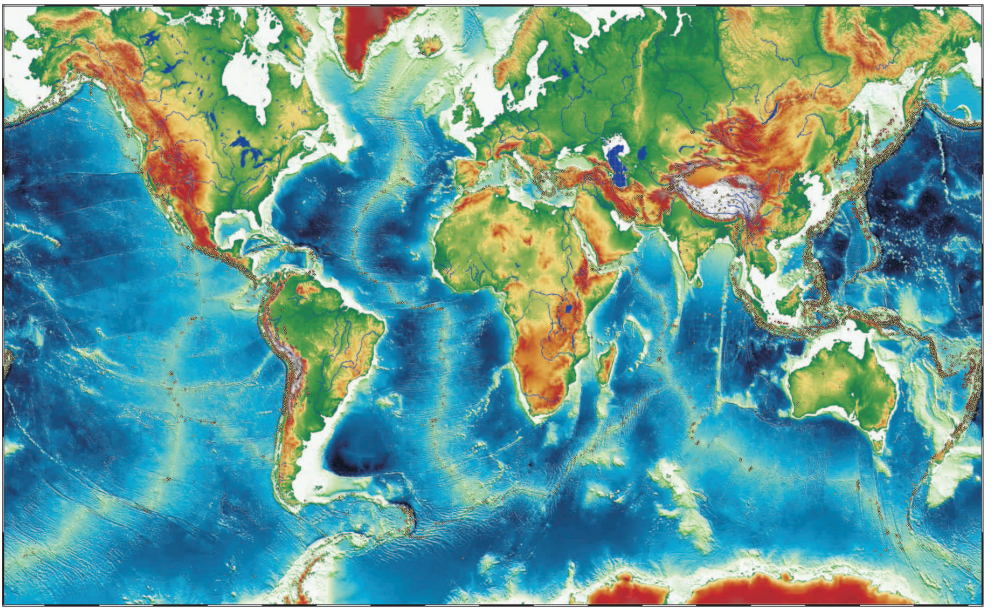
Physical Geodynamics

Innsbruck WS2021/2022

Kurt Stüwe

Manuskript version December 2021



Introduction

This course is a 1 hour/week one-semester course on geodynamics of the lithosphere taught at Innsbruck University in the winter term 2021/22. The course is compiled and shortened from a 2 hour/week course at Graz University which, in turn, is an extract from my textbook "Geodynamics of the Lithosphere" published in 2007. The content of this course can be found in a series of textbooks. For those of you who want to understand this subject in more depth, I recommend the following three books:

- Turcotte D.L. and Schubert G. (2014): Geodynamics, 3rd ed., Cambridge University Press, 496 p.
- Fowler C.M.R. (1990): The Solid Earth. An Introduction to Global Geophysics. Cambridge University Press. 472 p.
- Stüwe K. (2007): Geodynamics of the Lithosphere. Springer Verlag, Berlin Heidelberg, 493 p.

I hope you enjoy the course and look forward to work with you!

December 2021, Kurt Stüwe

Content

1 Unit: Vertical 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 is three: a crust, a mantle and a core (layers above the surface of solid earth like the 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 2 900 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. 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 $\mathrm{km\,s^{-1}}$.

2. A transition zone from about 400 km to the 650 km discontinuity.

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.

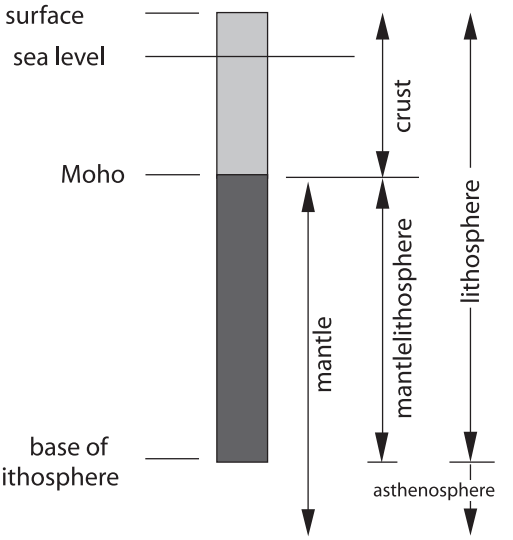


Figure 1: Nomenclature of different parts of the outer shells of the earth

1.1 Crust and Lithosphere

The lithosphere is the outer solid layer of the earth (s. sect. 4.1). 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 to be familiar with the nomenclature illustrated in Fig. 1. For comparison see also the original thoughts of Alfred Wegener illustrated in Fig. 2.

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. 54).

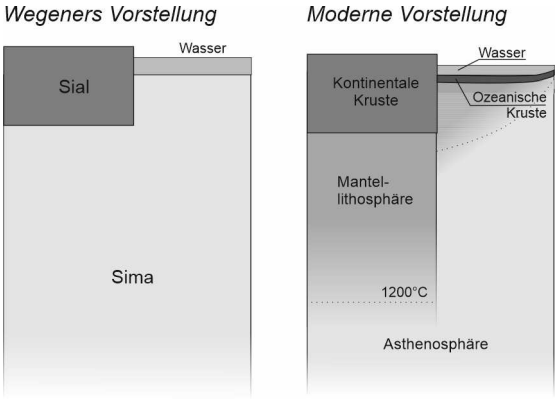


Figure 2: The outer shells of Earth according to Alfred Wegener and our modern understanding.

A schematic but characteristic thermal and density profile of the lithosphere is shown in Fig. 3 and Fig. 4. A very large number of the geodynamic processes 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 1 200°C or 1 300°C isotherm (sect. 4.1). 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 more or less constant below the depth z_1 for the next few hundrets of kilometers. The curves of Fig. 3 and Fig. 4 will be useful help throughout this course.

1.1.1 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*) (Parsons and

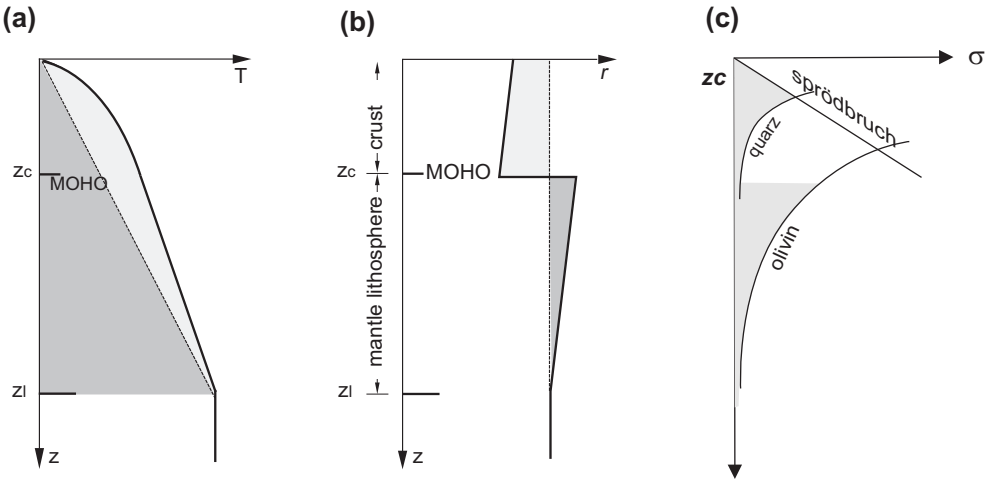


Figure 3: **a** Temperature **b** density and **c** failure stress of the continental lithosphere as a function of depth. The depth of the Moho below surface is z_c, that of the whole lithosphere is z_l. **a** The curvature of the geotherm within the continental crust is caused by radioactive heat production. 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. 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. **c** The distribution of deviatoric stresses in the lithosphere in the cartoon shown here is known as the *Brace-Goetze lithoshere* after the first geologists who suggested this model in the seventies. Detailed explanation of the curves needs to be done elsewhere.

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.

One thing can be said with certainty: the definition of the lithosphere depends on the question that is being asked. 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

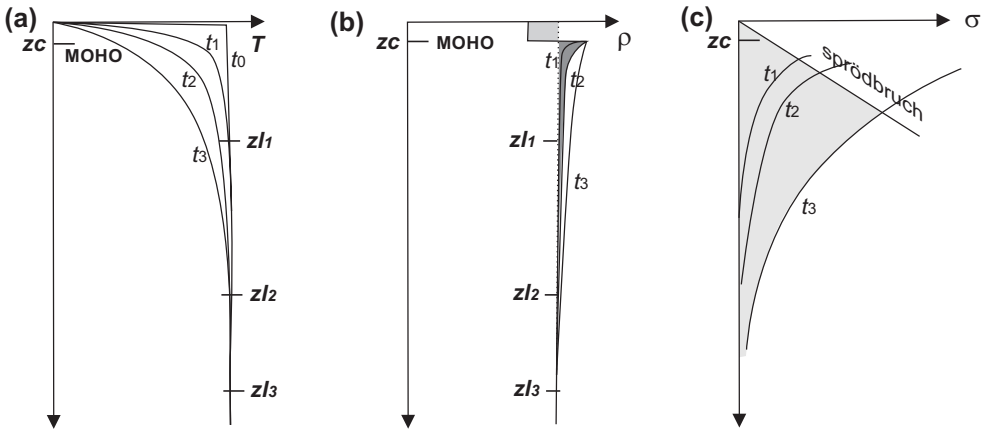


Figure 4: **a** Temperature **b** density and **c** failure stress of the oceanic lithosphere as a function of depth. The depth of the Moho below surface is z_c, that of the whole lithosphere is z_l. **a** The curvature of the geotherm is caused by successice cooling over time steps t1, t2, t3. **b** 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) for the shown time step. This is about 30 my. Older oceanic lithosphere is negatively bouyant. **c** The distribution of deviatoric stresses in the lithosphere shows that the total strength of the oceanic lithosphere (shaded area) is large than that of the continental lithoshere. although the oceanic lithosphere is thinner. This is because the dominating olivine rheology is relevant for temparatures as low as 100 degrees in oceans, where as olivien rheology is only relevant above about 600 degrees in continents.

is meaningful, because the mid-oceanic ridges may be interpreted as manometers of the upper mantle (s. p. 56, 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. 4.1 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.

1.1.2 Types of Lithosphere

There is 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.

• **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 mWm^{-2} (Pollack et al. 1993; Wei and Sandwell 2006).

• **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 ae 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 mWm^{-2} (Pollack et al. 1993; Wei and Sandwell 2006).

2 Plate Boundaries

The surface of earth can be divided into seven major lithospheric plates plus a number of smaller plates (Fig. 5). 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. Most geodynamically interesting processes occur along the plate boundaries. These boundaries can be divided according to (a) their kinematics or (b) according to whch plates 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,

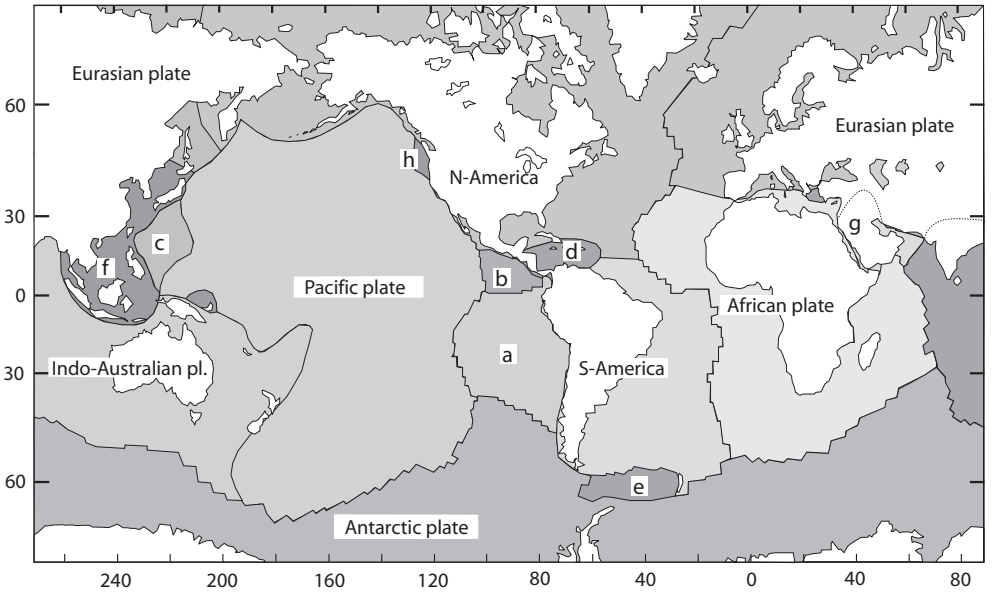


Figure 5: Plate tectonic division of the surface of earth. Continents are white and oceans are shaded. The difference between continental and oceanic lithosphere is not shown. Note that the plate boundaries coincide only in a few places with the coast lines that delineate the continents. The seven major 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

Table 1: 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 $\cdot 10^{-7o}$/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
Eurasic – 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 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.

2.1 Divergent Plate Boundaries

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 know examples (in order of progressive rift development) are the Rheingraben, the Central African rift system and the Red Sea.

Divergent plate margins between two oceanic plates (mid-oceanic ridges) are – in most cases – the last stage of a rift.

2.1.1 Mid ocean ridges

Oceanic lithosphere is created at the mid-ocen ridges. These ridges lie at 2500 m water depth and can be seen as a "hole" through the lithospheric plates where we can see the level of the asthenosphere. 5-7 km thick oceanic crust is created there and then stays of constant thickness with distance from the ridge. In contrast, the mantle part of the lithosphere successively thickens with distance from the ridge. In the course we have discussed them in detail

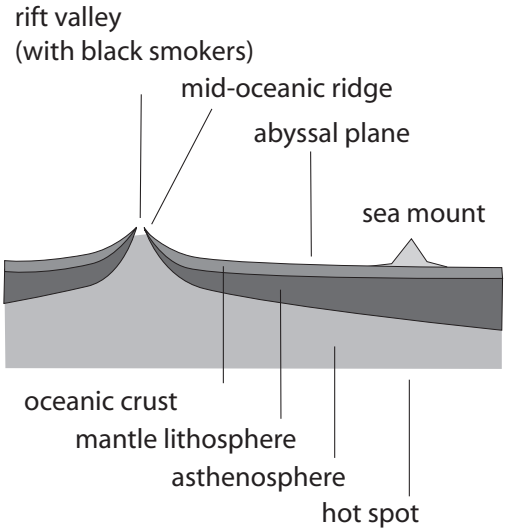


Figure 6: Nomenclature of some of the most important parts of oceanic lithosphere

• **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. 7a. 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. 7b. 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. 7b 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.

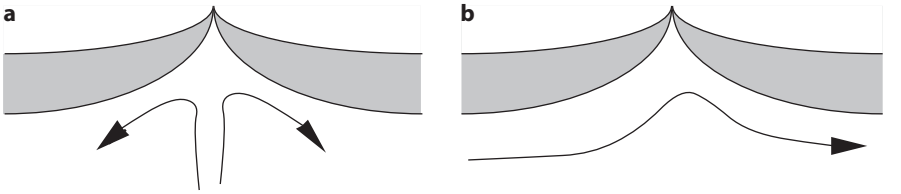


Figure 7: Cartoon showing two possible motions of the asthenosphere below mid oceanic ridges. **a** Asthenospheric material wells up below the mid ocean ridge in form of a *mantle plume*. During this process, adiabatic decompression of asthenosphere material will cause massive partial melting. It is 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. **b** shows the mantle motion that is thought to be representative for most mid oceanic ridges.

2.1.2 Passive margins

When continents rift, this is preceeded by an extensional phase int he continent and the extension of continents including basin formation is its own subject not discussed in this course. However, it should be said that rifting is usually associated with *more*thinning of the mantle part of the lithoshere than in the crust so that surface uplift occurs. This is called rift flank uplift and is seen in the African rift, the sides of the Rhine graben or both sides of the Red Sea. Rift flanks lead to the formation of a highly asymmetric mountain range that may form Great Escarpments (Fig. 9). Rifting can, however, occur asymmetrically (Fig. 8), so that one side has the preferential thinning in the crust and the other side in the mantle part of the lithoshere. This is the case for the rifting between Australia and New Zealand.

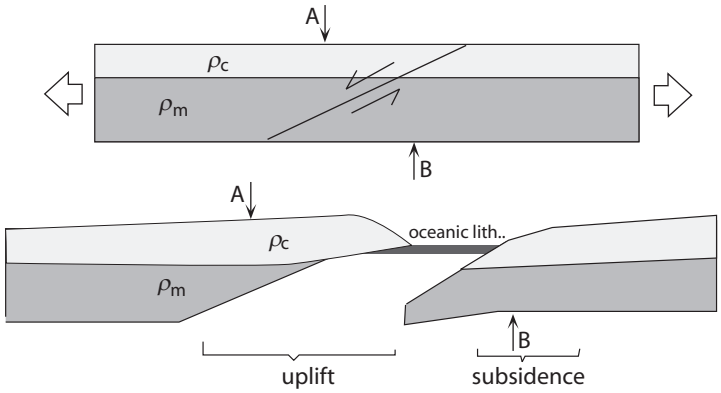


Figure 8: Extension of the lithosphere by normal faulting along a low angle normal fault that transects the entire lithosphere. The upper cartoon shows the situation at the onset of extension, the lower cartoon after full separation of the two plates and the development of passive continental margins. 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

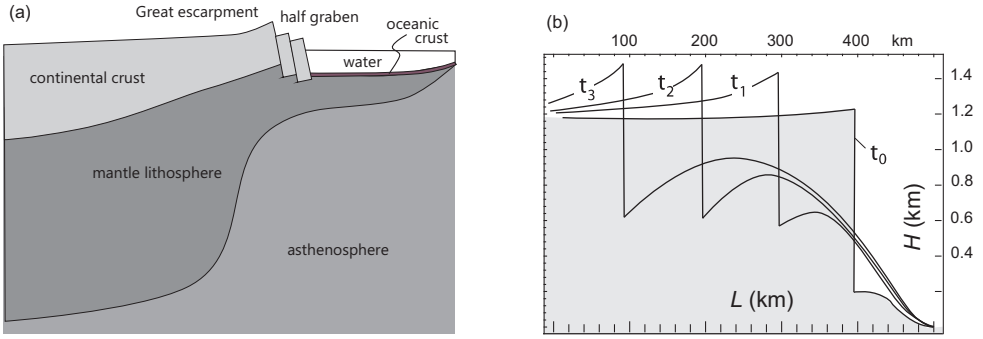


Figure 9: (a) geometry of a passive margin. Note that in the region of highest rift flank uplift, the mantle part of the lithosphere is more thinned than the crust. (b) Numerical model of the flexural response to a retreating Great Escarpment.

2.2 Convergent plate margins

Convergent pate boundaries may form between two continental plates, between two oceanic plates or between a continental and an oceanic plate. In convergent plate boundaries 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. 10). 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*.

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 to 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*.

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 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. This leads to the formation of the collisional mountain ranges that form most of the topographically high mountain belts of our globe. During this head on collision the crust gets typically pervasively deformed, while the mantle parts of the lithosphere override each other.

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. 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.

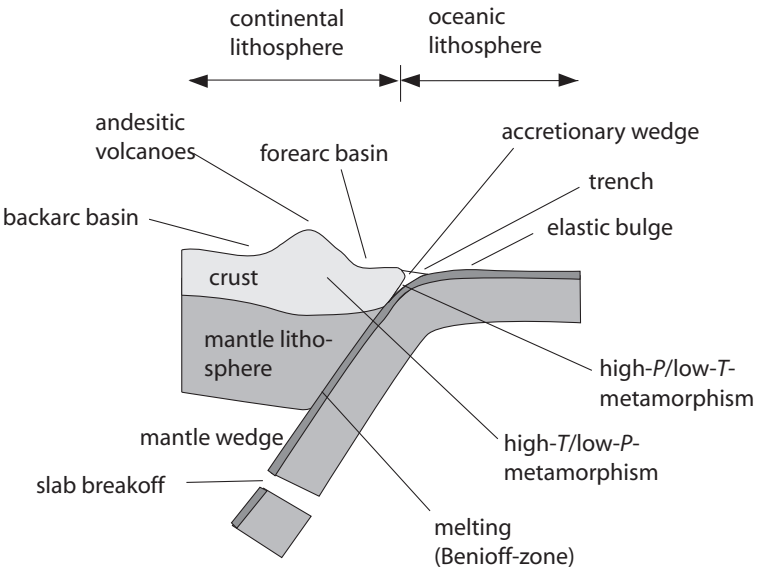


Figure 10: Nomenclature of important parts of subduction zones. Forearc- and backarc basins are even more clearly developed during subduction of oceanic lithosphere underneath another oceanic plate. Good descriptions of various phenomena on this figure can be found for slab break off by: Blanckenburg and Davies (1995); about the mantle wedge by: Spiegelman and McKenzie (1987) (sect. 2.2.1); about metamorphism by: Miyashiro (1973); about the elastic fore bulge: sect. 8.1

.

2.2.1 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. 11). 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.

• **Isotherms in Subduction Zones** Fig. 11 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. 15; 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). 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.

At a temperature of about 1 600 °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 $1.7 \cdot 10^5$ 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.

Fig. 11 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*.

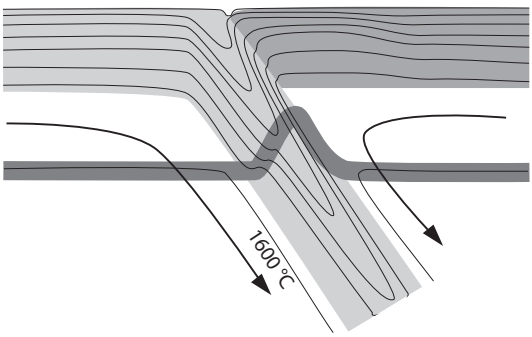


Figure 11: Schematic illustration of the temperature distribution in subduction zones. the subducted lithosphere is shaded light, the upper plate dark. The thick dark line that follows the 1 600 °C-isotherm outside the subduction zone is the Clapeyron-curve. It marks the olivine-spinel phase transition

• **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° 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. 12). Their model describes the motion of partial melts through the mantle

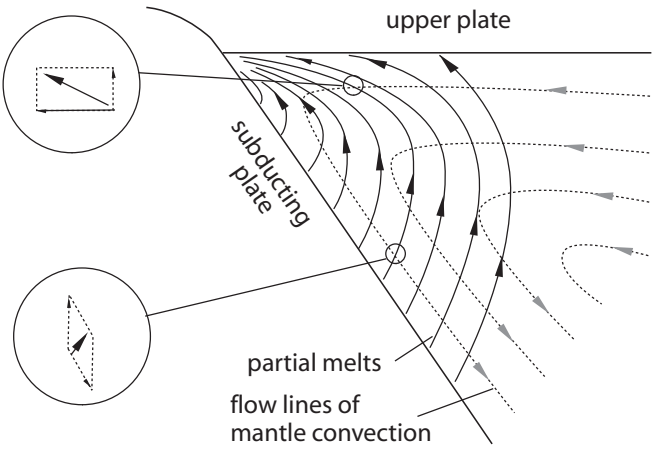


Figure 12: The motion of melts in and above the Benioff zone according to Spiegelman and McKenzie (1987). The dashed lines show the convective motion in the asthenosphere, the continuous lines are the motion of the partial melts. The enlarged sections show how the velocity field of the partial melts is given by the sum of the upward velocity of the melt and the motion in the mantle wedge

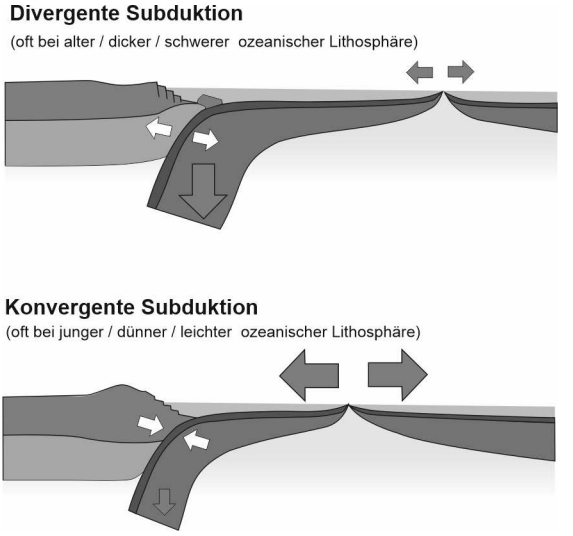


Figure 13: Cartoon illustrating the fact that subduction zones may be convergent or divergent, depending on the age of the oceanic lithosphere and the relative magnitude of ridge push and slab pull.

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. 12 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. 12). This elegant model is a beautiful example for a successful model description of fluid flow in deforming rocks.

3 Unit: The Diffusion 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. It turns out 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. Thus, the diffusion equation is the first equation in this course 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. The equation is a combination of two fundamental laws of heat conduction:

3.1 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:

$$q = -k\frac{\mathrm{d}T}{\mathrm{d}z} \quad . \tag{1}$$

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 $\mathrm{d}T/\mathrm{d}z$ 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: $\mathrm{m^3\,s^{-1}\,m^{-2} = m\,s^{-1}}$). 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 $\mathrm{m^2\,s^{-1}}$. In contrast, the flow of mass has the units $\mathrm{kg\,s^{-1}\,m^{-2}}$. 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. 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: $\mathrm{J\,s^{-1}\,m^{-2} = W\,m^{-2}}$). The thermal gradient now replaces the topographic gradient of the river. Because of

historical reasons heat flow is sometimes given in heat flow units, or hfu. One hfu corresponds to 10^{-6} cal s^{-1} cm^{-2} and can easily be converted into W m^{-2}. The units of the proportionality constant k, in eq. 1, follows now easily from the units of the other components of the equation: Because temperature has the units of K (or °C) and z has the unit m, k must have the units J s^{-1} m^{-1} 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. 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 $\mathrm{d}T/\mathrm{d}z$, 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. 1 states that the heat flow must also change.

3.2 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. 1 and may be written as:

$$\frac{\partial T}{\partial t} \propto -\frac{\partial q}{\partial z} \quad . \tag{2}$$

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. 14). If the heat flow into the cube of Fig. 14 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:

$$H = T\rho c_p \tag{3}$$

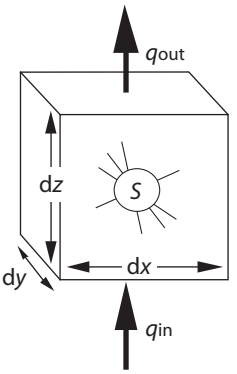


Figure 14: The flow of heat in a unity volume of rock. The heat production inside this volume S, is not considered until we discuss real continental geotherms.

where H is the volumetric heat content in J m^{-3}. 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 $J\,kg^{-1}\,K^{-1}$ 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. 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. 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 ρ, 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. 2 as:

$$\rho c_p \frac{\partial T}{\partial t} = -\frac{\partial q}{\partial z} \quad . \tag{4}$$

It should now be straight forward to understand eq. 4 intuitively using Fig. 14. 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. 2 to eq. 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.

3.3 The Diffusion Equation

If we substitute Fourier's law of heat conduction (eq. 1) into the thermal energy balance of eq. 4, we arrive at:

$$\rho c_p \frac{\partial T}{\partial t} = \frac{\partial \left(k \frac{\partial T}{\partial z}\right)}{\partial z} \quad . \tag{5}$$

Eq. 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. 5 significantly. k can then be taken out of the differential and we can write:

$$\rho c_p \frac{\partial T}{\partial t} = k \frac{\partial \partial T}{\partial z \partial z} \qquad \text{or :} \qquad \frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} \quad . \tag{6}$$

The constants k, ρ and c_p are now summarized to $\kappa = k/(\rho c_p)$. κ is called thermal diffusivity. Eq. 6 can also be understood intuitively, without following the detailed derivation given above. Eq. 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. 6, then remember that the first differential of a function describes its slope (or: "gradient", or: "rate") and the second its curvature.

Figure 15 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

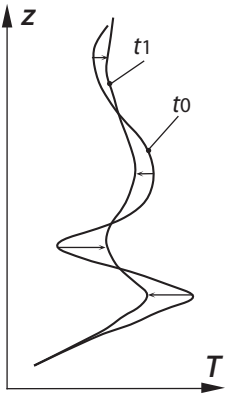


Figure 15: The thermal equilibration of a random temperature profile. The temperature profile is drawn at two different time steps t_0 and t_1. Note that the largest change in temperature between the two time steps has occurred in those places of the profile where the curvature of the profile is the largest (s. eq. 4). Where the curvature of the profile is zero (at the inflection points) the temperature does not change at all.

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. 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. Some of this will be done in the next units.

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

Table 2: Thermal conductivities and heat capacities of some rocks and common materials. k is given in J $\mathrm{s^{-1}}$ $\mathrm{m^{-1}K^{-1}}$ and c_p in J $\mathrm{kg^{-1}K^{-1}}$. 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).

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

4 Unit: Stable Continental Geotherms

Three fundamental processes create and redistribute heat in the continental lithosphere: *conduction*, *advection* and *production*. If we add heat advection and heat production terms to the diffusion equation discussed above, then a full one-dimensional description of the thermal energy balance for the lithosphere has the form:

$$\frac{\partial T}{\partial t} = \left(\frac{k}{\rho c_p}\right)\frac{\partial^2 T}{\partial z^2} + u\frac{\partial T}{\partial z} + \left(\frac{S}{\rho c_p}\right) \quad , \tag{7}$$

where the diffusivity is the ratio of conductivity and density $\times$ heat capacity: $\kappa = k/\rho c_p$. The central term on the right hand side of this equation describes advection (at the rate u) and advection may be due to transport of mass by *erosion, deformation, magma* or *fluid* and we will discuss some of these later in this course. In the most right term on the right side of the equation above, the heat production S may have *mechanical, chemical* and *radioactive* contributions. In this section we learn to describe aspects of the thermal structure of the continental lithosphere.

4.1 Thermal Definition of the Lithosphere

The lithosphere may be defined thermally or mechanically. According to the thermal definition, the lithosphere is the outer shell of 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 Earth. This boundary layer looses 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 $\mathrm{W\,m^{-2}}$. The total surface area of the continents is about $A_c = 2 \cdot 10^8$ $\mathrm{km^2}$. Thus, the total heat loss of earth from the continents is $1.3 \cdot 10^{13}$ $\mathrm{J\,s^{-1}}$. 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). In this section we calculate the quantitative shape of *stable continental geotherms*. That is, geotherms in which the temperature does not change anymore as is the case in most continental shields. This is in contrast to *transient geotherms* where the temperature changes because it is still equilibrating. This is the case in oceanic lithosphere and in most active orogens.

4.2 Stable Geotherms: The Relevant Equation

For the stable or steady state case, the heat conduction equation (eq. 6) or the full thermal energy balance (eq. 7) 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. Neglecting advection (because we deal with *stable* geotherms) eq.7 simplifies to:

$$\frac{\partial T}{\partial t} = \left(\frac{k}{\rho c_p}\right) \frac{\partial^2 T}{\partial z^2} + \frac{S}{\rho c_p} \quad . \tag{8}$$

We need the heat production term to account for radioactivity, which is of substantial importance to stable geotherms. For *steady state* geotherms, there is no change of the temperature with time. This means:

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

Eq. 8 simplifies to:

$$\left(\frac{k}{\rho c_p}\right) \frac{\mathrm{d}^2 T}{\mathrm{d}z^2} + \frac{S}{\rho c_p} = 0 \quad . \tag{9}$$

Note that eq. 9 is no *partial* differential equation anymore. By canceling out of the constants we get:

$$k \frac{\mathrm{d}^2 T}{\mathrm{d}z^2} = -S \quad . \tag{10}$$

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

4.3 The Contribution of Radioactivity to Stable Geotherms

The radioactive heat production rate of rocks is of the order of some microwatts per cubic meter. A typical value measured from samples at the Earth's surface is: $S = 2 - 5 \mu \mathrm{W\,m^{-3}} \equiv 2 - 5 \cdot 10^{-6}\ \mathrm{W\,m^{-3}}$. 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 ($z_\mathrm{c} = 30$ km) then the surface heat flow caused by radioactivity is:

$$q = q_\mathrm{s} = S \cdot z_\mathrm{c} = 0.03 \mathrm{W m^{-2}} \quad . \tag{11}$$

This can be converted into a temperature gradient using eq. 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 $k = 3\ \mathrm{W\,m^{-1}\,K^{-1}}$, then the assumptions from above indicate: $\mathrm{d}T/\mathrm{d}z = q/k = 0.05\,^\circ\mathrm{C\,m^{-1}} = 50\,^\circ\mathrm{C\,km^{-1}}$. This geothermal gradient of

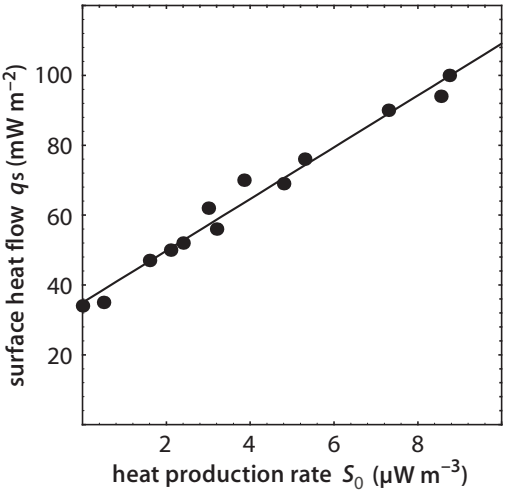


Figure 16: Measured data of surface heat flow q_s and surface heat production S_0 in the eastern US. The best line that fits the data is described by the equation $q_s = 0.035 + 7\,413 S_0$. Accordingly, the thickness of the layer that produces heat at a rate S is 7 413 m thick and the contribution of mantle heat flow to the total heat flow is 0.035 W m^{-2} (after Roy et al. 1968).

50 °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.

4.3.1 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_{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. 16). The straight line that fits these data has the form:

$$q_s = q_m + q_{\mathrm{rad}} = q_m + z_{\mathrm{rad}} S_0 \quad . \tag{12}$$

In this equation, q_s is the surface heat flow, q_m is the mantle heat flow, q_{rad} the radiogenically produced heat flow and z_{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 and the value of z_{rad} is given by its slope. The data of Roy et al. (1968) show that z_{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 $z_{rad}S_0$. This product corresponds to the area underneath the different model curves in Fig. 17.

In order to obtain a continuous function for the rate of heat production with depth, the most elegant assumption is the assumption that there is a continuous exponential drop off in radioactive heat production with depth (model *c* in Fig. 17). This model has the great advantage that there is no discontinuity in the heat production in the crust at z_{rad} and we do therefore not need several equations to describe a single geotherm. We assume that:

$$S_{(z)} = S_0 e^{\left(-\frac{z}{h_r}\right)} \quad . \tag{13}$$

The variable h_r is called the *characteristic drop off* or *skin depth* of heat production. According to eq. 13, 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 to calculate a geotherm is therefore:

$$k\frac{d^2T}{dz^2} = -S_0 e^{\left(-\frac{z}{h_r}\right)} \quad . \tag{14}$$

Before we use this equation, we make some further qualitative considerations of the behaviour of surface heat flow during lithospheric thickening or thinning.

• **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_{rad}:

$$q_s = q_m + q_{rad} \quad . \tag{15}$$

In this equation, the radiogenic heat flow is given by: $q_{rad} = S_{rad}z_{rad}$, as we explained when we discussed Fig. 16 (see also eq. 12). England and Thompson (1984) assumed that the radiogenically caused heat flow is comparable to the mantle heat flow ($q_{rad} \approx q_m$) and that the mantle heat flow remains unchanged, regardless of the thickness of the crust.

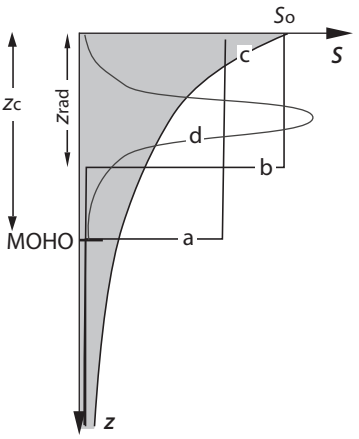


Figure 17: Four simple models describing the distribution of heat production with depth in the crust (s. Haack 1983). The total heat production of the crust is given by the area underneath the model curves. It is the same for all four models and is shaded for model *c*. *a* Constant concentration in the entire crust and no heat production in the mantle. *b* Constant concentration in the upper crust in a layer with the thickness $z_{\rm rad}$ and no heat production below that. *c* Exponential drop off of the heat production with depth. *d* Heat production peaking in the middle crust.

Thickening of the crust without thickening of the mantle part of the lithosphere doubles the radiogenic heat flow (because $z_{\rm rad}$ is doubled) but does nothing to the mantle heat flow. We can write:

$$q_{\rm s} = q_{\rm m} + 2q_{\rm rad} \quad . \tag{16}$$

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 $q_{\rm rad} = q_{\rm m}$ (Eq. 15).

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:

$$q_{\rm s} = \frac{q_{\rm m}}{2} + 2q_{\rm rad} \quad . \tag{17}$$

Thus, if $q_{\rm rad} = q_{\rm m}$ 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. 15. If $q_{\rm rad} = q_{\rm m}/2$, then thickening or thinning of the lithosphere as a whole does not change the surface heat flow at all.

4.4 Realistic Continental Geotherms

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:

$T = T_\mathrm{l}$ at the depth $z = z_\mathrm{l}$. This choice allows us to describe temperatures in the entire lithosphere.

4.4.1 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. 10 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. 18. 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.

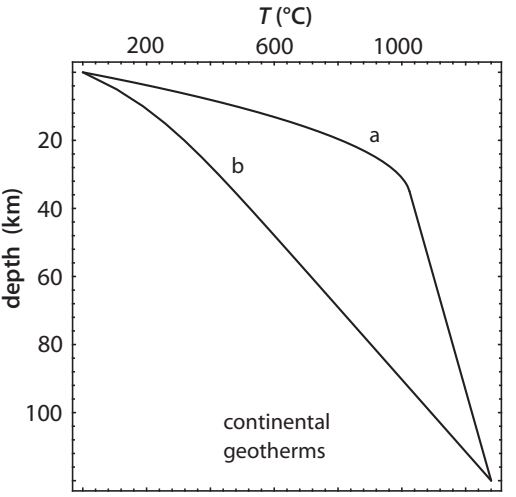


Figure 18: Examples of continental geotherms calculated with a lower boundary condition of a fixed temperature at the base of the lithosphere. Geotherm *a* was calculated assuming constant heat production in the crust and no heat production in the mantle lithosphere. Geotherm *b* was calculated for a continuous, exponentially decreasing heat production using eq. 21. The temperature T_l is assumed to be 1 280 °C.

4.4.2 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. 14 an elegant and simple description of stable continental geotherms. After two integrations we get:

$$kT = -h_\mathrm{r}^2 S_0 \mathrm{e}^{\left(-\frac{z}{h_\mathrm{r}}\right)} + C_1 z + C_2 \quad . \tag{18}$$

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. 18 goes to 1 so that C_2 must be:

$$C_2 = h_\mathrm{r}^2 S_0 \quad . \tag{19}$$

The lower boundary condition of $T = T_\mathrm{l}$ at the depth $z = z_\mathrm{l}$ can be evaluated by rearranging eq. 18:

$$C_1 = \frac{kT_\mathrm{l}}{z_\mathrm{l}} + \frac{h_\mathrm{r}^2 S_0 \mathrm{e}^{\left(-\frac{z_\mathrm{l}}{h_\mathrm{r}}\right)}}{z_\mathrm{l}} - \frac{C_2}{z_\mathrm{l}} \quad . \tag{20}$$

After inserting both constants into eq. 18 we get:

$$T = \frac{zT_\mathrm{l}}{z_\mathrm{l}} + \frac{h_\mathrm{r}^2 S_0}{k} \left(\left(1 - \mathrm{e}^{\left(-\frac{z}{h_\mathrm{r}}\right)}\right) - \left(1 - \mathrm{e}^{\left(-\frac{z_\mathrm{l}}{h_\mathrm{r}}\right)}\right) \frac{z}{z_\mathrm{l}} \right) \quad . \tag{21}$$

Curve *b* in Fig. 18 is an example of a geotherm calculated with this relationship. Eq. 21 provides a realistic and useful description of stable continental geotherms and has been presented and used by many authors.

4.4.3 More General Formulations

In order to use eq. 21 more efficiently, it is useful to introduce two new parameters: the vertical thickening (or thinning) strain of the crust f_c and that of the lithosphere f_l. A value of $f_\mathrm{c} = 2$ means that the crust is twice as thick as in the reference state. Using these parameters, eq. 21 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 get:

$$T = \frac{zT_\mathrm{l}}{f_\mathrm{l} z_\mathrm{l}} + \frac{f_\mathrm{c}^2 h_\mathrm{r}^2 S_0}{k} \left(\left(1 - \mathrm{e}^{\left(-\frac{z}{f_\mathrm{c} h_\mathrm{r}}\right)}\right) - \left(1 - \mathrm{e}^{\left(-\frac{z_\mathrm{l} f_\mathrm{l}}{f_\mathrm{c} h_\mathrm{r}}\right)}\right) \frac{z}{f_\mathrm{l} z_\mathrm{l}} \right) \quad . \tag{22}$$

This equation can be used to calculate the equilibrium temperatures at any depth for any thickness ratio of crust and mantle part of the lithosphere.

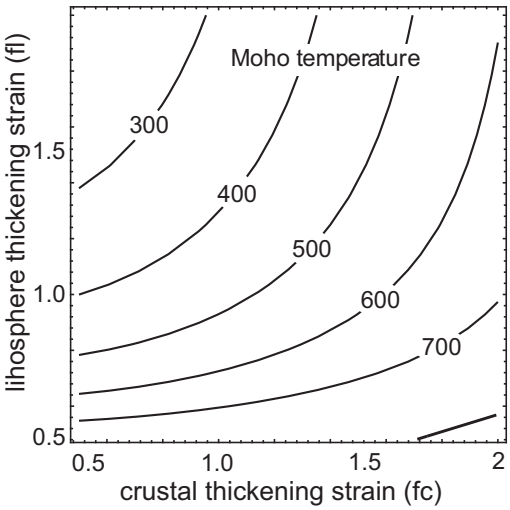


Figure 19: Moho-temperatures of continental lithosphere for different crustal thickening strains (expressed by f_c) and for different total thickening strains of the lithosphere (expressed by f_l). The diagram was calculated with eq. 22 assuming $z = z_c$. The assumption of the parameters are the same as in Fig. 18.

5 Unit: Heat in Oceanic Lithosphere

Oceanic lithosphere contains practically no radioactive elements. Thus, one could think that it is simple to describe stable oceanic geotherms. However, oceanic lithosphere is rarely old enough to be thermally stabilised. Geotherms in oceanic lithosphere are *transient*. The oldest oceanic lithosphere is about 150 my old (Fig. 20). However, we will soon show that the time for thermal equilibration of the lithosphere is of the order of 150 my or more! We can conclude that oceanic lithosphere is *not* thermally stabilized. As such, we must solve the time dependent diffusion equation (eq. 6). This is by far not as trivial as the simple integrations we have done for stable (time independent) geotherms in the last unit and we need to make a little excursion into the dealing with time dependent problems.

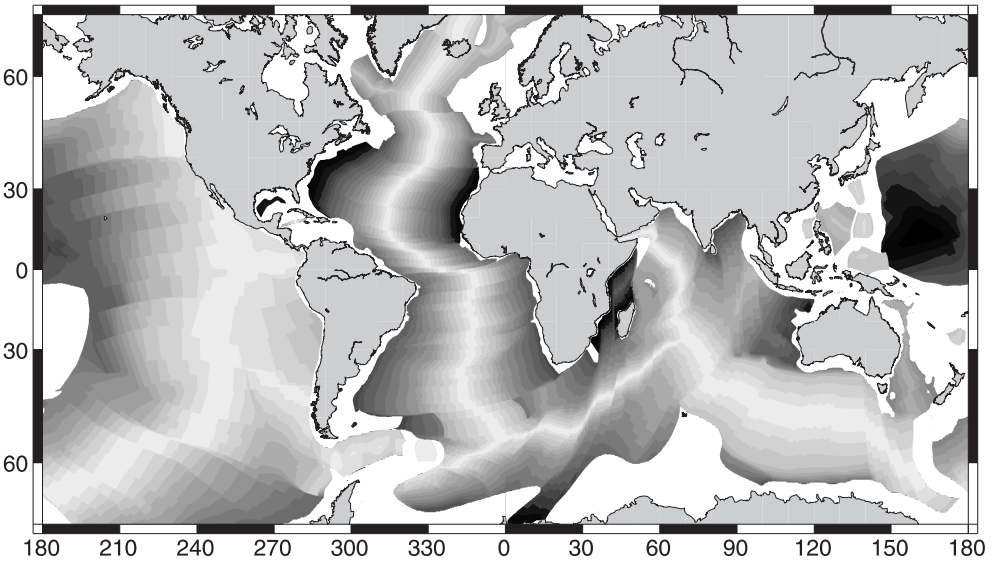


Figure 20: The age of the ocean floor (Müller et al. 1997). Shading intervals are every 10 my from 0 (white) to 160 my (black). Ocean floor older than 160 my is black. Areas with no data ar white. These regions are both on continental and on oceanic lithosphere. Landmasses are grey. Oldest known parts 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 a topography of the ocean floor and note the similarities.

5.1 Handling Time Dependent Problems

If we want to use eq. 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 (see unit 6). For most problems there are simply no solutions of eq. 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. 6 will contain a term of the form:

$$\frac{2}{\sqrt{\pi}} \int_0^n \mathrm{e}^{-n^2} \mathrm{d}n = \mathrm{erf}\,(n) \tag{23}$$

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. Fig. 21 shows the shape of the error function. In many solutions of eq. 6 the variable n from eq. 23 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. The *complementary* error function erfc is defined as:

$$\mathrm{erfc}\,(n) = 1 - \mathrm{erf}\,(n) \quad . \tag{24}$$

Solutions of the time dependent heat conduction equation very often contain error functions or complementary error functions of the term $\mathrm{erf}(l/\sqrt{4\kappa t})$, where l is a spatial coordinate.

5.1.1 Time Scales of Diffusion

We have seen already in the first unit that thermal equilibration is a process that is rapid at first (when the curvature of isotherms is still large) and then slows down more and more and the *complete* equilibrium is only reached after infinite time. This asymptotic form of the equilibration may also be seen in the shape of the error function on Fig. 21. Clearly, it is often useful to define some point in time when we call the equilibration to be "complete". In order to identify such a "time of complete equilibration", we use the fact that solutions of the heat flow equation often contain the term $\mathrm{erf}(l/\sqrt{4\kappa t})$. The shape of the error function in Fig. 21 shows that it reaches asymptotically 1 as n get very large. Correspondingly, the term $(l/\sqrt{4\kappa t})$ will reach

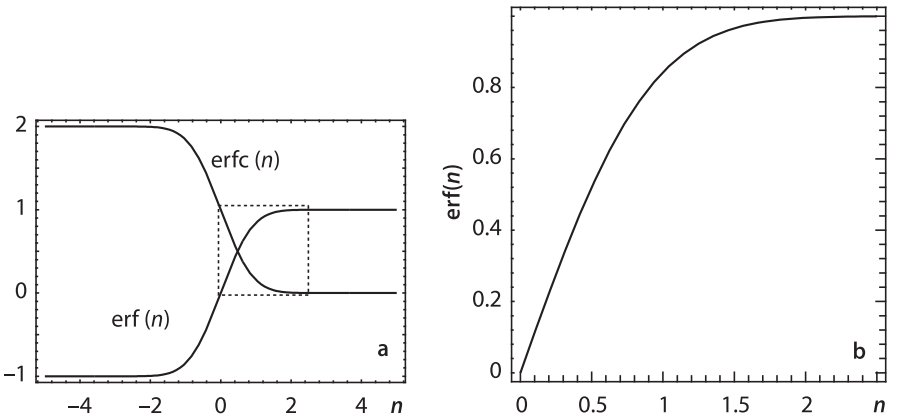


Figure 21: The error function and the complementary error function. The dashed frame in **a** shows the part of erf(n) that is shown enlarged in **b**.

1 for large l (regardless of t), or for 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. 23) is 1. This means that:

$$\left(\frac{l}{\sqrt{4\kappa t}}\right) = 1 \qquad \text{or :} \qquad t = \frac{l^2}{4\kappa} \quad . \tag{25}$$

Figure 21 illustrates that for the argument to be 1 (where $t = l^2/(4\kappa)$), the thermal equilibration is 84.3 % complete. 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". In fact, because this is a only a rough measure, the "4" is often left off this relationship and it is written:

$$t_{eq} = t = \frac{l^2}{\kappa} \quad . \tag{26}$$

The value of t_{eq} in eq. 26 is an important measure to estimate the duration of equilibration. It is often called *Thermal Time Constant* or: *Diffusive Time Scale of Equilibration.* In summary, the basic message of this section is:

– During conductive processes the duration of thermal equilibration increases with the square of the length scale of the equilibrating body and inverse proportionally with the diffusivity.

For diffusivities of the order of $\kappa = 10^{-6}\mathrm{m}^2\mathrm{s}^{-1}$ 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.

Table 3: Different values of the thermal time constant t_{eq} for a series of geologically relevant length scales l.

l	$\mathbf{t_{eq}} = l^2/2\kappa$	$\mathbf{t_{eq}} = l^2/\pi^2\kappa$
10 m	$5 \cdot 10^7$ s $\approx$ 1.58 y $\approx 10^0$ y	$1.01 \cdot 10^7$ s $\approx$ 16 weeks $\approx 10^{-1}$ y
100 m	$5 \cdot 10^9$ s $\approx$ 158 y $\approx 10^2$ y	$1.01 \cdot 10^9$ s $\approx$ 32 y $\approx 10^1$ y
1 km	$5 \cdot 10^{11}$ s $\approx$ 15 000 y $\approx 10^4$ y	$1.01 \cdot 10^{11}$ s $\approx$ 3 200 y $\approx 10^3$ y
10 km	$5 \cdot 10^{13}$ s $\approx$ 1.5 my $\approx 10^6$ y	$1.01 \cdot 10^{13}$ s $\approx$ 320 000 y $\approx 10^5$ y
100 km	$5 \cdot 10^{15}$ s $\approx$ 158 my $\approx 10^8$ y	$1.01 \cdot 10^{15}$ s $\approx$ 32 my $\approx 10^7$ y

5.2 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. 22). 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*.

5.2.1 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. 6, using a set of boundary and initial conditions. These conditions are provided by geological observation: The temperature at the surface $(T - s)$ of mid-oceanic ridges is that of the water temperature. For simplicity, we

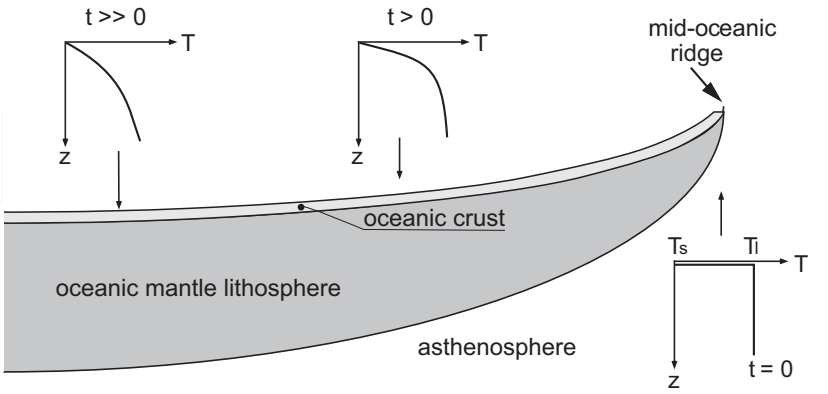


Figure 22: Thickness and thermal profile of oceanic lithosphere at a series of points..

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. 22. 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 $T = T_l$ (T_l is the temperature at the base of the lithosphere). 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. 22). The solution of the heat conduction equation for these boundary conditions is:

$$T = T_s + (T_l - T_s)\mathrm{erf}\left(\frac{z}{\sqrt{4\kappa t}}\right) \quad . \tag{27}$$

This solution is already a bit familiar to us from section 5.1. Fig. 23a shows temperature profiles through oceanic lithosphere of different ages, that were calculated with eq. 27. The curves correspond to the two sketches of thermal profiles in the middle and on the left of Fig. 22. Fig. 23b shows the depth of a series of isotherms as a function of age.

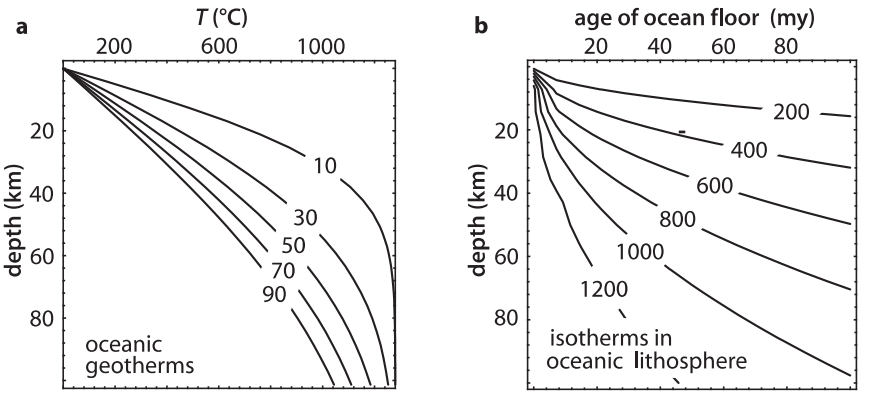


Figure 23: **a** Temperature T versus depth profiles through oceanic lithosphere at a number of different ages labeled in my. **b** The depth of isotherms (in °C) in oceanic lithosphere as a function of age between 0 and 100 my. The curves on both figures were calculated with eq. 27 assuming $T_\mathrm{l} = 1\,280\,^\circ\mathrm{C}$. 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.

5.2.2 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 to measure temperature in any representative way. 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 q_s (Pollack et al. 1993). The surface heat flow is the product of thermal conductivity and the thermal gradient at $z = 0$. This can be calculated from eq. 27 and can be compared with measured data in the oceans. To obtain surface heat flow we must differentiate eq. 27 with respect to depth and evaluate it at $z = 0$. From eq. 27 this is:

$$q_\mathrm{s} = k(T_\mathrm{l} - T_\mathrm{s}) \frac{\mathrm{d}\left(\mathrm{erf}\left(\frac{z}{\sqrt{4\kappa t}}\right)\right)}{\mathrm{d}z}_{,z=0} \quad . \tag{28}$$

As the error function itself is an integral (see eq. 23), it is easy to differentiate eq. 27 (sect. 5.1). We get:

$$q_\mathrm{s} = k(T_\mathrm{l} - T_\mathrm{s}) \sqrt{\frac{1}{\pi\kappa t}} \quad . \tag{29}$$

This equation can be rewritten 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

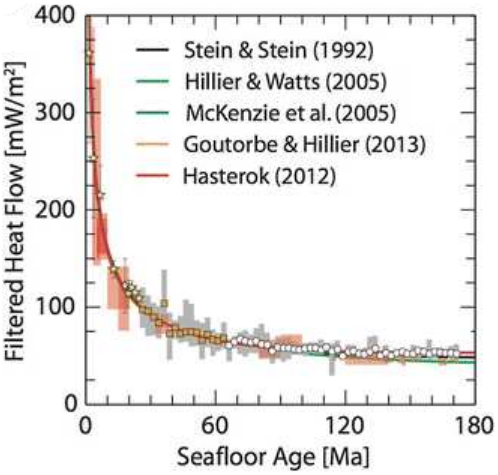


Figure 24: Measured surface heat flow of oceanic lithosphere (bars) as a function of age overlayed over curves calculated with the half space cooling model by different authors, for example eq. 30.

the distance from the mid-oceanic ridge and t is the age. If we replace t in eq. 29 by x/u, we get:

$$q_\mathrm{s} = k(T_\mathrm{l} - T_\mathrm{s})\sqrt{\frac{u}{\pi\kappa x}} \quad . \tag{30}$$

Eq. 30 shows us that the surface heat flow as a function of distance from the mid-oceanic ridge is a square root function of distance x from the ridge (as all other parameters in this equation are constants). Fig. 24 shows the surface heat flow in oceanic lithosphere as calculated with eq. 30. The heat flow data of Sclater et al. (1980) show that these curves correspond well with heat flow measured in the deep oceans. We will show later in this course 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. 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.

6 Unit: Production and Advection of Heat

We discern three fundamentally different geological mechanisms that produce heat:

– radioactive heat production,

– chemical heat production,

– mechanical heat production.

In the next sections we discuss the geological relevance of mechanical and chemical heat production. In general, the rate of temperature change due to heat production may be described by:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{S}{\rho c_p} \quad . \tag{31}$$

There, T, t, ρ and c_p correspond to temperature, time, density and heat capacity as discussed on p. 21 and S is the volumetric rate of heat production in $\mathrm{J\,s^{-1}\,m^{-3}} = \mathrm{W\,m^{-3}}$. 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 3.2. If S is positive, heat is produced, $\mathrm{d}T/\mathrm{d}t$ is positive and rocks heat up. If S is negative, heat is consumed, $\mathrm{d}T/\mathrm{d}t$ is negative and rocks cool. The heat production rate S can be of radioactive, chemical or mechanical origin so that we can write:

$$S = S_{\mathrm{rad}} + S_{\mathrm{chem}} + S_{\mathrm{mec}} \quad . \tag{32}$$

All three of these components may have a significant influence on the thermal evolution of rocks depending on the circumstances.

6.1 Radioactive Heat Production

Radioactive (or: radiogenic) heat is produced in the Earth predominantly by the naturally occurring radioactive isotopes ^{238}U, ^{235}U, ^{232}Th and ^{40}K. Of the two naturally occurring uranium isotopes 99.28% is ^{238}U and only 0.72% is ^{235}U. All of the naturally occurring thorium is ^{232}Th and only 0.0119% of the natural potassium is ^{40}K (Turcotte and Schubert 2002). As pure metals, these 4 isotopes produce the following amounts of heat: $^{238}\mathrm{U} = 9.46 \times 10^{-5}\mathrm{W\ kg^{-1}}$; $^{235}\mathrm{U} = 5.69 \times 10^{-4}\mathrm{W\ kg^{-1}}$; $^{232}\mathrm{Th} = 2.64 \times 10^{-5}\mathrm{W\ kg^{-1}}$ and $^{40}\mathrm{K} = 2.92 \times 10^{-5}\mathrm{W\ kg^{-1}}$. Fortunately, the concentrations of these elements in rocks are quite low so that substantially less heat is produced per cubic meter of rock. Table 4 lists some average concentrations of the heat producing

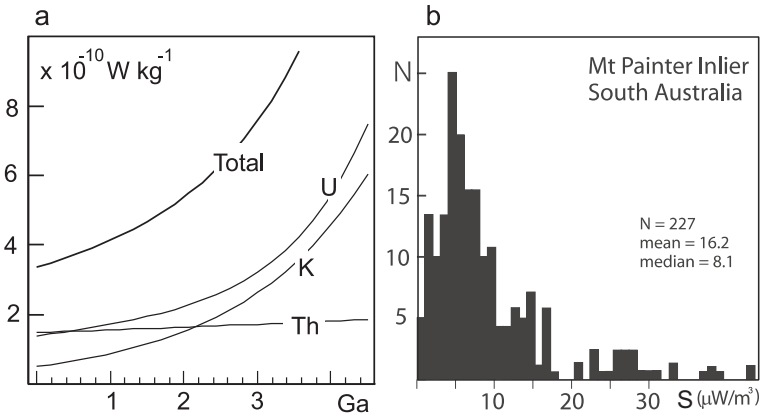


Figure 25: **a** Radioactive heat production in the crust through time. Note that the heat production in the Archaean 3 Ga ago was about twice as high as today. Also note that U and Th are the primary heat producing elements today, but it was U an K in the past. **b** Rates of radiogenic heat production 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.

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 3.4×10^{-10} W kg^{-1}, which corresponds roughly to a heat production rate of about one μW m^{-3} (see eq. 31). Using typical values for heat capacity and density of crustal rocks, $S = 1$ μW m^{-3} converts to a heating rate of of about 10°C per million years. 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 4 and there are many terrains around the world where radiogenic heat production rates is significantly higher than some μW m^{-3} (Fig. 25b; 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.

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

Table 4: 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.

6.2 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, 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*. We abbreviate this mechanical heat production with S_{mec}. The rate of mechanical heat production S_{mec} is given by the product of deviatoric stress τ and strain rate $\dot{\epsilon}$. Stress has the units of Pascal. One Pascal is one Joule per cubic meter ($1\ \mathrm{Pa} = 1\ \mathrm{J\,m^{-3}}$). 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:

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

The units of acceleration are $\mathrm{m\,s^{-2}}$ and those of force are therefore: $\mathrm{kg\,m\,s^{-2}}$. Stress and pressure therefore have the units of $\mathrm{kg\,m\,s^{-2}\,m^{-2}}$ or $\mathrm{Pa = kg\,m^{-1}\,s^{-2}}$ and energy has the units of $\mathrm{J = kg\,m^2\,s^{-2}}$. 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 (see e.g. Burg and Gerya 2005). However, if 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:

$$S_{\text{mec}} = \tau\dot{\epsilon} \quad . \tag{33}$$

In order to write the temperature change that arises from frictional heating we can write in analogy to eq. 31:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{\tau\dot{\epsilon}}{\rho c_p} \quad . \tag{34}$$

Note that eq. 34 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 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. $\mathrm{J\,s^{-1}\,m^{-2}}$ normal to the fault surface) which can be converted into a heating rate using the laws of heat conduction discussed in previous sections.

6.2.1 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. 34 to estimate the temperature increase a rock might experience for some realistic deviatoric stresses and strain rates.

Methods to measure geological strain rates show an upper limit of $\dot{\epsilon} = 10^{-12}$ to 10^{-14} 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 $\dot{\epsilon} = 10^{-6}$ 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} $\mathrm{N\,m^{-1}}$ and this implies a rock strength of 50–100 MPa, averaged over the thickness of the lithosphere. This estimate comes out of the strength that rocks need to support the weigth plateaus like the Tibetan plateau or the Altiplano and/or out of arguments about the Brace-Goetze lithosphere, that we discussed in the first unit of this course. (Note however, that we abbreviated deviatoric stresses with σ instead of τ there).

We could thus estimate shear heating as follows: If stress and strain rate remain constant over time, then integrating of eq. 34 gives:

$$T = t \times \frac{\tau\dot{\epsilon}}{\rho c_p} \quad . \tag{35}$$

The temperature at the end of deformation for a longitudinal strain of 1 is then: $T = \tau/(\rho c_p)$ ($\dot{\epsilon} \cdot t = \epsilon = 1$). Using standard values for the density and specific heat ($\rho = 2\,700$ $\mathrm{kg\,m^{-3}}$ and $c_p = 1\,000$ $\mathrm{J\,kg^{-1}\,K^{-1}}$) we can see that a rock that has a shear strength of 100 MPa will be heated by about 37 °C. We conclude that shear heating may be of significant importance to the thermal energy budget of the lithosphere.

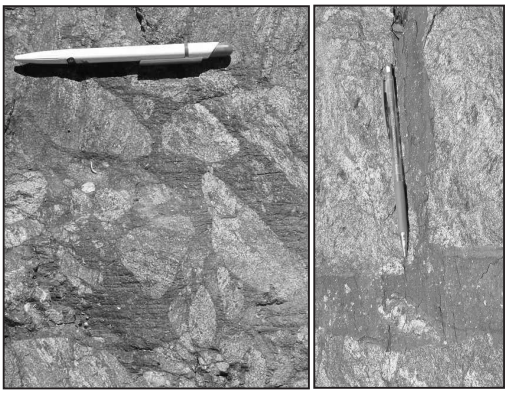


Figure 26: Visible evidence for frictional heating. The left photo shows network typical of many pseudotachylites, the right photo shows a pseudotachylite with chilled margin and recrystallised center.

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. 26). 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. For example, if a 100 m thick shear zone is active for 1 my, then the characteristic time scale of diffusion of this shear zone is of the order of only 1 000 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.

6.3 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_{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 % H_2O and this water is being released over quite a large temperature interval. Thus, the heat of reaction is fairly insignificant during regional metamorphism (about $5\text{–}10 \cdot 10^{-14}$ W 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 $L = 320\,000\ \mathrm{J\,kg^{-1}}$. Evaporation and condensation reactions are also strongly exothermic and endothermic respectively, but they are not very important in the geodynamics of the lithosphere.

6.3.1 Quantitative Description of Chemical Heat Production

The rate of reaction heat production S_{chem} has the same units as any other heat production rate: $\mathrm{W\,m^{-3}}$. It can be described by:

$$S_{\mathrm{chem}} = L\rho\frac{\mathrm{d}V}{\mathrm{d}t} \quad . \tag{36}$$

In this equation L is the latent heat of reaction in $\mathrm{J\,kg^{-1}}$. Since we think of the chemical heat production rate as a *volumetric* heat production rate, it is necessary to multiply L by the density ρ to convert it into a volumetric heat content. The expression $\mathrm{d}V/\mathrm{d}t$ is the volumetric proportion of the reaction that occurs per unit time (in $\mathrm{s^{-1}}$). 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. 36 into eq. 31 we can now formulate the temperature change during chemical heat production (and neglecting conduction) to be:

$$\frac{\partial T}{\partial t} = \frac{L}{c_p}\frac{\partial V}{\partial t} \quad . \tag{37}$$

Clearly, we could add a diffusion term to this equation to simultaneously consider chemial heat production and conductive distribution of this heat. However, as most up-temperature reactions are endothermic and most down temperature reactions exothermic, actual heating or cooling by chemical heat production rarely occurs. Instead, this chemical heat is more responsible for buffering the temperature increase or decrease.

6.3.2 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 although an estimate of the time of thermal equilibration would long have indicated that all ice must have melted: ice and water will both have a temperature of 0 °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 °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.

Let us illustrate the buffering effects that occur during cooling and crystallisation of a eutectic melts (i.e. all melt crystallises at the solidus T_s). If the crystallisation puts out $L = 320\,000$ J kg^{-1} and you need about 1000 J per Kilogram to heat a rock by one degree ($c_p = 1000$J kg^{-1}°C^{-1}) then clearly the buffering is equivalent to a temperature rise of:

$$\Delta T = L/c_p \tag{38}$$

In other words, if this melt cools with a constant cooling rate s of s =10°C per million years ($s = 10/(3.15 \times 10^{13})$s then the time of buffering is simply:

$$t = LV/sc_p \tag{39}$$

where we have inserted V now again as the volume fraction that actually crystallises (you can leve this out too). For a typical value of $L = 300000$ J kg^{-1} this gives $\approx 10^{14}$ s or about 3 million years. After this time interval is over, all melt is crystallises and the rock continues to cool at rate s. (Be careful not to confuse s (cooling rate) with s (SI unit seconds) here). In other words, it takes about three times as long to melt snow at zero degrees temperature to water at zero degrees temperature, than it takes to heat zero degree water to boiling.

6.4 Advection of Heat in the Lithosphere

Heat can be transported *actively* by the motion of warm rocks. Besides *conduction* and *production*, this is the thirs possible heat transfer meachanism. 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:

$$\frac{\partial T}{\partial t} = u\frac{\partial T}{\partial z} \quad . \tag{40}$$

In eq. 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. 40 describes transport *against* the spatial coordinate z: transport is from high z towards lower z. Eq. 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. 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.

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 and we have dealt with this in Unit 4.

6.5 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. 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. 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. 40 by the diffusion term from eq. 6. The equation that must be solved becomes:

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} + u \frac{\partial T}{\partial z} \quad . \tag{41}$$

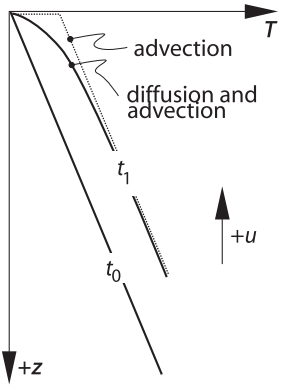


Figure 27: Schematic illustration of one-dimensional advection of heat by erosion. The coordinate system is fixed with $z = 0$ at the earth's surface. Temperature profiles through the crust are shown for two times: at the onset of erosion t_0, at which a linear geotherm is assumed and a later time t_1. The advection rate u is positive upwards. In the shown time interval the erosion process *advects* the geotherm by $u \times t_1$ meters upwards. Simultaneous *diffusion* causes the curvature of the temperature profile.

A schematic illustration how the two processes interact to shape a geotherm during erosion is shown in Fig. 27. You can see that diffusion and advection interact in shaping the temperature profile until a certain depth length scale), but that advection dominates at larger depths (length scales).

7 Unit: The Elevation of Mountains

Isostasy is a stress balance. 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. 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. 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. 28). 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.

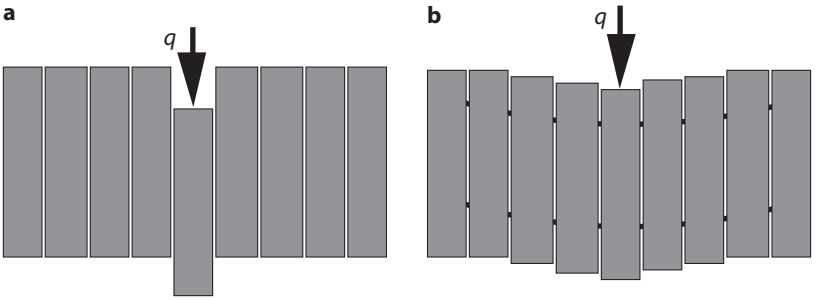


Figure 28: Illustration of the difference between **a** hydrostatic isostasy and **b** flexural isostasy. In **a** all vertical columns are considered independently of each other. In **b** the shear stresses between vertical columns are akso considered. q is the load.

7.1 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. 28a). 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. 29):

$$\sigma_{zz}^{\mathrm{A}}|_{z=z_{\mathrm{K}}} = \sigma_{zz}^{\mathrm{B}}|_{z=z_{\mathrm{K}}} \quad . \tag{42}$$

In this equation σ_{zz}^{A} and σ_{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 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:

$$\sigma_{zz}\Big|_{z=z_{\mathrm{K}}} = \int_0^{z_{\mathrm{K}}} \rho g \mathrm{d}z \quad . \tag{43}$$

Inserting eq. 43 into eq. 42 gives:

$$\int_0^{z_{\mathrm{K}}} \rho_{\mathrm{A}}(z) g \mathrm{d}z = \int_0^{z_{\mathrm{K}}} \rho_{\mathrm{B}}(z) g \mathrm{d}z \quad , \tag{44}$$

where $\rho_{\mathrm{A}}(z)$ and $\rho_{\mathrm{B}}(z)$ 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. 29, 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. 44 is the basis of all calculations of isostasy. When considering the isostatically supported elevation of a mountain belt, it is useful to divide the density variations in the lithosphere into two parts (a) density variations that are due to material differences and (b) density variations that are caused by thermal expansion.

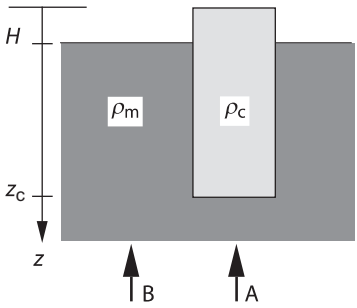


Figure 29: Illustration of isostatic equilibrium. Note that the z-axis is defined positively downwards and has its origin at the surface of the light shaded block (e.g. an iceberg or the lithosphere) that is assumed to float in a dark shaded region of higher density (e.g. water or the asthenosphere).

7.1.1 Isostasy due to Material Differences

We begin by solving the balance written out in eq. 44 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 $z_\mathrm{K} = z_\mathrm{c}$ (Fig. 29). The block in Fig. 29 has a constant density ρ_c (e.g. density of the crust) and floats in a denser medium of the constant density ρ_m (e.g. density of the mantle). We call its elevation above the surface of the denser medium H_mat, although it is just labeled as H in Fig. 29. We use the subscript $_\mathrm{mat}$ 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. 44 and integration is easy. By integrating the left half of the equation and splitting up the right half of eq. 44 we get according to Fig. 29:

$$\rho_\mathrm{c} g z \Big|_0^{z_\mathrm{c}} = g \int_0^{H_\mathrm{mat}} \rho_\mathrm{air} \mathrm{d}z + g \int_{H_\mathrm{mat}}^{z_\mathrm{c}} \rho_\mathrm{m} \mathrm{d}z \quad . \tag{45}$$

The density of air is negligible in comparison with ρ_m or ρ_c. Thus, the first integral on the right hand side of eq. 45 is also negligible. After finishing the integration, canceling out g and inserting the integration limits we get:

$$\rho_\mathrm{c} z_\mathrm{c} = \rho_\mathrm{m} z_\mathrm{c} - \rho_\mathrm{m} H_\mathrm{mat} \quad . \tag{46}$$

Solving for elevation H gives:

$$H = H_\mathrm{mat} = z_\mathrm{c} \left(\frac{\rho_\mathrm{m} - \rho_\mathrm{c}}{\rho_\mathrm{m}} \right) \quad . \tag{47}$$

This relationship describes the hydrostatically balanced elevation of the surface of a floating body above the medium it floats in. Remember that $H = H_\mathrm{mat}$ emphasizes

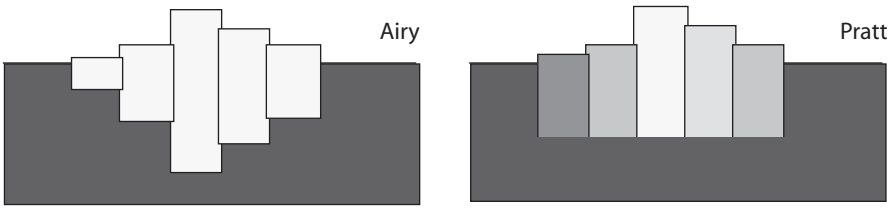


Figure 30: Comparison of the interpretations of the isostatic model according to Airy and Pratt. The shading indicates density. Darker shading means higher density.

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 ρ_c is zero, then this equation states that $H = z_1$: 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.

7.1.2 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. 30). 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. 47. 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. Seismic studies in many mountain belts show that most regions of high surface elevation are indeed compensated by significant roots at depth.

7.1.3 Isostasy Due to Thermal Expansion

In order to calculate the contribution of thermal expansion to surface elevation we need to introduce α: the *coefficient of thermal expansion.* α has the units of strain per temperature increment, which is K^{-1}. For most rocks the coefficient of thermal expansion is of the order $\alpha = 3 \cdot 10^{-5}\,\mathrm{K}^{-1}$. Using α and the density of the mantle ρ_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:

$$\rho(T) = \rho_\mathrm{m}(1 + \alpha(T_\mathrm{l} - T)) \quad . \tag{48}$$

There, T_l is the temperature at the base of the lithosphere at $z = z_\mathrm{l}$. According to eq. 48: $\rho = \rho_\mathrm{m}$, where $T = T_\mathrm{l}$. At lower temperatures, the density increases linearly. At the surface, where we can assume that the temperature is T_s =0 °C, eq. 48 becomes:

$$\rho_{(T=T_\mathrm{s})} = \rho_0 = \rho_\mathrm{m}\,(1 + \alpha T_\mathrm{l}) \quad . \tag{49}$$

If the density of the mantle is about ρ_m=3 200 $\mathrm{kg\,m^{-3}}$ at T_l, then the density at the surface is: ρ_0=3 300 $\mathrm{kg\,m^{-3}}$. Assuming a linear geotherm in the lithosphere, we can describe the mean density of the lithosphere with:

$$\overline{\rho} = \rho_\mathrm{m}\left(1 + \alpha\frac{T_\mathrm{l} + T_\mathrm{s}}{2}\right) \quad . \tag{50}$$

In order to estimate which proportion of the elevation of a mountain belt is due to thermal expansion (H_therm), we insert eq. 50 into the left hand side of eq. 44. The following algebra remains the same as in eq. 45 and eq. 46 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:

$$H_\mathrm{therm} = -z_\mathrm{l}\alpha(T_\mathrm{l} + T_\mathrm{s})/2 \quad . \tag{51}$$

The negative sign arises because $\overline{\rho}$ is larger than ρ_m.

7.1.4 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).

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. 51). The *material* contribution of the crust to the elevation, on the other hand, is positive and was derived in eq. 47. Density variations within the mantle part of the lithosphere are neglected here. Then, the isostatically supported surface elevation relative to the surroundings is given be the sum of the thermal and the material contributions:

$$H = H_\mathrm{mat} + H_\mathrm{therm} = z_\mathrm{c} \left(\frac{\rho_\mathrm{m} - \rho_\mathrm{c}}{\rho_\mathrm{m}} \right) - z_\mathrm{l} \alpha (T_\mathrm{l} + T_\mathrm{s})/2 \quad . \tag{52}$$

If we summarize all the material parameters into the constants:

$$\delta = (\rho_\mathrm{m} - \rho_\mathrm{c})/\rho_\mathrm{m} \qquad \text{and :} \qquad \xi = \alpha (T_\mathrm{l} + T_\mathrm{s})/2 \quad , \tag{53}$$

then this eq. 52 simplifies to:

$$H = \delta z_\mathrm{c} - \xi z_\mathrm{l} \quad . \tag{54}$$

If we insert meaningful numbers into eq. 52 (e. g. $\rho_\mathrm{m} = 3\,200\ \mathrm{kg\,m^{-3}}$, $\rho_\mathrm{c} = 2\,700\ \mathrm{kg\,m^{-3}}$), we get:

$$\delta \approx 0.15 \qquad \text{and :} \qquad \xi \approx 0.018 \quad . \tag{55}$$

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 3 600 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. 29. 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 3 600 m below

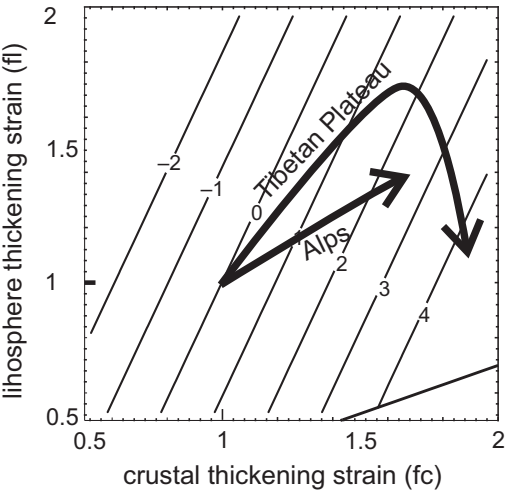


Figure 31: Isostatically supported surface elevation of mountain belts in the f_c-f_l-plane (in km with eq. 56). Following assumptions were used: $\rho_m = 3\,200$, $\rho_c = 2\,750$, $\alpha = 3 \cdot 10^{-5}$, $z_c = 35$ km, $z_l = 100$ km. Using these values, the two constants are: $\delta \approx 0.14$ and $\xi = 0.018$. Typical orogenic evolutions are superposed.

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. 52, so that the elevation is given as the elevation difference between a thickened (or thinned) lithosphere and an undeformed reference lithosphere:

$$H = (\delta f_c z_c - \xi f_l z_l) - (\delta z_c - \xi z_l) = \delta z_c (f_c - 1) - \xi z_l (f_l - 1) \quad . \tag{56}$$

The parameters f_c and f_l describe the thickening strains of the crust and the mantle lithosphere. The elevation of isostatically supported mountain belts above the undeformed reference lithosphere is shown in Fig. 31 (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. 31 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. 54 and eq. 56 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.

8 Unit: The Depth of the Oceans

The water depth of the oceans In isostatic equilibrium) is a direct function of the distance to the mid-oceanic ridges. 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 and we have discussed it already in unit 3.

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 this density variation to estimate the isostatically supported elevation of the ocean floor, we use the model sketched in Fig. 32. According to eq. 44 the vertical normal stresses of the columns A and B must be the same in the compensation depth $z = z_\mathrm{l}$. For column A the vertical normal stress at depth $z = z_\mathrm{l}$ is given by:

$$\sigma^{\mathrm{A}}_{zz}|_{z=z_\mathrm{l}} = \rho_\mathrm{w} g w + \int_0^{z_\mathrm{l}} \rho_{(z)} g \mathrm{d}z \quad . \tag{57}$$

There, w is the water depth in column A, ρ_w is the water density, g is the gravitational acceleration and $\rho_{(z)}$ is the density of the lithosphere as a function of depth. For column B we can formulate:

$$\sigma^{\mathrm{B}}_{zz}|_{z=z_\mathrm{l}} = \rho_\mathrm{w} g w + \rho_\mathrm{m} g z_\mathrm{l} \quad . \tag{58}$$

(see Fig. 32). After inserting eqs. 57 and 58 into eq. 44, the isostasy condition of gets the following form:

$$\rho_\mathrm{m} z_\mathrm{l} + w(\rho_\mathrm{m} - \rho_\mathrm{w}) = \int_0^{z_\mathrm{l}} \rho_{(z)} \mathrm{d}z \quad . \tag{59}$$

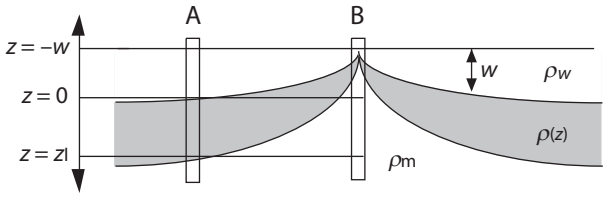


Figure 32: Schematic profile through a mid-oceanic ridge and the oceanic lithosphere as used for the calculation of water depth. The oceanic crust is neglected because it is everywhere of the same thickness.

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. 59 gets the form:

$$w(\rho_\mathrm{m} - \rho_\mathrm{w}) = \int_0^{z_\mathrm{l}} (\rho_{(z)} - \rho_\mathrm{m})\mathrm{d}z \quad . \tag{60}$$

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. Thus, if we know the temperature as a function of depth, then $\rho_{(z)}$ in eq. 60 is known, because we know already the relationship between density and temperature from eq. 48. Thus we can begin by inserting eq. 48 into eq. 60:

$$w(\rho_\mathrm{m} - \rho_\mathrm{w}) = \int_0^{z_\mathrm{l}} \rho_\mathrm{m}\alpha(T_\mathrm{l} - T(z))\mathrm{d}z \quad . \tag{61}$$

The variable $T(z)$ is the only unknown in this equation, but it is well-described by the half-space cooling model and we determined it in sect. 5.2. . Thus, the temperature profile of eq. 27 may be directly inserted into eq. 61. We get:

$$w(\rho_\mathrm{m} - \rho_\mathrm{w}) = \int_0^{z_\mathrm{l}} \rho_\mathrm{m}\alpha(T_\mathrm{l} - T_\mathrm{s})\mathrm{erfc}\left(\frac{z}{\sqrt{4\kappa t}}\right)\mathrm{d}z \tag{62}$$

or, after taking the constants out of the integral and solving for w:

$$w = \frac{\rho_\mathrm{m}\alpha(T_\mathrm{l} - T_\mathrm{s})}{(\rho_\mathrm{m} - \rho_\mathrm{w})} \int_0^{z_\mathrm{l}} \mathrm{erfc}\left(\frac{z}{\sqrt{4\kappa t}}\right)\mathrm{d}z \quad . \tag{63}$$

This is not too difficult to solve and results in:

$$w = \frac{2\rho_\mathrm{m}\alpha(T_\mathrm{l} - T_\mathrm{s})}{(\rho_\mathrm{m} - \rho_\mathrm{w})}\sqrt{\frac{\kappa t}{\pi}} \quad . \tag{64}$$

If we insert standard values for all the constants in this equation we get:

$$w \approx 5.91 \cdot 10^{-5}\sqrt{t} \quad . \tag{65}$$

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. 32). 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. 33 shows some water depth profiles calculated with this equation. The fantastic coincidence of these curves with bathymetric measurements in the oceans of the world confirm the model.

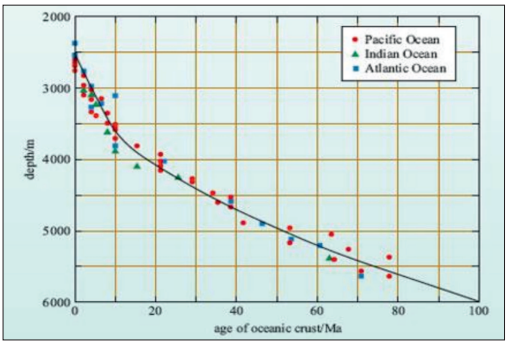


Figure 33: Profiles of water depth as a function of distance from the mid-oceanic ridge for three major oceans of this planet in comparison with water depth as calculated with eq. 64.

8.1 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. 28b). 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.

8.1.1 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. 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. They were created by *hot spots* that have their origin deep inside the mantle (Fig. 34). 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 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 know 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

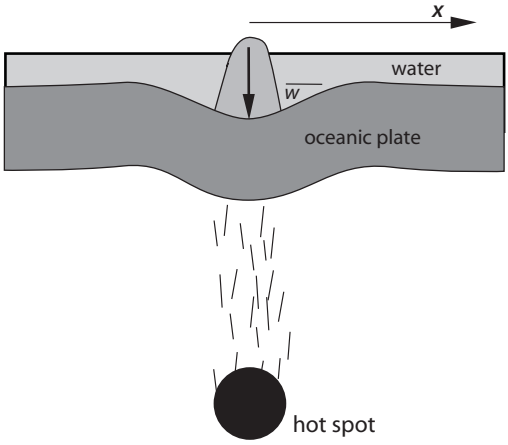


Figure 34: Flexure of oceanic lithosphere due to the loading of a sea mount.

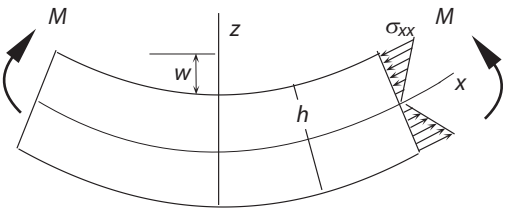


Figure 35: Bending of an ideal elastic plate in a simplified model view which is useful for the description of bending lithospheric plates.

Donau. However, in collisional orogens the *external* load applied by the weight of the mountain belt is partly compensated by an *internal* loads: the root of the mountain belt.

Passive continental margins also show often evidence for elastic bending of continental lithosphere. 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).

8.1.2 The Flexure Equation

Elastic deformation describes an empirically derived *constitutive relationship* in which stress and strain are proportional to each other. 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* ν.

Let us now consider the bending of a simple, ideal elastic plate like the one sketched in Fig. 35. We also neglect buoyancy forces for now. When integrating the horizontal normal stresses σ_{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. 35):

$$M = -D\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} \quad . \tag{66}$$

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.

Eq. 66 may be coupled with a force balance equation that relates bending moments to the vertical load q (s. Fig. 35) (s. Turcotte and Schubert 1982; Ranalli 1987). This

is called the one-dimensional flexure equation and (neglecting horizontal forces) is:

$$D\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} = q_x \quad . \tag{67}$$

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. 67 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:

$$D = \frac{Eh^3}{12(1-\nu^2)} \quad . \tag{68}$$

Thus, if the material constants E and ν are known and the flexural rigidity of a plate was derived from modeling its shape using eq. 67, then this may be converted directly into an elastic thickness of the lithosphere using eq. 68. Indeed, we only showed the last equation to illustrate the definition of the elastic thickness of the lithosphere. All descriptions of the bending of elastic plates are based on the integration of eq. 67, or its two-dimensional equivalent.

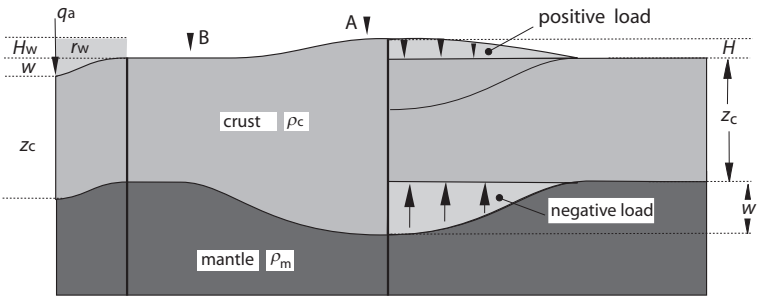


Figure 36: Distribution of loads during the elastic bending of lithospheric plates.

8.1.3 Application to the Lithosphere

Eq. 67 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 some important points:

1. 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. 68 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 is dominated by elastic behavior.

2. The distribution of loads on the plate must be thought through. The load as a function of distance q_x as used in eq. 67 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. 36. There, it may be seen that the downward force exerted by the mountain range on the plate is given by the vertical normal stress $q_{\text{ext}} = \rho_{\text{c}} g H$. 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. 36 with the upwards arrows. This internal load has the magnitude $q_{\text{int}} = (\rho_{\text{m}} - \rho_{\text{c}}) g w$, where w is the deflection of the plate. The net load that is applied to the plate is therefore:

$$q(x) = q_{\text{ext}} - q_{\text{int}} = \rho_{\text{c}} g H(x) - (\rho_{\text{m}} - \rho_{\text{c}}) g w \quad . \tag{69}$$

Note that the load is here already expressed as a function of horizontal distance x. If eq. 69 is inserted into eq. 67, this may be solved for w numerically or – for some simple boundary conditions - also analytically.

8.1.4 Applications to the Oceanic Lithosphere

A series of elastic bending problem in the oceanic lithosphere may be well-described with eq. 67 if two simplifying assumptions are made:

– 1. We assume that there are no horizontal forces applied to the plate.

– 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.

Based on the second assumption, and assuming that the downwards deflected region is filled with water, eq. 69 simplifies to:

$$q = q_{\text{a}} - (\rho_{\text{m}} - \rho_{\text{w}}) g w \quad . \tag{70}$$

as illustrated on the very left hand edge of Fig. 36 (ρ_w is the water density). Eq. 67 simplifies to:

$$D\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} = -(\rho_m - \rho_w)gw \quad . \tag{71}$$

Eq. 71 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, ρ_m and ρ_c often occur in the following relationship (which we do not need to remember):

$$\alpha = \left(\frac{4D}{g(\rho_m - \rho_w)}\right)^{1/4} \quad . \tag{72}$$

α is called the *flexure parameter* of the lithosphere and is not the be confused with the coefficient of thermal expansion, which we also abbreviated with α earlier on in theis script.

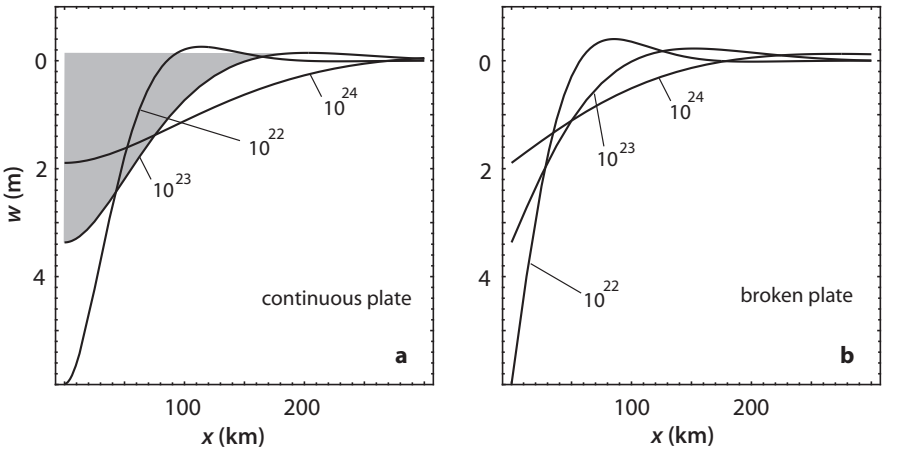


Figure 37: Shape of elastically bent plates. **a** Continuous plate loaded only at $x = 0$: the left margin of the diagram (eq. 73). Only half of the plate is shown. **b** Broken plate, also loaded only at $x = 0$ (eq. 75). The curves are labeled with the flexural rigidity of the plates in Nm.

• **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 formu-

lated boundary and initial conditions (e.g. the load applies only at $x = 0$, symmetry of the deflection so that $\mathrm{d}w/\mathrm{d}x{=}0$ at $x = 0$ and others) a solution of eq. 71 is:

$$w = w_0 \mathrm{e}^{-x/\alpha} \left(\cos(x/\alpha) + \sin(x/\alpha)\right) \quad . \tag{73}$$

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. 73 that $w \to w_0$ for $x \to 0$). Interestingly, the maximum deflection w_0 is given by:

$$w_0 = \frac{q\alpha^3}{8D} \quad . \tag{74}$$

Eq. 73 is a good approximation for the description of the water depth around the Hawaii and Emperor Island chains (s. Fig. 37a). 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. 71 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. 71 is:

$$w = w_0 \mathrm{e}^{-x/\alpha} \left(\cos(x/\alpha)\right) \quad . \tag{75}$$

The shape of plates as described by eq. 75 is illustrated in Fig. 37b. A comparison of the curves shown on Fig. 37b 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. 37b. 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.

9 Unit: Plate Driving Forces: Potential Energy

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.

On a plate tectonic scale, shear stresses clearly play a role in accretionary prisms. There, deformation occurs because of traction along the upper side of the lower plate. However, on a lithospheric scale shear stresses at both, the upper and the lower side of the plate are largely negligible. Just like the west wind has not enough traction at the surface of earth to move continents to the east, so does the mantle convection not have enough traction on the underside to move plates. At mid ocean ridges we have a series of arguments that it is NOT convection that moves the plates apart:

- The geometry of convection cells (mushroom shape) is impossible to reconcile with the geometry of the mid-ocean ridges on a global scale.
- By definition the lithosphere involves only those parts of the outer earth where forces can be transmitted.
- We know that dynamic topography produced by plumes creates up to 1 km of topography, whereas mid-ocean ridges are all at the same water depth, except in Iceland, where we know of a plume.
- From experimental petrology we know that decompression melts that are created during 600 km of decompression (the vertical scale of plumes) produce the equivalent of 20 km thick oceanic crust. In contrast, we know that oceanic crust is only 5-7 km - consistent with small scale convection on the lithospheric scale.

Sooo - if it is not mantle convection that drives plate motions, what is it then?

9.0.1 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 9.1 and 9.2.

In sect. 7.1 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 ρg between 0 and z, as we did in eq. 43. If the density over the thickness z remains constant, then this is simply $\rho g z$. This term has the units of Pa or $\mathrm{kg\,s^{-2}\,m^{-1}}$ or $\mathrm{J\,m^{-3}}$. 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:

$$E_\mathrm{p} = \int_0^z \sigma_{zz}\mathrm{d}z = \int_0^z \int_0^z \rho_{(z)} g \mathrm{d}z\mathrm{d}z \quad . \tag{76}$$

Very often the "depth of interest" is the isostatic compensation depth. If the density is independent of depth, then eq. 76 may be simplified to give:

$$E_\mathrm{p} = \int_0^z \sigma_{zz}\mathrm{d}z = \int_0^z \rho g z \mathrm{d}z = \frac{\rho g z^2}{2} \quad . \tag{77}$$

This integral corresponds to the gray shaded region in Fig. 38b. We want to remember that E_p has the units of energy *per* area and is, therefore, strictly speaking, no energy as such.

9.0.2 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. 38). It is true that:

$$\sigma_{zz} = \sigma_{xx} = \sigma_{yy} \quad . \tag{78}$$

The sum of all vertical stresses integrated over the 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

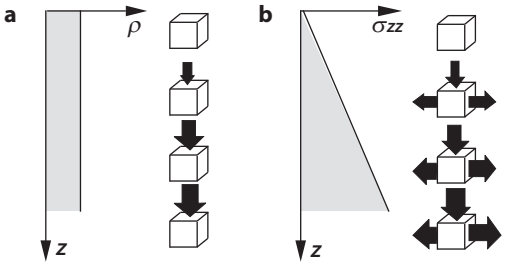


Figure 38: Density ρ and vertical normal stress σ_{zz} as a function of depth z. The value σ_{zz} is the vertically integrated density, times acceleration. Thus, the curve in **b** corresponds to the gray shaded region in **a**. The row of little unity cubes next to **a** illustrates how the vertical 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**.

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. 39):

$$\Delta E_\mathrm{p} = F_\mathrm{b} = \int_0^{z_\mathrm{K}} \int_0^{z_\mathrm{K}} \rho^\mathrm{A}(z) g \mathrm{d}z\mathrm{d}z - \int_0^{z_\mathrm{K}} \int_0^{z_\mathrm{K}} \rho^\mathrm{B}(z) g \mathrm{d}z\mathrm{d}z \quad . \tag{79}$$

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. 50. Below this depth there is no density differences between the vertical columns A and B (s. eq. 42). $\rho^\mathrm{A}(z)$ is the density of profile A as a function of depth z.

If density is a continuous function of depth, then eq. 79 may be usually integrated without too much trouble. However, in the lithosphere, the density distribution has (a the least) a discontinuity at the Moho so that it may be necessary to split the integral in eq. 79, even for very simple assumptions on the density distribution in the lithosphere.

The importance of the density *distribution* in the lithosphere for the potential energy may be illustrated nicely with an interesting example. Fig. 39 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. 39. This net force is exerted by column B towards column A.

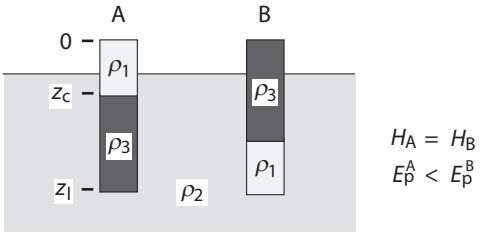


Figure 39: Schematic cartoon showing two columns in isostatic equilibrium ($\rho_1 < \rho_2 < \rho_3$). The surface of both columns has the same elevation above the liquid of density ρ_2, because both bodies consist of equally thick sections of the densities ρ_1 and ρ_3, i.e. they have the same weight. 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.

9.0.3 Force Balance Between Mountains and Foreland

In this section we estimate the forces exerted by a mountain range onto its foreland because of their potential energy difference (Fig. 40). 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. 40b. This choice for the vertical axis helps the intuitive understanding if the integration of eq. 79, as one of the integration limits is always zero. 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 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) and the geometry shown in Fig. 40. We can find this by integrating eq. 76. For the undeformed lithosphere in the foreland the potential energy above the Moho is simply:

$$E_{\mathrm{p}}^{\mathrm{foreland}} = \rho_{\mathrm{c}} g z_{\mathrm{c}}^2 / 2 \quad . \tag{80}$$

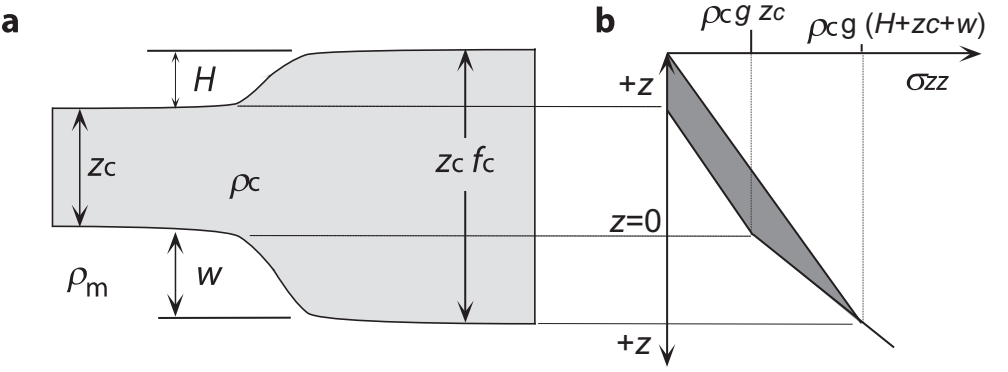


Figure 40: Cartoon contrasting the distribution of vertical stresses in mountain ranges relative to their foreland. **a** The thickness of the crustal root is w, the surface elevation relative to the reference lithosphere in the foreland H. In isostatic equilibrium it is true that: $H\rho_c = w(\rho_m - \rho_c) = w\Delta\rho$. 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.

Correspondingly, the potential energy of the thickened crust relative to the Moho is:

$$E_p^{\text{range}} = \rho_c g(H + z_c)^2/2 + \Delta\rho g w^2/2 \quad . \tag{81}$$

where $\Delta\rho = (\rho_m - \rho_c)$ and the thicknesses H, Z_c and w are as labeled on Fig. 40. 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. The potential energy difference per unit area is given by the difference of eq. 80 and eq. 81 (s. eq. 79). It is:

$$\begin{aligned} \Delta E_p &= F_b = E_p^{\text{range}} - E_p^{\text{foreland}} \\ &= \rho_c g H^2/2 + \rho_c g H z_c + \Delta\rho g w^2/2 \quad . \end{aligned} \tag{82}$$

Eq. 82 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. 82 to:

$$\Delta E_p = F_b = \rho_c g H\,(H/2 + z_c + w/2) \quad . \tag{83}$$

The force F_b corresponds to the dark shaded region in Fig. 40b. It is the difference between the vertically integrated vertical stresses σ_{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. 83 gives a force F_b of the order of $3\text{–}4 \cdot 10^{12}$ 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. 82 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. 82 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). This is the reason why mountain ranges do not grow infinitely on this planet and have a limiting elevation.

9.1 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:

9.1.1 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, we will neglect here the curvature of 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. 7a.

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. 79, using similar assumptions to those we have made when designing a model to explain the water depth of the oceans (s. Fig. 32). The density of oceanic lithosphere must be expressed in terms of temperature (eq. 48) and temperature as a function of depth (eq. 27; s. Turcotte and

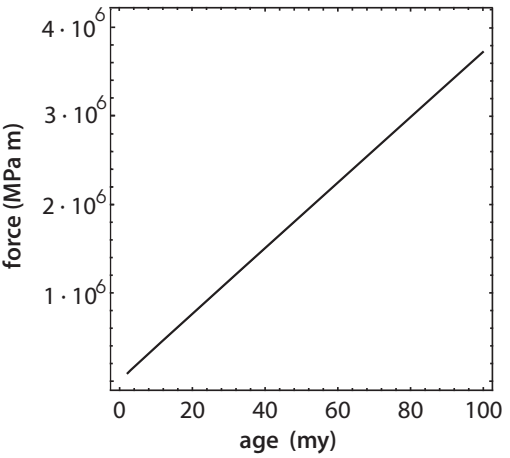


Figure 41: The force exerted by mid-oceanic ridges onto the surrounding plate per meter length of ridge, shown as a function of age of the oceanic lithosphere. Calculated with eq. 84.

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 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:

$$F_\mathrm{b} = g\rho_\mathrm{m}\alpha T_\mathrm{l}\kappa t \left(1 + \left(\frac{\rho_\mathrm{m}}{\rho_\mathrm{m} - \rho_\mathrm{w}}\right)\frac{2\alpha T_\mathrm{l}}{\pi}\right) \approx 1.19 \cdot 10^{-3} t \quad . \tag{84}$$

All parameters of this equation are explained in sect. 7.1. From eq. 84 we can see that the ridge push force is a *linear* function of age of the oceanic lithosphere (Fig. 41). As such it is different from water depth which - within this model is described by a square root function of age (Fig. 33). The numerical value of the proportionality constant between age and force in eq. 84 ($1.19 \cdot 10^{-3}$) is derived using the following constants: $T_\mathrm{l} = 1\,200°\mathrm{C}$; $\rho_\mathrm{m} = 3\,200\ \mathrm{kg\,m^{-3}}$; $\rho_\mathrm{w} = 1\,000\ \mathrm{kg\,m^{-3}}$; $\alpha = 3 \cdot 10^{-5}\ \mathrm{K^{-1}}$ and $\kappa = 10^{-6}\ \mathrm{m^2\,s^{-1}}$. Fig. 41 shows that ridge push is about an order of magnitude *smaller* than the integrated strength of continents at normal orogenic strain rates. Thus, we may conclude that ridge push alone is insufficient as the principal plate tectonic driving force.

9.1.2 Slab Pull and Trench Suction

Old oceanic lithosphere is denser that 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 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. 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} 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.

9.2 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. 40 we have already estimated the magnitude of these forces for a plate of constant density but variable thickness (eq. 82). In this section we want to refine these estimates. Fig. 42 illustrates two examples of potential energy differences between two lithospheric columns. Similar to Fig. 40 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. 79 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).

The considerations of Fig. 42 may be quantified by integrating eq. 79 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. 50, then the lateral buoyancy force is described by:

$$\frac{F_\mathrm{b}}{\rho_\mathrm{m} g z_\mathrm{c}^2} = \frac{\delta(1-\delta)}{2}(f_\mathrm{c}^2 - 1) - \frac{\alpha T_\mathrm{l}}{6(z_\mathrm{c}/z_\mathrm{l})^2}\left(f_\mathrm{l}^2 - 1 - 3\delta(f_\mathrm{c} f_\mathrm{l} - 1)\right)$$

$$+\frac{\alpha^2 T_\mathrm{l}^2}{8(z_\mathrm{c}/z_\mathrm{l})^2}(1 - f_\mathrm{l}^2) \qquad (85)$$

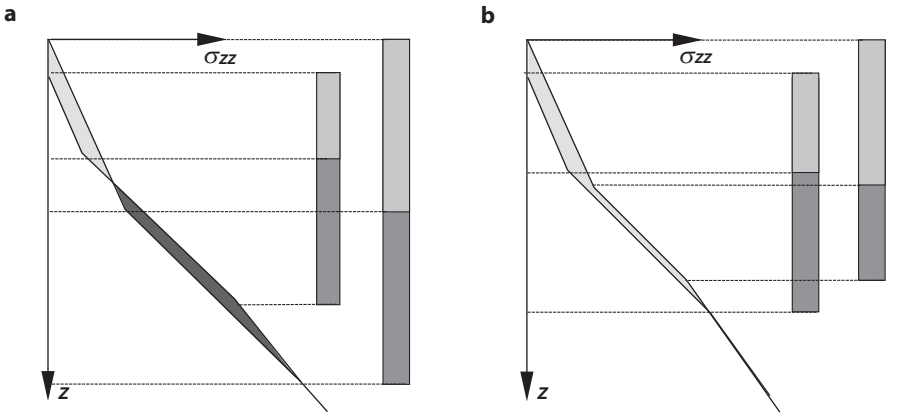


Figure 42: Illustration of vertical stresses and potential energy differences between two neighboring lithospheric columns. Vertical normal stress is plotted as a function of depth. The shaded region between the two curves is the potential energy difference per area between the two adjacent columns. 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.

(Turcotte 1983; Sandiford and Powell 1990). All parameters in this equation are the same as those we used in eq. 56 to calculate the elevation of mountain belts in isostatic equilibrium but the definition of δ differs from that of Sandiford and Powell (1990) and the way eq. 85 is written here differs therefore slightly from theirs as well. Here δ is the density ratio of crust and mantle lithosphere $\delta = (\rho_\mathrm{m} - \rho_\mathrm{c})/\rho_\mathrm{m}$, g is the gravitational acceleration, T_l the temperature at the base of the lithosphere and α is the coefficient of thermal expansion and f_c and f_l are the vertical thickening strains of the crust and the lithosphere, respectively. Lateral forces calculated with eq. 85 are shown in Fig. 43.

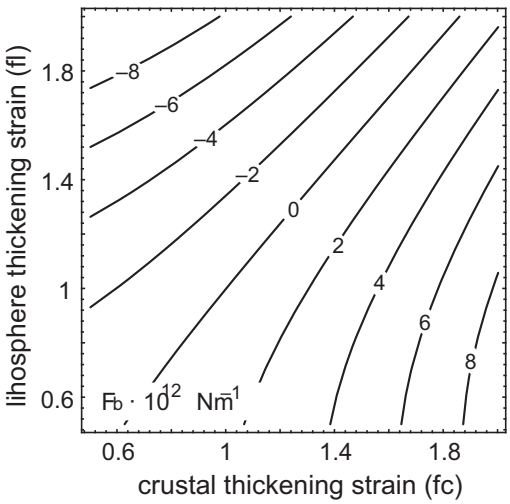


Figure 43: Diagram of lithospheric thickening strain f_l plotted against crustal thickening strain f_c and contoured for potential energy difference per area (equivalent to: "horizontal buoyancy force per meter" or: "lateral force"). The potential energy difference is always that between any point in f_c-f_l space and the reference lithosphere at f_c=f_l = 1. The diagram was calculated with eq. 85 and is contoured for F_b in 10^{12} N m^{-1}. Other assumptions are: ρ_m = 3 200 kg m^{-3}; ρ_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.

10 Unit: Dynamic Evolution of Orogens

Let us begin with an 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 orogens 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. 9.1.1.

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. **??**. It is the vertically *integrated* strength of the lithosphere, which has the units of force/meter and was explained on p. **??**. 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:

$$F_\mathrm{l} = F_\mathrm{d} - F_\mathrm{b} \quad . \tag{86}$$

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 (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".

• **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:

$$F_\mathrm{eff} = F_\mathrm{d} - F_\mathrm{b} \quad . \tag{87}$$

which we already introduced in eq. 86. 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_eff. Eq. 87 is often referred to as the "orogenic 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. 87 is often also written as:

$$F_\mathrm{eff} = F_\mathrm{d} - F_\mathrm{b} = F_\mathrm{l} \quad . \tag{88}$$

There, F_l is the vertically integrated strength of the lithosphere in Nm^{-1} and corresponds to the area under the failure envelope of the lithospere discussed in the 1st unit of this course (Fig. 3c and 4c. 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 $F_\mathrm{eff} < F_\mathrm{l}$, there is no deformation. However, we assume that active orogens are always on the point of failure so that $F_\mathrm{eff} = F_\mathrm{l}$. 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. 88, the integrated failure strength of the lithosphere is zero when the effective driving force is zero.

Because of the balance described by eq. 87 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. If the tectonic driving force is assumed to be constant, then such orogenic evolutions track towards an equilibrium where $F_\mathrm{b} = F_\mathrm{d}$ and $F_\mathrm{eff} = F_\mathrm{l} = 0$ (Fig. 44). Thus, collisional orogens are self limiting. As such, collisional orogens are fundamentally different from extensional orogens, which are not necessarily self limiting.

• **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. That is: if the horizontal and vertical principle stresses are of different magnitude (McKenzie 1972; Molnar and Lyon-Caen 1988). If the were no stress differences, then the surface of a plate subjected to lateral forces from the outside would lift everywhere

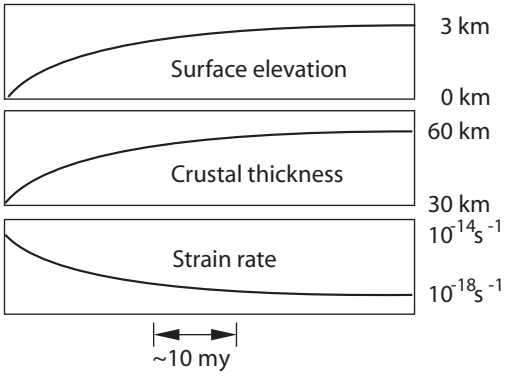


Figure 44: Schematic illustration of the evolution of a collisional orogen subject to the force balance of eq. 88. Surface elevation and crustal thickness converge to a steady state when the magnitude of the horizontal buoyancy force approaches the tectonic driving force

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) (see Unit 11).

Consider a mountain range which collapses under its own weight and to which there is no forces applied externally. then, there is no external driving force and we can reformulate eq. 88 to:

$$F_\mathrm{b} = -F_\mathrm{l} \quad . \tag{89}$$

The left hand side of eq. 89 is the potential energy difference between mountains and foreland per unit area and was evaluated in eq. 83 or, somewhat more precisely, with eq. 85 (s. also Fig. 40). The right hand side of eq. 89 is the integrated strength of the lithosphere. 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 $\sigma_\mathrm{d} = 69$ MPa (see eq. 83). For the Altiplano in the Andes similar estimates indicate a mean strength of $\sigma_\mathrm{d} = 52$ 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), 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.

We have now discussed the force balance of orogens in a qualitative way and have seen how the orogen will converge to a staeady state as the buoyancy forces that oppose the driving forves get larger. With our now-gained knowledge on potential energy we will quantify these considerations in this section.

• **Building up potential energy** In sect. 9.0.1 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 (eq.81). 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.

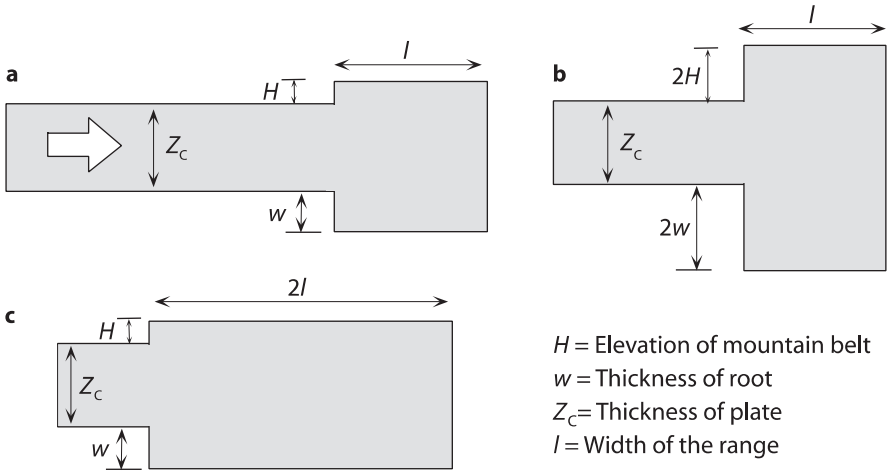


Figure 45: **a** Cartoon of a collisional orogen showing crust of normal thickness on the left and a mountain range on the right. Further displacement of the crust from left to right is compensated in **b** by further thickening and in **c** by lateral growth of the range. The difference in deformation style between *b* and *c* causes a significant difference of the potential energy of the mountain range (see eqs. 90 to 93) (s. also Fig. 40; after Molnar and Lyon-Caen 1988).

In order to understand how this equilibrium is reached, consider Fig. 45a, which illustrates a very simple model orogen. The left of this diagram shows normal thick crust of the thickness z_c and the density ρ_c. On the right, this diagram shows an elevated mountain range in isostatic equilibrium of the elevation H. The diagram is equivalent to Fig. 40. The difference in potential energy between the two mountain range and the foreland per square meter of area is given by eq. 82 and 83. Let us also recall that ΔE_p is a potential energy *per* area and has the units of $J\,m^{-2}$ 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. 83 we can derive directly that:

$$\Delta E_{p,m^{-1}} = \rho_c g H l \,(H/2 + z_c + w/2) \quad . \tag{90}$$

The subscripts are used to emphasize that we are dealing with the units of potential energy difference *per meter* , while the ΔE_p that we used in eq. 82 and eq. 83 has the units of potential energy difference *per area*. Further growth of the mountain range may now proceed either in the vertical direction (Fig. 45b) *or* in the horizontal direction (Fig. 45c). If the crust inside the orogen is doubled in *thickness*, then the potential energy of the range per meter grows to the following value:

$$\Delta E^{\text{high}}_{p,m^{-1}} = 2\rho_c g H l \,(H + z_c + w) \quad . \tag{91}$$

If the growth of the mountain range is by doubling its width (at constant thickness, as shown in Fig. 45c), then the potential energy per meter growth to the following value:

$$\Delta E^{\text{wide}}_{p,m^{-1}} = 2\rho_c g H l \,(H/2 + z_c + w/2) \quad . \tag{92}$$

The difference of the potential energy increases between the two deformation styles is given by the difference between eq. 91 and eq. 92:

$$\Delta E^{\text{high}}_{p,m^{-1}} - \Delta E^{\text{wide}}_{p,m^{-1}} = \rho_c g H l \,(H + w) = \left(\frac{\rho_c \rho_m}{\rho_m - \rho_c}\right) g l H^2 \quad . \tag{93}$$

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. 83. Eq. 93 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.

10.0.1 Mechanics on Vertical Sections

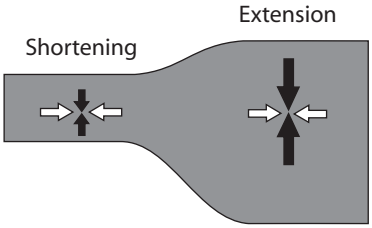


Figure 46: Distribution of horizontal and vertical stresses in a simple collisional orogen. If the topographic gradients at the surface and the base of the lithosphere are small, then the horizontal and vertical stresses σ_{xx} and σ_{zz} are parallel to the principal stresses. The horizontal stresses 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 σ_{xx}, while it is given by σ_{zz} in the orogen.

In the discussion of eq. 93 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. 46) are constant, regardless of thickness of the plate or surface elevation (Artyushkov 1973; Dalmayrac and Molnar 1981). In other words, the product of the mean horizontal stress σ_{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 σ_{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. 46 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 smallest where it is the thinnest (s. Fig. 40). As a consequence, the stress distribution in an orogen may be like that shown in Fig. 46. 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. 9.2), the largest principle stress is the vertical stress. The region is extending. In short: although the horizontal stress on Fig. 46 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. 47).

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.

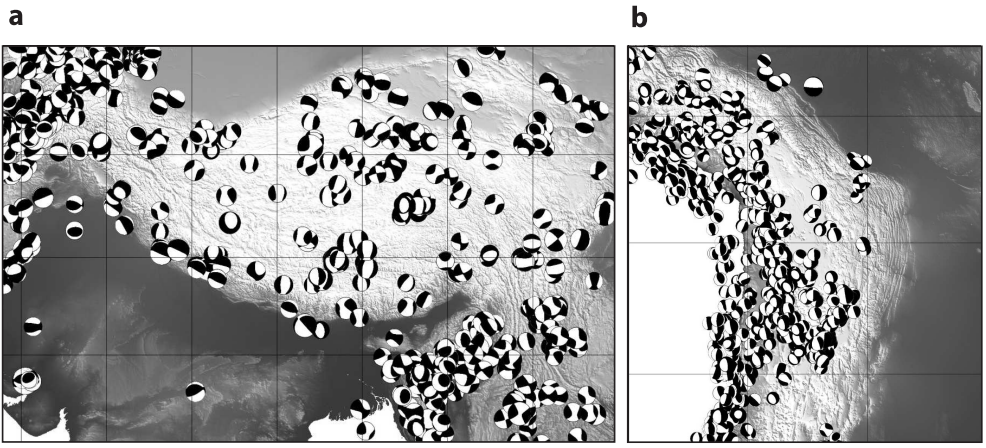


Figure 47: Fault plane solutions for the two great plateaus on this planet. **a** The Tibetan Plateau as the consequence of the India – Asia collision and: **b** the Altiplano as the consequence of the collision between the Pacific and the South American plates. Note that the majority of the fault plane solutions at low elevation regions indicate compression, while those on top of the plateau indicate largely extension.

Appendix: Working on a Spherical Surface

The earth is nearly a sphere 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.

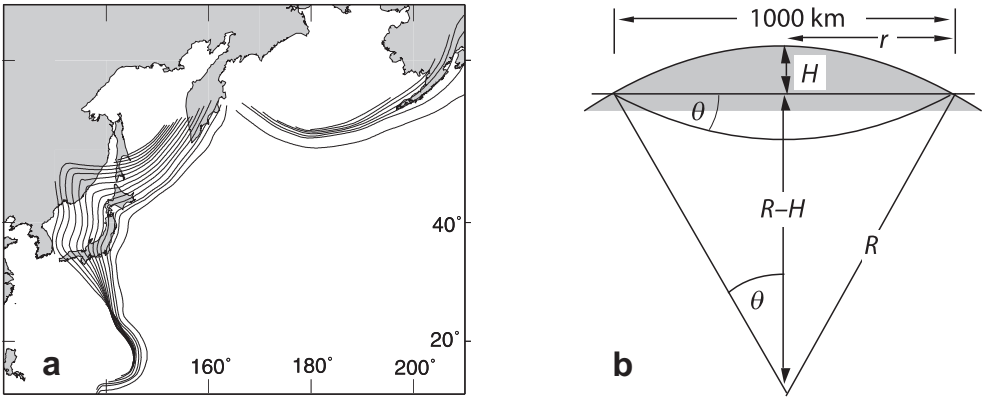


Figure 48: Illustration of the ping pong model. **a** Slab contours for the Aleutian slab and the Kuril-, Japan- and Mariana slabs. Depth contours are shown every 50 km from 0 km to 700 km (in the Japan Trench). Data from Gudmundsson and Sambridge (1998), (http://rses.anu.edu.au/seismology/projects/RUM). **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 $r = 1\,000$ km as shown here and the value for R, the deviation is $H \approx 80$ km

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 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

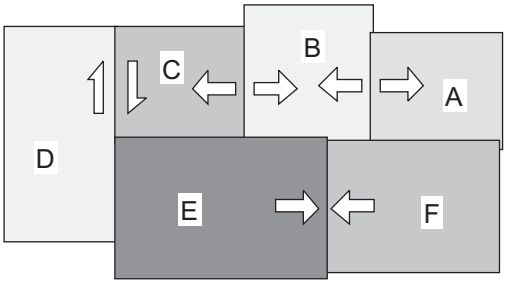


Figure 49: Six plates on a flat surface. The relative motions of some plates are shown by the arrows. However, the relative motions of plates C and E, D and E, B and E, B and F as well as that between A and F are completely unconstrained by the shown relative motions

in subduction zones. In fact, the model can be used to predict the subduction angle ($2\,\theta$), which should be given (according to Fig. 48) by:

$$\sin(\theta) = \frac{r}{R} \tag{94}$$

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. 94 (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 earths 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. 51).

10.1 Geometry on a Sphere

On a spherical surface the position of a point is described by its longitude ϕ, and latitude λ (Fig. 50). 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.

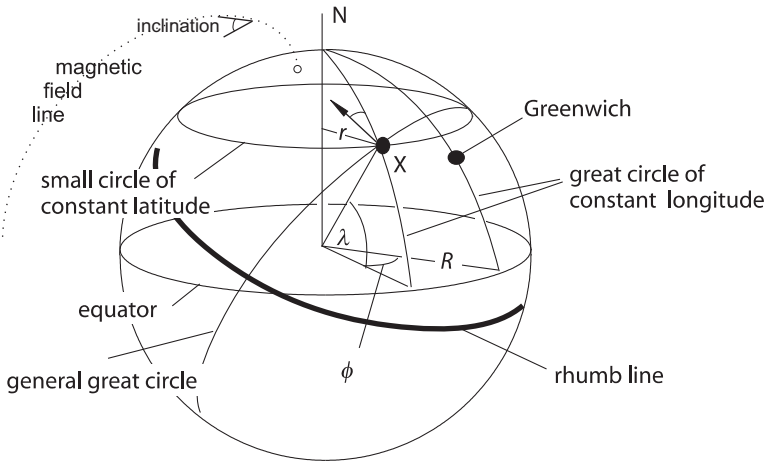


Figure 50: Definitions of important lines and angles on a spherical surface. The geographic longitude ϕ of point X is *west* of Greenwich. The white dot is the magnetic pole. The angle between magnetic and geographic north (labeled at point X) is the declination and the plunge angle of the magnetic field 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:

$$l \approx \cos(\lambda) \cdot 111 \quad . \tag{95}$$

In eq. 95 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 ≈ 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. 50). 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.

10.2 Kinematics on a Sphere

On a flat surface, velocity v and speed have the units of m s^{-1}. Velocity is a vector and speed is a scalar quantity. For example, the Indian Plate has a *speed* of 0.05 m s^{-1}, but a *velocity* of 0.05 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 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. 51). 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 m s^{-2}. Correspondingly, the angular acceleration has the units of 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. 51). In fact, even qualitative changes from divergent plate motion to convergent plate motion may occur along a plate margin because of this (compare Fig. 49 and 51). This is spectacularly illustrate by the transform faults on the ocean floors which solve the space problem caused by the angular rotation.

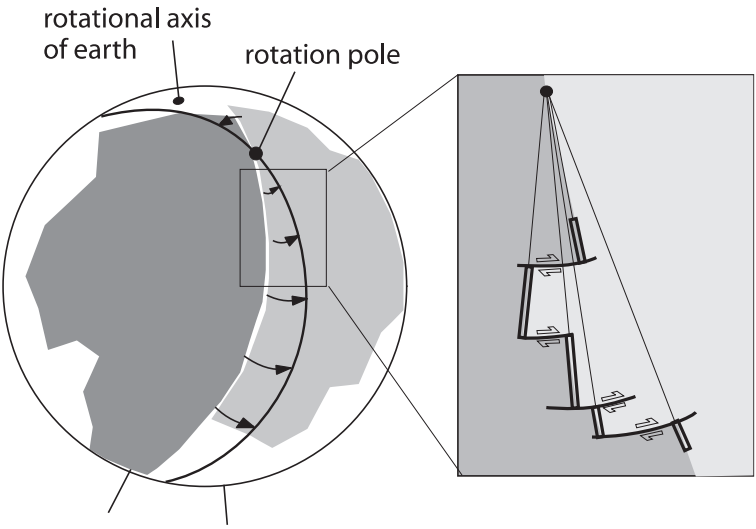


Figure 51: Illustration showing the meaning of rotation poles. The arrows are vectors showing the direction and magnitude of relative motion of the two plates (shaded regions). The thick line connecting the arrow heads is the new plate margin after some time. The axis of the earth is only shown to emphasize that it has nothing to do with the rotation pole of plate motion. The enlargement illustrates the origin of transform faults

10.3 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 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.

Force F is given in Newtons [N] and: $1\ \mathrm{N} = 1\ \mathrm{kg\,m\,s^{-2}}$. 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 $\mathrm{kg\,m^2\,s^{-2}}$. A force of 10^{12} Newton that acts in direction of a great circle on the earth's surface, corresponds to a torque of

Table 5: Important kinematical and mechanical parameters and their units. Each parameter is given with both, the linear and the spherical equivalents

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

$6.37{\cdot}10^{18}$ 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 with torques, which have the units of Newton times meters, there the meters are the distance to the rotation pole.

Force = mass × acceleration ($F = m \times \mathrm{d}v/\mathrm{d}t$) and similarly torque = mass × 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. 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.