusion temperature	$^{\circ}\text{C}; \text{K}$	Eq. 3.84
$T_l$	... at base of lithosphere	$\approx 1\,200\text{--}1\,300 \text{ }^{\circ}\text{C}$	Eq. 3.59
$T_l$	... liquidus	$^{\circ}\text{C}; \text{K}$	Eq. 3.38
$T_s$	... solidus	$^{\circ}\text{C}; \text{K}$	Eq. 3.38
$T_s$	... at the surface	$^{\circ}\text{C}; \text{K}$	Eq. 3.80
$T_0$	... initial temperature	$^{\circ}\text{C}; \text{K}$	Eq. 3.100
$u$	velocity (often: in $x$ direction)	$\text{m s}^{-1}$	Eq. 3.40
$U$	circumference	m	Sect. 4.2.1
$v$	velocity (often: in $y$ direction)	$\text{m s}^{-1}$	Eq. 4.10
$v_f$	... of fluids	$\text{m s}^{-1}$	Eq. 3.50
$v_{\text{ex}}$	exhumation rate	$\text{m s}^{-1}$	Eq. 4.10
$v_{\text{er}}$	erosion rate	$\text{m s}^{-1}$	Eq. 4.13
$v_{\text{ro}}$	uplift rate of rocks	$\text{m s}^{-1}$	Eq. 4.10
$v_{\text{up}}$	uplift rate of the surface	$\text{m s}^{-1}$	Eq. 4.10
$V$	volume	$\text{m}^3$	Eq. 3.2
$w$	angular velocity	$\text{s}^{-1}$	Sect. 2.2.4
$w$	water depth	m	Eq. 4.36
$w$	thickness of crustal root	m	Eq. 5.55
$w$	plate deflection	m	Eq. 4.47
$x$	spatial coordinate (horizontal)	m	Sect. 1.2
$X$	mole fraction	—	Eq. 7.7
$y$	spatial coordinate (horizontal)	m	Sect. 1.2
$z$	spatial coordinate (vertical)	m	Sect. 1.2
$z_c$	thickness of crust	m	Sect. 2.4.1; Eq. 3.63
$z_i$	initial depth	m	Eq. 3.49
$z_K$	isostatic compensation depth	m	Eq. 4.21
$z_l$	thickness of lithosphere	m	Sect. 2.4.1
$z_{\text{rad}}$	thickness of radioactive crust	$\approx 7\text{--}10 \text{ km}$	Eq. 3.61

Table D.2. Greek symbols

Symbol	Variable	Unit	1. Occurrence
$\alpha$	coefficient of thermal expansion	$\approx 3 \cdot 10^{-5} \text{ }^{\circ}\text{C}^{-1}$	Eq. 3.29
$\alpha$	general angle	radian	Eq. 3.99
$\alpha$	flexural parameter	m	Eq. 4.53
$\beta$	isothermal compressibility	$\text{Pa}^{-1}$	Eq. 3.28; 5.28
$\beta$	stretching factor (mantle lithosph.)	—	Eq. 6.10
$\delta$	stretching factor (crust)	—	Eq. 6.10
$\delta$	density ratio	—	Eq. 4.31
$\gamma$	shear strain	—	Eq. 4.2
$\eta$	viscosity	$\text{Pa s}$	Eq. 5.42
$\epsilon$	strain	—	Eq. 5.24
$\dot{\epsilon}$	strain rate	$\text{s}^{-1}$	Eq. 1.5
$\kappa$	diffusivity	$\text{m}^2 \text{ s}^{-1}$	Eq. 3.6
$\lambda$	longitude	degree	Eq. 2.2; Fig. 2.8
$\lambda$	wave length	m	Eq. 3.101
$\lambda$	pore fluid pressure ratio	—	Eq. 5.33
$\lambda_c$	... in the crust	—	Fig. 5.17
$\lambda_l$	... in the lithosphere	—	Fig. 5.17
$\mu$	coefficient of internal friction	—	Eq. 5.29
$\nu$	Poisson ratio	—	Eq. 4.49
$\phi$	latitude	degree	Fig. 2.8
$\phi$	latitude	degree	Fig. 2.8
$\phi$	angular shear strain	degree	Eq. 4.2
$\phi_0$	... at the surface	—	Eq. 6.1
$\phi$	friction angle	radian	Eq. 5.32
$\rho$	density	$\text{kg m}^{-3}$	Eq. 3.4
$\rho_c$	... of the crust	$\text{kg m}^{-3}$	Eq. 4.24
$\rho_g$	... of sediment grains	$\text{kg m}^{-3}$	Fig. 6.5; Eq. 6.7
$\rho_0$	... of the mantle at $0 \text{ }^{\circ}\text{C}$	$\approx 3300 \text{ kg m}^{-3}$	Eq. 4.28
$\rho_L$	... of a sedimentary pile	$\text{kg m}^{-3}$	Eq. 6.5
$\rho_m$	... of the mantle at $T_l$	$\approx 3200 \text{ kg m}^{-3}$	Eq. 4.24
$\rho_w$	... of water	$\approx 1000 \text{ kg m}^{-3}$	Eq. 4.36
$\sigma$	stress (tensor)	Pa	Eq. 4.3
$\sigma_1, \sigma_2, \sigma_3$	principal stress	Pa	Eq. 5.2
$\sigma_d$	differential stress	Pa	Eq. 3.25
$\sigma_s$	shear stress	Pa	Eq. 5.29
$\sigma_n$	normal stress	Pa	Eq. 5.1
$\sigma_D$	... critical, Dorn law creep	Pa	Eq. 5.47
$\tau$	deviatoric stress (tensor)	Pa	Eq. 5.7
$\theta$	dimensionless temperature	—	Eq. 1.3; 3.87
$\theta$	angle of a failure surface to $\sigma_1$	radian	Fig. 5.6
$\xi$	expansion ratio	—	Eq. 4.31

**Table D.3.** Important data of the earth

equatorial radius	6 378.139 km
polar radius	6 356.750 km
diameter of core	3 468 km
volume	$1.083 \cdot 10^{21} \text{ m}^3$
mass	$5.973 \cdot 10^{24} \text{ kg}$
surface area	$5.10 \cdot 10^{14} \text{ m}^2$
area of the continents	$1.48 \cdot 10^{14} \text{ m}^2$
area of continental lithosphere	$2.0 \cdot 10^{14} \text{ m}^2$
area of oceanic lithosphere	$3.1 \cdot 10^{14} \text{ m}^2$
mean elevation of the continents	825 m
mean depth of the oceans	3 770 m
total length of mid oceanic ridges	$\approx 60\,000 \text{ km}$
mean continental surface heat flow	$\approx 0.065 \text{ W m}^{-2}$
mean oceanic surface heat flow	$\approx 0.101 \text{ W m}^{-2}$

**Table D.4.** Important physical constants

Constant	Symbol	Value
gas constant	$R$	$8.3144 \text{ J mol}^{-1} \text{ K}^{-1}$
gravitational constant	$G$	$6.6732 \cdot 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$
speed of light	$c$	$2.99792 \cdot 10^8 \text{ m s}^{-1}$

**Table D.5.** SI-units

Physical Parameter	Symbol in Text	Unit	Abbreviation
distance	$x, y, z$	meter	m
time	$t$	second	s
mass	$m$	kilogram	kg
temperature	$T$	Kelvin	K

**Table D.6.** Important derived units

Physical Parameter	Symbol in Text	Unit	Abbreviation	SI-Unit
force	$F$	Newton	N	$\text{kg m s}^{-2}$
pressure	$P$	Pascal	$\text{Pa} = \text{N m}^{-2}$	$\text{kg m}^{-1} \text{s}^{-2}$
energy	$E$	Joule	$\text{J} = \text{N m}$	$\text{kg m}^2 \text{s}^{-2}$
power	–	Watt	$\text{W} = \text{J s}^{-1}$	$\text{kg m}^2 \text{s}^{-3}$

**Table D.7.** Conversions between derived units

Physical Parameter	Conversion
force	$= \text{mass} \times \text{acceleration}$ $= \text{pressure} \times \text{distance} = \text{Pa m}$
pressure	$= \text{force per area} = \text{N m}^{-2}$ $= \text{energy per volume} = \text{J m}^{-3}$
energy	$= \text{force} \times \text{distance} = \text{Nm}$ $= \text{mass} \times \text{velocity}^2 = \text{kg m}^2 \text{ s}^{-2}$
power	$= \text{work per time} = \text{J s}^{-1} = \text{W}$
velocity	$= \text{distance per time} = \text{m s}^{-1}$
acceleration	$= \text{velocity change per time} = \text{m s}^{-2}$

**Table D.8.** Important commonly used variables and their conversion into SI-units. Numbers are given to a maximum of 4 digits

Parameter	Unit	Conversion
length	1 angstrom	$= 10^{-10} \text{ m}$
	1 micrometer ( $\mu\text{m}$ )	$= 10^{-6} \text{ m}$
	1 millimeter (mm)	$= 10^{-3} \text{ m}$
	1 kilometer (km)	$= 10^3 \text{ m}$
	1 foot (ft)	$= 0.3048 \text{ m}$
	1 inch (in)	$= 2.54 \text{ cm}$
	1 mile (mi)	$= 1.6093 \text{ km}$
	1 yard (yd)	$= 0.9144 \text{ m}$
	1 nautical mile (nmi)	$= 1.852 \text{ km}$
	1° latitude	$= 60 \text{ nmi} = 111.12 \text{ km}$
area	1 hectar (ha)	$= 10^4 \text{ m}^2$
	1 acre	$= 4046.9 \text{ m}^2$
volume	1 liter (l)	$= 10^{-3} \text{ m}^3$
	1 gallon (US)	$= 3.7854 \text{ l}$
	1 gallon (UK)	$= 4.5461 \text{ l}$
	1 hectoliter (hl)	$= 100 \text{ l}$
	1 barrel (US)	$= 158.98 \text{ l}$
time	1 day	$= 8.64 \cdot 10^4 \text{ s}$
	1 million years (my)	$= 3.1557 \cdot 10^{13} \text{ s}$
temperature	1 °C	$= 1 \text{ K} (0^\circ\text{C} = 273.16 \text{ K})$
force	1 dyne	$= 1 \text{ g cm s}^{-2} = 10^{-5} \text{ N}$
pressure	1 bar	$= 10^5 \text{ Pa}$
	1 atmosphere (atm)	$= 1.0133 \cdot 10^5 \text{ Pa} = 760 \text{ mm Hg}$
	1 mm Hg (torr)	$= 1.3332 \cdot 10^2 \text{ Pa}$
	1 lb in <sup>-2</sup>	$= 6.8947 \cdot 10^3 \text{ Pa}$
energy	1 cal	$= 4.184 \text{ J}$
	1 erg	$= 1 \text{ dyne cm} = 10^{-7} \text{ J}$
	1 heat flow unit (hfu)	$= 10^{-6} \text{ cal cm}^{-2} \text{ s}^{-1} = 0.04184 \text{ W m}^{-2}$
	1 horse power (PS)	$= 746 \text{ W}$
viscosity	1 poise	$= 0.1 \text{ Pas}$

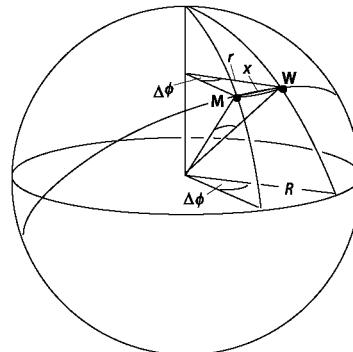
## E. Answers to Problems

**Problem 2.1.** According to Fig. 2.5, both the Aleute arc and the Java Trench appear to have small circle radii corresponding to roughly  $25^\circ$  latitude which is  $\approx 2700$  km. Thus, the ping pong ball model of eq. 2.1 predicts subduction angles of the order of  $25^\circ$ . The much steeper observed dip may be due to additional forces exerted onto the subducted slab by asthenospheric convection.

**Problem 2.2.** The experiment can be performed by starting at  $1 + (2\pi n)^{-1}$  km from the south pole, where  $n$  is an integer. Thus, there is not only a single point, but an infinite number of rings around the south pole from where this experiment can be started. 1 km south of each of these rings the circumference of the earth along a line going due east or west is  $1/n$  km.

**Problem 2.3.**  $(5/360) \cdot 24 \text{ hours} = 20 \text{ minutes}$ . Know edge of the latitude is not necessary.

**Problem 2.4.** The difference in geographic longitude is  $4^\circ 42' = 282'$ . The improved solution of Problem 2.3 is then  $282/(360 \cdot 60) \cdot 24 \text{ hours} = 18.8 \text{ minutes}$ .



**Figure E.1.** Sketch of the solution of Problem 2.5

**Problem 2.5.** a) The difference in longitude between Munich and Vienna is:  $\Delta\phi = 5^\circ$ . Therefore:  $we = \cos\lambda \frac{\Delta\phi}{360} 2R\pi = 371\,961.3$  m. The small circle radius  $r$  for the 48th degree of latitude is  $r = R\cos\lambda = 4\,262.362$  km. The great circle angular separation (using the formula in Table C.4) is:  $3.34507^\circ$ . Accordingly, the great circle distance is 371 896 m. This is about 65 m shorter than the distance along the small circle. The direct (Euclidian) distance along a hypothetical tunnel is:  $x = 371\,843$  m. It is only about 53 m shorter than the great circle distance.

**Problem 2.6.** At the given angular velocity, a full rotation takes  $2\pi/(10^{-15}\text{s}^{-1}) \approx 200$  my. The equatorial circumference of the earth is  $2R\pi \approx 40\,000$  km. Thus, the plate velocity at the equator is roughly 20 cm per year. At  $48^\circ$  north the small circle circumference is only 26 781 km (eq. 2.2). Thus, the relative plate velocity at this latitude is only 13 cm per year.

**Problem 2.7.** The torque at the equator is:  $M = R \cdot 10^{12}$  Nm. North and south of the equator we can write:  $M = r \cdot 10^{12}$  Nm. Using  $r = R\cos\lambda$  (e. 2.2) we get a latitude of:  $\lambda \approx 51^\circ$  either north or south.

**Problem 2.8.**  $x = C_1\phi$  and  $y = C_2\tg\lambda$ . The constants  $C_1$  and  $C_2$  can have any value for example 1. They are not really necessary, but are introduced here in order to show that the Mercator projection can be scaled to rectangles of any aspect ratio.

**Problem 2.9.** In thermally stabilized continental lithosphere, the base of the crust is at about  $z_c = 30 - 40$  km depth, the base of the lithosphere is at about  $z_l = 100 - 200$  km depth. The temperature at the Moho is about  $500^\circ\text{C}$  and at the base of the lithosphere it is about  $T_l \approx 1\,200 - 1\,300^\circ\text{C}$ . Thus, at least  $100 - 200^\circ\text{C}$  of the Moho-temperature may be attributed to radioactive heat. The mean density of crustal rocks is of the order of  $\rho_c = 2\,800\text{ kg m}^{-3}$  and that of mantle rocks in the asthenosphere about  $\rho_l = 3\,200\text{ kg m}^{-3}$ . The thermal expansion coefficient of rocks is about  $\alpha = 3 \times 10^{-5}\text{ }^\circ\text{C}^{-1}$  so that the density changes by several percent within both the crust and the mantle part of the lithosphere.

**Problem 2.10.** a) The relative velocities may be read from Table 2.3. b) A new plate boundaries is currently forming along the East African Rift system. However, how the world will look like in future is very much a trial and error game. Have a look at <http://www.scotese.com>. c) Ancient plate boundaries exist just about anywhere where there is ancient mountain belts. Fig. 2.3 gives an overview over the last 170 my.

**Problem 2.11.** There is a total of 10 possible triple junctions: RRR, FFF, TTT, RRT, RRF, FFT, FFR, TTF, TTR, RTF. RRR-triple junctions are always stable, FFF-triple junctions are always unstable. Most other triple junctions may occur in stable or in unstable configuration. Whether they

are unstable or not does not depend on the qualitative nature of the plate motion, but rather on the angles and relative velocities.

**Problem 3.1.** The amount of heat in one gram of mass is  $10^{-3} \text{ kg} \times 300\,000^2 \text{ km}^2 \text{ s}^{-2} = 9 \cdot 10^{13} \text{ J}$ , which is enough to light a 60 W globe for  $9 \cdot 10^{13} \text{ J} / 60 \text{ W} \approx 48\,000 \text{ years}$ .

**Problem 3.2.** See Table D.8.

**Problem 3.3.** It may be seen intuitively that the ore body will be nearly isothermal if it has a high conductivity. As a consequence, the isotherms above and below the ore body have a closer spacing (Fig. E.2).

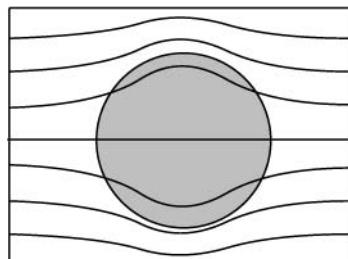
**Problem 3.4.** At the contact between two rocks of different conductivity the heat flow must be the same on both sides of the contact. Thus, for the given linear thermal gradient the temperatures at different depths will be:  $T_{5\text{km}} = 100^\circ\text{C}$ ;  $T_{7\text{km}} = 120^\circ\text{C}$  and  $T_{10\text{km}} = 180^\circ\text{C}$ .

**Problem 3.5.** The thermal time constant of the pile is roughly 40 my. This is significantly longer than the deformation in a), but significantly shorter than that in b). ( $\dot{\epsilon} = 10^{-12} \text{ s}^{-1}$  implies that the pile doubles in thickness in a mere 30 000 years). The implication of this result is that the deformation in b) occurs largely in thermal equilibrium. In contrast, in a) the crust will be out of thermal equilibrium at the end of deformation. The deformation will be followed by thermal equilibration. Metamorphic parageneses are expected to grow *across* the fabric in a) but grow syn-deformational in b).

**Problem 3.6.** Using eq. 3.17 it is easy to see that the thermal time constant (giving an estimate of the duration of heat conduction) of a 10 km large body is at least one order of magnitude longer than the questioned period of observation of  $10^5$  years. Heat conduction may therefore be neglected and the temperature is given (from eq. 3.21) by  $T = t \times S_{\text{rad}} / \rho c_p$ . Using the given values for density and heat capacity this implies a temperature rise of about  $T \approx 100^\circ\text{C}$ .

**Problem 3.7.** a) The product of deviatoric stress and strain rate gives minimum and maximum values for the heat production rate of  $3 \cdot 10^{-5}$  and  $3 \cdot 10^{-8} \text{ W m}^{-3}$ , respectively (eq. 3.25). These minimum and maximum values

**Figure E.2.** Answer to Problem 3.3



are significantly higher and smaller than typical values for radiogenic heat production rates, respectively. b) Multiplying these values with a deformation period of 1 my, the total shear heat production is  $10^9 \text{ J m}^{-3}$  for the largest stresses and strain rates and  $10^6 \text{ J m}^{-3}$  for the small stresses and strain rates. Thus, according to eq. 3.26, the maximum temperature that can be reached is given by the ratio of volumetric heat content and  $\rho c_p$ . This gives  $370^\circ\text{C}$  and  $0.37^\circ\text{C}$  for the maximum and minimum values, respectively. Whether or not these temperatures can be reached depends on the efficiency of heat conduction away from the site of heat production and on the change in strength of the shear zone during deformation. The former depends on the thickness of the heat producing shear zone.

**Problem 3.8.** None! All heat of the fire is reaction heat (chemically produced heat). It is released by the exothermic chemical reaction:  $\text{wood} \rightarrow \text{CO}_2 + \text{water}$ .

**Problem 3.9.** The duration for which the rock will remain at constant temperature is given by the ratio of *total amount of released fusion heat to rate of heat loss by conduction*. For example, if  $x \text{ J}$  are produced by the crystallization and the conductive cooling causes a heat loss of  $x \text{ Js}^{-1}$ , then if the cooling history will be halted for exactly 1 second. The volumetric amount of heat that is released at the solidus is  $0.3 \times L\rho$ . The rate of heat loss due to cooling is  $dT/dt\rho c_p$  (s. eq. 3.3). Thus, the duration of thermal buffering is  $= L/c_p(dt/dT) \approx 1 \text{ my}$ . The fact that the cooling history is halted for 1 my may imply that metamorphic parageneses re-equilibrate partially or fully at this temperature.

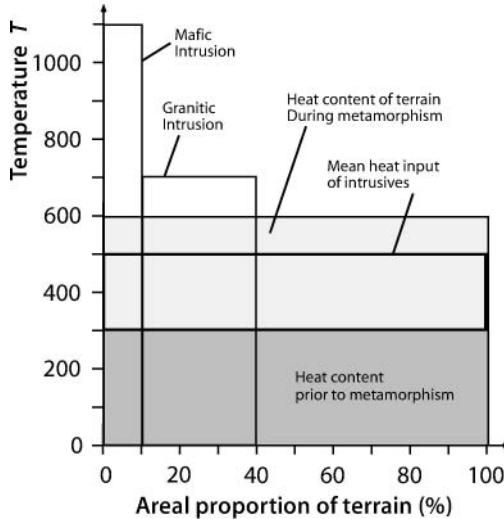
**Problem 3.10.** Using eq. 3.51 and a standard value for the thermal diffusivity of  $\kappa = 10^{-6} \text{ m}^2 \text{s}^{-1}$ , the three advection (erosion) rates give Peclet numbers of  $Pe \approx 0.1$ ,  $Pe \approx 1$  and  $Pe \approx 5$ , respectively. Thus, the slow erosion rate of a) warrant description of the regional thermal evolution with consideration of conduction only (e.g. eq. 3.6), while the intermediate rates of b) require consideration of both conductive and advective processes (e.g. eq. 3.42) and the rapid rates of c) would allow to consider the thermal evolution as a solely advective process (e.g. eq. 3.40).

**Problem 3.11.** The answer is discussed in eqs. 6.14 and 6.15.

**Problem 3.12.** a) Eq. 3.80 shows that the error function must have a value of 0.8333 so that the temperature is  $T = 1000^\circ\text{C}$ . For this, Fig. 3.6 shows that the argument of the error function must be  $\approx 0.98$  (thus:  $0.98 = z/\sqrt{4\kappa t}$ ). From this we get:  $z = 98 \text{ km}$ . b)  $0 \text{ km} / 0^\circ\text{C}$ ;  $10 \text{ km} / 371^\circ\text{C}$ ;  $20 \text{ km} / 689^\circ\text{C}$ ;  $30 \text{ km} / 921^\circ\text{C}$ ;  $40 \text{ km} / 1067^\circ\text{C}$ ;  $50 \text{ km} / 1144^\circ\text{C}$ ;  $75 \text{ km} / 1196^\circ\text{C}$ ;  $100 \text{ km} / 1199^\circ\text{C}$ .

**Problem 3.13.** a) Using  $t_{eq} = l^2/\kappa$  the result is  $\approx 80$  years. b) for symmetry reasons this temperature is half way between the intrusion and the host rock

**Figure E.3.** Solution of Problem 3.14



temperature if both have the same heat capacity:  $500\text{ }^{\circ}\text{C}$ , directly at the contact. c)  $50\text{ m} \cdot (T_i - T_b)\rho c_p = 5.4 \cdot 10^{10}\text{ J/m}^2$  dike surface. d) 40 years is roughly half of the total thermal history. The cooling of the dike is most rapid at the start. Thus, the dike will have cooled a bit more than half by then. The area underneath the  $T$ - $z$ -profile must remain constant as no heat may be lost. This information may be used to estimate the width of the contact metamorphic halo.

**Problem 3.14.** The depth of the rocks multiplied with the geothermal gradient gives the terrain temperature prior to metamorphism:  $300\text{ }^{\circ}\text{C}$ . If 10% of the terrain were at  $1100\text{ }^{\circ}\text{C}$  and 30% of the terrain area at  $700\text{ }^{\circ}\text{C}$ , and the remainder at  $300\text{ }^{\circ}\text{C}$ , then the mean temperature is  $\bar{T}_{cool} = 0.1 \times 1100 + 0.3 \times 700 + 0.6 \times 300 = 500\text{ }^{\circ}\text{C}$ . This is  $200\text{ }^{\circ}\text{C}$  above the host rock temperature. Thus, the heat of the intrusives is insufficient to explain metamorphism if latent heat of crystallization is neglected. The additional temperature contribution of the crystallization heat from 40% of the terrain is  $T_{cryst} = 0.4 \times L/c_p = 128\text{ }^{\circ}\text{C}$ . Thus, the potential metamorphic temperature is  $\bar{T}_{cool} + T_{cryst} = 628\text{ }^{\circ}\text{C}$ , as the mean potential temperature for the whole terrain. This is within 10%-error of the observed metamorphic temperature. Contact metamorphism is therefore a plausible model explanation for the observations. Fig. E.3 illustrates the problem answer graphically.

**Problem 3.15.** The depths and temperatures of the starting condition are:  $z[1] = 0, T[1] = 0$ ;  $z[2] = 10\text{ km}, T[2] = 125\text{ }^{\circ}\text{C}$ ;  $z[3] = 20\text{ km}, T[3] = 250\text{ }^{\circ}\text{C}$ ;  $z[4] = 30\text{ km}, T[4] = 375\text{ }^{\circ}\text{C}$ ;  $z[5] = 40\text{ km}, T[5] = 500\text{ }^{\circ}\text{C}$ ;  $z[6] = 50\text{ km}, T[6] = 0\text{ }^{\circ}\text{C}$ ;  $z[7] = 60\text{ km}, T[7] = 125\text{ }^{\circ}\text{C}$ ;  $z[8] = 70\text{ km}, T[8] = 250\text{ }^{\circ}\text{C}$ ;  $z[9] = 80\text{ km}, T[9] = 375\text{ }^{\circ}\text{C}$ ;  $z[10] = 90\text{ km}, T[10] = 500\text{ }^{\circ}\text{C}$ . The time stepping must be smaller than 0.7937 my. In order to perform the calculations with a round number it is useful to use something like  $\Delta t = 0.5\text{ my} = 1.575 \cdot 10^{13}\text{ s}$ , i.e.

$R = 0.1575$ . After  $1\Delta t$  the temperatures are:  $T[5] = 402^\circ\text{C}$  and  $T[6] = 98^\circ\text{C}$ . After  $2\Delta t$  the temperatures are:  $T[4] = 360^\circ\text{C}$ ,  $T[5] = 350^\circ\text{C}$ ,  $T[6] = 150^\circ\text{C}$  and  $T[7] = 140^\circ\text{C}$ . All other temperatures remain up to this time constant because there is no spatial curvature of the temperature profile around them (eq. 3.6).

---

**Problem 4.1.** Using eq. 4.8 we get a distance of the summit of Mt Everest of 6 382.207 km from the center of the earth and 6 384.32 km for Chimborazzo. Thus, Chimborazzo is actually about 2 100 m higher than Everest, if measured from the center of the earth.

**Problem 4.2.** a) There is *no* rock uplift at all. b) The rock is not exhumed at all. It is buried by 1 km. This example is actually what occurs to rocks on the Tibetan Plateau at the moment where extreme surface uplift and sedimentation occur at the same time.

**Problem 4.3.** a) Morphological equilibrium will be reached after a long time. Thus, eq. 4.18 may be used for  $t \rightarrow \infty$ :  $t_E = Hb/(\dot{\epsilon}(H + a)) = 1.834 \text{ my}^{-1}$ . b) If  $z = 0$  and  $t = 40 \text{ my}$  are inserted into eq. 4.19 and this is solved for  $z_i$  we get:  $z_i = 10\,202 \text{ m}$ . c) This may be evaluated by finding the derivative of eq. 4.19 with respect to  $t$ , setting the result to 0 and solving for  $t$ . The result is 11.8 my.

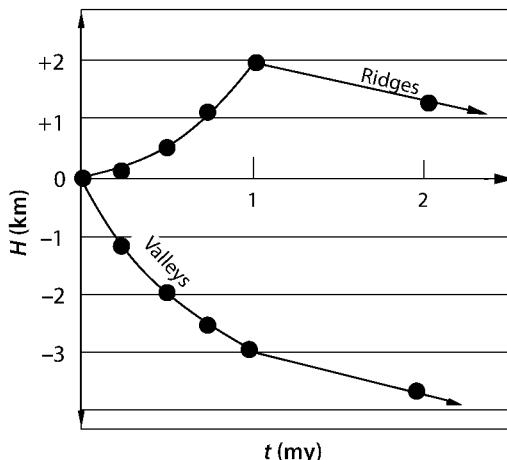
**Problem 4.4.** Eq. 4.23 is solved in eq. 4.26. Adapting the variables by inserting  $z_i = 1$  for a randomly chosen thickness of the iceberg of 1 and  $H = 0.1$ , because 10% of this thickness are above water and using the density of water for  $\rho_m$  we get for the density of ice:  $\rho_{\text{ice}} = 900 \text{ kg m}^{-3}$ .

**Problem 4.5.** The solution is shown graphically in Fig. E.4.

**Problem 4.6.** a)  $\Delta H = 10 \text{ km} (\rho_m - \rho_u)/\rho_m = 937 \text{ m}$ . Density and thickness of the crust do not change and therefore need not be considered. b)  $\Delta H = 391 \text{ m}$ . c) the thinning of the crust in b) causes a subsidence of  $(f_c - 1)(\rho_m - \rho_c)/\rho_m z_c = 2\,343 \text{ m}$ . This is partially balanced by an uplift of 391 m by the basaltic underplate. Relative to a) there is a total elevation difference of  $937 + (2\,343 - 391) = 2\,889 \text{ m}$ .

**Problem 4.7.** a) The answer is:  $2.5 \text{ cm y}^{-1}$ . From this and the given data the water depths are: 469, 878, 1 049, 1 483, 1 756, 1 906, 2 392, 2 633, 2 795, 2 930, 3 318, 3 709, 3 855 and 3 995 m for the given distances from the mid oceanic ridge. b) These model values correspond well with the observed values up to an age of roughly 100 my. From there on, sedimentary loading of the plate, its increasing strength and changes of the lower boundary condition by mantle convection cause the divergence between model and observation.

**Problem 4.8.** a) The units of  $D$  are Nm. The fourth derivative has the units of  $\text{m}^{-4}$  and  $q$  has the units of stress or force per area. The load is the vertical stress.  $F$  has therefore the units of  $\text{N m}^{-1}$ , or Newton per meter length of orogen. b)  $D_{h=10\text{km}} \approx 8.8 \times 10^{21} \text{ Nm}$  and  $D_{h=70\text{km}} \approx 3 \times 10^{24} \text{ Nm}$ .



**Figure E.4.** Solution of Problem 4.5. The mean vertical denudation  $x$  within 1 my is given by the relationship  $lx = (vt)^2$ . This relationship was derived geometrically from Fig. 4.49. Thermal expansion or the influence of the mantle lithosphere need not be considered as they remain constant. The mean subsidence at any time is given by:  $x(\rho_m - \rho_c)/\rho_m$ . Thus, the elevation of the ridges is after 0.25 my: 132 m, after 0.5 my: 527 m, after 0.75 my: 1 187 m and after 1 my: 2 110 m. Thereafter, the ridges subside with a rate of 781 m my $^{-1}$ . The subsidence of the valley floors may be easily derived from the product of incision rate and time subtracted from the results for the ridge elevations

**Problem 4.9.** a) The given hints are required to determine the constants of integration. For the first and second integration they the constants must be:  $C_1 = -qL/D$  and:  $C_2 = qL^2/(2D)$ , respectively. For the third and fourth integration the constants must be:  $C_3 = 0$  and  $C_4 = 0$ , respectively. Thus:  $w = qx^2/D (x^2/24 - Lx/6 + L^2/4)$ . b) At  $x = L$  the deflection is:  $w = 3qL^4/(24D)$ . Using the parameters given in the problem we get:  $D = 0.0008$  Nm.

**Problem 4.10.** a) For the lowest point where there is a density difference between the two columns we can write: column A:  $\sigma_{zz} = \rho_c g(H + z_c + w)$ ; column B:  $\sigma_{zz} = \rho_c g z_c + \rho_m g w$ . The buoyancy is given by the difference between these two values. This is the origin of eq. 4.50. b) If the load is  $q_a$  is not considered, then we can write for the downward force per unit area:  $(\rho_c - \rho_w)wg$ , and for the upward directed force:  $(\rho_m - \rho_c)wg$ . The net force per unit area is therefore:  $wg(\rho_w - \rho_m)$ . The total load is:  $q = q_a + wg(\rho_w - \rho_m)$ .

**Problem 4.11.**  $\alpha \approx 65$  km.

**Problem 4.12.** At the highest point of the elastic bulge the slope of the plate is  $dw/dx = 0$ . The slope of the plate as a function of distance is given by the first spatial derivative of eq. 4.54. If this is set to zero:  $0 = -2w_0/\alpha e^{-x/\alpha} \sin(x/\alpha)$ , we get:  $x = \pi\alpha$ . This is the highest point of

the fore bulge. In 250 km distance from this point we get from eq. 4.53:  $D = 2.2 \cdot 10^{23}$  Nm, and eq. 4.49:  $h \approx 33$  km. The value of  $w_0$  does not change the distance of the forebulge from the load and its knowledge is therefore not required.

**Problem 4.13.** The problem is easiest solved by plotting the data and graphing eq. 4.67 iteratively for different times until a best fit for the data is found. For this a half height of the scarp  $a \approx 7$  m should be used, which is roughly half the elevation difference between base and top. It is also useful to subtract roughly 150 m from all  $x$  values so that the inflection point of the scarp is at  $x = 0$ . Then, a best fit gives an age of about 30 000 years.

**Problem 4.14.** The two boulders have been exposed to weathering for the same time. This result is easiest arrived at by measuring  $l$  (for example in millimeters) off the photograph for the two different boulders and comparing the topographic shapes of the boulders with plots of eq. 4.69. Neither the magnitude of  $D$ , nor the actual length of  $l$  need be known as we perform only a comparison between the two profiles.

**Problem 5.1.** The stretch is 2, the elongation is 1. The vertical strain is 100%.

**Problem 5.2.** Force = mass  $\times$  acceleration. As weight is a force and the acceleration on the surface of the earth is  $g = 9.81$  m/s<sup>2</sup>, a mass of 1 kg hat has the weight of:  $1 \text{ kg} \cdot 9.81 \text{ m/s}^2 = 9.81 \text{ N}$ .

**Problem 5.3.** Energy is given by the product of force times distance. Thus, the collision process releases  $10^{13} \text{ N} \times 10^5 \text{ m} = 10^{18} \text{ J}$ . The most important forms of energy into which this mechanical energy is transformed are frictional heat and potential energy of the mountain belt that formed in response to the collision.

**Problem 5.4.** As there are no shear stresses and the problem is two dimensional, we can assume that  $\sigma_{zz} = \sigma_1$  and  $\sigma_{xx} = \sigma_2$  (there is no  $\sigma_3$ ). For the continent that is pulled apart by the subducting plate we can write:  $\sigma_{xx} = \sigma_2 = -A$  and  $\sigma_{zz} = \sigma_1 = 0$ . From this, the mean stress (pressure) is:  $\sigma_m = -A/2$ . According to eq. 5.10 the vertical and horizontal components of deviatoric stress are:  $+A/2$  and  $-A/2$ , respectively. For the continent that collapses under the weight of a mountain range we can write:  $\sigma_{xx} = \sigma_3 = 0$  and  $\sigma_{zz} = \sigma_1 = A$ . From this, the mean stress is given by:  $\sigma_m = A/2$ . The vertical and horizontal components of deviatoric stress are also:  $+A/2$  and  $-A/2$ . Thus, we can conclude that state of deviatoric stress is the same in both continents but the pressure is different.

**Problem 5.5.** As no shear stresses apply to the surfaces of the rock, all components of eq. 5.15 that describe shear stresses are zero. Eq. 5.15 simplifies to:  $d\sigma_{zz}/dz = \rho g$ . The corresponding equations for the other two

spatial directions in which there is no gravitational acceleration simplify to:  $d\sigma_{xx}/dx = 0$  and  $d\sigma_{yy}/dy = 0$ . Integration of these three equation gives:  $\sigma_{xx} = C_1$ ,  $\sigma_{yy} = C_2$  and  $\sigma_{zz} = \rho g z + C_3$ .  $C_1$ ,  $C_2$  and  $C_3$  are the constants of integration. They are determined by the boundary conditions of the problem. The normal stresses can be assumed to be zero on each free surface (Fig. 5.34):  $\sigma_{zz} = 0$  at  $z = -H$ . From this we get:  $C_3 = \rho g H$  and  $\sigma_{zz} = \rho g(H+z)$ . Furthermore, because there is no stresses on the sides it must be true that:  $C_2 = C_1 = 0$ ;  $\sigma_{xx} = \sigma_{yy} = 0$ . The mean principal stress (pressure  $P$ ) as a function of depth is:  $P = (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3 = \rho g(H+z)/3$ . The principal components of the deviatoric stress tensor are, according to eq. 5.7:  $\sigma_{xx} - P = \sigma_{yy} - P = -\rho g(H+z)/3$  and  $\sigma_{zz} - P = 2\rho g(H+z)/3$ .

**Problem 5.6.** Similar to problem 5.5, integration of the stress balance equations gives:  $\sigma_{xx} = C_1$ ;  $\sigma_{yy} = C_2$  and  $\sigma_{zz} = \rho g z + C_3$ . As fluids support only negligible differential stresses it is true that:  $\sigma_{xx} = \sigma_{yy} = \sigma_{zz}$ . With the same boundary condition for the vertical component of stress as in Problem 5.5 we get:  $\sigma_{xx} = \sigma_{yy} = \sigma_{zz} = \rho g(z+H) = P$ . Thus, at any given depth, the pressure in the fluid is three times as high as it is in the rock of Problem 5.5. All components of the deviatoric stress tensor are zero.

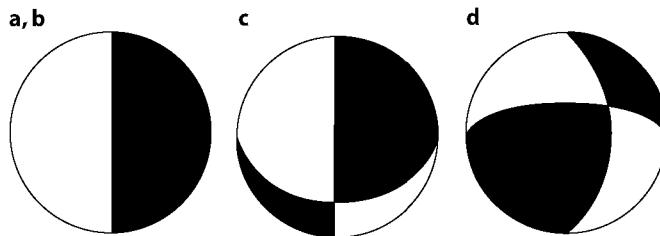
**Problem 5.7.** If the side wall is fixed, then the stain rate is:  $\dot{\epsilon} = 0$  and  $\sigma_{zz} = \sigma_{xx}$ . The mean force per area is given by the vertically integrated vertical stresses  $\rho g(z+H)$  using the limits of integration  $-H$  and  $0$  from  $F_b = \rho g H^2/2$ . The mean horizontal stress  $\bar{\sigma}_{xx}$  is therefore  $\bar{\sigma}_{xx} = F_b/H = \rho g H/2$ .

**Problem 5.8.** In the fluid:  $\sigma_{xx} = \sigma_{yy} = \sigma_{zz} = P = \rho_m g z$ . In the rock the horizontal stresses for  $z < 0$  are:  $\sigma_{xx} = \sigma_{yy} = 0$ , and for  $z > 0$  they are:  $\sigma_{xx} = \sigma_{yy} = \rho_m g z$ . The vertical stresses in the whole rock are:  $\sigma_{zz} = \rho_c g(H+z)$ . The pressure in the rock is given by the mean of the principal stresses. For  $z < 0$  it is:  $P = (\rho_c g(H+z))/3$ . For  $z > 0$  it is:  $P = (2\rho_m g z + \rho_c g(H+z))/3$ . All principal components of the deviatoric stress tensor in the fluid are 0. In the rock, for  $z < 0$  they are:  $\tau_{xx} = \tau_{yy} = \sigma_{xx} - P = \sigma_{yy} - P = -(\rho_c g(H+z))/3$  and:  $\tau_{zz} = \sigma_{zz} - P = (2\rho_c g(H+z))/3$ . For  $z > 0$  they are:  $\tau_{xx} = \tau_{yy} = \sigma_{xx} - P = \sigma_{yy} - P = (\rho_m g z - \rho_c g(z+H))/3$  and  $\tau_{zz} = \sigma_{zz} - P = -(2\rho_m g z - 2\rho_c g(z+H))/3$ .

**Problem 5.9.** From eq. 5.10 we get:  $\epsilon = \sigma/E = 5 \cdot 10^7 \text{ Pa} / 5 \cdot 10^{10} \text{ Pa} = 0.001 = 0.1\%$ .

**Problem 5.10.** The vertical stress at any given depth is:  $\sigma_{zz} = \rho g z$ . The strain of a unity cube at depth  $z$  in the lithosphere is therefore:  $\epsilon = \rho g z/E$ . From this, the total strain integrated over the entire lithosphere is given by:  $\int_0^{z_1} \rho g z/Edz = \rho g z^2/2E$ . Using the given numerical values for the parameters this is: 2500 m. Thus, the elastic thickness change of the total lithosphere due to its own weight is roughly 2.5 % of its total thickness.

**Problem 5.11.** The solution is graphically shown in Fig. E.5.



**Figure E.5.** Solution of Problem 5.11

**Problem 5.12.** For the quartz data the stresses are  $\sigma_d = 117 \text{ MPa}$ ;  $\sigma_d = 24 \text{ MPa}$  and  $\sigma_d = 11 \text{ MPa}$ , for  $n = 2$ ,  $n = 3$  and  $n = 4$ , respectively. For olivine the stresses are  $\sigma_d = 1.4 \times 10^8 \text{ MPa}$ ;  $\sigma_d = 2.7 \times 10^5 \text{ MPa}$  and  $\sigma_d = 1.2 \times 10^4 \text{ MPa}$ , for  $n = 2$ ,  $n = 3$  and  $n = 3$ , respectively.

**Problem 5.13.** When taking the logarithm of eq. 5.46 it becomes linear. It then has the form:  $\log(\dot{\epsilon}) = n \log(\sigma_d) + \log(A) - Q/RT$ . Thus, if the experimental results are plotted in a diagram of  $\log(\dot{\epsilon})$  against  $\log(\sigma_d)$  the slope of the data gives the power law exponent  $n$  and the intercept is  $(\log(A) - Q/RT)$ . Replotting the data from a constant stress on a diagram of  $\log(\dot{\epsilon})$  against  $1/T$  gives a line with the slope of  $-Q/R$  from which the activation energy may be determined.

**Problem 5.14.** It is because in oceanic lithosphere olivine dominates the lithospheric rheology at much shallower levels (and therefore colder temperatures) than in continental lithosphere.

**Problem 5.15.** According to eq. 5.50 it is:  $E_p = \int_0^{-H} \sigma_{zz} dz = \int_0^{-H} \rho g(H + z) dz = \rho g H^2 / 2$ .

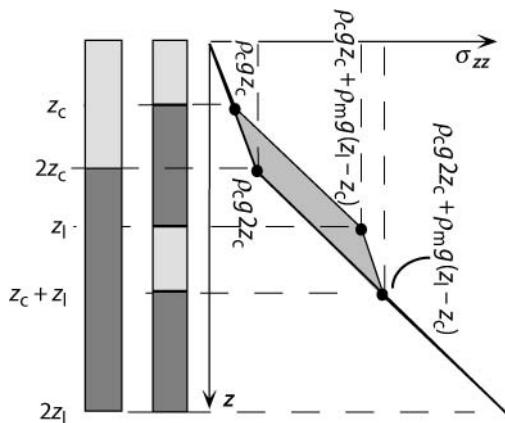
**Problem 5.16.** The solution is  $F_b \approx 10.5 \times 10^{12} \text{ N m}^{-1}$ . The result is the same as that from Problem 5.19 as can also be seen from Fig. E.6

**Problem 5.17.** Eq. 5.57 gives a buoyancy force of  $F_b \approx 6.6 \times 10^{12} \text{ N m}^{-1}$ .

**Problem 5.18.** Inserting the given values into eq. 5.59 give a horizontal buoyancy force of the mountain range of  $F_b \approx 7.71 \times 10^{12} \text{ N m}^{-1}$ . This result may also be read from Fig. 5.32. There, the point  $f_c = 2$ ,  $f_l = 1$  gives a value of about  $F_b \approx 8.3 \times 10^{12} \text{ N m}^{-1}$ . The difference arises from slightly different assumptions of the densities and thicknesses. The result is larger than that from eq. 5.57 because the refined equation (eq. 5.59) also considers thermal contraction of the mantle part of the lithosphere.

**Problem 5.19.** The graphical solution of this problem is shown in Fig. E.6. The net force between the two columns corresponds to the shaded area on this figure (see also Problem 5.16). Calculating this area algebraically, gives a net

**Figure E.6.** Graphical solution of Problem 5.19



force of  $10.5 \cdot 10^{12} \text{ N m}^{-1}$  towards the homogeneously thickened lithosphere. This is comparable to the magnitude of plate tectonic driving forces, even though there is *no* elevation difference between the two lithospheric columns.

**Problem 5.20.** The kinetic energy ( $E_{kin} = mv^2/2$ ) of the plate is:  $E_{kin} = 3 \cdot 10^{20} \text{ kg} \cdot (0.03 \text{ m/y})^2 = 271 \text{ J}$  (or  $\text{kg m}^2 \text{ s}^{-1}$ ). The potential energy per cubic meter is:  $E_p = \rho gh$ . Thus, with a kinetic energy of 271 J we can lift one cubic meter of rock (using  $\rho = 2700 \text{ kg m}^{-3}$  and  $g = 10 \text{ m s}^{-2}$ ) merely by about 1 cm. Consequently, the kinetic energy is completely insignificant to have an influence on the orogenic force balance. Plate motions are too slow!

---

**Problem 6.1.** Fig. 6.2 shows that the elevation contours have a slope of roughly  $(f_c/f_l) = 1.333$ . Thus, during extension (when both  $f_c$  and  $f_l$  become smaller than 1) uplift will occur when  $(f_c/f_l) > 1.333$ ; for the assumed thicknesses and densities of Fig. 6.2. Eq. 4.35 shows that the elevation does not change when  $\delta z_c(f_c - 1) = \xi z_l(f_l - 1)$  and subsidence will occur when  $\delta z_c(f_c - 1) < \xi z_l(f_l - 1)$ . This can be solved for any of the variables in that equation.

**Problem 6.2.** According to eq. 6.1 we can write for the cross over between sandstone and shale porosities:  $\phi_0^{sand} \exp(-c^{sand} z) = \phi_0^{shale} \exp(-c^{shale} z)$ . This can be solved for  $z$  by taking the logarithm of the equation giving:  $z = (\log(\phi_0^{sand}) - \log(\phi_0^{shale})) / (c^{sand} - c^{shale})$ . Using the parameter values from Fig. 6.5 we get  $z = 1116 \text{ m}$ .

**Problem 6.3.** Inserting eq. 6.1 into eq. 6.4 and using the parameter values from Fig. 6.5 gives  $L_0 = 130 \text{ m}$ .

**Problem 6.4.** According to eq. 4.23 we can formulate from Fig. 6.7:  $Lg\rho_L + z_c g \rho_c = z_s g \rho_w + z_c g \rho_c + (L - z_s) g \rho_m$ . Solving for  $z$  we get eq. 6.5 with a reversed sign. The different sign arises because the surface is *lower* before sedimentation, whereas we have automatically changed the sign in eq. 6.5 to

deal only with positive numbers. The equation that describes the change if water depth may be derived analogously.

**Problem 6.5.** The answer is shown in Fig. E.7. The first three columns show the field data: *a* illustrates thickness and lithology (sandstone: gray, slate: lines), *b* shows the age of the lithological boundaries and *c* the water depth. The strata are numbered from bottom to top from 1 to 5. For the following calculations the center points of each layer (shown in the 4th column) are used as reference point. For the porosities we use the data from Fig. 6.5. Using eq. 6.1 for the uppermost, 1 000 m thick layer we get a porosity of 34.4 % at the mean depth of 500 m (1st column of Table *d*). The thickness of each layer is based on field measurements and their densities are given by eq. 6.7. In the second column the data for porosity, thickness and density *after* removal of the uppermost layer (*i* = 5) are plotted. For the second layer from the top (*i* = 4) that has a mean depth of 1 250 m, we begin with a determination of porosity. Its decompactacted thickness is then calculated with eq. 6.4 using the porosities of the first column and the decompactacted porosities. The porosity of the third layer after removing the first two strata is given by eq. 6.1 at a depth of 655 m + 500 m and so on. The same principle is repeated for all columns using always the data from the fist column so that no methodical errors are propagated. The sum of the thicknesses and densities if the profile (in the bottom row) are calculated with eq. 6.8 and 6.9. *e* is a graphical illustration of the data from Table *d*.

**Problem 6.6.** To determine the ratio we take the subsidence in eq. 6.10 to be zero, i.e.:  $H = 0$ . Thus, after rearranging:  $z_c/z_l = \rho_l \alpha T_l / (2(\rho_l - \rho_w))$ . For the given numerical values of the parameters:  $z_c/z_l = 13.7$  km. At smaller starting ratios of the thicknesses, the contribution of the mantle part of the lithosphere is so large, so that homogeneous stretching of the lithosphere (when thinned) will lead to surface *uplift*.

**Problem 6.7.** Eq. 6.11 shows that the duration of the thermal sag period is only determined by the ratio  $t/t_{eq}$  in the last term of the equation. The time scale of thermal equilibration  $t_{eq}$  is the same we have discussed in eq. 3.17 and Table 3.2. It differs from that only by the factor  $\pi^2$ , but is in principle equivalent. We can see that when  $t = t_{eq}$ , then  $H_{sag}$  has reached  $(1 - e^{-1}) \times 100\%$  completion. 90% of the thermal sag is completed when:  $(1 - e^{-t/t_{eq}}) = 0.9$  or:  $t = -t_{eq} \ln(0.1)$ . This is  $\approx 75$  my for 100 km thick lithosphere and  $\approx 300$  my for 200 km thick lithosphere.

**Problem 6.8.** For the thickened lithosphere we can use eq. 6.15. We get:  $q_s = 0.06 \text{ W m}^{-2}$ ;  $q_s = 0.075 \text{ W m}^{-2}$  and  $q_s = 0.09 \text{ W m}^{-2}$ , for the three heat source distributions, respectively. For the thinned lithosphere we must use  $q_s = 2q_m + q_{rad}/2$  and get:  $q_s = 0.0525 \text{ W m}^{-2}$ ;  $q_s = 0.075 \text{ W m}^{-2}$  and  $q_s = 0.0975 \text{ W m}^{-2}$ , respectively. We can see that the first distribution results in an *increase* of the surface heat flow with increasing thickness, while the last distribution results in a *decrease* of the surface heat flow with increasing thickness.

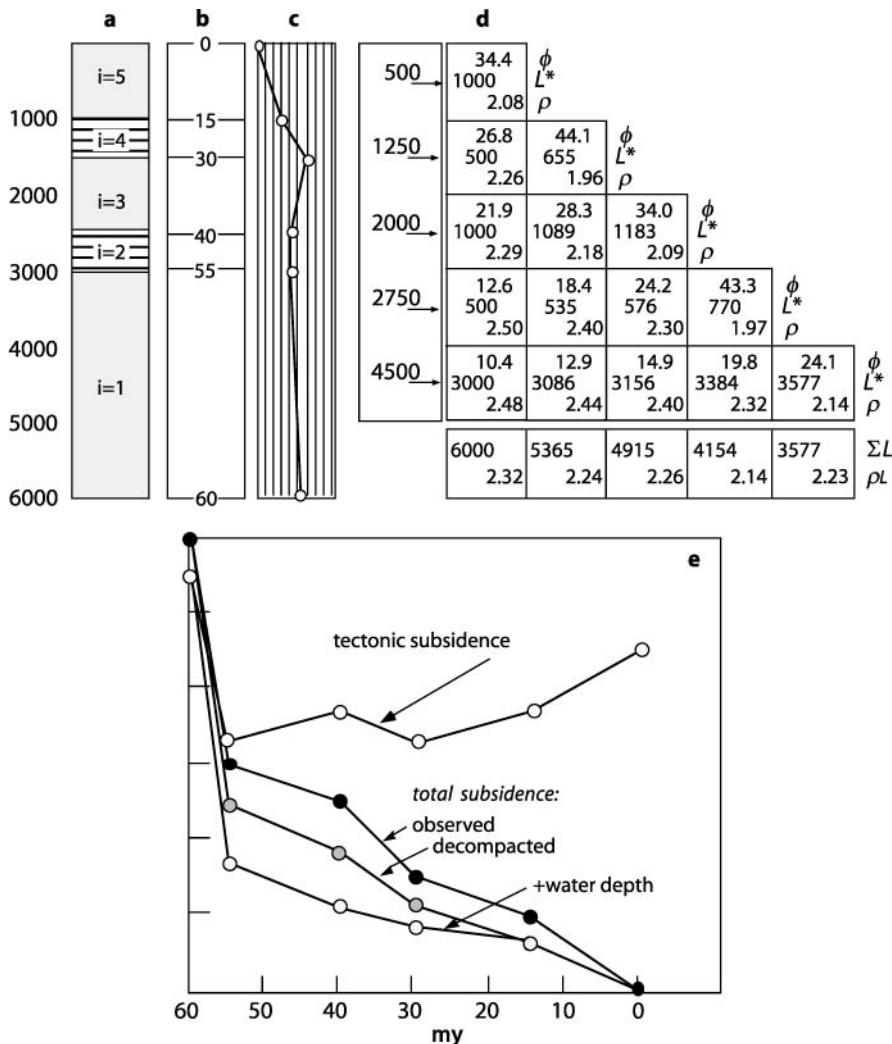


Figure E.7. Solution of Problem 6.5

**Problem 6.9.** (a) Fig. 5.32 shows that the buoyancy force  $F_b$  balances the driving force roughly at  $f_c \approx 1.65$  (while  $f_1 = 1$ ). Fig. 4.18 shows that this corresponds to a surface elevation of about  $H \approx 3.5$  km. (b) In order to use eq. 5.57, the thickness of the root  $w$  must be substituted by  $w = H\rho_c/\Delta\rho$ . It then can be solved for to give  $H = 4400$  m at a buoyancy force of  $F_b = F_d = 5 \cdot 10^{12}$  N m $^{-1}$ . (c) Eq. 5.59 simplifies significantly for  $f_1 = 1$ . It then can be used to derive  $f_c = 1.7$  to give  $F_b = F_d = 5 \cdot 10^{12}$  N m $^{-1}$ . From this, eq. 4.31 gives a surface elevation of  $H = 3800$  m. The estimate from

(b) is substantially higher than those from (a) and (c) because the thermal contraction in the mantle part of the lithosphere is not considered here.

**Problem 6.10.** Using the parameter values from Table 5.3 and eq. 6.31 gives:  $Ar^{(T=400)} \approx 0.27$  and  $Ar^{(T=400)} \approx 15$ . Thus, deformation in the “colder” orogen will be concentrated around the collision zone establishing substantial topography and potential energy contrasts, while deformation in the “warmer” orogen will be dissipated over a large region with little topography. However, note that the Argand number is used in geodynamics in order to *avoid* the use of ill-constrained material constants and temperatures in as much detail as we do here. Thus, it is not generally meaningful to use examples like this one for any quantitative consideration of a real orogen.

**Problem 6.11.** The lithostatic component of pressure in 10 and 15 km depth will be  $\sigma_{zz} = \rho g z = 275$  and 321 MPa, respectively. The temperature at this depth is 300 °C and 450 °C, respectively. The strain rate is: 1/5 my<sup>-1</sup>. From the power law we get differential stresses of 824 MPa and 40 MPa, respectively. According to eq. 6.33 the pressure component related to differential stress in 10 or 15 km depth 412 MPa or 20 MPa, respectively. This means that the non lithostatic component of pressure in 10 km depth may be larger than the stress exerted by burial. In 15 km depth pressure is roughly equal to the vertical normal stress.

**Problem 7.1.** a) We can insert into eq. 7.1 the values:  $z = 10\,000$  m;  $P = 1$  atm  $\approx 10^5$  Pa und  $g \approx 10$  m s<sup>-2</sup>. This gives a mean density of  $\rho = 1$  kg m<sup>-3</sup>.  
b) 1 atm  $\approx 1$  bar  $= 10^{-3}$  kbar.

**Problem 7.2.** One mole of almandine has the volume:  $V_{almandine} = 11.5$  J / 10<sup>5</sup> Pa  $= 1.15 \cdot 10^{-4}$  m<sup>3</sup>  $= 115$  cm<sup>3</sup>.

**Problem 7.3.** The cation diffusivities at the 3 temperatures are:  $D_{300} \approx 1.6 \cdot 10^{-30}$  m<sup>2</sup> s<sup>-1</sup>;  $D_{500} \approx 6.9 \cdot 10^{-25}$  m<sup>2</sup> s<sup>-1</sup>;  $D_{1000} \approx 1.5 \cdot 10^{-18}$  m<sup>2</sup> s<sup>-1</sup>. All these are many orders of magnitude smaller than the diffusivity of thermal energy. Thus, chemical zoning profiles in minerals will only develop because of the sluggishness of cation diffusion, but not because temperature gradients in a crystal.

**Problem 7.4.** Similar to problem 7.3 we can use the Arrhenius relationship to obtain a diffusivity of  $D_{600} \approx 4.9 \cdot 10^{-23}$  m<sup>2</sup> s<sup>-1</sup>. Inserting this diffusivity and  $l = 1$  mm (being careful to use SI units!) we obtain a time scale of more than 600 million years. This is much longer than the duration of most metamorphic events. We can conclude that 1 mm sized garnets are unlikely to be equilibrated at 600 °C by diffusion. Chemically equilibrated garnets of this size must have grown at this temperature.

**Problem 7.5.** The mean diffusivity between 700 °C and 400 °C is:  $\bar{D} \approx 1.1 \cdot 10^{-22}$  m<sup>2</sup> s<sup>-1</sup>.

**Problem 7.6.** a) The thermal maximum. b) Because the activation volume in eq. 7.5 is much smaller than the activation energy  $Q$ . c) The uppermost curve in Fig. 7.7a. d) Yes, for example if crustal shortening follows intrusion and contact metamorphism.

**Problem 7.7.** Inserting the given values into eq. 7.10 gives a cooling rate:  $s \approx 42^\circ\text{C my}^{-1}$ .

**Problem 7.8.** a) The characteristic time scale of cooling (the thermal time constant) of the dike is:  $4 \cdot 10^{12}$  s. The diffusivities at  $T_i$  and  $T_b$  are  $D_{1200} \approx 1.3 \cdot 10^{-14} \text{ m}^2 \text{s}^{-1}$  and  $D_{500} \approx 7.4 \cdot 10^{-20} \text{ m}^2 \text{s}^{-1}$ . The diffusive time constants of mica are  $t_{eq}^{1200} \approx 1.9 \cdot 10^9$  s and  $t_{eq}^{500} \approx 3.35 \cdot 10^{14}$  s. Thus, the chemical equilibration of the micas is much more rapid than the cooling of the dike at  $1200^\circ\text{C}$  but much slower than the cooling at  $500^\circ\text{C}$ . Thus, it may be - in principle - meaningful to use the micas to determine the intrusion age geochronologically. However, 50 m from the intrusion contact the maximum contact metamorphic temperatures are much lower than  $1200^\circ\text{C}$ . Thus it is necessary to obtain better information on the contact metamorphic temperatures to investigate if the dating of these micas is meaningful. b) Eq. 3.90 shows that the maximum contact metamorphic temperature of the micas is reached  $2.83 \cdot 10^{11}$  s after intrusion. Inserting this into eq. 3.88 we get the contact metamorphic peak temperature the micas experienced to be:  $T_{max} = 829^\circ\text{C}$ . Using this number as the starting temperature  $T_A$  and  $500^\circ\text{C}$  as the final temperature  $T_E$ , eq. 7.9 may be used to estimate the mean diffusivity to be:  $\bar{D} \approx 1.91 \cdot 10^{-17} \text{ m}^2 \text{s}^{-1}$  and  $\bar{t}_{eq} = 1.3 \cdot 10^{12}$  s. This value is only a bit shorter than the characteristic time scale of cooling of the dike. Thus, equilibration of the micas is likely to be at least affected by the contact metamorphism. Whether the micas are partially or fully equilibrated can not be answered and requires careful analysis.

**Problem 7.11.** The  $T$ - $t$ -curve passes following points:  $700^\circ\text{C}$ ;  $0 \text{ my} \rightarrow 700^\circ\text{C}$ ;  $3 \text{ my} \rightarrow 400^\circ\text{C}$ ;  $6 \text{ my} \rightarrow 0^\circ\text{C}$ ;  $26 \text{ my}$ . However, the last point will not be reached as the cooling curve intersects the stable geotherm before that. The corner points of the  $z$ - $t$ -curve are:  $20 \text{ km}$ ;  $0 \text{ my} \rightarrow 14 \text{ km}$ ;  $5 \text{ my} \rightarrow 0 \text{ km}$ ;  $40 \text{ my}$ . The intersection of the cooling curve with the stable geotherm may be found by finding the intersection of the linear equations  $T = z \times 20^\circ\text{C} / \text{km}$  and  $T = 280^\circ\text{C} - z \times 50^\circ\text{C} / \text{km}$ . It is at:  $T = 187^\circ\text{C}$  and  $z = 9333 \text{ m}$ . According to these results, the  $P$ - $T$ -path follows the following evolution:  $5.4 \text{ kbar}$ ;  $700^\circ\text{C} \rightarrow 4.43 \text{ kbar}$ ;  $700^\circ\text{C} \rightarrow 3.78 \text{ kbar}$ ;  $500^\circ\text{C} \rightarrow 3.67 \text{ kbar}$ ;  $400^\circ\text{C} \rightarrow 2.52 \text{ kbar}$ ;  $187^\circ\text{C} \rightarrow 0 \text{ kbar}$ ;  $0^\circ\text{C}$ .

---

**Problem B.1.** The solution of this problem is simply the sum of the approximations from eq. B.13 and eq. B.16:

$$T_i^+ = T_i^- \left( 1 - \frac{2\kappa\Delta t}{\Delta z^2} - u\Delta t \right) + T_{i+1} \left( \frac{2\kappa\Delta t}{\Delta z^2} + u\Delta t \right) + T_{i-1} \left( \frac{2\kappa\Delta t}{\Delta z^2} \right)$$

**Problem B.2.** The answer is:

$$w_i = \frac{-D}{6D + \Delta x^4(\rho_m - \rho_c)g} (w_{i+1} - 4w_{i+1} - 4w_{i-1} + w_{i-2})$$

This equation was derived by forming the difference between the curvature at neighboring grid points in an analogous way to the way eq. B.16 was formulated.

**Problem B.3.** Using  $u = \Delta x / \Delta t$  forward and backward finite differencing schemes give the following approximations:  $T_i^+ = T_{i+1}^-$  bzw.  $T_i^+ = 2T_i^- - T_{i-1}^-$ . If you draw these equations according to the scheme of Fig. B.5, it becomes visible that backward differencing schemes become unstable.

**Problem B.4.** Both data sets have a mean of  $\mu = 8$ . The first data set has a variance of  $Var = 4$ , but an unbiased estimate of the population variance of  $Var' = 4.57$ . For the second data set the variance is  $Var = 42$ , but the unbiased estimate of the population variance is  $Var' = 48$ .

## F. Internet Addresses

On the next few pages there is a random selection of internet addresses that are in some way related to geodynamics. The list was not created according to any important selective criteria and obviously claims no completeness. It is merely a number of addresses that were encountered during the writing of this book.

### Software and Online Tools:

ANSYS. Commercial FEM package

- <http://www.ansys.com/>

DimensionMG. General purpose 3D mesh generator

- <http://www.ahpcrc.org/johnson/SOFTWARE/MESHGEN/>

ESRI. widely used GIS software and data depot

- <http://www.esri.com/>

GeoVu. Software to simplify access to diverse data

- <http://www.ngdc.noaa.gov/seg/geovu/>

GMT. Software to produce maps from digital data

- <http://gmt.soest.hawaii.edu/>

MARC. Commercial FEM package

- <http://www.marc.com/>

MATLAB. Software to do mathematics

- <http://www.mathworks.com/products/matlab/>

MICRODEM. Software for imaging of digital elevation models

- <http://www.usna.edu/Users/ocean0/pguth/website/microdemdown.htm>

Online topographic map creation facility

- <http://www.aquarius.geomar.de/omc/>

Online palaeogeographic map creation (Fig. 2.3)

- <http://www.odsн.de/odsн/services/paleomap/paleomap.html>

Online geophysical map creation facility

- <http://jules.unavco.org>

SCOTese. Plate and climate reconstruction

- <http://www.scotese.com/>

Spectral analysis of data toolkit

- <http://www.atmos.ucla.edu/tcd/ssa/>

## Geological Data

### *Digital Elevation Models*

Shuttle Radar Topography Mission (SRTM)

- <ftp://e0srp01u.ecs.nasa.gov/srtm/version2/>

SRTM30 - 30 sec. resolution - global

- <ftp://e0srp01u.ecs.nasa.gov/srtm/version2/SRTM30>

SRTM30\_Plus

- <http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/arctic.html>

SRTM3 - 3 sec. resolution - global

- <ftp://e0srp01u.ecs.nasa.gov/srtm/version2/SRTM3>

SRTM1 - 1 sec. resolution - USA

- <ftp://e0srp01u.ecs.nasa.gov/srtm/version2/SRTM3>

GTOPO30 (better use SRTM30)

- <http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html>

GLOBE (30 sec similar to Gtopo30)

- <http://www.ngdc.noaa.gov/mgg/topo/gltiles.html>

ETOPO2 - 2 min. resolution including bathymetry

- <http://www.ngdc.noaa.gov/mgg/global/relief/ETOPO2/ETOPO2v2-2006/>

GEBCO-Sand (1minute blend of the Gebco bathymetry with Smith and Sandwell)

- <ftp://falcon.grdl.noaa.gov/pub/walter/>

GEBCO-BATHYMETRY (signed 16 bit integer )

- <http://snowy.arsc.alaska.edu/nasa/topography/>

Multibeam Bathymetric Data

- <http://map.ngdc.noaa.gov/website/mgg/multibeam/viewer.htm>

Coastal Relief

- <http://www.ngdc.noaa.gov/mgg/coastal/coastal.html>

Bathymetry of the Baltic Sea

- <http://www.pangaea.de/Software/PanMap/>

International Bathymetric Chart of the Arctic Ocean

- <http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/arctic.html>

Bathymetry, topography

- <http://www.ngdc.noaa.gov/mgg/bathymetry/relief.html>

Bathymetric Chart of the Arctic Ocean

- <http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/arctic.html>

Bathymetric Chart of the Oceans

- <http://www.ngdc.noaa.gov/mgg/gebco/gebco.html>

Seafloor topography from satellite

- [http://www.ngdc.noaa.gov/mgg/announcements/announce\\_predict.html](http://www.ngdc.noaa.gov/mgg/announcements/announce_predict.html)

Digital elevation models for the US

- <http://mcmcweb.er.usgs.gov>

Digital elevation model of the US

- [http://edcwww.cr.usgs.gov/Webglis/glisbin/guide.pl/glis/hyper/guide/7\\_min\\_dem](http://edcwww.cr.usgs.gov/Webglis/glisbin/guide.pl/glis/hyper/guide/7_min_dem)

### ***Satellite Images***

Global Land Cover Facility (free Landsat...)

- <http://glcf.umiacs.umd.edu/index.shtml>

Landsat Mosaic (MrSid - format)

- <http://www-wsm.physik.uni-karlsruhe.de/>

Free Orthoimages

- [http://www.worldwindcentral.com/wiki/Sources\\_of\\_free\\_orthoimagery](http://www.worldwindcentral.com/wiki/Sources_of_free_orthoimagery)

USGS Global Visualization Viewer

- <http://glovis.usgs.gov/>

Satellite images

- <http://earth.jsc.nasa.gov/>
- <http://visibleearth.nasa.gov/>
- <http://nix.nasa.gov/>
- <http://www.earthkam.ucsd.edu/>
- [http://www.spaceimaging.com/index\\_text.html](http://www.spaceimaging.com/index_text.html)

### ***Water Bodies, Rivers, Shorelines***

SRTM - Water Body Data

- <ftp://e0srp01u.ecs.nasa.gov/srtm/version2/SWBD/>

The Global Discharge Data Base

- <http://www.rivdis.sr.unh.edu/>

Global River Discharge Data Base

- <http://www.sage.wisc.edu/riverdata/>

GSHHS - A Global Self-consistent, Hierarchical, High-resolution Shoreline Database

- <http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html>

World Vector Shoreline (WVS)

- <http://www.pangaea.de/Software/PanMap/>

Shoreline Database

- <http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html>

Shoreline of the United States

- <http://www.csc.noaa.gov/products/shorelines/digdata.htm>

### ***Country Boundaries and City Coordinates***

City coordinates global

- <http://www.heavens-above.com>

Country Boundaries

- <http://www.cipotato.org/DIVA/data/MoreData.htm>

### ***Seismicity and Fault Plane Solutions***

Iris Earthquake Event Catalogue

- <http://www.iris.edu/quakes/eventsrch.htm>

Iris Moment Tensor Data

- <ftp://ftp.iris.washington.edu/pub/CMT/>

**Earthquake Catalog of Switzerland**

- [http://histserver.ethz.ch/simplequery\\_e.html](http://histserver.ethz.ch/simplequery_e.html)

**Earthquakes of Central, Northern and Northwestern Europe**

- <http://neic.usgs.gov/neis/epic/epic.html>

**Earthquake catalogue**

- <http://quake.geo.berkeley.edu/cnss/>

**Earthquake Database**

- [http://www.ngdc.noaa.gov/seg/hazard/sig\\_srch.shtml](http://www.ngdc.noaa.gov/seg/hazard/sig_srch.shtml)

**Earthquake Hazards Program**

- <http://earthquake.usgs.gov/research/>

**Advanced National Seismic System**

- <http://www.ncedc.org/anss/catalog-search.html>

**Southern California Earthquake Data Center**

- [http://www.data.scec.org/catalog\\_search/date\\_mag\\_loc.php](http://www.data.scec.org/catalog_search/date_mag_loc.php)

**Seismologic data bases**

- <http://www.seismology.harvard.edu/CMTsearch.html>

**World stress map project**

- <http://www-wsm.physik.uni-karlsruhe.de/>

***Precipitation and Heat Flow***

**Tropical Rainfall Measuring Mission (TRMM)**

- [http://daac.gsfc.nasa.gov/data/datapool/TRMM\\_DP/](http://daac.gsfc.nasa.gov/data/datapool/TRMM_DP/)

**Geothermal heat flow data**

- <http://www.noaa.gov/geothermal.html>

***Gravity and Magnetism***

**Land and Marine Gravity Data**

- <http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/arctic.html>

**GRIM5-S1/C1: A new global Earth's gravity field model**

- [http://www.gfz-potsdam.de/pb1/pg3/grim/grim5\\_e.html](http://www.gfz-potsdam.de/pb1/pg3/grim/grim5_e.html)

**Global Gravity and Magnetic Data**

- <http://www.getech.com/data/>

**Geomagnetic Data On-line**

- <http://www.ngdc.noaa.gov/seg/geomag/data.shtml>

**Marine Gravity From Sattelite Altimetry**

- [http://topex.ucsd.edu/WWW\\_html/mar\\_grav.html](http://topex.ucsd.edu/WWW_html/mar_grav.html)

**Gravity land**

- <http://www.noaa.gov/landgravity.html>

**Gravity data marine (Sandwell)**

- [http://topex.ucsd.edu/marine\\_grav/mar\\_grav.html](http://topex.ucsd.edu/marine_grav/mar_grav.html)

**Gravity data global**

- [ftp://topex.ucsd.edu/pub/global\\_grav\\_2min/](http://topex.ucsd.edu/pub/global_grav_2min/)

**Gravimetric data base**

- [http://bgi.cnrs.fr:8110/bgi\\_service\\_a.html](http://bgi.cnrs.fr:8110/bgi_service_a.html)

**Geomagnetics**

- <http://www.noaa.gov/geomagnetics.html>

**GPS**

## high-precision GPS data

- <http://sopac.ucsd.edu/>

## GPS - time series

- <http://sideshow.jpl.nasa.gov/mbh/series.html>

## GEOnet GPS- time series

- <http://www.geonet.org.nz/gps/TimeSeries.html>

## UNAVCO

- [http://sps.unavco.org/crustal\\_motion/dxdt/](http://sps.unavco.org/crustal_motion/dxdt/)

## Jules Verne Voyager: Earth

- <http://jules.unavco.org/Voyager/Earth>

## IGS. International GPS service

- <http://tonga.unavco.ucar.edu/>

## International GPS support association

- <http://igscb.jpl.nasa.gov/>

**Other Data**

## Earth sciences websites

- <http://www.websites.noaa.gov/guide/sciences/earth/earth.html>

## Geographic data download of USGS.

- <http://edcwww.cr.usgs.gov/doc/edchome/ndcddb/ndcddb.html>

## Geological time table

- <http://www.dinosauria.com/dml/history.htm>

## Geological time table

- <http://www.ucmp.berkeley.edu/help/timeform.html>

## Geophysical data base of the USGS

- <http://crustal.cr.usgs.gov/crustal/geophysics/index.html>

## GIS datadepot

- <http://www.gisdatadepot.com/>

## IERS. International earth rotation service

- <http://maia.usno.navy.mil/>

## Landform atlas of the US

- <http://fermi.jhuapl.edu/states/states.html>

## Marine geology and geophysics data bases

- <http://www.ngdc.noaa.gov/mgg/mggonline.html>

## National geophysical data center of the US

- <http://www.ngdc.noaa.gov/ngdc.html>

## Natural hazards

- <http://www.noaa.gov/hazards.html>

## Sediment thicknesses around the world

- <http://www.ngdc.noaa.gov/mgg/sedthick/sedthick.html>

Statistic data of the world

- <http://www.un.org/Depts/unsd>

World fact book of the CIA

- <http://www.odci.gov/cia/publications/factbook/index.html>

### **Important Earth Science Journals:**

Australian Journal of Earth Science (Blackwell)

- <http://www.blacksci.co.uk/products/journals/xajes.htm>

Computers and Geoscience (Elsevier)

- <http://www.elsevier.com/inca/publications/store/3/9/8/>

Earth and Planetary Science Letters (Elsevier)

- <http://www.elsevier.nl:80/inca/publications/store/5/0/3/3/2/8/>

Earth Science reviews (Elsevier)

- <http://www.elsevier.com/inca/publications/store/5/0/3/3/2/9/>

Electronic Geosciences

- <http://link.springer.de/link/service/journals/10069/index.htm>

EOS

- <http://www.agu.org/pubs/eos.html>

Geochemistry, Geophysics, Geosystems (electronic Journal)

- <http://146.201.254.53/>

Geophysical Journal International (Blackwell)

- <http://www.blacksci.co.uk/products/journals/gji.htm>

Geophysical Research Letters

- <http://www.agu.org/grl/>

GSA Bulletin, Geology, GSA today

- <http://www.geosociety.org/pubs/>

Journal of Applied Geophysics (Elsevier)

- <http://www.elsevier.com/inca/publications/store/5/0/3/3/3/3/>

Journal of Geophysical Research

- <http://www.agu.org/journals/jb/>

Journal of Metamorphic Geology (Australian site)

- <http://www.es.mq.edu.au/jmg/jmg.html>

Journal of African Earth Science (Elsevier)

- <http://www.elsevier.com/inca/publications/store/6/9/1/>

Journal of Geodynamics (Elsevier)

- <http://www.elsevier.com/inca/publications/store/8/7/4/>

Journal of Petrology (Elsevier)

- <http://petrology.oupjournals.org/>

Journal of Structural Geology(Elsevier)

- <http://www.elsevier.nl:80/inca/publications/store/5/3/9/>

Journal of the Virtual Explorer (electronic Journal)

- <http://www.virtualexplorer.com.au/>

**Mineralogy and Petrology**

- <http://link.springer.de/link/service/journals/00710/index.htm>

**Physics of the Earth and Planetary Interiors**

- <http://www.elsevier.com/inca/publications/store/5/0/3/3/5/6/>

**Precambrian Research (Elsevier)**

- <http://www.elsevier.com/inca/publications/store/5/0/3/3/5/7/>

**Reviews in Geophysics**

- <http://www.agu.org/rog/>

**Tectonics**

- <http://www.agu.org/journals/tc/>

**Important geological organizations:****American Geophysical Union (AGU)**

- <http://www.agu.org/>

**Austrian Geological Survey**

- <http://www.geolba.ac.at/>

**Austrian portal to academic Institutions Home pages**

- <http://www.portal.ac.at/>

**Australian Earth Science Departments**

- [http://www.earth.monash.edu.au/aus\\_depts.html](http://www.earth.monash.edu.au/aus_depts.html)

**Australian Geological Survey**

- <http://www.agso.gov.au/>

**British Geological Society**

- <http://www.geolsoc.org.uk>

**Geological Society of America**

- <http://www.geosociety.org/>

**German portal to academic Institutions Home pages**

- <http://www.bildungsserver.de/>

**International Earth Science Departments and Organizations**

- [http://dir.yahoo.com/Science/Earth\\_Sciences/Geology\\_and\\_Geophysics/Institutes/](http://dir.yahoo.com/Science/Earth_Sciences/Geology_and_Geophysics/Institutes/)

**NASA**

- [http://www.nasa.gov/NASA\\_homepage.html/](http://www.nasa.gov/NASA_homepage.html)

**Swiss Portal to academic Institutions Home pages**

- <http://www.switch.ch/>

**European Geoscience Department listing**

- <http://www.uni-mainz.de/FB/Geo/Geologie/GeoInst/Europa.html>

## References

Abbott DH, Hoffman SE (1984) Archaean plate tectonics revisited; 1, heat flow, spreading rate, and the age of subducting oceanic lithosphere and their effects on the origin and evolution of continents. *Tectonics* 3: 429–448

Abramowitz M, Stegun IA (1972) Handbook of mathematical functions with formulas, graphs mathematical tables. Dover Publications, New York, 1045 p

Ahnert F (1970) Functional relationships between denudation, relief and uplift in large mid-latitude drainage basins. *Am J Sci* 268: 243–263

Ahnert F (1976) Brief description of a comprehensive three-dimensional process-response model of landform development. *Z Geomorphol Suppl* 25: 29–49

Ahnert F (1984) Local relief and the height limits of mountain ranges. *Am J Sci* 284: 1035–1055

Allen PA, Allen JR (1990) Basin analysis. Principles and applications. Blackwell Scientific Publication, Oxford, 450 p

Anderson DA, Tannehill JC, Pletcher RH (1984) Computational fluid mechanics and heat transfer. Series in computational methods in mechanics and thermal sciences. McGraw Hill, New York, 599 p

Anderson EM (1951) The dynamics of faulting. Oliver and Boyd, Edinburgh, 206 p

Anderson GM, Crerar DA (1993) Thermodynamics in geochemistry, the equilibrium model. Oxford University Press, Oxford, 588 p

Andrews DJ, Bucknam RG (1987) Finite degradation of shoreline scarps by a non-linear diffusion model. *J Geophys Res* 92: 12857–12867

Angelier J (1984) Tectonic analysis of fault slip data sets. *J Geophys Res* 89: 5835–5848

Angelier J (1994) Fault slip analysis and palaeo-stress reconstruction. In: Hancock P (ed) Continental deformation. Pergamon Press, Oxford, pp 53–101

Angevine CL, Heller PL, Paola C (1990) Quantitative sedimentary basin modelling. AAPG Continuing Education Course Note Series 32, 133 p

Argus DF, Heflin MB (1995) Plate motion and crustal deformation estimated with geodetic data from the Global Positioning System. *Geophys Res Lett* 22: 1973–1976

Artyushkov EV (1973) Stresses in the lithosphere caused by crustal thickness inhomogeneities. *J Geophys Res* 78: 7675–7708

Atkins PW (1994) Physical chemistry, 5th edition. Oxford University Press, Oxford, 1031 p

Austrheim H, Erambert M, Boundy TM (1996) Garnets recording deep crustal earthquakes. *Earth Planet Sci Lett* 139: 223–238

Avigard D (1992) On the exhumation of coesite-bearing rocks in the Dora-Maira massif (Western Alps, Italy). *Geology* 18: 466–469

Avouac JP (1993) Analysis of scarp profiles: evaluation of errors in morphological dating. *J Geophys Res* 98: 6745–6754

Avouac JP, Peltzer G (1993) Active tectonics in southern Xinjiang, China: Analysis of terrace riser and normal fault scarp degradation along the Hotan-Qira Fault System. *J Geophys Res* 98: 21773–21807

Barr TD, Dahlen FA (1989) Brittle frictional mountain building. 2nd thermal structure and heat budget. *J Geophys Res* 94: 3923–3947

Barr TD, Houseman GA (1996) Deformation fields around a fault embedded in a non-linear ductile medium. *Geophys J Int* 125: 473–490

Barrell J (1914) The strength of the Earth's crust. *J Geol* 22: 441–468

Barton CM, England PC (1979) Shear heating at the Olympos (Greece) thrust and the deformation properties of carbonates at geological strain rates. *Geol Soc Am Bull* 90: 483–492

Barton MD, Hanson RB (1989) Magmatism and the development of low pressure metamorphic belts: implications from the western United States and thermal modelling. *Geol Soc Am Bull* 101: 1051–1065

Bassi G (1991) Factors controlling the style of continental rifting: insights from numerical modelling. *Earth Planet Sci Lett* 105: 430–452

Bassi G, Keen CE, Potter P (1993) Contrasting styles of rifting: models and examples from the eastern Canadian margin. *Tectonics* 12: 639–655

Batt GE, Braun J (1997) On the thermomechanical evolution of compressional orogens. *Geophys J Int* 128: 364–382

Beaumont C (1981) Foreland basins. *Geophys J Roy Astron Soc* 65: 291–329

Beaumont C, Fullsack P, Hamilton J (1992) Erosional control of active compressional orogens. In: McClay KR (ed) *Thrust tectonics*. Chapman and Hall, New York, pp 1–18

Beaumont C, Ellis S, Hamilton J, Fullsack P (1996) Mechanical model for subduct ion-collision tectonics of Alpine-type compressional orogens. *Geology* 24: 675–678

Begin SB, Meyer DF, Schumm SA (1981) Development of longitudinal profiles of alluvial channels in response to base level lowering. *Earth Surface Processes and Landforms* 6: 49–68

Bell JS, Gough DI (1979) Northwest-southeast compressive stress in Alberta: evidence from oil wells. *Earth Planet Sci Lett* 45: 475–482

Benfield AE (1949) The effect of uplift and denudation on underground temperatures. *J Appl Phys* 20: 66–70

Berman RG (1988) Internally-consistent thermodynamic data for minerals in the system  $\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{CaO}-\text{MgO}-\text{FeO}-\text{Fe}_2\text{O}_3-\text{Al}_2\text{O}_3-\text{SiO}_2-\text{TiO}_2-\text{H}_2\text{O}-\text{CO}_2$ . *J Pet* 29: 445–522

Bickle MJ, McKenzie D (1987) The transport of heat and matter by fluids during metamorphism. *Contrib Mineral Petrol* 95: 384–392

Bickle MJ, Hawkesworth CJ, England PC, Athey D (1975) A preliminary thermal model for regional metamorphism in the eastern Alps. *Earth Planet Sci Lett* 26: 13–28

Bird P (1979) Continental delamination and the Colorado Plateau. *J Geophys Res* 84: 7561–7571

Bird P, Piper K (1980) Plane stress finite element model of tectonic flow in southern California. *Phys Earth Planet Int* 21: 158–195

Blanckenburg F von, Davies JH (1995) Slab breakoff: a model for syncollisional magmatism and tectonics in the Alps. *Tectonics* 14: 120–131

Bohlen SR (1987) Pressure temperature time paths and a tectonic model for the evolution of granulites. *J Geol* 95: 617–632

Bond GC, Kominz MA, Devlin WJ (1983) Thermal subsidence and eustasy in the lower Paleozoic miogeocline of western North America. *Nature* 306: 775–779

Bons PD, Barr TD, ten Brink CE (1997) The development of delta-clasts in non-linear viscous materials: a numerical approach. *Tectonophysics* 270: 29–42

Bott MHP (1993) Modelling the plate-driving mechanism. *J Geol Soc London* 150: 941–951

Bott MHP, Waghorn GD, Whittaker A (1989) Plate boundary forces at subduction zones and trench-arc compression. *Tectonophys* 170: 1–15

Brace WF, Kohlstedt DL (1980) Limits on lithospheric stress imposed by laboratory experiments. *J Geophys Res* 94: 3967–3990

Braun J (1992) Dynamics of compressional orogens; beyond thin sheet and plane strain approximations. *EOS transactions* 73: 292

Braun J, Beaumont C (1995) Three dimensional numerical experiments of strain partitioning at oblique plate boundaries: Implications for contrasting tectonic styles in the southern coast ranges, California central south island, New Zealand. *J Geophys Res* 100: 18059–18074

Braun J, Sambridge M (1997) Modelling landscape evolution on geological time scales: a new method based on irregular spatial discretization. *Basin Res* 9: 27–52

Braun J (2002) Quantifying the effect of recent relief changes on age- elevation relationships. *Earth Planet Sci Lett* 200: 331–343

Braun J (2006) Quantitative constraints on the rate of landform evolution derived from low-temperature thermochronology. *Rev Min Geochem* 58: 351–374

Brown RW (1991) Backstacking apatite fission track “stratigraphy”: a method for resolving the erosional and isostatic rebound components of tectonic uplift histories. *Geology* 19: 74–77

Brown M (2006) Metamorphic conditions in orogenic belts: A record of secular change. *Int Geol Rev* in press.

Brun JP, Cobbold PR (1980) Strain heating and thermal softening in continental shear zones: a review. *J Struct Geol* 2: 149–158

Buck WR, Martinez F, Steckler MS, Cochran JR (1988) Thermal consequences of lithosphere extension: pure and simple. *Tectonics* 7: 213–234

Buck WR (1991) Modes of continental lithosphere extension. *J Geophys Res* 96: 20161–20178

Burbank DW, Anderson RS (2001) Tectonic geomorphology. Blackwell Science, Malden, Mass 274 p

Burg JP, Gerya TV (2005) The role of viscous heating in Barrovian metamorphism of collisional orogens: Thermomechanical models and application to the Leponetine Dome in the Central Alps. *J Met Geol* 23: 75–95

Byerlee JD (1968) Brittle ductile transition in rocks. *J Geophys Res* 73: 4741–4750

Byerlee JD (1970) Friction of rocks. In: Everden JF (ed) Experimental studies of rock friction with application to earthquake prediction. USGS, Menlo Park, pp 55–77

Camacho A, Vernon RH, Fitzgerald JD (1995) Large volumes of anhydrous pseudotachylite in the Woodroffe thrust, eastern Musgrave Ranges, Australia. *J Struct Geol* 17: 371–383

Camacho A, McDougall I, Armstrong R, Braun J (2001) Evidence for Shear Heating, Musgrave Block, Central Australia. *J Struct Geol* 23: 1007–1013

Carey SW (1976) The expanding Earth. Elsevier, Amsterdam, 488 p

Carslaw HS, Jaeger JC (1959) Conduction of heat in solids. Oxford Science Publications, Oxford University Press, Oxford, 510 p

Carson CJ, Powell R, Wilson CJL, Dirks PHMG (1997) Partial melting during tectonic exhumation of a granulite terrain: an example from the Larsemann Hills, East Antarctica. *J Met Geol* 15: 105–127

Carson MA, Kirkby MJ (1972) Hillslope form and processes, Cambridge University Press, Cambridge, 475 p

Cermak V, Rybach L (1989) Vertical distribution of heat production in the continental crust. *Tectonophys* 159: 217–230

Chamberlain CP, Sonder LJ (1990) Heat-producing elements and the thermal and baric patterns of metamorphic belts. *Science* 250: 763–769

Chapple WM (1978) Mechanics of thin-skinned fold-and-thrust belts. *Geol Soc Am Bull* 89: 1189–1198

Chase CG (1992) Fluvial landsculpting and the fractal dimension of topography. *Geomorphology* 5: 39–57

Chemenda A, Mattauer M, Bokun AN (1996) Continental subduction and a mechanism for exhumation of high-pressure metamorphic rocks: new modelling and field data from Oman. *Earth Planet Sci Lett* 143: 173–182

Christensen UR, Yuen DA (1984) The interaction of a subducting lithospheric slab with a chemical or phase boundary. *J Geophys Res* 89: 4389–4402

Christie JM, Ord A (1980) Flow stress from microstructures of mylonites: example and current assessment. *J Geophys Res* 85: 6253–6262

Cliff RA, Droop GTR, Rex DC (1985) Alpine metamorphism in the south-east Tauern Window, Austria: 2nd Rates of heating, cooling and uplift. *J Met Geol* 3: 403–415

Cloetingh S, van Wees JD, van der Beek PA, Spadini G (1995) Role of pre-rift rheology in kinematics of extensional basin formation: constraints from thermo-mechanical models of Mediterranean and intracratonic basins. *Mar Petrol Geol* 12: 793–807

Cloos M (1982) Flow melanges: numerical modelling and geologic constraints on their origin in the Franciscan subduction complex, California. *Bull Geol Soc Am* 93: 330–345

Cloos M (1984) Flow melanges and the structural evolution of accretionary wedges. *Geol Soc Am Spec Pap* 198: 71–79

Cloos M, Shreve RL (1988) Subduction channel model of accretion, melange formation, sediment subduction and subduction erosion at convergent plate margins 1. Background and description. *Pure and Applied Geophys* 128: 455–500

Coblentz D, Sandiford M (1994) Tectonic stress in the African plate: constraints on the ambient lithospheric stress state. *Geology* 22: 831–834

Coblentz D, Richardson RM, Sandiford M (1994) On the potential energy of the Earth's lithosphere. *Tectonics* 13: 929–945

Cochran JR (1982) The magnetic quite zone in the eastern gulf of Aden: implications for the early development of the continental margin. *Geophys J Roy Astron Soc* 68: 171–202

Cochran JR (1983) Effects of finite extension times on the development of sedimentary basins. *Earth Planet Sci Lett* 66: 289–302

Coffin MF, Eldholm O (1993a) Scratching the surface: Estimating dimensions of large igneous provinces. *Geology* 21: 515–518

Coffin MF, Eldholm O (1993b) Large igneous provinces. *Sci Am* 269: 42–49

Coffin MF, Eldholm O (1994) Large igneous provinces: crustal structure, dimensions, and external consequences. *Rev Geophys* 32: 1–36

Coffin MF (1997) Models for the emplacement of large igneous provinces; a review. In: International lithosphere program on volcanic margins, abstract volume, Geoforschungszentrum Potsdam.

Connolly JAD, Thompson AB (1989) Fluid and enthalpy production during regional metamorphism. *Contrib Mineral Petrol* 102: 347–366

Conrad CP, Molnar P (1999) Convective instability of a boundary layer with temperature- and strain-rate-dependent viscosity in terms of “available buoyancy”. *Geophys J Int* 139: 51–68

Cooke RA, OBrian PJ, Carswell DA (2000) Garnet zoning and the identification of equilibrium mineral compositions in high pressure-temperature granulites from the Moldanubian zone, Austria. *J Met Geol* 18: 551–569

Coulomb CA (1773) Sur une application des règles de maximis et minimis a quelques problèmes de statique relatifs a l'architecture. *Acad Roy Sci Mem de Math et de Phys* 7: 343–382

Cowan DS, Silling RM (1978) A dynamic scaled model of accretion and trenches and its implications for the tectonic evolution of subduction complexes. *J Geophys Res* 83: 5389–5396

Cox A (1972) Plate tectonics and geomagnetic reversals. Freeman and Company, San Francisco, 702 p

Crank J (1975) The mathematics of diffusion, 2nd edn. Oxford Science Publications. Clarendon Press, Oxford, 414 p

Craw D, Koons PO, Zeitler PK, Kidd WSF (2005) Fluid evolution and thermal structure in the rapidly exhuming gneiss complex of Namche Barwa-Gyala Peri, eastern Himalayan syntaxis. *J Met Geol* 23: 829–845

Creager KC, Jordan TH (1984) Slab penetration into the lower mantle. *J Geophys Res* 89: 3031–3050

Crough ST (1983) Hot spot swells. *Ann Rev Earth Planet Sci* 11: 165–193

Cull JP (1976) The measurement of thermal parameters at high pressures. *P Appl Geoph* 114: 301–307

Culling WEH (1960) Analytical theory of erosion. *J Geol* 68: 336–344

Cygan RT, Lasaga AC (1985) Self-diffusion of magnesium in garnet at 750°C to 900°C. *Am J Sci* 285: 328–350

Dahlen FA (1984) Non-cohesive critical Coulomb wedges: an exact solution. *J Geophys Res* 89: 10125–10133

Dahlen FA, Suppe J, Davis D (1984) Mechanics of fold-and-thrust belts and accretionary wedges: cohesive coulomb theory. *J Geophys Res* 89: 10087–10101

Dahlen FA, Barr TD (1989) Brittle frictional mountain building 1st Deformation and mechanical energy budget. *J Geophys Res* 94: 3906–3922

Dale J, Powell R, White RW, Elmer FL, Holland TJB (2005) A thermodynamic model for Ca-Na clinoamphiboles in Na<sub>2</sub>O-CaO-FeO-MgO-Al<sub>2</sub>O<sub>3</sub>-SiO<sub>2</sub>-H<sub>2</sub>O-O for petrological calculations. *J Met Geol* 23: 771–791

Dalmayrac B, Molnar P (1981) Parallel thrust and normal faulting in Peru and constraints on the state of stress. *Earth Planet Sci Lett* 55: 473–481

Davis D, Suppe J, Dahlen FA (1983) Mechanics of fold-and-thrust belts and accretionary wedges. *J Geophys Res* 88: 1153–1172

DeMets C, Gordon RG, Argus DF, Stein S (1990) Current plate motions. *Geophys J Int* 101: 425–478

Den Tex E (1963) A commentary on the correlation of metamorphism and deformation in space and time. *Geol Mijnbouw* 42: 17176

Dewey JF (1988) Extensional collapse of orogens. *Tectonics* 7: 1123–1139

DeYoreo JJ, Lux DR, Guidotti CV (1991) Thermal modelling in low-pressure/high-temperature metamorphic belts. *Tectonophys* 188: 209–238

Dickinson WR (1976) Plate tectonic evolution of sedimentary basins. AAPG Continuing Education Course Note Series 1: 1–62

Divins DL NGDC (2006) Total sediment thickness of the world's oceans and marginal seas, <http://www.ngdc.noaa.gov/mgg/sedthick/sedthick.html>

Dodson MH (1973) Closure temperature in cooling geochronological and petrological systems. *Contrib Mineral Petrol* 40: 259–274

Doin MP, Fleitout L (1996) Thermal evolution of the oceanic lithosphere: an alternative view. *Earth Planet Sci Lett* 142: 121–136

Doglioni C (1993) Some remarks on the origin of foredeeps. *Tectonophys* 228: 1–20

Doglioni C, Harabaglia P, Merlini S, Mongelli F, Peccerillo A, Piromallo C (1999) Orogenes and slabs vs. their direction of subduction. *Earth Sci Rev* 45: 167–208

Dumoulin C, Doin MP, Fleitout L (2001) On the interpretation of linear relationships between seafloor subsidence rate and the height of the ridge. *Geophys J Int* 146: 691–698

Dunning GH, Etheridge MA, Hobbs BE (1982) On the stress dependence of sub-grain size. *Textures and Microstructures* 5: 127–152

Edmond JM, Damm K von (1983) *Heiße Quellen am Grund der Ozeane*. Spektrum der Wissenschaft 6: 74–87

Ehlers K, Powell R (1994) An empirical modification of Dodson's equation for closure temperature in binary systems. *Geochim Cosmochim Acta* 58: 241–248

Ehlers K, Powell R, Stüwe K (1994a) The determination of cooling histories from garnet-biotite equilibrium. *Am Min* 79: 737–744

Ehlers K, Stüwe K, Powell R, Sandiford M, Frank W (1994b) Thermometrically inferred cooling rates from the Plattengneiss, Koralpe region, eastern Alps. *Earth Planet Sci Lett* 125: 307–321

Elison MW (1991) Intracontinental contraction in western North America: continuity and episodicity. *Geol Soc Am Bull* 103: 1226–1238

Ellis MA, Densmore AL, Anderson RS (1999) Development of mountainous topography in the Basin and Ranges, USA. *Basin Research* 11: 21–41

Engelder T (1993) Stress regimes in the lithosphere. Princeton University Press, Princeton, 451 p

Engelder T (1994) Deviatoric stressitis: a virus infecting the Earth science community. *EOS* 75: 18, 210–213

England PC, Richardson SW (1977) The influence of erosion upon the mineral facies of rocks from different metamorphic environments. *J Geol Soc London* 134: 201–213

England PC, Holland TJB (1979) Archimedes and the Tauern eclogites: the role of buoyancy in the preservation of exotic tectonic blocks. *Earth Planet Sci Lett* 44: 287–294

England PC (1981) Metamorphic pressure estimates and sediment volumes for the Alpine orogeny: an independent control on geobarometers? *Earth Planet Sci Lett* 56: 387–397

England PC, McKenzie D (1982) A thin viscous sheet model for continental deformation. *Geophys J Roy Astron Soc* 70: 295–321

England PC, Thompson A (1984) Pressure-temperature-time paths of regional metamorphism I. Heat transfer during the evolution of regions of thickened continental crust. *J Pet* 25: 894–928

England PC, Houseman GA, Sonder LJ (1985) Length scales for continental deformation in convergent, divergent and strike-slip environments: analytical and approximate solutions for a thin viscous sheet model. *J Geophys Res* 90: 3551–3557

England PC, Houseman G (1986) Finite strain calculations of continental deformation 2nd Comparison with the India-Asia collision zone. *J Geophys Res* 91: 3664–3676

England PC (1987) Diffusive continental deformation: length scales, rates and metamorphic evolution. *Phil Trans R Soc Lon* 321: 3–22

England PC, Houseman GA (1988) The mechanics of the Tibetan plateau. *Phil Trans Roy Soc Lon A326: 301–319*

England PC, Houseman G (1989) Extension during continental convergence, with application to the Tibetan plateau. *J Geophys Res* 94: 17561–17579

England PC, Jackson J (1989) Active deformation of the continents. *Ann Rev Earth Planet Sci* 17: 197–226

England PC, Molnar P (1990) Surface uplift, uplift of rocks and exhumation of rocks. *Geology* 18: 1173–1177

England PC, Molnar P (1991) Inferences of deviatoric stress in actively deforming belts from simple physical models. *Phil Trans R Soc Lond* 337: 151–164

England PC, Molnar P (1993) The interpretation of inverted metamorphic isogrades using simple physical calculations. *Tectonics* 12: 145–157

England PC (1996) The mountains will flow. *Nature* 381: 23–24

Ernst WG (1971) Do mineral paragenesis reflect unusually high-pressure conditions of Franciscan metamorphism? *Am J Sci* 270: 81–108

Eskola P (1915) On the relation between the chemical and mineralogical composition in the metamorphic rocks of the Orijarvi region. *Bull Comm Geol Finn* 44: 109–145

Evenden G (1990) Cartographic projection procedures for the UNIX environment – a user's manual. USGS Open File Report 90–284

Fischer GW (1973) Non-equilibrium thermodynamics as a model for diffusion controlled metamorphic processes. *Am J Sci* 273: 897–924

Fleitout L, Froidvaux C (1982) Tectonics and topography for a lithosphere containing density heterogeneities. *Tectonics* 1: 21–57

Fletcher CAJ (1991) Computational techniques for fluid dynamics 1st Fundamental and general techniques, 2nd edn. Springer series in computational physics. Springer, Berlin Heidelberg New York, 401 p

Forsyth DW (1985) Subsurface loading and estimates of the flexural rigidity of continental lithosphere. *J Geophys Res* 90: 12623–12632

Forsyth DW, Uyenda S (1975) On the relative importance of the driving forces of plate motion. *Geophys J Roy Astron Soc* 43: 163–200

Fortier SM, Giletti BJ (1991) Volume self diffusion of oxygen in biotite, muscovite and phlogopite micas. *Geochim Cosmochim Acta* 55: 1319–1330

Fourier JBJ (1816) Theorie de la chaleur. *Annales de Chimie et de Physique* 3: 350–376

Fourier JBJ (1820) Extrait d'un memoire sur le refroidissement seculaire du globe terrestre. *Annales Chimie et de Physique* 13: 418–437

Fowler CMR, Nisbet EG (1982) The thermal background to metamorphism II. Simple two-dimensional conductive models. *Geoscience Canada* 9: 208–214

Frohlich C, Coffin MF, Massell C, Mann P, Schuur CL, Davis SD, Jones T, Karner G (1997) Constraints on Macquarie Ridge tectonics provided by Harvard focal mechanisms and teleseismic earthquake locations. *J Geophys Res* 102: 5029–5041

Frost HJ, Ashby MF (1982) Deformation-mechanism maps. Pergamon Press, Cambridge, 260 p

Frottier LG, Buttles J, Olson P (1995) Laboratory experiments on the structure of subducted lithosphere. *Earth Planet Sci Lett* 133: 19–34

Garcia-Castellanos D, Fernandez M, Torne M (1997) Numerical modeling of foreland basin formation: a program relating subsidence, thrusting, sediment geometry and depth dependent lithosphere rheology. *Comp Geosci* 23: 993–1003

Gerya TV, Yuen DA (2003) Characteristics-based marker in cell method with conservative finite difference schemes for modelling geological flows with strongly variable transport properties. *Phys Earth Planet Int* 140: 293–318

Gilchrist AR, Kooi H, Beaumont C (1994) Post-Gondwana geomorphic evolution of southwestern Africa: implications for the controls on landscape development from observations and numerical experiments. *J Geophys Res* 99: 12211–12228

Gleason GC, Tullis J (1995) A flow law for dislocation creep of quartz aggregates determined with the molten salt cell. *Tectonophys* 247: 1–23

Goetze C (1978) The mechanisms of creep in olivine. *Phil Trans Roy Soc London* 288: 99–119

Goetze C, Evans B (1979) Stress and temperature in the bending lithosphere as constrained by experimental rock mechanics. *Geophys J Roy Astron Soc* 59: 463–478

Gordon RB (1965) Diffusion creep in the Earth's mantle. *J Geophys Res* 70: 56–61

Graham CM, England PC (1976) Thermal regimes and regional metamorphism in the vicinity of overthrust faults: an example of shear heating and inverted metamorphic zonation from southern California. *Earth Planet Sci Lett* 31: 142–152

Grasemann B, Mancktelow NS (1993) Two-dimensional thermal modelling of normal faulting: the Simplon Fault Zone, Central Alps, Switzerland. *Tectonophys* 225: 155–165

Greenfield JE, Clarke GL, White RW (1998) A sequence of partial melting reactions at Mt Stafford, central Australia. *J Met Geol* 16: 363–378

Greenwood HJ (1989) On models and modelling. *Canad Mineral* 27: 1–14

Griggs D (1939) A theory of mountain building. *Am J Sci* 237: 611–650

Groves DI, Condie KC, Goldfarb RJ, Hronsky JMA, Vielreicher RM (2005) Secular changes in global tectonic processes and their influence on the temporal distribution of gold-bearing mineral deposits. *Econ Geol* 100: 203–224

Gudmundsson O, Sambridge M (1998) A regionalized upper mantle (RUM) seismic model. *J Geophys Res* 103: 7121–7136

Haack U (1983) On the content and vertical distribution of K, Th and U in the continental crust. *Earth Planet Sci Lett* 62: 360–366

Hack JT (1957) Studies of longitudinal profiles in Virginia and Maryland. USGS Prof Pap 294-B.

Hacker BR, Abers GA, Peacock SM (2003) Subduction factory 1. Theoretical mineralogy, densities, seismic wave speeds, and H<sub>2</sub>O contents. *J Geophys Res* 108: doi: 10.1029/2001JB001127

Hanks TC, Andrews DJ (1989) Effect of far field slope on morphologic dating of scarp-like landforms. *J Geophys Res* 94: 565–573

Harker A (1939) Metamorphism. Lethuer S. Co., London, 462 p

Harley SL (1989) The origin of granulites: a metamorphic perspective. *Geol Mag* 126: 215–247

Harrison CGA (1994) Rates of continental erosion and mountain building. *Geol Rundsch* 83: 431–447

Harrison TM, Clark GK (1979) A model of the thermal effects of igneous intrusion and uplift as applied to Quottoon pluton, British Columbia. *Canad J Earth Sci* 16: 410–420

Harrison TM, Copeland P, Kidd WSF, Yin A (1992) Raising Tibet. *Science* 255: 1663–1670

Hawkesworth C, Turner S, Gallagher K, Hunter A, Bradshaw T, Rogers N (1995) Calc-alkaline magmatism, lithosphere thinning and extension in the basin and Range. *J Geophys Res* 100: 10271–10286

Hawkins SW (1988) A brief history of time. Space Time Publications. Guild Publishing, 213 p

Heezen BC (1962) The deep sea floor. In: Runcorn SK (ed) Continental drift. Academic Publishers, New York, pp 235–268

Heier KS, Brown GC (1978) The distribution and redistribution of heat-producing elements in the continents. *Phil Trans Roy Soc* 288: 393–400

Hergarten S (2002) Self organized criticality in Earth systems. Springer 272 p.

Hess H (1961) History of ocean basins. In: Engle AEJ, Petrologic Studies. A volume in honour of AE Buddington, *Geol Soc Am* pp 599–620

Hillier JK, Watts AB (2005) Relationship between depth and age in the North Pacific Ocean. *J Geophys Res* 110: doi: 10.1029/2004JB003406

Hillis RR, Hand M, Mildren SD, Morton J, Reid P, Reynolds S (2004) Hot dry rock geothermal exploration in Australia. *Petrol Expl Soc Australia Spec Pub* 2: 431–440

Hilst R van der, Engdahl ER, Sparkman W, Nolet G (1991) Tomographic imaging of subducted lithosphere below north-west Pacific island arcs. *Nature* 357: 37–43

Hodges KV (1996) Self-organization and the metamorphic evolution of mountain ranges. *MSG meeting Kingston University abs*

Hoffman PF, Grotzinger JP (1993) Orographic precipitation, erosional unloading and tectonic style. *Geology*, 21: 195–198

Hoffman PF, Kaufman AJ, Halverson GP, Schrag P (1998) A Neoproterozoic snow-ball Earth. *Science* 281: 1342–1346

Hoke L, Hilton DR, Lamb SH, Hammerschmidt K, Friedrichsen H (1994)  $^3\text{He}$  evidence for a wide zone of active mantle melting beneath the Central Andes. *Earth Planet Sci Lett* 128: 341–355

Holland TJB, Powell R (1990) An enlarged and updated internally consistent thermodynamic dataset with uncertainties and correlations: the system  $\text{K}_2\text{O}$ - $\text{Na}_2\text{O}$ - $\text{CaO}$ - $\text{MgO}$ - $\text{FeO}$ - $\text{Fe}_2\text{O}_3$ - $\text{Al}_2\text{O}_3$ - $\text{TiO}_2$ - $\text{SiO}_2$ - $\text{C}$ - $\text{H}_2\text{O}_2$ . *J Met Geol* 8: 89–124

Holmes A (1929) Radioactivity and Earth movements. *Trans Geol Soc Glasgow* 18: 559–607

Horton RE (1945) Erosional development of streams and their drainage basins: hydrophysical approach to quantitative geomorphology. *Bull Geol Soc Am* 56: 275–370

House MA, Wernicke BP, Farley KA (1998) Dating topography of the Sierra Nevada, California, using apatite (U-Th)/He ages. *Nature* 396: 66–69

House MA, Farley KA, Kohn BP (2000) An empirical test of helium diffusion in apatite: borehole data from the Otway basin, Australia. *Earth Planet Sci Lett* 170: 463–474

Houseman G, McKenzie DP, 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: 6115–6132

Houseman G, England P (1986a) Finite strain calculations of continental deformation 1st method and general results for convergent zones. *J Geophys Res* 91: 3651–3663

Houseman G, England P (1986b) A dynamical model of lithosphere extension and sedimentary basin formation. *J Geophys Res* 91: 719–729

Houseman G, England P (1996) A lithospheric-thickening model for the Indo-Asian collision. In: Yin A and Harrison M (eds) *The tectonic evolution of Asia*. Cambridge University Press 486 p

Houseman G, Gubbins D (1997) Deformation of subducted oceanic lithosphere. *Geophys J Int* 131: 535–551

Houseman G, Molnar P (1997) Gravitational (Rayleigh-Taylor) instability of a layer with non-linear viscosity and convective thinning of continental lithosphere. *Geophys J Int* 128: 125–150

Huppert HE, Sparks RSJ (1988) The generation of granitic magmas by intrusion of basalt into continental crust. *J Pet* 29: 599–624

Isacks BL, Barazangi M (1977) Geometry of Benioff zones: lateral segmentation and downward bending of the subducted lithosphere. In: Talwani M, Pitman III WC (eds) *Island Arcs, deep-sea trenches and back arc basins*. Am Geophys U, Maurice Ewing Series Washington 1: 99–114

Isacks BL, Oliver J, Sykes LR (1968) Seismology and the new global tectonics. *J Geophys Res* 73: 5855–5899

Issler D, McQueen H, Beaumont C (1989) Thermal consequences of simple shear extension of the continental lithosphere. *Earth Planet Sci Lett* 91: 341–358

Itayama K, Stüwe HP (1974) Mittlere Reaktionsgeschwindigkeit bei zeitlich veränderter Temperatur. *Zeitschr Metallk* 65: 70–72

Jackson J (2002) Strength of the continental lithosphere: Time to abandon the jelly sandwich? *GSA Today* 12: 4–10

Jackson ED, Shaw HR (1975) Stress fields in the central portions of the pacific plate: delineated in time by linear volcanic chains. *J Geophys Res* 80: 1861–1874

Jackson JA, McKenzie DP (1988) The relationship between plate motions and seismic moment tensors, and the rates of active deformation in the Mediterranean and the Middle East. *Geophys J Roy Astron Soc* 93: 45–73

Jaeger JC (1964) Thermal effects of intrusions. *Rev Geophys* 2: 443–466

Jaeger JC, Cook NGW (1979) Fundamentals of rock mechanics, 3rd edn. Science Paperbacks, Chapman and Hall, London, 585 p

Jarvis GT, McKenzie DP (1980) Sedimentary basin formation with finite extension rates. *Earth Planet Sci Lett* 48: 42–52

Jaupart C, Provost A (1985) Heat focusing, granite genesis and inverted metamorphic gradients in continental collision zones. *Earth Planet Sci Lett* 73: 385–397

Jeanloz R (1988) High-pressure experiments and the Earth's deep interior. *Physics Today* 41: 44–45

Jeanloz R, Richter FM (1979) Convection, composition and the thermal state of the lower mantle. *J Geophys Res* 84: 5497–5504

Jessel MW, Lister GS (1990) A simulation of the temperature dependence of quartz fabrics. In: Knipe RJ, Rutter EH (eds) Deformation mechanisms, rheology and tectonics. *Geol Soc Spec Publ* 54: 353–362

Joesten R (1977) Evolution of mineral assemblage zoning in diffusion metasomatism. *Geochim Cosmochim Acta* 41: 649–670

Jones RR (1997) Lateral extrusion in transpression zones; the importance of boundary conditions. *J Struct Geol* 19: 1201–1217

Jordan TH (1981a) Continents as a chemical boundary layer. *Phil Trans Roy Soc Lond A301*: 359–373

Jordan TE (1981b) Thrust loads and foreland basin evolution. Cretaceous, western United States. *Am Ass Pet Geol* 65: 2506–2520

Jordan TH (1989) Chemical boundary layers of the mantle and core. *Phil Trans Roy Soc Lond A328*: 441–442

Julien PY (1998) Erosion and Sedimentation Cambridge University Press, Cambridge UK, 298 p.

Jull M, Keleman PB (2001) On the conditions for lower crustal convective instability. *J Geophys Res* 106: 6423–6446

Karner GD, Watts AB (1983) Gravity anomalies and flexure of the lithosphere at mountain ranges. *J Geophys Res* 88: 10449–10477

Keen CE (1980) The dynamics of rifting: deformation of the lithosphere by active and passive driving mechanisms, *Geophys. J Roy Astron Soc* 62: 631–647

Keen C, Peddy C, de Voogd B, Mathew D (1989) Conjugate margins of Canada and Europe: results from deep reflection profiling. *Geology* 17: 173–176

Keller EA, Pinter N (1996) Active Tectonics - Earthquakes, Uplift and Landscape. Prentice Hall, Upper Saddle River, 338 p

Kelsey DE, White RW, Powell R (2005) Calculated phase equilibria in  $K_2O$ - $FeO$ - $MgO$ - $Al_2O_3$ - $SiO_2$ - $H_2O$  for silica undersaturated sapphirine bearing mineral assemblages. *J Met Geol* 23: 217–239.

Kelvin Lord (1864) The secular cooling of the Earth. *Trans Roy Soc Edinburgh* 23: 157

Kincaid C, Silver P (1996) The Role of Viscous Dissipation in the Orogenic Process. *Earth Planet Sci Lett* 142: 271–288

King LC (1953) Canons of landscape evolution. *Geol Soc Am Bull* 64: 721–753

King LC (1983) Wandering continents and spreading sea floors on an expanding Earth. John Wiley and Sons, Chichester, 232 p

King SD (1996) You don't need plumes to make flood basalts; recipes that don't require a plume. *EOS* 77: 768

Kirchner JW (1993) Statistical inevitability of Horton's laws and the apparent randomness of stream channel networks. *Geology* 21: 591–594

Kirkby MJ (1994) (ed.) *Process Models and Theoretical Geomorphology*. (British Geomorphological research group symposia) John Wiley and Sons. 432p

Kooi J, Beaumont C (1994) Escarpment evolution on high-elevation rifted margins insights derived from a surface processes model that combines diffusion, advection and reaction. *J Geophys Res* 99: 12191–12209

Koons PO (1990) Two-sided orogen: Collision and erosion from the sandbox to the southern Alps, New Zealand. *Geology* 18: 679–682

Kreemer C, Holt WE, Haines AJ (2003) An integrated global model of present day plate motions and plate boundary deformation. *Geophys J Int* 154: 8–34

Kuhlemann J, Frisch W, Dunkl I, Szekely B (2001) Quantifying tectonic versus erosive denudation by the sediment budget; the Miocene core complexes of the Alps. *Tectonophys* 330: 1–23

Kühn A, Stüwe K, Trouw RAJ (2004) Metamorphic evolution of the Ribeira Belt: Evidence from outcrops in the Rio de Janeiro area, Brazil. *J Pet* 11: 2303–2323

Kuznir NJ, Park RG (1986) Continental lithosphere strength: the critical role of lower crustal deformation. In: Dawson JB, Carswell DA, Hall J, Wedepohl KH (eds) *The nature of the lower continental crust*. Geol Soc Spec Pub 24: 79–93

Lachenbruch AH (1968) Preliminary geothermal model for the Sierra Nevada. *J Geophys Res* 73: 6977–6989

Lachenbruch AH (1970) Crustal temperature and heat production: implications of the linear heat flow relationship. *J Geophys Res* 75: 3291–3300

Lachenbruch AH, Bunker CM (1971) Vertical gradients of heat production in the continental crust. 2nd Some estimates from borehole data. *J Geophys Res* 76: 3852–3860

Lachenbruch AH (1980) Frictional heating, fluid pressure and the resistance to fault motion. *J Geophys Res* 85: 6097–6112

Lambeck K (1991) Glacial rebound and sea level change in the British Isles. *Terra Nova* 3: 379–389

Lambeck K (1993) Glacial rebound and sea level change, an example of a relationship between mantle and surface processes. *Tectonophys* 223: 15–37

Larson RL (1991) Geological consequences of superplumes. *Geology* 19: 963–966

Lasaga AC (1983) Geospeedometry: an extension of geothermometry. In: Saxena SK (ed) *Kinetics and equilibrium in mineral reactions*, Springer, Berlin Heidelberg New York, pp 82–114

Laske G, Masters G (1997) A global digital map of sediment thickness. *EOS Trans AGU* 78 F483

Le Pichon X (1983) Land-locked oceanic basins and continental collision: the eastern Mediterranean as a case example. In: Hsu KJ (ed) *Mountain building processes*. Academic, Orlando USA, pp 201–211

Le Pichon X, Francheteau J, Bonnin J (1976) Plate tectonics. *Developments in geotectonics* (2nd edn). New York, Amsterdam, 311 p

Le Pichon X, Angelier J, Sibuet JC (1982) Plate boundaries and extensional tectonics. *Tectonophys* 81: 239–256

Leeman (1989) Origin and development of the Snake River plain; an overview. In: Hanshaw PM (ed) *Volcanism and plutonism of western North America; vol 2 Snake River plain - Yellowstone province* Am Geophys Union, Washington 4–12

Lister GS, Etheridge MA (1989) Towards a general model. Detachment models for uplift and volcanism in the eastern highlands and their application to the origin of passive margin mountains. In: Johnson RW, Knutson J, Taylor SR (eds) *Intraplate volcanism in eastern Australia and New Zealand*. Cambridge Univ. Press. Cambridge, pp 297–313

Lister GA, Etheridge MA, Symonds PA (1986) Detachment faulting and the evolution of passive continental margins. *Geology* 14: 246–250

Lister GA, Etheridge MA, Symonds PA (1991) Detachment models for the formation of passive continental margins. *Tectonics* 10: 1038–1064

Lux DR, DeYoreo JJ, Guidotti CV and Decker ER (1986) Role of plutonism in low pressure metamorphic belt formation. *Nature* 323: 794–796

Lyon-Caen H, Molnar P (1983) Constraints on the structure of the Himalaya from an analysis of gravity anomalies and a flexural model of the lithosphere. *J Geophys Res* 88: 8171–8191

Lyon-Caen H, Molnar P (1989) Constraints on the deep structure and dynamic processes beneath the Alps and adjacent regions from an analysis of gravity anomalies. *Geophys J Int* 99: 19–32

Mackin JH (1948) Concept of the graded river. *Bull Geol Soc Am* 59: 463–512

Mahoney JJ, Coffin MF (1997) (eds) Large igneous provinces; continental oceanic, and planetary flood volcanicsm. *Geophysical Monograph* 100: 438 p

Malanson GP, Butler DR, Georgakakos KP (1992) Nonequilibrium geomorphic processes and deterministic chaos. *Geomorphology* 5: 311–322

Malniverno A, Pockalny RA (1990) Abyssal hill topography as an indicator of episodicity in crustal accretion and deformation. *Earth Planet Sci Lett* 99: 154–169

Mancktelow NS (1993) Tectonic overpressure in competent mafic layers and the development of isolated eclogites. *J Met Geol* 11: 801–812

Mancktelow NS (1995) Nonlithostatic pressure during sediment subduction and the development and exhumation of high pressure metamorphic rocks. *J Geophys Res* 100: 571–583

Mancktelow NS, Grasemann B (1997) Time-dependent effects of heat advection and topography on cooling histories during erosion. *Tectonophys* 270: 167–195

Mancktelow NS (1999) Finite-element modelling of single-layer folding in elasto-viscous materials: the effect of initial perturbation geometry. *J Struct Geol* 21: 161–177

Mandelbrot BB (1967) How long is the coast of Briotain? Statistical self similarity and fractional dimension. *Science* 156: 636–638

Mandelbrot B (1975) Stochastic models for the Earth's relief, the shape and fractal dimension of the coastlines and the number-area rule for islands. *Proc Nat Acad Sci, USA* 72: 3825–3828

Mandelbrot B (1982) *The fractal geometry of nature*. Freeman, New York, 460 p

Marsh JG, Lerch FJ, Putney BH, Felsentreger TL, Sanchez BV, et al. (1990) The GEM-T2 gravitational model. *J Geophys Res* 95: 22043–22071

Mastin L (1988) Effect of borehole deviation on breakout orientations. *J Geophys Res* 93: 9187–9195

McClay KR (1992) *Thrust tectonics*. Chapman and Hall, London, 447 p

McDonough WF, Rudnick RL (1998) Mineralogy and composition of the upper mantle. *Rev Mineral* 37: 139–164

McKenzie DP (1967) Some remarks on heat-flow and gravity anomalies. *J Geophys Res* 72: 6261–6273

McKenzie DP (1969a) The relationship between fault plane solutions for earthquakes and the directions of the principle stresses. *Bull Seismol Soc Am* 59: 591–601

McKenzie DP (1969b) Speculations on the consequences and causes of plate motions. *Geophys J Roy Astron Soc* 18: 1–32

McKenzie DP, Morgan WJ (1969) Evolution of triple junctions. *Nature* 224: 125–133

McKenzie DP (1972) Active tectonics in the Mediterranean region. *Geophys Roy Astron Soc* 30: 109–185

McKenzie DP, Roberts JM, Weiss NO (1974) Convection in the Earth's mantle: towards a numerical simulation. *J Fluid Mech* 62: 465–538

McKenzie DP (1977a) Surface deformation, gravity anomalies and convection. *Geophys J Roy Astron Soc* 48: 211–238

McKenzie DP (1977b) The initiation of trenches: a finite amplitude instability. In: Talwani M, Pitman III WC (eds) *Island arcs, deep-sea trenches and back arc basins*. Am Geophys U, Maurice Ewing Series 1: 57–61

McKenzie DP (1978) Some remarks on the development of sedimentary basins. *Earth Planet Sci Lett* 40: 25–32

McKenzie DP (1984) The generation and compaction of partially molten rock. *J Pet* 25: 713–765

McKenzie DP, Bickle MJ (1988) The volume and composition of melt generated by extension of the lithosphere. *J Pet* 29: 625–679

McKenzie K (2001) Evidence for a plate tectonic debate. *Earth Science Reviews* 55: 235–336

McLaren S, Sandiford M, Hand M (1999) High radiogenic heat producing granites and metamorphism - an example from the western Mt Isa Inlier, Australia. *Geology* 27: 679–682

Means WD (1976) Stress and strain – basic concepts of continuum mechanics for geologists. Springer, Berlin Heidelberg New York, 339 p

Medvedev SE, Podladchikov YY (1999) Ne extended thin-sheet approximation for geodynamic applications: 1. Model formulation. *Geophys J Internat* 136: 567–585

Menard HW (1964) *Marine Geology of the Pacific*. McGraw Hill, New York, 271 p

Meyerhoff AA (1995) Surge tectonic evolution of southeastern Asia: A geohydrodynamics approach. *Journal of SE Asian Earth Science* 12: 145–247

Michael AJ (1987) The use of focal mechanisms to determine stress: a control study. *J Geophys Res* 92: 357–368

Mildren SD, Sandiford M (1995) Heat refraction and low pressure metamorphism in the northern Flinders Ranges, South Australia. *Austral J Earth Sci* 42: 241–247

Miyashiro A (1973) Metamorphism and metamorphic belts. Allen and Unwin, New York, 492 p

Mohr O (1900) Welche Umstände bedingen die Elastizitätsgrenze und den Bruch eines Materials? *Z Ver dt Ing* 44: 1524–1530, 1572–1577

Molnar P (1992) Brace-Goetze strength profile, the partitioning of strike-slip and thrust faulting at zones of oblique convergence and the stress heat flow paradox of the San Andreas Fault. In: Evans B, Wong T (eds) *Fault mechanics and transport properties of rocks. A festschrift in honour of WF Brace*. Acad. Press, San Diego, pp 435–459

Molnar P, England PC (1990a) Temperatures, heat flux and frictional stress near major thrust faults. *J Geophys Res* 95: 4833–4856

Molnar P, England PC (1990b) Late Cenozoic uplift of mountain ranges and global climatic change: chicken or egg? *Nature* 346: 29–34

Molnar P, England PC (1995) Temperatures in zones of steady-state underthrusting of young oceanic lithosphere. *Earth Planet Sci Lett* 131: 57–70

Molnar P, Gibson JM (1996) A bound on the rheology of continental lithosphere using very long baseline interferometry: the velocity of South China with respect to Eurasia. *J Geophys Res* 101: 545–553

Molnar P, Lyon-Caen H (1988) Some simple physical aspects of the support, structure and evolution of mountain belts. *Geol Soc Am Spec Pap* 218: 179–207

Molnar P, Lyon-Caen H (1989) Fault plane solutions of earthquakes and active tectonics of the Tibetan plateau and its margins. *Geophys J Roy Astron Soc* 99: 123–153

Molnar P, Tapponier P (1975) Cenozoic tectonics of Asia: effects of a continental collision. *Science* 189: 419–426

Molnar P, Tapponier P (1978) Active tectonics of Tibet. *J Geophys Res* 83: 5361–5374

Molnar P, England P, Martinod J (1993) Mantle dynamics, uplift of the Tibetan plateau and the Indian monsoon. *Rev Geophys* 31: 357–396

Molnar P, Houseman G, Conrad C (1998) Rayleigh-Taylor instability and convective thinning of mechanically thickened lithosphere: effects of non-linear viscosity decreasing exponentially with depth and of horizontal shortening of the layer. *Geophys J Int* 133: 568–584

Montgomery DR (1994) Valley incision and the uplift of mountain peaks. *J Geophys Res* 99: 13913–13921

Morgan J (1968) Rises, trenches, great faults and crustal blocks. *J Geophys Res* 73: 1959–1982

Morgan WJ (1971) Convection plumes in the lower mantle. *Nature* 230: 42–42

Müller RD, Roest WR, Royer JY, Gahagan LM, Sclater JG (1997) Digital isochrons of the world's ocean floor. *J Geophys Res* 102: 3211–3214

Müller B, Reinecker J, Fuchs K (2000) The 2000 release of the world stress map (available online at <http://www-wsm.physik.uni-karlsruhe.de>)

Nabelek PI, Liu M (1999) Leucogranites in the Black Hills of South Dakota; the consequence of shear heating during continental collision. *Geology* 27: 523–526

Newman WI (1983) Nonlinear diffusion: Self-similarity and traveling waves. *Pure Appl Geophys* 121: 417–441

North GR, Cahalan RF, Coakley JA (1981) Energy balance climate models. *Rev Geophys Space Phys* 19: 91–121

Oertel G (1996) Stress and deformation. A handbook on tensors in geology. Oxford University Press, New York, 292 p

Ollier CD (1985) Morphotectonics of continental margins with great escarpments. In: Morrison M, Hack JT (eds) *Tectonic geomorphology*. Allan & Unwin, Boston, 390 p

Onsager L (1931) Reciprocal relations in irreversible processes. I. *Phys Rev* 37: 405–426

O'Rouke J (1993) Computational Geometry in C. Cambridge University Press, 346 p

Oxburgh ER, Turcotte DL (1974) Thermal gradients and regional metamorphism in overthrust terrains with special reference to the eastern Alps. *Schweiz Mineral Petrograph Mitt* 54: 641–662

Oxburgh ER (1980) Heat flow and magma genesis. In: Hargraves RB (ed) *Physics of magmatic processes*. Princeton University Press, Princeton NJ, pp 161–199

Oxburgh ER (1982) Heterogeneous lithospheric stretching in early history of orogenic belts. In: Hsü K (ed) *Mountain building processes*. Academic Press, London, pp 85–94

Parsons B, McKenzie D (1978) Mantle convection and the thermal structure of the plates. *J Geophys Res* 83: 4485–4496

Parsons B, Richter FM (1980) A relationship between the driving force and the geoid anomaly associated with the mid-ocean ridges. *Earth Planet Sci Lett* 51: 445–450

Parsons B, Sclater JG (1977) An analysis of the variation of ocean floor bathymetry with age. *J Geophys Res* 82: 803–827

Passchier CW, Myers JS, Kröner A (1990) Field geology of high-grade gneiss terrains. Springer Verlag, Berlin 150 p

Pavlis TL (1986) The role of strain heating in the evolution of megathrusts. *J Geophys Res* 91: 12407–12422

Peacock SM (1989) Numerical constraints on rates of metamorphism and fluid production fluid flux during regional metamorphism. *Geol Soc Am Bull* 101: 476–485

Perry J (1895) On the age of the Earth. 51: 224–227 & 341–342 & 582–585

Pfiffner OA, Lehner P, Heitzmann P, Mueller S, Steck A (eds) (1997) Deep structure of the Swiss Alps. Results of NRP20. Birkhäuser, Basel, 380 p

Pierce KL, Coleman SM (1986) Effect of height and orientation (microclimate) on geomorphic degradation rates and processes, late glacial terrace scarps in central Idaho. *Geol Soc Am Bull* 97: 869–885

Pinet C, Jaupart C (1987) The vertical distribution of radiogenic heat production in the Precambrian Crust of Norway and Sweden: Geothermal implications. *Geophys Res Lett* 14: 260–263

Pinter N, Brandon MT (1997) How erosion builds mountains. *Sci Am*, April 1997, pp 74–79

Pitman WC, Andrews JA (1985) Subsidence and thermal history of small pull-apart basins. In: Biddle K, Christie-Blick N (eds) Strike-slip deformation, basin formation and sedimentation. *Spec Publ Soc Econ Palaeont Mineral* 37: 45–119

Platt JP (1990) Thrust mechanics in highly overpressured accretionary wedges. *J Geophys Res* 95: 9025–9034

Platt JP (1993a) Mechanics of oblique convergence. *J Geophys Res* 98: 16239–16256

Platt JP (1993b) Exhumation of high-pressure rocks: a review of concepts and processes. *Terra Nova* 5: 119–133

Platt JP, England PC (1994) Convective removal of lithosphere beneath mountain belts: thermal and mechanical consequences. *Am J Sci* 294: 307–336

Pluim B, Marshack S (1997) Earth structure – an introduction to structural geology and tectonics. McGraw Hill, New York, 495 p

Pollack HN, Chapman DS (1977) On the regional variation of heat flow, geotherms and lithosphere thickness. *Tectonophys* 38: 279–296

Pollack HN, Hurter SJ, Johnson JR (1993) Heat flow from the earth's interior – Analysis of the global data set. *Rev Geophys* 31: 267–280

Pollard DD, Fletcher RC (2006) Fundamentals of Structural Geology. Cambridge University Press, Cambridge, 512 p

Powell R (1978) Equilibrium thermodynamics in petrology. Harper and Row, London, 284 p

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1989) Numerical receipes in Pascal. The art of scientific computing. Cambridge University Press, Cambridge, 759 p

Pribnow D, Hurter S (1998) Distribution of heat production in the crust. *EOS, Amer. Geophys* 79, 838

Price PH, Slack MR (1954) The effect of latent heat on numerical solutions of the heat flow equation. *Br J Appl Phys* 3: 379–384

Putnis A, McConnell JDC (1980) Principles of mineral behaviour. Geoscience Texts 1, Elsevier, New York, 257 p

Ramsay JG, Huber M (1983) The techniques of modern structural geology, Strain analysis. Academic Press, San Diego, 1: 1-300

Ramsay JG, Huber M (1987) The techniques of modern structural geology, Folds and fractures 2: 301-700.

Ramsay JG, Lisle RJ (2000) The techniques of modern structural geology, Applications of continuum mechanics in structural geology. Academic Press, San Diego, 3: 701–1061

Ranalli G (1987) Rheology of the Earth, 2nd ed. Chapman and Hall, London, 413 p

Ranalli G (1994) Nonlinear flexure and equivalent mechanical thickness of the lithosphere. *Tectonophys* 240: 107–114

Ratschbacher L, Frisch W, Linzer HG, Merle O (1991) Lateral extrusion in the Eastern Alps, part 2: structural analysis. *Tectonics* 10: 245–256

Reece G (1986) Microcomputer modelling by finite differences. MacMillian. Basingstoke UK, 125 p

Reiner M (1964) The Deborah number. *Physics Today* 17: 62

Reiner M (1969) Deformation, strain and flow. HK Lewis, London, 360 p

Richardson RM (1992) Ridge forces, absolute plate motions and the intra-plate stress field. *J Geophys Res* 97: 11739–11748

Richardson SW, England PC (1979) Metamorphic consequences of crustal wadelite production in overthrust orogenic zones. *Earth Planet Sci Lett* 42: 183–190

Ridley J (1989) Vertical movement in orogenic belts and the timing of metamorphism relative to deformation. In: Daly JS, Cliff RA, Yardley BWD (eds) Evolution of metamorphic belts. *Geol Soc Lond Spec Pub* 43: 103–115

Riesco M, Stüwe K, Reche J, Martinez FJ (2004) Silica depleted melting of pelites. Petrogenetic grid and application to the Susqueda Aureole, Spain. *J Met Geol* 22: 475–494

Ring U, Brandon MT, Willett SD, Lister GS (1999) Exhumation processes. In: Ring U, Brandon MT, Willett SD, Lister GS (eds) Exhumation processes: normal faulting, ductile flow and erosion. *Geol Soc Spec Pub* 154: 1–29

Ringwood (1988) Phase transformations and their bearing on the constitution and dynamics of the mantle. *Geochim Cosmochim Acta* 55: 2083–2110

Robinson P (1990) The eye of the petrographer, the mind of the petrologist. *Am Mineral* 76: 1781–1810

Robinson AH, Sale RD, Morrison JL, Muehrke PC (1984) Elements of Cartography, 5th edn. John Wiley and Sons, New York, 544 p

Robl J, Stüwe K (2005) Continental collision with finite indenter strength: 2. European Eastern Alps. *Tectonics* 24: doi: 10.1029/2004TC001741

Roering JJ, Kirchner JW, Dietrich E (2001) Hillslope evolution by nonlinear, slope dependent transport: Steady state morphology and equilibrium adjustment time scales. *J Geophys Res* 106: 16499–16513

Roy RF, Decker ER, Blackwell DD, Birch F (1968) Heat flow in the United States. *J Geophys Res* 73: 5207–5221

Royden L, Keen CE (1980) Rifting processes and thermal evolution of the continental margin of eastern Canada determined from subsidence curves. *Earth Planet Sci Lett* 51: 343–361

Royden LH (1993a) The tectonic expression of slab pull at continental convergent boundaries. *Tectonics* 12: 303–325

Royden LH (1993b) The steady-state thermal structure of eroding orogenic belts and accretionary prisms. *J Geophys Res* 98: 4487–4507

Ruppel C (1995) Extensional processes in continental lithosphere. *J Geophys Res* 100: 24187–24216

Rutland RWR (1965) Tectonic overpressures. In: Pitcher WS, Flinn GW (eds) Controls of metamorphism. Oliver and Boyd, New York, pp 119–139

Sabodini R, Lambeck K, Boschi E (eds) (1991) Glacial isostasy, sea level and mantle rheology. NATO ASI Series Serie C: Mathematical and physical sciences, 334 p

Sambridge M, Braun J, McQueen H (1995) Geophysical parameterization and interpolation of irregular data using natural neighbours. *Geophys J Int* 122: 837–857

Sandiford M (1989) Secular trends in the thermal evolution of metamorphic terrains. *Earth Planet Sci Lett* 95: 85–96

Sandiford M, Powell, R (1990) Some isostatic and thermal consequences of the vertical strain geometry in convergent orogens. *Earth and Planet Sci Lett* 98: 154–165

Sandiford M, Powell R (1991) Some remarks on high-temperature-low-pressure metamorphism in convergent orogens. *J Met Geol* 9: 333–340

Sandiford M, Martin N, Zhou S, Fraser G (1991) Mechanical consequences of granite emplacement during high-T, low-P metamorphism and the origin of anticlockwise PT paths. *Earth Planet Sci Lett* 107: 164–172

Sandiford M, Coblenz DD (1994) Plate-scale potential-energy distribution and the fragmentation of ageing plates. *Earth Planet Sci Lett* 126: 143–159

Sandiford M, Coblenz DD, Richardson RM (1995) Ridge-torques and continental collision in the Indo-Australian plate, *Geology* 23: 653–656

Sandiford M, Hand M (1998a) Controls on the locus of intraplate deformation in central Australia. *Earth Planet Sci Lett* 162: 97–110

Sandiford M, Hand M (1998b) Australian Proterozoic high-temperature, low pressure metamorphism in the conductive limit. In: Treloar P, O'Brian P (eds.) *What controls metamorphism*. Geol Soc London Spec Publ 138: 103–114

Sandiford M, Hand M, McLaren S, (1998) High geothermal gradient metamorphism during thermal subsidence. *Earth Planet Sci Lett* 163: 149–165

Sandiford M (1999) Mechanisms of basin inversion. *Tectonophys* 305: 109–120

Sandiford M (2002) Low thermal Peclet number intraplate orogeny in central Australia. *Earth Planet Sci Lett* 201: 309–320

Sandiford M, McLaren S, Neumann N (2002) Long term thermal consequences of the redistribution of heat producing elements associated with large scale granitic complexes. *J Met Geol* 20: 87–98

Saunders AD, Tarney J, Kerr AC, Kent RW (1996) The formation and fate of large oceanic igneous provinces. *Lithos* 37: 81–95

Sawyer DS (1985) Brittle failure in the upper mantle during extension of continental lithosphere. *J Geophys Res* 90: 3021–3025

Schaltegger U, Amundsen H, Jamtveit B, Frank M, Griffin WL, Gronvold K, Tronnes R, Torsvik T (2002) Contamination of OIB by underlying ancient continental lithosphere: U-Pb and Hf isotopes in zircons question EM1 and EM2 mantle components. *Geochim Cosmochim Acta* 66: p 673

Schatz JF, Simmons G (1972) Thermal conductivity of Earth materials at high temperatures. *J Geophys Res* 77: 6966–6983

Schmalholz SM, Podladchikov Y (1999) Buckling versus folding: Importance of Viscoelasticity. *Geophys Res Lett* 26: 2641–2644

Schmid DW (2005) Rigid polygons in shear. *Geol Soc Spec Pub* 245: 421–431

Scholz CH (1980) Shear heating and the state of stress on faults. *J Geophys Res* 85: 6174–6184

Schumm SA (1956) The evolution of drainge basins and slopes in badlands at Perth Amboy, New Jersey. *Bull Geol Soc Am* 67: 597–646

Schumm SA, Dumont JF, Holbrook JM (2000) *Active Tectonics and Alluvial Rivers*. Cambridge University Press. 276p.

Sclater JG, Christie PA (1980) Continental stretching: an explanation of the post-mid Cretaceous subsidence of the central North Sea basin. *J Geophys Res* 85: 3711–3739

Sclater JG, Jaupart C, 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

Silverstone J (1988) Evidence for east-west crustal extension in the Eastern Alps: implications for the unroofing history of the Tauern Window. *Tectonics* 7: 87–105

Seth HC (1999) Flood basalts and large igneous provinces from deep mantle plumes: fact, fiction and fallacy. *Tectonophys* 311: 1–29

Shemenda AI (1994) Subduction: Insights from physical modelling. Kluwer Academic Ser. *Modern Approaches in Geophysics*. Netherlands pp 215

Shreve RL (1967) Infinite topologically random networks. *J Geol* 75: 178–186

Sleep NH (1971) Thermal effects of formation of Atlantic continental margins by continental breakup. *Geophys J Roy Astron Soc* 24: 325–350

Sleep NH (1979) A thermal constraint on the duration of folding with reference to Acadian geology, New England (USA). *J Geol* 87: 583–589

Sleep NJ (1992) Hotspot volcanism and mantle plumes. *Ann Rev Earth and Planet Sci* 20: 19–43

Smith GD (1985) Numerical solutions of partial differential equations: finite difference methods, 3rd edn. Oxford Applied Mathematics and Computing Science Series. Clarendon Press, Oxford, 337 p

Smith WHF, Sandwell DT (1997) Global sea floor topography from satellite altimetry and ship depth soundings. *Science* 277: 1956–1962

Smoot NC (1997) Magma floods, micro plates and orthogonal intersections In Choi DR, Dickins JM (eds) *New concepts in global tectonics* Newsletter 5: 8–13

Snyder JP (1987) Map projections – a working manual. USGS Prof Pap 1395, 383 p

Snyder JP, Voxland RM (1989) An album of map projections: USGS Prof Pap 1453, 249 p

Sonder LJ, Chamberlain CP (1992) Tectonic controls of metamorphic field gradients. *Earth Planet Sci Lett* 111: 517–535

Sonder LJ, England P (1986) Vertical averages of rheology of the continental lithosphere: relation to „thin sheet“-parameters. *Earth Planet Sci Lett* 77: 81–90

Sonder LJ, Jones CH (1999) Western United States extension: How the west widened. *Ann Rev Earth Planet Sci* 27: 417–462

Spear FS (1993) Metamorphic phase equilibria and pressure-temperature-time paths. Min Soc Am Monograph, Book Crafters, Chelsea, Michigan, 799 p

Spear FS, Florence FP (1992) Thermobarometry in granulites: pitfalls and new approaches. *Precamb Res* 55: 209–241

Spear FS, Peacock SM (1989) Metamorphic pressure-temperature-time paths. Am Geophys Union Short Course in Geology AGU press 7, 102 p

Spiegel MR (1968) Mathematical handbook of formulas and tables. Schaums outline series in mathematics. McGraw Hill, New York, 271 p

Spiegelman M, McKenzie D (1987) Simple 2-D models for melt extraction at mid-ocean ridges and island arcs. *Earth Planet Sci Lett* 83: 137–152

Steckler MS, Watts AB (1978) Subsidence of Atlantic-type continental margin off New York. *Earth Planet Sci Lett* 41: 1–13

Steckler MS, Watts AB (1981) Subsidence history and tectonic evolution of Atlantic-type continental margins. In: Scrutton, RA (ed) *Dynamics of passive margins*. Am Geophys U, Geodynamics Series 6: 184–196

Stefan J (1891) Über die Theorie der Eisbildung, insbesondere über die Eisbildung im Polarmeere. *Annalen der Physik und Chemie* 42: 269–286

Stephansson O (1974) Stress-induced diffusion during folding. *Tectonophys* 22: 233–251

Stockmal GS (1983) Modelling of large scale accretionary wedge deformation. *J Geophys Res* 88: 8271–8287

Strahler AN (1964) Quantitative geomorphology of drainage basins and channel networks. In: Chow VT (ed) *Handbook of applied hydrology*, McGraw-Hill, New York Section 4-II

Strang G (1988) Linear algebra and its applications, 3rd edn. Harcourt Brace Jovanovich International Edition, 505 p

Strömgard KE (1973) Stress distribution during formation of boudinage and pressure shadows. *Tectonophys* 16: 215–248

Stüwe K (1991) Flexural constraints on the denudation of asymmetric mountain belts. *J Geophys Res* 96: 10401–10408

Stüwe K, Sandiford M, Powell R (1993a) On the origin of repeated metamorphic-deformation events in low-P high-T terranes. *Geology* 21: 829–832

Stüwe K, Will TM, Zhou S (1993b) On the timing relationships between fluid production and metamorphism in metamorphic piles. Some implications for the origin of post-metamorphic gold mineralisation. *Earth Planet Sci Lett* 114: 417–430

Stüwe K (1994) Process and age constraints for the formation of Ayers Rock / Australia – An example for two-dimensional mass diffusion with pinned boundaries. *Z Geomorph* 38: 435–455

Stüwe K, Sandiford M (1994) On the contribution of deviatoric stresses to metamorphic PT paths: an example appropriate to low-P, high-T metamorphism. *J Met Geol* 12: 445–454

Stüwe K, White L, Brown R (1994) The influence of eroding topography on steady-state isotherms. Application to fission track analysis. *Earth Planet Sci Lett* 124: 63–74

Stüwe K (1995) The buffering effects of latent heat of fusion on the equilibration of partially melted metamorphic rocks. *Tectonophys* 248: 39–51

Stüwe K, Sandiford M (1995) Mantle-lithospheric deformation and crustal metamorphism with some speculations on the thermal and mechanical significance of the Tauern Event, Eastern Alps. *Tectonophys* 242: 115–132

Stüwe K (1998a) Heat sources for Eoalpine metamorphism in the Eastern Alps. A discussion. *Tectonophys* 287: 251–269

Stüwe K (1998b) Keyword: Diffusion. In: Marshall CP, Fairbridge RW (eds.) *Encyclopedia of Geochemistry*. Chapman and Hall, London 133–140

Stüwe K (1998c) Tectonic constraints on the timing relationships of metamorphism, fluid production and gold-bearing quartz vein emplacement. *Ore Geol Rev* 13: 219–228

Stüwe K, Barr T (1998) On uplift and exhumation during convergence. *Tectonics* 17: 80–88

Stüwe K, Barr T (2000) On the relationship between surface uplift and gravitational extension. *Tectonics* 19: 1056–1064

Stüwe K, Hintermüller M (2000) Topography and isotherms revisited: the influence of laterally migrating drainage divides. *Earth Planet Sci Lett* 184: 287–303

Suess E (1885) *Das Antlitz der Erde*, 778 p

Summerfield MA (1991) *Global Geomorphology*. Longman House, Burnt Mill, England, 537 p

Summerfield MA, Hutton NJ (1994) Natural controls of fluvial denudation rates in major world drainage basins. *J Geophys Res* 99: 13871–13883

Suppe J (1981) Mechanics of mountain building and metamorphism in Taiwan. *Mem Geol Soc China* 4: 67–89

Suppe J (1985) *Principles of structural geology*. Prentice Hall, Englewood Cliffs, N.J., 537 p

Suppe J (1987) The active Taiwan mountain belt. In: Schaeer JP, Rodgers J (eds) *The anatomy of mountain ranges*. Princeton University Press, Princeton N. Y., pp 277–293

Tapponier P, Molnar P (1976) Slip-line field theory and large scale continental tectonics. *Nature* 264: 319–324

Tapponier P, Peltzer G, Le Dain AY, Armijo R, Cobbolt P (1982) Propagating extrusion tectonics in Asia: new insights from simple experiments with plasticine. *Geology* 10: 611–616

Tapponier P, Meyer B, Avouac JP, Peltzer G, Gaudemer Y, Shunmin G, Hongfa X, Kelun Y, Zhitai C, Shuhua C, Huagang D (1990) Active thrusting and folding in the Qilian Shan, and decoupling between upper crust and mantle in northeasternt Tibet. *Earth Planet Sci Lett* 97: 382–403

Tarboton DG (1996) Fractal rivers networks, Horton's laws and Tokunga cyclicity. *J Hydrology* 187: 105–117

Taylor FB (1910) Bearing of the Tertiary mountain belt on the origin of the Earth's plan. *Bull Geol Soc Am* 21: 179–226

Taylor B, Martinez F (2003) Back-arc basin basalt systematics. *Earth Planet Sci Lett* 210: 481–497

Tenczer V, Stüwe K, Barr T (2001) Pressure anomalies around cylindrical objects in simple shear. *J Struct Geol* 23: 777–788

Tenczer V, Stüwe K (2003) The metamorphic field gradient in the eclogite type locality. *J Met Geol* 21: 377–393

Thelin GP, Pike RJ (1991) Landforms of the conterminous United States; a digital shaded-relief portrayal. *Misc Invest Ser USGS Rep I-2206* 16p

Thompson AB, England PC (1984) Pressure-temperature-time paths of regional metamorphism Part II: some petrological constraints from mineral assemblages in metamorphic rocks. *J Pet* 25: 929–955

duToit A (1937) Our wandering continents. Oliver and Boyd, Edinburgh, 366 p

Tong H (1983) Threshold models in non-linear time series analysis. Lecture notes in statistics, Springer, Berlin Heidelberg New York, 323 p

Tucker GE, Slingerland R (1994) Erosional dynamics, flexural isostasy and long-lived escarpments: A numerical modeling study. *J Geophys Res* 99: 12229–12243

Tucker GE, Slingerland R (1996) Predicting sediment flux from fold and thrust belts. *Basin Res* 8: 329–349

Turcotte DL (1979) Flexure. *Advances in Geophys* 21: 51–86

Turcotte DL (1983) Mechanisms of crustal deformation. *J Geol Soc London* 140: 701–724

Turcotte DL (1997) Fractals and chaos in geology and geophysics, 2nd edn. Cambridge University Press, Cambridge, 398 p

Turcotte DL, Oxburgh ER (1973) Mid-plate tectonics. *Nature*, 244: 337–339

Turcotte DL, Schubert G (1982; 2002 2nd ed) Geodynamics. Applications of continuum physics to geological problems. John Wiley and Sons, New York, 450 p

Turcotte DL, Haxby WF, Ockendon JR (1977) Lithospheric instabilities. In: Talwani M, Pitman III WC (eds) Island arcs, deep sea trenches and back-arc basins. Am Geophys U, Maurice Ewing Series Washington 1: 63–69

Turner S, Arnaud N, Liu J, Rogers N, Hawkesworth C, Harris N, Kelley S, Van Calsteren P, Deng W (1995) Post collision shoshonitic volcanism in the Tibetan plateau: Implications for convective thinning of the lithosphere and the source of ocean island basalts. *J Pet* 37: 45–71

Twiss RJ, Moores EM (1992) Structural geology. Freeman and Company, New York 532 p

Vilotte JP, Daignieres M, Madariaga R (1982) Numerical modeling of intraplate deformation: simple mechanical models of continental collision. *J Geophys Res* 87: 10709–10728

Vine F, Mathews D (1963) Magnetic anomalies over oceanic ridges, *Nature* 199: 947–949

Voorhoeve H, Houseman G (1988) The thermal evolution of lithosphere extending on a low angle detachment zone. *Basin Res* 1: 1–9

Waschbusch PJ, Royden LH (1992) Episodicity in foredeep basins. *Geology* 20: 915–918

Watts AB, Ryan WBF (1976) Flexure of the lithosphere and continental margin basins. *Tectonophys* 36: 24–44

Watts AB (1976) Gravity and bathymetry in the central Pacific ocean. *J Geophys Res* 81: 1533–1553

Wees JD, Jong KC, Cloetingh S (1992) Two dimensional  $P$ - $T$ - $t$ -modelling and the dynamics of extension and inversion in the Betic Zone (SE Spain). *Tectonophys* 302: 305–324

Wegener A (1912a) Die Entstehung der Kontinente. Petermanns Geograph Mitt 127: 185–308

Wegener A (1912b) Die Entstehung der Kontinente. *Geol Rundsch* 3: 276–292

Wegener A (1915) Die Entstehung der Kontinente und Ozeane. Sammlung Vieweg 23, 94 p

Wei M, Sandwell D (2006) Estimates of ridge-axis heat flow from depth and age data. *Tectonophys* 417: 325–335

Weinberg R (1997) Diapir-driven crustal convection; decompression melting, renewal of the magma source and the origin of nested plutons. *Tectonophys* 271: 217–229

Weijermars R (1997) Principles of rock mechanics. Alboran Science Publishing, Amsterdam, 359 p

Wells PRA (1980) Thermal models for the magmatic accretion and subsequent metamorphism of continental crust. *Earth Planet Sci Lett* 46: 253–265

Wernicke B (1985) Uniform sense normal simple shear of the continental lithosphere. *Canad J Earth Sci* 22: 108–125

Wessel P, Smith WHF (1995) New version of the generic mapping tool released, EOS, Am Geophys U 76, 329

Wessel P, Smith WHF (1998) New, Improved Version of Generic Mapping Tools Released, EOS Trans. AGU, 79 (47), 579

Wheeler J (1991) Structural evolution of a subducted continental sliver: the northern Dora Maira massif, Italian Alps. *J Geol Soc London* 148: 1103–1113

White R, McKenzie D (1989) Magmatism at rift zones: the generation of volcanic continental margins and flood basalts. *J Geophys Res* 94: 7685–7729

White RW, Powell R, Holland TJB (2001) Calculation of partial melting equilibria in the system  $\text{Na}_2\text{O}$ - $\text{CaO}$ - $\text{K}_2\text{O}$ - $\text{FeO}$ - $\text{MgO}$ - $\text{Al}_2\text{O}_3$ - $\text{SiO}_2$ - $\text{H}_2\text{O}$ . *J Met Geol* 19: 139–153

Wickham SM (1987) The segregation and emplacement of granitic magmas. *J Geol Soc London* 144: 281–297

Wignall PB (2001) Large igneous provinces and mass extinctions. *Earth Sci Rev* 53: 1–33

Wilde M, Stock J (1997) Compression directions in southern California (from Santa Barbara to Los Angeles basin) obtained from borehole breakouts. *J Geophys Res* 102: 4969–4983

Will TM (1998) Phase equilibria in metamorphic rocks. Lecture notes in Earth Science, Springer 315 p

Willet SD (1992) Dynamic and kinematic growth and change of a Coulomb wedge. In: McClay KR (ed) Thrust tectonics. Chapman and Hall, pp 19–31

Willet SD, Beaumont C, Fullsack P (1993) Mechanical model for the tectonics of doubly vergent compressional orogens. *Geology* 21: 371–374

Willet SD (1999) Orogeny and orography: The effects of erosion on the structure of mountain belts. *J Geophys Res* 104: 28957–28981

Willet SD, Brandon MT (2002) On steady states in mountain belts. *Geology* 30: 175–178

Willett SD, Schlunegger F, Picotti V (2006) Messinian climate change and erosional destruction of the central European Alps. *Geology* 34: 613–616

Willgoose G, Bras RL, Rodriguez-Iturbe I (1991) Results from a new model of river basin evolution. *Earth Surface Processes and Landforms* 16: 237–254

Wilson JT (1963) A possible origin of the Hawaiian islands. *Canad J Phys* 41: 863–868

Wilson JT (1972) Continents adrift. Readings from Scientific American Freeman WH and Company, San Francisco, 172 p

Wilson M (1993) Plate-moving mechanisms: constraints and controversies. *J Geol Soc London* 150: 923–926

Winter J (2001) Introduction to igneous and metamorphic petrology. Prentice Hall, 699 p

Wintsch RP, Andrews MS (1988) Deformation induced growth of sillimanite: “stress”minerals revisited. *J Geol* 96: 143–161

Xu G, Will TM, Powell R (1994) A calculated petrogenetic grid for the system  $K_2FeO$ - $MgO$ - $Al_2O_3$ - $SiO_2$ - $H_2O$ , with particular reference to contact metamorphosed pelites. *J Met Geol* 12: 99–119

Yale LB, Carpenter SJ (1998) Large igneous provinces and giant dike swarms; proxies for supercontinent cyclicity and mantle convection. *Earth Planet Sci Lett* 163: 109–122

Yardley BWD (1989) An introduction to metamorphic petrology. Longman Harlow UK, 248 p

Zen E-an (1966) Construction of pressure-temperature diagrams for multi component systems after the method of Schreinemakers: a geometric approach. USGS Bull 1225, 56 p

Zhou S, Sandiford M (1992) On the stability of isostatically compensated mountain belts. *J Geophys Res* 97: 14207–14221

Zhou S, Stüwe K (1994) Modeling of dynamic uplift, denudation rates thermomechanical consequences of erosion in isostatically compensated mountain belts. *J Geophys Res* 99: 13923–13939

Ziegler PA (1992) Plate tectonics, plate moving mechanisms and rifting. *Tectonophysics* 215: 9–34

Ziegler PA (1993) Plate-moving mechanisms: their relative importance. *J Geol Soc London* 150: 927–940

Zill DG (1986) A first course in differential equations with applications (3rd edition). PWS-Kent Publishing, Boston, 520 p

Zoback ML (1992) First- and second-order patterns of stress in the lithosphere: the world stress map project. *J Geophys Res* 97: 11703–11728

Zoback ML, Haimson BC (1983) Hydraulic fracturing stress measurements US Nat Comm for Rock Mechanics, Nat Acad Press, Washington D. C., 270 p

# Index

---

## A

- Acceleration
  - angular, 25
  - gravitational, 26, 75
  - of plates, 26
- Accretionary wedge, 106, 246, 309
- Accuracy
  - definition, 5
  - numerical, 401, 402
- Acoustic energy, 69
- Activation
  - energy, 231, 345
  - volume, 345
- ADI-method, 404
- Adiabatic
  - compressibility, 75
  - gradient, 76
  - melting, 34
  - processes, 74
- Advection, 80
- Africa
  - lithosphere, 35
  - rift, 40
- Airy, 164
- Altiplano, 300
- Amontons law, 226
- Analytical solution, 396
- Andersons theory, 226
- Angle of friction, 225
- Angular
  - acceleration, 25
  - momentum, 27
  - velocity, 25
  - velocity of plates, 41
- Arc
  - island, 108
  - magmatic, 108
- Argand number, 305
- Arrhenius relationship, 231, 344

## Asthenosphere

- definition, 36
- Atmosphere CO<sub>2</sub>, 329
- Australia, 178, 281
- Ayers rock, 9

---

## B

- Back stripping, 274
- Back thrust, 315
- Back-arc basin, 42, 256, 270
- Barometry, 346
- Barrier reef, 178
- Barrow zone, 93, 292
- Basin
  - back-arc, 42, 256
  - fore-arc, 270
  - foreland, 42, 270
  - intra-continental, 270
  - McKenzie model, 277
  - molasse, 172
  - origin, 266
  - types, 269
- Bathymetry, 16
- Bending
  - of continental plates, 172
  - of oceanic plates, 171
- Benioff zone, 42, 108
- Bessel Ellipsoid, 144
- Black smokers, 42
- Boundary condition
  - at infinity, 102, 110
  - definition, 398
  - mechanical, 246
  - orogenic, 28
  - thermal, 120
- Boundary layer
  - mechanical, 36
  - thermal, 36, 319
- Brace-Goetze lithosphere, 235

## Brittle

- deformation, 223
- ductile transition, 226, 234, 327

Buchan zones, 341

Bulk modulus, 223

Buoyancy force, 249, 296

Butterfly effect, 334

Byerlees law, 225

---

**C**

Chaos, 200, 202, 330

Cheese tectonics, 20

Chemical boundary layer, 36

Chinese Loess Plateau, 203

Clapeyron curve, 107

Coast line

- fractal dimension, 204

Cohesion, 224

Collision

- momentum, 218
- of continental plates, 41
- of oceanic plates, 44
- of plates generally, 41

Compaction, 271

Compensation depth, 162, 249

Complementary error function, 61

Compressibility

- adiabatic, 75

Conductivity

- changes, 56
- thermal, 53

Confidence interval, 420

Constitutive relationship, 220

Contact metamorphism, 57, 82, 114, 322, 352

Continent

- mean elevation, 147

Continental

- collision, 285
- elevation, 434
- geotherm, 90
- lithosphere, 39

Convection, 80

Convergence

- of plates, 44, 298

Cooling

- of earth, 59
- of half space, 103
- of oceanic lithosphere, 102

Cooling rate, 114

Coordinates

- cartesian, 22

– cylindrical, 59

– polar, 23

– reference frame, 144

– spherical, 59

Core

- of earth, 34

Corner flow model, 313

COST-drilling program, 267

Coulomb criterion, 224, 310

Crank-Nicolson, 403

Crust

- continental, 39

– oceanic, 102, 255

Crystallization

- heat, 120

– of dikes, 120

Curvature

- definition, 394

---

**D**

Deborah number, 306

Deccan, 253, 317

Declination, 25

Deformation

- laws, 220
- maps, 230
- permanent, 223

deformation tensor, 141

Degradation

- of scarps, 187

Delamination, 319

Density

- of the crust, 36, 56, 166
- of the mantle lithosphere, 36, 166

Derivatives

- order, 395

Dextral, 228

Differential

- partial, 395
- total, 395

Differential equation, 393

- analytical solution, 397

– method of solution, 413, 416

- numerical solution, 397

– order, 395

– ordinary, 396

- partial, 396

Differential stress, 74, 213

Diffusion

- cation, 349

– equation, 55

- mass, 184

- nonlinear, 186
- numerical, 81, 405
- radial, 59
- time scale, 62
- Diffusion age, 192
- Diffusivity, 231
  - erosional, 185
  - thermal, 56, 62
- Dike
  - swarms, 116
  - thermal evolution, 113
- Dimension
  - dimensionless, 9
  - one-dimensional, 13
  - spatial, 6
  - three-dimensional, 8
  - two-dimensional, 11
  - units, 25
  - zero-dimensional, 8
- Dislocation energy, 69
- Displacement, 295
- Dissipative heat production, 68
- Distribution
  - Gaussian, 419
  - normal, 419
- Divergence
  - definition, 412
- Dorn law, 232
- Driving force
  - plate tectonic, 242, 245, 295
- Ductile deformation, 216
- Dynamic
  - collision, 294
  - extension, 282
- Dynamics
  - on a sphere, 25

---

- E**
- Earth
  - age of, 59
  - core, 34
  - layer structure of, 33
  - radius, 143
  - shape, 143
  - total heat loss, 89
- Earthquake, 228
  - distribution, 244
- Eclogite, 325
- Elastic
  - bulge, 172
  - thickness, 173
- Elasticity
  - law, 221
  - of continents, 178
  - of oceanic plates, 176
- Elevation
  - of mountain belts, 165
- Ellipticity, 143
- Elongation, 73
  - definition, 140
- endogenic process, 179
- Endothermic
  - reaction, 76
- Energy
  - acoustic, 69
  - activation, 231
  - equilibrium, 54
  - Gibbs, 347
  - kinetic, 219
  - potential, 69, 282
- Equator
  - radius, 143, 431
- Equilibrium
  - coefficient, 347
  - force, 218
  - isostatic, 161
  - steady state, 61, 89, 157, 199, 312
  - thermodynamic, 347
  - transient, 89
- Erosion model, 157, 180, 181
- Error function, 61, 427
  - derivative of, 104
- Escarpment, 172
- Euler
  - reference frame, 82
- Euler pole, 25
- Eulerian
  - description, 81
  - reference frame, 84, 144, 179
- Exhumation, 149
  - mechanism, 324
- exogenic process, 179
- Exothermic
  - reaction, 76
- Expansion
  - thermal, 75, 165
- Explicit methods, 402
- Extension
  - active, 265, 282, 316
  - brittle, 226
  - continental, 148
  - force, 296
  - lateral, 303
  - passive, 316
  - passive, 265, 282

- rheology change, 236
- self limiting, 282
- two-dimensional, 280

Extrusion

- lateral, 303, 308
- vertical, 325

---

**F**

Fault

- orientation, 226
- plane solution, 228
- temperatures, 132, 150
- transform, 44

Feedback, 329

Ficks law, 52

Ficks law, 185

Finite

- difference method, 137, 399
- element method, 397

Finite difference method

- stability, 401

Flexural

- parameter, 176
- rigidity, 173

Flexure

- of plates, 170

Flood basalt, 255

Flood basalts, 316

Flow

- of heat, 52
- of mass, 52
- viscous, 231
- volumetric, 52

Flow law, 220

Fluid

- flow, 85, 108
- heat transport, 85
- Newtonian, 231
- pressure, 225

Fold and thrust belt, 309

Folding

- of isotherms, 127
- rate, 127

Force, 210

- applied to oceanic plates, 253
- driving plates, 248
- extension, 296
- transmission, 245
- units, 25

Force balance

- extension, 282
- in orogens, 298

Fore-arc basin, 42, 256, 270

Foreland basin, 256, 270

Fourier

- cell number, 401
- first law, 52, 185, 343
- second law, 54
- series, 415

Fractal, 329

- geometrical, 203
- statistical, 203

Friction coefficient, 224

Friction heat, 68, 132, 338, 360

Fusion heat, 77

---

**G**

Gas constant, 345

Gaussian distribution, 419

Geobarometer, 213

Geobarometry, 346

Geochrononology, 345

Geoid

- anomaly, 144, 250
- definition, 143

Geometry

- on a sphere, 23

Geosyncline, 46

Geotherm

- continental, 90
- definition, 89
- oceanic, 103

Geothermometry, 346

Gondwana, 15

GPS, 227

Graded rivers, 199

Gradient, 58

- definition, 412

Granitic weathering, 188

Gravitational

- acceleration, 75, 170, 218, 339
- constant, 4
- stress, 249

Great barrier reef, 178

Great circle, 23

Great Escarpment, 281

Great escarpment, 172, 184

Greenwich, 23

---

**H**

Hack Law, 196

Half-space cooling model, 102, 112, 169

Hawaii, 176, 208

Heat

- capacity, 54
- conductivity, 53
- content, 78
- content of intrusions, 124
- content of terrains, 123
- dissipative, 132
- friction, 360
- of fusion, 77
- refraction, 56

Heat flow

- at the surface, 101, 242
- definition, 53
- from mantle, 92
- mantle, 289, 293
- oceanic, 105
- surface, 289
- two-dimensional, 58, 118, 129
- unit, 53

Heat production

- by friction, 68
- by radioactivity, 67, 92
- by reaction, 76
- distribution, 94

Heat transfer

- by active transport, 80
- by diffusion, 52
- types of, 51

Himalaya, 57, 194, 219

Histogram, 419

Hooks law, 222

Hydrostatic isostasy, 161

Hydrostatic pressure, 162, 201, 220

---

**I**

Iceland, 255

Implicit methods, 402

Inclination, 25

Incompressibility, 223

Indenter, 306

Indus River, 198

Initial condition

- definition, 398

Interferometry, 227

Intra-continental basins, 270

Intrusion

- cooling of, 114
- heat content, 124
- spherical, 117
- thermal evolution, 110

Invariant

- tensor, 212

Island arc, 44, 108, 310

Isograds

- stretching, 132
- telescoping, 132

Isostasy, 162

- Airy, 164
- equilibration rate, 161
- flexural, 170
- hydrostatic, 161
- Pratt, 164

Isotherms

- around faults, 132
- folding of, 126
- topography, 128

---

**J**

Jacobian matrix, 141

Julia set, 329

---

**K**

Karoo, 253

Kelvin model, 59

Kinematics

- on a sphere, 25
- vertical, 149

Kinetic

- energy, 219

Kinetics

- of reaction, 198

Knick point, 197

Kochs snow flake, 203

Kurtosis, 420

---

**L**

La Place equation, 60

Lagrangian

- coordinates, 397
- description, 81, 191
- reference frame, 144

Landscape model, 199

Landslide, 200

Large igneous provinces, 316

Latent heat

- of evaporation, 76
- of fusion, 77, 122

lateral extrusion, 306

Laurasia, 15

Layers of earth, 34

Least squares, 417

Liquidus, 78

Lithosphere  
 – definition, 36  
 – mantle part, 36, 97  
 – mechanical definition, 246  
 – subdivision, 35  
 – temperature at base, 36  
 – thermal definition, 88  
 – types of, 38  
 Lithosphere thickness  
 – elastic, 165, 238  
 – mechanical, 37  
 – thermal, 37  
 Lithostatic pressure, 220  
 Load  
 – external, 175, 238  
 – internal, 175, 238  
 Longitudinal wave, 228  
 Lorentz attractor, 330  
 Loxodrome, 25  
 LPHT-terrains, 321  
 Luzon, 310

---

**M**

Main Central Thrust, 57  
 Mantle  
 – definition, 34  
 – density, 165  
 – discontinuity, 257  
 – heat flow, 92, 288  
 – lithosphere, 36, 97, 102, 235  
 – plume, 255, 259, 260  
 – viscosity, 161, 162  
 – wedge convection, 108  
 Mantle lithosphere, 164, 265  
 Mantle plume, 282  
 Mantle plumes, 315  
 Map projection, 28  
 – cylindrical, 29  
 – true to angle, 29  
 – true to area, 29  
 Mariana Trench, 15, 22  
 Mass extinctions, 318  
 Mass flux, 52, 185  
 Matrix, 212  
 Mean  
 – depth, 147  
 – elevation, 147, 434  
 – surface heat flow, 434  
 Mercator projection, 29  
 Meridian, 23  
 Metamorphic field gradient  
 – definition, 358

Metamorphic geotherm  
 – definition, 359  
 Metamorphic gradient  
 – contact, 116, 362  
 – inverted, 133  
 – regional, 291, 357  
 Metamorphism  
 – Barrovian, 292, 321  
 – contact, 57, 322  
 – facies, 340  
 – LPHT, 321  
 – regional, 285, 291, 359  
 – space-time relations, 359  
 – space-time relationship, 291  
 Michigan Basin, 282  
 Mid-oceanic ridge, 16, 102, 167, 253  
 – manometer, 36, 166  
 Model  
 – concept, 2  
 – requirements, 5  
 Moho  
 – definition, 34  
 – temperature, 100  
 Mohr-Coulomb criterion, 224  
 Molasse, 172  
 Moment  
 – of inertia, 27  
 Momentum, 25, 27, 218  
 – angular, 27  
 Mt Blanc, 10  
 Mt Everest, 10

---

**N**

Nabla operator, 58  
 Neotectonics, 179  
 Network law, 194  
 New degrees, 23  
 Newton  
 – units, 26  
 Newtonian fluid, 231  
 Nodal plane, 229  
 Normal fault, 226  
 Normal stress, 211  
 Normalize, 9  
 Nucleation rate, 343  
 Numerical  
 – diffusion, 81, 405  
 – solution, 396

---

**O**

Obduction, 42

Ocean  
– mean depth, 147  
Ocean floor age, 19  
Oceanic  
– CHABLIS model, 106  
– cooling, 102  
– geotherms, 101  
– water depth, 167, 434  
Olivine, 34, 71, 106, 232, 240, 255, 283, 430  
Orogenic force balance, 298

---

**P**

P-wave, 228  
Paired metamorphic belt, 108  
Palaeopiezometry, 346  
Palaeostress, 227  
Pannonian Basin, 282  
Parametric diagram, 114, 338  
Passive margin, 172, 269  
Peclet number, 86  
Periodic problems, 125  
Petrology, 342  
Phase  
– transition, 77  
Phase rule, 348  
Piezometry, 210  
Piezothermal array, 361  
Ping-pong ball model, 22  
Plane strain, 11, 306  
Plate  
– graveyards, 257  
– reconstruction, 18  
– tectonic driving force, 245  
– tectonics, 15  
Plate boundary  
– convergent, 41, 285  
– diffuse, 20  
– divergent, 40  
– passive, 40, 269, 283  
– transform, 44  
Plume  
– mantle, 255  
Poisson constant, 173, 222  
Porosity, 85, 272  
Post orogenic collapse, 307  
Potential  
– surface, 8, 412  
Potential energy, 69, 248, 282  
– continents, 257  
Power law  
– definition, 231  
– exponent, 233  
Pratt, 164  
Pre-exponent constant, 231  
Precision  
– definition, 5  
Pressure, 364  
– chemical, 213  
– hydrostatic, 220  
– mechanical, 213  
– tectonic, 220, 326  
Pressure-depth relationship, 338  
Projection  
– of maps, 28  
Pseudotachylite, 74  
PT-path, 337  
– anticlockwise, 353  
– interpretation, 292  
– parts of, 345  
– shape, 353  
Pure shear, 154  
Pythagoras, 411

---

**Q**

Quadratic equation, 427

---

**R**

Radioactive heat, 67  
Radius  
– of earth, 143  
Reaction  
– endotherm, 76  
– exotherm, 76  
– rate, 343  
Reaction heat  
– dehydration, 77  
– melting, 77  
Red Sea, 46, 282  
Reference frame, 144  
– Eulerian, 84  
– Lagrangian, 144  
Reference levels, 142  
Refraction  
– of heat, 56  
Retro-arc foreland basin, 270  
Rheology  
– definition, 233  
– of continents, 234  
– of oceans, 243  
– olivine, 238  
– quartz, 238  
Rhumb line, 25

- Ridge push, 253
- Rift, 40
  - basin, 269
- Rifted margin sequence, 316
- Rifting, 316
- Ring of Fire, 22
- Roll back, 256
- Rotation pole, 25

---

- S**
- San Andreas Fault, 44
- Scalar
  - definition, 410
- Scarp
  - degradation, 187
- Schmidt net, 29, 229
- Sea level
  - mean, 147
- Sea mount, 171
- Sedimentary basin
  - origin, 266
- Seismicity, 244
- Shear
  - pure, 281
  - simple, 154, 281
  - strength, 74
  - stress, 211
- Shear strength
  - definition, 216
- Shear zone, 70, 153, 281, 314
- Shields, 152, 321
  - continental, 96
- Simple shear, 154
- Skewness, 420
- Skin depth, 96
- Slab breakoff, 257
- Slab pull, 255
- Slope
  - of a function, 393
- Snowball earth, 329
- Solar system, 4
- Solidus, 78
- Solution
  - analytical, 397
  - numerical, 396
- South Georgia, 256
- Specific heat, 54, 56
- Speed
  - definition, 25
- Sphere
  - geometry, 20
  - latitude, 23
  - longitude, 23
  - meridian, 23
- Spheroid, 143
- Stability
  - numerical, 401, 402
- Stable geotherm, 89
- Standard deviation, 420
- Steady state
  - equilibrium, 61
  - geotherm, 89
  - surface elevation, 157
- Stefan problem, 120
- Strain
  - definition, 140, 222
  - hardening, 214
  - longitudinal, 73
  - plane, 11, 303, 306
  - rate, 13, 141
- Strength
  - continents, 239
  - definition, 216
  - integrated, 239
  - mean, 73, 300
  - of the continents, 235
  - of the oceans, 243
- Stress
  - deviatoric, 69, 213
  - differential, 74, 213
  - mean, 213
  - principal, 212
  - shear, 211
  - tensor, 210
- Stress field, 17
- Stress profile
  - quantitative, 239
- Stretch
  - definition, 140
- Strike slip fault, 226, 230
- Subduction Erosion, 153
- Subduction zone, 42, 106
  - accretionary wedge, 309
  - force balance, 42
  - geometry, 7
  - isotherms, 107
  - roll back, 256
  - shape, 171, 176
  - slab break off, 257
- Subsidence
  - analysis, 271
  - isostatic, 268
  - mechanisms, 268
  - tectonic, 267
  - total, 267

- Surface
  - heat flow, 101, 105, 242, 289, 331
- Surface elevation, 165
  - evolution, 157
- Surface heat flow
  - mean, 434

---

**T**

- Taiwan, 309
- Tectonic
  - driving force, 296, 298
  - pressure, 220, 326
- Temperature
  - absolute, 345
  - dimensionless, 10
  - fluctuations, 126
  - steps, 111
- Tensor
  - calculation, 410
  - stress, 210
- Thermal
  - boundary layer, 36, 64, 319
  - conductivity, 53
  - expansion, 75, 165
- Thermodynamic
  - degree of freedom, 348
  - equilibrium, 347
- Thermomechanics, 294
- Thermometry, 346
- Thickness
  - elastic, 174
  - of elastic lithosphere, 173
  - of mechanical lithosphere, 37
  - of thermal lithosphere, 37, 88
- Thin sheet approximation, 13, 303
- Thomas algorithm, 404
- Threshold, 200
  - mechanism, 332
- Thrust, 226
- Tien Shan, 246
- Time constant
  - diffusive, 344
  - thermal, 62
- Topography
  - dynamic, 170
- Torque, 25, 253
- Traction
  - definition, 210
- Transform
  - basin, 269
  - plate boundary, 44
- Transform fault, 25

- Trench, 176
  - suction, 255
- Triple junction, 45
- Two-dimensional
  - deformation, 146

---

**U**

- Under plating, 81
- Units
  - in this book, 429
  - mistakes with, 25
- Uplift, 149
  - of rocks, 151
  - surface, 156
- UTM projection, 30

---

**V**

- Variance, 420
- Vector
  - calculation, 410
  - definition, 410
- Velocity
  - angular, 25
  - definition, 25
- Vertical
  - kinematics, 149
- Viscosity
  - dynamic, 230
  - effective, 231
  - kinematic, 230
  - mantle, 161
- Volcano
  - shape, 201

---

**W**

- Wegener, 15
- WGS84, 143
- Wilson cycle, 46, 282
- Wulff net, 29

---

**Y**

- Young's modulus, 173, 222