Tectonophysics, 38 (1977) 119—136

ON THE SCALE OF MANTLE CONVECTION

RICHARD J. O'CONNELL

Department of Geological Sciences, Harvard University, Cambridge, Massachusetts (U.S.A.)

(Manuscript received March 24, 1976; revised version received May 4, 1976)

ABSTRACT

O'Connell, R.J., 1977. On the scale of mantle convection. In: J. Bonnin and R.S. Dietz (editors), Present State of Plate Tectonics. Tectonophysics, 38 (1—2): 119—136.

Observational evidence from glacial rebound and gravity anomalies shows that the mantle has a relatively uniform viscosity. Such a structure is consistent with the effects of temperature and pressure on microscopic mechanisms for plastic flow of solids. Thus the asthenosphere may extend to the base of the mantle. Mantle phase changes do not inhibit mantle-wide convection, and any compositional differences between the upper and lower mantle are not resolved at present. Thus there is no reason not to expect convection extending into the lower mantle. The stress state of descending slabs is consistent with mantle-wide convection, and although the slabs may encounter an increased resistance to sinking at ~700 km depth there is no evidence that they are stopped by the increased resistance. Seismic heterogeneity in the lower mantle suggests large-scale convection, which is perhaps driven by heat from the earth's core. Plate motions may interact with large-scale flow in the mantle, and are probably not completely decoupled from flow in the lower mantle. Lateral variations in flow properties in the upper mantle may be more important than vertical variations, and may strongly influence plate motions.

INTRODUCTION

The relative motion of lithospheric plates must be accompanied by some sort of flow of mantle material at depth; in fact thermal convection in the mantle is widely accepted as the most likely driving mechanism for the plate motions. The scale of any such convection is at present unknown, but it should correspond to the size of the plates. Since these vary in scale from very large plates, such as the Pacific plate (~6000 km), to smaller plates, such as the Caribbean (~1000 km), we might expect a corresponding variation in scale of mantle flow. Nevertheless, the largest plates indicate that there will be large-scale components of flow with lateral dimensions of several thousand kilometers.

Most models of mantle convection have limited convection to the upper mantle (e.g. Turcotte and Oxburgh, 1972; Richter, 1973a, b; McKenzie et al.,

1974) with a maximum depth of ~1000 km or less. This choice was based on a number of arguments, including the belief that the 'viscosity' of the lower mantle is considerably greater than that of the upper mantle and that sinking lithospheric slabs do not penetrate deeper than the deepest earthquakes. While such models can produce reasonable agreement with geophysical observations this does not mean that the models are correct or that other models could not satisfy the observations equally well.

The object of this paper is to review data that indicate that there is no compelling reason to believe that mantle convection should be confined to the upper mantle, and to suggest that models of mantle-wide convection should be systematically investigated. In fact, there is considerable evidence that convection does indeed extend to the lower mantle, which has considerable implications for the evolution of the earth.

Mantle-wide convection has not been totally neglected. Tozer (1972) has maintained that thermal convection in the earth is necessarily mantle-wide owing to the self-regulating feedback effect of the extreme temperature dependence of the flow laws for solids, and McKenzie and Weiss (1975) have used this approach in their discussion of the earth's thermal history. Nevertheless, the idea that the lower mantle does not participate in large-scale flow has become sufficiently entrenched in the literature that it is deemed worthwhile to systematically review the evidence bearing on the possibility or not of convection in the lower mantle.

We do not presume to "prove" that the lower mantle does convect, but we do hope to show that the evidence that it does not convect is not very strong at all. We will first discuss the flow properties of the mantle as deduced from geophysical observations and considerations of plastic flow mechanisms in solids and see that there is no indication that a large increase in 'viscosity' with depth in the mantle exists or should be expected to exist.

We will also consider whether phase changes in the mantle present a barrier to convection, and the extent that the absence of earthquakes deeper than 700 km indicates that no flow penetrates deeper than that. Recent observations of seismic heterogeneities in the lower mantle will be discussed, and lastly, some implications of mantle-wide convection on the forces moving lithospheric plates will be discussed.

MANTLE VISCOSITY

Several different investigations of the viscosity of the mantle indicated that it increased markedly with depth. More recent work has indicated, though, that this is not necessarily the case, McConnell (1965) deduced the variation of viscosity with depth in the upper mantle from a Fourier analysis of post-glacial rebound of Fennoscandia. His resulting model had a low viscosity channel at 200—400 km depth with an increase in viscosity by roughly $1\frac{1}{2}$ orders of magnitude at a depth of 1000 km. The coincidence of the low-viscosity channel with the seismic low-velocity zone led Anderson

(1966) to suggest that there was a physical correlation between the high attenuation of seismic waves in the low-velocity zone and the low viscosity at the same depth. This correlation has also led to the widespread indentification of the seismic low-velocity zone (and region of high attenuation) with the asthenosphere or low-viscosity zone, in spite of the fact that quite different phenomena are involved with the two characterizations. It further suggested that the asthenosphere was a relatively thin layer overlying a more viscous lower mantle.

The increase in viscosity with depth that McConnell found was dictated by the long-wavelength components of the post-glacial rebound. Yet he later recognized (McConnell, 1968) that the long-wavelength data were affected by his assumption about the deformation at large distances from ice sheet and was therefore unreliable. He attempted to constrain the long-wavelength data by assuming that the non-hydrostatic bulge of the earth was due to incomplete recovery from deglaciation. However, this assumption was shown to be unwarranted by O'Connell (1971), so that the increase of viscosity with depth below about 600 km is relatively unconstrained by reliable data. In addition Parsons (1972) showed that the uplift data used by McConnell were not sufficient to unambiguously resolve *any* low-vicosity channel, and O'Connell (1971) showed that the increase of viscosity with depth that McConnell did find was artificial owing to his use of a plane rather than spherical model. The net result is that the post-glacial uplift data from Fennoscandia neither resolve a distinct low-viscosity channel in the upper mantle, nor require a significant increase of viscosity below a depth of several hundred kilometers.

Post-glacial rebound in North America depends more on the viscosity of the lower mantle owing to the greater size of the Laurentide ice sheet. Cathles (1975) has pointed out that the presence of a peripheral bulge is a direct indication that the rebound is accomplished by flow that extends into the lower mantle, and finds that a model with uniform viscosity of roughly 10^{22} poise (10^{21} Ns/m^2) throughout the mantle can satisfy the uplift data. Peltier (1974) has reached a similar conclusion, but has pointed out the need for a more thorough and consistent analysis of the North American data, as have Farrell and Clark (1976). It should be noted that Walcott (1973) finds that considerations including residual gravity anomalies associated with the present unrecovered rebound indicate that the viscosity increases with depth beneath a low-viscosity channel in the upper mantle. Cathles (1975, p. 247) has shown that much of the uplift data is consistent with either model, and that the choice depends on the significance one assigns to the present gravity anomalies in Canada and to the indications of a peripheral bulge. We may expect that the extension of Peltier's (1974) work will resolve the problem, but at present we must conclude that the rebound data from North America are perfectly consistent with a mantle of essentially uniform viscosity.

The viscosity of the lower mantle has also been estimated by interpreting the earth's non-hydrostatic bulge. McKenzie (1966) attributed the bulge to the delayed response of the earth to secular angular deceleration from tidal friction, after MacDonald (1963), and concluded that the lower-mantle viscosity was several orders of magnitude greater than that of the upper mantle. McKenzie's attribution of the bulge was questioned by Goldreich and Toomre (1969) who placed an upper limit of 10^{24} poise on the viscosity of the lower mantle from the rate of polar wander deduced from paleomagnetic data. Although McKenzie et al. (1974) argued that both estimates were based on logically consistent arguments, there is no evidence whatsoever that the bulge is a transient associated with the earth's secular decleration, whereas there are several other indications that the lower mantle need not have a very high viscosity.

The non tidal secular angular acceleration of the earth, was used by Dicke (1969) and O'Connell (1971) to estimate the change in the earth's moment of inertia due to global post-glacial rebound; both found that a uniform mantle viscosity of the order of 10^{22} poise was indicated. In addition O'Connell showed that the absence of large-scale gravity anomalies due to deglaciation ruled against a high-viscosity lower mantle. Although Peltier (1974) has criticized O'Connell's (1971) assumption of an exponential rebound response, inspection of Peltier's calculated response curves shows that O'Connell's assumption was a perfectly reasonable approximation for the average long-term response.

In summary, the estimates of mantle viscosity from post-glacial rebound and the associated changes in the earth's moment of inertia give no clear indication of a large increase in viscosity with depth in the mantle. In fact, granting the uncertainties inherent in the models used to interpret the data, the indications are that the mantle may be adequately characterized by a roughly uniform viscosity of the order of 10^{22} poise, with possible local variations and fine structure in the upper mantle.

PLASTIC FLOW MECHANISMS

The flow properties of the mantle can also be estimated by considering the effects of temperature and pressure on the microscopic mechanisms responsible for the plastic deformation of solids (Weertman, 1970; Stocker and Ashby, 1973). Some of the first of such estimates (Gordon, 1965) indicated that the lower mantle should have a much higher viscosity than the upper mantle, and the agreement of this result with McKenzie's (1966) calculated lower-mantle viscosity may be responsible for the widely held belief that the viscosity of the lower mantle is very high.

At low or moderate stresses and high temperature, plastic deformation of solids is accommodated by the diffusion of vacancies (Nabarro-Herring and Coble creep) and the climb of dislocations over obstacles (Weertman or power-law creep). In either case the rate-limiting process is the diffusion of vacancies either to grain boundaries or to dislocations. The effects of pressure and temperature on the rate of flow, or effective viscosity, enter primarily through their effects on thermally activated vacancy diffusion. For example, the viscosity that results from Nabarro-Herring creep is (Stocker and Ashby, 1973):

$$\eta = \frac{RTd^2}{14V_0D}$$

where R is the gas constant (8.314 J/g-mole °K), T the temperature, d the grain size, V_0 the molar volume and D is the diffusivity of vacancies given by:

$$D = D_0 \exp\left(-\frac{E^* + pV^*}{RT}\right)$$

where E^* is the energy for the formation and activation of a vacancy, p is the pressure, V^* is the activation volume and D_0 a constant. The important parameter for the effect of pressure is the activation volume V^* which may be considered to represent the volume increase of the sample that accompanies the creation and motion of a vacancy, or alternatively the logarithmic pressure derivative of the diffusivity. One expects V^* to be of the order of the atomic (or molar) volume of the vacant atom, and indeed for metals it is slightly less than this (cf. Lazarus and Nachtrieb, 1963), presumably owing to the relaxation of the lattice around the vacancy. Whether this is also the case for oxides is not known, but there is no reason to expect that they would behave much differently.

Since the activation volume represents an actual volume increase of the solid due to the creation and motion of a vacancy, it will itself depend on pressure. The volume of the solid will decrease with pressure owing to bulk compression; the volume of a vacancy should decrease proportionately at least as much as the atomic volume. Furthermore, since a vacancy results in a local 'soft spot' in the atomic lattice, it should be more compressible than the lattice itself. As a first approximation we may represent a vacancy as a spherical cavity in an elastic continuum. This approximation is suggested by the success of similar models in treating lattice defects (e.g. Narayan and Washburn, 1973), and its accuracy depends on the extent to which the distortion due to the defect is accommodated by elastic deformation of the surrounding lattice rather than by extremely short range interactions immediately around the defect.

To illustrate the possible effects of mantle pressures and temperatures on plastic flow we consider Nabarro Herring creep of MgO, for which some reliable data is available and which may be an important constituent of the lower mantle (Ringwood, 1975). The purpose of the calculation is to show that such flow mechanisms can result in a mantle with little variation of viscosity with depth, and it is not to be considered as a realistic calculation of the effective viscosity of the mantle. Nevertheless it may indicate the order of magnitude of possible bounds that can be placed on mantle flow properties. In order to estimate the effect of pressure on the activation volume, we treat it as a spherical cavity in a homogeneous continuum. The effective compressibility of such a cavity is (Timoshenko and Goodier, 1951):

$$\frac{1}{V}\frac{\partial V}{\partial p} = \frac{1}{K}\frac{3}{2}\left(\frac{1-\nu}{1-2\nu}\right)$$

For Poisson ratio $\nu = \frac{1}{4}$, the effective bulk modulus of the cavity, K_c, is then $\frac{4}{9}$ of the bulk modulus of the matrix, K. The size of the cavity at pressure can be found by integration:

$$\frac{V}{V_0} = \left(1 + \frac{K_c'}{K_c}\right)^{-1/K_c'}$$

where $K_c' = \partial K_c/\partial p$ is taken as $\frac{4}{9}$ of the corresponding value for the matrix.

Table I lists the parameters used. The activation energy and constant D_0 are from Narayan and Washburn (1973), and correspond to intrinsic volume diffusion of oxygen ions in MgO. The activation volume V^* is just the molar volume of MgO, and may be an overestimate if there is significant relaxation of the lattice around a vacancy. The bulk modulus K and its pressure derivative K' are appropriate for MgO. The pressure and temperature gradients taken for the mantle are rough figures only, but are realistic representative figures.

Table II shows the diffusivity of vacancies and effective viscosities for MgO under mantle conditions. Although the effect of pressure more than doubles the effective activation energy, this is more than compensated by the increase in temperature. Note also that pressure reduces the activation volume by a factor of two at the base of the mantle. Regardless of the absolute values of the effective viscosities (which depend on the grain size), Table II clearly shows that the effect of pressure will not necessarily increase the effective viscosity of the mantle with depth. This result was obtained using plausible values for parameters and temperatures in the mantle (3000°C at the core—mantle boundary). Since the effect of pressure

TABLE I

Creep parameters used

E^*	460 kJ/mole
V^*	11.25 cm^3/mole = 11.25 · 10^{-6} m^3/mole
D_0	1.37 · 10^{-6} m^2/sec
R	8.314 J/mole °K
ρ	3.584 Mg/m^3
dp/dz	0.333 kbar/km = 33.3 MPa/km
dT/dz	0.6°K/km
K	1600 kbar = 160 GPa
K'	4
ν	0.25

TABLE II

Creep properties at mantle conditions

Depth (km)	T (°K)	pV^*/E^*	V^*/V_0^*	D (m^2/s)	η (poise)*
500	1773	0.33	0.82	$1.1 \cdot 10^{-24}$	$2.4 \cdot 10^{24}$
1000	2073	0.58	0.71	$6.7 \cdot 10^{-25}$	$4.6 \cdot 10^{24}$
1500	2373	0.77	0.63	$1.5 \cdot 10^{-24}$	$2.4 \cdot 10^{24}$
2000	2673	0.94	0.58	$5.1 \cdot 10^{-24}$	$7.7 \cdot 10^{23}$
2500	2973	1.08	0.53	$2.1 \cdot 10^{-23}$	$2.1 \cdot 10^{23}$
3000	3273	1.21	0.49	$8.3 \cdot 10^{-23}$	$5.8 \cdot 10^{22}$

* Nabarro-Herring creep with grain size of 0.1 mm.

enters through the vacancy diffusivity, the conclusion should hold equally for dislocation climb controlled creep. Whether the calculation is realistic for silicate phases is not known, but the activation volume should not be very different. Since it is this that determines the effect of pressure, our conclusion may well apply to silicate phases as well. If the lower mantle is a mixture of magnesio-wüstite and stishovite, sufficient flow could possibly be accommodated by (Mg,Fe)O alone even if stishovite was much more resistant to flow. The effect of changes to denser phases at greater depths is not known; a denser phase may have a larger activation energy owing to its more closely packed atomic structure. On the other hand the activation volume may be smaller owing to the greater density and more lattice relaxation around a vacancy. These two effects may offset one another at high pressure, and the net result is unknown.

In any event, it is clear that one can not predict an increase in viscosity with depth in the mantle on the basis of micro-mechanisms of plastic flow; in fact one could predict a *decrease* in viscosity with depth by choosing appropriate, yet plausible, parameters.

MANTLE PHASE BOUNDARIES AND CHEMICAL HETEROGENEITY

The seismic discontinuities at 450 and 670 km depth have been identified with the transformation of olivine to a spinel structure and then to a denser phase (Anderson, 1967), most likely a mixture of the oxides magnesio-wüstite and stishovite (Ming and Bassett, 1975; Kumazawa et al., 1974). Further phase transformations may exist at greater depths (Ringwood, 1975), but none have yet been clearly identified with seismic structure. It has been argued that such phase transitions would present a barrier to convection across them (cf. Ringwood, 1972); nevertheless it has been shown that phase transitions can promote convection in some circumstances (cf. Schubert et al., 1975). Also Richter (1973b) has clearly shown that the olivine—spinel transition would not present a barrier to finite-amplitude convection, but would enhance it. Since this transition has a positive Clapeyron slope, the phase boundary will be elevated in a descending lithospheric slab relative to the surrounding mantle (Toksöz et al., 1973; Solomon and Paw U, 1975) and the resulting sinking force (or negative buoyancy) may well be the most easily identified of the forces responsible for mantle flow.

It has been thought that the disproportion of olivine spinel to mixed oxides would have distinctively negative Clapeyron slope (Ahrens and Syono, 1967) and that this phase boundary would be deeper in the cooler slab than the adjacent mantle, thus providing buoyancy resisting the further sinking of the slab. That this is not necessarily the case has been shown by Jackson et al., (1974), who demonstrated that the reaction may well have a negligible entropy change (and thus be independent of temperature) or even a distinctly positive Clapeyron slope if the olivine spinel has a large inverse character at high temperature. In this case the transition would be elevated in a sinking slab, and would promote further sinking of the slab similar to the olivine—spinel transition. Furthermore Richter (1973b) has shown that even with a negative Clapeyron slope, the transition would not inhibit convection. Thus there is no clear evidence that the phase boundary at 670 km depth presents a barrier to convection extending into the lower mantle, and it may even promote such convection.

Although the phase changes by themselves may not inhibit mantle-wide convection, it has been suggested that the lower mantle is enriched in iron relative to the upper mantle (Anderson et al., 1972; Press, 1972). Such chemical differences between the phases may present a barrier to convection (Richter and Johnson, 1974) and, in fact such differences would present direct evidence that large-scale mixing and homogenization of the mantle has not taken place. The evidence for chemical differences between the upper and lower mantle comes from interpreting seismic velocities and the density of the mantle in terms of the elastic properties of minerals and theoretical equations of state for extrapolating the properties to different pressures and temperatures. There are considerable uncertainties involved in the determination of small variations in the composition of the mantle by such methods, and it has been more recently shown (Davies, 1974; Watt et al., 1975) that such variations in iron content as have been claimed are beyond the resolution of the available data. In addition, the lack of definite knowledge of the crystal structure and phase assemblages of the major constituents of the lower mantle (Liebermann et al., 1976) indicate that extreme caution should be used in deducing chemical composition from the seismic properties of the lower mantle. Since at present either chemical homogeneity or iron enrichment in the lower mantle is allowed, but not required, by the data (Davies, 1974), there is no reason to rule out mantle-wide convection on the basis of chemical inhomogeneity of the mantle. In addition, any heterogeneities that are resolved in the future should be compared with heterogeneities in the upper mantle before assessing how strong a constraint they impose on the likelihood of lower-mantle convection.

EARTHQUAKES IN DESCENDING PLATES

Both the occurrence of earthquakes in descending lithospheric plates and the orientation of stress deduced from focal mechanisms have been taken as evidence that the plates do not penetrate deeper than 700 km (McKenzie et al., 1974; McKenzie and Weiss, 1975). This is not the only possibility however; the absence of earthquakes deeper than 700 km could be due to any one of: (1) the failure of slabs to penetrate deeper than 700 km; (2) the rise of temperature in the slab and the consequent reduction in creep strength of the material (McKenzie, 1969; Griggs, 1972); or (3) the change of phase of lithospheric material to a material with different flow properties. The latter two possibilities account for the cessation of seismicity, yet still permit the slab to sink deeper. That the slab will heat as it descends and eventually lose its mechanical integrity seems beyond question, and this may well account for the cessation of earthquakes. Nevertheless the rather sudden drop off in seismicity that occurs at around 700 km (Isacks and Molnar, 1971) in some regions suggests that the phase transition at that depth in the mantle may be responsible. If the transition is from predominately olivine to mixed oxides, there may be a marked reduction in the creep strength of the material, thus relaxing shear stresses before catastrophic failure takes place. Such a change in flow properties could be due to the introduction of a non silicate, magnesio-wüstite, as a major phase, which might have a lower creep strength than olivine spinel. In addition the change of phase could result in a much smaller grain size, which might be stabilized against grain growth by the presence of two mutually insoluble major phases in roughly equal amounts. The small grain size would enhance diffusional creep (Stocker and Ashby, 1973) and could dramatically reduce the effective viscosity of the material. Thus the absence of earthquakes could be due to the slab becoming less rigid, or penetrating a less rigid region, rather than running into a more rigid lower mantle.

The distribution of stresses in sinking slabs has also been taken as evidence that the slabs encounter increased mechanical resistance at 700 km depth. In particular, the stresses in the slabs can be characterized by the observation (Isacks and Molnar, 1971) that slabs that extend to 700 km exhibit down-dip compression for their entire length, whereas slabs that do not extend so deep exhibit down-dip extension above 300 km, and compression below 300 km if they extend that far. The stress distribution in a slab is determined by the distribution of forces on the slab. If the slab is sinking at a uniform velocity into the mantle, the body force due to the greater density of the slab must be exactly balanced by surface tractions due to the resistance of the mantle to penetration by the slab. These resisting forces can be roughly divided into shear tractions on the sides of the slab due to viscous drag, and normal tractions at the front edge of the slab due to the local increase in pressure necessary to force the fluid out of the way of the slab. In addition there may be extensional normal tractions at the top of the slab resisting the sinking of the slab. These result from a reduction of the normal hydrostatic stress due to the motion of the slab. If the pressure on the front edge is the predominant resisting force, then a slab sinking under its own weight should be under down-dip compression for its entire length, in the same way that a brick standing on end on a table is under vertical compression, owing to the body force distributed along the length of the slab being balanced by an equal and opposite force at the front edge of the slab. (This is contrary to the assertion by Isacks and Molnar (1971), that such a sinking slab should be under extension.) Thus down-dip compression of the slabs is consistent with sinking slabs resisted primarily by normal forces at the leading edge.

The extension of the slabs at depths less than 300 km indicates that the sinking of the slab under its own weight is resisted by shear tractions concentrated near the top of the slab or by extensional tractions from the shallower parts of the slab, or else the slab is being "pulled down" by the part of the slab that is deeper. The last alternative could result from a concentration of the sinking force in the region where the olivine—spinel phase change is elevated in the slab, namely at 300—400 km depth. In this case the lower part of the slab would tend to sink under its own weight more rapidly than the upper part. This would lead to extensional deviatoric stresses in the upper part of the slab owing to the reduction in the normal stress parallel to the slab. Note that such extensional stresses do not require that tensile stresses be supported by the slab; since the whole slab is under considerable hydrostatic stress, a reduction in the down-dip component of the compressive stress leads to a *deviatoric* stress state that is extensional even though the stress is everywhere compressive. (It is in this way that Bridgeman measured the tensile strength of materials at very high pressures.) Thus compressive stresses below 300 km, and extensional stresses above 300 km are entirely what one would expect for a slab sinking under its own weight, with a large fraction of the excess density of the slab concentrated between 300 and 400 km owing to the elevation of the olivine—spinel phase boundary in the cooler slab. There is no need to postulate that the slab is penetrating more resistant material at depth in order to account for the compression of the lower part of the slab; this can be accounted for by just the resistance to penetration of a mantle with relatively uniform viscosity.

The observation that slabs that reach 700 km depth exhibit compression along their entire length seems to indicate that the slabs encounter increased resistance to sinking at this depth. Such resistance could result from either an increased viscosity below 700 km, or from an increased buoyancy of material that reached that depth; both alternatives were suggested by Isacks and Molnar (1971). We note that a modest increase in viscosity (say by a factor of 2) might well be sufficient to account for the increased resistance to penetration; this would not be enough to prevent the slabs from penetrating the lower mantle, nor would it inhibit convection in the lower mantle.

Alternatively, the increased resistance could be due to buoyancy due to the downwarping of the spinel—mixed oxides phase boundary to a depth of

700 km. This is suggested by the presence of the earthquakes down to 700 km, and absence immediately below that. The downwarping could result from either a negative Clapeyron slope for the transition, or from the failure of the transition to occur at the equilibrium depth, with olivine spinel existing metastably between 670 and 700 km where it finally all converts to mixed oxides. Either would account for an increase in the force at the tip of the slab resisting the downward sinking of the slab. This force would increase that already present, and result in the slab going into compression for its entire length; the increased resisting force should also slow the descent of the slab. But it is not necessary to postulate a rigid obstruction that stops the slab sinking in order to account for the stress state of the slab. In fact Forsyth and Uyeda (1975) have suggested that compression along the whole length of a slab could be the result of slabs being constrained by being attached to the same plate; thus even increased resistance to sinking at 700 km is not required.

In summary we see that the distribution of earthquakes and stresses in sinking slabs do not necessarily indicate large increase in viscosity at 700 km depth, nor do they indicate that sinking slabs stop sinking at that depth. One can reasonably account qualitatively for the observations by a simple model of a sinking slab propelled to a large extent by the sinking force due to an elevated olivine—spinel phase boundary, and resisted primarily by hydrodynamic resistance at the leading edge of the slab, with this resistance perhaps increasing somewhat at 700 km depth. Such a picture suggests convection extending into the lower mantle.

LOWER-MANTLE HETEROGENEITY

We have seen that there is little evidence that convection cannot extend into the lower mantle. There is at the same time accumulating evidence of significant heterogeneities in both the upper and lower mantle that may be directly related to temperature differences and flow, and which indicate that convection extends into the lower mantle. Some of this evidence has recently been reviewed by Jordan (1975); he notes: (1) that seismic velocity differences between continents and oceans appear to extend to depths greater than 400 km, which suggests that the region that translates coherently in the course of horizontal plate motions is at least that thick; (2) substantial ScS-S differential travel-time residuals indicate significant seismic heterogeneity in the lower mantle; and (3) marked azimuthal dependence of travel-time residuals may be associated with velocity anomalies below 800 km depth in the mantle, which may be associated with subduction zones. Seismic anomalies below 700 km beneath subduction zones have also been noticed by Engdahl (1975).

Large-scale seismic velocity anomalies in the lower mantle have been found by Dziewonski et al., (1976); moreover the anomalies detected were shown to correlate remarkably well with the long-wavelength components of the gravity field. The anomalies could be interpreted in terms of compositional variations, or in terms of temperature variations and distortions of the earth's surface and core—mantle boundary associated with large-scale flow. Either alternative suggests chemical or thermal convection in the lower mantle. Seismic anomalies in the lower mantle should come as no great surprise. The existence of substantial gravity anomalies of low spherical harmonic degree indicates that there are density anomalies in the mantle, and the lower mantle is a likely location for those that contribute to the low-degree components of the gravity field (Dziewonski et al., 1976). Associated with the density anomalies are deviatoric stresses of the order of tens of bars (Kaula, 1963); these must be supported either by the finite strength of the material or by dynamical effects of fluid flow. The mantle is unlikely to have a finite strength sufficient to support the stresses (Weertman, 1970), which leaves only fluid flow, which if not maintained, would quickly decay, relaxing the stresses and reducing the gravity anomalies. Even the estimate of a high-viscosity lower mantle of McKenzie (1966) (which precluded convection in the lower mantle) was based on an estimate of $\sim 10^7$ years for the time for the C_2^0 term of the nonhydrostatic gravity field to decay. Other, equally large, components of the gravity field, such as that associated with the ellipticity of the equator, would also decay in 10^7 years. Thus their existence, even with a high-viscosity lower mantle, requires some active process in order to maintain them over geologic time. Since the evidence indicates that the lower mantle has a relatively low viscosity, the requirement for fluid motions in order to support the nonhydrostatic stresses becomes even more severe, and the most likely cause of the flow would be large-scale thermal convection.

Jordan (1975) has mentioned that mantle-wide convection poses a geochemical problem, however. Since mantle-wide convection would bring to the surface material that was once near the core—mantle boundary, one would expect that the oxidation state and nickel content of surface rocks would reflect equilibrium with the metallic, nickel-rich core of the earth. But that would be the case only if a large enough volume of the lower mantle could equilibrate with the core, and such equilibration would require the diffusion of chemical species in the solid-mantle material. The magnitude of diffusion constants for solid-state diffusion is such that one would expect only local equilibrium in a thin boundary layer at the base of the mantle. To estimate its thickness, take the value for Fe—Mg interdiffusion in olivine (Buening and Buseck, 1973) extrapolated to 3000°C : $D \approx 10^{-6}$ cm^2/s; this value must certainly be an upper limit (by several orders of magnitude) for lower-mantle material. Using a residence time at the core—mantle boundary of 500 million years, one calculates that the thickness of a diffusional boundary layer $l = (Dt)^{1/2}$ is about one kilometer. This is considerably thinner than a similar boundary layer for thermal conduction. Thus chemical equilibrium will be achieved only over extremely small distances, and the fraction of mantle material that can have equilibrated with the core is very small indeed. Consequently, it appears that the disequilibrium between the mantle and the core is consistent with mantle-wide convection.

DISCUSSION

The evidence we have reviewed indicated that neither the viscosity of the lower mantle, nor phase changes, nor the state of sinking slabs preclude convection in the lower mantle. We must therefore conclude that lower-mantle convection is as likely as upper-mantle convection. The heat sources for deep mantle convection could come from distributed radio-active heat sources, but an important heat source must be the earth's core.

Gubbins (1976) has placed an absolute lower bound on the heat flux out of the core from consideration of the minimum ohmic dissipation involved with the generation of the magnetic field. He finds that the heat flux is at least $2 \cdot 10^{10}$ watts, and most probably is considerably greater than this. There is little difficulty in accounting for this heat flow from the core. It could arise from radio-active ^{40}K in the core (Lewis, 1971) which could account for as much as one third of the observed surface heat flux; even a much lower concentration of potassium would still produce a substantial heat flux into the lower mantle (Verhoogen, 1973). Alternatively, the heat could come from the gradual cooling and freezing of the inner core (Verhoogen, 1961), which could account for $\sim 3 \cdot 10^{12}$ watts (10% of the surface flux) if the inner core froze over $3 \cdot 10^9$ years (Stacey, 1969 p. 264). This estimate assumes that the core has cooled by 300°C; if the mantle cooled proportionally, then its stored heat would also have to be transported to the surface by convection. Consequently a substantial fraction of the present surface heat flux could be due to the earth's "primordial" heat. In addition, if the solid inner core has a different composition than the liquid, compositional differences in the liquid will arise as the inner core freezes out; these may give rise to instabilities that promote flow which will involve dissipation of energy. The heat available from this mechanism may be comparable to that from the latent heat of freezing (D. Gubbins, personal communication, 1976; Braginsky, 1964).

If convection extends into the lower mantle, one might expect that its horizontal scale would be of the order of the depth, namely ~3000 km. If a substantial amount of heat is supplied by the core, the upwelling flow may be more localized, and only a few hundred kilometers across, as suggested by the numerical experiments of McKenzie et al. (1974). The flow in the upper mantle could have two different scales, as suggested by Richter and Parsons (1975) and McKenzie and Weiss (1975), with the difference that the large-scale flow would extend into the lower mantle rather than be confined to the upper mantle as they assumed. The large-scale flow might be expected to change more slowly with time than the small-scale flow, owing to its greater turnover time, and may be associated with volcanic hot spots (Morgan, 1971, 1972) (although not necessarily by forming concentrated jets of hot material reaching the surface).

If the mantle has relatively uniform flow properties and is convecting at all depths, then it does not seem correct to regard the asthenosphere as a weak plastic layer in the upper mantle. Rather the entire mantle constitutes the asthenosphere, and there is no need for a more rigid "mesosphere" beneath the asthenosphere. The identification of the asthenosphere with the seismic low-velocity zone may be misleading, since different mechanisms account for long-term creep and shorter anelastic relaxation. The low-velocity zone, LVZ is thought to be due to partial melting (Anderson and Sammis, 1970), although other interpretations exist (Guegen and Mercier, 1973). A very small amount of partial melt can account for the observed velocity reduction and high attenuation. The effect of a partial melt on the plastic flow properties is not precisely known; but if it were distributed as thin films along grain boundaries it would enhance grain boundary diffusion and deformation by Coble creep (Stocker and Ashby, 1973). The magnitude of this effect depends on the diffusivity of the solid phase in the fluid, and is not accurately known. But it has been shown that the presence of a fluid phase itself will not enhance steady-state creep rates over that of the solid grains (Auten et al., 1974). Thus the boundary between the mechanically rigid lithosphere and the plastic asthenosphere may bear no intrinsic relation to the top of the seismic low-velocity zone, although they may coincide (where the LVZ exists) by chance.

With large-scale flow throughout the mantle, the interaction between lithospheric plates and the flow may by quite complex, and the assignment of the forces driving the plates to specific mechanisms difficult. For example Forsyth and Uyeda (1975) and Solomon and Sleep (1974) have analyzed the forces acting on the plates by assuming (among other things) a viscous drag on the bottom of plates that is proportional to velocity. But this implicitly assumes that the plates are sliding over a relatively thin viscous layer over a relatively stationary lower mantle. It does not necessarily take into account the interaction of one smaller plate with the flow associated with a larger, adjacent plate for example, or with variations in the flow at depth in the mantle.

Forsyth and Uyeda (1975) found that the drag under continental plates was about eight times that under oceanic plates, which is consistent with Jordan's (1975) observations of substantial differences between oceans and continents at depth. Such observations suggest substantial lateral variations in the flow properties of the mantle, most probably related to temperature differences; such lateral differences may be more important than vertical differences, especially in the upper mantle. Thus the flow associated with continental plate motions may extend deeper into the mantle than that associated with oceanic plates, owing to the latter overlying hotter, more plastic mantle material. Such differences would indicate that the analysis of the forces on plates may have to consider the variations in the flow at depth in the mantle in order to determine the viscous drag (or pull) on each plate.

CONCLUSIONS

Direct estimates of the effective viscosity of the mantle, based on geophysical observations, indicate that the viscosity is relatively uniform with depth; there is no evidence for a large increase of viscosity in the lower mantle that cannot be explained equally well by a model with no such increase, while there is evidence that requires no large increase in viscosity, namely the absence of large-scale gravity anomalies correlated with deglaciation. A similar conclusion comes from consideration of plastic flow mechanisms in solids; the effect of pressure, which increases creep resistance, may be offset by the increase in temperature with depth in the mantle, so that the effective viscosity remains fairly constant with depth in the mantle. Such behaviour is consistent with recent models of the earth's thermal history, as well as with plausible temperatures in the lower mantle. Previous estimates of a large increase in viscosity in the mantle may have overestimated the activation volume for vacancy formation and diffusion.

The seismic evidence for the stress state of descending lithospheric slabs indicates that the slabs may encounter increased resistance at 700 km depth; this may come from a slight increase in flow resistance or be due to the spinel—post-spinel phase change. The fact that such slabs do not stop or significantly slow down indicates that the resistance is not sufficient to prevent material flowing into the lower mantle. Since mantle phase changes will also not necessarily inhibit convection, it is unlikely that convection will be confined to the upper mantle, but will be mantle-wide.

Such a convection pattern is also suggested by heterogeneities in the lower mantle. The more precise determination of these may permit more firm conclusions about the pattern of mantle flow, and may shed considerable light on the distribution of heat sources and the thermal state of the earth's core.

The asthenosphere, where significant flow can take place, is not a thin layer in the upper mantle, but includes the whole mantle. Furthermore the asthenosphere should not be identified with the seismic low-velocity zone, since it extends deeper, and quite different physical mechanisms probably account for the properties of each region. This suggests that there is no weak decoupling layer under the plates, although lateral variations of flow properties may produce local decoupling or unusually strong coupling to the lower mantle in some regions, such as oceanic ridges or under continents.

The scale of mantle-wide convection may preclude a simple analysis of the forces driving plate motions based on a model of plates strongly decoupled from the flow beneath. Rather, an analysis that takes into account the flow at depth may be required, and should be investigated.

ACKNOWLEDGEMENTS

This work was partially supported by National Science Foundation grant NSF GP 34723. I thank D. Gubbins, K. Lambeck and D.P. McKenzie for helpful comments, and gratefully acknowledge the hospitality of the Department of Geodesy and Geophysics, Cambridge University, where much of this paper was written.

REFERENCES

Ahrens, T.J. and Syono, Y., 1967. Calculated mineral reactions in the earth's mantle. J. Geophys. Res., 72: 4181.

Anderson, D.L., 1966. Earth's viscosity. Science, 151: 321.

Anderson, D.L., 1967. Phase changes in the upper mantle. Science, 157: 1165.

Anderson, D.L. and Sammis, C., 1970. Partial melting in the upper mantle. Phys. Earth. Planet. Inter., 3: 50—51.

Anderson, D.L., Sammis, C. and Jordan, T.H., 1972. Composition of the mantle and core. In: E.C. Robertson (editor), The Nature of the Solid Earth. McGraw-Hill, New York, pp. 41—66.

Auten, T.A., Gordon, R.B. and Stocker, R.L., 1974. *Q* and mantle creep. Nature, 250: 317—318.

Braginsky, S.I., 1964. Kinematic models of the earth's hydromagnetic dynamo. Geomagn. Aeron., 4: 572—583 (Engl. transl.).

Buening, B.K. and Buseck, P.R.; 1973. Fe—Mg Lattice diffusion in olivine. J. Geophys. Res., 78: 6852—6862.

Cathles, L.M., 1975. The Viscosity of the Earth's Mantle. Princeton Univ. Press. 386 pp.

Davies, G.F., 1974. Limits on the constitution of the lower mantle. Geophys. J.R. Astron. Soc., 38: 479—504.

Dicke, R.J., 1969. Average acceleration of the Earth's rotation and the viscosity of the deep mantle. J. Geophys. Res., 74: 5895—5902.

Dziewonski, A.M., Hager, B. and O'Connell, R.J., 1976. Large-scale heterogeneities in the lower mantle. J. Geophys. Res., in press.

Engdahl, E.R., 1975. Effects of plate structure and dilatancy on relative teleseismic P-wave residuals. Geophys. Res. Lett., 2: 420—422.

Farrell, W.E. and Clark, J.A., 1976. On post-glacial sea level. Geophys. J.R. Astron. Soc., in press.

Forsyth, D. and Uyeda, S., 1975. On the relative importance of the driving forces of plate motion. Geophys. J.R. Astron. Soc., 43: 163—200.

Goldreich, P. and Toomre, A., 1969. Some remarks on polar wandering. J. Geophys. Res., 74: 2555—2567.

Gordon, R.B., 1965. Diffusion creep in the Earth's mantle. J. Geophys. Res., 70: 2413.

Griggs, D.T., 1972. The sinking lithosphere and the focal mechanism of deep earthquakes. In: E.C. Robertson (editor), The Nature of the Solid Earth. McGraw-Hill, New York, pp. 361—384.

Gubbins, D., 1976. Observational constraints on the generation process of the earth's magnetic field. Geophys. J.R. Astron. Soc., in press.

Gueguen, Y. and Mercier, J.M., 1973. High attenuation and the low-velocity zone. Phys. Earth Planet. Inter., 7: 39—46.

Isacks, B.L. and Molnar, P., 1971. Distribution of stresses in the descending lithosphere from a global survey of focal mechanism solutions of mantle earthquakes. Rev. Geophys., 9: 103—174.

Jackson, I.N.S., Liebermann, R.C. and Ringwood, A.E., 1974. Disproportionation of spinels to mixed oxides: significance of cation configuration and implications for the mantle. Earth Planet. Sci. Lett., 24: 203—208.

Jordan, T.H., 1975. Lateral heterogeneity and mantle dynamics. Nature, 257: 745—750.

Kaula, W.M., 1963. Elastic models of the mantle corresponding to variations in the external gravity field. J. Geophys. Res., 68: 4967—4978.

Kumazawa, M., Sawamoto, H., Ohtaki, E. and Masaki, K., 1974. Post-spinel phase of forsterite and evolution of the Earth's mantle. Nature, 247: 356—358.
Lazarus, D. and Nachtrieb, N.H., 1963. Effect of high pressure on diffusion. In: W. Paul and D.M. Warschauer (editors); Solids Under Pressure. McGraw-Hill, New York.
Lewis, J.S., 1971. Consequences of the presence of sulphur in the core of the earth. Earth Planet. Sci. Lett., 11: 130—134.
Liebermann, R.C., Ringwood, A.E. and Major,A., 1976. Elasticity of stishovite, Preprint.
MacDonald, G.J.F., 1963. The deep structure of the continents. Rev. Geophys., 1: 587—665.
McConnell, R.K., 1965. Isostatic adjustment in a layered Earth. J. Geophys. Res., 70: 5171—5188.
McConnell, R.K., 1968. Viscosity of the mantle from relaxation time spectra of isostatic adjustment. J. Geophys. Res., 73: 7089—7105.
McKenzie, D.P., 1966. The viscosity of the lower mantle. J. Geophys. Res., 71: 3995—4010.
McKenzie, D.P., 1969. Speculations on the causes and consequences of plate motions. Geophys. J.R. Astron. Soc., 18: 1—32.
McKenzie, D.P. and Weiss, N., 1975. Speculations on the thermal and tectonic history of the earth. Geophys. J.R. Astron. Soc., 42: 131—174.
McKenzie, D.P., Roberts, J.M. and Weiss, N.O., 1974. Convection in the earth's mantle: towards a numerical simulation. J. Fluid Mech., 62: 467—538.
Ming, L.C. and Bassett, W.A., 1975. The post-spinel phases in the Mg_2SiO_4—Fe_2SiO_4 system. Science, 187: 66—68.
Morgan, W.J., 1971. Convection plumes in the lower mantle. Nature, 230: 42—43.
Morgan, W.J., 1972. Deep mantle convection plumes and plate motions. Am. Assoc. Pet. Geol. Bull., 56: 203—213.
Narayan, J. and Washburn, J., 1973. Self diffusion in magnesium oxide. Acta Metall., 21: 533—538.
O'Connell, R.J., 1971. Pleistocene glaciation and the viscosity of the lower mantle. Geophys. J.R. Astron. Soc., 23: 299—327.
Parsons, B.D., 1972. Changes in the Earth's Shape. Unpubl. Ph.D. Dissertation, Downing College, Cambridge University.
Peltier, W.R., 1974. The impulse response of a Maxwell earth. Rev. Geophys. Space Phys., 12: 649—670.
Press, F., 1972. The Earth's interior as inferred from a family of models. In: E.C. Robertson (editor), The Nature of the Solid Earth. McGraw-Hill, New York, pp. 147—171.
Richter, F.M., 1973a. Dynamical models for sea-floor spreading. Rev. Geophys. Space Phys., 11: 223—287.
Richter, F.M., 1973b. Finite-amplitude convection through a phase boundary. Geophys. J.R. Astron. Soc., 35: 265—276.
Richter, F.M. and Johnson, C.E., 1974. Stability of a chemically layered mantle. J. Geophys. Res., 79: 1635—1639.
Richter, F.M. and Parsons, B., 1975. On the interaction of two scales of convection in the mantle. J. Geophys. Res., 80: 2529.
Ringwood, A.E., 1972. Phase transformations and mantle dynamics. Earth Planet. Sci. Lett., 14: 233—241.
Ringwood, A.E., 1975. Composition and Petrology of the Earth's Mantle. McGraw-Hill, New York.
Schubert, G., Yuen, D.A. and Turcotte, D.L., 1975. Role of phase transitions in a dynamic mantle. Geophys. J.R. Astron. Soc., 42: 705—735.
Solomon, S.C. and Paw U, K.T. 1975. Elevation of the olivine—spinel transition in subducted lithosphere: seismic evidence. Phys. Earth Planet. Inter., 11: 97—108.

Solomon, S.C. and Sleep, N.H., 1974. Some simple physical models for absolute plate motions. J. Geophys. Res., 79: 2557—2567.

Stacey, F.D., 1969. Physics of the Earth. Wiley, New York.

Stocker, R.L. and Ashby, M.F., 1973. On the rheology of the upper mantle. Rev. Geophys. Space Phys., 11: 391—426.

Timoshenko, S. and Goodier, J.N., 1951. Theory of Elasticity, McGraw-Hill, New York.

Toksöz, M.N., Sleep, N.H. and Smith, A.T., 1973. Evolution of the downgoing lithosphere and the mechanisms of deep-focus earthquakes. Geophys. J.R. Astron. Soc., 35: 285—310.

Tozer, D.C., 1972. The present thermal state of the terrestrial planets. Phys. Earth Planet. Inter., 6: 182.

Turcotte, D.L. and Oxburgh, E.R., 1972. Mantle convection and the new global tectonics, In: Annual Review of Fluid Mechanics, 4: 33—68. Annual Reviews inc., Palo Alto. Calif.

Verhoogen, J., 1961. Heat balance of the earth's core. Geophys. J.R. Astron. Soc., 4: 276—281.

Verhoogen, J., 1973. Thermal regime of the earth's core. Phys. Earth Planet. Inter., 7: 47—58.

Walcott, R.I., 1973. Structure of the earth from glacio-isostatic rebound. Ann. Rev. Earth Planet. Sci., 1: 15—37.

Watt, J.P., Shankland, T.J. and Mao, N., 1975. Uniformity of mantle composition. Geology, 3: 91—94.

Weertman, J., 1970. The creep strength of the Earth's mantle. Rev. Geophys., 8: 145—168.