The Cathles Parameter (Ct): A Geodynamic Definition of the Asthenosphere and Implications for the Nature of Plate Tectonics

Mark A. Richards1 and Adrian Lenardic2

1Department of Earth and Planetary Science, University of California, Berkeley, CA, USA, 2Department of Earth, Environmental, and Planetary Sciences, Rice University, Houston, TX, USA

Abstract A weak asthenosphere, or low-viscosity zone (LVZ), underlying Earth’s lithosphere has played an important role in interpreting isostasy, postglacial rebound (PGR), and the seismic LVZ, as well as proposed mechanisms for continental drift, plate tectonics, and postseismic relaxation. Consideration of the resolving power of PGR, postseismic relaxation, and geoid modeling studies suggests a sublithospheric LVZ perhaps ~100–200 km thick with a viscosity contrast of ~100–1,000. Ab initio numerical models of plate-like boundary layer motions in mantle convection also suggest a key role for the LVZ. Paradoxically, a thinner LVZ with a strong viscosity contrast is most effective in promoting plate-like surface motions. These numerical results are explained in terms of the reduction in horizontal shear dissipation due to an LVZ, and a simple scaling theory leads to somewhat nonintuitive model predictions. For example, an LVZ causes stress magnification at the base of the lithosphere, enhancing plate boundary formation. Also, flow within the LVZ may be driven by the plates (Couette flow), or pressure-driven from within the mantle (Poiseuille flow), depending upon the degree to which plates locally inhibit or drive underlying mantle convection. For studies of the long-wavelength geoid, PGR, and mantle convection, a simple dimensionless parameter controls the effect of the LVZ. This “Cathles parameter” is given by \( \text{Ct} = \frac{\eta^* D}{\lambda^3} \), where \( \eta^* \) is the viscosity contrast, \( D \) is the thickness of the LVZ, and \( \lambda \) is the flow wavelength, emphasizing the tightly coupled trade-off between LVZ thickness and viscosity contrast.

Plain Language Summary The Earth’s global system of tectonic plates move over a thin, weak channel (“low-viscosity zone”) in the mantle immediately underlying the plates. This weak channel is commonly referred to as the asthenosphere, and its presence accounts for a number of important Earth observations, including isostasy (e.g., support for the uplift of large mountain ranges), the shape of the Earth’s gravity field, the response of the Earth’s surface to the removal of large ice sheets (“postglacial rebound”), and the relationship between plate motions and underlying thermal convection in the mantle. In this paper, we show that these phenomena can be understood in terms of a single unifying parameter consisting of the viscosity contrast between the asthenosphere and the underlying mantle, and the cube of the thickness of the asthenosphere. We propose to call this the “Cathles parameter” in recognition of the author who first recognized its importance in postglacial rebound studies.

1. Introduction

Plate tectonics, along with quantum mechanics and general relativity, is one the foundational scientific theories developed during the twentieth century. As formulated, it is a kinematic theory (Cox, 1973; Morgan, 1968), and developing a dynamic theory of plate tectonics remains an intensely active area of research (e.g., Bercovici et al., 2000, 2015; Coltice et al., 2017).

A first step toward a dynamic theory of plate tectonics was a remarkably straightforward plate cooling model that explains, to first order, the relationship between bathymetric and heat flow observations and oceanic plate age (Parsons & McKenzie, 1978; Parsons & Sclater, 1977; Stein & Stein, 1992; Turcotte & Oxburgh, 1967). The stunning success of this model, which explains a broad range of observations and provides a causal link between them, is direct evidence that the oceanic lithosphere is the active upper thermal boundary layer of mantle convection. The ability to treat the structure of oceanic lithosphere within a thermal convection framework directly linked plate tectonics to the energy sources that drive plate motion—a key step toward a dynamic theory. There is no longer serious debate that plate tectonics is the surface...
manifestation of convection within the Earth's mantle. What is less remarked upon is the fact that the cooling plate model, in combination with observational data, also provided a more precise definition of “lithosphere” than had existed prior to its development.

The concept of the lithosphere extends back well before the development of plate tectonics, including early notions of isostasy (Barrell, 1914). However, beyond the idea that the lithosphere is a layer of long-term strength, defining the lithosphere in a more precise manner (e.g., material properties and thickness) had to wait until the plate tectonics revolution. Connecting gravity, topography, and heat flow observations to a dynamic model of lithosphere formation allowed the data to directly constrain the thermal structure of oceanic lithosphere (Stein & Stein, 1992) and provided a fairly rigorous definition of the thickness of the “thermal lithosphere.” The more traditional concept of the lithosphere as a region of strength, the “mechanical lithosphere,” was also constrained with the addition of rock deformation experiments that determined the temperature dependence of mantle strength (e.g., Kohlstedt et al., 1995). The critical notion we wish to stress here is that observational data connected to a dynamic model allowed the definition of oceanic lithosphere to advance from a conceptual one to a more quantitative one, incorporating not only the material properties but also the effective thickness of the lithosphere.

With a first-order dynamic theory of the lithosphere in place, a logical next step would be to develop a more precise dynamic understanding of plate margins. This remains an active area of research in geodynamics (e.g., Bercovici & Ricard, 2014). However, an equally valuable step would be to provide a data-driven definition for the more fluid mantle layer below plates, the “asthenosphere,” and to elucidate its dynamic role in plate tectonics. The latter goal is the focus of this paper.

We begin by considering the ability of two branches of geodynamics—postglacial rebound (PGR) and geoid modeling—to independently resolve the radial viscosity structure of the Earth’s sublithospheric mantle. These are both established “classical” problems in geodynamics, which provide the two most direct constraints on mantle viscosity structure. We demonstrate that neither PGR data nor combined geoid/seismic tomography data are able to independently resolve asthenosphere thickness and viscosity. Instead, the asthenosphere viscosity contrast and its thickness cubed may be combined as a single, critical parameter that determines to first order the behavior of both PGR models and dynamic geoid models. We also show that the same parameter may be used to characterize, to first order, the response of mantle convection models to the presence of a low-viscosity zone (LVZ), or asthenosphere (e.g., Busse et al., 2006; Hoink et al., 2011; Lenardic et al., 2006). We review the results of these models, showing how this particular combination of asthenosphere properties can affect mantle flow and plate driving forces and, in turn, the length scale and stability of the plate tectonic regime of mantle convection. Moreover, we speculate that the three essential ingredients for a dynamic theory for plate tectonics on Earth must account for (1) the existence of a strong (cold) upper boundary layer, or lithosphere; (2) failure of the lithosphere and consequent formation of plate boundaries; and (3) the existence of a thin, low-viscosity asthenosphere beneath the lithosphere. Regarding the latter, we propose that the dynamic effects of the asthenosphere are best understood by considering the asthenosphere’s effective thickness and viscosity in combination. We further propose that the particular combination be termed the Cathles parameter (Ct) in honor of the person who first foresaw its potential importance (Cathles, 1975) in PGR studies.

The form of this paper is that of a review, although it is perhaps better described as a “perspective” paper, as it draws heavily upon our own previously published work. The reason we feel this perspective paper is needed is that although many researchers are aware that there is often a modeling trade-off between the thickness and viscosity contrast of the asthenosphere, it has not been previously shown that several distinct branches of geodynamic modeling involving the asthenosphere can be unified via the Cathles parameter, which is the main focus of this paper.

2. Geodynamic Constraints

The nature of the asthenosphere relates directly to the viscosity structure of the Earth; however, wide-ranging estimates for mantle viscosity structure have been (and continue to be) published, so an Earth scientist not specializing in studies of mantle viscosity structure might well ask: If this is such a central question in mantle dynamics, then why, given the wide variety of data that have been brought to bear, has it not been satisfactorily answered? In this section we address this question with emphasis upon an essential modeling
ambiguity that arises in the two most prominent classes of geodynamic constraints on mantle viscosity structure, namely, PGR and geoid studies. As it turns out, the fluid-mechanical explanation for this ambiguity is also key to understanding the effects of a LVZ on mantle convection and plate tectonics (section 3) and therefore essential to understanding the role of the LVZ in Earth evolution and dynamics.

2.1. Postglacial Rebound

Large-scale Ice Age glaciation on continents, that is, the Fennoscandian and Laurentide ice sheets, depress the Earth’s surface. When these ice sheets melt and retreat, the Earth’s surface rebounds on timescales of ~10,000–100,000 years. Modeling this rebound is the classical method for estimating mantle viscosity (e.g., Cathles, 1975; Haskell, 1935). The physics of this PGR problem (also known as “glacio-isostatic rebound”) is straightforward, with the rate of rebound proportional to the remaining un-rebounded (negative) topography, resulting in an exponentially decaying signal for the simplest models, or multiple exponentially decaying components for more complicated models.

Resolving radial mantle viscosity structure, including an LVZ, from PGR-type data suffers from the following: (1) uncertainties in the data (ice loading history, residual sea level [or RSL], measurements, geodetic data, etc.); (2) uncertainties regarding appropriate rheologies (linear versus nonlinear viscous creep, elasticity, plastic failure, etc.); (3) geographical variability of the mechanical properties of the lithosphere and mantle (continental versus oceanic mantle, lateral viscosity variations due to temperature differences); and (4) inherent modeling trade-offs. All of these effects have been studied to varying degrees. However, even under the most optimistic and simplifying assumptions (limited data uncertainties, laterally constant, linear-viscous rheology, no elastic lithosphere), modeling trade-offs of a physical (fluid mechanical) nature probably account for much of the apparent disagreement regarding the existence and strength of the LVZ in the PGR literature.

Any multiparameter, viscous relaxation forward or inverse modeling exercise involves ambiguities in resolution, but the most important modeling trade-off is that between the thickness and viscosity contrast of a low-viscosity channel (LVZ) beneath the lithosphere. Simply stated: The thicker the channel, the smaller the viscosity contrast necessary to explain a given data set, and vice versa. Recognition of this trade-off was implicit in some of the earliest speculations and studies regarding PGR, going back at least to Daly (1934) and Van Bemmelen and Berlage (1935). However, Cathles (1975), in his seminal work on The Viscosity of the Earth’s Mantle, gives a succinct and quantitative explanation for this trade-off as follows.

Consider a vertical sinusoidal load of horizontal wavelength \( \lambda \) (wave number \( k = 2\pi/\lambda \)) imposed at the free surface of a viscous half-space of Newtonian uniform viscosity \( \eta \) and density \( \rho \). This load will relax exponentially toward isostatic equilibrium with a characteristic time constant given by

\[
\tau_{\text{uniform}} = (4\pi/\lambda) \cdot (\eta/\rho g) = (2k) \cdot (\eta/\rho g)
\]

(1)

where \( g \) is the acceleration of gravity. Note that this relaxation time is independent of the load amplitude, that is, the thickness of the ice sheet. This “textbook” result (see Turcotte & Schubert, 1982) is a primary point of departure for most analyses of PGR-related data.

We now consider the equally interesting limiting case in which a viscous channel (LVZ) of thickness \( D \) overlies a rigid half-space (deep mantle) and is again loaded by a sinusoidal vertical load. In the long-wavelength limit (thin channel) in which \( \lambda \gg 2\pi D \) (or \( KD \ll 1 \)), straightforward analysis yields (Cathles, 1975)

\[
\tau_{\text{LVZ}} = \left(3/k^2 D^3\right) \cdot (\eta/\rho g) = \left(3\lambda^2/4\pi^2 D^3\right) \cdot (\eta/\rho g) = \tau_{\text{uniform}} \left(3/\pi\right) \cdot \left(\eta/\rho g\right) \cdot (\lambda/D)^3
\]

(2)

The factor \( (\eta/\rho g) \) is common to both the half-space and thin-channel cases, confirming the obvious: that the relaxation time is proportional to the viscosity and inversely proportional to both the acceleration of gravity and the density contrast across the free surface. However, comparison of equations (1) and (2) reveals two key additional insights: First, the viscous relaxation time decreases linearly with increasing load wavelength in the half-space (uniform mantle viscosity) case but increases with the square of the load wavelength in the thin-channel (LVZ) case. The reason for the latter case is that the horizontal shear stress in the thin channel increases rapidly as the flow wavelength \( \lambda \) becomes large relative to the channel thickness \( D \). This effect will be shown (section 3) to also largely explain how the LVZ affects mantle convection and plate tectonics.
Figure 1. Relaxation decay time as a function of spherical harmonic degree (relaxation spectrum) for three Earth models, each with a 100-km-thick lithosphere and a lower-mantle viscosity of $2.2 \times 10^{21}$ Pa·s. The solid line has a uniform mantle viscosity ($2.2 \times 10^{21}$ Pa·s). The dashed lines show models with a low-viscosity zone (LVZ) beneath the lithosphere. The bottom (long dashed) line is for a model with a thinner and weaker LVZ (75 km, $7.9 \times 10^{17}$ Pa·s), and the middle (short dashed) line shows the spectrum of a thicker and stronger LVZ (375 km, $1.3 \times 10^{20}$ Pa·s). The gray regions show the measured relaxation times (and uncertainties) of two sites on Hudson Bay. (Figure is from Paulson & Richards, 2009.)

The second insight is that independent of the loading wavelength, there is an ensemble, or family, of thin-channel (LVZ) models such that equal values of $\eta/D^3$ yield identical relaxation times $\tau_{LVZ}$, for given values of $\lambda$, $\rho$, and $g$ (see equation (2)). In more sophisticated viscous relaxation models, this means in practice that the LVZ viscosity contrast with respect to the deep mantle and the cube of the LVZ thickness cannot be independently resolved by data whose effective horizontal wavelengths are large compared to the thickness of the LVZ ($\lambda \sim 2\pi D$), a condition that applies to most of the PGR data commonly used in global inversions for mantle viscosity structure. In section 3 we will show that the ratio $\eta/D^3$ also determines the degree to which an LVZ influences the pattern of mantle convection.

We have explored this modeling trade-off by comparing three forward models for the relaxation spectrum of a spherical Earth model with a 100-km-thick elastic lithosphere and a lower-mantle viscosity of $2.2 \times 10^{21}$ Pa·s, separated by a channel of lower viscosity $\eta$ and thickness $D$ (Paulson & Richards, 2009). Both the Earth’s surface and core-mantle boundary contribute relaxation modes for each component wavelength, or spherical harmonic degree (horizontal axis). These “mantle” and “core” modes are known, respectively, as the M0 and C0 modes, which in turn involve primarily deflection of the Earth’s surface and the core-mantle boundary. However, the C0 modes are of little importance in most PGR models, so we focus here on the M0 modes as a function of the viscosity and thickness of the LVZ layer.

Figure 1 compares the relaxation spectra of two models with almost identical values of $\eta/D^3$ (~2,000 Pa·s/m$^3$). The dotted line is for an LVZ of thickness 375 km and viscosity $1.3 \times 10^{20}$ Pa·s, and the dashed line is for an LVZ thickness of only 75 km and a much lower viscosity of $7.9 \times 10^{17}$ Pa·s. The two gray bands show the postglacial relaxation times (and uncertainties) measured at two sites on Hudson Bay (Mitrovica et al., 2000), which have been widely used in PGR models. (The horizontal extent of these gray bands has no meaning, but see later description of the spectral content of various data sources in Figure 2.) The solid curve in Figure 1 is for a uniform-viscosity mantle.

The comparisons in Figure 1 demonstrate that up to spherical harmonic degree $l = 20$ (horizontal wavelength $\lambda \sim 2,700$ km), the difference in the relaxation spectra for the two LVZ models remains smaller than the uncertainties in measured postglacial relaxation times, even though the two models have LVZ viscosities that differ by more than 2 orders of magnitude! Therefore, data with such large uncertainties (not to mention mutual disagreement) that do not constrain spherical harmonic degrees of at least $l \sim 20$ or higher cannot distinguish between a thin channel of extremely low viscosity and a thicker channel of much higher viscosity. On the other hand, we note that the uniform-viscosity model (solid line) and the two LVZ models differ substantially for $l \gtrsim 8$, so that the ability of PGR data to distinguish uniform mantle viscosity from the existence of an upper-mantle LVZ is relatively robust. Note also that the no-LVZ relaxation curve behaves qualitatively in accord with equation (1), whereas the two LVZ curves invert the dependence of relaxation time upon spherical harmonic degree, or inverse wavelength, in accord with equation (2).

Given the fundamental ambiguity in terms of $\eta$ and $D^3$ described above, what then are we to make of numerous attempts to constrain the nature of a putative LVZ from PGR data? Historically, the most influential paper published on PGR was Haskell’s (1935) analysis of the rebound of Fennoscandia following the most recent Ice Age. Treating the mantle as a viscous half-space, Haskell determined that the viscosity of the “upper mantle” beneath Fennoscandia must be $\sim 10^{21}$ Pa·s, a classic result that has stood the test of time remarkably well for over eight decades in the light of more sophisticated (spherical, layered) models and much

Figure 2. Approximate ranges of effective wavelengths (spherical harmonic degrees) measured by various postglacial rebound data sets (see text). The vertical axis is only used to separate the data sets and has no physical significance. PGR = postglacial rebound; GRACE = Gravity Recovery and Climate Experiment; RSL = residual sea level. (Figure is from Paulson & Richards, 2009.)
additional data. In this work, Haskell clearly understood that the restricted lateral dimension of the Fennoscandian ice sheet meant that there was little sensitivity in the data to the "deep mantle."

In a masterful update of this and other classic work on PGR, Mitrovica (1996) reanalyzed Haskell's work in a more modern context, making four important points about PGR studies in general: (1) geophysical definitions of what exactly is meant by the "upper mantle" or LVZ vary considerably and have led to confusion in the interpretation of geodynamic models, especially PGR models; (2) the Haskell (1935) constraint on the viscosity of the upper mantle applies, in fact, to mantle depths of order 1,000–1,200 km, and not to later definitions of the upper mantle more tied to the seismic discontinuity at about 670-km depth; (3) the addition of Hudson Bay residual sea level (RSL) constraints to the Fennoscandian data introduces a modest preference for models that include 1–2 orders of magnitude increase in viscosity with depth in the mantle; and (4) these data offer few independent constraints on the details of upper-mantle viscosity structure, consistent with the thin-channel ambiguity from Cathles (1975) we have reviewed above.

Further insight into this dilemma is gained by examination of the effective horizontal loading wavelengths represented by data sets commonly used in PGR modeling. Paulson and Richards (2009) compared the approximate spectral content of five types of PGR-related data (Figure 2); $J_2$ (true polar wander); residual sea level, or RSL, data (e.g., the Hudson Bay curves; see Mitrovica & Forte, 2004); Gravity Recovery and Climate Experiment (GRACE) satellite gravity changes (see Paulson et al., 2007a; Tamisiea et al., 2007); Fennoscandian rebound data (e.g., "strand lines"; McConnell, 1968; Mitrovica & Peltier, 1993; Wieczerkowski et al., 1999); and work with hybrid data sets by Nakada and Lambeck (1987, 1989) incorporating both RSL and continental margin tilting. Note that the shorter-wavelength ($l \geq 20$) data for Fennoscandia are geographically distinct from the longer-wavelength RSL data (mainly Hudson Bay), so that the temptation to combine these two data sets for joint inversion is fraught with peril, since the viscosity structure and lithospheric properties beneath the Laurentide and Fennoscandian ice sheets may be quite different. It is also noteworthy that the studies by Nakada and Lambeck (1987, 1989), whose data set includes short-wavelength contributions ($l \sim 40–80$) from ocean islands and small-scale continental embayments, stands out as inferring a considerably larger viscosity contrast for the LVZ than many other studies, at least beneath the oceanic lithosphere, a conclusion supported by Kendall and Mitrovica (2007) and consistent with an increasing ability to distinguish LVZ relaxation spectra from uniform mantle viscosity spectra with increasing spherical harmonic degree (decreasing effective wavelength).

To illustrate even more explicitly the ambiguity involved in modeling constraints that only contain long-wavelength ($l \leq 20$) information, Paulson and Richards (2009) performed a joint inversion of what at the time were considered the most reliable RSL data from Hudson Bay (Richmond Gulf and James Bay sites; Mitrovica et al., 2000) and GRACE data on the time-varying gravity field over Canada from the 53 months prior to 2006 (Paulson et al., 2007a). A grid search was performed using many Earth models via a spectral Green's function code (Paulson et al., 2007b) to compute the isostatic response of a viscously stratified spherical Earth to the loads given by the ICE-5G glacialization model (Peltier, 2004) and an ocean governed by the sea level equation (Farrell & Clark, 1976; Mitrovica, 2000). All models were overlain by a 100-km-thick elastic lithosphere. In order to examine the effects of varying the viscosity and thickness of an LVZ, all models included a lower-mantle (beneath the LVZ) reference viscosity of $1.4 \times 10^{21}$ Pa·s, compatible with previous estimates (Haskell, 1935; Mitrovica, 1996; and many others), which for uniform mantle viscosity (no LVZ) yielded a $\chi^2 = 2.8$ misfit to the data.

Figure 3 shows the misfit to the data as a function of the log of the ratio of the lower-mantle viscosity to the LVZ viscosity, $\eta^*$ (horizontal axis), and the LVZ thickness $D$ (vertical axis). The band of blue colors gives the lowest misfit, and the dashed line shows the relation $D \sim (\eta^*)^{-1/3}$ (equation (2)), which passes directly through the models of lowest misfit. Although these models are relatively simple, and the parameter search is limited, these results clearly illustrate the fundamental trade-off described by Cathles (1975), and given by equation (2).

There are many published modeling studies involving PGR-related data, mostly concluding that the data are compatible with an LVZ and some appearing to demand an LVZ 1–3 orders of magnitude less viscous than the underlying mantle. It is not our purpose here to review this literature, but a few comments are in order: First, Cathles's (1975) conclusion that it is likely that a pronounced LVZ of viscosity $\sim 10^{20}$ Pa·s and thickness ~250 km was strongly influenced by data on the rebound of Pleistocene Lake Bonneville (Utah, USA), which
in fact were subsequently shown to require an uppermost mantle viscosity as low as \(-10^{18}\) Pa·s (Passy, 1981). This latter inference is noteworthy also because it involves weak, tectonically active lithosphere and hot (mantle plume-influenced?) upper mantle. (The sensitivity of PGR studies to lithospheric thickness was described in detail by Kendall & Mitrovica, 2007, but in no way negates the fundamental $\eta^* \text{ versus } D^3$ trade-off emphasized above.)

Second, it is not entirely clear that a “global” upper mantle or LVZ model is appropriate to begin with, and not only because of the fundamental difference between ancient and thick continental lithosphere and relatively thin oceanic lithosphere. For example, Hudson Bay and Fennoscandian PGR are both occurring on old continental lithosphere. However, Fennoscandian relaxation time estimates (Wieczorkowski et al., 1999) are significantly longer than those for the Hudson Bay region, and uncertainties in the Fennoscandian data do not allow them to discriminate between a thin, weak LVZ and a relatively thick, strong LVZ. The incorporation of more “regional” or “local” data into models further exacerbates the question of general applicability, although studies involving shorter-wavelength constraints (tilting of shorelines, Global Positioning System (GPS), etc.) also appear compatible with a pronounced LVZ, but again with little real resolution on thickness (Nakada & Lambeck, 1987).

Thus, PGR studies leave us echoing Cathles’s (1975) main two conclusions regarding mantle viscosity and the LVZ: (1) The “deep-mantle” viscosity beneath the LVZ (but not necessarily at depths greater than \(-1,000\) km, as per Mitrovica, 1996) is of order $10^{21}$ Pa·s. (2) Most studies of PGR-related data are compatible with, and perhaps demand, an LVZ \(-1–3\) orders of magnitude lower in viscosity than the underlying mantle, and with only poorly constrained thicknesses ranging from less than 100 km to the entire upper mantle. We are not aware of any substantial progress on this lack of model resolution from PGR studies.

In the sections that follow we will demonstrate that the parameter $\eta^* D^3$, or, alternatively, the dimensionless version $\eta^* (D/\lambda)^3$ in the spectral domain, figures prominently not only in studies of PGR but also in geoid modeling and understanding the fundamental nature of mantle convection and plate tectonics on Earth (and possibly the other terrestrial planets). We will refer to this parameter as the “Cathles parameter,” or $Ct$.

### 2.2. The Geoid

Although the long-wavelength “geoid problem,” like PGR, has been understood theoretically since the 1930s (Morgan, 1965; Pekeris, 1935), it was not until the advent of global seismic tomography in the late 1970s and early1980s that modeling the geoid to constrain mantle heterogeneity structure and dynamic topography became truly feasible (Dziewonski et al., 1977; Hager et al., 1985).

This important development in geodynamics can be understood in reference to the simple two-dimensional schematic model of Figure 4. At some depth $D$ in the mantle, a sinusoidal lateral density contrast of $\sigma_D$ and wavelength $\lambda$ (or wave number $k = 2\pi/\lambda$) causes sinking and rising flow, respectively, depressing and elevating the topography of the surface (lithosphere) and core-mantle boundary, as well as any internal nonadiabatic (chemical) boundary layers within the mantle. The total gravity signal, usually expressed for long wavelengths in terms of elevation or depression of the gravitational equipotential surface (roughly sea level) or “geoid,” is thus composed of the signal from the density contrast itself and the contributions from the deflection of the lithosphere and core-mantle boundary from their equilibrium values for an otherwise hydrostatic, rotating Earth.

For whole-mantle flow models (no interior nonadiabatic layering), contributions from these surface deflections are opposite in sign from that of the internal density contrast, so that the geoid signal measured at the surface is often the small difference among signals of comparable magnitude. In fact, for models with constant viscosity, the surface deformation, being closer to where the geoid is measured, “wins,” (aided also by the core-mantle boundary deflection), so that the net geoid signal from a positive density contrast is
actually negative. For layered viscosity, the signal may be positive or negative, depending upon the viscosity structure; the wavelength, or spherical harmonic degree; and the depth of the density contrast. Thus, the geoid is sensitive not only to the mantle’s internal lateral density structure but also to the viscosity structure.

Employing three-dimensional spherical Earth geometry, Figure 5 compares geoid “kernels” $G_l(r)$ for low spherical harmonic degrees $l = 2, 3, \text{and} 6$, corresponding to wavelengths of approximately 20,000, 13,000, and 7,000 km, respectively. These kernels give the geoid measured at the surface as a function of the depth of the driving internal density contrasts for each spherical harmonic degree. (The geoid kernels $G_l(r)$ are normalized to the geoid that would be obtained if a density contrast was placed at the Earth’s surface with no deformation of either the surface or core-mantle boundary. The variable $r$ is the depth of the density contrast.) The top panels are for uniform mantle viscosity, the middle panels are for a viscosity contrast of 10 between the upper and lower mantle, and the bottom panels are for a viscosity contrast of 30, all with the upper/lower mantle boundary at 670-km depth (Richards & Hager, 1984). The panels on the left (models U1, U10, and U30) are for an adiabatic mantle (no internal flow boundaries), and the panels on the right (models C1, C10, and C30) include a flow boundary at 670-km depth, driving the geoid kernels at that depth level to zero due to perfect compensation for a density contrast placed at that level. We note that these six models result in very different geoid kernels and hence very different inferences regarding mantle viscosity structure when using, for example, seismic tomography to map internal density contrasts and model the geoid (Hager et al., 1985; Hager & Richards, 1989).

The geoid is, of course, a nonunique reflection of density contrasts and surface deflections from throughout the mantle, but the advent of seismic tomography made it possible to greatly reduce this nonuniqueness, in principle, by providing independent estimates for internal density structure. Following on the initial theoretical work for spherical Earth models (Ricard et al., 1984; Richards & Hager, 1984), a large number of studies modeling the geoid for various tomographic density models were undertaken, with a principal goal being that of determining the radial viscosity structure of the Earth. As discussed in more detail below, the general conclusion from these studies was that the upper mantle (surface to 670-km depth) was on average about a factor of 10 or so less viscous than the lower mantle, but different models yielded different results for the more detailed viscosity structure of the upper mantle, some including a relatively thin LVZ (e.g., Hager & Richards, 1989), while others used a more coarse parameterization in terms of upper versus lower mantle, for example, model U10 in Figure 5 (Hager et al., 1985).

Although the geoid is determined at a level of precision for long wavelengths that is entirely sufficient for the kind of modeling discussed here, geoid modeling, like PGR, suffers from a number of important (and similar) limitations: uncertainties regarding appropriate rheologies (linear versus nonlinear viscous creep, plastic failure, and faulting at plate boundaries); expected strong lateral viscosity variations and imperfect constraints on internal density (e.g., subducted slabs versus seismic tomography); inherent nonuniqueness of gravity; and what turns out to be the same nonuniqueness as for PGR regarding viscosity contrast and layer depth for the LVZ.

As indicated in Figure 4, the long-wavelength geoid depends most critically upon the vertical normal stress due to fluid mechanical tractions acting on the base of the lithosphere. Although both the lithosphere and core-mantle boundary deflections can be important, it is the lithospheric deflection that usually dominates the net geoid signal, somewhat analogous to the above-noted dominance of the $M_0$ mode over the $C_0$ mode for the PGR problem. As it turns out (see Appendix A), the effect of an LVZ on the dynamic topography of the upper boundary (lithosphere) in the long-wavelength limit scales, like PGR, as the dimensionless Cathles parameter:

$$\tau_{ZZ} = 1 - (kD)^3 \eta^*$$  \hspace{1cm}(3)

where $\tau_{ZZ}$ is the vertical normal stress due to a lateral density contrast within the mantle of wave number $k = 2 \pi / \lambda$, $D$ is the thickness of the LVZ, and $\eta^*$ is the ratio of the viscosity of the underlying mantle to that of the LVZ. (Note that the derivation in Appendix A uses $D$ as both the depth of the LVZ and the depth of
Figure 5. Dynamic response functions (geoid kernels) for surface density contrasts of spherical harmonic degrees 2 (solid lines), 3 (long dashes), and 6 (short dashes) plotted against radius for six Earth models. (left) Models U have uniform composition, which permits mantle-wide flow; that is, they assume a purely adiabatic mantle. (right) Models C have a chemical discontinuity at 670-km depth, causing stratification into separate upper and lower mantle flow systems. Models in the top row have uniform viscosity; models in the middle row have a factor of 10 viscosity increase below 670 km; models in the bottom row have a factor of 30 viscosity increase below 670 km. The geoid kernels are the surface geoid signal normalized by that which would be obtained for an equivalent uncompensated density contrast placed at the Earth’s surface.
The lateral density contrast, but only for analytical convenience, as the basic scaling result remains the same, as will be apparent in Figure 6.) In other words, for a given long-wavelength mantle density contrast, the effect of flow on upper boundary deformation is identical for identical values of the ratio \((kD)^3\eta^*\). This result holds in the long-wavelength (small wave number) limit such that \(kD \ll 1\). Although a trade-off between layer thickness and viscosity contrast has been implicitly recognized in earlier studies (King & Masters, 1992; Richards & Hager, 1988; Thoraval & Richards, 1997), we are not aware of the scaling relation of equation (3) having been made explicit in any previous studies. We have therefore included a derivation in Appendix A. The fact that this scaling is identical to the modeling trade-off derived by Cathles for PGR modeling (Paulson & Richards, 2009) is a general limitation on geodynamic modeling of the radial viscosity structure of the Earth and is unlikely to be resolved, for example, by joint modeling of PGR and geoid data.

To make this result concrete, Figure 6 compares the spherical harmonic degree \(l = 2\) and 6 geoid kernels (computed as in Richards & Hager, 1984) for two seemingly very different LVZ models that have identical Cathles parameters. The red and yellow curves are geoid kernels as a function of mantle depth for spherical harmonic degrees \(l = 2\) and 6, respectively, for a model in which a 100-km-thick lithosphere of relative viscosity 1,000 overlies a lower mantle of viscosity 1. (The absolute viscosity does not enter the geoid problem explicitly, and only viscosity contrasts are relevant in these models.) The blue and gray curves are again for \(l = 2\) and 6, respectively, but for a model in which the LVZ is only 100 km thick and of viscosity 1. As can be seen in Figure 6, the geoid kernels for the two different models are very similar, as predicted by equation (3), and clearly not sufficiently distinct for any inversion procedure relying upon imperfect data, such as seismic tomography, to resolve. And yet, to the untrained eye, these two viscosity models would appear to be radically different! Thus, the fundamental dilemma involving the Cathles parameter trade-off that we identified for PGR studies is fully present, in exactly the same form, when using models for the geoid to constrain mantle viscosity structure.

(In Figure 6 the geoid kernels are computed at ~150-km-depth intervals rather than as continuous functions. This emphasizes the limited vertical resolution of seismic tomography, so that the differences in the kernels that manifest mainly in the uppermost mantle contribute relatively little to the overall model geoid and occur in the uppermost regions of the mantle where the relationship between density and seismic velocity are highly uncertain.)

3. Mantle Convection

From the earliest days of the plate tectonics revolution it was suspected that a low-viscosity layer below plates, the asthenosphere, plays a role in the operation of plate tectonics. The essential idea was that a low-viscosity layer would lower the resistance to plate motions (e.g., Cox, 1973). Under this view, the role of the asthenosphere, for the operation of plate tectonics, is to provide basal "lubrication." Relative to the themes of this paper, some aspects of the lubricant interpretation are worth noting: (1) The defining property of the asthenosphere is considered to be its viscosity relative to that of the lithosphere. (2) Flow in the asthenosphere is assumed to be passive, that is, driven by plate motions from above (this equates to the idea that the asthenosphere is associated with plate resisting forces). (3) And the role of asthenosphere thickness, if any, is not made explicit or even treated as a matter of importance.

3.1. Plate Lubrication Versus an Asthenosphere

The lubricant interpretation of the asthenosphere remains a prevalent one to this day. From the mantle dynamics perspective, signs that it might be incomplete can be traced back to another long-standing observation from the plate tectonics revolution: The lateral dimensions of major tectonic plates (e.g., the Pacific...
Plate) are considerably larger than the depth of the mantle (and much larger than the depth of the upper mantle, to the extent that layered convection remains a relevant concept). The implication is that convective cells in the Earth’s mantle are of relatively long wavelength. This inference was supported by seismic tomography studies that mapped long-wavelength structure in the mantle (e.g., Su & Dziewonski, 1992) and modeling of the Earth’s geoid, which was also consistent with long-wavelength mantle structure (e.g., Hager & Richards, 1989). The dominant mantle heterogeneity structure appears to occur at spherical harmonic degrees 2–3, corresponding to horizontal wavelengths of order 13,000–20,000 km (Richards & Engebretson, 1992). This presented a dynamic problem in that long-wavelength cells are not the norm at the level of convective vigor inferred for the Earth’s mantle, that is, for Rayleigh numbers of order $10^7$–$10^9$ (Busse, 1985; Turcotte & Schubert, 1982). A potential solution to the long-wavelength flow problem came from numerical simulations of mantle convection that showed that a high-viscosity lower mantle below a low-viscosity upper mantle could lead to long-wavelength flow (e.g., Bunge et al., 1996, 1997; Hansen et al., 1993; Tackley, 1996; Zhang & Yuen, 1995; Zhong et al., 2000; Zhong & Zuber, 2001). Low-viscosity upper mantle is a coarse model analog of the asthenosphere, and the numerical results are consistent with the idea that an asthenosphere facilitates long-wavelength mantle convection.

Numerical simulations alone did not elucidate the physical mechanism(s) by which a depth-variable viscosity could generate long-wavelength flow. They could show the viability of that result, in the way laboratory tank experiments can show the viability of a flow configuration, but theory was required to get at the essential physics. The application of the boundary layer theory to this problem suggested that the same combination of variables that defines the Cathles parameter, from PGR and geoid considerations, is also critical for mantle dynamics (Busse et al., 2006; Lenardic et al., 2006). As a small aside, we note, as coauthors on both of these studies, that the significance of this connection was not at all obvious at the time (nor has it been explicitly remarked upon, to the best of our knowledge, to this day). An appreciation of these boundary layer results will benefit from a review of why long-wavelength cells are not the norm in vigorously convecting layers.

In a mantle layer with no internal viscosity variations, long-wavelength cells become unstable because lateral viscous dissipation, associated with horizontal flow in the convecting layer, dominates over vertical dissipation. This notion is implicit in earlier work by Turcotte and Oxburgh (1967) and Busse (1985) but was derived explicitly by Busse et al. (2006). Lateral dissipation increases rapidly with cell wavelength, largely as a simple matter of geometry. The associated viscous resistance slows the lateral motion of the upper thermal boundary layer. This leads to boundary layer thickening and the initiation of cold sinking instabilities that break up would-be long aspect-ratio cells. The dominance of lateral over vertical dissipation is independent of absolute mantle viscosity—lowering the viscosity of the mantle uniformly does not change the conclusion above.

Introducing a high-viscosity lower mantle means that two viscosities determine bulk mantle dissipation. This changes the relative effects of lateral versus vertical dissipation (Busse et al., 2006; Lenardic et al., 2006). With a high-viscosity lower mantle, vertical dissipation can become dominant in terms of total energy dissipation over a broader wavelength range. As the depth of the mantle is fixed, the vertical term does not increase in the same way as the lateral term does with increasing aspect ratio. This allows mantle convection cells to remain stable over broader wavelength bands. The boundary layer theory formalizes these ideas through a global work-energy balance that equates the rate at which work is done on sinking/rising boundary layers (analogs to subducting slabs and hot plumes) via gravitational body forces to the rate at which work is done on the central portion of a convection cell by viscous forces. The total viscous work term is associated with energy dissipation (viscous resistance) and is composed of vertical and horizontal shear components. A key assumption of the boundary layer analysis, applied to the case of depth-variable viscosity, is that lateral flow in the mantle below plates channelizes into a low-viscosity, relatively thin region in the upper mantle (Busse et al., 2006; Lenardic et al., 2006). This analysis led to the prediction that a low-viscosity upper-mantle channel should increase flow wavelength. The analysis further predicted that the increase should scale as the product of the thickness of the channel cubed and the ratio of the lower-mantle viscosity to that of the low-viscosity channel, that is, the Cathles parameter, $Ct$.

Figure 7 compares boundary layer theory predictions to results from suites of numerical simulations. Before and beyond the cell aspect ratio that maximizes surface heat flux, convective cells become prone to instability (e.g., Ahmed & Lenardic, 2010). Thus, tracking how the maximizing wavelength scales with system parameters provides a metric for the degree to which an asthenosphere can favor long-wavelength flow. A
key prediction, consistent with numerical solutions that solve the full conservation equations for thermal convection, is that absolute viscosity is not the critical factor for altering flow wavelength trends (Busse et al., 2006; Lenardic et al., 2006). Keeping this in mind will prevent one from assuming that long-wavelength flow can be achieved by “making the mantle below plates weaker”—in fact, a whole mantle sublithospheric viscosity increase will increase the potential of small-scale boundary layer instabilities that can break up long aspect ratio convection cells (Busse, 1985). What is critical is the viscosity increase from the upper to lower mantle that allows lateral flow below plates to channelize. As well as effecting the balance of vertical to lateral viscous dissipation, channelized flow can suppress small-scale upper boundary layer instabilities (Hoink & Lenardic, 2010). A viscosity increase with depth can be implemented by making the lower mantle more viscous without any change in the viscosity of the upper mantle. That is, a channelization view of the asthenosphere is really quite different from the classic lubrication view—lubrication alone does not lead to longer-wavelength flow, nor does it favor plate-like surface motions in and of itself.

3.2. Asthenosphere-Plate Coupling

The above analysis assumed that flow resistance is derived principally from the sublithospheric mantle. That is, potential flow resistance at plate margins was assumed to be small (a weak margin assumption). When this assumption is relaxed, dissipation from plate margin zones enters into the energetics (e.g., Conrad & Hager, 1999). The initial numerical experiments of Lenardic et al. (2006) explored this potential by allowing a relatively low viscosity asthenosphere to reside below a higher viscosity near surface layer (an analog for a strong plate). Those results indicated that even longer-wavelength flow could be stabilized. Subsequent numerical experiments by Hoink and Lenardic (2008, 2010) confirmed that conclusion. Figure 8 replots some of those results along with results from a configuration that did not allow for the modeling of plate strength. Figure 9 shows select numerical cases from Hoink and Lenardic (2010) along with average horizontal velocities from one experimental suite.

Of note in Figure 9 is that flow in the asthenosphere is no longer passive (i.e., driven by plate shear from above) across all cell wavelengths. Shear flow in a viscous channel is associated with a near-constant velocity gradient (Couette flow). The longest aspect ratio cases do show this flow type but at shorter aspect ratios the velocity field exhibits pressure-driven flow (Poiseuille flow). Such a flow configuration was not allowed for in the boundary layer analysis of Busse et al. (2006) and Lenardic et al. (2006). The transition from Poiseuille to Couette flow is associated with a change in the sign of the shear stress at the base of the model plate. That is, basal tractions at the lithosphere-asthenosphere boundary can transition from plate driving to plate resisting terms. This is another distinct break from the lubricant view of an asthenosphere.

Physical insights into the model behavior discussed above came from two independent studies: (1) a semi-analytic boundary layer formulation based on an integral form of the global energy balance (Crowley & O’Connell, 2012) and (2) a scaling analysis based on an effective channel Rayleigh number (Hoink et al., 2011). The two studies were consistent in terms of key implications for mantle dynamics. Both indicated that as the effective viscosity of a plate relative to the asthenosphere increases, convective cells can operate in a sluggish-lid mode. In that mode, surface plate velocity is finite (i.e., not a stagnant lid) but lower than that of the mantle below.
This regime occurs when plate resistance is comparable to viscous drag on its base (Crowley & O’Connell, 2012; Solomatov, 1995). If mantle dissipation is low relative to that associated with plate strength, then the plate dissipation term can be dominant in the global energy balance. For a low-viscosity channel below plates, the lateral pressure gradient associated with upper boundary layer thickening can drive a channelized flow that exerts a drag on the plate above (Hoink et al., 2011). The pressure gradient in the low-viscosity channel increases with cell wavelength (Crowley & O’Connell, 2012; Lenardic et al., 2006). As the work done on the plate by the pressure-driven flow (and associated basal traction) increases, it can overcome plate resistance. At that stage the surface velocity exceeds internal velocities and the system transitions to active-lid convection, with the mantle now dominating global dissipation (Crowley & O’Connell, 2012; Hoink et al., 2011). This is favored for a configuration with a relatively weak upper mantle, which insures that lateral mantle dissipation does not dominate the energetics (Crowley & O’Connell, 2012; Hoink et al., 2011). Collectively, this allows longer cells to be stabilized.
Returning to our main theme, the studies discussed above show that relative asthenosphere viscosity and the thickness work in combination to influence mantle dynamics. The wavelength beyond which convection cells become unstable occurs near the point at which plate and asthenosphere velocities become equal. This is the point at which asthenosphere flow switches from driving plate motion to resisting plate motion. Figure 10 plots numerical simulation results and predicted theoretical trends for the plate-asthenosphere velocity ratio (Hoink et al., 2011). The theory predicts that the Cathles parameter combination once again plays the key role. Although in the studies of Hoink et al., 2011, and Crowley & O’Connell, 2012, developed independently, once the principals became aware of each other’s work they shared all results to check consistency; the analysis of Crowley & O’Connell, 2012, also predicted the numerical trends with good accuracy. The ratio of effective plate viscosity (a parameterization of margin strength) to asthenosphere viscosity also enters into predicted scaling trends (Crowley & O’Connell, 2012; Hoink et al., 2011). We have plotted predicted trends based only on asthenosphere properties in Figure 10 to emphasize that despite this added complexity, the Cathles combination remains first order in its effect on system dynamics. Recall that the viscosity of the asthenosphere relative to the lower mantle, not the lithosphere, is used for the Cathles parameter.

It is noteworthy that \( \psi \) remains a key factor determining convection wavelength when plate strength is allowed for—again, the key for understanding the energetics is the effective plate strength at margins. Even more noteworthy is that \( \psi \) is now also predicted to be a key factor for the balance of plate driving and plate resisting forces—a critical consideration for any efforts to move plate tectonics from a kinematic to a dynamic theory. Many recent efforts in that direction have focused on, and continue to focus on, the rheology and/or strength of plate margin zones. However, we argue that the repeated appearance of a particular combination of asthenosphere properties, in realms that at first blush might not seem connected, is a signpost that the dynamic role of the asthenosphere not only cannot be ignored, but that it can also be understood in a remarkably straightforward fashion.

4. Concluding Remarks

The radial viscosity structure of the Earth’s mantle is critical to understanding many facets of Earth’s internal dynamics and evolution. Most geodynamic studies are consistent with a significant increase in viscosity with depth in the sublithospheric mantle but do not, as a whole, resolve the amount of this viscosity increase or the thickness of the LVZ, or asthenosphere. Increasingly sophisticated models seeking to explain Earth’s unique plate tectonic style of mantle convection show that a key ingredient is a pronounced asthenosphere (Richards et al., 2001; Tackley, 2000). The main results of all of these studies appear to be largely characterized via the Cathles parameter, which can be defined in dimensionless form as

\[
\psi = \frac{\eta^*}{D/\lambda},
\]

where \( \eta^* \) is the ratio of deep mantle to asthenosphere viscosity, \( D \) is the depth of the asthenosphere, and \( \lambda \) is the effective wavelength of the relevant flow field. We do not mean to suggest that \( \psi \) captures all interesting aspects of mantle viscosity structure or that it characterizes all important geodynamics effects of the asthenosphere, but for the types of studies discussed above it serves as a working geodynamic definition of the asthenosphere.

An important exception to the applicability of \( \psi \) is that of postseismic relaxation (PSR). PSR is the slow after-creep that follows large earthquakes, which can persist for years to decades. The characteristic timescale of PSR for an elastic lithosphere overlaying a viscous channel (asthenosphere) is given by \( t_{PSR} \sim Hh/\mu \), where \( H \) is the lithospheric thickness, \( \eta \) is the channel viscosity, \( h \) is the channel thickness, and \( \mu \) is the elastic rigidity of the lithosphere (Segall, 2010, Chapter 6, “Postseismic Relaxation”). Ignoring the lithosphere, the effect of the asthenosphere is characterized by \( t_{PSR} \sim \eta/h \), which is essentially what most models of PSR seek to fit in terms of the after-creep data. This scaling suggests that PSR data have a different dependence upon the...
viscosity and thickness of the asthenosphere, so that comparing PSR to, for example, geoid or PGR modeling results might help to resolve $h$ and $\eta$ independently.

To this point, PSR data are generally sensitive to the viscosity of the mantle immediately underlying the fault zone and constitute a “short-wavelength” data set relative to the geoid and PGR data discussed above. Importantly, PSR studies generally suggest very low sublithospheric viscosities of order only $10^{18}$–$10^{19}$ Pa·s, consistent, perhaps tellingly, with estimates derived from the “short-wavelength” rebound of Lake Bonneville (Passy, 1981). In a recent study, Hue et al. (2016) derived an asthenosphere viscosity of only $\sim 10^{18}$ Pa·s from analysis of GPS data from the $M_w = 8.6$ Indian Ocean earthquake of 2012. These authors also concluded that the PSR times in their models scale approximately as $\tau_{\text{PSR}} \sim \eta/h^{1.5}$. Because their models included many complications, including a continental back-arc mantle distinct from the sub-oceanic fore-arc mantle, as well as a high-viscosity subducted slab, it is not clear whether this scaling is significantly different from that of Segall (2010). However, both results yield scalings that are significantly different from that of the Cathles parameter for the types of geodynamic problems discussed in the foregoing sections. Moreover, the most important message from PSR studies is that they imply that the asthenospheric layer is probably very thin. For example, PGR models with an asthenospheric layer of viscosity only $10^{18}$ Pa·s would suggest that the asthenosphere is only $\sim 100$ km thick. As we showed in section 2, such a thin, high-viscosity-contrast LVZ does not appear to be resolvably inconsistent with either PGR or geoid modeling as long as the Cathles parameter trade-off between viscosity contrast and channel thickness is properly taken into account.

At the same time, it is also important to keep in mind that the subcontinental and sub-oceanic asthenospheres may be quite distinct, as suggested by seismic evidence that strong shear wave anisotropy aligned with plate motions beneath the oceans may be confined to a layer of order $\lesssim 200$-km thickness (Nettles & Dziewonski, 2008), whereas continental regions exhibit a seismic LVZ and anisotropic asthenosphere occurring at great depths of order 200–400 km (French & Romanowicz, 2014). PGR studies are dominated by data from continental areas, mainly in proximity to where large Pleistocene ice sheets have formed, so it appears quite possible that a thin, high-viscosity-contrast asthenosphere may apply more to the mantle beneath oceanic lithosphere. This would not by any means nullify the Cathles parameter concept, but it does emphasize that strong lateral viscosity contrasts in the mantle render it difficult to resolve details of radial viscosity structure in a global sense from either PGR or geoid modeling studies (see, e.g., Paulson et al., 2007b, regarding effects of radial viscosity variations on PGR, and Thoraval & Richards, 1997, regarding effects of lateral viscosity variations on the geoid). This said, virtually all geodynamic studies appear to be consistent with a pronounced asthenosphere, and we believe it likely that the asthenosphere is best characterized as being of order only $\sim 100$–200 km thick and a factor of order $\sim 100$–1,000 lower viscosity than that of the deeper mantle, at least beneath the oceanic lithosphere.

We have not attempted here to shed any light upon the physical nature of the asthenosphere, which may result from the presence of sublithospheric partial melt (approach of the geotherm to the mantle solidus), the presence of volatiles due to recycling at subduction zones ($\text{H}_2\text{O}$ and $\text{CO}_2$), or perhaps also mineralogical effects. From a strictly geodynamic modeling standpoint, and especially to the extent that such a low-viscosity layer is ubiquitous, it is not necessary to know exactly how or why the asthenosphere arises.

We note that Venus does not appear to have an asthenosphere, based on gravity/topography studies (Kiefer et al., 1986), and the absence of water on, and presumably within the upper mantle of, Venus figures prominently in explanations for the lack of plate tectonics there (Richards et al., 2001). Here we have mainly been concerned with showing that many of the principal geodynamic effects of the asthenosphere may be understood via a remarkably simple parameterization, which we propose to name the Cathles parameter and which we hope will be useful to future modelers.

**Data Availability**

All the data used in this paper are from previously published work (see references).

**Appendix A: Geoid Anomaly Due to a Density Contrast in a Layered Half-Space**

In an adiabatic, Newtonian viscous mantle, the geoid anomaly generated by a density contrast at depth consists of signals due to the density contrast itself plus the fluid-dynamically-induced deformations of the
upper surface and the core-mantle boundary (Morgan, 1965; Ricard et al., 1984; Richards & Hager, 1984). In fact, for a uniform-viscosity mantle a net negative geoid signal results from a positive mass anomaly, mainly because the upper surface (lithosphere) deformation is also the reference surface for the geoid itself (radius of observation). This paradoxical result, however, does not hold for cases where the viscosity increases markedly with depth in the mantle, because much of the compensating boundary deformation is transferred to the core-mantle boundary, which is much more remote from the reference surface. The geoid signal being the “small difference of larger numbers” is therefore extremely sensitive to radial viscosity structure in the mantle.

In this section we examine a simple but instructive case that illustrates how the Cathles parameter controls, to first order, the trade-off between viscosity layer thickness and viscosity contrast. We undertake this analysis following the two-dimensional Cartesian half-space analysis described in detail in the Appendix of Hager and O’Connell (1981). This analysis uses a Fourier analysis of Newtonian viscous flow, driven by internal density contrasts, and allowing for vertically layered viscosity. A standard propagator-matrix approach is used to obtain analytical solutions for the resulting system of linear differential equations.

Referring to Figure 4 of the main text, we consider a sinusoidal density contrast with amplitude \( \sigma_0(k) \) placed at a depth \( D \) in a half-space. For convenience, we set the ratio of the viscosity of the fluid above \( D \) to the viscosity below \( D \) to \( \eta^* \), so that a value of \( \eta^* \ll 1 \) models a low-viscosity upper mantle between the free surface and the deep mantle. All variables in the model (velocities and stresses) occur at the relevant wave number \( k = 2\pi/\lambda \), where \( \lambda \) is the Fourier wavelength. The relevant variables are as follows:

\[
\begin{align*}
u_1 &= v_z \quad \text{(vertical velocity)} \\
u_2 &= v_x \quad \text{(horizontal velocity)} \\
u_3 &= \tau_{x'\theta}/2\eta_k \quad \text{(normalized vertical normal stress)} \\
u_4 &= \tau_{x'\theta}/2\eta_k \quad \text{(normalized shear stress)}
\end{align*}
\]

where \( \eta_0 \) is an arbitrary reference viscosity. Hager and O’Connell (1981) define a 4 \( \times \) 4 propagator matrix for “propagating” a velocity/stress solution vector \( \mathbf{u} = (u_1, u_2, u_3, u_4) \) from a level \( z_1 \) to a level \( z_2 \) within the fluid half-space. This propagator matrix (their equation A27), consists of terms that include only the viscosity ratio \( \eta^* \) and hyperbolic sine and cosine functions of the wave number and vertical coordinates \( \cosh(k(z_2 - z_1)) \) and \( \sinh(k(z_2 - z_1)) \) such that \( P(z_2, z_1) = f(\eta^*, k, z_1, z_2) \) only.

One can then write the solution vector \( \mathbf{u}_0 = (u_{10}, u_{20}, u_{30}, u_{40}) \) at the surface \( (z = 0) \) as

\[
\mathbf{u}_0 = P(D, 0) \mathbf{u}_0 + P(D, 0)\sigma_0 = (A1)
\]

where \( \mathbf{u}_D \) is the velocity/stress solution vector at \( z = D \), and the vector

\[
\sigma = (0, 0, \sigma_D/2\eta_k, 0)
\]

contains the Fourier density contrast component that drives the flow and stresses, which is also (for convenience) placed at depth \( D \).

Expression \( (A1) \) is a set of four algebraic equations for the components of the two velocity/stress solution vectors \( \mathbf{u}_0 \) and \( \mathbf{u}_D \) at the top surface and depth \( D \) within the fluid half-space. The eight unknowns in these solution vectors can be reduced to four via application of boundary conditions: At the free-slip upper surface, the vertical velocity and horizontal shear stress are zero, so that

\[
\mathbf{u}_0 = (0, u_{20}, u_{30}, 0)
\]

As \( z \to \infty \), the solution vector must vanish, requiring that the solution vector at \( z = D \) must be of the form

\[
\mathbf{u}_D = (u_{1D}, u_{2D}, u_{1D}, u_{2D})
\]

Applying these solution vector forms, taking the long-wavelength limit \( 2\pi/k \gg D \), and keeping only the leading terms in \( kD \), expression \( (A1) \) yields (after considerable algebra) the leading terms for the vertical normal stress term at the level \( z = 0 \), normalized to the driving density contrast
Expression (A2) tells us that for long-wavelength loads, the vertical normal stress at the upper free surface \( \sigma_D \) depends to first order upon a dimensionless version of the Cathles parameter. The geoid signal is solely a function of the driving density contrast \( \sigma_D \) and the upper surface deformation, which in turn is controlled by the vertical normal stress. Therefore, expression (A2) shows that the geoid signal normalized to the driving density contrast \( \sigma_D \) is, to first order, a function of only of the Cathles parameters as reflected in expression (3) in the main text and in the results of Figure 6 comparing long-wavelength geoid signals from radically different viscosity models with identical Cathles parameters.

\[
\sigma_D \frac{1-a}{\sqrt{a}}
\]

(A2)

Acknowledgments
This paper represents a long-term collaborative effort, and we are grateful for generous support from the U.S. National Science Foundation via several grants to us individually over the past decade. Our previously published work with Archie Paulson and Tobias Hoink was critical to the thread of the arguments presented herein. Carolina Lithgow-Bertelloni kindly provided an updated code to compute the geoid kernels in Figure 6. We thank Yanick Ricard and Thorsten Becker for thoughtful reviews of the original manuscript.

References


