PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Heat Transfer and Convection Currents

D. C. Tozer

Phil. Trans. R. Soc. Lond. A 1965 **258**, 252-271
doi: 10.1098/rsta.1965.0038

XXI. Heat transfer and convection currents

By D. C. Tozer

Department of Physics, University of Newcastle upon Tyne

'Es gibt nichts praktischeres als die Theorie'

(LUDWIG BOLTZMANN—Vorlesungen 1895)

The possible causes of convection in the Earth's mantle are examined, and it is concluded that radiogenic heating together with thermal conduction is most likely to provide the driving force for any convection present at this stage of the Earth's history.

The theory of convection in a medium with internal heat generation is discussed semi-quantitatively. It is concluded that a mantle convection pattern based on calculations of marginal stability in an incompressible fluid is a gross oversimplification of the real situation. A more realistic theory indicates that thermal convection is probably confined to the outer mantle and is an unstable flow. There is evidence that it would have a character that would explain the narrowness and linearity of the pattern presented by the ocean ridges and some orogenic belts.

Notation

C	specific heat (unspecified type)
C_p	specific heat at constant pressure
f	fraction of convected heat generated within region of mantle convection
$\mathbf{f}$	vector function describing the spatial variation of velocity
g	acceleration due to gravity
h	distance that convection penetrates a region of sub-adiabatic temperature gradient
H	radioactive heat generation/unit volume
J_2	coefficient of second axial harmonic in series for the Earth's external gravitational potential
k	thermal conductivity
L	a scaling factor for length
P	pressure
Pe	Péclet number
Pe_c	Péclet number for transition to unsteady flow
Pr	Prandtl number
Pr′	Prandtl number for which shear flow instability and Péclet instability occur simultaneously
r	radius
Ra	Rayleigh number
Ra_c	Rayleigh number for onset of convection
Ra′	Rayleigh number for onset of unsteady convection
Re	Reynolds number
Rh	parameter characterizing slow convection in system with internal heat generation
Rh_c	value of Rh at which convection begins

Rh′ value of Rh at which unsteady convection begins
S entropy
t time
T temperature
$\left(\frac{\partial T}{\partial r}\right)_s = \frac{g\beta T}{C_p}$ adiabatic temperature gradient
$\mathbf{v}$, v velocity
z co-ordinate in direction of gravity
β coefficient of expansion
η coefficient of viscosity
κ thermal diffusivity
ν kinematic viscosity
ρ density
τ_M, τ_T characteristic times
χ compressibility
ω angular velocity of the Earth

In recent years there has been a considerable growth of evidence supporting continental drift, but this hypothesis has remained largely empirical in nature and has not been very successfully related to the general body of physical knowledge about the Earth. It is still a fashion among many geophysicists to treat the subject entirely empirically and to reject any attempt to find a physical basis for continental drift as valueless. If one remembers the impossibility of proving by subsequent observation that a certain unitary event or sequence of events did or did not occur, the limited returns of such an outlook will be obvious. Besides, a purely empirical approach does not normally lead to an efficient solution of scientific questions. More data are hardly likely to remove all the difficulties that exist in connexion with continental drift and the hypothesis will gain wider acceptance if theory can remove the objections that have been raised and suggest a plan of observation.

Theories of continental drift may be divided into two classes depending on whether they do or do not involve convection in the earth. At the present time, theories in the former class are the more favoured since they do not conflict strongly with any notions that may be reasonably held about the physical properties and state of the Earth's interior. On the other hand, theories that attribute continental drift to a general expansion of the Earth have an *ad hoc* nature since they require either rather fantastic physical properties for the major constituents of the planet, or a premature meddling with the foundations of physics. They should therefore be dismissed until their assumptions can be independently justified or the inadequacy of convection theories demonstrated.

The adoption of a convective explanation of continental drift immediately raises the question of the origin of the forces driving the motions. No entirely satisfactory answer can be given to this at the present time, but it is widely accepted that the actively driven region lies within the Earth's mantle and that the motions at the Earth's surface may be passive in the sense that they arise either from hydrodynamic drag on a relatively thin and much more solid surface layer, or by the gravitational spreading of surface features produced by the underlying convection. The driving forces could arise from chemical or physical inhomogeneity. Examples of the former are the fractionation of sialic minerals to form the crust and the separation of iron to create the core; physical inhomogeneity is created by radioactive heating combined with thermal conduction. An attempt to decide which is the most important of these possibilities may be made by considering the energy released in each process. Making the common assumption that chondritic meteorites represent a fairly good approximation to the earth's average composition, one may demonstrate that radioactivity will have produced approximately 10^{31} J throughout the earth over the past 4×10^9 years, and that the present rate of heat loss from the earth is in balance with radioactive heat production. About 10^{31} J have also been released during the formation of the Earth's core if the iron were initially uniformly distributed. If one examines core formation in some detail (Tozer 1965), it is difficult to avoid concluding that this gravitational energy has been dissipated uniformly throughout the fractionating region by viscosity. Current theories discussing the kinetics of core formation (Runcorn 1962) (Munk & Davies 1964) have been based primarily on geometrical considerations. However, if one takes into account the variation of physical parameters, in particular viscosity, that result from the viscous heating, one may reasonably conclude that core formation can easily have 'run away' and have been virtually completed in an extremely short and remote geological time. The smallness of the non-hydrostatic components of gravity and the balance of radioactive heat generation with the geothermal flux support this view. This balance is readily explained if convection driven only by radioactive heating allows the heat generated at great depths to reach the surface in times of the order of 10^8 to 10^9 y, rather than the 10^{10} to 10^{11} y required by conduction theory.

The energy released by the separation of sialic minerals from the whole mantle to form the Earth's crust is about 3×10^{29} J, which is about 10^4 J/g of crustal material. There is uncertainty in this figure due to the effect of phase transformations (negative correction), chemical reactions (positive correction?)† and a change of thermal conditions near the surface, but this is hardly likely to produce more than a 10 % error in this estimate of the energy released. This energy is about 2·5 % of the energy released by core formation or radioactivity in 4×10^9 y and from this we may conclude that the formation of the crust can at most have had a dominant effect on convection for very short periods of geological time or in small regions of the mantle. Of course, it may be plausibly argued that orogenic activity and continental drift has been episodic and that the activity today and in the past has been concentrated in belts. No argument based on energy has been found against attributing all orogenic and seismic activity to convection currents confined to these belts and driven exclusively by the energy released by the gravitational separation of the crust. However, that would mean that the forces shaping the surface of the Earth are a relatively minor part of those causing the evolution of the outer parts of the mantle, and that almost everywhere little connexion exists between heat flow and convection currents.

The strength of the argument that crustal fractionation is entirely responsible for the general pattern of activity would be weakened if it could be demonstrated that the convection arising from radiogenic heating has the required episodic and localized character.

† Phase transformations will mainly proceed in the sense that gives expansion and absorption of latent heat; chemical reactions like the serpentization of olivine and basalt formation from peridotite are exothermic.

A clue that some other process is involved besides crustal fractionation may be found by noting that the pattern of activity has changed on a number of occasions throughout geological time and has occurred at approximately the same place on more than one occasion. The processes involved in chemical fractionation do not contain the time explicitly (unlike radioactivity) and it is difficult to see how the decay of a gravitationally unstable state would lead to a quasiperiodic pattern of activity without some process such as radioactivity to reactivate stable states of the mantle.

It seems convenient, therefore, to divide the convection problem into two parts, although from a physical point of view it is a rather arbitrary division. The first deals with the initiation and character of mantle convection under the influence of radiogenic heating, while the second is concerned with the effects of crustal fractionation and the chemical inhomogeneities created by it in the mantle. It is emphasized that once any convection has begun, chemical inhomogeneity may be created and may play an important role in determining subsequent behaviour.

Most of this paper is devoted to a study of mantle convection using a model of the mantle in which it is assumed that the convecting region is chemically homogeneous down to a reasonably small scale, and that the radiogenic heating is not appreciably augmented (or decreased) by the free energy of chemical reactions or gravitational differentiation occurring as a result of any convection.

Convection in a medium with internal heat generation

Some recent writers have proposed extremely simple pictures of the convective motions in the mantle that are based on the marginal stability theory of convection for incompressible fluids in spherical shells. The only factor determining the pattern of convection in this theory is the ratio of the core and surface radii. Such theories are almost as unconnected to our knowledge of the earth's physical properties as the continental drift they are meant to explain. They rest on a precise but very limited mathematical theory, great simplification of the mantle properties and the *ad hoc* hypothesis that the core has been growing continuously throughout geological time. A slightly more detailed look at the factors influencing mantle convection reveals a very complicated problem and we would indeed be fortunate if the use of such a simple theory could ever be justified. Apart from the difficulty of theoretical justification, MacDonald (1963) has drawn attention to a number of geophysical facts that do not fit comfortably into such theories.

A considerable drawback in discussing mantle convection at the present time is the absence of experimental data on systems that contain the essentials of the mantle convection problem. Almost all work on convection has been concerned with the flow that occurs when fixed temperature gradients are applied to the fluid (e.g. Bénard cell experiments) or when hot bodies are immersed in a cooler fluid. In the case of the mantle a close approximation to the thermal boundary conditions is made by assuming that the outer surface temperature is fixed and that any convected heat is being generated within the mantle. No experimental work appears to have been done on fluid systems with appropriate geometry and internal heat generation.

The fraction f of the heat being convected in the outer mantle that comes from within the region of convection obviously depends on the depth to which convection extends, which is itself dependent on the level of heat generation. The highest estimates for the heat entering the mantle from the core assume that the core motions are also a thermal convection, and give about a tenth of the surface heat flow. There are geochemical difficulties in finding this amount of radioactive heating in the core (Verhoogen 1961) and other forces have been suggested to drive the core motions. The heat input to the mantle is then the average dissipation rate in the core, which is estimated to be 10^{11} J/s (Bullard & Gellman 1954); this is less than 1 % of the surface heat flow. Since this is so small compared with the heat generated in the mantle, an estimate of f can be based on the ratio of the convecting and non-convecting masses of the mantle. Values of f found in this way assume quasi-steady thermal conditions and should probably be regarded as upper limits, since they disregard the heat input from the core and take no account of the preferential loss of radioactive material from the convecting upper mantle to the crust.

The differential equations governing thermal convection in the mantle, treated as a chemically homogeneous fluid medium, are five in number—the equations of motion, heat transfer and mass continuity.† Formal solution of these equations is at present impossible on mathematical grounds, but in any case unnecessary to a geophysicist wishing to understand only those major features that survive in the geological record or dominate the present pattern of activity. Rather what is required is a demonstration of the adequacy of the assumed driving forces and a general description of the character of convection in space and time. We shall see that by making a number of assumptions and utilizing relevant geophysical data, the equations can be greatly simplified and these more modest objectives brought within reach of solution.

The assumptions that are made may be divided into two groups, that refer respectively to the form of the differential equations and their solution. In the first group are the Boussinesq approximations appropriate for the mantle convection:

(*a*) The convecting region may be treated as incompressible.

(*b*) Except for the density and temperature differences responsible for buoyancy forces, the convecting region is homogeneous.

(*c*) The convecting region is close enough to the convectively stable state for the second and higher powers of the density and temperature differences from the stable state to be neglected.

(*d*) The driving forces (internal heat generation and heat conducted to and from the convecting region) may be assumed quasi-steady.

We now attempt to justify the assumptions (*a*) and (*b*); (*c*) and (*d*) will be discussed in connexion with the solution of the convection equations.

For convective instability problems of the type in which convection is the result of temperature differences impressed across a fluid (e.g. Bénard cell convection), Jeffreys (1930) has shown that provided we measure temperature and density with respect to an adiabatic i.e.

† Poisson's equation for the gravitational potential is omitted because the density differences associated with mantle convection are sufficiently small to neglect the effects of self gravitation. (Gravity data indicate that large scale density differences are no greater than 1:10^4 on a level surface in the mantle.) Geophysical calculation indicates that the acceleration of gravity is virtually constant in the mantle; a value of $g = 10^3$ cm/s^2 is taken for this 'externally' generated gravity.

isentropic distribution of these variables, the theory approximates to that for an incompressible fluid. Knopoff (1964) has examined this approximation for a Bénard cell problem in which the compressibility and thickness of the fluid layer are given values quoted for the mantle. He concludes that the Jeffreys approximation gives the correct answer to the stability calculation within 5 %. Similarly it may be shown that a formal solution of the present problem is to use an approximating adiabatic distribution as a fiducial distribution. The theory of convection in an incompressible medium is used with a heat source distribution given by the difference of the real heat source distribution and that which would maintain the fiducial temperature distribution.† Unfortunately, the full effects of compressibility are not as easily understood for a system as large as the mantle with internal heat generation. The difficulty arises because with any reasonable distribution of the heat sources only parts of the system have superadiabatic temperature gradients, so that the effect of compressibility is tied with the question of how far into the mantle convection extends. There is also difficulty in knowing which adiabatic distribution to use as a fiducial distribution.

Let us first consider what is known of convection in incompressible fluids contained within spherical shells or spheres. We suppose that there is radial gravity, quasi-steady and uniform internal heat generation, and that the outer surface temperature is fixed. One might predict that since more heat flows in the outer parts of the system, convection will begin there and spread inwards as the heating rate is slowly increased. However, such a simple argument ignores the greater inhibiting effect of viscosity and thermal conductivity on small scale convection, and the mechanical forces the supposed partial convection would exert on the non-convecting region. In fact, the marginal stability calculations show (Chandrasekhar 1961) that the whole of the fluid begins to convect as the heating is increased above a certain threshold.

Do these calculations mean that all the mantle is involved in any convection? It is important to realize that the strictly incompressible fluid considered in these calculations is either neutrally stable or unstable with respect to any 'external' mechanical disturbance, for all values of the heat generation; there are no mechanically or convectively stable states corresponding to subadiabatic temperature gradients in a compressible fluid. Therefore, the liquid does not oppose any extension of partial convection, which would decrease the effect of dissipation. For systems as big as the mantle the stabilization provided by a subadiabatic temperature gradient is important (see below).

A reasonable approximate solution to the problems presented by compressibility is now outlined. It involves either explicitly or tacitly all the assumptions stated above.

(i) Use steady-state conduction theory to find the temperature gradient at any depth and compare this with the local adiabatic temperature gradient.

(ii) Apply convection theory for a uniform incompressible fluid with internal heat generation to those parts of the mantle that have both super adiabatic temperature gradients and a reasonable homogeneity of the physical parameters appearing in the convection equations. (The temperature and density variables at any depth are to be measured with respect

† As with the case considered by Knopoff, the adequacy of this approximation depends on the size of the system, L. This should be small enough for $\rho g \chi L \ll 1$. If one puts appropriate values of $L = 10^8$ cm, $\rho = 4g/\text{cm}^3$, $g = 10^3\ \text{cm/s}^2$, $\chi = 5 \times 10^{-13}\ \text{cm}^2/\text{dyne}$, $\rho g \chi L = 0{\cdot}2$.

to a state obtained by adiabatic compression of the material at the top of the convecting region to the pressure existing at that depth.)

(iii) The heat source distribution effective in driving convection is the actual heat source distribution minus that distribution which will, with the heat conducted to and from the region, maintain the adiabatic distribution described in (ii).

In applying this procedure one assumes that thermal conduction in the mantle may be treated by a linear theory. Otherwise the modification of the conduction theory temperature distribution due to convection, alters the computed conduction gradient everywhere which would necessitate a successive approximations technique to locate the boundary of the convecting zone. Fortunately, it appears that the non-linearities are small enough and the factors responsible for limiting convection radially are such that only small changes are introduced by recalculation.

The validity of this procedure depends on the convection not being strong enough to drive motions far inside a region with a subadiabatic temperature gradient. An estimate of this distance of penetration h into a stable region may be made by equating the hydrodynamic force acting on the boundary to the restoring buoyancy forces set up when disturbing the thickness h. We have

$$\eta v/L \sim g\delta\rho h, \tag{1}$$

$$\delta\rho \sim \frac{\mathrm{d}}{\mathrm{d}r}\left[\left(\frac{\partial T}{\partial r}\right)_s - \frac{\mathrm{d}T}{\mathrm{d}r}\right] h^2\beta\rho, \tag{2}$$

$$h \sim \sqrt[3]{\left[\frac{\nu v}{\beta L g}\left\{\frac{\mathrm{d}}{\mathrm{d}r}\left[\left(\frac{\partial T}{\partial r}\right)_s - \frac{\mathrm{d}T}{\mathrm{d}r}\right]\right\}^{-1}\right]}. \tag{3}$$

If we first consider the situation when the boundary of the convecting region lies outside a region of major phase change, it may be shown that the derivative of the adiabatic temperature gradient is small compared with that of the conduction gradient $\mathrm{d}T/\mathrm{d}r$. We then have

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\left(\frac{\partial T}{\partial r}\right)_s - \frac{\mathrm{d}T}{\mathrm{d}r}\right] \sim -\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\mathrm{d}T}{\mathrm{d}r}\right) \sim \frac{H}{k}, \quad h \sim \sqrt[3]{\frac{\nu v k}{\beta L g H}}. \tag{4}$$

The values of the quantities in (4) are taken to be (see below) $\nu = 10^{20}\,\mathrm{cm^2/s}$, $v = 10^{-7}\,\mathrm{cm/s}$, $\beta = 10^{-5}/\mathrm{degC}$, $L = 10^8\,\mathrm{cm}$, $H/k = 10^{-13}\,\mathrm{degC/cm^2}$. These give $h \sim 50\,\mathrm{km}$.

In a region of phase transformation the adiabatic temperature distribution still serves as a criterion for separating convectively stable states from the possibly unstable ones.† However, the normal expansion coefficient and specific heat in the expression for the temperature gradient $(\partial T/\partial r)_s$ are effectively modified by the volume change and latent heat associated with the phase change (Tozer 1959). It now appears plausible that the rapid rise of density with depth between 500 and 1000 km inferred from seismological data

† The discussion of Knopoff (1964) on the effect of phase transitions is most misleading. The model he chooses is of a univariant transition, but his restriction that it all occurs at the same fixed depth is only appropriate if the volume change is infinite! The normal situation, even for univariant transitions, is for the latent heat to be supplied by a change of temperature in passing through the phase transition region.

Univariant transitions are peculiar in that they lead to an indeterminate expression for the adiabatic temperature gradient. Their stability can always be determined by comparing adiabatic and actual density gradients.

may be mainly attributed to a number of phase changes in the system MgO–FeO–SiO_2. The information necessary for a precise estimate of the adiabats in this region is not yet available but it seems quite possible that the adiabatic gradient may be several degrees per kilometre and the effective expansion coefficient greater than 10^{-4}/degC. Putting

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\partial T}{\partial r}\right)_s = 10^{-12}\,\mathrm{degC/cm^2} \quad \text{and} \quad \beta = 5\times 10^{-4}/\mathrm{degC},$$

we have $h \sim 6$ km. These estimates of h are so small compared with mantle dimensions that the condition of subadiabaticity of the geothermal gradient provides a good criterion to judge the extent of convection.

Another factor that determines the usefulness of the above procedure and the validity of the assumption (*b*), is the constancy of various parameters to be used in the convection theory (viscosity, expansion coefficient, thermal diffusivity) along the fiducial adiabat. For a chemically homogeneous mantle and outside any regions of major phase change, the variation of the expansion coefficient with depth is sufficiently small to be neglected in the present theory. If, however, the convecting region penetrates a zone of phase change the inhomogeneity of the expansion coefficient may be more important. The variation of thermal diffusivity to be expected along an adiabat is also small enough to be neglected at this stage of the development of a mantle convection theory.

Before we attempt to explain the significance of the viscosity variation with physical conditions, it is necessary to discuss flow mechanisms in nominally solid materials and comment on the use of a Newtonian viscosity in convection theories. There is no justification from solid state physics for the common assertion that solids (including the mantle) have a 'finite strength', meaning a capacity to withstand the application of a finite, non-hydrostatic stress without a permanent deformation after the stress is removed. It follows, without approximation, that because all solids have only finite binding energy, processes exist that give rise to 'creep' with vanishingly small shear stresses. The illusion of finite strength has arisen because there are creep processes contributing to the total creep rate that increase rapidly with stress and therefore suddenly become observable above a certain stress threshold (van Bueren 1961). These creep processes involve dislocation movements and their strong dependence on stress is understood in terms of the forces required to move dislocations and activate dislocation sources. It may be said that many complications of rheological behaviour observed in the laboratory for ionic and covalently bonded solids, which include the mantle minerals, result from the ability of large stresses to build up thermodynamically metastable concentrations of dislocations. The mechanism of creep at low stresses ($<$ *ca.* 10^{-4} to 10^{-5} shear modulus) is of a kind that does not significantly involve dislocations, but the diffusion of individual atoms or ions either within or on the surfaces of crystals. The departure from the equilibrium concentration of these point defects under the low stress conditions is small enough for the creep mechanism to be a linear Markov diffusion process that is phenomenologically described by a Newtonian viscosity. The exponential temperature dependence of these creep processes and their dominance only at low stress means that it is normally only feasible to observe them at temperatures close to the melting point. The values of non-hydrostatic stress in the mantle inferred from satellite gravity data are in the low stress range defined above, and it might therefore be thought appropriate to adopt low stress mechanisms of creep, and hence a Newtonian viscosity, for all mantle convection theory. However, the gravity data give a very blurred picture of the stress distribution and the occurrence of earthquakes is a positive indication that the low stress situation does not occur everywhere in the mantle. In these circumstances, it seems permissible to use a Newtonian viscosity mechanism when discussing the onset of convection, but that we should be careful in applying to the mantle any experimental or theoretical results for finite amplitude convection in a Newtonian fluid. Since little progress can be made by a direct discussion of the equations of convection in a non-Newtonian fluid, the convection theory and experimental results for Newtonian fluids will be reviewed and applied to the mantle convection problem on the assumption that providing we only inquire about the general or averaged characteristics of the flow, the predictions will be approximately correct. The manner in which non-Newtonian behaviour might specifically alter the instantaneous pattern of flow inferred from the Newtonian theory is necessarily discussed qualitatively.

Since one of the main purposes of this paper is to examine the compatibility of the convection hypothesis with other knowledge of the Earth's interior, it seems best to choose a low stress creep mechanism to calculate the Newtonian viscosity that satisfies the following conditions:

(i) It is virtually certain to occur.

(ii) It will provide an estimate of the upper limit of the mantle viscosity.

(iii) It is reasonably easy to evaluate from known physical parameters.

The mechanism that has been chosen is vacancy creep (Nabarro 1948; Herring 1950) in olivine. Theory indicates that the kinematic viscosity due to this mechanism is given by

$$\nu = \nu_0 \, e^{E/kT}. \tag{5}$$

The quantity ν_0 is related to a diffusion process,† and E is approximately proportional to the melting point (Zharkov 1963). Olivine was chosen because it is the most refractory of the common mantle minerals and therefore gives the greatest estimates of this viscosity. It may also be true that the viscosity of the heterogeneous mantle material is determined by its most viscous constituents. The values taken for the parameters are

$$\nu_0 = 10^4\,\mathrm{cm^2/s}, \quad E = 4\,\mathrm{eV}.$$

It is important to note that pressure tends to increase E but that the effect of pressure on ν_0 is relatively unimportant.

A number of interesting facts of considerable significance to developing a convection theory emerges from this model of the creep behaviour of the mantle:

(i) There is an enormous reduction of this viscosity (a factor of *ca.* 10^{40}) with increasing depth in the outer 100 km or so of the earth where the temperature gradient more than compensates for the increase in E with pressure. Of course, few people would maintain that olivine determines the rheological behaviour throughout this region, but an enormous decrease of the viscosity in this range of depths would seem to be typical of any material chosen.

† Ionic conductivity is also a guide to the value of this parameter, but the value of E determined from conductivity data is likely to be systematically low. The estimation of the parameters ν_0 and E is discussed in a paper now in preparation.

(ii) If we examine the variation of this viscosity along an adiabat, but not in a region of phase change, the viscosity variation is relatively small; the amount of variation depends somewhat on the adiabat chosen. The temperature of the hottest lava (*ca.* 1500 °K.) gives some idea of the position of the adiabat that approximates the thermal conditions in the upper mantle and along this, between depths of 100 and 500 km, this viscosity increases by a factor between ten and a hundred.† Although the method of investigating the viscosity variation is not capable of very precise estimates, there are reasons for believing that in the phase transition region of the mantle the viscosity will increase by several powers of ten along this particular adiabat.

(iii) Along the adiabat identified from lava temperatures, the average kinematic viscosity above the region of phase change is about 10^{20} cm^2/s, which is in fair agreement with mantle viscosities estimated directly from geophysical observations (Munk & MacDonald 1960).

From these characteristics of the viscosity variations, it seems a reasonable approximation to divide the discussion of the thermal state of a possibly convecting mantle into two parts. In the convection theory proper, we treat the various parameters as constants, while in the second part we investigate those factors that determine the parameter values.

With these various approximations and assumptions the equations for thermal convection in the mantle, expressed in co-ordinates rotating with the Earth are

$$\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\,.\nabla)\,\mathbf{v}+2\boldsymbol{\omega}\wedge\mathbf{v}=-\frac{1}{\rho}\nabla p+\nu\nabla^2\mathbf{v}-\beta T\mathbf{g}, \tag{6}$$

$$\frac{\partial T}{\partial t}+\mathbf{v}\,.\nabla T=\kappa\nabla^2 T+\frac{\nu}{2C}\left[\frac{\partial v_i}{\partial x_k}+\frac{\partial v_k}{\partial x_i}\right]^2+\frac{H}{C\rho}, \tag{7}$$

$$\nabla\,.\mathbf{v}=0. \tag{8}$$

The heat source distribution H to be used in equation (7) is calculated in the manner indicated above. However, in the regions of the mantle without phase changes, the curvature of the adiabatic temperature distribution is so small that the modification of the actual heat source distribution may be neglected. It is only where convection penetrates an important zone of phase changes that the adiabats become so curved that appreciable modification of the actual heat source distribution may be necessary. It can be shown that if the convection is limited at its lower boundary by the steepening of the adiabatic gradient to a few degrees per kilometre somewhere in the region 500 to 1000 km deep, there is in effect a considerable augmentation of the heat source distribution near this lower boundary. It also follows that if thermal conditions in the convecting zone are unsteady, the effective heat source distribution is also unsteady. We shall not consider further these effects of inhomogeneity and unsteadiness in H, but take H to be sensibly uniform and equal to a third of the value for chondritic meteorites i.e. 3×10^{-8} ergs cm^{-3} s^{-1}. This makes an allowance for the fractionation of radioactive minerals to the crust. There may be systematic differences in H between continental and oceanic mantle but these horizontal

† In a convecting region the temperature gradient is greater than the adiabatic gradient, which reduces or reverses the viscosity variation along an adiabat. This particular viscosity is constant along a temperature curve with about twice the adiabatic temperature gradient.

gradients of H are small enough for the assumption of the uniformity of H to be acceptable in discussing upper mantle convection occurring in any one place.

Equations (6) and (7) can immediately be simplified by examining the relative magnitude of the various terms comprising them. To do this we need to introduce the assumptions referring to the solution of these convection equations. We take values for some of the parameters and variables that have been directly calculated from geophysical observations.

Assumption (*e*). A time average for the velocity of convection is set by the mean rate of movement on transcurrent faults and continental drift as inferred from palaeomagnetism. This gives $\overline{|v|} \sim 10^{-7}$ cm/s. This estimate of $\overline{|v|}$ should be regarded as a lower limit since it is not clear how convection velocities are related to surface movements, but they can hardly be less.

Assumption (*f*). The kinematic viscosity of the convecting region of the mantle to be used in the approximating theory of convection in a Newtonian fluid is of the same order as that found by considering the damping of the variation of latitude as an effect of mantle viscous dissipation.

The variation of latitude only produces stresses in the low stress range defined above. It is believed that a better average value for the viscosity of the least viscous parts of the mantle can be derived from this phenomenon than from glacial uplift, which does depend on the properties of the very inhomogeneous† outer 100 km of the earth. Using a Maxwell solid model of the mantle, one obtains $\nu \sim 10^{20}$ cm^2/s. Munk & MacDonald (1960) criticize the Maxwell solid model on the grounds that a kinematic viscosity as low as 10^{20} cm^2/s would mean the disappearance of all surface features in a few tens of years. However, we have seen that an enormous increase of viscosity occurs towards the surface and unless the surface features extend laterally for several times the thickness of this inhomogeneous layer, i.e. several hundred kilometres, their relaxation time can give no indication of the rheological properties of the convecting region of the mantle. Isostatic compensation is generally good for such large features. It is also most important to realize that viscosities as low as 10^{20} cm^2/s are quite explicable in terms of the properties of mantle minerals (see above).

Assumption (*g*). The convection occurs in the upper mantle in a region that extends approximately from the low velocity layer to about the depth of the deepest earthquakes.

Some of the evidence for this has already been described. It may be summarized by stating that large increases in viscosity are to be expected above and below this region and that the temperature gradient obtained from conduction theory is likely to be less than the adiabatic temperature gradient in the transition zone. We take L, a factor representing the scale of convection, as 6×10^7 cm. With L so small in comparison with the earth's radius, the effects of curvature are negligible and we may treat the region as a plane layer.

If we consider the possibility of convection in the lower mantle, we can recognize three important physical differences from the upper mantle. They are a lower heat flow and a higher thermal conductivity and viscosity. All these factors work against the occurrence of thermal convection in this region. Whether a large fraction of the lower mantle has a sub-adiabatic temperature gradient depends on the amount of heat coming from the core. More than a thousand kilometres of the lower mantle is convectively stable if the only heat supplied to the mantle is the result of viscous and electrical dissipation in the core.

† With respect to viscosity.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Some observational support for concluding that there is at present no convection below the transition region of the mantle may be found from the satellite gravity data and the distribution of earthquakes. It has been pointed out (MacDonald 1963) that the departure of the earth's gravitational field from that predicted by hydrostatic theory has serious consequences for convection theories that include the whole of the mantle in any convective motion. By considering the discrepancy in J_2 to arise from a lag of the earth in accommodating itself to a secular lengthening of the day, Macdonald obtains a kinematic viscosity $\nu \sim 10^{26}\,\text{cm}^2/\text{s}$, which he attributes to the whole mantle. We have already noted that the assumption of viscous homogeneity of the mantle leads to inconsistency between the viscosity inferred from the damping of the latitude variations and that needed to explain the permanence of surface features. It is not surprising that its use for this problem leads to a further inconsistency. Such a relaxation of the Earth's figure would be controlled by the most viscous parts of the earth and since the thin surface shell is not rigid enough to explain the discrepancy, it is much more attractive to assign a kinematic viscosity as large as $10^{26}\,\text{cm}^2/\text{s}$ to the lower mantle. We have seen that theory suggests a greater viscosity for this region than in the upper mantle, and it would not conflict with the other viscosity data. Such a large viscosity would quite effectively suppress convection in the lower mantle.

The absence of earthquakes in the lower mantle is either due to the lack of a stress accumulating mechanism or a more efficient process of stress relaxation than exists in the upper mantle. The second of these alternatives has been favoured up to the present (Gutenberg 1959) but with the data reviewed above it is more satisfactory to attribute the absence of earthquakes to the inhibition of convection, which is a plausible mechanism of stress accumulation in the upper mantle. It may also be added that seismic waves appear to be much less attenuated in the lower mantle than near the surface. This can be interpreted as the effect of an increase of viscosity with depth.

Finally in support of convection only in the upper mantle, one may cite the geochemical evidence on the distribution of nickel in crustal rocks. Assuming a chondritic composition for the earth as a whole, one may demonstrate that the nickel concentration in the crust is not compatible with complete mixing of mantle material.

Comparing terms in equation (6) we have:

$$\frac{2\boldsymbol{\omega}\wedge\mathbf{v}}{\nu\nabla^2\mathbf{v}} \sim \frac{2\omega L^2}{\nu} \quad (\text{Taylor number})^{\frac{1}{2}} \sim 10^{-8}, \tag{9}$$

$$\frac{(\mathbf{v}\,.\,\nabla)\,\mathbf{v}}{\nu\nabla^2\mathbf{v}} \sim \frac{vL}{\nu} = \text{Re} \quad (\text{Reynolds number}) \sim 6\times 10^{-20}. \tag{10}$$

From equation (9) we conclude that the Earth's rotation has no important dynamical effect on the convection, and there should be no observable tendency to axial symmetry.† The extremely small Reynolds number is a clear indication that the convection is not unstable because of a shear flow instability, but this alone is not sufficient to guarantee that the flow is steady if the boundary conditions are fixed. Thermal convection can become

† Possibly a more important coupling to the rotation is through the latitudinal variation of surface temperature, but this can also be classed as a small effect.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

unsteady due to the effects of either mechanical or thermal inertia, and it is important to examine both sources of instability in any convection problem. It seems wise to reserve the term 'turbulence' for those unstable flows caused by mechanical inertia (Reynolds number criterion), since there is the possibility that in some circumstances the unsteady flows that result from thermal inertia may have a more regular character that does not require the stochastic description given to flows at high Reynolds number †

More difficult to evaluate is the importance in determining the flow of heating by viscous dissipation. We have from equation (7):

$$\frac{\nu}{2C}\left[\frac{\partial v_i}{\partial x_k}+\frac{\partial v_k}{\partial x_i}\right]^2\frac{C\rho}{H}\sim\frac{\nu\rho v^2}{HL^2}=3\times 10^{-2}. \tag{11}$$

This ratio is not sufficiently small that we can confidently neglect the viscous heating on the strength of equation (11) since the values chosen for the physical parameters used in the calculation of this ratio are rather inaccurate. We shall, however, neglect viscous heating, providing a better reason for doing so later.

In accordance with our desire to investigate, for the moment, only statistical aspects of the flow, we now examine assumption (*d*), which implies that at least in a statistical sense, the upper mantle convection may be regarded as quasi-steady. This demands at least that the various boundary conditions do not change appreciably in certain characteristic times of the convecting region. Two such characteristic times that may be conveniently termed the motional and thermal time constants appear to be of considerable importance. For a convecting layer of thickness L these are respectively $\tau_M = L/v$ and $\tau_T = 0{\cdot}1L^2/\kappa$; physically they represent the time for a fluid mass to cross the system, and the time it takes a thermal disturbance to be conducted across the system. Taking the parameter values given above, together with $\kappa = 2\times 10^{-2}\,\text{cm}^2/\text{s}$ we obtain $\tau_M = 2\times 10^7$ y and $\tau_T = 6\times 10^8$ y.

As far as one can judge, the most important change in boundary conditions is the variation of H with time owing to radioactive decay and the fractionation of radioactive minerals to the crust. Using the chondrite radioactivity data given by MacDonald (1959), one can estimate that the rate of radioactive heating decays by 25 % in the longer characteristic time τ_T, with most of the change being due to ^{40}K. MacDonald (1963), after a study by Tilton & Reed (1963), has suggested that the ratio of potassium to uranium is less in the earth than in the chondrites by an amount which reduces the change in heating during τ_T to less than 10 %. The rate of loss of radioactivity from the mantle by fractionation is also small. Something of the order of a third to a half of the radioactive minerals have moved to the crust over several billion years, which represents a small fractional loss in the time τ_T. This also shows that the assumption of uniformity for the real heat source distribution is not disturbed by rapid fractionation.

From these arguments it appears reasonable to regard the convection as at least statistically stationary, so that if we form the time averages (over some suitably long period of

† Knopoff (1964), using a Bénard cell model of mantle convection, concludes that mantle convection is turbulent because the Rayleigh number used in his theory is so great. This is probably correct if the term turbulence is applied to any unsteady flow, but if used in the restricted sense given above is incorrect. The onset of high Reynolds number instability is determined by both the Rayleigh number and the Prandtl number (Pr $= \nu/\kappa$).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

time) of the various terms remaining in equations (6), (7) and (8), we have (bars denoting time averaging):

$$\overline{\frac{1}{\rho}\nabla p} = \overline{\nu\nabla^2\mathbf{v}} - \overline{\beta T\mathbf{g}}, \tag{12}$$

$$\overline{\mathbf{v}.\nabla T} = \overline{\kappa\nabla^2\mathbf{v}} + \overline{H/C\rho}, \tag{13}$$

$$\overline{\nabla.\mathbf{v}} = 0. \tag{14}$$

Dimensional analysis of these equations shows that the solutions are of the form

$$\overline{\mathbf{v}} = \frac{\kappa}{L}\overline{\mathbf{v}}(z/L, \mathrm{Rh}), \tag{15}$$

$$\mathrm{Rh} = \frac{g\beta L^5 H}{\kappa^2 \nu C\rho}. \tag{16}$$

The dimensional group Rh has the significance for fluid systems with internal heat generation that the Rayleigh number Ra has for systems with a characteristic temperature gradient impressed on them. One expects the onset of convection and changes in the modes of convection to be characterized by various critical values of Rh, providing inertia forces remain unimportant.

It has been shown from general arguments (Landau 1944) that if the onset of an instability is characterized by some number Q_c say, then the non-dimensional velocity of convection increases as $(Q-Q_c)^{\frac{1}{2}}$. Assuming this result also applies to the statistically averaged velocity if the flow is unsteady,† we have:

$$|\overline{\mathbf{v}}| = A(\kappa/L)\,(\mathrm{Rh}-\mathrm{Rh_c})^{\frac{1}{2}}\,|\mathbf{f}(z/L, \mathrm{Rh})|, \quad (|\mathbf{f}| \sim 1), \tag{17}$$

where A is a numerical constant of the order unity. Using the values of convection parameters given above, we obtain from equation (16) $\mathrm{Rh} \sim 10^6$. It is known (Chandrasekhar 1961) that $\mathrm{Rh_c} \sim 10^3$ for such a layer, so that we may put $(\mathrm{Rh}-\mathrm{Rh_c}) \sim \mathrm{Rh}$. Assuming that $A \sim 0{\cdot}3$ (see footnote) we find from equation (17) $|\overline{\mathbf{v}}| \sim 10^{-7}$ cm/s, which is in excellent agreement with the more direct estimates of the mean velocity given above.

If we now reconsider the importance of viscous heating relative to radioactive heating, we have from equations (11) and (17):

$$\frac{\overline{\nu\rho v^2}}{HL^2} \sim \frac{A^2 g\beta L}{C}. \tag{18}$$

From this result we see that the relatively large uncertainties involved in estimating ν and v are self-cancelling in their effects on this ratio and that the principal factor governing the importance of viscous heating is the size of the system. By substitution, we find

$$A^2 g\beta L/C = 6\times 10^{-3},$$

† Malkus (1954) has derived a formula of this form (replacing Q by Ra) for the r.m.s. velocity in the fully 'turbulent' state of Bénard cell convection—$\overline{v} = \frac{1}{3}(\kappa/L)\mathrm{Ra}^{\frac{1}{2}}$ ($\mathrm{Ra} \gg \mathrm{Ra_c}$). Chandrasekhar (1961) has given expressions for the velocity of steady Bénard cell convection in the epimarginal range of Rayleigh numbers which is also of the same form; he calculates A (equation (17) with Rh replaced by Ra) to be in the range 0·22–0·3 depending on boundary conditions. The dependence of **f** on Rh is not expected to be strong, and will not appreciably affect the result $|\mathbf{f}| \sim 1$ for most of the convecting region, and for all Rh. Dependence of **f** on variables in the horizontal direction disappears on averaging.

from which it may be computed, by using the value of the geothermal flux ($2{\cdot}4\times10^{13}$ J/s), that the average rate of viscous dissipation is $1{\cdot}4\times10^{11}$ J/s. Since the lateral extent of the convection zone around the earth is so large compared with its thickness (*ca.* $70L$), it seems permissible to consider quantities space averaged through the convection region to be equal to their time average. This implies that the average rate of energy release by seismic waves is about 5% of the viscous dissipation,† and indicates that non-Newtonian flow may be quite important in mantle convection.

It is interesting to consider briefly what might occur in a system for which $A^2g\beta L/C > 1$. General thermodynamic considerations forbid that the average viscous heating is greater than the radiogenic heating and we would conclude that for such large systems, the velocity increases as $(\mathrm{Rh}-\mathrm{Rh_c})^{\frac{1}{2}}$ for a limited range of Rh, and then more slowly. The convection cells may also be less than the vertical thickness of the region in such a way as to keep $A^2g\beta L/C < 1$ for any cell.

The identification of Rh as the parameter characterizing the statistical properties of the convection (or the flow itself if it is steady) provides a useful prescription for the correlation of experimental results, but for mantle convection the time required for averaging is so long that unless it can also be demonstrated that the convection is steady, the results are not of any great value in understanding the pattern of activity we see today. It has been demonstrated above that the equations relating the unaveraged convection variables in the mantle are to an adequate approximation as follows:

$$\frac{\partial \mathbf{v}}{\partial t} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{v} - \beta T\hat{\mathbf{g}}, \tag{19}$$

$$\frac{\partial T}{\partial t} + \mathbf{v}.\nabla T = \kappa\nabla^2 T + H/C\rho, \tag{20}$$

$$\nabla.\mathbf{v} = 0. \tag{21}$$

It is reasonable to suppose (see below) that beyond the marginally stable state ($\mathrm{Rh} > \mathrm{Rh_c}$) there are solutions of this set of equations corresponding to régimes of both steady and unsteady finite amplitude convection. It follows that any transition from steady to unsteady convection would be characterized by Rh only, since we know that all steady solutions of these equations are specified by the values of this number.‡ If we now rearrange equation (17), we see that such an instability would occur at a certain value of vL/κ. This dimensionless group is known as the Péclet number Pe and it plays a role in the thermal part of the convection problem analogous to that played by the Reynolds number in the mechanical part. It is a measure of the extent to which the effects of thermal inertia are dissipated by thermal conduction. Alternatively, one may see that the Péclet number is a measure of the importance of the non-linear term $\mathbf{v}.\nabla T$ compared with $\kappa\nabla^2 T$ in equation (20). This is brought out clearly by the following relations:

$$\mathrm{Pe} = \frac{vL}{\kappa} = \frac{L^2}{\kappa}\frac{v}{L} = 10\frac{\tau_T}{\tau_M}, \tag{22}$$

$$\mathrm{Pe} = \mathrm{Re}\,\mathrm{Pr}. \tag{23}$$

† Seismic energy released in convecting region $\sim 6\times10^9$ J/s.

‡ Compare this situation with the onset of turbulence (high Reynolds number). Because of the importance of the term $(\mathbf{v}.\nabla)\mathbf{v}$ in the steady régime near the instability, the transition is determined by both Rh and Pr.

Using the interpretations of τ_T and τ_M given above, we see that such instability, which will be henceforth referred to as Péclet instability, will arise if the time of convective overturn τ_M becomes too short in comparison with the time it takes heat to be conducted across the layer. This seems very reasonable because it is just the condition for a fluid layer in an unstable state to be inverted to a stable configuration. From an examination of equation (23), we see that the condition for the Péclet instability to precede the shear flow instability and therefore not to be masked by it, is that the Prandtl number must be greater than some value $\mathrm{Pr}' = \mathrm{Pe_c}/\mathrm{Re_c}$. If one remembers that $\mathrm{Re_c}$ is probably greater than 10^2, it seems likely that Pr′ is likely to be of the order or less than unity.

The Prandtl number of the fluids used in laboratory experiments on thermal convection are in the range 1 to 5000 and by the above argument we may look for evidence of Péclet instability in these experiments. The work of Silveston (1958) on Bénard cell convection shows quite clearly that this is the cause of the transition to unsteady flow. If we recast the above theory to apply to Silveston's experiments, we expect the transition to unsteady flow to be determined by a critical Rayleigh number Ra′ if it is a Péclet instability, or a critical value of $(\mathrm{Ra}/\mathrm{Pr}^2)$ if it is a shear flow instability. Silveston's experiments were performed with several liquids having Prandtl numbers that differed by as much as a factor 10^3, and he was able to show that the onset of unsteady convection was characterized by the relation:

$$\mathrm{Ra}' = 18\,000\,\mathrm{Pr}^{0\cdot 2}. \qquad (24)$$

The weak dependence on Prandtl number probably arises because, for the liquids with the lowest Prandtl numbers, the Péclet and shear flow instabilities are not well separated $(\mathrm{Pr} \sim \mathrm{Pr}')$ and they will interact. That such interaction is taking place is revealed by a transitional régime of flow in a range of Rayleigh numbers just less than those in the unsteady régime, that decreases as the Prandtl number increases. One expects that at high Prandtl numbers Ra′ has an asymptotic value of the order 4×10^4. From the expression (Chandrasekhar 1961) for the velocity of Bénard cell convection with the boundary conditions used in Silveston's experiments, we have a critical Péclet number given by†:

$$\mathrm{Pe_c} = 0\cdot 22(\mathrm{Ra}' - \mathrm{Ra_c})^{\frac{1}{2}}. \qquad (25)$$

Substitution gives $\mathrm{Pe_c} \sim 40$ and from equation (22) we see that unsteady flow sets in when the time of convective overturn τ_M becomes less than about $\frac{1}{4}\tau_T$.

Returning to the mantle convection problem, we may definitely expect a Péclet instability to occur first, since the Prandtl number of the mantle is *ca.* 10^{22}. We expect the flows to become unsteady at about the same or a lesser (see below) value of the Péclet number than that deduced above for Bénard convection, which would mean a critical Rh′ of *ca.* 10^4. This strongly suggests that mantle convection is unsteady since we know that $\mathrm{Rh} \sim 10^6$. We have seen that even if we calculate τ_T/τ_M directly from geophysical observations that give only the average convection velocity, we find this ratio is ~ 30 and it will, of course, be higher if unsteadiness is allowed for. We shall see that this predicted instability of the flow is able to relate many diverse geophysical observations.

There is reason to believe that convection resulting from internal heat generation may be more prone to Péclet instability than Bénard cell convection, where the fluid is heated

† $\mathrm{Ra_c} = 1700$, which was also confirmed by Silveston's experiments.

at its boundaries. In the steady régime of Bénard convection, one observes that most of the heat transport is accomplished by the part of the fluid that comes into closest proximity to the boundaries and which lies on the outside of convection cells. The rest of the fluid is moved more by viscosity than buoyancy forces. As a result, the scale factor to be substituted in the definition of τ_T is now the thickness of the thermal boundary layer. We therefore conclude that the critical Péclet number is likely to be reached quicker in a system with internal heat generation than in Bénard cell convection, i.e. $\mathrm{Rh}' < \mathrm{Ra}'$. On examination, it will be realized that the concept of convection 'cells' automatically implies *both* inhomogeneous heating and transport of heat by the fluid so that it is more difficult to reconcile any finite convection in closed streamlines with a uniform heat generation in the fluid. Chandrasekhar (1961) has, however, demonstrated that the principle of the exchange of stabilities is valid at marginal stability for some cases of internal heat generation, presumably because the fluid is still asymmetrically cooled.

It is interesting to inquire as to the character of the unsteady convection at super critical Péclet numbers. It can be shown that the unsteady solutions of equations (19), (20) and (21) are, strictly speaking, functions of two parameters (Rh, Pr) or (Ra, Pr) depending on heating conditions. However, it may also be shown that the dependence on Prandtl number is very weak until Rh or Ra becomes large enough to cause an additional shear flow instability. The dependence on two parameters is confirmed by measurements of the heat transport in Silveston's experiments. The nature of the unsteady flow in systems with internal heat generation has not yet been investigated experimentally, so that for the moment one assumes that they will be similar to the flows observed by Silveston. If we are to use these flow patterns as a guide to mantle convection ($\mathrm{Rh} \sim 10^6$, $\mathrm{Pr} \sim 10^{22}$), we must, on account of the much smaller Prandtl number of the laboratory fluids, consider their behaviour only in the immediate vicinity of the onset of unsteady flows. Here one observes a shifting pattern of the elongated cells or 'rolls' that are well developed in the steady régime (photographs in Chandrasekhar 1961). In the case of the mantle, it seems very likely that the unsteadiness of such a linear pattern would be modified by non-Newtonian effects. The effective viscosity falls rapidly with increasing rates of strain if the stress is in the 'high stress' range defined above, and this would encourage the development of a relaxation oscillation.† An overturn of the convecting layer would occur in a time of the order τ_M or less followed by a quiescent period of the order τ_T.

Another feature to be expected for mantle convection is a large difference in the lateral extent of the ascending and descending streams. This arises partly because of the rapid variation of viscosity with temperature. The requirement that the flow in a system with steady boundary conditions adjusts itself so that the rate of entropy production in the fluid is a minimum‡ will mean that the least viscous parts of the system have the maximum shear rates. Such asymmetries are observed in Bénard cell experiments (Silveston 1958) with normal fluids, and would be particularly marked in the mantle. It may be calculated that

† The type of behaviour observed when two bodies are rubbed together with the dynamic friction less than the static friction.

‡ Glansdorff & Prigogine (1964) have obtained this result only for a steady flow—it is not clear from their communication how the principle is applicable to unsteady flows, but if the result may also be applied to the time averaged entropy production, the above result conjectured for the instantaneous flow seems physically plausible.

the viscosity in the rising stream is perhaps more than ten times less than in the descending stream, which would result in the ascending stream being much the narrower. The effect is likely to be further accentuated by non-Newtonian properties. The effective viscosity decreases as the shear stress increases, due to the occurrence of additional creep processes and it may be seen that the result of this is to concentrate convective shearing into narrow zones. Fracture may be the final result of such concentration.

Before concluding, we discuss briefly the factors that determine the temperature distribution in the mantle if convection is allowed for. This may be understood qualitatively by considering the sequence of quasi steady states that would occur in the earth if H were increased slowly from zero. At first, the temperature gradients are everywhere subadiabatic and the conduction equation solution is appropriate. When H becomes great enough for any region to be super adiabatic, it is necessary to examine the value of Rh. Providing this is less than Rh_c, the conduction equation solution is still correct. In the mantle the viscosity is so large when super adiabatic conditions are first reached that considerable super adiabatic gradients would be stabilized. Remembering the enormous variation of viscosity with temperature, it may be demonstrated that it is chiefly the viscosity decrease that determines the rise of Rh as the heating rate is further increased. Eventually, the viscosity at some intermediate range of depths becomes small enough for $\mathrm{Rh} > \mathrm{Rh}_c$ and convection commences. There is relatively little change in temperature and viscosity with further increases of H. For a body the size of the mantle, the viscosity would never be much less than $10^{20}\,\mathrm{cm^2/s}$ with any rate of radiogenic heating conceivable in the past.

Conclusions

The main points that have been suggested by this study of thermal convection theory and its application to the mantle are:

(i) The properties of mantle minerals are such that convection can be driven at an average velocity of about $10^{-7}\,\mathrm{cm/s}$ ($3\,\mathrm{cm/y}$) by the amount of radiogenic heating that is likely to be available.

(ii) Convection is confined to the outer few hundred kilometres of the mantle.

(iii) The convection is unsteady, probably with an active phase of movement separated by periods of quiescence several times longer.

(iv) There is evidence that convective overturn will take place in a very elongated convection pattern, i.e. ‘rolls’.

(v) The ascending streams are much narrower than the descending ones.

One of the restrictions on convection theories pointed out by MacDonald (1963) is the evidence from seismology that the oceanic and continental mantle is significantly different down to a depth of a few hundred kilometres. The equality of oceanic and continental heat flow, combined with the evidence of the distribution of radioactive elements in the crust, also supports this evidence of mantle inhomogeneity. The restriction of convection to the upper mantle would give poor mixing of regions thousands of kilometres apart laterally so that this type of convection is not incompatible with these observations. Its compatability with the non-hydrostatic shape of the Earth has already been noted.

The prediction from convection theory that mantle convection is unsteady with relatively long periods of quiescence, is a particularly satisfying feature. As indicated in the text, for this scale of convection, spatial and temporal averaging of the convection should give the same results, and as a consequence, the variations in time should be matched by variations in space. Hence, the large ratio of quiescent and active phases of convection explain also why the active belts, such as the present ocean ridge system or the orogenic belts of geology occupy such a small fraction of the Earth's surface. The predicted linear pattern of the convection is also in agreement with the pattern of the active belts. Joly (1925), Holmes (1944), and Umbgrove (1947) have all emphasized the episodic character of mountain building and their estimates of the duration of active and quiescent phases is in satisfactory agreement with the value of τ_M and τ_T. At the completion of overturn, the temperature at a depth of about 100 km will have increased by 100 to 200 °C, which would provide an explanation of the creation of vast quantities of magma in orogenic belts. Heat flow at the surface would eventually regionally increase to about 20% greater than the average. The more concentrated and hence more rapidly moving ascending convection streams may well be the explanation of why it is relatively so difficult to find any evidence at the surface of the descending streams.

It seems premature to attempt an explanation of continental drift with only a thermal convection theory. The very deep structure of continents (MacDonald 1963) and concentration of earthquake activity around the margin of the Pacific, suggests that convection driven by a lateral chemical and physical inhomogeneity may play an important role in this process. The average depth of earthquakes is different in the circum-Pacific belt from elsewhere so that it seems safest to apply the ideas developed above only to phenomena occurring in either strictly oceanic or continental regions. That this type of convection is probably also involved in any continental drift is indicated by the far from random distribution of ridges in ocean basins.

It is a pleasure to record the helpful discussion and criticism of Dr D. J. Tritton.

References (Tozer)

Bullard, E. C. & Gellman, H. 1954 *Phil. Trans.* A, **247**, 291–322.
Chandrasekhar, S. 1961 *Hydrodynamic and hydromagnetic stability.* Oxford University Press.
Glansdorff, P. & Prigogine, I. 1964 *Physica*, **30**, 351–375.
Gutenberg, B. 1959 *Physics of the Earth's interior.* New York: Academic Press.
Herring, C. 1950 *J. Appl. Phys.* **21**, 437–445.
Holmes, A. 1944 *Principles of physical geology.* London: Nelson Press.
Jeffreys, H. 1930 *Proc. Camb. Phil. Soc.* **26**, 170–172.
Joly, J. 1925 *The surface history of the Earth.* Oxford University Press.
Knopoff, L. 1964 *Rev. Geophys.* **2**, 89–122.
Landau, L. D. 1944 *C.R. Dokl. Acad. Sci. U.R.S.S.* **44**, 311–314.
MacDonald, G. J. F. 1959 *J. Geophys. Res.* **64**, 1967–2000.
MacDonald, G. J. F. 1963 *Rev. Geophys.* **1**, 587–665.
Malkus, W. V. R. 1954 *Proc. Roy. Soc.* A, **225**, 196–212.
Munk, W. H. & Davies, D. 1964 *Isotopic and cosmic chemistry*, pp. 341–346. Amsterdam: North-Holland Publishing Company.
Munk, W. H. & MacDonald, G. J. F. 1960 *The rotation of the Earth.* Cambridge University Press.
Nabarro, F. R. N. 1948 *Rep. Conf. on strength of solids, Bristol*, p. 75. London: The Physical Society.
Runcorn, S. K. 1962 *Continental drift*, pp. 1–39. New York: Academic Press.

Silveston, P. L. 1958 *Forsch. Ing. Wes.* **24**, 29–32, 59–69.
Tilton, G. R. & Reed, G. W. 1963 *Earth science and meteoritics*, pp. 31–44. Amsterdam: North-Holland Publishing Company.
Tozer, D. C. 1959 *Phys. and Chem. Earth*, **3**, 414–436.
Tozer, D. C. 1965 *Geophys. J.* **9**, nos. 2 and 3, pp. 95–112.
Umbgrove, J. H. F. 1947 *The pulse of the Earth.* Nijhoff.
Van Bueren, H. G. 1961 *Imperfections in crystals.* Amsterdam: North-Holland Publishing Company.
Verhoogen, J. 1961 *The Earth today*, pp. 276–281. Edinburgh: Oliver and Boyd.
Zharkov, V. N. 1963 *Problems of theoretical seismology, Physics of the Earth's interior*, pp. 29–49. Israel Programme for Scientific Translations.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES