Chapter 2.
Production and Conduction of Heat.
September_10_2016.

2.1. Introduction.

As indicated in Chapter 1**,** hydrothermal systems are part of a cascade of processes that dissipates energy at progressively finer scales. This cascade is driven ultimately by the cooling of the Earth. Energy arising from processes associated with heat flow, fluid flow and tectonic displacements at a large spatial scale (the scale of a tectonic plate) cascades down the spatial scales until it is finally dissipated by fine scale fracturing or exothermal chemical reactions (alteration) or is stored in endothermic chemical reactions (mineralisation). Thus the transport of heat is an important process to understand. Such transport can occur by two processes that are important for mineralising systems, namely, conduction and advection in a moving fluid or rock mass. This chapter is concerned only with conduction and some aspects of advection of heat in deforming rocks; advection of heat in fluids, including the important case of advection of heat in a network of permeable fractures, is discussed in Chapter 3. Since the sources and the mechanisms of heat input are fundamental to any understanding of hydrothermal systems we begin this chapter with a brief discussion of heat sources and of heat transport in the mantle of the Earth. It is important in this regard that we arrive at a thermal model of the Earth's mantle based on what we can actually observe and on interpretations of these data sets based on constraints derived from the physics and chemistry of the processes involved rather on models based on the latest fad. Hence we set the scene by considering the planetary scale processes of heat production and transport. We then proceed to discuss finer scale details of heat transport in the lithosphere of the Earth. At the end of the chapter (Section 2.8) we consider the heat generated during the development of a typical hydrothermal alteration system and the implications for the evolution of such systems.

2.2. The Earth as a heat engine.

The second law of thermodynamics says that the energy dissipation of a closed system is always greater than zero (or equal to zero at equilibrium). We have seen in Chapter 1 that dissipation is the conversion of work done by mechanical, hydraulic, thermal and chemical processes into heat. Thus the Earth, which is heated essentially from within, is cooling and, if we consider the Earth and the immediate vicinity of space as a closed system, the various processes that operate within the system are a global response to the second law. A number of alternative modes of cooling are available to the planet and of course it is of immediate concern to us to try and understand which mode operates now and if different modes operated in the past. One possibility is that the large scale dissipative processes that operate could be simply conduction of heat from the core to the crust and subsequent radiation of heat into space. Probably such a process is supplemented at great depths by *radiative* transport of heat through the crystalline solids at lower mantle conditions (Matyska and Yuen, 2005).

The simple conduction of heat is a process driven by the temperature gradient between the core and the surface of the Earth. However such a process has its repercussions in that it generates gradients in density from the hotter to the colder parts of the planet. Such a

density gradient is stable in a gravitational field as long as the stress state generated by small fluctuations in the density distribution results in only elastic deformations. If the temperature gradient is large enough, the density fluctuations result in the yield stress being exceeded throughout the mantle. The system is now unstable so that the coldest, densest material tends to sink whilst the hottest, less-dense material tends to rise.

However since matter cannot interpenetrate, sinking and rising cannot occur simultaneously everywhere. The compromise reached within the system is to form a pattern of sinking and rising motions known as a *convection system*. This is an example of a non-equilibrium stationary state where ultimately the work done by the forcing of the system (arising from the temperature gradient) is balanced by new forms of dissipation, namely those arising from mechanical processes that enable the solid state flow of the mantle and those arising from exothermic phase transitions that operate as the material moves through specific temperature and pressure conditions. This means that the system is now highly nonlinear because new temperature and density gradients in the system can arise from local phase transitions and from energy dissipated by plastic flow particularly if the deformation localises in the form of shear zones.

The types of convection patterns that develop in such nonlinear, non-equilibrium systems are highly sensitive to the temperature gradient(s), the density gradient(s), the geometry of the system and the boundary conditions for the system. The patterns that form are described by a mathematical equation known as the Swift-Hohenberg equation (Cross and Hohenberg, 1993) which has been studied intensely over the past 40 years or so. This equation is one of several non-linear equations that result in pattern formation (Cross and Greenside, 2009). Thus a convecting system is a pattern forming system[1]. This is important because the pattern forming process is a system-wide characteristic process so that a small change anywhere in the system can result in changes to the pattern everywhere in the system. Thus an endothermic phase change (such as the spinel-perovskite transition) or an exothermic phase change (such as the post-perovskite → perovskite transition) can lead to changes in the overall pattern of convection all the way from the core-mantle boundary to the surface of the Earth so that the system does not necessarily resemble the system that develops in the absence of these phase changes. As an example, the post-perovskite → perovskite transition did not occur until the Earth cooled enough for post-perovskite to be stable and form above the core-mantle boundary. Modelling by Maruyama et al. (2007) indicates that this occurred at 2.3 Ga and resulted in increased vigour of mantle convection and intensification of the supercontinent cycle.

As indicated above, the pattern of convective motions is quite sensitive to the boundary conditions. It is common in geodynamic models to impose fixed temperatures at the

[1] We use the term *pattern forming system* where many others would use the term *self-organising system*. Our reason is that the term *self-organising* has an anthropomorphic connotation to it as though the system has a role in planning what happens. The term also encourages unclear thinking in that if one does not understand what causes the pattern to develop one can always say the system self-organised as though such a statement is some form of explanation. Most systems in nature form spatial or temporal patterns because the processes that operate within them comprise competing material transport and material production mechanisms that are described by so called *reaction-diffusion equations* (Cross and Greenside, 2009) which can be shown to produce patterns of various kinds. The Swift-Hohenberg equation is a special form of these equations.

top and bottom of the system. Thus the top of the convecting Earth is commonly fixed at 0^oC or 25^oC and the core-mantle boundary at 6000^oC or some temperature (which might change with time as the Earth cools) based on a selected Earth model. A moment's thought will indicate that these boundary conditions are very difficult to establish from an experimental point of view. The fixed top temperature would require an array of thermometers linked to a control pumping system that circulates constant temperature fluids continuously chilled by a cooling system. The core-mantle temperature can only be kept at constant temperature again by installing an array of thermometers that control a system that pumps very hot fluids that are maintained at constant temperature over the boundary. In other words, fixing the top and bottom temperatures of the system requires some-one to do work. In a closed thermodynamic system this requires intervention from outside the system. No such work is done when a modeller fixes these temperatures for a computer simulation; the model is not compatible with the second law of thermodynamics. The relevant boundary conditions are *radiative* heat boundaries at the top of the system, so that the Earth radiates heat into space, and heat *flux* conditions at the base of the system where heat is added from the core (perhaps in a non-uniform manner). Entirely different convective systems develop with these different boundary conditions. The reason for this is that the constant temperature boundary conditions constrain the isotherms near the boundaries to be strictly parallel to the top and bottom surfaces whereas the radiative/flux boundaries induce no such constraint.

A second issue concerns the shape of the system within which convection occurs. The solutions to the Swift-Hohenberg equation for small perturbations of the fluid displacement are commonly periodic with a wavelength that for a given system and material is set by the temperature gradient. This means that an integral number of wavelengths need to fit into the space available. If this is not geometrically possible the pattern that develops is as close as possible to the right number of wavelengths but it is possible that the system will continuously switch between the two or more close alternatives. This also means that the geometrical necessity to fit an integral number of wavelengths into the space available leads to different patterns of convection for two and three dimensional systems and for planar or spherical geometries. Thus convection models presented in planar two and three dimensional systems need not be representative of three dimensional spherical geometries.

The real Earth however exhibits even greater complexity in that the material properties of the material that convects are sensitive to temperature and pressure. The properties of importance here are the yield stress, viscosity, the density, the thermal conductivity and the thermal expansion. Models that do not take such dependencies into account are severely deficient. Moreover there are a number of mineralogical phase changes that occur within the Earth that need to be taken into account. These include the eclogite transition, the olivine-spinel transition, the post perovskite $\rightarrow$ perovskite transition (exothermic) and many others in the mid mantle involving pyroxene phase changes. As we have indicated, neglect of these phase changes makes significant differences to the patterns of convection that develop when they are included. Increases in viscosity make the system more sluggish to convection whilst increases in density and decreases in thermal expansion decrease the buoyancy of the material; exothermic/endothermic phase transitions add/subtract

heat at specific parts of the mantle and hence influence the local physical properties but have an effect throughout the entire mantle.

An additional issue concerns internal heat production. The Earth's mantle is not simply heated by fluxes of heat from the core, the decay of radiogenic isotopes leads to significant internal heating as well. Thus the fundamental issues that need to be incorporated into any realistic model of the thermal structure of the Earth are:

- Realistic boundary conditions that are compatible with the laws of thermodynamics and do not require some-one (a Maxwell Demon) to do work on the system in order to maintain the boundary conditions; this means heat flux boundaries at the base and radiative boundaries at the top of the system.
- Realistic geometry; this means 3D, spherical models need to be studied.
- The temperature and pressure dependencies of yield stress, viscosity, density, thermal expansion and thermal conductivity.
- The influence of phase changes (thermal and mechanical).
- Internal heating due to radioactive decay.
- Contributions to thermal transport from radiative transfer.
- Secular variations in basal heat flux and heat production from radioactive decay.
- Constitutive relations need to be realistic. This means inclusion of elasticity, plasticity and viscous effects together with the realistic evolution of microstructure including CPO development and anisotropy. Without incorporation of elasticity and a yield surface, self-consistent localisation during deformation does not occur in materials. Commonly localisation is induced by some artificial means in geodynamic models.

The last and most important issue that demands attention is to understand the temperature gradient in the Earth's mantle. Knowledge of this gradient is fundamental for any self-consistent theory of hydrothermal mineralising systems. Ever since the classical paper by McKenzie and Bickle (1988) the geological community has embraced the concept that the thermal gradient in the Earth's mantle is adiabatic. This concept of an adiabatic gradient is illustrated in Figure 2.1. If we take a small parcel of rock at a particular depth with a temperature, T_1, and raise the parcel to a higher level whilst maintaining the heat content of the parcel then the volume of the parcel will increase as the pressure drops and this increase in volume at constant heat content results in a drop in temperature to T_2. If T_2 is less than the temperature prescribed by the local temperature gradient the gradient is sub-adiabatic. If the temperatures are equal then the gradient is adiabatic; if T_2 is greater than that given by the local temperature gradient the gradient is super-adiabatic. If the gradient is super-adiabatic the parcel is less dense than the surrounding material and buoyancy is enhanced. If the gradient is sub-adiabatic then the density of the parcel is larger than the surrounding material and buoyancy is inhibited. Notice that an isothermal gradient is sub-adiabatic.

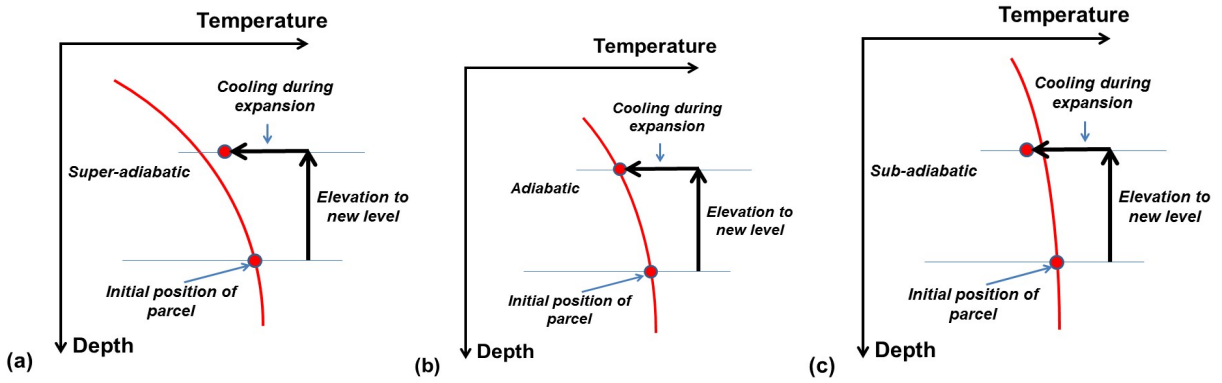


Figure 2.1. Definitions of temperature gradients. (a) super-adiabatic, (b) adiabatic and (c) sub-adiabatic temperature gradients.

The thermal gradient that develops in the mantle depends on the pattern of convective motions that form rather than being imposed by boundary conditions. These convective motions, in turn, are controlled by the dot points listed above. Numerical models that consider all of these dot points do not exist but many come close. The basic conclusion from the most recent models (Matyska and Yuen, 2005; Schuberth et al., 2009; Zhang et al., 2010) is that the thermal profile in much of the mantle is not adiabatic (figure 2.2 a) but instead is sub-adiabatic (figure 2.2 b) with some of the highest temperatures in the upper 400 to 500 km of the mantle.

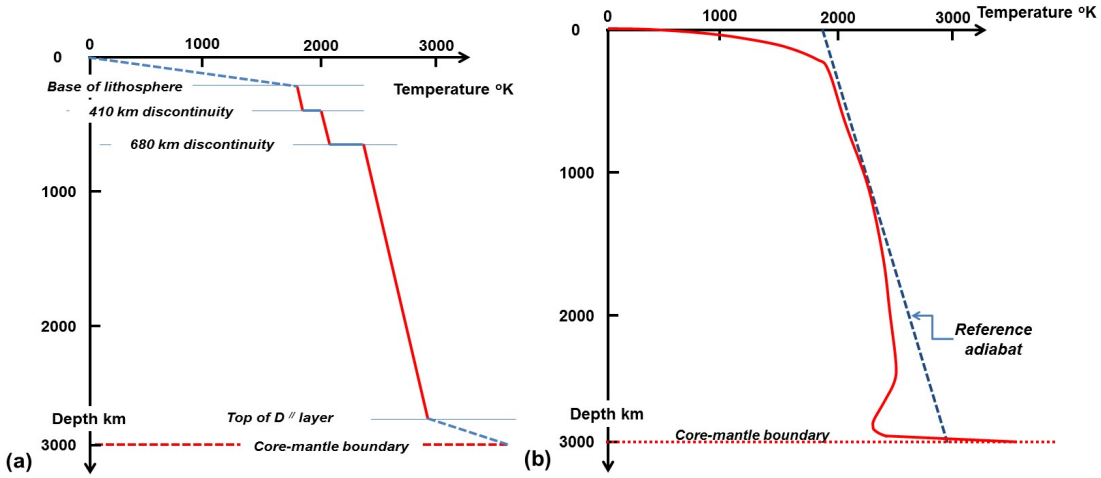


Figure 2.2. Models for temperature distributions through the mantle of the Earth. (a) Distribution proposed by Turcotte and Schubert (1982). The full red lines correspond to adiabatic gradients. The dashed blue lines correspond to super-adiabatic gradients. (b) Temperature distribution with internal heating from radiogenic decay and heat flux from the core at the base. Modified from Schuberth et al. (2009).

2.2.1. The observations.

A new paradigm can arise by revisiting the basic assumptions made in the old paradigm. Following this reasoning we lay out below the fundamental observations we can make about the Earth and compare them with the current paradigm of *Plate Tectonics* and with what is sometimes posed as a competing process called the *Plume Hypothesis*. The basic observations that we can make concerning the Earth and how it works can be divided into two classes:

(i) Direct observation of the make-up of the surface of the Earth including its structure and mineralogical composition together with any secular variations in these properties.
(ii) Geophysical observations including gravity, seismic observations including wave speeds and anisotropy, free oscillations, heat flow, and GPS or satellite observations of surface displacements.

These are the observations (the phenomenology) that constrain any hypothesis that is put forward to explain them. Everything else is part of that hypothesis. Any hypothesis however, is constrained by the laws of physics and chemistry, including mechanics and thermodynamics, and hence must, in particular, be consistent with the second law of thermodynamics. Any reasonable hypothesis should also be based on realistic material properties and behaviour of rocks rather than be based on convenience or the ability to run computer codes that might represent the system. This means to us that simulations of the Earth based on fluid dynamics in which elasticity and any yield phenomena are neglected have to be suspect. We take as a starting point that a fundamental observation is that the present surface of the Earth is divided into a number of segments that move and behave internally as coherent entities. These are the *plates.*

Direct structural and petrological observations.

The first order observation we make here is that the continents display a long history of intense non-elastic deformations. *Thus the plates are not rigid*[2] and never have been. There seems to be a secular change in the temperatures of the preserved plates in that the first record of bi-modal metamorphism (co-existing belts of high and low pressure metamorphic rocks) dates from ~3 Ga (Figure 2.3 b from Brown, 2014). Before that time (and in some cases such as the Yilgarn before ~2.8 Ga) the surface exposed rocks exhibit relatively low grades of metamorphism. This generates a paradox (called by some the Archean paradox (Moresi, 2013)) where high temperature komatiites are extruded onto apparently cold lithosphere. Brown (2014) proposes that modern plate tectonics begins at ~3.0 Ga. A switch to lawsonite-bearing blueschists and eclogites and ultra-high pressure metamorphic rocks occurs at 0.55 Ga associated with significant cooling of the subducting slabs (Figure 2.3 a).

Maruyama et al. (2007) propose, based on a range of geological and geophysical observations, a thermal history for the mantle as follows: (i) Birth of the solid inner core at 3.6 Ga. (ii) Double layer mantle convection prior to ~2.8 Ga. (iii) Whole mantle overturn between 2.8 and 2.7 Ga with widespread eruption of basalts at the surface and development of whole mantle convection. (iv) Formation of post-perovskite at the core-mantle boundary at 2.3 Ga associated with increase in the vigour of mantle convection and increased rates of crustal growth; possible intensification of the supercontinent cycle. (v) Massive water transport into the mantle begins at 750 Ma when subducting slabs cooled enough to transport

[2] A rigid body is one where deformations are such that the distance between any two points in the body does not change (Truesdell, 1966). The concept has nothing to do with strength defined as the level of stress that the body can support. A very weak body can undergo a rigid deformation if constrained by boundary conditions to do so. Both structural observations and GPS data indicate that the plates are not rigid. They undergo internal non-reversible deformations

significant volumes of water. This resulted in the appearance of lawsonite bearing eclogites and blueschists on Earth.

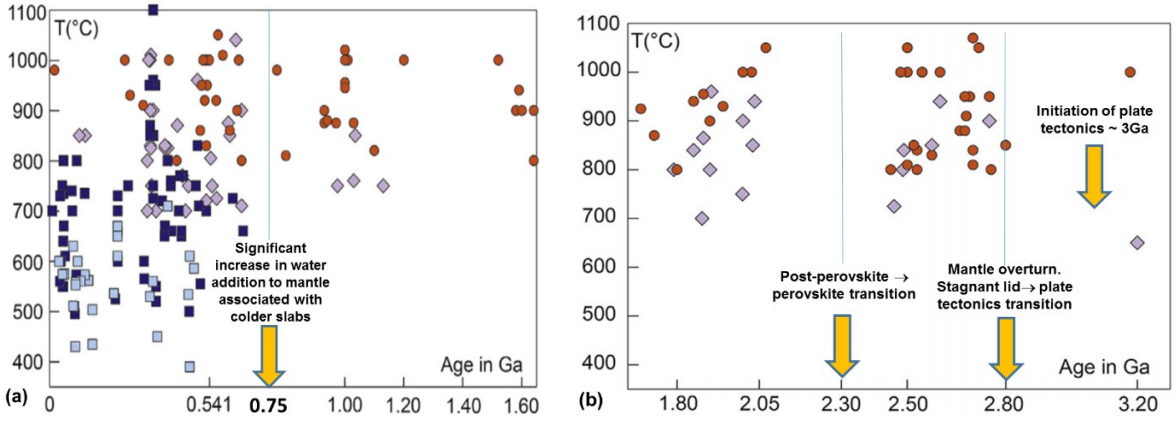


Figure 2.3. The distribution of metamorphic conditions throughout the known history of the Earth from Brown (2014). (a) The time span between modern and 1.8 Ga (b) The time span between 1.8 Ga and 3.2 Ga. Legend: orange circles: granulite to ultra high temperature metamorphism; light purple diamonds: eclogite to high pressure granulite metamorphism; pale blue squares: lawsonite bearing blueschists and eclogites; dark blue squares: eclogites and ultra-high pressure metamorphic rocks. The orange arrows indicate proposed transitions from stagnant lid → plate tectonics at ~3 Ga (Brown, 2014) , the post perovskite → perovskite transition at 2.3 Ga (Maruyama et al., 2007) and the cooling of the subducting slabs that allowed incresed amounts of water to be added to the mantle (Maruyama et al., 2007).

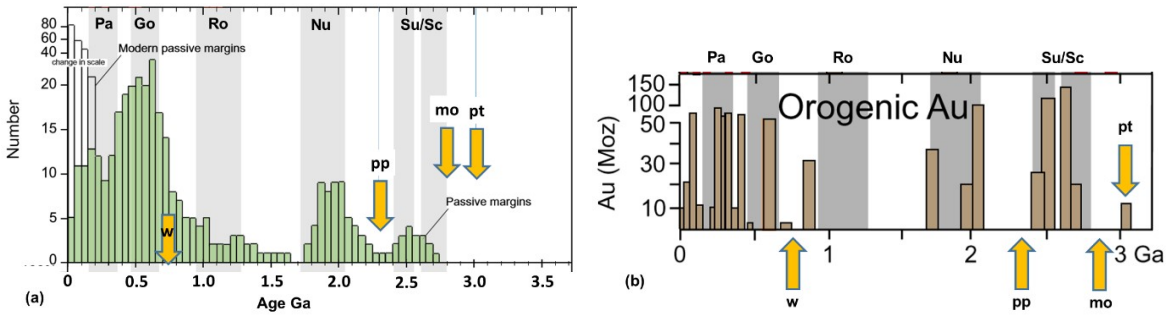


Figure 2.4. Secular changes in passive margin development and orogenic gold mieralisation. Compiled from Cawood and Hawkesworth (2014). (a) Passive margin development through time. (b) Orogenic gold mineralisation through time. In each diagram the gray bars represent supercontinent development. Pa: Pangaea; Go: Gondwana; Ro: Rodinia; Nu: Nuna; Su/Sc: Superia/Sclavia. The orange arrows represent various transitions postulated in the mantle. w: significant increase in water added to mantle; pp: post-perovskite → perovskite transition initiated at core/mantle boundary; mo: whole mantle overturn; pt: postulated start of plate tectonics.

In Figure 2.4 we show the evolution of the number of passive margins preserved in the geological record and the distribution of orogenic gold mineralisation with time superimposed on the postuated supercontinent cycle and with significant mantle events indicated by arrows. These distributions and the various gaps in both Figures 2.3 and 2.4 seem to tell a fairly consistent story of significant mantle events influencing passive margin development, suepercontinent assembly and breakup, styles of metamorphism and orogenic gold mineralising events.

Geophysical observations.

The first order features derived for gravity maps of the Earth reveal two broad equatorial gravity lows, one situated in the Pacific Ocean and the other 180^{o} away under

Africa. These gravity anomalies correspond to elevated regions on the geoid and regions of high seismic attenuation at moderate depths in the mantle (Figure 2.5). These are interpreted as broad thermal up-wellings within the mantle of the Earth. Also imaged are narrow gravity highs that coincide with ocean trenches and with zones of deep seismic activity (the Benioff zones). Seismic tomography reveals that these zones are slabs of (presumably cold) material that extend to depths of about 650 km and occasionally deeper. The results of seismic reflection and refraction studies over the past 100 years have arrived at an internally-consistent model of the density structure of the Earth comprised of a number of layers: the crust, the lithosphere, discontinuities at 410, 660 and approximately 3000 km (the $D^{//}$ layer), the core mantle boundary and the transition between the outer and inner cores (Figure 2.7). Arguments based on thermodynamics indicate that the density distribution (including discontinuities at phase transitions) is best explained by the increase in pressure with depth and (for at least the lower half of the mantle) a sub-adiabatic temperature gradient.

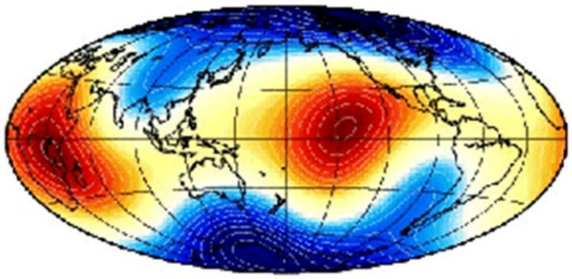


Figure 2.5. Seismic attenuation within the Earth at 500 km depth. The high attenuation regions (dark red) are on the equator 180 degrees apart and correspond to gravity lows and elevated regions on the geoid.

The two gravity lows (which correspond to two equatorial elevated regions on the geoid 180^{o} apart) under the Pacific Ocean and Africa have now been correlated, using seismic tomography (Romancwicz et al., 2015), with hot upwellings originating in the vicinity of the core-mantle boundary (Figures 2.5 and 2.6). These upwellings have been christened *superplumes* and commonly become dispersed above about 1000 km depth (Figure 2.6). Other hot regions exist high in the mantle, mainly above the 660 km discontinuity with no necessary connection to the major upwellings. There is not always a one-to-one correspondence between hot spots on the surface of the Earth and hot regions in the upper mantle. For instance there is no hot region immediately below Yellowstone (Figure 2.6 d).

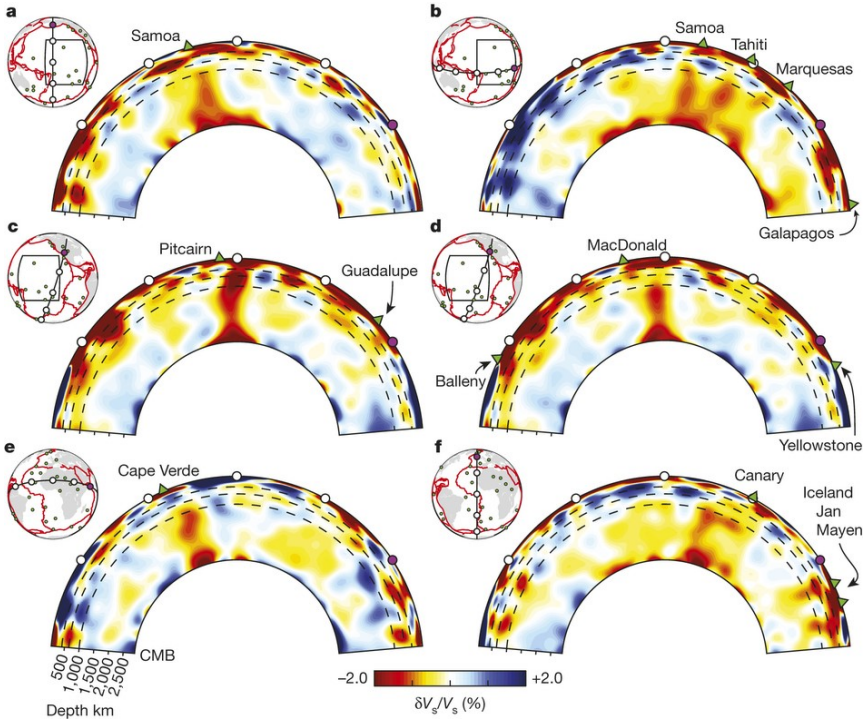


Figure 2.6. Cross sections (positions shown on insets) across the Earth based on seismic tomography. The Pacific superplume is shown in (a), (b), (c) and (d) and the African superplume in (e) and (f). From Romancwicz et al. (2015). Superplumes are generally restricted to the mantle below 1000 km. The example beneath Pitcairn, (c), is an exception. In other cases the superplume is dispersed at about 1000 km depth. Hot regions in the upper mantle are, in general, separated from the superplumes and form flat layers commonly above the 660 km discontinuity and sometimes only above the 410 km discontinuity. Hotspots at the surface are shown as green dots in the insets.

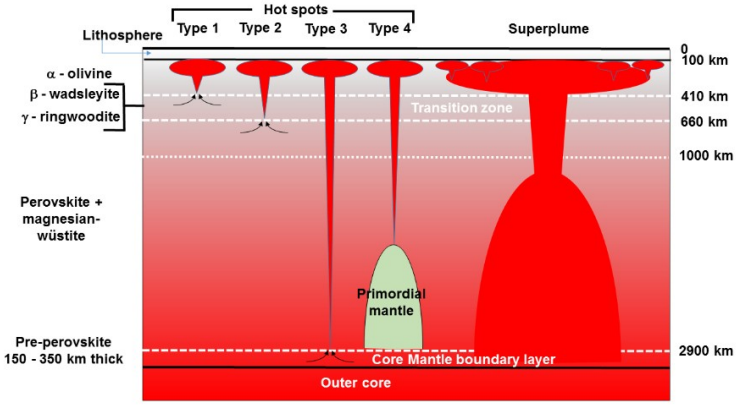


Figure 2.7. A classification of plumes after Cserepes and Yuen (2000). Type 1, 2, 3 and 4 respectively originate at the 410 km, 660 km, core mantle boundary layer and a relic part of primordal mantle. These plumes typically have a lifetime of ~100 my and result in hot spots. Superplumes act as engines that drive the dynamics of the Earth and have a typical lifetime of ~800-1000 my. Although hot regions not directly associated with superplumes exist in the uppermost mantle (Figure 2.6), the long narrow feeeder zones for plumes of type 1 to 4 are not resolvable by present day seismic methods and so have not been observed.

Although one can show that plume structures can nucleate on a interface within a layered fluid the anlysis is for small displacements of the interface and says nothing about the evolution of the plume at large displacements. In addition there is at present no compelling seismic evidence for small plumes that might form at the 410 km, 660 km discontinuities or at the top of the $D^{//}$ layer. However the resolution of seismic surveys is not high enough to

detect the narrow tails if they are present. There are hot more or less horizontal layers above the 410 km and 660 km discontinuities but they generally do not resemble plumes. Thus although plume type structures have been postulated and classified (Figure 2.7), at present there is little in the way of direct observational evidence for these structures except of course at the very large scale in the form of the Pacific and African superplumes.

A seemingly well accepted model for the evolution of superplumes is that they form in the "graveyard" of old subduction slabs where these slabs disturb the $D^{//}$ layer and generate heat by promoting the exothermal post-perovskite → perovskite transition (Yuen et al., 2007). In the mature evolution of a period of subduction activity this graveyard lies beneath a supercontinent that has developed during an extended period of global subduction. The blanketing effect of the supercontinent also promotes the development of a superplume (Anderson, 1982). The activity of the superplume nucleates rifts in the supercontinent and a new period of supercontinent break-up initiates. The process then repeates itself with a period of approxiamtely 800 My (Li and Zhonge, 2009). Clearly this cyclic amalganation and breakup of supercontinents has important potential for controlling heat supply to the crust. In this model the Pacific and African superplumes are thought to be left over from the Rodinia and Pangea supercontinent episodes (Utsunomiya et al. 2007). There is some debate on this issue (Li and Zhong, 2009).

A common model for introducing heat into the lithosphere is to propose the impingement of a hot (say 1500°C) plume head on the base of the lithosphere. The models then need to incorporate mechanisms to transport heat through the lithosphere if the lithosphere remains intact. We will see in Section 2.3 that for lithospheric thicknesses of 100 to 300 km this involves thermal time scales of 316 my to 2.85 Gy which would seem to be too long to explain the metamorphic and hydrothermal systems we observe. One mechanism to overcome this, involving rapid thermal erosion of the lithosphere (Figure 2.8), has been proposed by Cloetingh et al. (2013) with a similar model proposed by Sobolev et al. (2011). The thermal time constant is then reduced to that of a thinned crust (say 30 My or less).

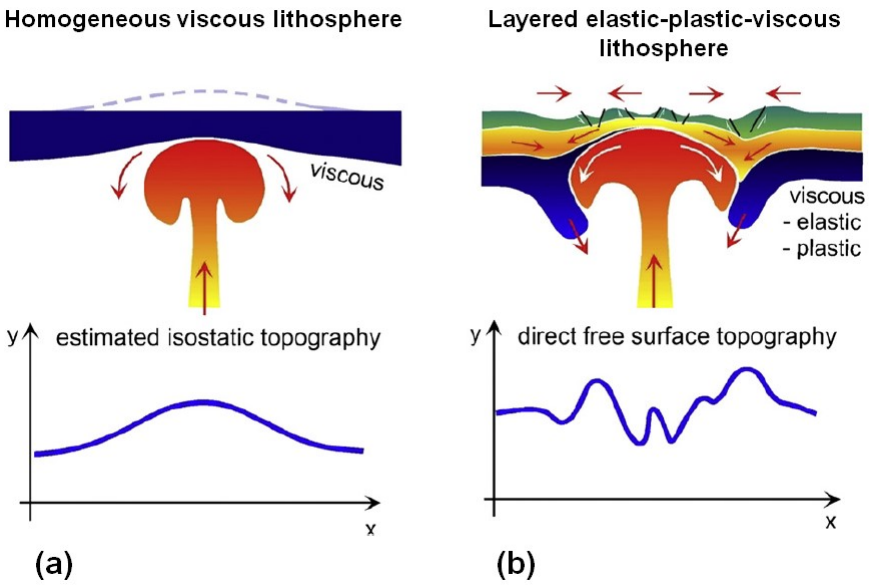


Figure 2.8. Plume impingement at the base of the lithosphere (From Cloetingh et al., 2013). (a) Impingement at the base of a purely viscous lithosphere. (b) Impingement at the base of an elastic-plastic-viscous lithosphere with rapid erosion of the sub-continental lithosphere.

Perhaps a more realistic example of the influence of a plume on lithospheric structure is the model of Gerya et al. (2016) for the breakup of a supercontinent arising from plume emplacement (Figure 2.9). This model is proposed as representing the first initiation of subduction accompanying rifting in a supercontinent.

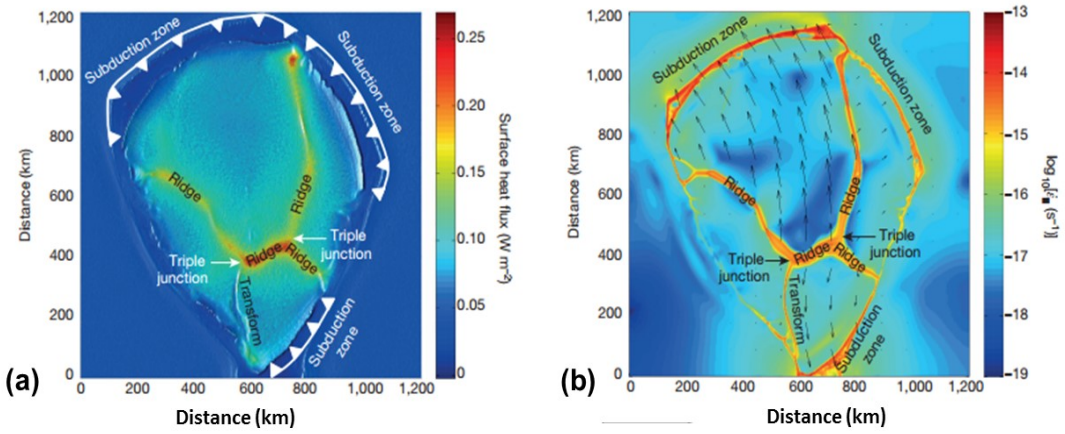


Figure 2.9. The breakup of a supercontinent and initiation of subduction zones arising from plume emplacement beneath the supercontinent. (a) Surface heat flux. (b) Distribution of strain rate. From Gerya et al. (2016).

2.3. Fourier's law of heat conduction. Diffusion of heat.

If a temperature gradient exists in a solid then heat is observed to flow from the hotter to the cooler parts of the body. This is an example of a system not at equilibrium. The thermodynamic force that drives the system away from equilibrium in this case is the gradient of (1/T) where T is the absolute temperature. The thermodynamic flow that dissipates energy and tends to move the system towards equilibrium is heat flow. If the temperature gradient is maintained a non-equilibrium stationary state eventually evolves where the thermal flux is such as to generate a constant temperature gradient in the body. Such a state is sometimes, erroneously, referred to as thermal equilibrium. Since heat continues to flow, this state, comprising a constant temperature gradient, is a non-equilibrium stationary state and not an equilibrium state. Thermal equilibrium corresponds to a constant temperature throughout the Earth.

If the material is isotropic then the flow of heat is proportional to the temperature gradient. This relationship is known as *Fourier's Law* and is expressed as:

$$\mathbf{q} = -\mathrm{k}\,(grad\mathrm{T}) \qquad (2.1)$$

where **q** is the heat flow vector expressed in the units Joules m^{-2} s^{-1}, and k is a material constant known as the *thermal conductivity* with units Joules m^{-1} s^{-1} K^{-1}; T is the absolute temperature. The minus sign indicates that the heat flows *down* the temperature gradient. If the material is anisotropic the thermal conductivity may also be anisotropic (Nye, 1957).

The flow of heat can be expressed as a diffusion equation. The argument is developed in Carslaw and Jaeger (1959). The result is:

$$\rho c_P \frac{\partial \mathrm{T}}{\partial t} = \mathrm{k}\left(\frac{\partial^2 \mathrm{T}}{\partial x^2} + \frac{\partial^2 \mathrm{T}}{\partial y^2} + \frac{\partial^2 \mathrm{T}}{\partial z^2}\right) \equiv \mathrm{k}\nabla^2 \mathrm{T} \qquad (2.2)$$

or,

$$\frac{\partial T}{\partial t} = \kappa\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) \equiv \kappa\nabla^2 T \qquad (2.3)$$

where $\kappa = \frac{k}{\rho c_P}$ is known as the *thermal diffusivity* and has the units $m^2\ s^{-1}$. As with any diffusive process, if temperature changes occur with a characteristic time scale τ then they diffuse a distance $\sqrt{\kappa\tau}$ in that time. On the other hand a time L^2/κ is required for a temperature change to diffuse over the distance L. The average value of the thermal diffusivity for rocks is $10^{-6}\ m^2\ s^{-1}$. Thus the characteristic time scale for a temperature change imposed at the base of a 30km thick crust is $(30000)^2/10^{-6}$ seconds or approximately 28.5 my (taking 1 year = 3.16 x 10^7 seconds). Equally a plume arriving at the base of a lithosphere 150 km thick is associated with a characteristic thermal time scale of 712 my. These very long time scales for thermal effects to be felt far from the application of a temperature change need to be taken into account when one postulates the impingement of plumes or delamination events at the base of the lithosphere. A plot of the characteristic time, τ, against the length scale, L, of the system is shown in figure 2.10 for $\kappa = 10^{-6}\ m^2\ s^{-1}$.

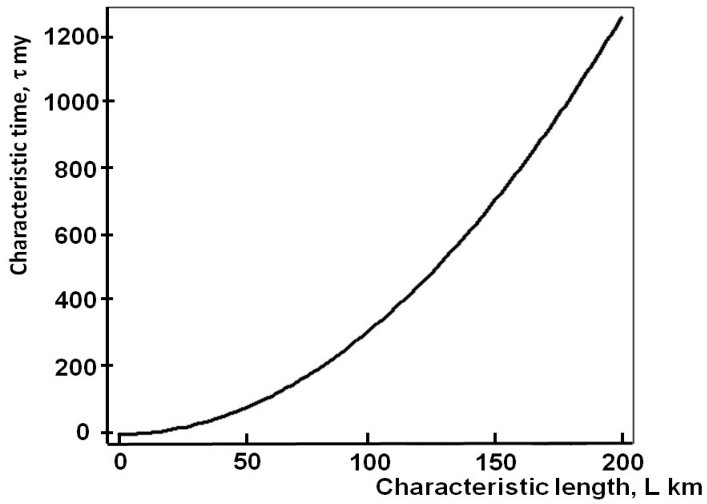
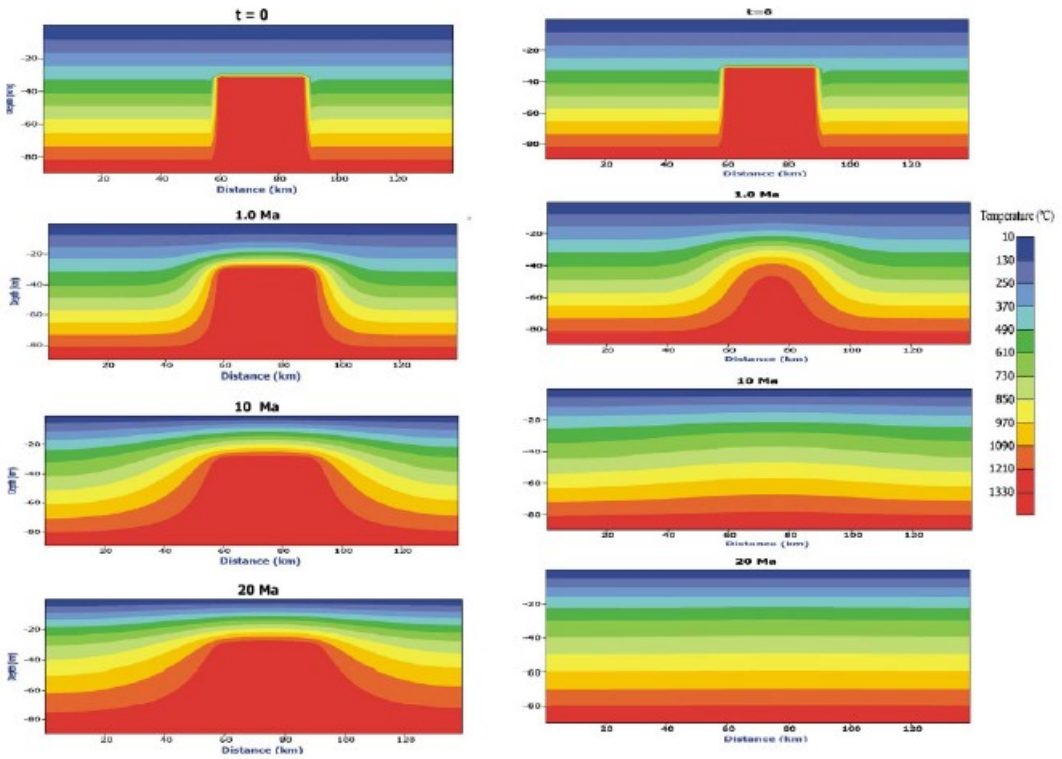


Figure 2.10. Plot of characteristic time for heat diffusion over a characteristic length scale. Drawn for $\kappa = 10^{-6}\ m^2\ s^{-1}$.

If heat is produced at a point, P, at a rate *H* (expressed as Joules $kg^{-1}\ s^{-1}$) then, for a conductivity independent of temperature, (2.2) becomes

$$\nabla^2 T - \frac{1}{\kappa}\frac{\partial T}{\partial t} = -\frac{H(x,y,z,t)}{\kappa c_P} \qquad (2.4)$$

Examples of temperature distribution evolution following a mantle upwelling into the lithosphere is shown on Figure 2.11. Notice that if the upwelling maintains its temperature the thermal effects are felt 20 km away in about 20 million years. If the upwelling is allowed to cool its effects vanish in 20 million years.



Intrusion at constant T (1330°C) Intrusion allowed to cool

Figure 2.11. Scenarios of thermal structure due to conductive heat transfer of an asthenospheric upwelling (Ranalli and Rybach 2005). Models are shown that illustrate the impact of an asthenospheric upwelling on the change of thermal structure within the Earth's crust depending on whether the upwelling maintains its temperature (left) or cools (right).

2.4. Internal heat production.

Heat is produced within deforming chemically reacting rocks from the deformation and chemical processes that operate during metamorphism, from the circulation of hot fluids (including magmas) and also from the decay of radioactive isotopes. A large part of the heat that drives many metamorphic processes originates from the radioactive decay of uranium, thorium and potassium mainly within the crust. Here we concentrate solely on this process and discuss the effects of heat produced by deformation and mineral reactions in Section 2.8.

The rate of decay of a radioactive isotope is given by

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -\lambda^{(1/2)} N \qquad (2.5)$$

where N is the number of atoms of the radionuclide present at time t and $\lambda^{(1/2)}$ is known as the *decay constant*. The solution to (2.12) gives the number of atoms present at time t when the number of atoms present at time $t = 0$ is N_0:

$$N = N_0 \exp\left(-\lambda^{(1/2)} t\right) \qquad (2.6)$$

The *half-life*, $\tau^{(1/2)}$, of a radionuclide is the time required for half of the atoms present at time $t = 0$ to decay. We can obtain this by writing

$$\frac{N}{N_0} = 0.5 = \exp\left(-\lambda^{(1/2)} \tau^{(1/2)}\right)$$

or,

$$\tau^{(½)} = \frac{\ln 2}{\lambda^{(½)}} = \frac{0.69325}{\lambda^{(½)}}$$

The decay constants and half-lives for the common heat producing radio-nuclides are given in Table 2.1 together with the heat production for each isotope or element.

Table 2.1. The decay constants, $\lambda^{(½)}$, half-lives, $\tau^{(1/2)}$ and rates of heat release, *H*, of the important radioactive isotopes in the Earth. Values taken from Turcotte and Schubert, 1982).

Isotope	**Decay constant, $\lambda^{(½)}$ (yr^{-1})**	**Half-life, $\tau^{(1/2)}$ (years)**	***H*** **($W\ kg^{-1}$)**
^{238}U	1.55×10^{-10}	4.47×10^{9}	9.37×10^{-5}
^{235}U	9.85×10^{-10}	7.04×10^{8}	5.69×10^{-4}
U			9.71×10^{-5}
^{232}Th	4.9×10^{-11}	1.40×10^{10}	2.69×10^{-5}
^{40}K	5.54×10^{-10}	1.25×10^{9}	$2,79 \times 10^{-5}$
K			3.58×10^{-9}

As an example, one can confirm from Table 2.1 that an Ordovician rock with a current composition (by weight) of 3.5% K_2O, 3.5 ppm uranium and 16.9 ppm thorium has a current heat production rate of 2.4 μWm^{-3}. In the Ordovician this same rock would have had a heat production rate approximately 8% higher.

2.5. Tectonic models of thermal evolution.

Metamorphic and metasomatic systems are regions of enhanced temperature gradients within the crust of the Earth. The processes that drive such enhanced gradients have been attributed to at least four classes of events namely: (i) enhanced thickening of the crust in which heat is produced by radioactive decay; (ii) various contributions from regions of enhanced heat production due to radioactive decay that are moved during deformation; (iii) plume impingement on the base of the lithosphere; and (iv) delamination of the lithosphere associated with Rayleigh-Taylor instabilities.

The temperature of a particular elemental volume of rock in a metamorphic system at a point P and its evolution with time is dependent on a number of factors which can be considered under two headings:

(i) ***Internal heat production at P*** due to internal processes such as radioactive decay, heat production or absorption arising from chemical reactions, mechanical deformation and fluid flow at P. We examine the influence of radioactive decay below but defer a discussion of the influence of other forms of internal heat production to later chapters.

(ii) ***The relative passage of isotherms past P.*** Such movement of isotherms relative to P may be due to end member processes: The first is thermal conduction, which we have seen is a relatively slow process. The second is advection of heat within a deforming body. In this process heat is carried with the deforming body as it moves; this is a relatively fast process.

In the case where it is proposed that chemical equilibrium has been achieved a metamorphic petrologist would sometimes adopt a spatial or Eulerian representation and be only interested in the temperature and pressure at a particular spatial point where the current bulk chemical composition is defined. If the history of metamorphism is considered in the form of PT*t* diagrams (England and Thompson, 1984) then sometimes a Lagrangian approach is adopted although commonly it is not clear whether a Lagrangian or Eulerian approach is considered. The following examples are meant as illustrations of the distinction between Eulerian and Lagrangian representations of thermal evolution and of the different time scales associated with thermal conduction and thermal advection. We briefly examine four models for the production of heat that drives metamorphic systems as examples of heat flow models.

2.5.1 ***Thickening of a radioactive crust.***

The temperature of a body of rock within the crust with internal heat production depends on four factors: (i) the internal heat production and its distribution with depth; (ii) the flow of heat into the crust at the Moho; (iii) the depth of the rock below the surface and (iv) if the rock mass is deforming, the thermal Peclet number, $Pe^{thermal}$, which is the ratio of the velocity of the rock to the rate of heat transport by thermal conduction and is given by

$$Pe^{thermal} = \frac{vL}{\kappa} = \frac{vL\rho c_P}{k} \qquad (2.7)$$

where v is the velocity and L is a characteristic length scale. These four factors are in general poorly constrained; some discussions are given by Turcotte and Schubert (1982) and Sandiford and McLaren (2005). The history of temperature evolution is commonly treated as a function of the thickening and thinning of the crust.

As an example of the effect of internal heat production and its distribution with depth we present Figure 2.12 taken from Jamieson et al. (1998). Three different heat distributions of K, U and Th are considered for a crust of thickness 35km and heat flow from the mantle of 30 mWm^{-2}. These three different assumptions result in three different temperature gradients and three different temperatures at the Moho.

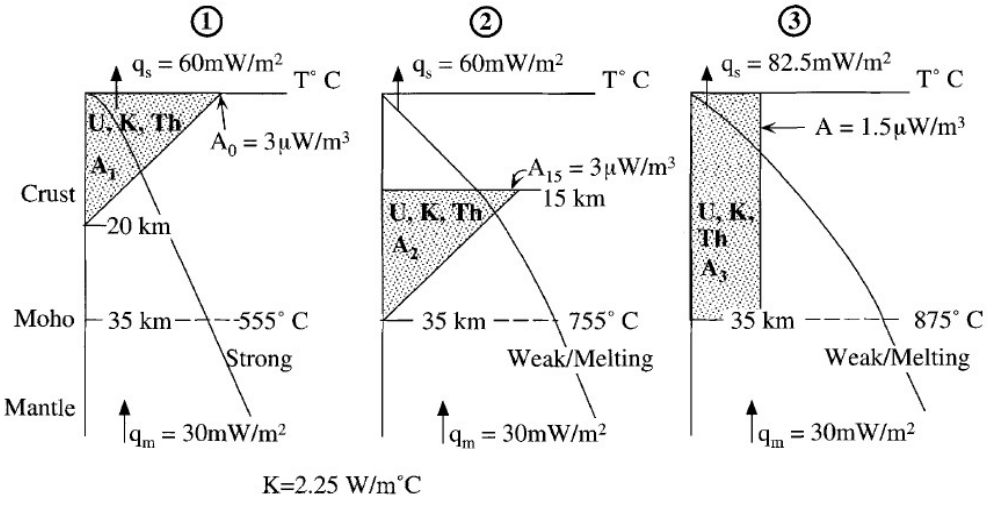


Figure 2.12. Different models for incorporation of internal heat production isotopes in the crust. The crust is 35 km thick and the heat flow from the mantle is 30 mW m^{-1} K^{-1} in each model. (1) Heat producing elements are distributed linearly in the upper 20 km of crust with a heat production of 3μWm^{-3}at the surface. This represents a situation where heat producing isotopes are concentrated in igneous intrusions high in the crust. Surface heat flow is 60 mWm^{-2} with a temperature at the Moho of 555°C. (2) Heat producing elements are distributed

linearly in the lower 20 km of crust with a heat production of 3μWm^{-3}at a depth of 15 km. This represents a situation where heat producing isotopes are concentrated in an old subducted complex. Surface heat flow is 60 mWm^{-2}, with a temperature at the Moho of 755^{o}C. (3) Heat producing elements are distributed uniformly throughout the crust with a uniform heat production of 1.5μWm^{-3}. This represents a situation where heat producing isotopes are distributed within an accreted complex. Surface heat flow is 82.5 mWm^{-2}, with a temperature at the Moho of 875^{o}C (From Jamieson et al., 1998).

The modelling of England and Thompson (1984) is the classical work for understanding the thermal evolution of a thickening crust. The example we present in Figure 2.13 is similar to their models in that it involves a one dimensional situation but where crustal thickening results from the superposition of new igneous and sedimentary units on a crust with internal heat producing radio-nuclides followed by erosion of these units. We begin (Figure 2.13 a) with a crust 35 km thick with a lower crust (5km thick) comprised of mafic granulites with an internal heat production of 1μWm^{-3}. This is overlain by a 30km thick sequence of feldspar rich sediments with a uniformly distributed heat production of 1.5 μWm^{-3}. The heat flux from the mantle is 20 mWm^{2} (Figure 2.13a). The model results in a steady state temperature at the Moho of 683 ^{o}C.

The sequence is thickened by another 10 km due to thrusting and folding; the material comprising the thickened volume has an internal heat production of 1.5 μWm^{-3} and is added instantaneously as on the left of figure 2.13 (a). Erosion of the thrust pile then takes place in six steps at an average erosion rate of 0.5 mm y^{-1}. The resulting temperature evolution for a point in the crust just above the Moho is shown in red in figure 2.13 (b) and the clockwise PT*t* history for this point in figure 2.13 (c); the equivalent history for a point 10 km above the Moho is shown in blue. Notice that the time taken for this evolution is approximately 200 my. The convention is adopted that PT*t* paths are represented on diagrams where the pressure, P, is the axis of ordinates and the temperature, T, is the axis of abscissae. The paths shown in Figure 2.13 (c) are, by this convention, clockwise. In Figure 2.13 (c) the conditions for production of granites by dehydration melting are shown with the PT*t* paths superposed. The granite types follow classification proposed by Wyborn et al. (2002):

- Group 1: These are relatively low temperature granites inferred to form early by the break-down of quartz, albite, K-feldspar and water at the source: biotite and hornblende are common within these melts and there are abundant xenoliths.
- Group 2: consists of the fractionated granites that are low in incompatible elements. These granites have rare amphiboles and it is proposed that in their source regions dehydration melting of biotite has occurred.
- Group 3: The Subgroup 3_1 granites are probably formed by dehydration melting of F-enriched biotites, whilst Subgroups 3_2 and 3_3 may involve dehydration melting of amphibole. Temperatures of formation of some of the Group 3 granites may have been as high as 1000°C.

The systematic progression through these groupings indicates that the source region temperature has increased progressively with time as indicated in the model.

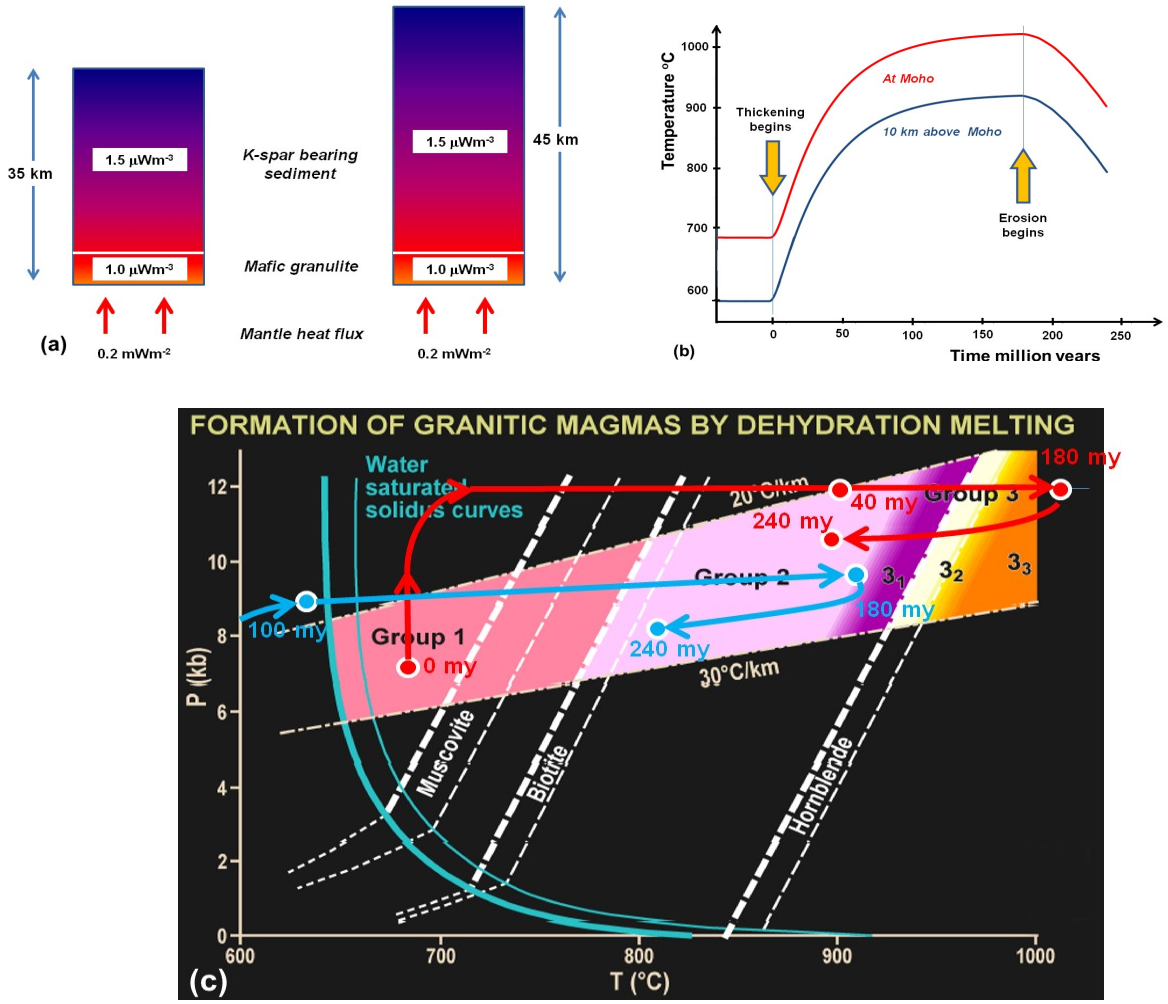


Figure 2.13. One dimensional model of PT changes associated with thickening of the crust by thickening. (a) Details of the model. (b) Temperature-time plot for two levels in (a). (c) Fields of granite melts corresponding to the history shown in (b). Heat flow into base of model is 20 mWm^{-2}. Internal heat production in the crust is shown in (a).

2.5.2 *Enhanced heat production due to radioactive decay.*

We take as an example here the work of Jamieson et al. (1998) where two dimensional models of a subducting crust are examined and in which various spatial distributions of heat producing concentrations are incorporated. The models are shortened with various rates and characteristic lengths which lead to thermal Peclet numbers varying from about 50 to 5. We present the results for one of these models in Figure 2.14 where $Pe^{thermal} = 16$. Crustal and lithospheric heat producing regions, (A_3 and A_2 in Figure 2.14 b) each with $H = 1.5$ μWm^{-3}, correspond respectively to an accreted wedge and a previously subducted wedge of material. The resulting deformation is shown in figure 2.14 (a) with a crustal shear zone developing in the hottest region. The PT*t* histories of various material points are shown in Figure 2.15 together with the Lagrangian trajectories of these points. In some instances material points have been transported over 100 km. Adopting the convention proposed by metamorphic petrologists, these PT*t* paths are clockwise. Notice that this approach is Lagrangian; although the PT*t* paths record the total displacement history of each material point, much of this history is not reflected in the deformation fabrics since much of the displacement history involves rigid body displacements.

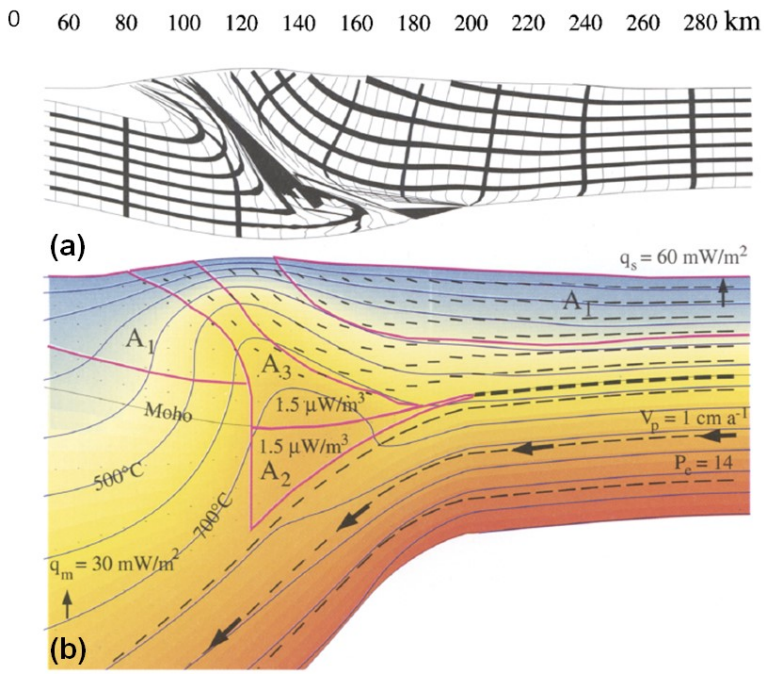


Figure 2.14. Deformation of crust and lithosphere during subduction at 10 mm a^{-1}. Initial radioactve decay sites with elevated heat production of 1.5 μW m^{-3} are marked as A_2 and A_3. Mantle heat flow is 30 mWm^{-2}. (a) Deformed grid showing development of shear zone in the crust. This corresponds to the site of elevated temperatures as shown in (b). From Jamieson et al. (1998).

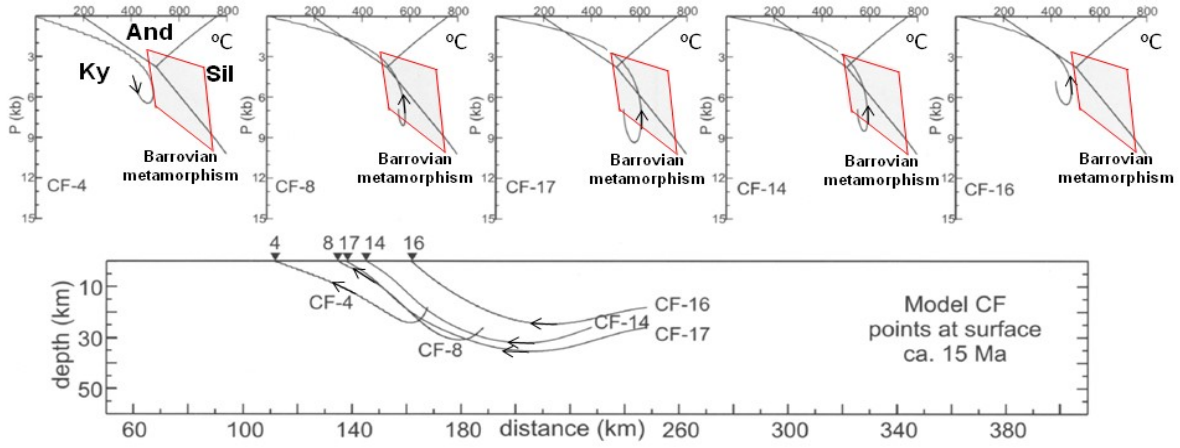


Figure 2.15. PT*t* paths corresponding to figure 2.13. The lower frame shows the trajectories of various material points during 15 Ma of deformation relative to undeformed coordinates. The top frames show the corresponding PT*t* paths in PT space with the field of Barrovian metamorphism marked along with the aluminosilicate stability field. Notice that according to the convention discused in the text these are clockwise paths. Because of the high thermal Peclet number temperature changes are quite rapid corresponding to thermal transport dominantly by advection. From Jamieson et al. (1998).

2.5.3 *Delamination.*

Delamination either in the form of Rayleigh-Taylor instabilities or as "peeling-off events" or combinations of both (Houseman and Molnar, 1997; Elkins-Tanton, 2007; Gorczyk et al., 2013) is a mechanism for introducing hot aesthenopheric material into a position close to the base of the crust. The time scale for heat conduction is then reduced from that associated with a relatively thick lithosphere (say 100 to 200 km thickness and a time scale of 316 my to 1265.8 My) to that associated with a crustal thickness of perhaps 40 km (a time scale of 50.6 My). PT*t* paths are mainly clockwise but some can be anti-

clockwise. An example of a delamination sequence is shown in Figure 2.16 (a) and the resulting PT*t* paths in Figure 2.16 (b). The evolution of topography is shown in Figure 2.16(c). Other examples of delamination are given in Chapter 4.

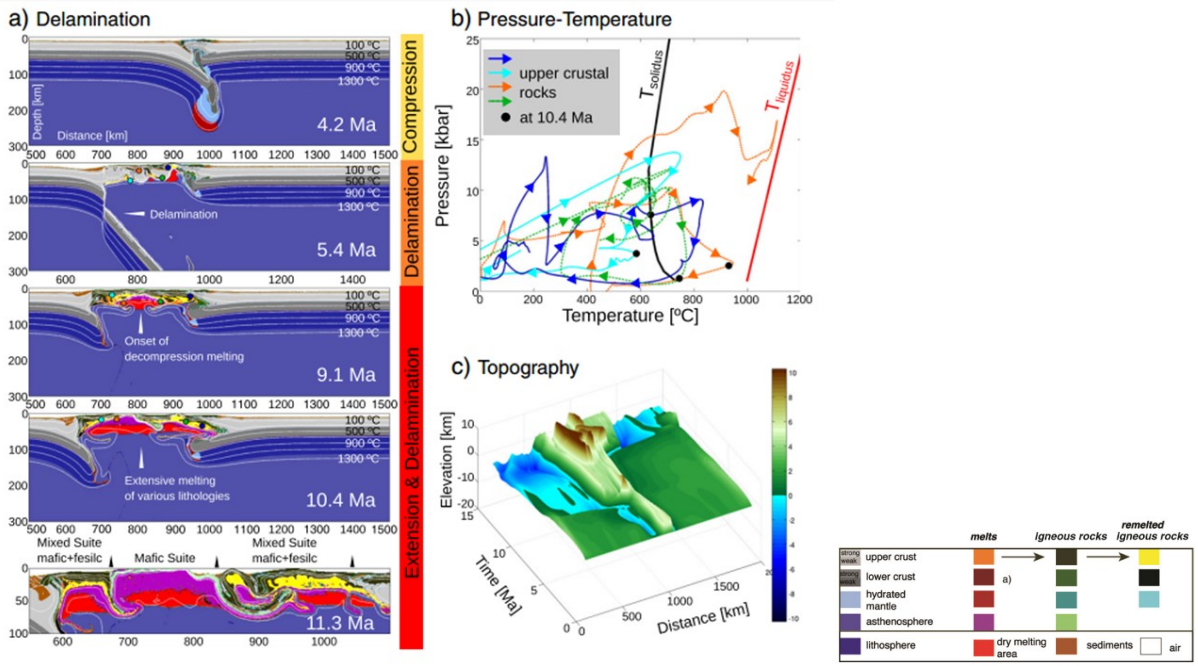


Figure 2.16. Delamination–mechanical removal/peeling away of the lithosphere. (a) Dynamic evolution of the lithosphere after 6 Ma of convergence at a rate of 2 cm/year and an initial lithospheric thickness of 100 km. Detachment of dense lithosphere develops an asymmetric geometry and allows for hot inflow of asthenospheric material. Hence, decompression melting is triggered and results in ultrahigh temperatures at shallow levels and multiple melting and remelting events of crustal and mantle derived rocks. The PT*t* paths of upper crustal rocks are shown in (b). Notice that all paths are clockwise. (c) Topographic evolution with time. The initial uplift caused by compression is overprinted by a strong topographic response above the melt zone. Areas distant from the major melting event are characterized by the development of deep intra-continental basins. The colour legend is shown to the lower right. From Gorczyk and Vogt (2015).

2.5. Heat produced in hydrothermal systems.

The time scale involved in producing mineralised systems is at least an order of magnitude smaller than in metamorphic systems; 1 to 2 million years seems an upper limit for the formation of a mineralised hydrothermal system (Silberman et al., 1972; Goldfarb et al., 1991; McInnes et al. 2005 a, b; Simmons and Brown, 2006; Fu et al., 2010; Yardley and Cleverley, 2016) whereas 10 to 40 million years seems a lower limit for metamorphic systems (Connolly and Thompson, 1989). The importance of this issue becomes apparent when we consider heat production/absorption rates within hydrothermal systems arising from exothermic/endothermic mineral reactions. In metamorphic systems the rates of release or absorption of heat from mineral reactions integrated over time scales greater than a few million years is comparable to or less than the background radiogenic crustal heat production (Connolly and Thompson, 1989). In hydrothermal systems the heat production rates integrated over time scale less than or equal to a million years exceed the background

radiogenic heat production rates and have an important influence on temperature and hence on reaction rates and on fluid pressure.

In order to emphasise the importance of time scale in controlling the evolution of hydrothermal systems we present in Figure 2.17 a linear, uncoupled model of heat generated by an alteration system. In many discussions of hydrothermal systems the system is assumed to be isothermal and that no chemical reactions occur that act as sinks or sources for heat and fluids. For many geological situations, and especially those involving hydrothermal systems, these assumptions are not valid; the system may be open to fluxes of both heat and fluids and heat and fluids may be generated or consumed internally to/from the system as in hydration/devolatilising or melting reactions. Or fluid might be added to/consumed in the system from external/internal sources and sinks and heat generated within the system as in retrograde metamorphism (Haack and Zimmermann, 1996).

One of the fundamental controlling factors in the formation of a mineralising system is the initiation of alteration mineral reactions such as the conversion of K-feldspar to sericite or of mafic minerals to chlorite. Most, if not all, reactions of this type, involving the hydration of silicates, the dissolution of quartz, and the deposition of carbonates and iron oxides are exothermic. These temperature changes arising from exothermic reactions tend to be neglected in the study of metamorphic reactions and are commonly dismissed as unimportant (Connolly and Thompson, 1989) since the heat production rates integrated over the lifetime of the metamorphic system are small compared to the crustal radiogenic heat production rates. However in hydrothermal systems the effects cannot be neglected since the time scale involved in the evolution of a hydrothermal system is at least an order of magnitude smaller than for metamorphic systems. Some implications of this are discussed for retrograde metamorphism by Haack and Zimmermann (1996).

As an example of the influence of these exothermic reactions we present in Figure 2.17 the thermal evolution resulting from the alteration of quartz diorite to a quartz-sericite assemblage. We start with the assumption that the reaction is continuous in time and that the reaction rate is independent of temperature and of the supply of fluid. The thermal budget involved in this alteration process is shown in Table 2.2. Some reactions are exothermic whilst others are endothermic. The net result of the set of reactions is a heat output from the system of 7.68 x 10^8 J m^{-3} of rock over the lifetime of the alteration process.

Table 2.2. Heat produced during hydrothermal alteration of quartz diorite to quartz sericite at 300^{o}C (after Norton and Cathles, 1979).

Minerals in original rock	Weight %	Mass of minerals destroyed (kg/100kg rock)	Heat of hydrolysis (x 10^8Jm^{-3} of rock)
Albite	38	34	-2.034
Anorthite	17	17	-4.52
Quartz	16	-30	-2.825
Annite	11	10	-1.356
K-feldspar	18	-11	0.565
Muscovite	-	-20	3.277
Pyrite	-	-1	-0.791
Totals	100	-49	-7.684

In order to understand what this heat production means for the temperature history of the alteration system we need to know how long the alteration assemblage takes to form and if the reactions continue at constant rate or start rapidly and dwindle with time. If the reaction lasts 10^3 years and is uniform in rate throughout, the heat production rate for the system in Table 2.1 is a massive 2.43 x 10^{-3} W m^{-3}. If the reaction lasts 10^6 years and is uniform in rate throughout the heat production rate is 2.43 x 10^{-6} W m^{-3} which is comparable to typical crustal heat production rates of 1 to 8 x 10^{-6} W m^{-3}. If the reactions begin rapidly and dwindle with time the heat production rate is higher. As indicated above, mineralised hydrothermal systems are thought to form in a time period ≤ 1 million years whereas metamorphic systems form in a time period ≥ 10 million years. Thus for a hydrothermal system the heat production rate represented by Table 2.1 is larger than background crustal radiogenic heat production rates if the exothermal alteration part of the system forms in less than 1 million years.

Finite difference modelling for systems that span the range of potential heat production scenarios for the quartz-diorite alteration system show the magnitudes of the temperature changes involved in hydrothermal alteration systems. The crustal system is taken to be 45 km thick and 100 km long (Figure 2.17 a) with an internal background radiogenic heat production rate of 1 x 10^{-6} W m^{-3}. A heat flux from the mantle of 0.02 Wm^{-2} is applied at the Moho. This gives a steady state thermal gradient through the crust of ~18.2 C^o km^{-1}, a Moho temperature of 820°C and a surface heat flux of 0.065 Wm^{-2}. These values represent an average modern crustal situation. An alteration system 10 km x 10 km in size corresponding to the quartz-diorite system in Table 1 is initiated at mid-upper crustal levels as shown by the black square in Figure 2.17 (a). This corresponds to a large alteration system but is smaller than those proposed beneath Olympic Dam and similar areas on the basis of seismic and magneto-telluric surveys (Drummond et al., 2005; Heinsen et al., 2006; Maier et al., 2007). The time for the alteration system to reach completion is modelled at 10^4, 10^5 and 10^6 years (Figures 2.17 b). At the end of these time periods the alteration reaction is assumed to be complete so that heat production rates within the alteration system are returned to the ambient crustal value and the system temperature is left to decay towards the crustal background value. The temperature histories are recorded in Figure 2.17 (b). These models assume all heat transport is by conduction with a thermal diffusivity of 10^{-6} m^2 s^{-1}.

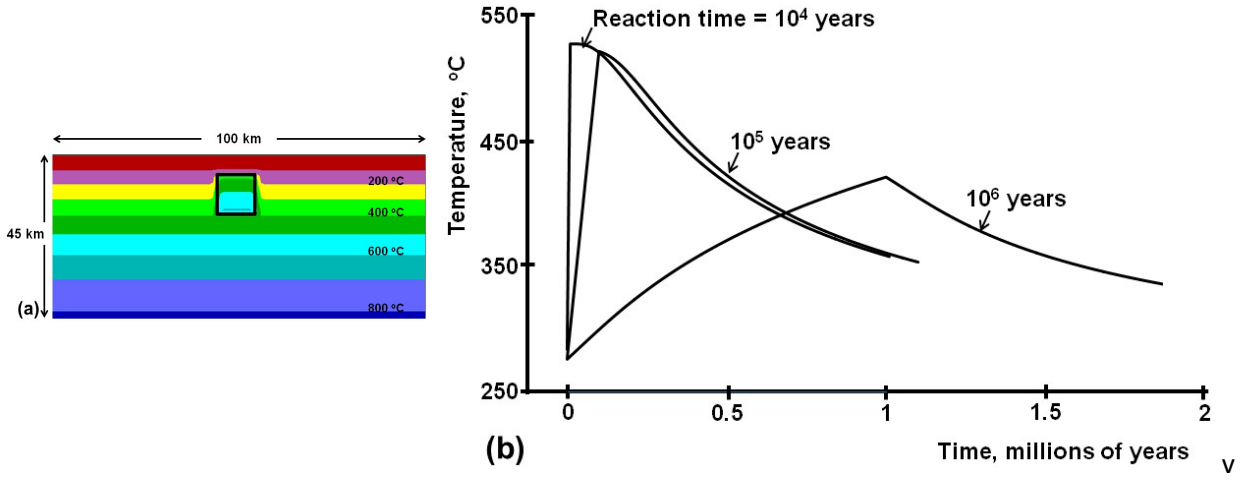


Figure 2.17. Temperature changes arising from exothermic reactions in the conversion of a quartz diorite to a muscovite-quartz assemblage. (a) The model geometry. The reactions in Table 1 take place in a 10 km x 10 km

region as shown by the black square. This model assumes the reaction takes place in 10^4 years and the temperature distribution is shown at the end of the reaction time. For other details see the text. (b) Results for three reaction times, 10^4, 10^5 and 10^6 years. After each reaction completion the system cools for 10^6 years as shown.

We see that if the time for the alteration reaction to proceed to completion is 10^4 years then the peak temperature in the alteration system is 220°C above ambient and is still 80°C above ambient 1 million years after the completion of the reaction. If the time to completion is 10^6 years then the peak temperature is 140°C above ambient and is still 60°C above ambient 1 million years after completion of the alteration reactions. With the heat budget proposed in Table 2.1 the time for completion of the alteration reactions needs to exceed 10^6 years before the heat effects become unimportant; this is a time scale more relevant to metamorphic systems rather than hydrothermal systems. As we have indicated the possibility of such a long time period for the formation and completion of a hydrothermal alteration system seems marginal. We emphasise that these model results assume no coupling between deformation, mineral reaction rates and the supply/consumption of heat. We show later in Chapter 8 that if such coupling is included then the high temperature rises shown in Figure 2.17 do not occur and instead the system oscillates in a chaotic mannre between two temperatures defined by the kinetics of the mineral reactions involved.

Recommended reading.

Boyce, W E., and DiPrima, R. C. 2005. *Elementary Differential Equations and Boundary Value Problems*. John Wiley. A readable treatment of the mathematics involved in the solution of heat transport problems.

Carslaw, H. S., and Jaeger, J. C. 1959. *Conduction of Heat in Solids*. Oxford University Press. This is the definitive text on thermal transport. It contains a large number of solutions for specific boundary conditions and system geometries.

Turcotte, D. L., and Schubert. G. 1982. *Geodynamics*. Chapter 4 in this book is an important treatment of heat flow in geological systems

Yuen, D. A., Maruyama, S., Karato, S., and Windley, B. F. 2007. *Superplumes: Beyond Plate Tectonics*. Springer. This publication contains many excellent summaries of the nature of superplumes, their origins and evolution.