

Mantle Convection and Viscosity.

W. R. PELTIER (*)

Department of Physics, University of Toronto - Toronto, Ont., Canada M5S 1A7

1. - Introduction.

The «renormalization» of the Earth Sciences, which began in earnest during the decade of the 60's, was based upon the increasingly widespread acceptance of the notion of continental drift—or at least the form of continental drift which came to be called sea floor spreading [1, 2]. This idea, and later refinements and elaborations of it, now serves as the guiding principle in terms of which much of geological and geophysical thinking is organized.

At the heart of this paradigm is the understanding that the Earth's mantle behaves as a viscous fluid when it is subject to a shear stress which is applied over a sufficiently long time scale. Given this characteristic behaviour, it follows that, if the radial temperature gradient in the mantle is sufficiently in excess of the adiabatic, then the mantle will be convectively unstable. The observed spreading of the sea floor away from the mid-ocean ridges is presumably a surface manifestation of deep-seated mantle convection. This and other aspects of the pattern of surface motions associated with convection has been described kinematically within the framework of the set of ideas which has come to be called «plate tectonics» [3, 4]. In this picture the surface of the Earth is divided into a relatively small number of rigid «plates» which interact with each other only at their boundaries. Continents are carried passively by the plates and drift with them over the Earth's surface. Although the practitioners of this view agree that plates are created by the rise of hot material at the ocean ridges and are destroyed at trenches where cold material sinks, they are not unanimous in their willingness to allow that there may be an almost 1:1 relation between surface plates and deep-mantle convection cells. This has led to a great deal of confusion concerning the connections between the ideas of plate tectonics and mantle convection. Not the least of these confusions concerns the question as to what makes the plates go [5-7].

That there may be a direct connection between mantle convection cells

(*) Alfred P. Sloan Foundation Fellow.

and surface plates does not seem at present to be widely accepted, although it was certainly an explicit part of the spreading hypothesis advanced by HESS [1]. It appears that one argument against acceptance of this simple model has been based upon the belief that the viscosity of the lower mantle below a depth of about 650 km was extremely high [8, 9]. This has led to the construction of models in which convection is confined to the upper mantle [10-12], but such models have several problems associated with them, not the least of which concerns the aspect ratio of the circulation which is required if each plate is to correspond roughly to the «roof» of a convection cell. Yet it is only with such approximate correspondence that the reason for plate motion ceases to be mysterious. Rather than abandoning the upper-mantle convection models on account of this failure, in favour of the whole-mantle view of Hess, Holmes [13] and others, their advocates have felt obliged to make them tortuously complex. The extent of this complexity is best illustrated by the «two scale» models introduced in [14, 15]. It is tempting to draw an analogy between such complex models of mantle dynamics and models of planetary orbits in an Earth-centered solar system with their profusion of Ptolemaic epicycles!

In the discussion which follows, I have tried to review the main questions which are at issue in attempting to connect the descriptive kinematic ideas of plate tectonics with models of mantle dynamics. Since there have been several recent and extensive surveys of the literature in this field [16-18], I have not tried here to provide a similarly complete coverage, but rather have focussed upon the questions which seem to me to be most important. In sect. 2 the observational data which have combined to make the convection hypothesis so compelling are briefly summarized. Section 3 is concerned with a discussion of models of postglacial rebound and of their application to the problem of inferring the profile of mantle viscosity. Sections 4 and 5 are respectively devoted to discussions of convection in the laboratory and convection in the Earth. Section 6 is concerned with a discussion of the importance of convection in understanding the Earth's thermal history and the variation of temperature with depth in the mantle.

2. - Evidence for a mantle general circulation.

Although even the earliest geographers were impressed with the similarities between the coastlines of South America and Africa, it remained for BULLARD *et al.* [19] to describe the goodness of fit statistically and this was done subsequent to the advance of the spreading hypothesis in [1]. A reconstruction of the original supercontinent Pangaea [20] is shown in fig. 1. That the present continents may indeed have been produced by the break-up of Pangaea was established using paleomagnetic methods [21, 22]. Reconstruction of the

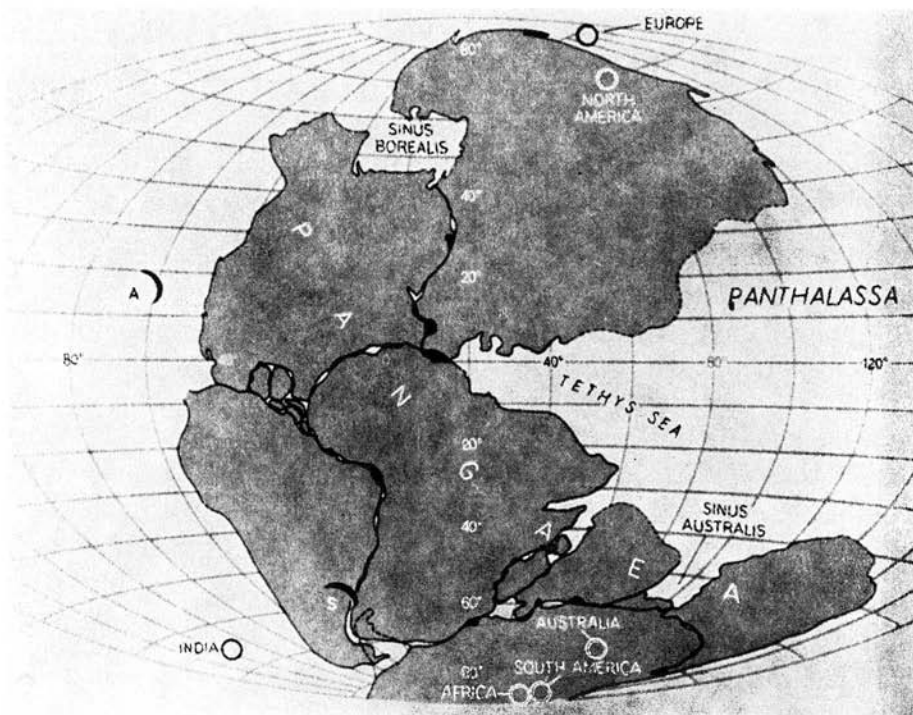


Fig. 1. — Reconstruction of the supercontinent Pangaea as it may have looked 200 million years ago. The relative positions of the continents are based upon computer best fits (from ref. [20]).

paleo pole paths from Europe and North America shows that they were similar until the Triassic, when they diverged, an event interpreted as coinciding with continental break-up.

2'1. *Paleomagnetic evidence.* — Paleomagnetic methods have in fact played a dominant role in establishing the validity of the spreading hypothesis. In 1963 MORELY [unpublished] and VINE and MATHEWS [23], in the course of studies of the pattern of magnetic anomalies over an ocean ridge, independently realized that the patterns observed could be understood in terms of Hess' theory of sea floor spreading. An example of the characteristic magnetic-anomaly pattern which obtains over a ridge is shown in fig. 2 which is from [24]. MORELY and VINE and MATHEWS proposed that such characteristic patterns were produced when hot material rising from the mantle beneath the ridge crest was cooled below the Curie point, acquiring in the process an induced magnetization in the local Earth's magnetic field. The material reaching the surface is normally or reversely magnetized depending upon the polarity of the field at the time. The analogy has often been drawn between the sea floor,

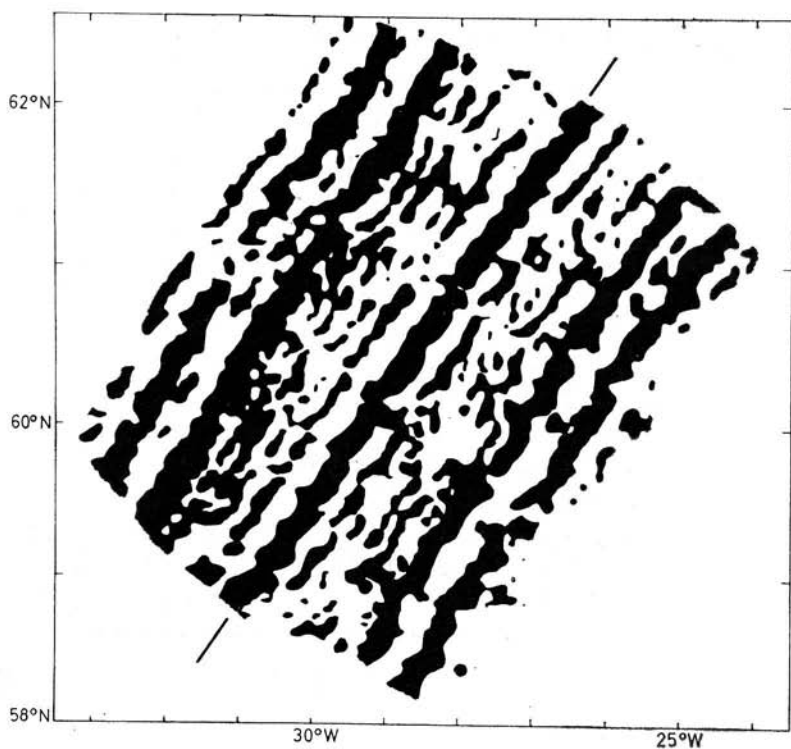


Fig. 2. - Pattern of magnetic anomalies over the Reykjanes Ridge southwest of Iceland. The straight lines indicate the ridge axis and the central positive anomaly (from ref. [24]).

spreading away from a mid-ocean ridge, and a strip of magnetic tape on which the record of spreading has conveniently been written.

When the characteristic pattern of magnetic anomalies in the vicinity of a ridge crest is combined with the time scale of geomagnetic reversals established on land [25, 26], it is then possible to deduce the rate of spreading as a function of location. Examples of such data [27] are shown in fig. 3 and illustrate spreading rates of the sea floor which vary of (1÷5) cm/y.

2.2. Seismic evidence and the ideas of plate tectonics. - If the methods of paleomagnetism were responsible for establishing that hot mantle-derived material is continuously rising under and spreading horizontally from the mid-ocean ridges, it remained for seismology to demonstrate in an equally convincing fashion that cold and relatively old material was elsewhere sinking into the mantle. Prior to the advance of the sea floor spreading hypothesis it had been suggested by BENIOFF [28] that the deep ocean trenches were regions where ocean floor was downthrust into the mantle along what are now

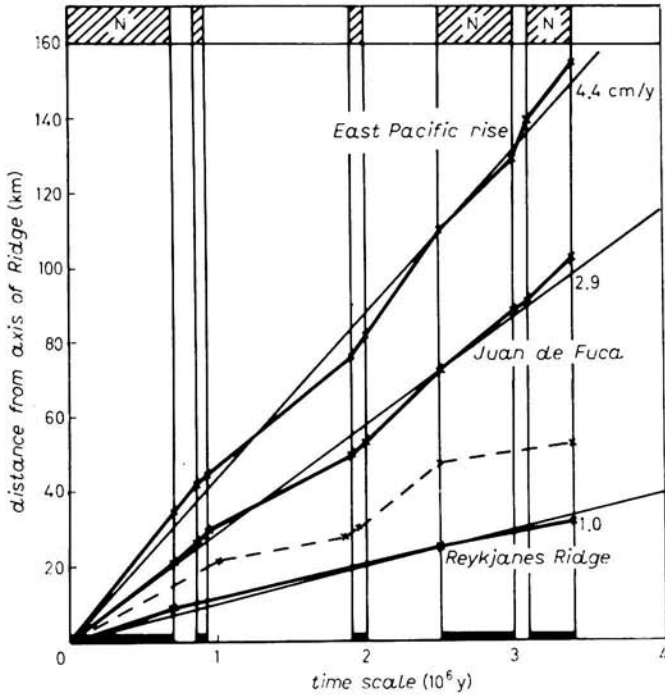


Fig. 3. — Inferred normal and reversed boundaries in the crust plotted against the time scale of geomagnetic reversals for various spreading ridges. The slopes are the local spreading rates (from ref. [27]).

called Benioff zones which are defined by the locus of deep seismicity. Figure 4 shows Benioff's data for the Kurile-Kamchatka region illustrating the geometry of the process which is now referred to as the «subduction» of oceanic lithosphere. In fact, it remained until the occurrence of the Alaska earthquake in 1969 before Benioff's suggestion was established as fact [29, 30]. The analysis of the earthquake source mechanism as a function of depth along the Benioff zone using first motion data [31] has played an important role in constraining mantle convection models. The cessation of deep seismicity beyond a depth of (650 ÷ 700) km, where a discontinuity of the seismic parameters occurs, and the compressive nature of the deep focal mechanisms have been interpreted as implying that the cold oceanic slab does not penetrate below this depth. This unsubstantiated interpretation has been the main observational evidence invoked to support the notion that convection is confined to the upper mantle.

On the basis of global distributions of seismicity such as discussed by ISACKS *et al.* [32], it was recognized that the system of ridges and trenches, and their important interconnections via transform faults [33], was continuous over the surface of the Earth and divided this surface into a relatively small number of «plates», the interiors of which were relatively devoid of seismicity.

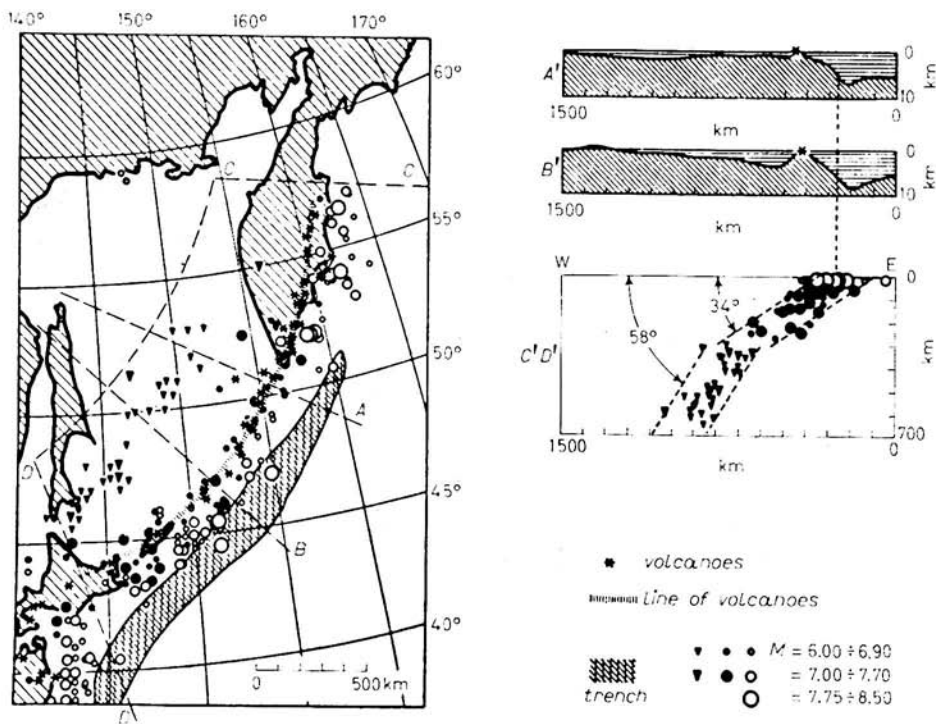


Fig. 4. - Location map and composite profile of Kurile-Kamchatka earthquake sequences (from ref. [28]).

An example of a recently compiled global seismicity map is shown in fig. 5 and a map of plate boundaries is illustrated in fig. 6. The idea of plate tectonics is basically to describe geological processes such as mountain building and volcanism as consequences of the interactions among this relatively small number (12) of plates. It is supposed that all interactions take place at plate boundaries which consist of segments of the three characteristic forms: spreading ridges, deep ocean trenches (subduction zones) and transforms. Plate tectonics, as developed in [3, 4, 34, 35], is a purely kinematic though nevertheless useful description of the present pattern of horizontal motion of material on the Earth's surface. Within the context of this «theory» neither the existence of the plates nor the reason for their relative motion are explained. Although the approximately 1:1 relation between convection cells and plates envisioned by HESS would appear to provide a natural explanation for the motion of the plates, and their existence can be readily understood in terms of convection in a fluid whose viscosity is a strong function of temperature, this is not a view which is accepted by some of those who espouse both convection and plate tectonics [14, 15].

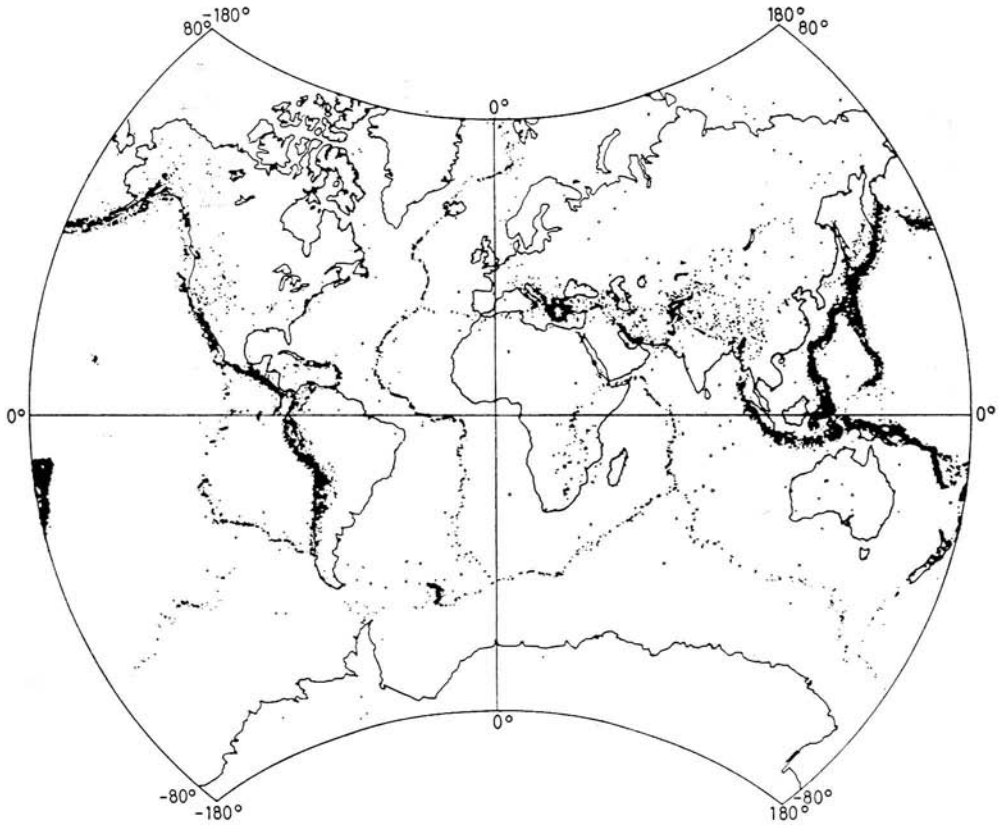


Fig. 5. — Global map of seismic epicentres for all earthquakes in the magnitude range $4.5 \div 5.5$ with depths $(0 \div 700)$ km. Note the correlation between seismicity and plate boundaries. (Based on NOAA data tape, Van der Grinten projection, Goddard Space Flight Center, 1978.)

2'3. *Other signatures of the circulation.* — Besides the information from paleomagnetism and seismology there are other geophysical data which are able to contribute usefully in providing constraints upon the nature of the circulation in the mantle. One of the most interesting of these sets of data consists of the observed variations of heat flow and ocean bathymetry as a function of ocean floor age away from the ridge crest. Examples of such data taken from [36] are shown in fig. 7 and 8 for bathymetry and heat flow, respectively. When plotted as a function of the square root of the age of the lithosphere, both of these sets of data fall essentially on straight lines, although there is a systematic departure for ages $> 10^8$ years. MCKENZIE and SCLATER [37] have shown that this characteristic variation is expected on the basis of a simple model consisting of a plate (the lithosphere) of constant thickness spreading at a uniform rate and subject to a constant temperature at its end (the ridge crest). In this model the topography is assumed to be isostatically compensated and this

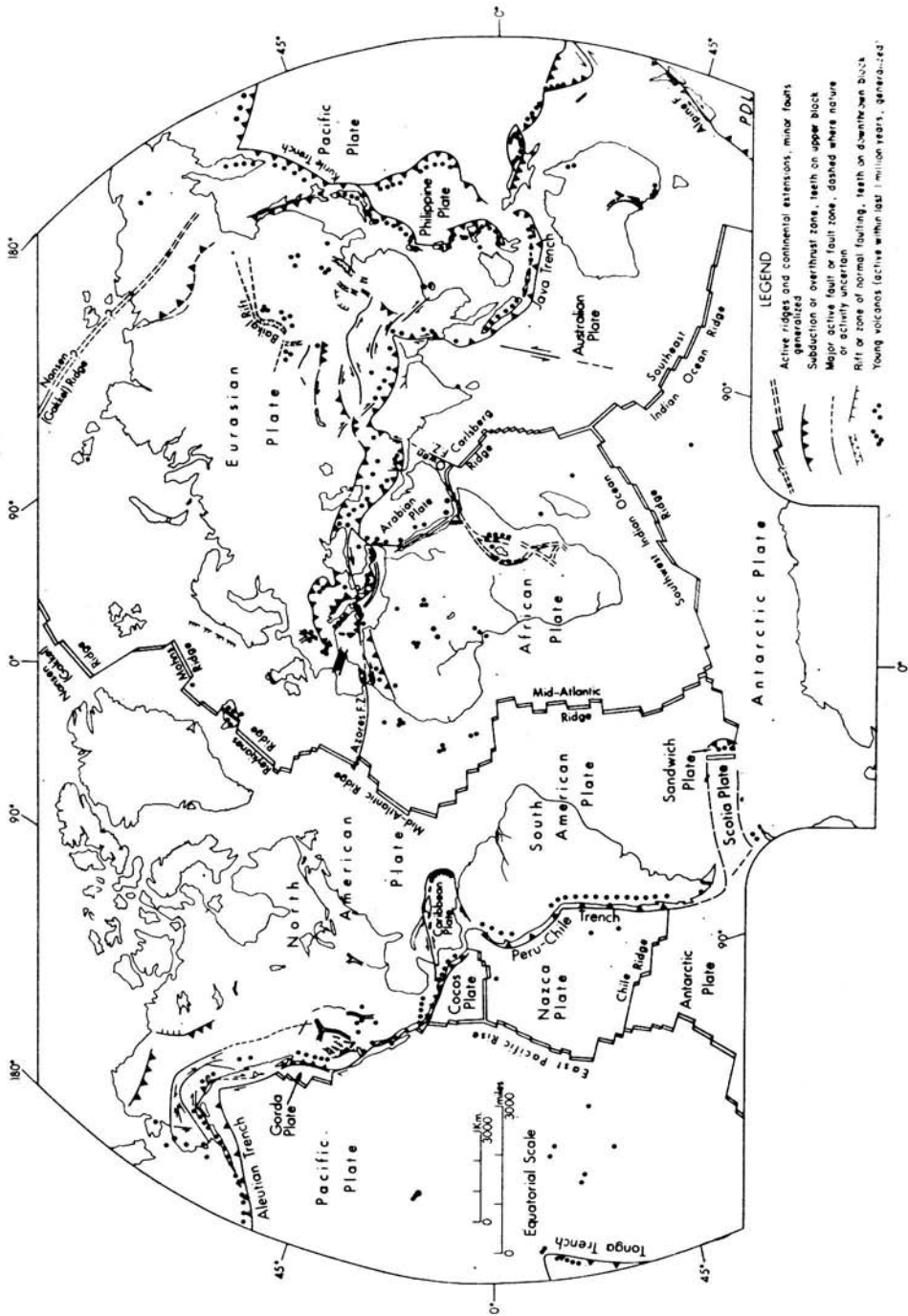


Fig. 6. - Global map of plate boundaries and of tectonic and volcanic activity. (Goddard Space Flight Center, 1978, based on *The Physical World*, © 1975 by the National Geographic Society, Van der Grinten projection, P. D. LOWMAN jr.).

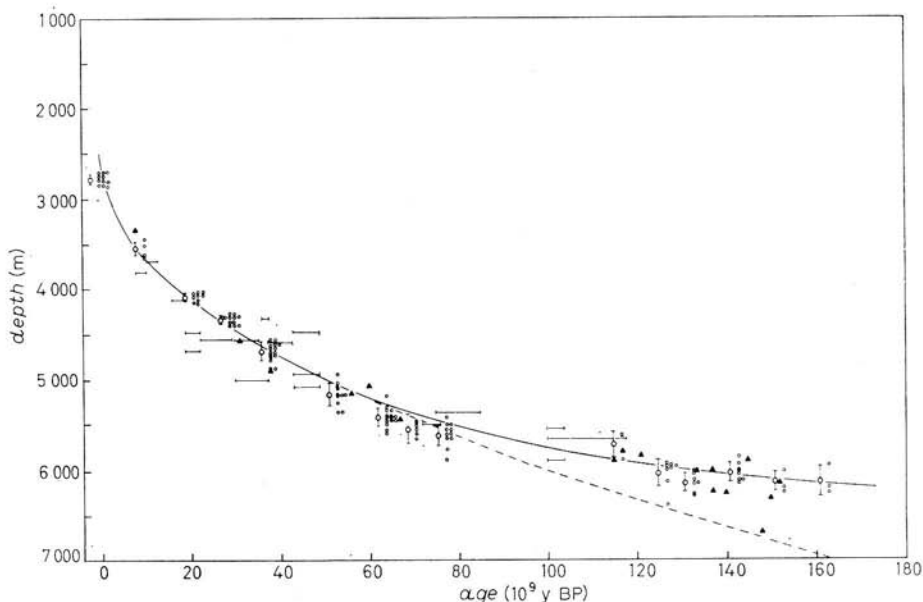


Fig. 7. — Plot of individual depth measurements used in calculating mean depths and depths at DSDP sites for the North Pacific (from ref. [36]): ϕ mean and standard deviation; \circ individual depth values from profiles; \blacktriangle DSDP sites with ages from magnetic anomalies; \blacksquare DSDP sites with biostratigraphic ages; — theoretical depth, plate model; - - - linear $t^{1/2}$ relation.

assumption is born out by the observed gravity field. Why the lithosphere is taken to be of constant thickness is not explained and there is no explicit hypothesis made regarding the reason for the spreading. It is not recognized explicitly in either [36] or [37] that the \sqrt{t} dependence of both heat flow and bathymetry also follows from the assumption that the ocean floor constitutes the cold thermal boundary layer of large-scale mantle convection cells. This follows from the boundary layer theory for convection at high Rayleigh and Prandtl numbers which will be discussed further in subsect. 4.4. Regarding the heat flow data shown in fig. 8, the reason for the large scatter of these data in the vicinity of the ridge crest is now understood to be due to the fact that a considerable fraction of the vertical heat transport may be effected by thermal convection of sea water though its near-surface host rock.

In the vicinity of a trench, where oceanic lithosphere is descending into the mantle, perhaps the most striking feature is the gravity anomaly on a profile perpendicular to the trench axis. In fig. 9 we show data from TALWANI *et al.* [38] for Tonga with superimposed model calculations due to GRIGGS [39]. The modelling shows that the characteristic form of the observed anomaly cannot be fitted by the topography alone, but is easily explained by the existence beneath the trench of a slab of cold high-density material dipping at the angle

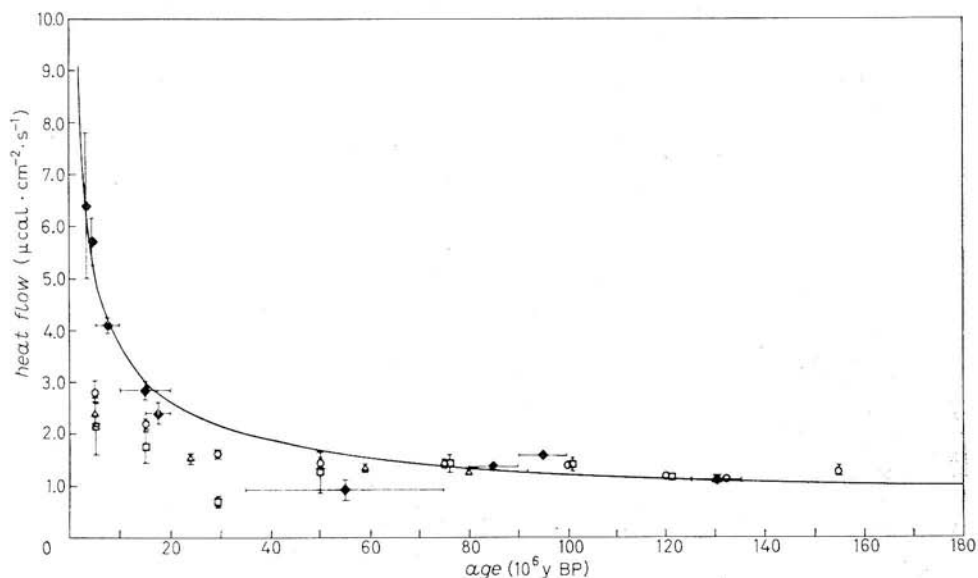


Fig. 8. — Heat flow as a function of age of the ocean floor indicating initial decay as $t^{-1/2}$ (from ref. [36]): \circ Pacific, \square South Atlantic, \triangle Indian, \blacklozenge « reliable » means from Pacific, — theoretical heat flow, plate model.

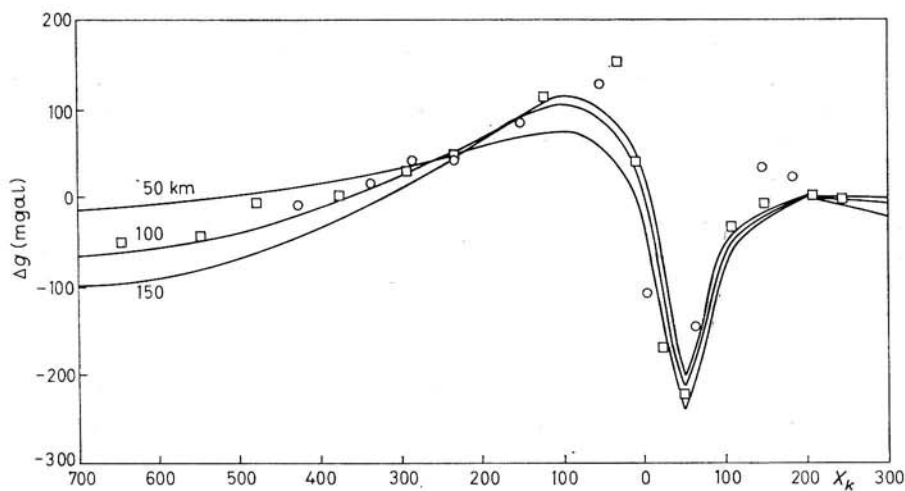


Fig. 9. — Gravity anomaly in milligal as a function of distance from the trench axis. A model with a dipping cold slab under the trench fits the data (from ref. [39]); no transition, $\psi = 45^\circ$, $v = 8$ cm/y, $\lambda = 2000$ km, \square traverse $\sim 20^\circ$ S, \circ traverse $\sim 25^\circ$ S.

of the observed seismic Benioff zone. Models such as these were introduced originally by MCKENZIE [40] and MINEAR and TOKSÖZ [41]. The form of the gravity anomaly over the Tonga trench is similar to those found in other such locations.

On a more global scale the free-air anomaly data have a far less diagnostic character. Figure 10 is a global free-air anomaly map taken from KAULA [42]

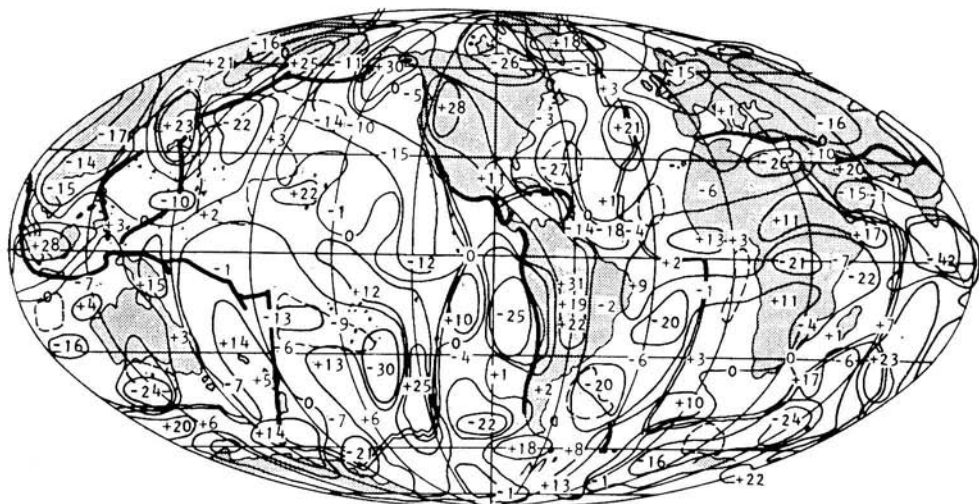


Fig. 10. — Global free-air anomalies in milligal referred to a fifth-degree figure (from ref. [42]): ——— compression, = = = = tension, - - - - approximately 3000 fathom.

referred to a fifth-degree figure and calculated from the spherical harmonic coefficients of the gravitational field using degrees 6-16 from GAPOSCHKIN and LAMBECK [43]. There is no clear correlation between the pattern of anomalies for such large wavelengths and the global system of ridges, trenches and transforms. One feature which is apparent on this scale, however, is the sequence of negative anomalies over the northern-hemisphere land masses near the pole. There is an anomaly of -26 mGal over the Hudson Bay and one of -15 mGal over Fennoscandia. These negative anomalies are presumably due to present-day isostatic disequilibrium produced by the disintegration of the large-scale Pleistocene ice sheets which existed in these locations, a process which began ca. 20 000 years BP (before present). We will see in the next section that the interpretation of these data, and of other information related to them, provides fundamental information which is required in the construction of mantle convection models.

One last piece of observational evidence which we should call attention to is clear by inspection of fig. 6. Although most of the global volcanic activity is closely associated with plate boundaries, there is nevertheless considerable activity in plate interiors and the explanation of this activity has remained

enigmatic. The Hawaiian chain of volcanic islands is the standard example, but considerable activity is also found in the centre of the African plate. The appearance of the Hawaiian and Emperor chains is such that they seem to have been formed by the motion of a rigid plate over a «hot spot» fixed in the mantle. The spreading rate which one infers from the relative ages of the islands does in fact agree with that deduced by independent means. MORGAN [44-46] has attached special significance to the global distribution of plume traces.

2.4. *The convection hypothesis.* — The convection hypothesis of drift and spreading is clearly able to explain *a priori* the most important of the observations discussed above, which is that hot material rises in the mantle and cold material sinks! That it is some form of convective circulation that maintains the plates in their relative motion, I would suggest, is indisputable. The questions which remain are more subtle ones, such as the extent to which the mantle participates in the circulation, and whether the circulation is maintained by the decay of radioactive heat sources, as is most often assumed [16, 17], or whether it is driven substantially by the primordial heat content of the planet following a hot formation, as has recently been suggested in [47]. Another issue which has never been resolved concerns the degree of coupling which exists between the plates and the underlying mantle. One idea due to EL-SASSER [48], which has persisted in the literature, is that the plates are essentially decoupled from the mantle beneath them. Once this idea is accepted, the question of what drives the plates can become perplexing indeed! This is clearly not the situation envisioned by HESS [1], who initially at least seemed to advocate the idea of whole-mantle convection with a roughly 1:1 relation between plates and convection cells. Figure 11, taken from Hess' article, illustrates this view. In the sections which follow this issue will provide one of the strands of continuity.

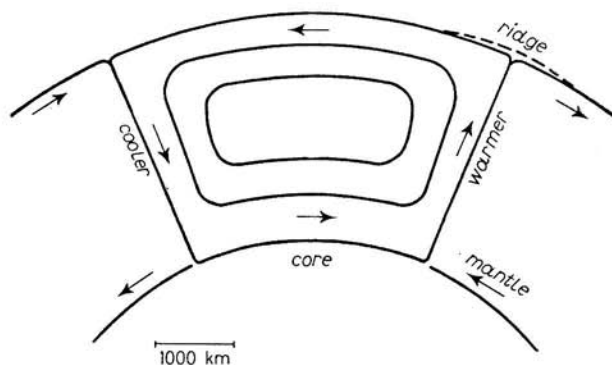


Fig. 11. — Sketch of the convective circulation envisioned by HESS [1]. Convection currents fill the entire mantle.

3. – The mantle as a fluid: viscosity estimates from postglacial rebound.

As mentioned in the introduction, the ideas of plate tectonics and mantle convection depend upon the assumption that the Earth's mantle deforms as a viscous fluid for processes of sufficiently long time scale. In order to verify this assumption, we require independent data and such is available by consideration of the response of the Earth to the massive continental deglaciations which occurred at the end of the last ice age. We show here that the relaxation data can be understood in terms of a Newtonian viscous Earth model and employ the model to infer the characteristic viscosity profile of the planetary interior. The question as to whether the magnitude of the viscosity which we obtain is compatible with the convection hypothesis is of course the crucial one and is addressed in detail in sect. 4 and 5.

3.1. *Glacial isostatic adjustment: the data.* – The most reliable information concerning the response of the Earth to deglaciation consists of relative-sea-level data which constitute local records of the time-dependent separation between the geoid and the surface of the solid Earth. Such data are obtained from flights of uplifted or drowned beaches, such as that shown in fig. 12, which is found in the Richmond Gulf of Hudson Bay. The age of each beach in the



Fig. 12. – Photograph of a staircase of raised beaches in the Richmond Gulf of Hudson Bay (photo courtesy of Prof. C. HILLAIRE-MARCEL).

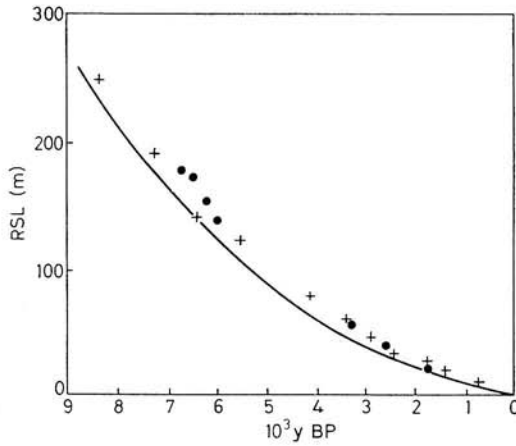


Fig. 13. — Relaxation curve from the ^{14}C dates for the Richmond Gulf beaches compared with a theoretical calculation (solid curve). The data are from ref. [49].

sequence can be determined using radiocarbon dating and, if the height of each horizon above present-day sea level is plotted as a function of its age, then one obtains a relaxation curve such as the one shown in fig. 13 for the beaches in fig. 12 [49]. The solid curve on this figure is the prediction of RSL (relative sea level) for this location from a theoretical model which will be described below.

When one considers such relaxation data from a global perspective, a striking pattern emerges. In fig. 14 we show the geographical distribution of ^{14}C dated

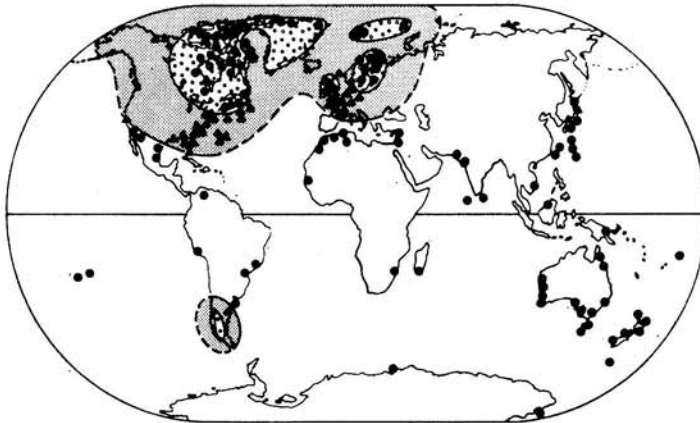




Fig. 14. — Geographical distribution of beach material in the age range $(0-5) \cdot 10^3$ BP plotted according to its disposition with respect to present-day sea level. Note that the regions of emergence correspond to negative anomalies in the global free-air anomaly map shown in fig. 10 (from ref. [50]).  region of postglacial rebound,  peripheral zone of submergence, • dated marine shells above present sea level, ▲ dated terrestrial peats below present sea level.

beach material in the age range $(0 \div 5) \cdot 10^3$ BP plotted on the basis of whether this material is found above or below present-day sea level [50]. In the regions of open circles over the Hudson Bay and Fennoscandia the present sense of the vertical motion is up (rebound), while in the immediately peripheral regions shown as solid triangles the sense is down (submergence). This pattern is one which any model of glacial isostasy must be able to reproduce. It is a pattern which may be easily understood on the basis of mass conservation if the Earth behaves as an incompressible viscous fluid on such long time scales. While the Pleistocene ice sheets were at their maximum extent, the gravitational interaction between them and the planet resulted in a depression of the surface beneath them. The matter forced from beneath the ice sheets collected in the immediately peripheral region where the local radius of the planet was increased. When the ice sheets melted, mantle material flowed from the bulge towards the depression in order to restore gravitational equilibrium. This resulted in the collapse of the peripheral region (submergence of the land there) and the uplift of the surface of the solid Earth where it was initially ice covered (rebound).

From the relaxation curve shown in fig. 13 we can obtain an estimate of mantle viscosity immediately if we assume the Earth to be a uniform Newtonian viscous sphere. For such a sphere, with constant density and viscosity, the free decay time for a spherical harmonic deformation of degree l is just $\tau = (\nu/\rho g_s)(2l^2 + 4l + 3)/la$ [51], where ν and ρ are viscosity and density, g_s is the surface gravitational acceleration, and a is the radius. Using parameters appropriate for the Earth, taking $l \simeq 5$ from the scale of the Laurentide ice sheet [52] and using the observed relaxation time $\tau \simeq 2.5 \cdot 10^3$ years from fig. 13, we obtain $\nu \simeq 10^{22}$ poise (CGS units) for the viscosity of the mantle. This estimate may be crude and we are obliged to refine it.

3.2. Anelastic Earth models. — The first step in this process of refinement is to improve the Earth model. Clearly the Earth does not behave as a viscous fluid on all time scales, it rather behaves as an elastic body for short-time-scale deformations and as a viscous fluid for processes of sufficiently long time scale. The simplest viscoelastic model which can accommodate such behaviour is that for a linear Maxwell solid. The stress-strain relation for this material [51] is

$$(1) \quad \tilde{\tau}_{kl} = \lambda(s) \tilde{\epsilon}_{kk} \delta_{kl} + 2\mu(s) \tilde{\epsilon}_{kl}$$

in the Laplace transform domain with s the Laplace transform variable. The compliances $\lambda(s)$ and $\mu(s)$ are

$$(2a) \quad \lambda(s) = \frac{\lambda \cdot s + \mu K/\nu}{s + \mu/\nu},$$

$$(2b) \quad \mu(s) = \frac{\mu \cdot s}{s + \mu/\nu},$$

where λ and μ are the usual elastic Lamé parameters and $K = \lambda + \frac{2}{3}\mu$ is independent of s . From (2) it is clear that for $t \ll T_m = \nu/\mu$ the material behaves as a compressible solid, while for $t \gg T_m$ it behaves as an incompressible viscous fluid to the extent that $\lambda(s) \cdot \tilde{\epsilon}_{kk} \rightarrow -p$, a constant. For the Earth, the Maxwell time $T_m \simeq 200$ years in a region where the viscosity $\nu \simeq 10^{22}$ poise. Visco-elastic processes which act on time scales less than this, such as those which produce seismic velocity dispersion and determine the Q 's of elastic gravitational normal modes, will not be correctly described by the constitutive relation (1).

The dynamical behaviour of an Earth model which consists of material described by (1) is governed by the following linearized and Laplace transformed equations for momentum balance and for the perturbation of the gravitational potential

$$(3a) \quad \nabla \cdot \tilde{\boldsymbol{\tau}} - \nabla(\rho g \tilde{\mathbf{u}} \cdot \mathbf{e}_r) - \rho \nabla \tilde{\varphi} + g \nabla \cdot (\rho \tilde{\mathbf{u}}) \mathbf{e}_r = 0,$$

$$(3b) \quad \nabla^2 \tilde{\varphi} = -4\pi G \nabla \cdot (\rho \tilde{\mathbf{u}}),$$

where $\rho = \rho(r)$ is the density field in the basic hydrostatic equilibrium configuration, $g = g(r)$ is the corresponding gravitational acceleration, \mathbf{u} is the displacement field, φ the perturbation of the ambient gravitational potential and G the gravitational constant. In the momentum equation (3a) the stress tensor $\boldsymbol{\tau}$ is given by (1) and the inertial force has been suppressed because of the long time scale of the phenomenon which concerns us.

We may find primitive solutions to (3) for (\mathbf{u}, φ) when the Earth is deformed by gravitational interaction with a point mass load which is placed on its surface at $t = 0$ and instantaneously removed. If the physical properties of the interior are functions of r only, then we may expand \mathbf{u} and φ as

$$(4a) \quad \mathbf{u} = \sum_{n=0}^{\infty} \left(U_n(r, s) P_n(\cos \theta) \mathbf{e}_r + V_n(r, s) \frac{\partial P_n}{\partial \theta}(\cos \theta) \mathbf{e}_\theta \right),$$

$$(4b) \quad \varphi = \sum_{n=0}^{\infty} \varphi_n(r, s) P_n(\cos \theta).$$

Subject to the appropriate boundary condition on $r = a$ we may construct solutions for the spectral amplitudes U_n, V_n, φ_n by solving a set of six coupled ordinary differential equations using a shooting method [51]. This set of equations is the same as that which describes the elastic gravitational (spheroidal) free oscillations and is shown explicitly in the notes by GILBERT in this volume. A useful representation of these solutions on the Earth's surface is in the following form:

$$(5) \quad \begin{bmatrix} U_n \\ V_n \\ \varphi_{1,n} \end{bmatrix} = \varphi_{2,n}(r) \begin{bmatrix} h_n(r, s)/g \\ l_n(r, s)/g \\ k_n(r, s) \end{bmatrix},$$

where $\varphi_n = \varphi_{1,n} + \varphi_{2,n}$ and $\varphi_{2,n}$ is the perturbation of the gravitational potential due to the point mass on the surface, which is independent of s because the point mass is applied as a Dirac delta-function in the time domain. The triplet of dimensionless scalars (h_n, l_n, k_n) constitute the nondimensional spectral form of the impulse response of the system. They are the viscoelastic analogues of the surface load « Love » numbers of elasticity. An example of the spectral surface $h_n(a, s)$ is shown in fig. 15 for an Earth model in which $\nu = \infty$ to a depth of 112.5 km (the lithosphere), $\nu = 10^{22}$ poise between the base of the lithosphere and the core-mantle boundary, and $\nu = 0$ throughout the core. The elastic structure of the model $(\rho(r), \mu(r), \lambda(r))$ fits the free-oscillation data [52]. In fig. 15 I have in fact plotted $h_n^v(a, s) = h_n(a, s) - h_n^E$, where h_n^E is the large- s elastic asymptote for each value of n , as discussed in [51].

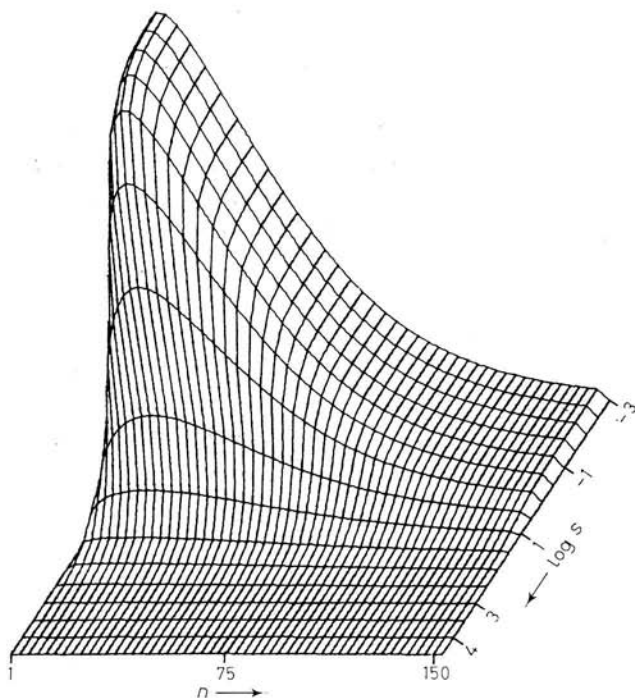


Fig. 15. — Laplace transform domain relaxation surface. Note that, as $n \rightarrow \infty$, $h_n^v \rightarrow 0$ and the viscous relaxation is entirely suppressed.

It may be shown directly [53] that these spectra have exact normal-mode expansions of the form

$$(6) \quad h_n(a, s) = \sum_j \frac{\gamma_j^n}{s + s_j^n} + h_n^E,$$

where s_j^n are a set of poles (a different set for each n) which lie on the negative real axis in the complex s -plane. The τ_j^n are simply the residues at these poles and thus measure the extent to which a given normal mode is excited by the point forcing. In fig. 16 is shown a sequence of relaxation diagrams for a series of illustrative Earth models in order to demonstrate the effect upon the modal structure of each of the main features of the planetary model [54].

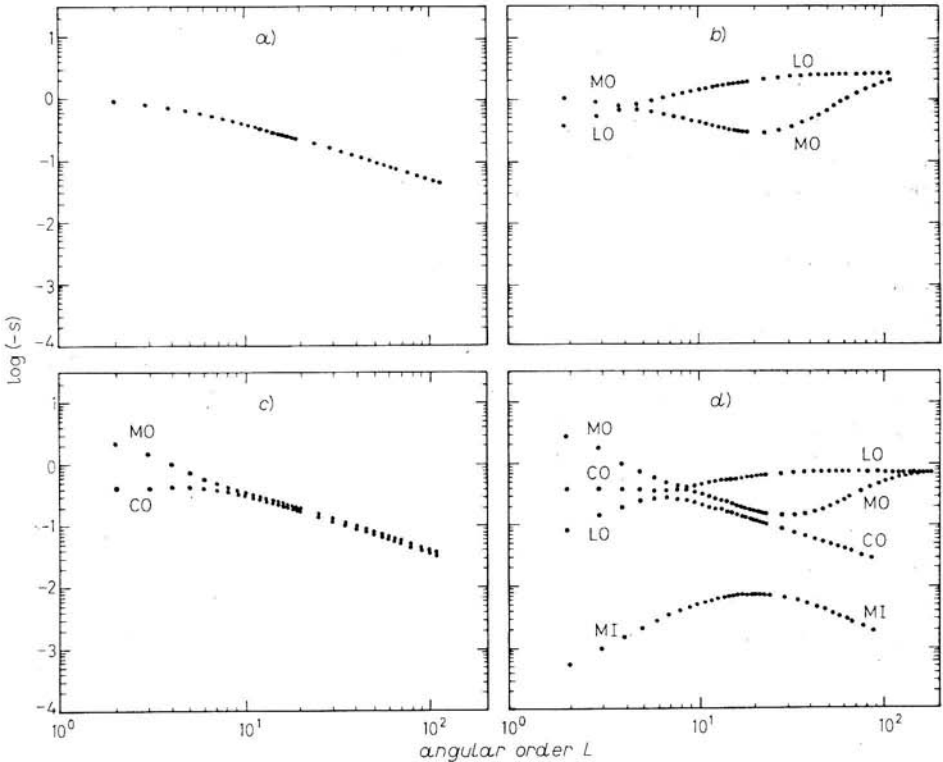


Fig. 16. - Relaxation diagrams for the four Earth models described in the text. The relaxation time $\tau = s^{-1}$ and times are measured in units of 10^3 years.

Figure 16a) is the relaxation diagram for an Earth model in which all of ρ , λ , μ , ν are constant and illustrates the relationship between relaxation time $\tau = s^{-1}$ and spherical wave number $n = l$ discussed above (all times are nondimensionalized with a characteristic time of 10^3 years). For sufficiently large l , τ increases linearly with l in accord with the prediction of the equivalent half-space model which gives $\tau = 2\nu k_H / g_s \cdot \rho$, where k_H is the horizontal wave number. In fig. 16b) we show the effect upon the relaxation spectrum of including a lithosphere at the surface of the planet and for illustration the the viscosity in this region has been taken to be infinite and the thickness equal to 112.5 km. Two effects are apparent. Firstly the presence of the lithosphere

reduces the relaxation times for the short deformation wavelengths (large l). Secondly it introduces a completely new relaxation time for each value of l , so that for each wavelength there are now two accessible modes. The two modes have been labelled M and L on the figure, the first corresponding to mantle, the second to lithosphere. The two modal lines coalesce at large l , a mathematical manifestation of the physically intuitive result that for sufficiently short wavelength all viscous gravitational relaxation is suppressed. Such short deformation wavelengths are supported elastically. In fig. 16c) we show a relaxation diagram for an Earth model with an inviscid core, a constant-viscosity mantle and no lithosphere. Again, the effect of the presence of the core, with a strong density jump across the core-mantle boundary, is to introduce a second relaxation time for each value of l . For this model both modes have relaxation time increasing as the wavelength decreases. In fig. 16d) we show the relaxation diagram for a complete Earth model. This has a λ, μ, ρ structure which is identical to 1066A of Gilbert and Dziewonski [52]. In addition it has an inviscid core and a lithosphere which is 112.5 km thick. Throughout the mantle the viscosity is 10^{22} poise. Inspection of this figure shows that the relaxation diagram contains each of the three modal branches discussed above, but that there are additional modes due to the strong density jumps in the mantle.

The viscoelastic nature of the real Earth is extremely important insofar as the understanding of observed relaxation data is concerned. The effect of the elastic lithosphere provides an interesting case in point and we will proceed here to demonstrate that such a layer exists and to determine its thickness. The thickness of the surficial elastic lithosphere is a very useful datum in the convection hypothesis, in fact we shall argue in later sections that it provides an accurate measure of the thickness of the thermal boundary layer of the mantle convective circulation.

Figure 17 shows the relaxation spectrum determined by McCONNELL [8] from

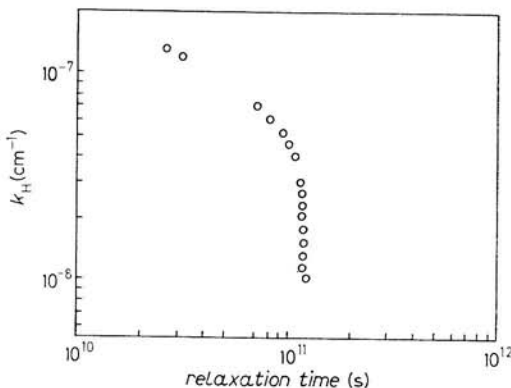


Fig. 17. — Relaxation spectrum from the Fennoscandian rebound data from ref. [8].

the Fennoscandian rebound data. The points to note are that the longest wavelengths show a relaxation time of about $4 \cdot 10^3$ years and that there is a knee in the relaxation spectrum such that for wavelengths less than about 500 km there is a sharp decrease in relaxation time. This is precisely the effect expected of the presence of a thin elastic layer (lithosphere) on the surface, as demonstrated in fig. 16. We may further constrain the thickness of the elastic layer by invoking Crittenden's [54] data for Pleistocene Lake Bonneville, which imply a relaxation time $\tau \simeq 5000$ years for this feature, which has an effective l of about 150, *i.e.* a horizontal scale which is an order of magnitude less than in Fennoscandia. These data together fix the effective elastic thickness of the continental lithosphere to be somewhat in excess of 100 km.

Knowing the relaxation surfaces $h_n(a, s)$, $k_n(a, s)$ and the relaxation spectra s_j^n , we can invert (6) into the time domain to get

$$(7) \quad h_n(a, t) = \sum_j r_j^n \exp[-s_j^n t] + h_n^E \delta(t),$$

from which we may construct space-time Green functions of the problem for various signatures of the response. We find the Heaviside form of the functions most useful and the spectral amplitudes for this form may be obtained from (7) by convolution with a unit step. This yields (suppressing a)

$$(8) \quad h_n^H(t) = \sum_j \frac{r_j^n}{s_j^n} (1 - \exp[-s_j^n t]) + h_n^E = h_n^H(t) + h_n^E.$$

We have previously shown [53] that the Green functions for radial displacement, gravity anomaly and the perturbation of the gravitational potential are

$$(9a) \quad u_r^H(\theta, t) = \frac{a}{M_e} \sum_{n=0}^{\infty} h_n^H(t) P_n(\cos \theta),$$

$$(9b) \quad \Delta g^H(\theta, t) = \frac{g}{M_e} \sum_{n=0}^{\infty} [n - 2h_n^H(t) - (n + 1)k_n^H(t)] P_n(\cos \theta),$$

$$(9c) \quad \varphi^H(\theta, t) = \frac{ag}{M_e} \sum_{n=0}^{\infty} (1 + k_n^H(t) - h_n^H(t)) P_n(\cos \theta),$$

respectively. The viscous part of $u_r^{H,v}(\theta, t)$ is shown in fig. 18 for an Earth model which has constant mantle viscosity and a lithosphere which is 112.5 km thick. Inspection of this Green function shows that it does contain the symmetry required to explain the rebound pattern of fig. 14. At small θ , $u_r^{H,v} < 0$, while, for sufficiently large θ , $u_r^{H,v} > 0$, which is the region of the peripheral bulge discussed previously. We proceed to show how (9) may be employed to make RSL predictions.

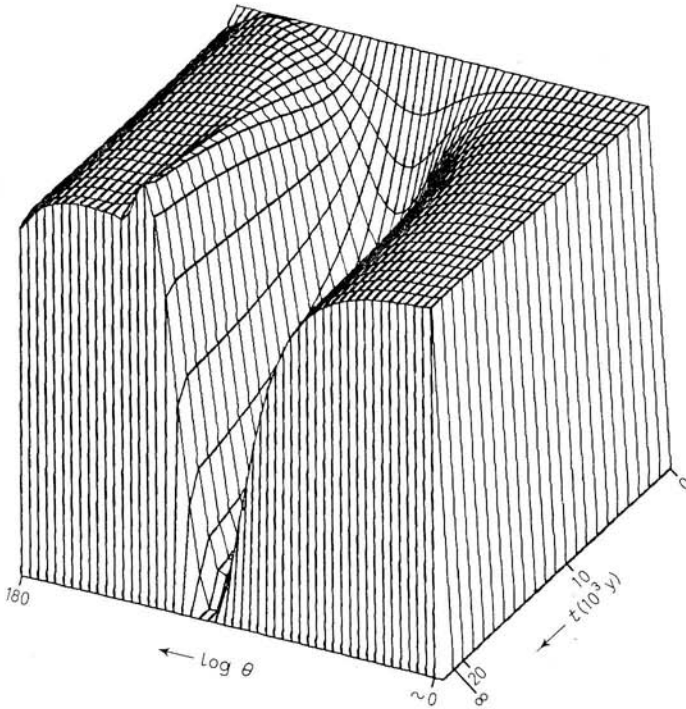


Fig. 18. — Viscous part of the Green function for radial displacement for an Earth model with constant mantle viscosity and a lithosphere which is 112.5 km thick. The response has been normalized by multiplication with $a\theta$ to remove the effect of the geometric singularity at $\theta=0$ for plotting.

3.3. Gravitationally self-consistent sea level calculations. — We need a theory which will calculate relative sea level directly, so that we may predict relaxation curves such as that shown in fig. 13. The problem of glacial isostasy consists of describing the gravitational interaction between the three component parts of the system: the anelastic Earth, the ice sheets on its surface and the oceans into which the ice sheets discharge their meltwater. The strategy which we adopt to solve this problem is as follows. Suppose we knew the melting histories of the major Pleistocene ice sheets exactly and let us enquire as to where in the global oceans the meltwater from these ice sheets must go when they disintegrate. So long as the perturbations of the gravitational potential produced by this melting event are sufficiently small, it is in fact possible to answer this question exactly. The key idea is the recognition that water must redistribute itself in the oceans in such a way that the oceanic surface remains an equipotential surface [55-57].

We can calculate the perturbations of the potential field which are produced by both the ice and the ocean parts of the surface load simply by convolving this surface load with the potential perturbation Green's function (9c). Insisting

that the ice-water surface load conserves mass and that the potential on the surface of the ocean remains constant at all times (the value of the constant changes), we may obtain an integral equation for the direct calculation of relative sea level. The construction of this integral equation and a description of the numerical algorithm used for its solution has been described in [56]. A schematic form of this equation is

$$(10) \quad S = \varrho_{\text{I}} \frac{\varphi^{\text{H}}}{g} *_{\text{I}} L + \varrho_{\text{w}} \frac{\varphi^{\text{H}}}{g} *_{\text{O}} S + C,$$

where S is exactly the relative-sea-level variation, an example of which was seen in fig. 13. In writing (10) I have suppressed the melting history for convenience. ϱ_{I} is the density of ice, ϱ_{w} that of water, and C is a constant which is fixed by the requirement of mass conservation. Clearly (10) is an integral equation since the convolution operations $*$ are integral relations (I means « over the ice », and O means « over the oceans ») and thus the unknown relative sea level S appears both on the left and under the integral on the right.

The input to (10) is just L , the deglaciation history, and the methods which are employed to obtain *a priori* estimates of this field are discussed in [58]. Given L , we seek to find $\nu(r)$ such that the solutions to (10) fit the global set of relaxation data. Since both L and ν are imperfectly known initially, the inverse problem is nonlinear, but, if we can find a solution $(L, \nu(r))$ which is linearly close to a solution, then we can converge iteratively using the formal inverse theory developed in [53]. Most of our effort to date has been expended in getting linearly close, since our first guess to the L field turned out to be grossly in error, particularly for the Laurentide ice sheet. This led to large misfits between the observed and predicted RSL curves in this region [57] because the Laurentide ice sheet had been taken to be much too thick. The analysis leading to this conclusion is discussed in [59].

The solution to (10) for a realistic deglaciation history consists of a sequence of maps showing the rise (fall) of sea level relative to the surface of the solid Earth as a function of time since melting commenced. An example of such a solution is shown in fig. 19, where relative sea level is contoured in metres for four times (given in thousands of years BP). From the complete sequence of such maps we may extract a relative-sea-level curve for any site for which we have RSL data available. A sequence of comparisons between observation and theory is shown in the next subsection.

3.4. Comparisons between observation and theory. — On the basis of the global solutions to (10), an example of which we have just seen in fig. 19, we may divide the surface of the oceans into a number of distinct zones in each of which the relative-sea-level histories are characteristic. These regions have been discussed in detail in [57, 58] and we shall not reproduce this discussion here

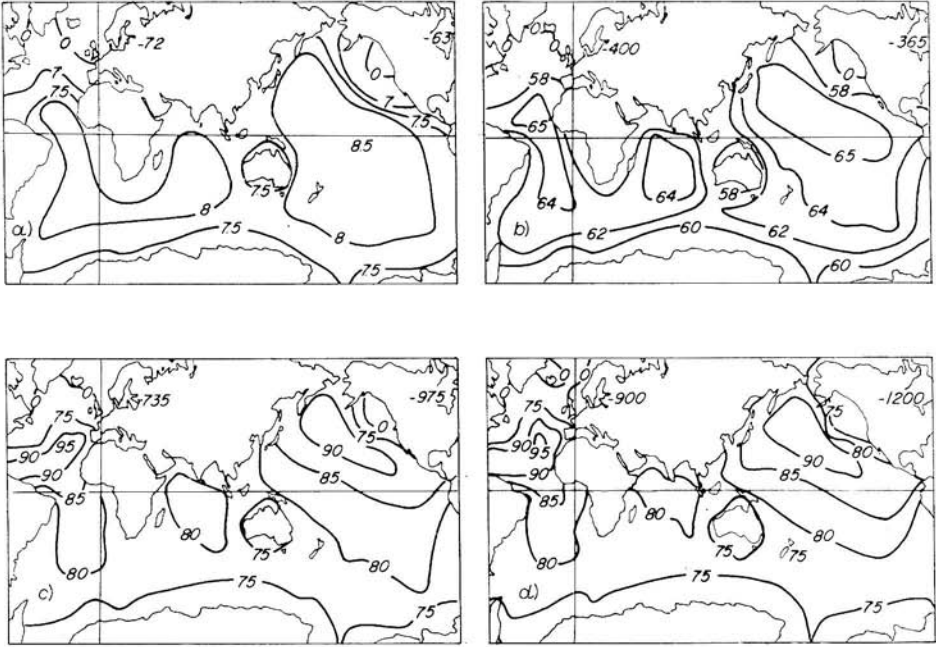


Fig. 19. — Four time slices through a solution of the equation for the gravitationally self-consistent variation of relative sea level. The contours are marked in metres of RSL. Times are a) 13 000 y BP, b) 8 000 y BP, c) 5 000 y BP, d) present.

except to say that the predictions fit the global data set remarkably well in all regions, particularly after refinement of the deglaciation history as discussed in [59]. Using the new deglaciation history and an Earth model which has a lithosphere which is 112.5 km thick and a mantle in which the viscosity is a constant 10^{22} poise between the base of the lithosphere and the core-mantle boundary, we shall proceed to illustrate the extent of the agreement between prediction and observation.

The first such comparison has already been illustrated in fig. 13 for the Richmond Gulf which is located in Hudson Bay near the centre of the Laurentide rebound. This is one of the best controlled relaxation curves yet obtained in the central region and it is clear that the uniform-viscosity model fits it very well. In fig. 20 we show six further comparisons for more sites in the central region and for a series of locations extending along the east coast of North America from the North (Newfoundland) where the response is pure uplift, through the transition region governed by the migration of the forebulge (Bay of Fundy), southward to Carolina in the region of peripheral bulge collapse and then offshore to Bermuda. At each of these locations the uniform-viscosity model fits the relaxation data beautifully providing strong circumstantial evidence that it is correct.

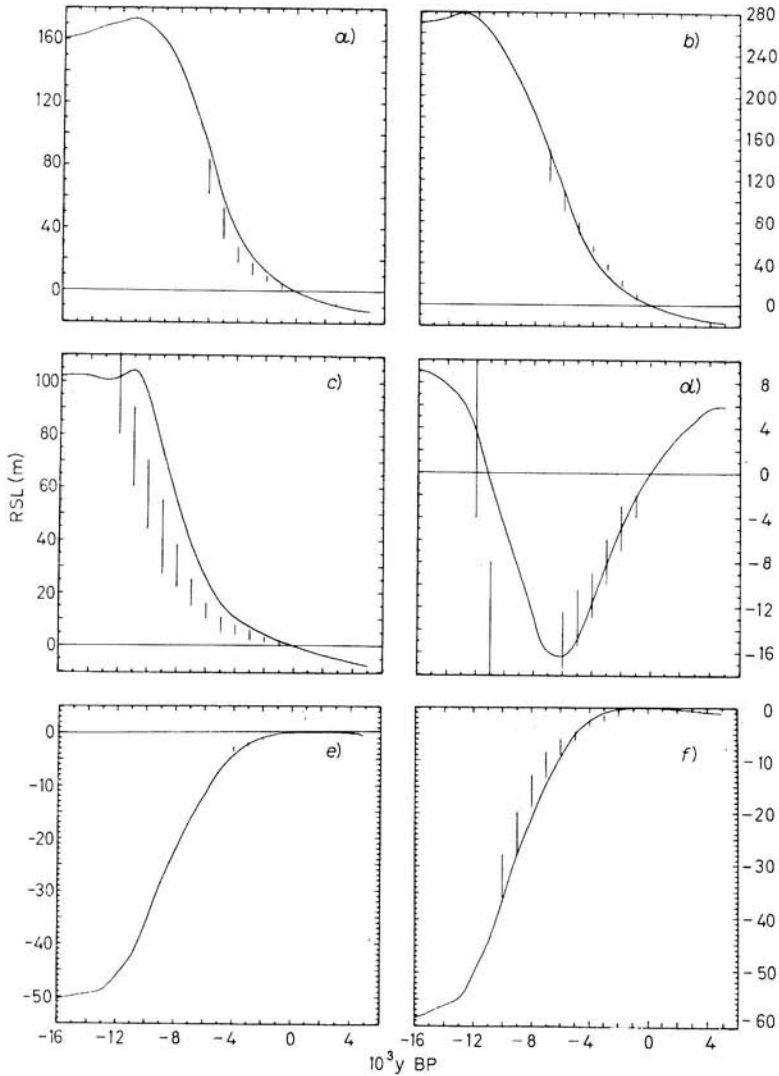


Fig. 20. — Comparison of observed and theoretically predicted RSL variations at a series of sites in North America. The locations are a) Ipiq Bay, b) Churchill, c) North West Newfoundland, d) Bay of Fundy, Nova Scotia, e) Southport North Carolina, f) Bermuda. The solid lines are the predicted relaxation curves and the vertical bars are the data. The length of each vertical line represents an estimate of experimental error. Predictions are for the model with uniform mantle viscosity.

Such evidence is, however, only circumstantial and, in order to establish the uniform-viscosity model as *the* correct model of mantle viscosity, we must establish that other perhaps radically different models are not equally able to fit the observational constraints. The main such alternative in which we are

interested here is a model which is everywhere identical to that discussed above except in the lower mantle, where the viscosity is considerably higher than 10^{22} poise. In [59] it is shown that at least one version of this model is not strongly rejected by the relative-sea-level data and this has a jump in ν from 10^{22} P to 10^{24} P at a depth of 10^3 km. Previous indications to the contrary [58] appear to have been due to inadequacies of both the melting history and the RSL model.

Our reasons for believing that models with some increase of viscosity with depth may be relevant are basically two in number; that is besides the new evidence that such models are not strongly rejected by the RSL data. In the first instance, since mantle viscosity is a consequence of a creep process which is thermally activated and since the activation energy is expected to increase through a phase change [60] such as the olivine \rightarrow spinel or the spinel \rightarrow post spinel phase change, we have a physical reason to expect such an increase. Even if there were no phase changes, some increase of viscosity might be expected on the basis of Weertman's [61] phenomenological creep law which gives the viscosity ν in the form $\nu = \nu_0 \exp [g^* T_m / T]$, where g^* is a constant, $T_m(r)$ is the melting temperature and $T(r)$ the geotherm. If the spherically averaged geotherm $T(r)$ is adiabatic, as would be expected if the mantle were convecting, then, since $T_m(r)$ is steeper than adiabatic, ν should increase with depth. Whether this empirical relation is in fact valid, however, remains a question. The second reason for preferring some increase of mantle viscosity with depth concerns the inability of the uniform-viscosity model to fit a further observational datum; that is the free-air gravity anomaly over the central depression. Although this datum has been previously ignored in discussions of isostasy [58, 62], evidence is accumulating that it should not be. We previously drew attention to the negative anomalies over Hudson Bay and Fennoscandia in the free-air anomaly map of fig. 10. In the past there has been a great deal of controversy as to whether these anomalies should be entirely attributed to deglaciation-induced deviations from isostatic equilibrium or whether they were related to some other cause. Two pieces of evidence suggest an origin in glacial isostasy: 1) the zero-anomaly contours are almost exactly coincident with the original boundaries of both ice sheets which would be expected if the origin were glacial, and 2) within this boundary of the Laurentide depression, even smaller-scale local minima appear now to be correlated with original local maxima of ice sheet thickness. This suggests that we should try simultaneously to fit both the relative-sea-level and the gravity data.

Given a solution S to (10) we may combine it with L to produce a completely self-consistent surface mass land distribution as a function of space and time on the surface of the Earth. This may be convolved with the Green function for the gravity anomaly in (9b) to produce maps of the free-air gravity anomaly as a function of time. In fig. 21 we compare the predicted present-day free-air anomaly map, obtained using the same Earth model and load

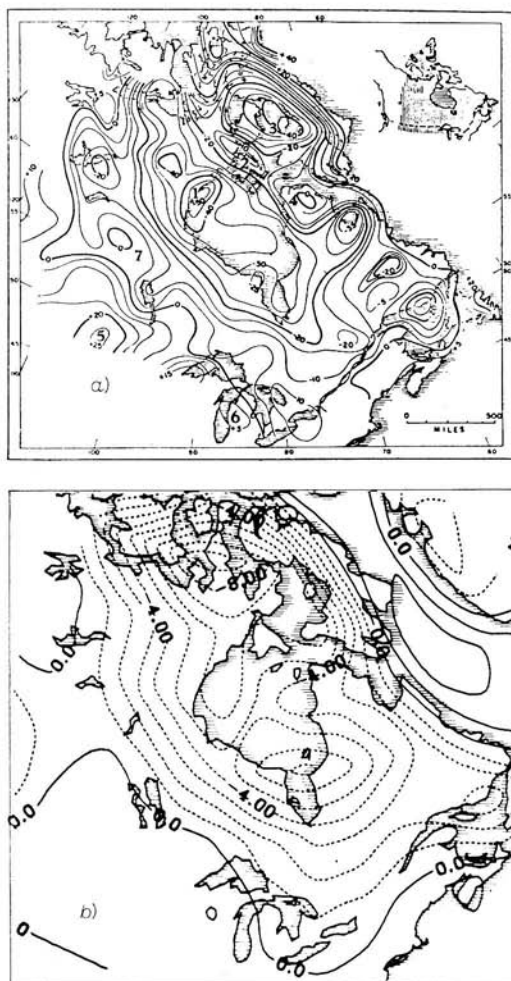


Fig. 21. - Comparison of the observed free-air gravity anomaly map (a) for the Laurentide region with the prediction using the model with uniform mantle viscosity (b). Contours are in milligal. Station density in $1^\circ \times 2^\circ$ squares: $\text{|||||} > 50$, $\text{||||} 10 \div 50$, $\text{||} 1 \div 10$, $\square 0$.

history as led to the sea level data shown in fig. 20, with the free-air anomaly map from [63]. The observed maximum negative free-air anomaly is about 35 mGal, while the prediction with the uniform-viscosity model is about $(8 \div 10)$ mGal. Note that the zero-anomaly contours in the two maps are almost exactly coincident. The general pattern of the observed anomaly is, therefore, exactly what one expects if it is due to deglaciation, the transition from the interior negative anomaly to the exterior positive anomaly is almost exactly coincident with the edge of the ice sheet and this is strongly constrained by end moraine data.

On the basis of the free-air data it is clear that the model with uniform mantle viscosity which fits the relative-sea-level data so accurately is not able simultaneously to reconcile the observed gravity anomaly. As pointed out in [59], however, it should be possible to reconcile both data sets with a model in which the viscosity increases from the upper to the lower mantle sufficiently to prevent the too rapid relaxation of the free-air anomaly, but not so much that the fit to the relative-sea-level data is destroyed. We require a complete analysis of the trade-offs involved in this process of model fitting and this is a topic of current research using the formal inverse theory developed in [53]. Once such a model has been constructed, we will be able to employ it to remove the glacial signal from the global free-air map and, after this filtering has been accomplished, it could well be that the signature of mantle convection in the global gravity field will become much more apparent.

3.5. Implications of the viscosity profile for the convection hypothesis. — In the first instance the above analysis provides direct confirmation of the fact that on time scales in excess of a few hundred years the mantle of the Earth does indeed behave like a viscous fluid, at least at depths in excess of about 100 km. Below this lithosphere, in which the viscosity may be so high that the concept loses meaning, ν is on the order of 10^{22} poise. This number is extremely important, as we shall see. If the observed free-air gravity anomaly is due to glacial isostatic disequilibrium, then the viscosity of the mantle must increase through the transition zone by perhaps 1–2 orders of magnitude. Such an increase would be completely insufficient to prevent whole-mantle convection [64], but it would probably be sufficient to explain why the deep earthquake focal mechanisms are compressive [31]. We will see in the following discussion of convection that the thickness of the lithosphere and the magnitude of the viscosity beneath it place important constraints on the convection hypothesis of drift and spreading.

4. — Convection in the laboratory and Oberbeck-Boussinesq theory: a preliminary model of mantle convection.

Before attempting to discuss the problem of convection in the Earth it is probably wise to begin with a discussion of convection under the circumstances in which it is best known—that is in the controlled conditions of the laboratory. In addition we shall concentrate in this section entirely on the problem of convection in a fluid with constant physical properties. There is a vast literature on this topic and we will certainly not be able to review it here but rather direct the interested reader to the recent discussion by BUSSE [65]. The treatment presented here will be made somewhat novel by its attempt to describe the heated below and heated internally problems in parallel. In this way we

wish to direct immediate attention to the importance of the question as to how the convection is driven. The relevance of the internally heated problem to convection in the Earth's mantle should be clear. It is now conventionally assumed that this convection may be driven primarily by the decay of the long-lived radioactive isotopes and these clearly act as a volume heat source.

4.1. *The basic equations and nondimensional parameters—linear theory.* —

Consider a plane layer or spherical shell of fluid of thickness d subject to constant gravitational acceleration g directed everywhere normal to the layer or shell. The fluid has constant coefficient of thermal expansion α . We begin by nondimensionalizing the hydrodynamic equations in the manner first introduced in [64], choosing the scales for the field variables listed in table I which are themselves functions of the natural temperature scale A which depends upon the heating configuration. The variables T, ρ, u_i, x_i, p, t are, respectively,

TABLE I. — *Scaling: liquids.*

$T = AT'$	$p = \rho_0 g d p'$
$\rho = \rho_0 \rho'$	$t = ((\mu_0/\rho_0)g\alpha d A) t'$
$u_i = (g\alpha d^2 A^{1/2}(\mu_0/\rho_0)) u'_i$	heated below: $A = T_l - T_u$
$x_i = dx'_i$	heated internally: $A = d^2(Q_0/K_0)$

temperature, density, velocity components, length, pressure, time. Zero subscripts imply evaluation at some reference level. In the definitions of A in table I, T_l and T_u are, respectively, the constant temperatures on the lower and upper surfaces in the heated below problem and Q is the rate of internal heat addition per unit volume for the internally heated problem. Substituting these scales into the hydrodynamic equations and completing the system with a simple linear equation of state, we obtain the following nondimensional system [64]:

$$(11) \quad \rho \frac{\text{Ra}}{\text{Pr}} \frac{du_i}{dt} = \frac{1}{\delta} (-\partial_i p + \rho \lambda_i) + \partial_j \left[2f \left(e_{ij} - \frac{1}{3} \Delta \delta_{ij} \right) \right],$$

$$(12) \quad \partial_i \rho + \partial_j (\rho u_j) = 0,$$

$$(13) \quad \frac{dT}{dt} - \frac{T\tau}{\rho} \frac{dp}{dt} = \frac{1}{\rho \text{Ra}} (\partial_j k \partial_j T + h) + \frac{2\tau f}{\rho} \left(e_{ij} e_{ij} - \frac{1}{3} \Delta^2 \right),$$

$$(14) \quad \rho = 1 - \delta(T - T_0) + \Gamma(p - p_0),$$

where $e_{ij} = (\partial_j u_i + \partial_i u_j)/2$ is the strain rate tensor, $\Delta = \partial_j u_j$ is the divergence of the velocity field, and $\lambda = (0, 0, 1)$ is a unit vector in the z -direction. The nondimensional parameters Ra, Pr, τ, δ, Γ are defined in table II for both the heated below and heated internally problems. In (11)-(14) the functions f, k and h represent variations of molecular viscosity, thermal conductivity

and volume content of heat sources relative to a fixed reference level, *i.e.* $Q = Q_0 h$, etc.

TABLE II. - *Nondimensional parameters: liquids.*

	Heated below	Heated internally
Ra (Rayleigh number)	$g\alpha d^3 \Delta T / (K_0 / \varrho_0 C_p)(\mu_0 / \varrho_0)$	$g\alpha d^5 (Q_0 / \varrho_0 C_p) / (K_0 / \varrho_0 C_p)^2 (\mu_0 / \varrho_0)$
Pr (Prandtl number)	$(\mu_0 / \varrho_0) / (K_0 / \varrho_0 C_p)$	same
$\delta = \alpha t$	$\alpha \Delta T$	$\alpha d^2 (Q_0 / K_0)$
τ	$g\alpha d / C_p$	same
Γ	$\varrho_0 \gamma g d$	same

The nondimensionalization of the hydrodynamic equations employed here, which was first introduced in the context of the mantle convection problem in [64], is not that which is usually used in the discussion of thermal convection. Normally, *e.g.* [65], one employs scales for length, time and temperature which are, respectively, d , d^2/ν and $(T_1 - T_u)/\text{Ra}$. This conventional scaling leads to a ratio of the inertial force to the viscous force in the nondimensional system which is Pr^{-1} rather than RaPr^{-1} as we have obtained. Furthermore, the ratio of the nonlinear thermal-convection term to the thermal-diffusion term is Ra in our scaling rather than unity as obtains with the conventional one. The scaling employed here is natural for convection at high Rayleigh number when the circulation is vigorous, since it is physically motivated by balancing the viscous force against the buoyancy force and this balance of forces is precisely that which obtains in the thin vertical plumes which dominate the heat transport at large Ra . The fact that the inertial force goes like RaPr^{-1} rather than Pr^{-1} in this scaling may be important with regard to the construction of laboratory models of the mantle circulation which attempt to ensure dynamic similarity. It is not sufficient, for this purpose, simply to use a fluid with $\text{Pr} \gg 1$; one must ensure $\text{RaPr}^{-1} \ll 1$ in order to eliminate the influence of the inertial force from the laboratory experiments and thus to ensure dynamic similarity with the mantle. The new scaling is also natural for convection in thick systems for which $\tau = O(1)$, since for such systems at high Rayleigh number the two terms in the energy equation which scale like τ are indeed of the same order of magnitude and balance in a global sense. This will be discussed in detail in sect. 5.

The Rayleigh number Ra is the principal nondimensional group in the thermal-convection problem. It measures the balance between the dissipation of kinetic energy by viscosity and the release of potential energy by the buoyancy force. The Prandtl number Pr is the ratio of the kinematic viscosity to the thermal diffusivity and for a steady flow is important only in the non-linear regime, where it governs the relative thicknesses of the viscous and thermal boundary layers. The remaining nondimensional parameters τ , Γ , δ measure

the degree of non-Boussinesq behaviour of the fluid. δ and Γ together control the run of density through the layer in the static-equilibrium state, while from (13) τ governs the importance of both compression work (the second term on the left of (13)) and the viscous dissipation (the last term on the right). The dissipation number τ was introduced by PELTIER [64] and, although it will be considered in detail in sect. 5, it is perhaps instructive to point out its physical meaning immediately. Consider an adiabatic process which according to the second law must satisfy

$$(15) \quad dS = c_p dT - \frac{\alpha T}{\varrho} dp = 0,$$

where S is the entropy. It, therefore, follows that

$$(16) \quad \frac{dT}{dr} = \frac{\alpha T}{\varrho c_p} \frac{dp}{dr},$$

but the radial stratification is hydrostatic if there is no motion, thus $dp/dr = -\varrho g$ and, therefore,

$$(17) \quad \frac{dT}{dr} = \frac{-\alpha g T}{c_p} \rightarrow T = T_0 \exp[(r_0 - r)/(c_p/\alpha g)].$$

The adiabatic scale height is thus $H_{\text{Tad}} = c_p/\alpha g$ and this is the depth over which the temperature changes by a factor of e if it varies adiabatically. The ratio of the depth d of the convecting region to this scale height is just

$$(18) \quad \tau = \frac{d}{H_{\text{Tad}}} = \frac{g\alpha d}{c_p},$$

the dissipation number introduced in [64].

The Oberbeck-Boussinesq equations are obtained from (11)-(14) by taking the limit $d \ll H_{\text{Tad}}$ or $\tau \ll 1$ and $\delta \ll 1$. From this follows the fact that density variations may be neglected everywhere except in the buoyancy term of the momentum balance equation (11). Furthermore, the transport properties are assumed to be independent of both pressure and temperature. These approximations reduce the set of equations (11)-(14) to the new system

$$(19) \quad \frac{\text{Ra}}{\text{Pr}} \frac{du_i}{dt} = \frac{1}{\delta} (-\partial_i p + \varrho \lambda_i) + \nabla^2 u_i,$$

$$(20) \quad \partial_j u_j = 0,$$

$$(21) \quad \frac{dT}{dt} = \frac{1}{\text{Ra}} (\nabla^2 T + h),$$

$$(22) \quad \varrho = 1 - \delta(T - T_0),$$

which applies remarkably well to laboratory scale flows.

The first step in analyzing the convection equations (19)-(22) is to employ them in a linear stability analysis of the state-of-rest conduction solutions. This analysis proceeds by expanding each of the hydrodynamic fields ψ in the form $\psi = \psi_0 + \psi'$. When these expansions are substituted into (19)-(22) and the resulting equations linearized in deviations from the time-independent conduction solutions ψ_0 , we may seek normal-mode solutions of the form

$$(23) \quad \psi' = \left[\sum_n a_n \exp [i \mathbf{k}_n \cdot \mathbf{r} + \sigma_n t] \right] g(z, l \mathbf{k}_n l).$$

So long as the fluid has no « horizontal » mean flow in the conduction state, then it can usually be established that $\sigma_n = 0$ will mark the boundary between unstable (growing) and stable (decaying) solutions. In terms of the vertical component of the perturbation velocity w the stability of the layer is governed by the solution of the following o.d.e.

$$(24) \quad \left(\frac{d^2}{d\zeta^2} - k_H^2 \right)^3 w = -k_H^2 \text{Ra} \frac{dT_c}{d\zeta} w,$$

where $k_H^2 = k_x^2 + k_y^2$ and $T_c(\zeta)$ is the nondimensional conduction temperature profile. Equation (24) constitutes a two-point boundary-value problem on the interval $0 \leq \zeta \leq 1$, the eigenvalues of which are Ra and k_H . An example of the stability boundary $\text{Ra}(k_H)$ from [66] is shown in fig. 22 for the heated below problem when the horizontal boundaries are rigid as they would be in a laboratory experiment. The main idea is that for $\text{Ra} > \text{Ra}_c$ the conduction solution is unstable and that the circulation which replaces the state-of-rest solution will be periodic and characterized by a horizontal scale with wave number $k_{H \text{crit}}$. These critical parameters, which are the co-ordinates of the minimum on the stability boundary, are functions of the boundary conditions and the heating configuration, as is well known.

Linear theory leaves many questions unanswered. It is doubly degenerate in the sense that it is neither able to predict what the amplitude of convection will be for a given $\text{Ra} > \text{Ra}_c$, nor is it able to provide any information on the geometry of the planform to be expected—only that the width of a cell should be on the same order as its depth (*i.e.* aspect ratio ~ 1). These are questions which only nonlinear theory can resolve. However, before we discuss this theory we shall describe in the next subsection a few of the laboratory results which nonlinear theory must be able to explain.

4.2. Laboratory observations of high-Prandtl-number convection: heated below and internally. — In the general convection problem posed by (19)-(21), nonlinear flows are functions of both Ra and Pr. This parametric complexity is eliminated when one focuses on the problem of convection in a planetary mantle because of the magnitude of the Prandtl number. Indeed, since the kinematic

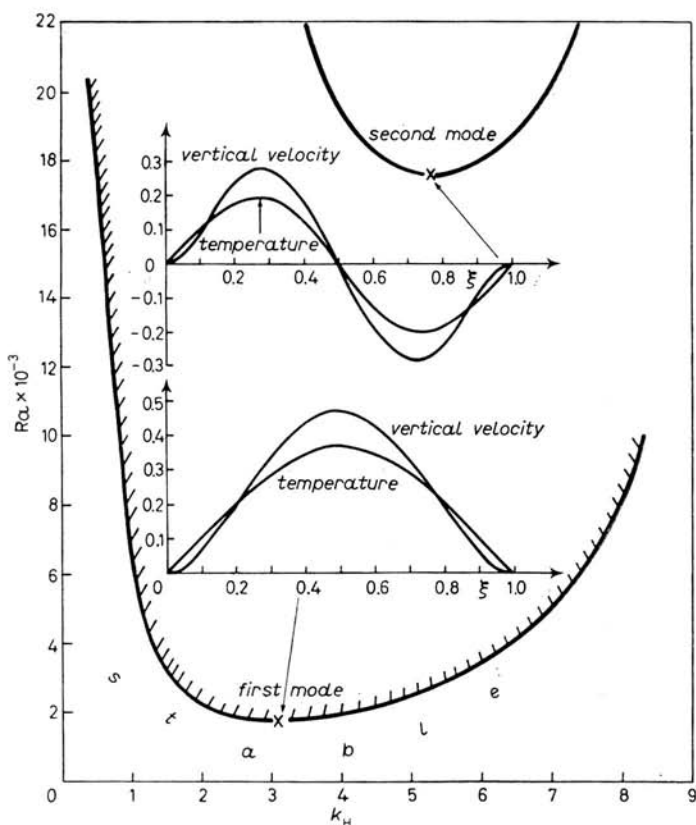


Fig. 22. — Linear stability boundary $Ra(k_H)$ for the heated below convection problem with two rigid boundaries. Eigenfunctions for the vertical velocity w and the temperature fluctuation θ are inset for the first two modes.

viscosity is $\sim 3 \cdot 10^{21} \text{ cm}^2 \text{ s}^{-1}$ (from postglacial rebound) and the thermal diffusivity is $\sim 10^{-2} \text{ cm}^2 \text{ s}^{-1}$ (from laboratory measurement), it follows that $Pr = O(10^{23})$. From eq. (19) it therefore follows that for any sensible Rayleigh number the inertial force will be completely negligible compared with the viscous force. Since this is the only term in the equations where the Prandtl number enters, it then ceases to be an important parameter in the problem. In discussing the laboratory observations we shall, therefore, focus on the properties of convection at infinite Prandtl number, since only such flows could conceivably be dynamically similar to those in the Earth.

4.2.1. The heated below (Bénard) problem. The linear stability boundary for heated below convection shown in fig. 22 suggests that a rather broad band of horizontal wavelengths may be realizable above the critical Rayleigh number. This expectation based on linear theory is not born out by

laboratory experiment. If one attempts to induce a two-dimensional form of convection (rolls) with wavelength much in excess of the critical wavelength at onset (Ra not too much in excess of Ra_c) the induced flow goes unstable via the so-called zig-zag instability [67], which transforms the induced flow into a stable two-dimensional flow with wavelength again on the order of the critical value λ_c . Similarly, if one attempts to induce two-dimensional convection with $\lambda \ll \lambda_c$, it is modified by the cross roll instability to a new set of rolls with increased wavelength which is again such that $\lambda \simeq \lambda_c$ [67]. These two instabilities are illustrated in fig. 23a, b) from [65]. For Rayleigh numbers $Ra \geq 2.26 \cdot 10^4$

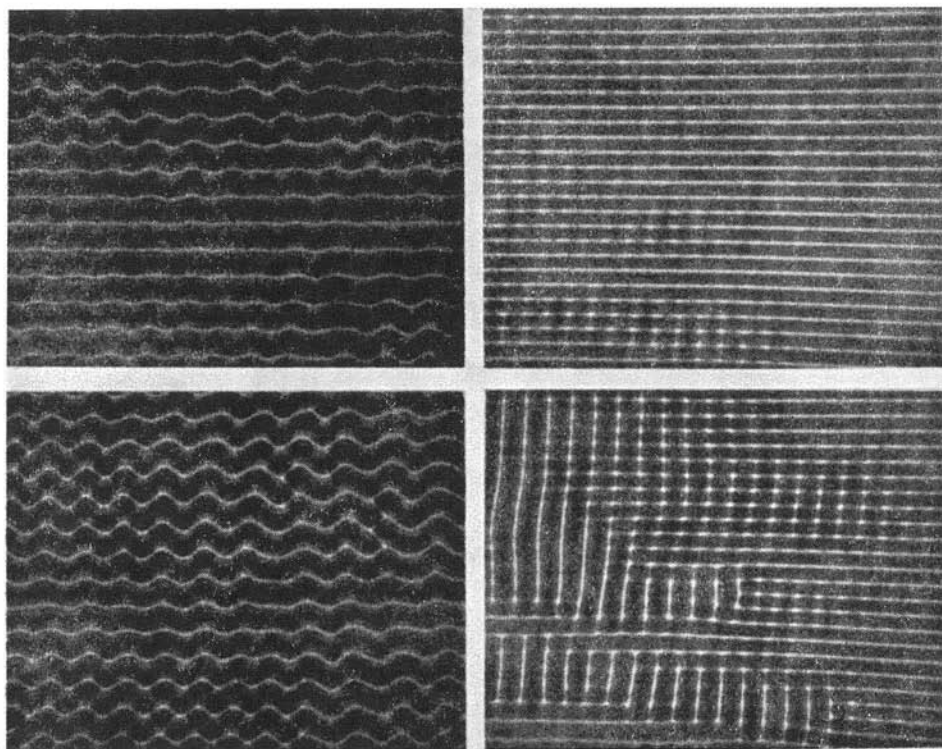


Fig. 23. — The two figures on the left illustrate the development of a zig-zag instability which transforms the roll planform into one with smaller wavelength, while the two figures on the right illustrate the cross roll instability which transforms the roll planform into one with larger wavelength (from ref. [65]).

two-dimensional convection rolls of any wavelength are not physically realized, but rather are replaced by a three-dimensional flow which BUSSE and WHITEHEAD [67] have called bimodal. The planform of the bimodal flow is illustrated in fig. 24. Under certain conditions this steady bimodal flow may itself become unstable to a time-dependent flow in which the planform consists of a spoke pattern [68]. An example of the transition from bimodal flow

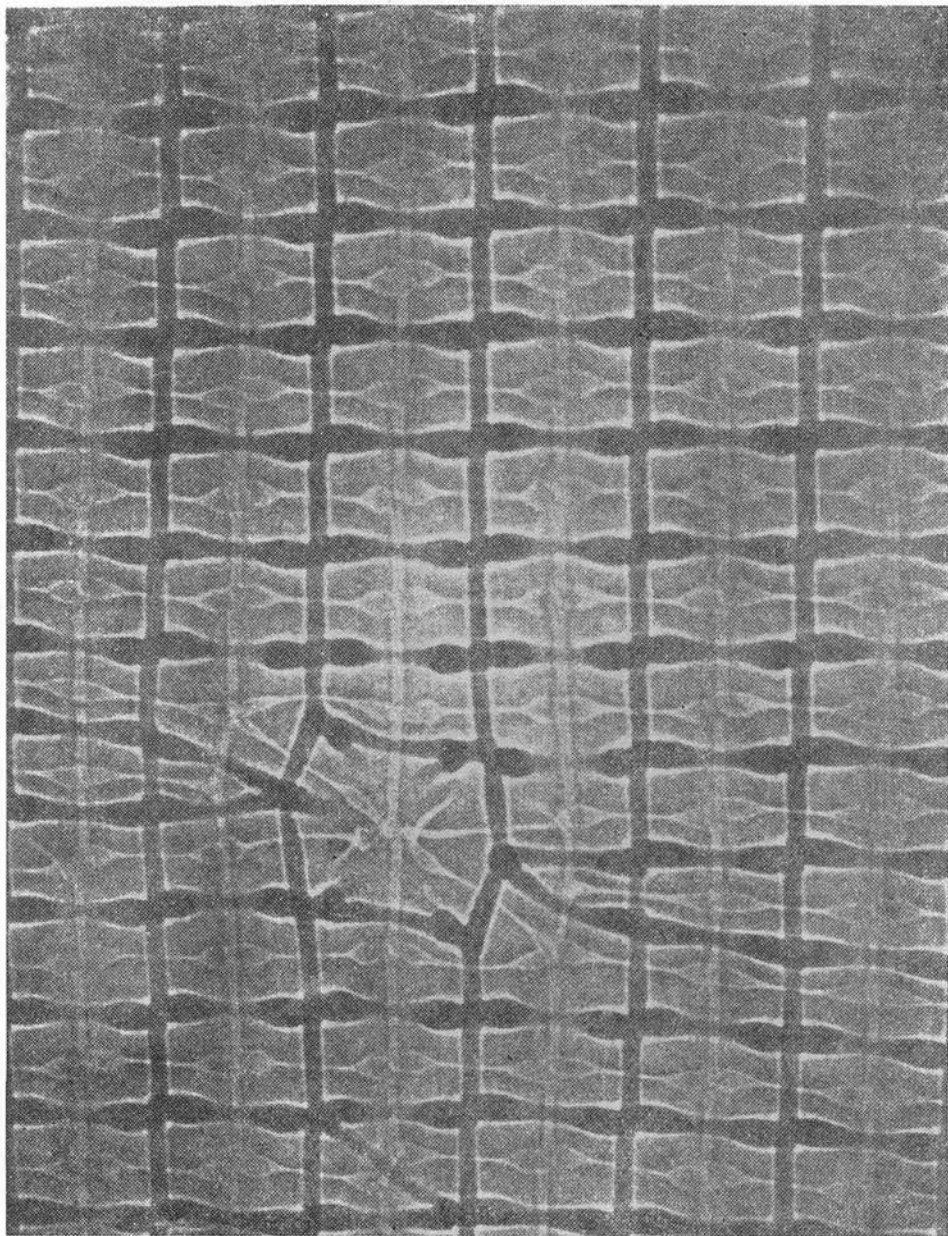


Fig. 24. - Planform of bimodal convection containing a flaw (from ref. [65]).

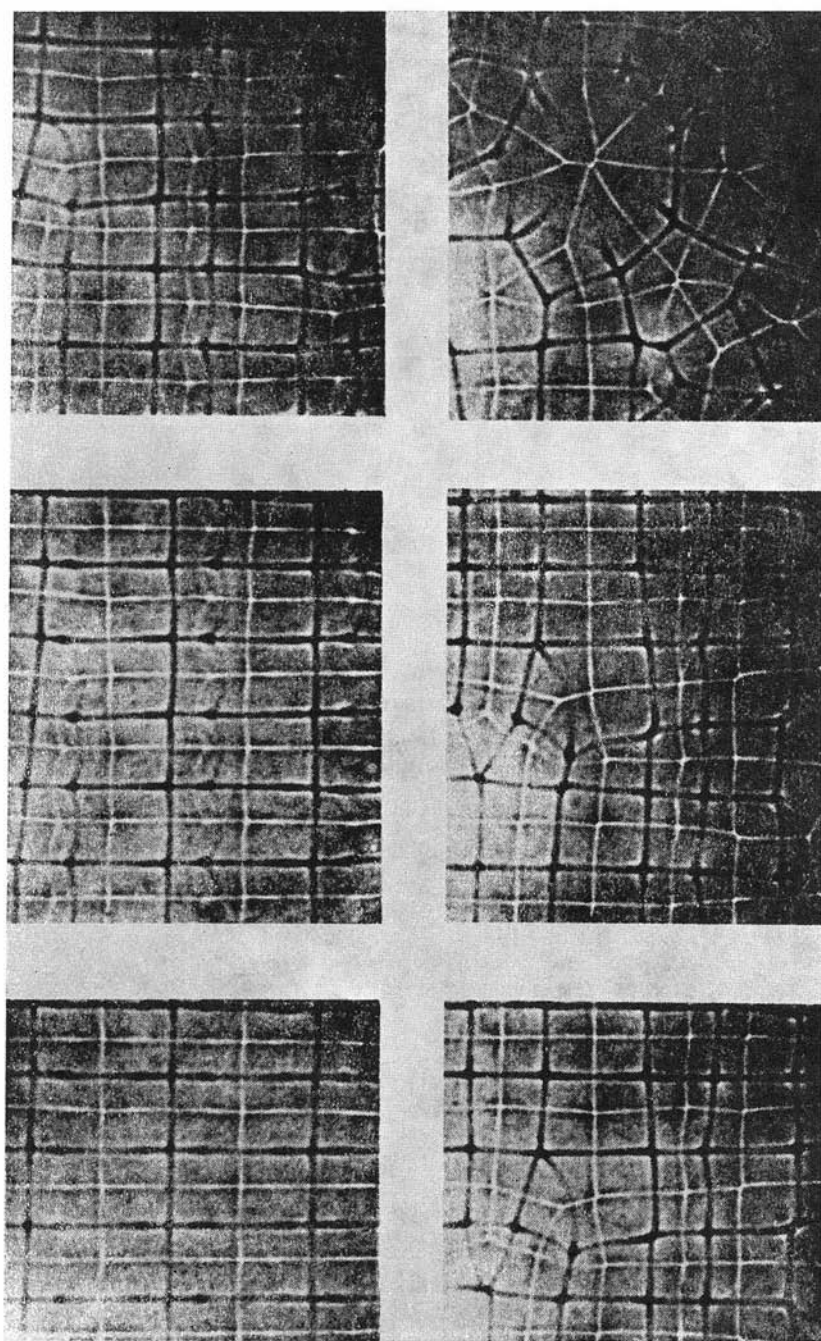


Fig. 25. - The transition from bimeron flow to spokes at $R_2 = 3.72 \cdot 10^5$ (from ref. [69]).

to spokes is shown in fig. 25 [69] for $Ra = 3.72 \cdot 10^5$. As we shall discuss further in the next section, these higher-order transitions in convection (*e.g.* rolls \rightarrow \rightarrow bimodal) which occur as the Rayleigh number is increased are due to secondary instabilities in the thermal boundary layers which form adjacent to the horizontal plates. Their occurrence or nonoccurrence is strongly dependent upon the mechanical boundary conditions (*e.g.* stress free or rigid), a point to which we shall also return.

Complementing the above-described observations of planform are quantitative measurements of the dependence of the convective heat flow upon the Rayleigh number. These measurements are usually expressed in terms of the Nusselt number which is a nondimensional measure of the total heat transport. For heated below, steady-state convection it is defined as

$$(25) \quad Nu = 1 + \frac{q}{K \Delta T/d},$$

where $K \Delta T/d$ is the heat flux which would be produced in the absence of convection and q is the convective heat flux. The heat transfer curve $Nu(Ra)$ shows distinct changes of slope at the critical point and for the value of Ra at which the transition from rolls to bimodal flow occurs. At high Rayleigh number, boundary layer theory may provide an explanation of the observed $Nu(Ra)$ relation. This will be discussed in subsect. 4.4 and will be used as the basis of a thermal-history model in sect. 6.

4.2.2. Convection with internal heating. This heating configuration has been less carefully studied than has the heated below case, the main reason being that the experiments are very much more difficult to perform. An example of an apparatus which has been used successfully for such measurements at the University of Toronto is shown in fig. 26 [70]. The fluid is heated by the Joule dissipation of an alternating current which is forced to flow through the fluid between the electrodes which form its upper and lower boundaries. The lower boundary actually consists of a wire mesh as the lower electrode, below which is a relatively thick glass plate which serves to enforce a no heat flux lower boundary condition by virtue of its low conductivity. For these experiments the flow was visualized not by the shadowgraph technique used above for the Bénard flows, but by streak photography. In this method the fluid is doped with neutrally buoyant particles (polystyrene beads) and the flow visualized through the light which they scatter during a time exposure of the film. In fig. 27*a*), *b*) we show, respectively, a sample planform and a vertical cross-section for a flow at moderately supercritical Rayleigh number. The planform is much less regular than those employed to illustrate the Bénard case, a consequence of the fact that no regular form was induced initially. It does, however, show a rather remarkable degree of two-dimen-

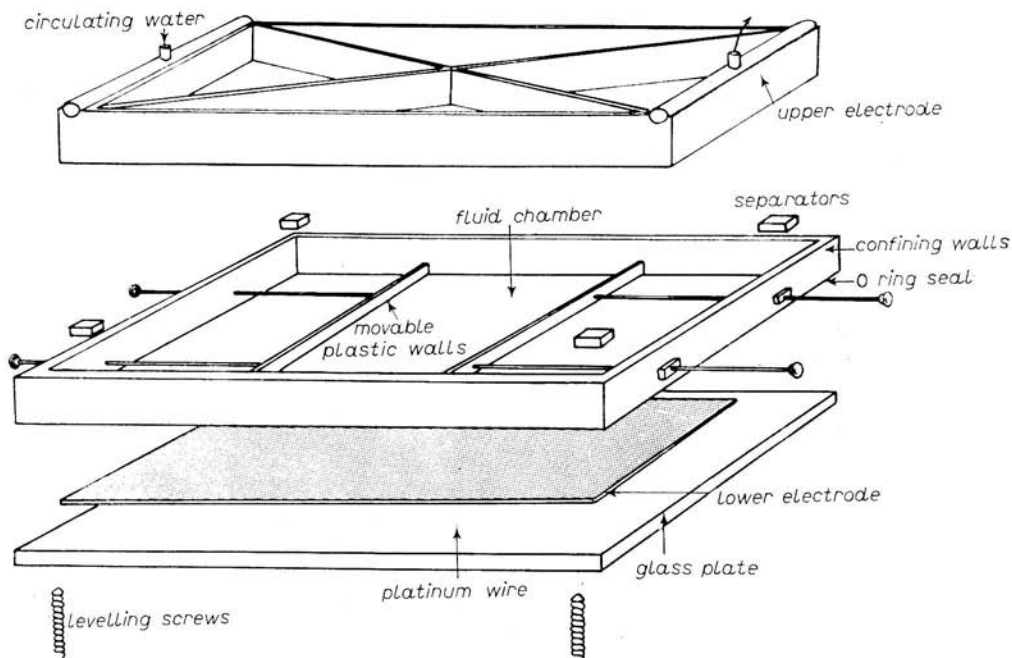


Fig. 26. — Schematic diagram of an apparatus for the study of convection with internal heat generation (from ref. [70]).

sionality nevertheless. The vertical cross-section in fig. 27b) illustrates the most striking feature of convective flows generated by internal heating. Notice that the stagnation point in the core of each cell is shifted dramatically away from the cell centre both horizontally and towards the upper corner. The circulation in these cells is such that the descending stream is in the narrowest region between two adjacent stagnation points. This asymmetry is due to the fact that there is only one thermal boundary layer in an internally heated layer and this is located adjacent to the cold upper surface which is held at constant temperature. We will return to this point below.

The nondimensional heat transfer curve for the internally heated configuration has also been measured in a preliminary fashion. This is shown in fig. 28, where we plot the Nusselt number for these flows as a function of the Rayleigh number. In the internally heated case the Nusselt number is defined as

$$(26) \quad \text{Nu} = \frac{\Delta T_{\text{cond}}}{\Delta T_{\text{conv}}},$$

which is just the ratio of the temperature drop across the layer which would obtain in the absence of convection to that which obtains in the presence of

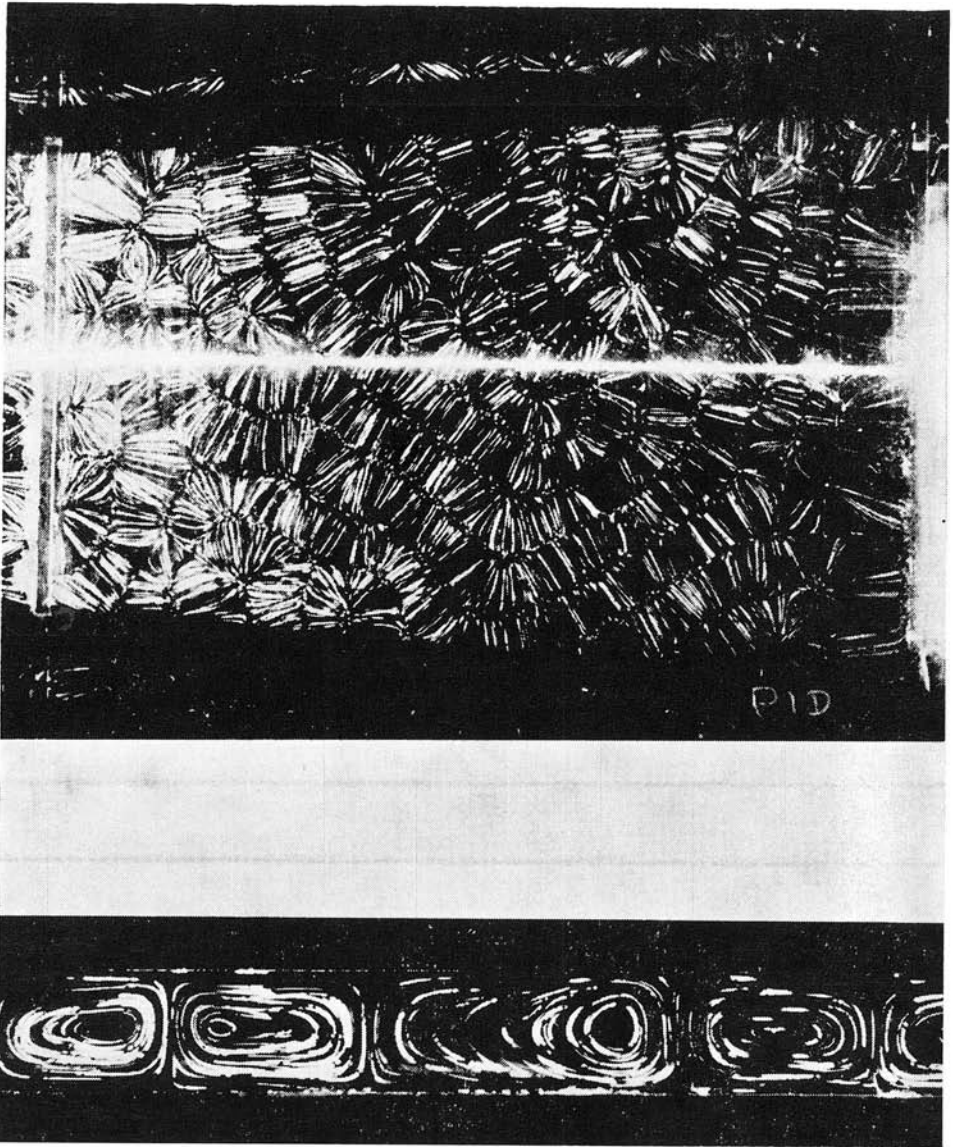


Fig. 27. — The upper streak photograph is a planform of convection with internal heat generation for $Ra \simeq 7.9 \cdot 10^3$. Note the strong tendency towards two-dimensional flow in much of the domain. The lower streak photograph is a vertical cross-section which illustrates very well the asymmetric streamline pattern in the cell core at higher Rayleigh number $Ra \simeq 5 \cdot 10^5$. Note that the stagnation point is shifted substantially away from the cell centre (from ref. [70]).

convection. Superimposed upon the experimental data, which are shown as heavy circles, are the results from a sequence of theoretical calculations (crosses) which will be discussed below. Theory and experiment clearly agree very

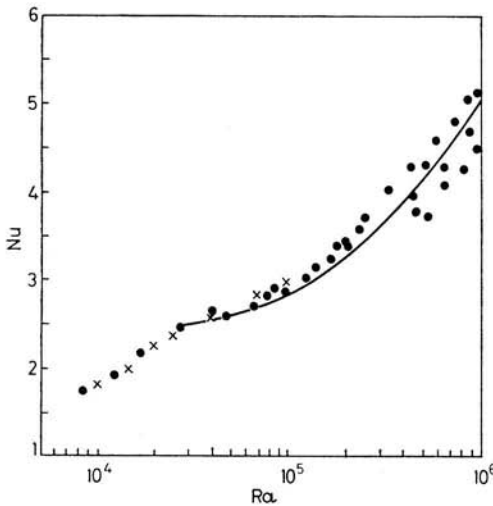


Fig. 28. — Nusselt number vs. Rayleigh number curve for the internally heated experiments. The solid circles are experimental points, while the crosses are results of numerical calculation for two-dimensional steady-state flows.

well insofar as the heat transfer characteristics of the circulation are concerned. The scatter in $Nu(Ra)$ at large Ra is a consequence of noise in the measuring system. This has been corrected in a new apparatus [71].

4.3. *Nonlinear solutions of the Oberbeck-Boussinesq equations.* — If we can obtain exact solutions to the nonlinear system of equations (19)-(22), we should be able to correctly predict the experimental results described above. These observations suggest that we should first seek two-dimensional steady-state solutions to the equations, at least for moderately supercritical Rayleigh numbers, since such solutions are the ones which are realized physically. Although such solutions may be constructed using finite-difference methods (*e.g.* [10]), probably the most elegant way to proceed is via the Galerkin formalism which has been employed so successfully by BUSSE and his co-workers [72, 73]. We will illustrate the method here by using it to solve the internally heated problem, which has apparently not been done previously.

We begin by expanding the temperature field in the form $T = T_c + \theta$, where T_c is the conduction solution and θ the deviation from it. When this is substituted into (19)-(21), and the conduction solutions subtracted, we obtain

the following p.d.e.'s in \mathbf{u} , θ , π , where π is the scaled pressure:

$$(27) \quad \nabla \cdot \mathbf{u} = 0,$$

$$(28) \quad \frac{\text{Ra}}{\text{Pr}} \frac{d\mathbf{u}}{dt} = -\nabla\pi + \theta\boldsymbol{\lambda} + \nabla^2\mathbf{u},$$

$$(29) \quad \frac{d\theta}{dt} = \frac{1}{\text{Ra}} \nabla^2\theta + (\boldsymbol{\lambda} \cdot \mathbf{u})(z + \frac{1}{2}).$$

The horizontal boundaries are on $z = \pm \frac{1}{2}$. By expressing the solenoidal velocity field \mathbf{u} in terms of poloidal and toroidal scalars as

$$(30) \quad \mathbf{u} = \boldsymbol{\delta}\varphi + \boldsymbol{\varepsilon}\psi,$$

the continuity equation (27) is satisfied exactly. The operators $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are defined such that

$$(31a) \quad \boldsymbol{\delta}\varphi = \nabla \times (\nabla \times \boldsymbol{\lambda}\varphi),$$

$$(31b) \quad \boldsymbol{\varepsilon}\psi = \nabla \times \boldsymbol{\lambda}\psi.$$

By operating on (28) with $\boldsymbol{\lambda} \cdot \nabla \times$ and $\boldsymbol{\lambda} \cdot \nabla \times (\nabla \times)$, eqs. (28) and (29) are equivalent to the following equations in the scalars φ , ψ , θ (following [74]):

$$(32) \quad \nabla^4 \Delta_2 \varphi - \Delta_2 \theta = \frac{\text{Ra}}{\text{Pr}} \left\{ \boldsymbol{\delta} \cdot (\boldsymbol{\delta}\varphi + \boldsymbol{\varepsilon}\psi) \cdot \nabla (\boldsymbol{\delta}\varphi + \boldsymbol{\varepsilon}\psi) + \frac{\partial}{\partial t} \nabla^2 \Delta_2 \varphi \right\},$$

$$(33) \quad \nabla^2 \Delta_2 \psi = \frac{\text{Ra}}{\text{Pr}} \left\{ \boldsymbol{\varepsilon} \cdot (\boldsymbol{\delta}\varphi + \boldsymbol{\varepsilon}\psi) \cdot \nabla (\boldsymbol{\delta}\varphi + \boldsymbol{\varepsilon}\psi) + \frac{\partial}{\partial t} \Delta_2 \psi \right\},$$

$$(34) \quad \frac{1}{\text{Ra}} \nabla^2 \theta - (z + \frac{1}{2}) \Delta_2 \varphi = (\boldsymbol{\delta}\varphi + \boldsymbol{\varepsilon}\psi) \cdot \nabla \theta + \frac{\partial \theta}{\partial t}.$$

In (32)-(34) $\Delta_2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ and we assume rigid boundaries in order to accord with the laboratory conditions. On the horizontal boundaries at $z = \pm \frac{1}{2}$ the boundary conditions are $\varphi = \partial_z \varphi = \psi = 0$ on $z = \pm \frac{1}{2}$ and $\theta = 0$ on $z = \frac{1}{2}$ and $\partial_z \theta = 0$ on $z = -\frac{1}{2}$, since the top surface is held at constant temperature, while the bottom surface has no heat flux through it.

Two-dimensional solutions to (32)-(34) have $\psi \equiv 0$ and are independent of x , say. Steady two-dimensional solutions satisfy the coupled equations (slightly modified from [74] for the present problem)

$$(35) \quad \partial_y (\nabla^4 \varphi - \theta) = \frac{\text{Ra}}{\text{Pr}} \{ \partial_{yz}^2 \varphi \partial_{yzz}^4 \varphi - \partial_{yy}^2 \varphi \partial_{yzz}^4 \varphi + \partial_{yz}^2 \varphi \partial_{yyy}^4 \varphi - \partial_{yy}^2 \varphi \partial_{yyz}^4 \varphi \},$$

$$(36) \quad \frac{1}{\text{Ra}} \nabla^2 \theta - (z + \frac{1}{2}) \partial_{yy}^2 \varphi = \partial_{yz}^2 \varphi \partial_y \varphi - \partial_{yy}^2 \varphi \partial_z \theta.$$

To obtain a Galerkin solution to (35) and (36) we expand the scalars θ and φ in terms of orthogonal functions as (e.g. [73])

$$(37a) \quad \theta = \sum_{\lambda, \nu} b_{\lambda\nu} \exp[i\lambda\alpha y] f_\nu(z) = \sum_{\lambda, \nu} b_{\lambda\nu} \theta_{\lambda\nu},$$

$$(37b) \quad \varphi = \sum_{\lambda, \nu} a_{\lambda\nu} \exp[i\lambda\alpha y] g_\nu(z) = \sum_{\lambda, \nu} a_{\lambda\nu} \varphi_{\lambda\nu},$$

where the $f_\nu(z)$ and $g_\nu(z)$ each satisfy the boundary conditions on θ and φ , respectively, and constitute a complete set of functions on $-\frac{1}{2} \leq z \leq \frac{1}{2}$. For the internally heated problem they are

$$(38) \quad f_\nu(z) = \sin(\nu - \frac{1}{2})\pi(z - \frac{1}{2}),$$

$$(39) \quad g_\nu(z) = \begin{cases} \frac{\sinh(\beta_{\frac{1}{2}\nu} z)}{\sinh(\frac{1}{2}\beta_{\frac{1}{2}\nu})} \frac{\sin(\beta_{\frac{1}{2}\nu} z)}{\sin(\frac{1}{2}\beta_{\frac{1}{2}\nu})}, & \nu \text{ even,} \\ \frac{\cosh(\lambda_{\frac{1}{2}(\nu+1)} z)}{\cosh(\frac{1}{2}\lambda_{\frac{1}{2}(\nu+1)})} \frac{\cos(\lambda_{\frac{1}{2}(\nu+1)} z)}{\cos(\frac{1}{2}\lambda_{\frac{1}{2}(\nu+1)})}, & \nu \text{ odd.} \end{cases}$$

The $g_\nu(z)$ are the so-called Chandrasekhar functions which are defined in [74], where the constants β_ν , λ_ν are listed.

To solve (35)-(36) using the Galerkin formalism, we simply substitute the expansions (37) and then multiply (35) and (36), respectively, by $\theta_{\kappa\mu}$ and $\varphi_{\kappa\mu}$ and integrate over the layer. This leads to the following set of simultaneous algebraic equations in the $a_{\lambda\nu}$, $b_{\lambda\nu}$:

$$(40a) \quad I_{\kappa\mu\lambda\nu}^{(1)} a_{\lambda\nu} + I_{\kappa\mu\lambda\nu}^{(2)} b_{\lambda\nu} + \frac{\text{Ra}}{\text{Pr}} I_{\kappa\mu\lambda\nu\varrho\pi}^{(3)} a_{\lambda\nu} a_{\varrho\pi} = 0,$$

$$(40b) \quad \frac{1}{\text{Ra}} I_{\kappa\mu\lambda\nu}^{(4)} b_{\lambda\nu} + I_{\kappa\mu\lambda\nu}^{(5)} a_{\lambda\nu} + I_{\kappa\mu\lambda\nu\varrho\pi}^{(6)} a_{\lambda\nu} b_{\varrho\pi} = 0,$$

where the I are interaction matrices the elements of which are all constants obtained from the layer integrals. To solve the set of algebraic equations (40) we use a triangular truncation such that $|\kappa| + \mu \leq N$ and a standard subroutine for the solution of simultaneous algebraic equations.

Examples of solutions which we have obtained (SHARPE and PELTIER, in preparation) using these methods, for both the heated below and the heated internally problems, are shown in fig. 29, where the Rayleigh number in each case is $\sim 10\text{Ra}_c$ and $\text{Pr} = \infty$. Each solution is represented by contour maps for the stream function and the isotherms, and also by a graph of the mean temperature field as a function of height through the layer. These solutions illustrate the fundamental differences between the heated below and internally problems quite clearly. In the former case two boundary layers form, one adjacent to each horizontal surface, while in the latter case only a single thermal boundary layer exists adjacent to the top constant-temperature boundary.

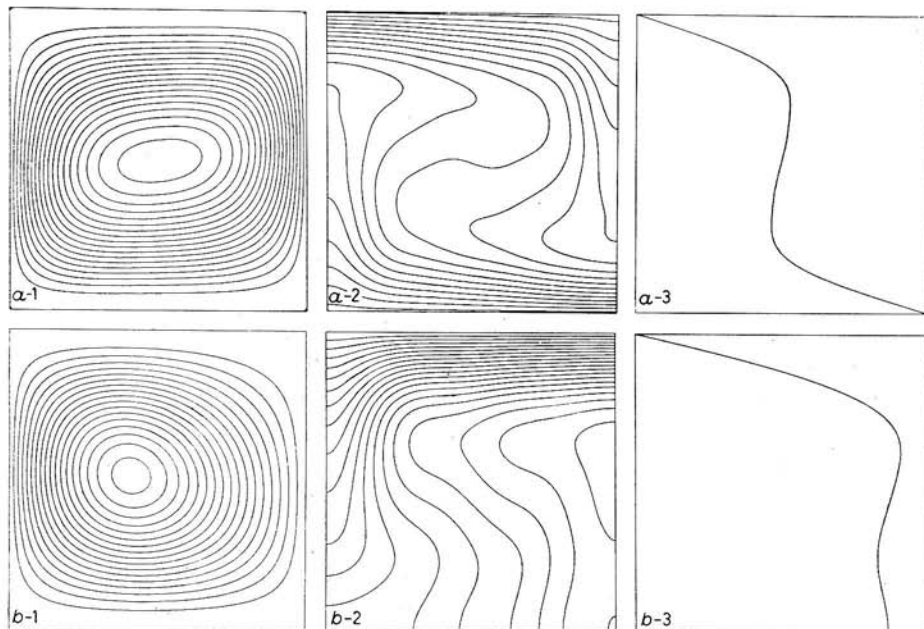


Fig. 29. — Results of the Galerkin calculation for both heated below and heated internally flows at $Ra \simeq 10 Ra_c$. Sequence *a* is for the heated below case and *b* for heated internally. In each sequence plate 1 is the streamline pattern, 2 the pattern of isotherms and 3 the horizontally averaged temperature field as a function of height. Comparing *a*-3 and *b*-3 it is clear that no thermal boundary layer forms adjacent to the lower adiabatic boundary in the internally heated case.

Also evident in the heated internally solution is the shift of the stagnation point in the streamfield away from the centre of the cell which we saw in the experimental cross-section of fig. 27*b*). We can obtain a quantitative comparison of these solutions for the internally heated flow with the laboratory observations by calculating the Nusselt number Nu as a function of the Rayleigh number. In terms of the Galerkin coefficients the Nusselt number (26) is given by

$$(41) \quad Nu = \frac{1}{1 + 2 \sum_{\nu} b_{0\nu} \sin(\frac{1}{2} - \nu)}$$

The comparison is shown in fig. 28 and it is clear that, for sufficiently small Ra at least, the two-dimensional steady-state solutions accurately predict the observed heat transfer.

The success of the Oberbeck-Boussinesq theory in reconciling laboratory observations has now been extended far beyond the mere reproduction of observed heat transfer data. At least insofar as the heated below problem is concerned, all of the basic symmetry transitions involving the stability of two-

dimensional solutions have been carefully enumerated. A recent and comprehensive review of these results has been given by BUSSE [65]. The secondary instabilities which mark these higher-order transitions involve the thermal stability of the thermal boundary layers which form on the horizontal surfaces as the Rayleigh number is increased. For sufficiently high Rayleigh number we may invoke boundary layer theory to deduce the heat transport characteristics of convection. This theory is in fact very useful in understanding mantle convection and we review it in the next subsection.

4.4. *Boundary layer theory for convection at high Rayleigh and Prandtl number: a simple model of convection in the planetary mantle.* — As Ra increases, the thermal boundary layers illustrated in fig. 29 continue to become thinner and for very high Ra it becomes prohibitively expensive computationally to construct complete numerical solutions. However, in this regime the high-Prandtl-number flows are amenable to boundary layer analysis from which the most important characteristics of the circulation may be determined. This analysis was first applied in the context of mantle convection by TURCOTTE and OXBURGH [75]. It has recently been exploited to some advantage in the construction of thermal-history models by SHARPE and PELTIER [47], who also demonstrated its consistency with the main observational data when whole-mantle convection was assumed. Here we shall briefly review the theory and compare its predictions to the geophysical observations discussed in sect. 2. Following [47] we will focus on the idea of whole-mantle convection as opposed to the Turcotte-Oxburgh view which has usually been restricted to the upper-mantle form.

In the thin thermal boundary layers of thickness δ there is a steady-state balance between the horizontal advection of heat and its vertical diffusion *i.e.*, if ΔT is the temperature drop over the length scale δ and if u is the horizontal velocity and κ the diffusivity, then

$$(42a) \quad u \frac{\Delta T}{d} \sim \kappa \frac{\Delta T}{\delta^2},$$

where d is the layer depth and the horizontal boundaries are assumed stress free. In the vertical plumes, on the other hand, the buoyant generation of vorticity ω is balanced by viscous diffusion (from the curl of (28) in the limit $Pr \rightarrow \infty$). In dimensional form we, therefore, have

$$(42b) \quad g \frac{\alpha \Delta T}{\delta} \sim \nu \nabla^2 \omega \sim \frac{\nu \omega}{\delta^2}.$$

But, since $u \sim w \sim \omega d$ away from the plumes and boundary layers in the cell

interior, it, therefore, follows by eliminating u from (42a), (42b) that

$$(43a) \quad \delta = a_1(\Delta) d (\text{Ra}_c / \text{Ra})^{\frac{1}{3}},$$

$$(43b) \quad u = a_2(\Delta) \text{Ra}^{\frac{2}{3}} z / d,$$

$$(43c) \quad w = a_3(\Delta) \text{Ra}^{\frac{2}{3}} z / d,$$

$$(43d) \quad q = a_4(\Delta) (\text{Ra} / \text{Ra}_c)^{\frac{1}{3}} z \Delta T / d.$$

The constants $a_i(\Delta)$ are functions of the aspect ratio of convection Δ and may be determined by a detailed matching of the boundary layer solutions for the plumes and horizontal thermal boundary layers to the isothermal Stokes flow which obtains in the core of the cell. This may be a relatively simple exercise [75] or a complex one [76] depending upon the degree of rigour imposed in the matching process. The analysis in [75] in fact turns out to be quite accurate, although it has been improved in [77] and for an aspect ratio slightly larger than that at the onset of instability yields $a_1(\Delta_M) \simeq 0.50$, $a_2(\Delta_M) \simeq 0.143$, $a_3(\Delta_M) \simeq 0.251$ and $a_4(\Delta_M) \simeq 1$. The result (43d) is one which may be written in the form $\text{Nu} = (\text{Ra} / \text{Ra}_c)^{\frac{1}{3}}$ and compared to laboratory experiment and numerical computation. Although the fit to the finite-difference computations of Moore and Weiss [78] is very good, the comparison with laboratory data is not. The discrepancy between the boundary layer theory and experimental data is due to the fact that the experiments are of necessity performed with no-slip boundary conditions.

The consistency of the boundary layer theory with the idea of whole-mantle convection was first pointed out in [47] and the discussion is worth reiterating here. All of the quantities (43a), (43b) and (43d) are geophysical observables. u is just a typical plate speed, q is the observed mean mantle contribution to the surface heat flux [79], and δ is the thickness of the thermal boundary layer at the surface. It is an observable because of the strong temperature dependence of viscosity due to the thermally activated nature of the creep process. The thermal boundary layer at the surface of the planet is, therefore, a region through which there is an enormous variation of viscosity, because it is in just this region that the temperature gradient is strong. The thermal boundary layer of the mantle convective circulation and the lithosphere are, therefore, spatially coincident. In sect. 3 we showed that the lithosphere was about 100 km thick and this is, therefore, also the boundary layer thickness. This interpretation is radically different from most interpretations of mantle convection in the literature, *e.g.* [11, 12, 14, 15]. It depends upon the assumption that the lithosphere, once broken, rides passive and platelike on the uniform-viscosity mantle beneath it and that, in spite of the large viscosity contrast through it, it is still perfectly coupled to the mantle. This implies the absence of any substantial velocity shear beneath the plates.

Since each of the predictions (43) is in terms of only two unknowns, the Rayleigh number Ra and the depth of convection d , we may choose any two of the expressions for the three observables and eliminate one unknown in terms of the other. The unique solution to (43a) and (43b) with $\delta \simeq 100$ km and $u \simeq 7$ cm/y is $Ra \simeq 3.4 \cdot 10^6$ and $d \simeq 3 \cdot 10^8$ cm ($Ra_c = 10^3$), *i.e.* whole-mantle convection! We may tentatively conclude from this analysis that the boundary layer theory provides a remarkably accurate explanation of the main observables of mantle convection.

Another way of looking at the expressions (43) is to eliminate ΔT between any two of them to get an explicit expression for the depth of convection d in terms of observed quantities. There are in fact two ways in which this can be done. If we combine (43b) and (43d), we get

$$(44a) \quad d = \frac{a_4^{\frac{1}{2}}}{a_2} \left(\frac{C_p v}{Ra_c^{\frac{1}{2}} g \alpha g} \right)^{\frac{1}{2}} u,$$

while, if we combine (43a) and (43b), we get

$$(44b) \quad d = \frac{u \delta^2 / \kappa}{a_2 \cdot a_1^2 \cdot Ra_c^{\frac{2}{3}}}.$$

If in (44a) we substitute $u = 4$ cm·y⁻¹ as the average plate speed, $Ra_c = 10^3$, $g = 10^3$ cm·s⁻², $q = 0.5$ erg·cm⁻² s⁻¹ [79], $\nu = 10^{22}$ g·cm⁻¹·s⁻¹ (poise), $\alpha = 3 \cdot 10^{-5}$ K⁻¹, $C_p = 1.2 \cdot 10^7$ erg·g⁻¹·K⁻¹, then we obtain $d \simeq 4.6 \cdot 10^3$ km. ELSASSER *et al.* [80] attribute the result (44a) to GOLITSYN, but, as we have shown above, it follows trivially from the usual boundary layer theory which was first employed to demonstrate compatibility of observations with whole-mantle convection in [47]. Although this prediction of the depth scale is quite suggestive of whole-mantle convection, the prediction of the equally relevant (44b) which has apparently not been derived previously is $d \simeq 3600$ km, where we have again used the moderate $u = 4$ cm·y⁻¹ and $\delta = 100$ km.

On the basis of the above arguments we may be tempted to conclude that whole-mantle convection is not only compatible with the observations, but is more or less demanded by them. This conclusion is dependent upon at least four fundamental assumptions: i) that the two-dimensional steady-state boundary layer scaling for plane-layer convection is not fundamentally altered for three-dimensional convection in a spherical shell, ii) that the viscosity of the mantle is a constant everywhere or that the system behaves as if it were constant, iii) that the enormous viscosities which obtain in the lithosphere are entirely «passive» in that they produce no essential effect upon the circulation (this could conceivably be due to nonhydrodynamic effects such as melting beneath the ridge) and that there is no substantial shear in the horizontal velocity between the lithosphere and the asthenosphere (*i.e.* no decoupling),

iv) the mantle convective circulation is driven entirely from below by the efflux of heat from the core. It is unlikely that all of these assumptions are correct (iv) is certainly not correct), but the internal consistency of the observations with convection scaling is nevertheless well demonstrated by the boundary layer analysis. Perhaps the most interesting point to be made on the basis of the above exercise is that, if the viscosity of the mantle were orders of magnitude different from the value deduced from postglacial rebound, the convection hypothesis could not be made compatible with the observations, no matter what the choice for d !

Given the excellent fit of the boundary layer model to the observations from a global perspective, it is tempting to enquire as to whether it is also able to fit any detailed local data. This is in fact the case. In the boundary layer approximation the steady-state form of the energy equation becomes

$$(45) \quad u \frac{\partial T}{\partial x} = \kappa \frac{\partial^2 T}{\partial z^2},$$

where we have assumed $\partial^2/\partial x^2 \ll \partial^2/\partial z^2$. Under the similarity transformation $\eta' = z/x^{\frac{1}{2}}$ (45) is reduced to an o.d.e. in η' which may be nondimensionalized by writing $\eta' = \eta \cdot 2(\kappa/u)^{\frac{1}{2}}$ to give the equation

$$(46) \quad \frac{d^2 T}{d\eta^2} + 2\eta \frac{dT}{d\eta} = 0,$$

so that subject to boundary conditions $T = T_s$ on $x = 0$ and $T \rightarrow T_m$ as $z \rightarrow \infty$ the dimensional solution is

$$(47) \quad T = T_s + (T_m - T_s) \operatorname{erf} \left[\frac{z}{2} \left(\frac{u}{\kappa x} \right)^{\frac{1}{2}} \right].$$

This is the boundary layer solution for the temperature field in the boundary which cools horizontally as it moves away from the ridge crest. T_s is the temperature at the ridge crest $x = 0$ and T_m is the temperature of the mantle material in the core of the convection cell beneath the thermal boundary layer. The surface heat flux associated with the temperature distribution (47) is

$$(48) \quad q_s = K \left. \frac{\partial T}{\partial z} \right|_{z=0} = K(T_m - T_s) \left(\frac{u}{\pi \kappa x} \right)^{\frac{1}{2}}.$$

Since the age of the lithosphere is just x/u , (48) predicts that the heat flow will decrease inversely as the square root of the age of the ocean floor and this is the dependence noted in the observations discussed in sect. 2 and shown in fig. 8. From the solution (47) for the temperature field we may also calculate the topography of the sea floor under the assumption that it is isostatically

compensated and that the density of the boundary layer material is a function of temperature only. If one assumes compensation, the depth of the sea floor beneath the ocean surface will be [81]

$$(49) \quad D = D_{\text{ref}} - \frac{\alpha}{1 - \rho_w/\rho} \int_0^{\infty} dz (T - T_{\text{ref}}),$$

where D_{ref} is the reference depth at the ridge crest, say, and $T_{\text{ref}} = T_s$ is the temperature there. If we substitute (47) into (49), the result is

$$(50) \quad D = D_{\text{ref}} - \frac{2\alpha(T_m - T_s)}{1 - \rho_w/\rho} \left(\frac{\kappa}{\pi}\right)^{\frac{1}{2}} \left[\left(\frac{x_{\text{ref}}}{u}\right)^{\frac{1}{2}} - \left(\frac{x}{u}\right)^{\frac{1}{2}} \right],$$

so that the depth of the sea floor should increase linearly as the square root of the age of the oceanic lithosphere and this was demonstrated to be the case observationally, as seen in fig. 7. The boundary layer model which assumes complete coupling between the lithosphere and the mantle is, therefore, in very close accord with the observations. Much has been made [14, 36] of the deviation from the square root of the age dependence which is observed in the oldest oceanic lithosphere for the bathymetry field. It has been argued that this might be explained by the existence beneath the old ocean floor of a second scale of convection derivative of a secondary instability of the upper boundary layer. Although this scale of motion might exist, it cannot but represent a relatively innocuous perturbation of the large-scale flow associated with plate creation and destruction, since the temperature dependence of viscosity forces the instability to reside in a region in which the available potential energy is minimal. In terms of the whole-mantle convection model, the observed flattening of the sea floor may be attributed to partial internal heating of the flow due to distributed radioactivity in the mantle [82].

If whole-mantle convection is indeed occurring in the Earth, as the boundary layer theory may be taken to suggest, then there are several physical effects which we have so far omitted which may become important because of the depth of the system. We investigate these and other hydrodynamic complexities in the next section.

5. - Convection in the Earth and its hydrodynamic complexities.

If we are obliged to consider convection in a deep layer like the mantle, then certain non-Boussinesq effects neglected in the preceding discussion may become important, even if we continue to assume that the transport coefficients are constant. These effects are associated with the dissipation number

τ which appears in the energy equation (13), where it measures the importance of both compression work and viscous dissipation. The possible influence of these effects on mantle convection was demonstrated in [64]. This section is concerned with these and other complexities which are not encompassed within the Oberbeck-Boussinesq theory, including the influence of phase transitions, of non-Newtonian rheology and of temperature-dependent viscosity.

5.1. Anelastic convection: the adiabatic gradient and viscous dissipation. — An approximation which leads to considerable simplification of the complete set of hydrodynamic equations (11)-(14) is the so-called anelastic approximation which has been widely used in meteorological applications for many years. Examples of its use in this context may be found in [83, 84]. This approximation is based upon the neglect of $\partial\rho/\partial t$ in the continuity equation (12), a device which filters compression waves completely from the equations of motion and, therefore, requires that the Mach number M of the flow be such that $M^2 \ll 1$. This is certainly a valid approximation for the mantle circulation! If $\partial\rho/\partial t = 0$, then $\rho = \bar{\rho}(z)$ if the density field is initially in hydrostatic equilibrium (*i.e.* a function of height only). A second approximation which is useful in studies of deep-mantle convection pertains to the compression work term in the energy equation which contains the factor dp/dt . A simple scale analysis demonstrates that $dp/dt \simeq -\bar{\rho}w$, where w is the vertical component of velocity and we have used the fact that $dp/dz \simeq -\bar{\rho}g$. Subject to these two very good approximations the field equations (11)-(14) are reduced to

$$(51) \quad \frac{\text{Ra}}{\text{Pr}} \bar{\rho} \frac{du_i}{dt} = \frac{1}{\delta} (-\partial_i p + \rho \lambda_i) + \partial_j \left[2f \left(e_{ij} - \frac{1}{3} \Delta \delta_{ij} \right) \right],$$

$$(52) \quad \partial_j (\bar{\rho} u_j) = 0,$$

$$(53) \quad \frac{dT}{dt} + \tau T w = \frac{1}{\bar{\rho} \text{Ra}} [\partial_j (k \partial_j T) + h] + \frac{2\tau f}{\bar{\rho}} \left(e_{ij} e_{ij} - \frac{1}{3} \Delta^2 \right),$$

$$(54) \quad \rho = \bar{\rho} [1 - \delta(T - \bar{T}) + \Gamma(p - \bar{p})],$$

where $\rho = \bar{\rho}(z)$ in the state in which $T = \bar{T}$ and $p = \bar{p}$ and for the description of deep convection it is natural to take this state to be adiabatic, since, only if the temperature gradient is superadiabatic, can convection occur. In the adiabatic reference state the hydrostatic density field (nondimensional) is

$$(55) \quad \bar{\rho}(z) = \exp[(1 - \zeta) \tau/\gamma],$$

where $\gamma = \alpha/\rho c_p \Gamma$ is Grüneisen's parameter which has been assumed constant in deriving (55) and where $\tau = g\alpha d/C_p$ is the dissipation number [64] which we showed in subsect. 4.1 to be equal to the ratio of the layer depth and the adiabatic temperature scale height H_T .

The first and, as we shall show, most important effect of τ may be revealed by linearizing (51)-(54) about a steady-state conduction solution. Expanding the first-order fields in normal modes, as previously, and invoking the exchange-of-stabilities principle, the stability equation which replaces (24) is

$$(56) \quad \mathcal{L}^6 w = -k_H^2 \text{Ra} \left(\frac{dT_c}{d\zeta} + \tau T_c \right) \rho_c w,$$

where \mathcal{L}^6 is a complicated ordinary differential operator the explicit form for which is given in [64]. If the density variations are neglected, *i.e.* $\rho_c \rightarrow 1$, then $\mathcal{L}^6 \rightarrow (d^2/d\zeta^2 + k_H^2)^3$ as before. The effect of τ is then seen to be in the modulation of Ra by the factor $dT_c/d\zeta + \tau T_c$ rather than the factor $dT_c/d\zeta$ which obtains in the Boussinesq limit. Since τT_c is the nondimensional adiabatic gradient, it is clear from (56) that it is the difference between the actual and adiabatic gradients which drives convection in a deep layer and not the temperature gradient itself.

Mathematically the effect of τ is to introduce « turning points » into the stability equation [64] because the sign of $dT_c/d\zeta + \tau T_c$ may change from + to - through the layer. Where the sign is + the temperature gradient is superadiabatic and that fraction of the layer is potentially unstable, and where - it is subadiabatic and stable. In liquidlike fluids such as planetary interiors, the effect of the adiabatic gradient for constant α is to stabilize the deepest layers relative to the near-surface region [64]. This is because the adiabatic gradient increases with temperature and the temperature increases with depth. In gaslike fluids, such as stars, on the other hand, this relative stabilization of the hot interior does not occur, a consequence of the fact that the thermal-expansion coefficient $\alpha = T^{-1}$, so that the adiabatic gradient is independent of temperature.

If the dissipation number is sufficiently large that the lower part of the layer is stabilized, then the style of convection which obtains at onset may be « penetrative » [64] and the flow consists of a stack of cells in the vertical, only the topmost of which is driven by instability, while the rest are driven viscously from above. Whether this is important in the Earth depends upon the magnitude of $\tau = g\alpha d/e_p$, which is an increasing function of the depth of convection d . The largest effects are, therefore, expected for whole-mantle convection, which, as we have shown using boundary layer arguments, may be suggested by the data. Substituting $g = 10^3 \text{ cm s}^{-2}$, $\alpha = 3 \cdot 10^{-5} \text{ }^\circ\text{C}^{-1}$, $d = 3 \cdot 10^8 \text{ cm}$, $e_p = 1.2 \cdot 10^7 \text{ erg g}^{-1} \text{ }^\circ\text{C}^{-1}$ into $\tau = g\alpha d/e_p$, we get $\tau \simeq 0.75$ as an upper limit, since α may be somewhat lower in the deep mantle. Another independent method of deducing τ is to make use of the seismically observed density variation through the mantle in conjunction with eq. (55). This approach leads to a similar value of $\tau \simeq 0.6$, as pointed out by JARVIS and MCKENZIE [85]. Given this value of τ for whole-mantle convection, it is not

clear what the effects of τ might be beyond moderating the intensity of convection through the adiabatic gradient. It does not follow from (53), for example, that for $\tau = O(1)$ viscous dissipation is of the same order of importance as thermal advection. To discover the full impact of finite τ , we must actually solve the nonlinear problem. Before discussing such solutions there are several interesting observations which can be made on the basis of the volume integral of the energy equation (53).

For steady convection ($\partial T/\partial t \equiv 0$) this volume integral gives the simple dimensional result

$$(57) \quad \Phi = \tau \int_V \bar{\rho} T w \, dv,$$

where Φ is the total viscous dissipation. This is simply a statement of the fact that for steady convection the shear heating is entirely offset by the compression work (work done against the adiabatic gradient). Under the assumption that the convection is vigorous, this expression may be manipulated as in [86] to give an expression for the ratio of Φ to the total heat flux F across the top surface of the convecting region as

$$(58) \quad \frac{\Phi}{F} = \tau \left(1 - \frac{\mu}{2} \right) = E,$$

where μ is the ratio of the internal heat generation to the total heat flux. The ratio $\Phi/F = E$ may be interpreted as the efficiency of conversion of heat into work within the convecting region following MALKUS [87]. As pointed out by HEWITT *et al.*, (58) clearly implies that the importance of viscous dissipation is completely independent of the Rayleigh number and this was already clear from the scaling of Peltier [64]. In the Boussinesq limit $\tau \rightarrow 0$ and viscous dissipation can be neglected, but for $\tau = O(1)$ both the effect of shear heating and of the adiabatic gradient *may* become simultaneously important. To determine how important and for what values of τ we must solve the nonlinear problem in (Ra, τ) parameter space. The anelastic equations (51)-(54) have been rederived in [85] and solved using finite-difference methods in two spatial dimensions for moderate values of Ra and τ . We turn to a brief discussion of these solutions here.

In fig. 30 we show a sequence of mean-temperature profiles for a large number of heated below numerical calculations which were integrated to steady or statistically steady state for various values of Ra and $\tau = d/H_x$ [85]. For a fixed value of τ , Ra increases with the mean temperature. As is clear by inspection, as τ increases the definition of the lower boundary layer is erased, particularly at low Rayleigh number. This is a consequence of the fact, stated above, that the stabilizing effect of τ is strongest near the bottom of the layer,

where the temperature is highest. Near the bottom boundary, thermal conduction, therefore, plays a larger role. In addition, the mean temperature field in the interior of the layer is no longer isothermal as was the case for the Boussinesq flow shown in fig. 29. It is rather adiabatic as may be established by employing the second law in the form

$$(59) \quad \frac{dS}{dz} = \frac{c_p}{T} \frac{dT}{dz} - \frac{\alpha}{\bar{\rho}} \frac{dp}{dz}.$$

Since the pressure field is hydrostatic, then to a good approximation $dp/dz = -\bar{\rho}g$ and we may integrate (59) to obtain

$$(60) \quad S(z) - S(\bar{d}) = \alpha g(z - \bar{d}) + c_p \ln(T/T_0),$$

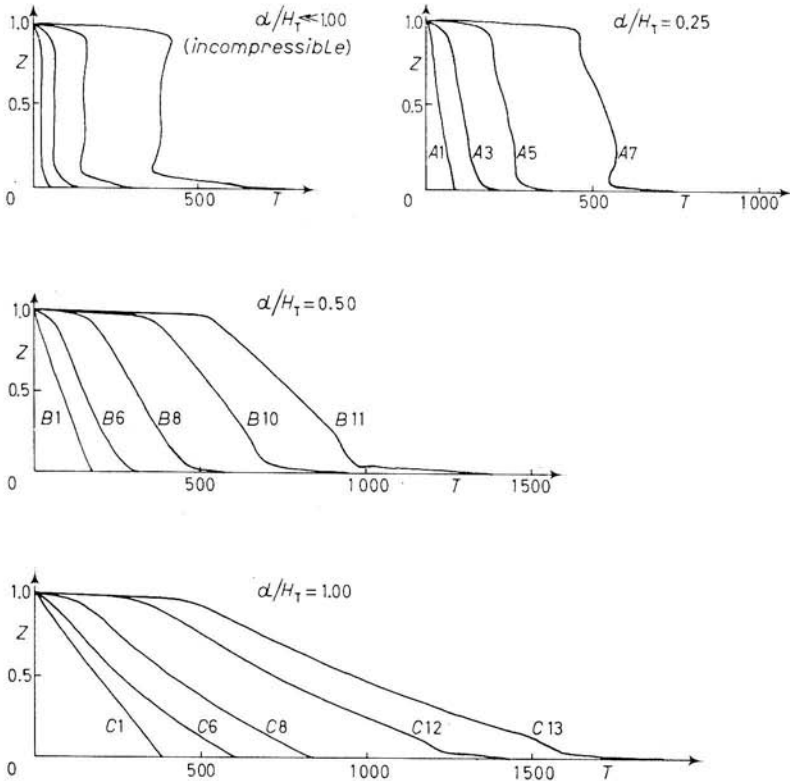


Fig. 30. — Mean-temperature profiles for a sequence of numerical simulations of heated below convection in deep layers with finite $d/H_T = \tau$. The τ value is marked beside each sequence and within each sequence the Rayleigh number increases with the mean temperature. Note the adiabatic cores which develop in the cell interiors and the increased effect of the adiabatic gradient with increasing depth (from ref. [85]).

where T_0 is the absolute temperature on $z = d$, the top surface, and is an important additional independent variable in deep-layer calculations, as was pointed out in [64] and subsequently verified in [85]. If the temperature field were adiabatic, then T would have the form (17) and S would be constant. We can use (60) to transform the temperature profiles shown in fig. 30 into profiles of S and the result is shown in fig. 31 for a sequence of experiments at

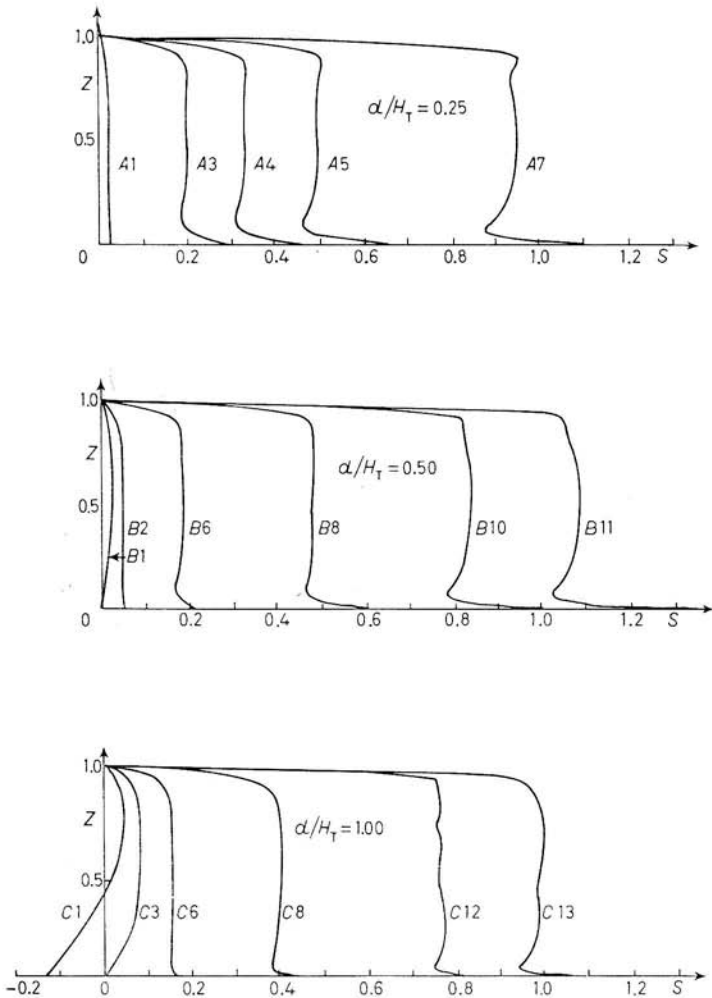


Fig. 31. - Entropy profiles for the heated below flows showing that the cell interiors are indeed adiabatic (from ref. [85]).

various values of τ and Ra [85]. Clearly the core region of the layer is adiabatic and the entropy profiles for the deep layer look quite like the temperature profiles for the Boussinesq flows. The most obvious effect of τ is thus to make the lapse rate adiabatic in the interior, an effect which is well known.

Inspection of the mean profile $B1$ in fig. 31 shows the characteristic asymmetry between the bottom and the top of the layer, which is produced by the fact that the adiabatic gradient depends linearly upon temperature. For this model, which has $Ra \simeq 2.8 \cdot 10^5$, the bottom of the layer is completely stabilized and inspection of the two-dimensional contour representations of the field variables in fig. 32 shows (see the streamfield ψ) that the motion at the bottom is driven viscously from above, as was demonstrated by the linear analysis in [64]. As the Rayleigh number increases, $Ra = 4.12 \cdot 10^5$, $1.65 \cdot 10^5$, $6.59 \cdot 10^5$, $1.65 \cdot 10^7$, $2.64 \cdot 10^7$ for $B2$, $B6$, $B8$, $B10$, $B11$, respectively, this stable region predicted by linear theory is eliminated and a well-developed lower thermal boundary layer is observed. With finite τ , however, the temperature drop across the bottom boundary layer is always less than the temperature drop across the top boundary layer. Also evident in fig. 32, where contours of constant viscous

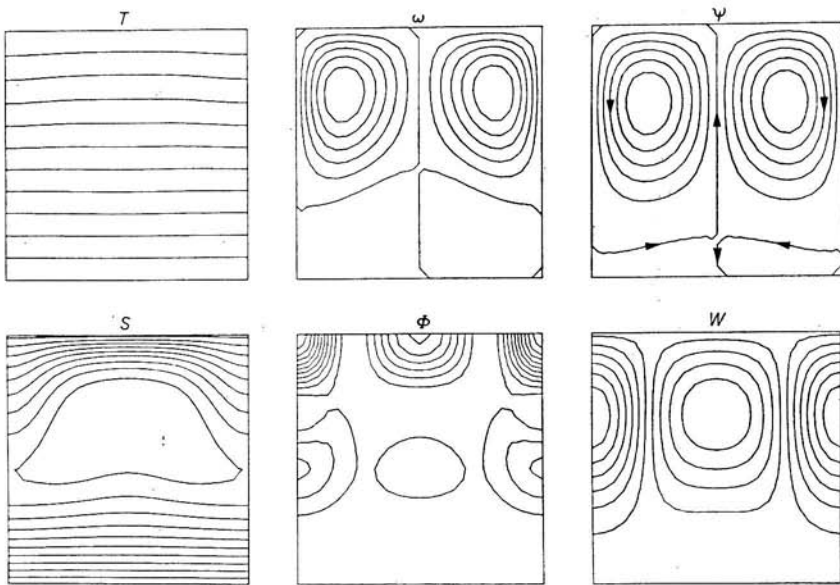


Fig. 32. — Different representations of an anelastic solution at $\tau=0.5$ and $Ra=2.8 \cdot 10^5$. Note particularly the streamfield ψ which demonstrates the penetrative style of convection discussed in the text and the dissipation field Φ which maximizes in the regions of large deformation which occur at the terminations of the hot and cold vertical plumes (from ref. [85]).

dissipation are also shown, is the fact that the shear heating maximizes at the stagnation points where the deformation is largest, as must be the case on intuitive grounds.

In fig. 33 we show the very nice result from [85] illustrating the variation of the « efficiency » of convection as a function of Rayleigh number for various

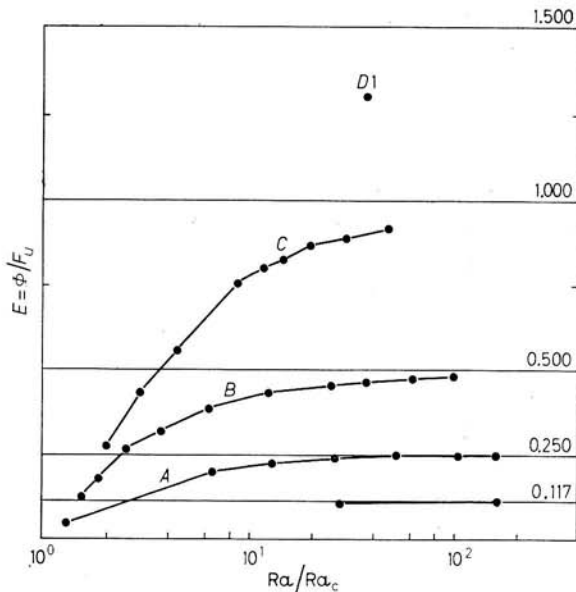


Fig. 33. — The «efficiency» of convection as a function of Rayleigh number showing that τ is indeed an upper bound which is obtained only when the convection is very vigorous.

values of τ . From (58) we have $E = \tau$ for the heated-from-below case and this appears to be a rigorous upper bound upon the efficiency of convection which is only achieved in the limit $Ra \rightarrow \infty$. This is entirely expected, since (58) was derived on the basis of the assumption of vigorous convection and this, of course, occurs at high Rayleigh number.

The geophysically relevant result of the calculations in [85] is that, with τ on the order of that appropriate for the Earth's mantle, the only significant effect is that of the adiabatic gradient. In a global sense, viscous dissipation remains unimportant and, therefore, in the boundary layer theory we may simply replace the temperature drop ΔT by the adiabatic excess $\Delta T - \Delta T_s$ and the same analysis holds for a thick layer. Thus neither of the expressions derived for d in terms of observable quantities, (44a), (44b), are affected, since both were obtained by eliminating ΔT . It has been suggested in [85] that the effect of τ actually leads to a reduction of the aspect ratio of convection which is significant, and, therefore, that one cannot account for the spatial scale of the surface plates simply by assuming whole-mantle convection as I have advocated here. As was pointed out in [64], in order for a significant rescaling of the solutions to occur, the magnitude of τ must be sufficiently large to enforce a penetrative style of convection and the aspect ratio is expected to remain approximately constant until this decoupling of the lower levels occurs. This requires values of τ significantly greater than those estimated for the mantle

and, therefore, the natural explanation for the mean plate size (~ 4000 km) remains the mantlewide convection model in which the circulation has an aspect ratio of ~ 1 .

5.2. *Phase transition, non-Newtonian rheology and sphericity.* — Other factors which may impact upon the simple boundary layer scaling include the presence of phase transitions, possible nonlinearity of the constitutive relation for the mantle fluid and the influence of spherical geometry. The first of these influences which we will discuss here is that of the real divariant solid-solid phase transitions which exist in the mantle. There are two such major discontinuities at depths of 400 km and 670 km which are associated respectively with the transformation of the mineral olivine to the more closely packed spinel phase and of spinel to an oxide configuration [88]. It should be pointed out here that the interpretation of the 670 km seismic discontinuity as a phase change is not universally accepted. It had been suggested by KNOPOFF [89] that phase transitions might inhibit the penetration of convection through them. Subsequent calculation by BUSSE and SCHUBERT [90], SCHUBERT and TURCOTTE [91] and PELTIER [64], which treated the phase transitions as univariant, clarified the physical processes. These involve a competition between a stabilizing influence (the latent-heat release) and a destabilizing influence (the distortion of the phase boundary by thermal advection) when the phase transition is exothermic as in the case of olivine \rightarrow spinel. These calculations showed that this transition would in fact contribute constructively to the instability and thus provide modest enhancement of the convection rather than any inhibition. Finite-amplitude analysis by RICHTER [92] supported the conclusions of the stability calculations.

SCHUBERT *et al.* [93] have considered the additional effects produced by divariance in the case that the phase change is exothermic and show that additional destabilization is liable to occur. We do not presently know whether the spinel-oxide transition (if indeed it exists!) is endothermic or exothermic [94], but, if it were endothermic, it could offer some resistance to penetration of the slab (although in [93] it is shown that this could only be slight) and perhaps contribute to the explanation of the compressive nature of the deep seismic focal mechanisms. A larger stabilizing influence of the seismic discontinuity at 670 km depth could be produced if there were some increase in viscosity associated with it, as has been suggested by MCKENZIE and WEISS [95]. They have stated that an increase of 2 eV in the creep activation energy across the boundary would make the viscosity in the lower mantle so high, in fact, that convection through the boundary would be completely prevented. Given the inability of the constant-viscosity model to simultaneously reconcile both the relative-sea-level and the gravity data, as pointed out in sect. 3, it seems clear that some increase in lower-mantle viscosity will be necessary. As pointed out in [60], however, based upon a systematic relation between oxygen ion

packing and activation energy, it seems that only a very small increase of activation energy is expected, implying a viscosity increase of not much more than one order of magnitude. Such an increase will probably allow us to reconcile both the free-air and sea-level data, but would be completely insufficient to inhibit whole-mantle convection. This notion has recently been reinforced by SABADINI and PELTIER [96], who have inferred a lower-mantle viscosity of about $5 \cdot 10^{23}$ poise by inverting polar-motion data for the nontidal acceleration of rotation and for the polar wander evident in the ILS pole path.

An additional non-Boussinesq effect which we will briefly touch upon here concerns the possibility of nonlinearity of the stress-strain relation, that is the question as to whether non-Newtonian behaviour might be important. Such behaviour is strongly suggested by laboratory data on the creep of olivine, which show that, at high temperature and pressure and at laboratory strain rates, the steady-state regime is controlled by the motion of dislocations by glide and climb. The experimental data are consistent with the power law

$$(61) \quad e = \frac{B_n}{T} \exp \left[-\frac{(E^* + pv^*)}{RT} \right] t^n,$$

where e is the strain rate and t the differential shear stress. The values for the power law exponent n and the activation energy E^* have been determined by KOHLSTEDT and GOETZE [97] and KOHLSTEDT *et al.* [98] to be equal to $n = 3$, and $E^* = 125$ kcal/mol. The pre-exponential dependence upon T^{-1} has not been observed experimentally, but is inferred from theoretical considerations [99]. ROSS *et al.* [100] have recently measured the activation volume and obtained $10.6 \text{ cm}^3/\text{mol} < V^* < 15.4 \text{ cm}^3/\text{mol}$ with a mean of $13.4 \text{ cm}^3/\text{mol}$. If we believe that data taken at the high creep rates which obtain in the laboratory can be extrapolated to the much lower creep rates which occur in the mantle (10^{-6} s^{-1} compared with 10^{-16} s^{-1}), then it would appear that the constitutive relation for at least the upper mantle is quite strongly non-Newtonian and the important question here is whether the influence upon convection will be extreme. PARMENTIER *et al.* [101] have studied non-Newtonian convection with $n = 3$ in the case in which the temperature dependence in (61) is suppressed. This work shows very nicely that, so long as one defines an appropriate equivalent Newtonian viscosity, the properties of the convection are indistinguishable from those which obtain in the constant-viscosity Newtonian case. The stress exponent n must be very much larger than 3 in order to produce a significant variation.

The final complexity which we wish to discuss in this subsection concerns the effects of geometry. Although very little of the work which has been completed to date on the problem of convection in a spherical shell is completely relevant to the mantle, there has been some which is nevertheless quite useful, including that by HSUI *et al.* [102], YOUNG [103], BUSSE [104] and SCHUBERT

and YOUNG [103]. Most of the finite-amplitude work has been concerned with axisymmetric flows which are the spherical equivalent of rolls in a plane layer and most have assumed constant transport coefficients. A single recent example of the results which have been obtained in such work will demonstrate the main effects of geometry for axisymmetric flow in a shell heated from below which has a ratio of inner to outer radius of 0.5 and which is thus close to that for the Earth's mantle. In fig. 34 we show a sequence of mean-temperature

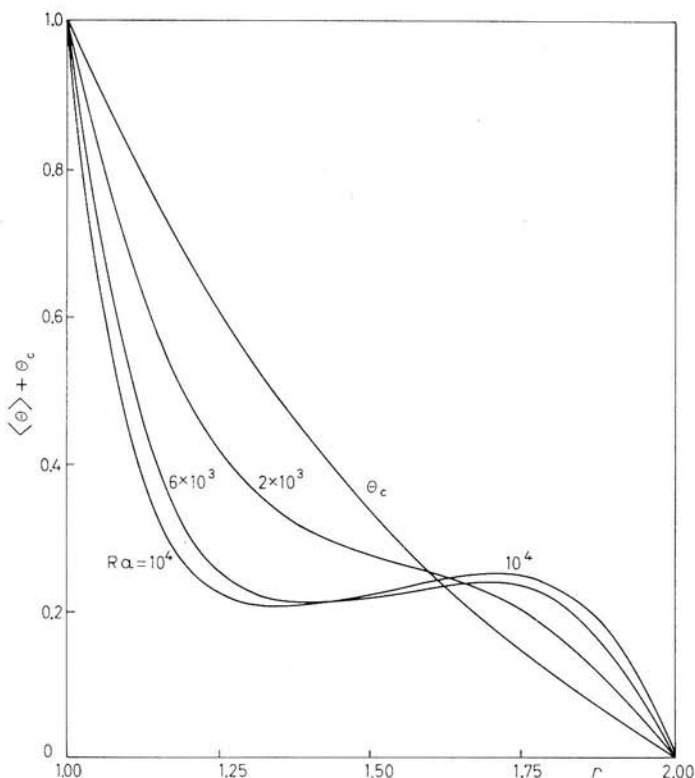


Fig. 34. — Radial variations of mean temperature for convection in a spherical shell with a radius ratio of 0.5 as for the Earth. Note the development of an isothermal core as in the plane-layer case but also the large asymmetry in temperature drops across the two boundary layers which is forced by the spherical geometry (from ref. [106]).

profiles across the shell for a variety of Rayleigh numbers up to about ten times critical. The calculations are from [107] and illustrate the geometric affects quite nicely. On this figure θ_c is the steady-state conduction solution and it will be observed that the conduction profile has a $1/r^2$ curvature rather than being linear as in the plane-layer case. This is simply a consequence of

the fact that in steady state the total heat flow through each boundary of the shell is the same and since the inner shell has smaller surface area and $(K \partial T / \partial r) 4\pi r^2$ must be constant, therefore $\partial T / \partial r$ is greater adjacent to the inner boundary $r = 1$. For these Boussinesq calculations an isothermal core develops in the interior as the Rayleigh number increases, just as in the plane-layer case, but we observe a dramatic difference between the two thermal boundary layers as geometry enforces the radial asymmetry. This effect might not be nearly so dramatic in the Earth because the adiabatic gradient could mitigate against it and this has been neglected in the Boussinesq model. Figure 35

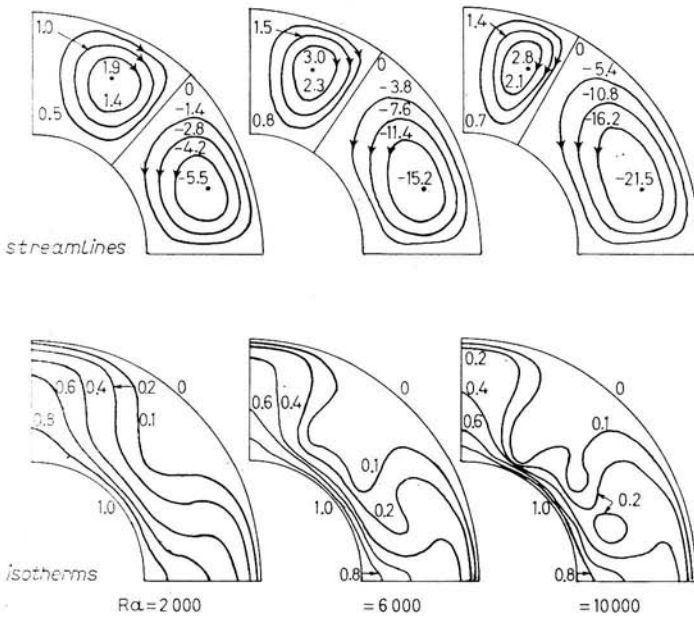


Fig. 35. — Meridional plane projections for axisymmetric solutions in the spherical shell as a function of Rayleigh number. Only even axisymmetric solutions are shown (from ref. [106]).

shows a sequence of meridional plane projections of streamlines and isotherms for these axisymmetric solutions [107] at slightly supercritical Rayleigh number. The motion is dominated by the $l = 4$ mode, there being one fast cell and one slow cell in each hemisphere, and upwelling both at the equator and at the poles. As Ra increases, the fast cell near the equator grows and the counter-rotating cell is progressively more strongly confined to the polar region. These calculations were performed using the spherical equivalent of the Galerkin formulation discussed in sect. 4.

5.3. Temperature dependence of viscosity. — This is by far the most important non-Boussinesq feature of the convective flows to be expected in the

mantle, but there are severe problems associated with the direct incorporation of the temperature dependence in (61) in a thermal-convection calculation. These problems are associated with the high Rayleigh number appropriate to mantle convection; the sharp thermal boundary layers develop huge viscosity contrasts within them and the resolution required in a numerical calculation becomes formidable. These difficulties have been responsible for several attempts to simplify the problem, one of which has involved the introduction of models in which the high-viscosity upper thermal boundary layers are treated as rigid plates which are in a sense decoupled from the mantle beneath them. RICHTER and PARSONS [14] modelled the lithosphere in the laboratory with a mylar sheet, and by dragging this sheet across the upper surface of a heated-from-below convection chamber showed that convection took the form of rolls aligned by the shear in the mean flow into the direction of motion of the sheet. They suggested that such shear-aligned convection could be occurring in the mantle and it has become customary to call the «small scale» of convection beneath the sheet (read lithosphere) the «second» scale of mantle convection. There is no evidence that such convection exists. Indeed, in terms of the simple model of mantle convection which I have stressed here this second scale of convection could obtain only minimal amplitude. The reason for this is clear. If the oceanic lithosphere is the thermal boundary layer of the global convective circulation, then the temperature field beneath it is adiabatic and the boundary layer is not decoupled from the mantle. There can be no strong convective instability beneath the plate because there is no buoyancy to drive it and there is no shear between the plate and the mantle. A recent variation on this theme by PARSONS and MCKENZIE [106] appears to concede that the original idea was suspect, but attempts rather to extract a second scale of convection not from below the boundary layer but rather from within it. Such boundary layer instabilities do occur in constant-viscosity fluids and in fact are responsible for the transition from rolls to bimodal flow in heated below convection as pointed out in sect. 4. In the mantle, however, instability of the upper boundary layer is strongly suppressed by the temperature dependence of viscosity. This is demonstrated explicitly elsewhere in this volume [108]. It is in fact the high viscosity of the upper thermal boundary layer which gives the surface expression of the flow its platelike character. Indeed, the viscosity contrast is so high that the surface material cannot legitimately be called viscous and it is necessary to invoke nonhydrodynamic processes such as melting in order to explain how it is that the hot upcoming stream is able to reach the surface.

If the mantlewide convective circulation is at least partly driven by heating from below (*i.e.* by the efflux of primordial heat from the core), then a lower thermal boundary layer also exists and this will behave in a completely different manner from the upper boundary layer since in it the viscosity will be extremely low. One should, therefore, expect that this boundary layer might

be strongly destabilized and in consequence serve as a source of small-scale plumes, the ascent of which to the surface could well provide an explanation of the observations of volcanism and of hot spots in plate interiors which were discussed in sect. 2 in reference to fig. 6. This idea is made plausible by virtue of the fact that material leaving the lower boundary layer not only has a buoyancy excess, but it also has a viscosity which is much less than that which obtains in the core of the circulation. The ascent of such plumes has been studied in the laboratory by WHITEHEAD and LUTHER [109]. Convective disruption of the low-viscosity lower boundary layer could be the source of the thin fast-rising plumes which they have envisaged and should be seriously considered, therefore, as the source of small-scale thermal anomalies in plate interiors.

It is interesting to note that there is strong seismological evidence for the existence of a boundary-layer-like feature at the base of the planetary mantle. The seismic D'' layer with a thickness of approximately 100 km (about equal to the surface lithospheric thickness) is a region in which the body wave velocities show a rather abrupt departure from the depth variations of a smooth model such as PEM1. The sense of this departure (a velocity decrease) is such as to imply a temperature increase in this region with increasing depth and the simplest explanation of D'' is, therefore, that it exists in consequence of the existence of a thermal boundary layer above the core-mantle interface. Such a boundary layer is expected if the lower mantle is also involved in the convective circulation and its intensity is determined by the heat flux across the core-mantle boundary. The larger this heat flux (weighted by the surface area of the core) as a fraction of the total heat loss at the surface of the planet, the greater the extent to which mantle convection is driven by heating from below. We have already seen that the effect of geometry on steady-state heated-from-below convection in a spherical shell is accommodated in such a way that the thickness of the lower boundary layer remains approximately equal to the thickness of the upper boundary layer for convection in a spherical shell for which the ratio of radii is $\frac{1}{2}$. This means that the temperature drop across the bottom boundary layer is 4 times that across the top boundary layer and that the effect of spherical geometry favours the instability of the lower boundary layer relative to the upper. Even if mantle convection is driven substantially by internal heating, the geometry of the whole-mantle convection model ensures the prominence of the lower boundary layer.

A second class of simplified models which has been introduced in consequence of the complexities associated with the strong temperature dependence of viscosity is distinguished by the fact that they treat the observed surface plate motions as an externally prescribed quantity rather than as one which is to be explained in a self-consistent fashion. Motions in the mantle are assumed to be a consequence of plate motion rather than the cause of them and one is inevitably left asking «why do the plates move?». An example of such a model is that employed by HAGER and O'CONNELL [110], who solve the Stokes

flow equation in a sphere with depth-dependent viscosity subject to surface boundary conditions on tangential velocity appropriate to the observed present-day pattern of plate motions. By fitting the model to the observed Benioff zone dips in convergent regions they show that the purely kinematic model favours whole-mantle convection. It is, however, precisely in the vicinity of the downgoing slab (or the ridge) that the isentropic Stokes flow model is most seriously in error because here the buoyancy force is extremely important and this cannot be encompassed in a Stokes flow model except by *a priori* specification and this certainly cannot be accomplished uniquely. Although such kinematic models provide a useful insight into the constraints imposed upon the flow by geometry, they provide no explanation for the flow itself. The circulations implied by the Stokes flow model modified to include an *a priori* buoyancy field may be quite different from those which actually obtain in the mantle.

Aside from the important boundary layer effects which would be expected in the whole-mantle style of convection which I have advocated here, for a mantle circulation whose rate is governed essentially by the mean viscosity in its interior, an equally important role of temperature-dependent viscosity will be due to the variation of the mean viscosity with time as the planet cools or warms on the time scale of the Earth's age. This brings us to the last major topic which we shall consider in these lectures, which is the thermal history of the Earth. As we shall see, it is in this connection that the temperature dependence of viscosity assumes crucial importance and it is only by considering the thermal history that we may fairly assess the extent to which the model of whole-mantle convection is capable of reconciling the observational data.

6. - Mantle climate and the thermal history of the Earth.

The main characteristic of the simple mantle convection model which we have been concentrating on above is that the circulation fills the entire mantle. One question which this raises concerns the extent to which such efficient convection might lead to solidification of the core and therefore to suppression of the magnetic field. The most stringent test of this possibility is clearly in the case that convection is driven entirely by core cooling (*i.e.* by heating from below). In [47, 111] we have constructed simple «parameterized» convection models in order to explore this and other questions pertaining to the mantle convection hypothesis elaborated previously.

The basic idea is to design a thermal-history model which incorporates the effect of mantle convection upon the internal-energy budget but does not require explicit solution of the Navier-Stokes equations. This is possible to the extent that the convection may be considered quasi-steady and this depends

upon the separation of the «fast» dynamic time scale associated with the circulation and the «slow» time scale on which the mean temperature of the interior changes. With such models we may focus our attention upon the long-time-scale «climatic» variability of the system and ignore the short-time-scale fluctuations in «weather» which doubtless occur, but which are only of local interest. This is an approach to the thermal-history problem which TOZER [112] has long advocated and we have developed a simple scheme for realizing such a model objectively. It is based upon the boundary layer theory for heated below convection which was elaborated at the end of sect. 4 and which we used to demonstrate the compatibility of the observations with the whole-mantle convection hypothesis.

The simplest version of the model is based upon the following relations:

$$(62a) \quad \frac{d\bar{T}}{dt} = -\frac{3q_c}{r_c \bar{\rho}_c C_{pc}} = -\frac{6q_m}{\bar{\rho}_m C_{pm}} \frac{r_p^2}{r_p^2 - r_c^2},$$

$$(62b) \quad q_c \simeq \frac{\alpha \Delta T}{d} \cdot \text{Nu} \cdot \frac{r_c^2}{(r_c + (r_p - r_c)/2)^2},$$

$$(62c) \quad \text{Nu} = \left(\frac{\text{Ra}}{\text{Ra}_c} \right)^{\frac{1}{2}},$$

$$(62d) \quad \text{Ra} = \frac{g \alpha d^3 (\Delta T - \Delta T_s)}{\nu},$$

$$(62e) \quad \nu = \nu_0 \exp [g^* T_m / \bar{T}],$$

$$(62f) \quad q_s = q_c \left(A + \frac{r_c^2}{r_p^2} \right),$$

$$(62g) \quad A = \frac{1}{2} \frac{C_{pm} \bar{\rho}_m}{C_{pc} \bar{\rho}_c} \left(\frac{r_p}{r_c} - \frac{r_c}{r_p} \right)^2,$$

where \bar{T} is the mean temperature of the core, q_c is the heat flux through the core-mantle boundary and q_m that through the surface of the planet due to direct convection, r_c and r_p are the radii of the core and of the planet, $\bar{\rho}_c$ and $\bar{\rho}_m$ are the mean densities of the core and the mantle, ΔT is the temperature drop across the mantle and ΔT_s the adiabatic drop, and q_s is the total heat flux through the surface due to the combined effects of heat loss from the core and the secular cooling of the mantle. ν is the mean mantle viscosity which depends exponentially upon temperature through Weertman's empirical relation (62e) in which \bar{T} is the temperature in the mantle at some mid-level (*i.e.* outside the boundary layers) where the melting temperature is T_m . The mantle temperature profile is constructed by matching conduction solutions in the upper and lower boundary layers, whose thicknesses are determined by the boundary layer theory when the heat flux is known, to an adiabatic profile in the interior as discussed in the last section and in [111]. Given the mantle

temperature profile we can compute \bar{T} and then ν for use in the Rayleigh number. The factors involving r_c and r_p are corrections introduced in the boundary layer theory to account for the geometry of the sphere. That we can do this is a rather strong assumption which should be checked directly in a complete shell model. The single prognostic equation (62a) is time-stepped using a Δt sufficiently small that further reduction does not change the time history.

As can be readily seen by inspection of the simple system (62), it is one which will be strongly controlled by the negative feedback due to the exponential dependence of viscosity upon temperature. If for some reason the mantle is becoming warmer in the mean, then from (62e) the viscosity will drop and from (62d) the Rayleigh number will increase. From (62c) this causes the Nusselt number to increase and from (62b) the heat flux will likewise increase and thus also the rate of cooling which will then act contrary to the original tendency. The opposite reaction will occur if the mantle is cooling. The system is, therefore, self-regulating as TOZER has suggested in the context of somewhat different models.

Since a description of the thermal history involves the solution of an initial-value problem, an important ingredient in the calculation is a model for planetary formation from which one may deduce an initial temperature profile. In [47, 111] we used simple Mizutani models of accretion with short accretion times such that the initial states were hot enough that most of the primordial material was melted and thus gravitational differentiation of iron from silicates was complete by the end of formation. KAULA [113] has recently shown, using the more realistic Safronov model, that it is indeed very difficult to avoid a hot initial gravitationally differentiated state. Such models may have their mantles largely depleted of radioactivity since the ions U^{IV} , Th^{IV} , K^I would be squeezed upwards in the remaining melt during the primary crystallization of the mantle because they are incompatible with the major phases. This is based upon the notion that the silicate solidus is superadiabatic everywhere and therefore that the mantle, cooling on an adiabat, solidified from the bottom upwards. As argued in [111] it is, therefore, quite plausible that the thermal state of the planet 4.5 billion years ago was similar to that which exists today in the sense that it had a liquid iron core and a solid silicate mantle and it is also plausible perhaps that the mantle contained little radioactivity. We will illustrate the thermal-history model embodied in (62) by using it to see whether we can obtain a rational thermal history in which mantle convection is driven entirely from below and the initial state is hot.

In fig. 36 we show an example of such a calculation from [111] in which, although the mantle is depleted of radioactivity, the core contains 0.2% potassium as has been suggested by GOETTEL [114]. Inspection of the thermal evolution shows that mantle convection is initially unable to keep up with the radioactive heating of the core so that the core warms at first. Eventually,

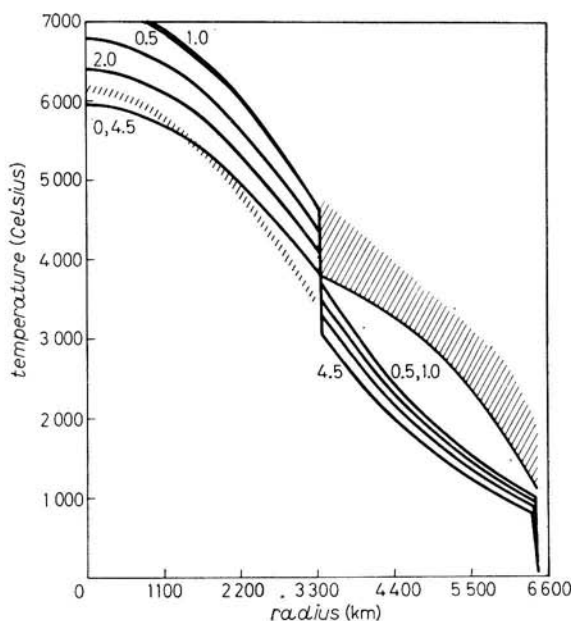


Fig. 36. — Evolution of the mean radial temperature profile with time for a thermal-history model with parameterized convection in the mantle.

however, it begins to cool and the core geotherm intersects the superadiabatic core solidus and the inner core begins to grow in radius. The latent heat released upon solidification is taken into account in the calculation. As the planet continues to cool, the viscosity of the mantle rises and the Rayleigh number falls, which inhibits further cooling. As suggested by the high-Rayleigh-number asymptotic analysis, the boundary layers of the mantle convective circulation begin to thicken. In fig. 37 we show the variation with time of the mean mantle viscosity, the mantle Rayleigh number and the mean mantle heat flow for this model which is labelled 6*b*) in the figure. The curve marked 4*b*) is for a model which differs from the present one in that it has no potassium in the core, while the dashed curve is for a model of Venus which also has potassium in its core but which actually cools so quickly that the core solidifies. Model 6*b*) clearly fits the observed present-day mantle viscosity and the mantle contribution to the observed heat flow and it also fits the observed present-day boundary layer thickness. Of course, whether this is the way the Earth evolved in fact, we are in no position to say!

Simple parameterized convection models, such as the one discussed here, would seem to provide a very useful tool for the future investigation of alternative thermal histories. They can easily be modified to include the contribution to the convective heat flux due to radioactivity in the mantle, and alternative models in which this effect is included should certainly be studied.

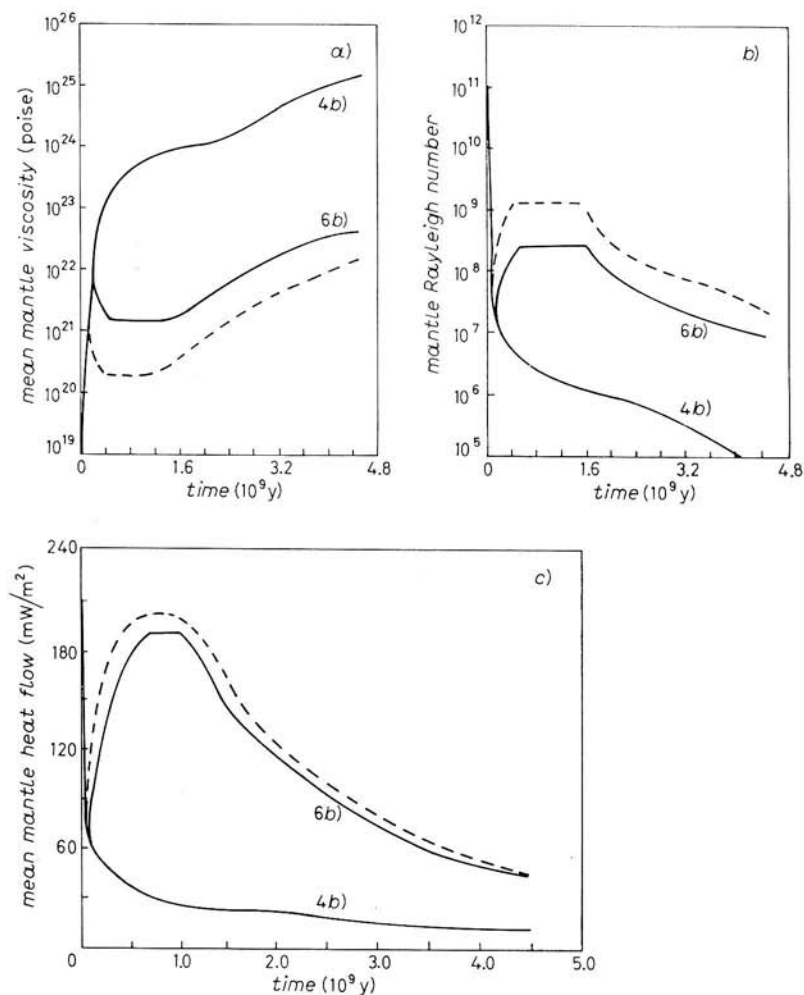


Fig. 37. — Evolution with time of the mean mantle viscosity (a), the mantle Rayleigh number (b) and the mean mantle heat flow (c) for a sequence of thermal-history models with parameterized convection in the mantle.

In fact, on the basis of the geochemical arguments presented in O'NIONS *et al.* [115], there is now clear reason to believe that the mantle is at present largely depleted of radioactivity, but that it did not begin to differentiate this material into stable continental crust much before about 3 billion years ago. It seems quite likely that when this process is included in the thermal-history model there will be no necessity of including any radioactivity in the Earth's core, the initial mantle radioactivity being sufficient to prevent too rapid cooling of the planet.

7. - Conclusions.

The simple model of mantle convection advocated in the preceding sections differs considerably from the view which seems to prevail in most of the current literature. Although it is almost certainly oversimplistic in that it neglects mantle radioactivity entirely, it is nevertheless able to meet all of the observational constraints quite nicely. These include the mean plate size, the plate speeds, the mean surface heat flow and the thermal-boundary-layer thickness. The model does not appear to violate the constraints upon the thermal history if at least some radioactivity is included in the mantle and the contribution to the surface heat flux due to the secular cooling of the mantle is properly accounted for. The main assumption on which the model is based is that the convective circulation is mantle wide, filling the entire region between the core-mantle boundary and the Earth's surface.

Perhaps the most important ingredient in this theory is the magnitude of the viscosity of the mantle which is inferred from studies of postglacial rebound. Although this work has yet to provide a clear and unambiguous picture of the detailed variations of mantle viscosity with depth, there are two extremely useful pieces of information from it which contribute in an important way to the internal consistency of the whole-mantle convection picture. These are the nominal mean mantle viscosity of $\sim 5 \cdot 10^{22}$ poise and the inferred thermal boundary layer thickness $\delta \simeq 100$ km. Both of these numbers are in fact required in the convection hypothesis and do establish the internal consistency of the model.

On the basis of the recent numerical work discussed in sect. 5, we are in a reasonably good position today to be able to assess the importance of two non-Boussinesq effects upon mantle convection. Neither viscous dissipation nor non-Newtonian rheology can be expected to contribute in an important way to the properties of the circulation. The only important effect of the dissipation number, τ , is through the adiabatic temperature gradient which reduces the vigour of convection and, because of the linear dependence of the adiabatic gradient upon T , introduces some radial asymmetry of the mean fields. The most important effect of the presence of phase transitions (?) is liable to be associated with the seismic discontinuity at 670 km depth across which there should be some increase of the creep activation energy and therefore an increase of viscosity. Some increase of viscosity at this depth also seems necessary to reconcile the magnitude of the observed free-air gravity anomaly over the Canadian Shield and the observed polar wander and nontidal acceleration of the Earth's rotation [96]. It does not appear that the maximum increase of viscosity allowed by the data would be sufficient to inhibit the penetration of convection through the 670 km discontinuity, but it should be sufficient to perhaps explain why the deep seismic focal mechanisms are compressive.

In this picture of whole-mantle convection there is an implicit interpretation of the oceanic lithosphere. It is to be considered the cold thermal boundary layer of the mantlewide convective circulation, whose platelike behaviour occurs in consequence of its enormous viscosity. Because of the great contrast in viscosity between the lithosphere and the mantle beneath it, it is in fact possible to approximately measure the thermal-boundary-layer thickness using glacial-rebound data as we have described. The Prandtl number of the mantle fluid is so high that any very significant shear between the lithosphere and the mantle is unlikely and strong coupling of the motion field is to be expected. It seems unlikely in this connection that any second scale of convection of large amplitude could occur in the near-surface region although a weak instability is not inconceivable. It should not develop beneath the plates, because beneath the thermal boundary layer the temperature gradient is adiabatic. It cannot develop within the plates because the viscosity is enormous and this suppresses the thermal instability. The hypothesis of in approximately 1:1 relation between plates and convection cells seems to fit all of the observational constraints rather well. Perhaps the most interesting idea to emerge from the thermal-history analyses is an understanding of the extent to which the primordial heat content of the planet may contribute to driving its convective circulations if the system is cooling.

* * *

I am very much indebted to my colleagues S. DE LA CRUZ, G. JARVIS, J. SCHUBERT and H. SHARPE for their permission to draw upon some of their unpublished materials in the course of writing these lecture notes.

REFERENCES

- [1] H. H. HESS: *Petrological Studies*, edited by A. E. J. ENGEL, H. L. JANES and B. L. LEONARD (New York, N. Y., 1962), p. 599.
- [2] R. S. DIETZ: *Nature (London)*, **190**, 854 (1961).
- [3] W. J. MORGAN: *J. Geophys. Res.*, **73**, 1959 (1968).
- [4] D. P. MCKENZIE and R. L. PARKER: *Nature (London)*, **216**, 1276 (1967).
- [5] D. W. FORSYTH and S. UYEDA: *Geophys. J. R. Astron. Soc.*, **43**, 163 (1975).
- [6] W. M. CHAPPEL and T. E. TULLIS: *J. Geophys. Res.*, **82**, 1967 (1977).
- [7] F. RICHTER and D. P. MCKENZIE: *J. Geophys. Res.*, **44**, 441 (1978).
- [8] R. K. MCCONNELL: *J. Geophys. Res.*, **73**, 7089 (1968).
- [9] D. P. MCKENZIE: *J. Geophys. Res.*, **71**, 3995 (1966).
- [10] D. P. MCKENZIE, J. M. ROBERTS and N. O. WEISS: *J. Fluid Mech.*, **62**, 465 (1974).
- [11] F. M. RICHTER: *Rev. Geophys. Space Phys.*, **11**, 223 (1973).
- [12] F. M. RICHTER: *Ann. Rev. Earth Planet. Sci.*, **6**, 9 (1978).
- [13] A. HOLMES: *Principles of Physical Geology* (New York, N. Y., 1965).

- [14] F. M. RICHTER and B. PARSONS: *J. Geophys. Res.*, **80**, 2529 (1975).
- [15] D. P. MCKENZIE and F. RICHTER: *Sci. Am.*, **235**, 72 (1976).
- [16] D. C. TOZER: *Sci. Prog. (London)*, **64**, 1 (1977).
- [17] E. R. OXBURGH and D. L. TURCOTTE: *Rep. Prog. Phys.*, **41**, 1249 (1978).
- [18] G. SCHUBERT: *Ann. Rev. Earth Planet. Sci.*, **7**, 289 (1979).
- [19] E. C. BULLARD, J. E. EVERETT and A. G. SMITH: *Philos. Trans. R. Soc. London Ser. A*, **258**, 41 (1965).
- [20] R. S. DIETZ and J. C. HOLDEN: *Sci. Am.*, **229**, 103 (1970).
- [21] S. K. RUNCORN: *Geol. Assoc. Canada Proc.*, **8**, 77 (1956).
- [22] M. W. MCELHINNY: *Nature (London)*, **241**, 523 (1973).
- [23] F. J. VINE and D. W. MATHEWS: *Nature (London)*, **199**, 947 (1963).
- [24] J. R. HERTZLER, X. LE PICHON and J. G. BARON: *Deep Sea Res.*, **13**, 427 (1966).
- [25] A. COX, G. B. DALRYMPLE and R. R. DOELL: *Sci. Am.*, **216**, 44 (1967).
- [26] A. COX: *Science*, **163**, 237 (1969).
- [27] F. J. VINE: *Science*, **154**, 1405 (1966).
- [28] H. BENIOFF: *Geol. Soc. Am. Bull.*, **65**, 385 (1954).
- [29] G. PLAFKER: *Science*, **148**, 1675 (1965).
- [30] W. STAUDER and G. A. BOLLINGER: *J. Geophys. Res.*, **71**, 5283 (1966).
- [31] B. ISACKS and P. MOLNAR: *Rev. Geophys. Space Phys.*, **9**, 103 (1971).
- [32] B. ISACKS, J. OLIVER and L. R. SYKES: *J. Geophys. Res.*, **73**, 5855 (1968).
- [33] J. T. WILSON: *Nature (London)*, **207**, 343 (1965).
- [34] D. P. MCKENZIE and W. J. MORGAN: *Nature (London)*, **224**, 125 (1969).
- [35] X. LE PICHON: *J. Geophys. Res.*, **73**, 3661 (1968).
- [36] B. PARSONS and J. G. SCLATER: *J. Geophys. Res.*, **82**, 803 (1977).
- [37] D. P. MCKENZIE and J. G. SCLATER: *J. Geophys. Res.*, **73**, 3173 (1968).
- [38] M. TALWANI, J. L. WORZEL and M. EWING: *J. Geophys. Res.*, **66**, 1265 (1961).
- [39] D. T. GRIGGS: in *The Nature of the Solid Earth*, edited by E. C. ROBERTSON *et al.* (New York, N. Y., 1972), p. 36.
- [40] D. P. MCKENZIE: *Geophys. J. R. Astron. Soc.*, **18**, 1 (1969).
- [41] J. W. MINEAR and M. N. TOKSÖZ: *J. Geophys. Res.*, **75**, 1397 (1970).
- [42] W. M. KAULA: in *The Nature of the Solid Earth*, edited by E. C. ROBERTSON *et al.* (New York, N. Y., 1972), p. 385.
- [43] E. M. GAPOSCHKIN and K. LAMBECK: *J. Geophys. Res.*, **76**, 4855 (1971).
- [44] W. J. MORGAN: *Nature (London)*, **230**, 42 (1971).
- [45] W. J. MORGAN: *Bull. Am. Assoc. Petr. Geol.*, **56**, 203 (1972).
- [46] W. J. MORGAN: *Studies in Earth and Space Sciences, Geol. Soc. Am. Mem.*, **132** (Boulder, Colo., 1972), p. 7.
- [47] H. N. SHARPE and W. R. PELTIER: *Geophys. Res. Lett.*, **5**, 737 (1978).
- [48] W. M. ELSASSER: *J. Geophys. Res.*, **76**, 1101 (1969).
- [49] C. HILLAIRE-MARCEL and R. W. FAIRBRIDGE: *Geology*, **6**, 117 (1978).
- [50] R. I. WALCOTT: *Rev. Geophys. Space Phys.*, **10**, 849 (1972).
- [51] W. R. PELTIER: *Rev. Geophys. Space Phys.*, **12**, 649 (1974).
- [52] F. GILBERT and A. DZIEWONSKI: *Philos. Trans. R. Soc. London Ser. A*, **278**, 187 (1975).
- [53] W. R. PELTIER: *Geophys. J. R. Astron. Soc.*, **46**, 669 (1976).
- [54] M. D. CRITTENDEN: *J. Geophys. Res.*, **68**, 5517 (1963).
- [55] W. E. FARRELL and J. A. CLARKE: *Geophys. J. R. Astron. Soc.*, **46**, 647 (1976).
- [56] W. R. PELTIER, W. E. FARRELL and J. A. CLARK: *Tectonophysics*, **50**, 81 (1978).
- [57] J. A. CLARK, W. E. FARRELL and W. R. PELTIER: *Quat. Res. (N. Y.)*, **9**, 265 (1978).
- [58] W. R. PELTIER and J. T. ANDREWS: *Geophys. J. R. Astron. Soc.*, **46**, 605 (1976).
- [59] W. R. PELTIER and P. WU: *Geophys. J. R. Astron. Soc.*, submitted (1980).

- [60] C. G. SAMMIS, J. C. SMITH, G. SCHUBERT and D. YUEN: *J. Geophys. Res.*, **82**, 3747 (1977).
- [61] J. WEERTMAN: *Rev. Geophys. Space Phys.*, **8**, 145 (1970).
- [62] L. M. CATHLES: *Viscosity of the Earth's Mantle* (Princeton, N. J., 1976).
- [63] R. I. WALCOTT: *Can. J. Earth. Sci.*, **7**, 716 (1970).
- [64] W. R. PELTIER: *Geophys. Fluid Dyn.*, **3**, 365 (1972).
- [65] F. H. BUSSE: *Rep. Prog. Phys.*, **41**, 1929 (1978).
- [66] W. R. PELTIER: *Thermal stability of non-Boussinesq configurations*, unpublished Ph. D. dissertation, Department of Physics, University of Toronto (1971).
- [67] F. H. BUSSE and J. A. WHITEHEAD: *J. Fluid Mech.*, **47**, 305 (1971).
- [68] F. H. BUSSE and J. A. WHITEHEAD: *J. Fluid Mech.*, **66**, 67 (1974).
- [69] J. A. WHITEHEAD and B. PARSONS: *Geophys. Astrophys. Fluid Dyn.*, **9**, 201 (1978).
- [70] S. DE LA CRUZ: *Convection with internal heat generation*, unpublished M. Sc. dissertation, Department of Physics, University of Toronto (1973).
- [71] S. DE LA CRUZ: *Convection with internal heat generation and the dynamics of the Earth's mantle*, Ph. D. dissertation, Department of Physics, University of Toronto (1980 expected).
- [72] J. M. STRAUS: *J. Fluid Mech.*, **56**, 353 (1972).
- [73] R. M. CLEVER and F. H. BUSSE: *J. Fluid Mech.*, **65**, 625 (1974).
- [74] S. CHANDRASEKHAR: *Hydrodynamic and Hydromagnetic Stability* (Oxford, 1961).
- [75] D. L. TURCOTTE and E. R. OXBURGH: *J. Fluid Mech.*, **28**, 29 (1967).
- [76] G. O. ROBERTS: *Geophys. Astrophys. Fluid Dyn.*, in press (1979).
- [77] P. OLSON and G. M. CORCOS: *Geophys. J. R. Astron. Soc.*, submitted (1979).
- [78] D. R. MOORE and N. O. WEISS: *J. Fluid Mech.*, **58**, 289 (1973).
- [79] H. N. POLLACK and D. S. CHAPMAN: *Earth Planet. Sci. Lett.*, **34**, 174 (1977).
- [80] W. M. ELSASSER, P. OLSON and B. D. MARSH: *J. Geophys. Res.*, **84**, 147 (1979).
- [81] D. L. TURCOTTE and E. R. OXBURGH: *Ann. Rev. Fluid Mech.*, **4**, 33 (1972).
- [82] G. T. JARVIS and W. R. PELTIER: *Nature (London)*, in press (1980).
- [83] W. R. PELTIER, J. HALLÉ and T. L. CLARK: *Geophys. Astrophys. Fluid Dyn.*, **10**, 53 (1978).
- [84] W. R. PELTIER and T. L. CLARK: *J. Atmos. Sci.*, **36**, 1498 (1979).
- [85] G. JARVIS and D. P. MCKENZIE: *J. Fluid Mech.*, **96**, 515 (1980).
- [86] J. M. HEWITT, D. P. MCKENZIE and N. O. WEISS: *J. Fluid Mech.*, **68**, 721 (1975).
- [87] W. V. R. MALKUS: *Geophys. Fluid Dyn.*, **4**, 267 (1973).
- [88] A. E. RINGWOOD: *Phys. Earth Planet. Inter.*, **3**, 109 (1970).
- [89] L. KNOPOFF: *Rev. Geophys.*, **2**, 89 (1964).
- [90] F. H. BUSSE and G. SCHUBERT: *J. Fluid Mech.*, **46**, 801 (1971).
- [91] G. SCHUBERT and D. L. TURCOTTE: *J. Geophys. Res.*, **76**, 1424 (1971).
- [92] F. M. RICHTER: *Geophys. J. R. Astron. Soc.*, **35**, 265 (1973).
- [93] G. SCHUBERT, D. A. YUEN and D. L. TURCOTTE: *Geophys. J. R. Astron. Soc.*, **42**, 705 (1975).
- [94] R. C. LIEBERMANN, I. JACKSON and A. E. RINGWOOD: *Geophys. J. R. Astron. Soc.*, **50**, 553 (1977).
- [95] D. MCKENZIE and N. WEISS: *Geophys. J. R. Astron. Soc.*, **42**, 131 (1975).
- [96] R. SABADINI and W. R. PELTIER: *Geophys. J. R. Astron. Soc.*, in press (1980).
- [97] D. L. KOHLSTEDT and C. GOETZE: *J. Geophys. Res.*, **79**, 2045 (1974).
- [98] D. L. KOHLSTEDT, C. GOETZE and W. B. DURHAM: *The Physics and Chemistry of Minerals and Rocks*, edited by R. STRENS (London, 1976), p. 35.
- [99] J. WEERTMAN: *Trans. ASME*, **61**, 681 (1968).
- [100] J. V. ROSS, H. G. AVÉ LALLEMANT and N. L. CARTER: *Eos*, **59**, 374 (abstract) (1978).

- [101] E. M. PARMENTIER, D. L. TURCOTTE and K. E. TORRANCE: *J. Geophys. Res.*, **81**, 1839 (1976).
- [102] A. T. HSUI, D. L. TURCOTTE and K. E. TORRANCE: *Geophys. Fluid Dyn.*, **3**, 35 (1972).
- [103] R. E. YOUNG: *J. Fluid Mech.*, **63**, 695 (1974).
- [104] F. H. BUSSE: *J. Fluid Mech.*, **72**, 67 (1975).
- [105] G. SCHUBERT and R. E. YOUNG: *Tectonophysics*, **35**, 201 (1976).
- [106] B. PARSONS and D. MCKENZIE: *J. Geophys. Res.*, **83**, 4485 (1978).
- [107] A. ZEBIB, J. SCHUBERT and J. M. STRAUS: *J. Fluid Mech.*, in press (1979).
- [108] D. A. YUEN and W. R. PELTIER: this volume, p. 432.
- [109] J. A. WHITEHEAD and D. S. LUTHER: *J. Geophys. Res.*, **80**, 705 (1975).
- [110] B. H. HAGER and R. J. O'CONNELL: *Tectonophysics*, **50**, 111 (1978).
- [111] H. N. SHARPE and W. R. PELTIER: *Geophys. J. R. Astron. Soc.*, **59**, 171 (1979).
- [112] D. C. TOZER: *Phys. Earth Planet. Inter.*, **6**, 182 (1972).
- [113] W. M. KAULA: *J. Geophys. Res.*, **84**, 999 (1979).
- [114] K. A. GOETTEL: *Geophys. Surv.*, **2**, 369 (1976).
- [115] R. K. O'NIONS, N. M. EVENSEN, P. J. HAMILTON and S. R. CARTER: *Philos. Trans. R. Soc. London Ser. A*, **258**, 547 (1978).