
1. Introduction

The Theory of Plate Tectonics Is Celebrating Its Fiftieth Anniversary. Since Its Inception, the Theory Has Served to Explain How the Earth’s Surface Arrived at Its Present State, Including Its Geological Expression and the Connection to Processes Such as Earthquakes, Volcanic Activity, and Climate Evolution. The Forces Driving Plate Tectonics Have Since Been Recognized to Arise from Mantle Convection, Linking Earth’s Surface Motion to the Dynamics of Its Deep Interior and Planetary Scale Evolution. At the Heart of Our Search for Understanding How and Why Earth, Alone Among the Known Terrestrial Planets, Has Plate Tectonics, Is the Physics of Rock Deformation.

The Plate Tectonic Style of Convective Motion at the Earth’s Surface Involves Severe Deformation of the Lithosphere to Form Narrow Plate Boundaries, Even as Most of the Lithosphere Is Left Largely Undeformed to Comprise Plate Interiors. However, the Cause for Plate Boundary Formation Remains Somewhat Elusive. The Mechanical Strength of Plate Interiors Appears to Be Too High to Deform Under the Available Forcing from Mantle Convection (Cloetingh et al., 1989; Solomatov, 1995). The Occurrence of Plate Tectonics Thus Requires a Mechanism for Localized Weakening to Form Plate Boundaries with Which to Mobilize the Surface (Bercovici, 1993, 1995, 2003; Bercovici et al., 2015; Kaula, 1980; Montési, 2013; Tackley, 2000, 2000). Plate Boundaries Span the Entire Depth of the Lithosphere and Thus Require Strain Localization Across a Range of Rheological Regimes: From Brittle Faulting in Shallow Regions to Ductile Shear at Depth. The Strength of the Lithosphere Peaks in Its Cold Ductile Region around 40–80 km Depth (e.g., Karato, 2008 pp. 338–358; Kohlstedt et al., 1995) and Thus Presents the Main Impediment to Shear Localization and the Formation of Tectonic Plate Boundaries. Many Modern Plate Boundaries Appear to Form Over Preexisting Weak Zones (Gurnis et al., 2000; Hall et al., 2003; Lebrun et al., 2003; Tommasi et al., 2009; Toth & Gurnis, 1998; Stern, 2004; Stern & Bloomer, 1992), Which Implies That the Mechanism Responsible for Lithospheric Weakening Is Long-Lived and Allows for the Inheritance and Reactivation of Older Plate Boundaries.

The Occurrence of Strain Localization Requires a Self-Weakening Positive Feedback Mechanism, Whereby Deformation Causes Weakening, Which Leads to Increased Deformation, Further Weakening, and So On.
Lithosphere and crustal rocks can weaken through a number of phenomena, such as an increase in temperature, the presence of fluids, and the changes in microstructure. Frictional heating in a deforming rock and the associated strain localization due to thermal weakening allows for some strain localization in the lithosphere. However, the weakening effects of shear heating are mitigated by thermal diffusion, which rapidly erases the warm weak regions over a few million years (Bercovici, 1998; Bercovici & Karato, 2003; Bercovici et al., 2015; Lachenbruch & Sass, 1980). In addition, the accelerated grain growth at higher temperatures and the resulting increase in viscosity further curbs the weakening effect of frictional heating (Foley, 2018; Kameyama et al., 1997). The presence of fluids, such as water, in the lithosphere may induce weakening through lubrication of plate boundaries at subduction zones or by reducing friction through increasing pore pressure (Bercovici, 1998; Gerya et al., 2008). While the rheological effects of fluids may be long lived, because of relatively slow diffusion of hydrogen in minerals, they are restricted to shallow lithospheric depths (i.e., the upper tens of kilometers), since it is prohibitively difficult to deliver water to the deepest regions of the lithosphere (e.g., Korenaga, 2017).

Microstructural influences on rock strength include viscous anisotropy, for example, by crystal preferred orientation, or CPO (Durham & Goetze, 1977; Hansen et al., 2012; Zhang & Karato, 1995), mineral grain size reduction through dynamic recrystallization, microcracking or other forms of damage (see; Bercovici et al., 2015 for a recent review). Lithospheric weakening through changes in microfabric can be persistent, since, once deformation ceases, erasing the microstructure (e.g., randomizing crystal orientations in the case of CPO, or grain growth in the case of grain size reduction) is much slower than convective mantle overturn, thus allowing for inheritance and reactivation of lithospheric weak zones (Bercovici & Ricard, 2014; Tommasi et al., 2009).

In areas where ductile lithosphere is exposed at the Earth’s surface near plate boundaries, such as in ophiolites (Hansen et al., 2013; Jin et al., 1998; Linckens et al., 2011; Kohlstedt & Weathers, 1980; Skemer et al., 2010; Warren & Hirth, 2006), severe deformation is coincident with highly reduced grain size, as is evident in mylonites and ultramylonites (Furusho & Kanagawa, 1999; White et al., 1980). Strain weakening by grain size reduction is also commonly seen to accompany rock deformation in experimental studies (De Bresser et al., 1998; Green & Radcliffe, 1972; Hansen et al., 2012; Karato et al., 1980). CPO microfabric is observed in lithospheric shear zones as well but is restricted to the coarse-grained and less deformed parts of the rock (Ebert et al., 2007; Linckens et al., 2011). Thus, the rheological feedbacks due to CPO development are probably more important in the earlier stages of moderate rock deformation (Durham & Goetze, 1977; Ebert et al., 2007; Hansen et al., 2012, 2012; Zhang & Karato, 1995), while the effect of grain size reduction dominates for extreme deformations evident in mylonites. Moreover, seismic anisotropy (which is indicative of CPO) occurs within plate interiors, and not just at plate boundaries (Long & Becker, 2010; Nicolas & Christensen, 1987; Tommasi et al., 2000). However, upon extreme deformation exclusive to plate boundaries, mineral grains recrystallize to smaller sizes, erasing the preexisting viscous anisotropy in the process (Boneh et al., 2017; Ebert et al., 2007; Linckens et al., 2011). Finally, while viscous anisotropy may induce about an order of magnitude weakening (Hansen et al., 2012, 2012) for rocks deforming by dislocation creep, grain size reduction, and the resulting activation of grain size-sensitive creep may allow for weakening over multiple orders of magnitude (Bercovici & Ricard, 2003, 2012; Karato et al., 1980; Mulyukova & Bercovici, 2017; Poirier, 1980; Van der Wal et al., 1993).

The mechanisms responsible for mylonitization and grain size reduction are thought to play a key role in lithospheric shear localization and grain damage theory provides a theoretical framework for understanding these mechanisms. Specifically, the theory stipulates the positive feedback between rock deformation, grain size reduction and resultant mechanical weakening, which together may lead to lithospheric shear localization and ultimately plate boundary formation (Austin & Evans, 2007; Bercovici & Ricard, 2005, 2012, 2014, 2016; Rozel et al., 2011). In what follows, we outline the physics of grain damage theory, its applications to some leading questions in the origin, and operation of plate tectonics and lastly to recent developments and future directions.

2. Grain Damage Physics

2.1. Grain Damage in Single Phase Materials

A rock’s rheological behavior is determined by the micromechanics of the creep mechanisms by which it deforms under deviatoric or differential stress. In the deeper ductile portion of the lithosphere, rock deforms predominantly by diffusion and dislocation creep, whose rheological relations are given by, respectively,
Table 1

<table>
<thead>
<tr>
<th>Property</th>
<th>Symbol</th>
<th>Value/Definition</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas constant</td>
<td>$R_G$</td>
<td>8.3144598</td>
<td>J·K$^{-1}$·mol$^{-1}$</td>
</tr>
<tr>
<td>Shear modulus</td>
<td>$G$</td>
<td>70</td>
<td>GPa</td>
</tr>
<tr>
<td>Length of Burgers vector</td>
<td>$b$</td>
<td>0.50</td>
<td>nm</td>
</tr>
<tr>
<td>Surface tension</td>
<td>$\gamma$</td>
<td>1</td>
<td>J/m$^2$</td>
</tr>
<tr>
<td>Phase volume fractiona</td>
<td>$\phi_i$</td>
<td>$\phi_1 = 0.4$, $\phi_2 = 0.6$</td>
<td></td>
</tr>
<tr>
<td>Phase distribution functiona</td>
<td>$\eta$</td>
<td>$3\phi_1\phi_2 \approx 0.72$</td>
<td></td>
</tr>
<tr>
<td>Shear modulus</td>
<td>$G$</td>
<td>70</td>
<td>GPa</td>
</tr>
<tr>
<td>Length of Burgers vector</td>
<td>$b$</td>
<td>0.50</td>
<td>nm</td>
</tr>
<tr>
<td>Surface tension</td>
<td>$\gamma$</td>
<td>1</td>
<td>J/m$^2$</td>
</tr>
<tr>
<td>Phase volume fractiona</td>
<td>$\phi_i$</td>
<td>$\phi_1 = 0.4$, $\phi_2 = 0.6$</td>
<td></td>
</tr>
<tr>
<td>Phase distribution functiona</td>
<td>$\eta$</td>
<td>$3\phi_1\phi_2 \approx 0.72$</td>
<td></td>
</tr>
</tbody>
</table>

Dislocation creep$^b$

$$\dot{\varepsilon}_{\text{disl}} = A\tau^n$$

Activation energy

$$E_{\text{disl}} = 530$$

Prefactor

$$A_0 = 1.1 \cdot 10^5$$

Stress-Exponent

$$n = 3$$

Compliance

$$A = A_0 \exp(-\frac{E_{\text{disl}}}{RGT})$$

Diffusion creep$^b$

$$\dot{\varepsilon}_{\text{diff}} = B r^{-m}$$

Activation energy

$$E_{\text{diff}} = 300$$

Prefactor

$$B_0 = 13.6$$

Grain size exponent

$$m = 3$$

Compliance

$$B = B_0 \exp(-\frac{E_{\text{diff}}}{RGT})$$

Grain growth$^c$

Activation energy

$$E_G = 200$$

Prefactor

$$G_0 = 2 \cdot 10^4$$

Exponent

$$p = 2$$

Grain growth rate

$$G_G = G_0 \exp(-\frac{E_G}{RGT})$$

Interface coarsening$^d$

Exponent

$$q = 4$$

Interface coarsening rate

$$G_I = \frac{q}{p} (\mu m)^{q-p} \frac{G_0}{250}$$

<table>
<thead>
<tr>
<th>Property</th>
<th>Symbol</th>
<th>Value/Definition</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation energy</td>
<td>$E_{\text{disl}}$</td>
<td>530</td>
<td>kJ/mol</td>
</tr>
<tr>
<td>Prefactor</td>
<td>$A_0$</td>
<td>$1.1 \cdot 10^5$</td>
<td>MPa$^{-n}$·s$^{-1}$</td>
</tr>
<tr>
<td>Stress-Exponent</td>
<td>$n$</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Compliance</td>
<td>$A$</td>
<td>$A_0 \exp(-\frac{E_{\text{disl}}}{RGT})$</td>
<td>MPa$^{-n}$·s$^{-1}$</td>
</tr>
<tr>
<td>Activation energy</td>
<td>$E_{\text{diff}}$</td>
<td>300</td>
<td>kJ/mol</td>
</tr>
<tr>
<td>Prefactor</td>
<td>$B_0$</td>
<td>13.6</td>
<td>$\mu$m$^m$·MPa$^{−1}$·s$^{−1}$</td>
</tr>
<tr>
<td>Grain size exponent</td>
<td>$m$</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Compliance</td>
<td>$B$</td>
<td>$B_0 \exp(-\frac{E_{\text{diff}}}{RGT})$</td>
<td>$\mu$m$^m$·MPa$^{−1}$·s$^{−1}$</td>
</tr>
<tr>
<td>Activation energy</td>
<td>$E_G$</td>
<td>200</td>
<td>kJ/mol</td>
</tr>
<tr>
<td>Prefactor</td>
<td>$G_0$</td>
<td>$2 \cdot 10^4$</td>
<td>$\mu$m$^p$/s</td>
</tr>
<tr>
<td>Exponent</td>
<td>$p$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Grain growth rate</td>
<td>$G_G$</td>
<td>$G_0 \exp(-\frac{E_G}{RGT})$</td>
<td>$\mu$m$^p$/s</td>
</tr>
<tr>
<td>Interface coarsening rate</td>
<td>$G_I$</td>
<td>$\frac{q}{p} (\mu m)^{q-p} \frac{G_0}{250}$</td>
<td>$\mu$m$^q$/s</td>
</tr>
</tbody>
</table>

$^a$Phase distribution function from Bercovici and Ricard (2012), assuming a peridotite mixture with volume fractions of 40% and 60% pyroxene and olivine, respectively.

$^b$Our model values for rheological parameters are representative of experimentally determined olivine creep laws (Karato & Wu, 1993; Hirth & Kohlstedt, 2003).

$^c$Olivine grain growth law from Karato (1989); note that the value for $E_G$ has been debated (e.g., Evans et al., 2001).

$^d$Interface coarsening law from Bercovici and Ricard (2013), based on the analysis done in Bercovici and Ricard (2012).

\[
\dot{\varepsilon}_{\text{diff}} = \frac{B}{R m} \tau \quad (1)
\]

\[
\dot{\varepsilon}_{\text{disl}} = A\tau^n \quad (2)
\]

where $\dot{\varepsilon}$ and $\tau$ are the square root of the second invariant of the strain rate and stress tensors, $B$ and $A$ are the diffusion and dislocation creep compliances, $m$ and $n$ are the grain size and the stress exponents, respectively, and $R$ is the mean grain size of a rock sample (see Table 1 for typical parameter values). When deforming by diffusion or dislocation creep, the rock’s strength is strongly affected by temperature, which enters the relations (1) and (2) through the rheological compliances $B$ and $A$ (see Table 1). The effect of temperature on lithospheric strength is profound: cooling by a few hundred Kelvin can induce orders of magnitude increase in viscosity, with the exact numbers depending on the activation energies of the creep mechanisms. The peak strength in the ductile lithosphere, and thus the main impediment to plate boundary formation, falls within the temperatures 800–1400 K (Kohlstedt et al., 1995). The increase in pressure with depth across the lithosphere is assumed to have a smaller effect on rheology than that of temperature and is therefore not considered (i.e., we set the activation volumes of $B$ and $A$, shown in Table 1, to zero), for simplicity. Viscosity is defined as $\tau/(2\dot{\varepsilon})$, which for diffusion creep is sensitive to grain size $R$, whereby grain size reduction leads to softening (1). In contrast, dislocation creep viscosity is grain size insensitive, but depends nonlinearly on stress, such that an increase in stress leads to a decrease in viscosity.
For a given set of deformation conditions (i.e., temperature, stress, grain size, etc.), it is common to assume that the dominant creep mechanism is the one that induces fastest strain, or is most efficient at releasing stress. In a simple rheological model where only two creep mechanisms are considered, diffusion and dislocation creep, this assumption leads to the following composite rheological law:

\[ \dot{\varepsilon} = A \tau^n + \frac{B}{R^m} \tau = A \tau^n \left( 1 + \left( \frac{R_F}{R} \right)^m \right) \]  

(3)

where we defined the rheological field boundary grain size

\[ R_F = \left( \frac{B}{A \tau^n - 1} \right)^{1/m} \]  

(4)

at which \( \dot{\varepsilon}_{\text{diff}} \) and \( \dot{\varepsilon}_{\text{disl}} \) are equal. At a given stress, grains that are smaller than \( R_F \) deform predominantly by diffusion creep, and grains larger than \( R_F \) deform by dislocation creep.

For our chosen rheological law (3), the rate at which external forces do work to deform a rock is

\[ \Psi = 2 \dot{\varepsilon} \tau = 2A \tau^n \left( 1 + \left( \frac{R_F}{R} \right)^m \right) \]  

(5)

There are two ways in which deformational work can be deposited: as irrecoverable energy, in which work is dissipated as heat and contributes to entropy production, and as recoverable energy, in which work is stored as new grain boundary (or surface) energy upon formation of new grains and subgrains (Austin & Evans, 2007; Bercovici & Ricard, 2005; Ricard & Bercovici, 2009). The splitting of grains into new smaller grains is part of the dynamic recrystallization (DRX) process, which is accommodated by the migration and accumulation of dislocations to form new grain boundaries, which eventually leads to a reduction in average grain size. Storing mechanical work as new surface energy on grain boundaries is referred to as grain damage.

While deformational work acts to reduce grain size by damage, it is counteracted by the process of grain growth. Grain growth occurs because smaller grains have higher internal energies than larger grains (due to the small grains’ higher grain boundary curvature and surface tension, which imbues them with greater internal pressure and chemical potential); this energy contrast drives mass diffusion out of the smaller grains and into the larger ones, causing an increase in average grain size. This process of normal grain growth is equivalent to healing, since it effectively reverses the grain size reduction by damage.

Although rock strength is partially governed by average grain size, grain evolution actually occurs to the grain size distribution, whereby changes in the individual populations of grain sizes affect whether the average grain size goes up (net grain coarsening) or down (net grain size reduction). To simplify the analysis of grain size evolution, Rozel et al. (2011) approximated the grain size distribution with a self-similar lognormal function, which retains the same shape through time such that its mean, variance, and amplitude are uniquely determined by one characteristic grain size \( R \) (see Ricard & Bercovici, 2009 for the full analysis of grain size evolution, where the self-similarity assumption is relaxed).

The resulting theoretical model of grain size evolution in a continuum is derived from the considerations of energy conservation (i.e., how deformational work is distributed between dissipated heat and new grain boundaries) and the positivity of entropy production (i.e., the second law of thermodynamics; Ricard & Bercovici, 2009; Rozel et al., 2011) and leads to

\[ \frac{dR}{dt} = \frac{G_G}{p R^{p-1}} - \lambda R^2 f 2A \tau^{n+1} \frac{3}{\gamma} \]  

(6)

The first term on the right side of (6) describes coarsening by grain growth, where \( G_G \) is a temperature-dependent rate and \( p \) the exponent of grain growth, respectively. When no deformational work is done on the system (i.e., \( \Psi = 0 \)), the standard static grain growth relation is recovered (Evans et al., 2001), and the range of values for \( G_G \) and \( p \) have been determined by fitting experimental data for grain growth to this power law relation (see Table 1).

The second term on the right side of (6) describes grain size reduction by damage, where \( \gamma \) is the surface tension (see Table 1), \( \lambda = \exp \left( \frac{5e_1}{2} \right) \) is specific to the lognormal grain size distribution (Rozel et al., 2011, with
dimensionless variance \( \sigma \) typically set to 0.8), and \( 2A r^{n+1} \) is the amount of mechanical work that goes toward creating new grain boundary area and energy. New grain boundaries are created through DRX, in which dislocations merge to form subgrains and eventually new grains. DRX can only occur in grains where the strain is predominantly accommodated by dislocation creep. Therefore, out of the total mechanical work \( \Psi \), only the work done by dislocation creep \( (2A r^{n+1}) \) can be used to form new grains (Rozel et al., 2011), and out of this work only a fraction \( f \), called the damage partitioning fraction, goes toward creating new grain boundary area and energy.

Grain size reaches a dynamical equilibrium when the rates of coarsening and damage are in balance. As inferred by (6), the steady state grain size is a function of the deformational work rate and can therefore serve as a paleowattmeter (Austin & Evans, 2007; Rozel et al., 2011): that is, the measured grain size can be translated into work rate, based on the theoretical considerations above. Analogously, the observed grain size can be translated into the driving stress, thus serving as a paleopiezometer, based on the empirically deduced correlation between grain size and stress (Jung & Karato, 2001; Karato et al., 1980; Post, 1977; Van der Wal et al., 1993). These piezometric relations were used in Rozel et al. (2011) and Mulyukova and Bercovici (2017) to determine the range of possible values for \( f \), by comparing them to the steady state grain size predicted by the grain damage model for different stress and temperature conditions; \( f \) is found to be temperature dependent and, starting off roughly at 0.1 for coldest temperatures of 800 K, decreases several orders of magnitude with temperature increasing to 1600 K, depending on the assumed activation energies for grain growth and rheological compliances, among others.

As noted earlier, shear localization by grain damage requires a positive feedback whereby grain size reduction induces weakening, which focuses deformation, which in turn accelerates grain size reduction, and so on. However, in monomineralic materials, grain size reduction occurs during dislocation creep (2), while the necessary weakening by grain size sensitive (GSS) viscosity only occurs during diffusion creep (1). That these two creep mechanisms are somewhat exclusive (except over a narrow range of grain sizes near the rheological field boundary, e.g., by dislocation accommodated grain boundary sliding; see; Hansen et al., 2011; Hirth & Kohlstedt, 1995) precludes the coexistence of grain reduction and self-weakening, and hence the localizing feedback leading to mylonites. However, actual lithospheric rocks consist of at least two phases (olivine and pyroxene), and the presence of secondary phases is known to have a significant effect on grain size evolution and the resulting material strength (Herwegh et al., 2005, 2011; Linckens et al., 2011, 2015). Moreover, field studies suggest that mylonitization preferentially occurs in the polymineralic domains of the rocks, especially for upper mantle peridotites (Skemer et al., 2010; Warren & Hirth, 2006). To that end, Bercovici and Ricard (2012) developed the two-phase grain damage theory to infer how the interaction between phases facilitates grain reduction, inhibits healing and enhances shear localization.

### 2.2. Grain Damage in Polymineralic Materials

The physical description of a grained two-phase material involves the volume fraction \( \phi_i \) and the mean grain size \( R_i \) for each of the phases (where the subscript \( i = 1 \) or 2 denotes the individual phases, e.g., \( i = 1 \) is for the secondary phase like pyroxene, and \( i = 2 \) is for the primary phase like olivine), which are both functions of space and time. When the phases are unmixed, the interface that separates them is smooth, meaning that it has a large radius of curvature \( r \) (\( r \) is also referred to as interface roughness), and the interface area is minimized. When the phases are well mixed, in which one phase is well dispersed through the other phase, the interface roughness \( r \) is also referred to as interface roughness, and the interface area is minimized. When the phases are well mixed, in which one phase is well dispersed through the other phase, the interface is rougher and has a small radius of curvature \( r \). The two-phase grain damage theory tracks the coupled evolution of the interface roughness \( r \) and the grain size \( R_i \) of each of the phases \( i \) (see; Bercovici & Ricard, 2012):

\[
\frac{dR_i}{dt} = \frac{G_i}{pqr^{n+1}} Z_i - \frac{\lambda R_i^2}{3r_i} f_i 2A_i r_i^{n+1} Z_i^{-1}
\]

(8)

where \( G_i \) (\( G \)) and \( q \) (\( p \)) are the rate and the exponent of interface (grain boundary) coarsening, respectively, \( Z_i \) is the Zener pinning factor, to be discussed shortly, and \( r_i \) and \( \gamma_i \) are the interface and the grain boundary surface energies, respectively. The phase distribution function \( \eta = 3\phi_1\phi_2 \) ensures that the interface area vanishes in the limit of \( \phi_i \to 0 \) or 1 (Bercovici et al., 2001, Section 2.2). The partitioning fractions \( f_i \) and...
Figure 1. Cartoon of how Zener pinning distortion increases the mean curvature and surface and internal energy of large grains and thus impedes grain growth driven by energy contrasts (a) and facilitates grain boundary damage (b).

\( f_G \) determine how much of the deformational work is partitioned into new interface and grain boundary area, respectively (see also Bercovici & Ricard, 2016), where the work to produce more grain boundary area is restricted to dislocation creep (see Section 2.1). The work to produce more interface area is tapped from the phase-volume average of the total work rate (\( \dot{\Psi} = \sum \phi_i \dot{\Psi}_i \)), since any mode of deformation will cause interface distortion by stretching, rending or phase mixing. The range of possible values for \( f_I \) and \( f_G \) were determined in Bercovici and Ricard (2016) and Mulyukova and Bercovici (2017), by comparing the theoretically predicted steady state grain sizes to the ones reported from experimental (Jung & Karato, 2001; Karato et al., 1980; Linckens et al., 2011, 2015) and field data (e.g., Herwegh et al., 2011; Linckens et al., 2011, 2015) for different stress and temperature conditions. As before, the factor \( \lambda \) is specific to the lognormal grain size distribution.

The Zener pinning factor \( Z_i \) in (8), which couples the grain size and the interface roughness evolution, is defined as (Bercovici & Ricard, 2012):

\[
Z_i = 1 - h_i \left( \frac{R_i}{r} \right)^2
\]

where

\[
h_i = \lambda^* (1 - \phi_i)
\]

accounts for the fact that the pinning of grains of phase \( i \) depends on the presence of the other phase, and where \( \lambda^* = (3/160) \exp(6\sigma^2) \) depends on the assumed shape of the grain size distribution (Bercovici & Ricard, 2012; Appendix F.4).

Grain evolution in each phase (8) is similar to that for the single phase case (Section 2), except that they are now affected by the interface between the phases. One of the interface effects involves the obstruction of grain growth (Figure 1a). Specifically, grain growth arises from the grains’ tendency to lower their internal energy: larger grains have lower grain boundary curvature and lower surface tension and therefore are more energetically favorable. However, if a grain can only grow by wrapping itself around an obstacle (i.e., a grain of a different phase) or by wedging itself between two grains of opposite phase, then the associated grain boundary distortion involves an increase in interface area and curvature (i.e., a decrease in \( r \)) that requires additional surface energy, and therefore no longer becomes energetically favorable. Obstruction of grain growth by the interface is also known as Zener pinning; as \( r \) becomes smaller and \( Z_i \) approaches 0 (9), pinning impedes coarsening (first term in (8)), as is classically inferred for Zener pinning theory (e.g., Smith, 1948; Hillert, 1965; 1988; Manohar et al., 1998). While the classical approach of Zener pinning assumes that grain boundary migration is impeded due to small particles and impurities on the grain boundaries, (2012; Appendix C.3) demonstrated that mineral assemblages with commensurate volume fractions and grain sizes experience an analogous effect. The smaller the secondary phase particles, and thus the radius of curvature of the interface, the stronger the effect of Zener pinning (Bercovici & Ricard, 2012, and references therein).

The second effect of Zener pinning on grain size evolution, as shown by Bercovici and Ricard (2012), is that when \( Z_i \) decreases, the grain size reduction by damage gets amplified. The theory thus predicts that pinning
amplifies the contribution of dislocation creep through DRX to grain boundary damage (second term in (8)), even as the grains shrink into the diffusion creep regime and dislocation creep is no longer the dominant deformation mechanism (Figure 1b). Specifically, new grain boundary energy is created from some fraction of deformatonal work when grains split into smaller grains by DRX, and the amount of energy stored in the process is equal to the difference in total surface energy between the grains before and after splitting. If, prior to splitting, a grain’s boundary is distorted by Zener pinning, then its surface energy is elevated relative to when its boundary is undistorted. The smaller grains created after splitting have less boundary distortion by Zener pinning because they are comparable in size, as well as being in less contact with, pinning bodies; hence, their surface energy is less elevated by pinning. As a result, the difference in net surface energy between grains before and after grain splitting is smaller than the case without Zener pinning. Therefore, it requires less energy to cause grain splitting in a two-phase medium, compared to the single-phase case, which suggests that grain size reduction occurs more readily in polymineralic rocks.

In total, (7) and (8) show that in a polymineralic medium, grain damage is expressed in two forms, grain boundary damage (i.e., reduction in $R_i$) and interface damage (i.e., reduction in $r$). Specifically, a fraction of work can be stored both as grain boundary energy within a given phase (as in the monomineralic case), as well as interface energy between phases, through rending, stretching and stirring of the interface when the rock deforms. The interface between phases gets distorted and its radius of curvature $r$ decreases, while its area and energy increase at the cost of deformational work. Importantly, work can be transformed into interface damage regardless of creep mechanism, since such damage only requires that the medium deforms, one way or another. Thus, the roughness of the interface $r$, or the size of the pinning bodies, can shrink to sizes below the field boundary grain size, into the GSS diffusion creep regime. Shrinking pinning bodies drive down the grain size with them, through the dual Zener pinning effects, that is, enhancing damage by DRX even for small grains and suppressing grain growth (8). Interface damage thus allows for grain size reduction and grain size-sensitive rheology to coexist, providing a positive feedback mechanism for self-weakening of polymineralic rocks.

The deformatonal work rate in (7) and (8) uses similar rheological relations to the composite rheology (3), but for each phase $i$:

$$\dot{\varepsilon}_i = A_i \tau_i^n + B_i \frac{\tau_i}{R_m}$$

where $A_i$ and $B_i$ are again the rheological compliances for dislocation and diffusion creep, respectively, $\dot{\varepsilon}_i$ and $\tau_i$ are the square root of the second invariant of the strain rate and stress tensors, respectively, and $R_i$ is the mean grain size, with the grain size distribution assumed to be lognormal in each of the phases.

Grain size partially governs the rheological properties of the rock, as it determines the creep regime (4) and the rock’s viscosity in the GSS regime (1). Thus, the microphysics of grain size evolution affects the macroscopic processes of rock deformation. At the same time, the grain size evolution laws (in single-phase (6), as well as two-phase (7) and (8) materials) depend on physical parameters (such as stress and temperature) that are governed by continuum-scale processes. In a full system of equations that describe rock deformation, the microscale and macroscale processes are connected by coupling the governing continuum equations, including conservation of mass, momentum, and energy, through the rheological relations, to the grain size evolution laws (for monomineralic (6) or polymineralic (7) and (8) materials). The two-phase continuum model includes conservation laws for both phases (Bercovici & Ricard, 2012). The complete system of coupled equations can be applied to geodynamic models featuring grain damage (see examples in Section 3).

### 2.3. Pinned State Limit

The two-phase grain damage model predicts that as the grain size $R_i$ in each phase and the interface roughness $r$ evolve, they approach relative sizes in which $Z_i \to 0$ (the exact size-ratios being determined by the value of $h_i$ (10)), as is also supported by experimental observations of Hiraga et al., 2010 (2010; at least after the system has had some time to evolve, see; Bercovici & Ricard, 2012). Without any deformation (i.e., when the second term in (8) is 0), pinning by definition would be such that $R_i$ grows until $Z_i$ reaches 0, and then stops. But with deformation, the damage term grows as $Z_i$ shrinks (8), and thus the grains shrink, potentially below the field boundary grain size, following the decrease in interface roughness $r$ (7). In the asymptotic limit $Z_i \to 0$, known as the pinned state limit, the grain size becomes slaved to $r$, as the grains within each phase are blocked by the presence of the other phase. To model grain size evolution in the pinned state limit,
it is sufficient to only track the evolution of $r$ (7), while the grain size in each phase can be approximated by the solution to $Z_i \approx 0$, or specifically

$$R_i \approx \frac{r}{\sqrt{h_i}}$$

(12)

Due to its simplicity, relative to the full set of equations (7) and (8), the pinned state limit is frequently used in geodynamic applications of grain damage theory (Bellas et al., 2018; Bercovici et al., 2015, 2018; Mulyukova & Bercovici, 2017, 2018).

3. Applications

3.1. Generation of Plate Tectonics

The physical models of lithospheric rock deformation can be used to explain several important aspects of plate tectonics. Besides being tessellated into broad strong plates that move relative to each other, the Earth’s surface motion is peculiar in that the poloidal and toroidal components of its velocity field contain comparable amounts of kinetic energy (Bercovici et al., 2015; Hager & O’Connell, 1978, 1979, 1981; O’Connell et al., 1991; Ribe, 1992). The poloidal flow component is associated with convective motion and distinguished by convergent and divergent surface flow, which accommodate mantle upwelling and downwelling. The large contribution to surface flow from toroidal motion, which is associated with strike-slip shear (as well as a modest amount of plate spin) is more enigmatic, since it does not have a direct energy source (like release of gravitational potential energy by convection). The generation of plate like flow from mantle convection requires nonlinear rock rheology to indirectly couple toroidal flow to convective motion (Bercovici, 1993; 1995; Bercovici et al., 2015; Cadek & Ricard, 1992; Kaula, 1980), as well as its reduction of viscous dissipation in mantle and lithospheric flow (Bercovici, 1995). The strongly localizing effect of damage on the rheology of polymineralic materials, according to the two-phase grain damage theory, as opposed to the weaker localizing effect of the standard dislocation creep or grain damage in monomineralic materials, has demonstrably given rise to surface deformation with Earth-like toroidal and poloidal components (Bercovici & Ricard, 2015).

Deformation at plate boundaries occurs in a wide range of modes, including convergent, divergent, and strike-slip motion, as well as in the form of active and passive margins. Moreover, tectonic plate boundaries are long-lived: they can become dormant and reactivated again in response to changes in the underlying mantle flow. The fact that the forces that drive surface motions arise from the same source, that is, the large-scale convective mantle flow, suggests that at least some of the observed variety is due to the heterogeneity in the lithosphere itself and that this heterogeneity is long-lived. Variations in the strength of the lithosphere induced by changes in grain size are preserved for long periods of time, and can affect subsequent deformation. Using a simple model of lithospheric deformation driven by time-variable pressure field (i.e., a parameterized slab pull), Bercovici and Ricard (2014) demonstrated that as the lithosphere inherits one damaged weak zone after another, it can eventually form a network of localized weak zones, which appear as the full range of tectonic plate boundaries, including passive spreading centers, and passive strike-slip boundaries, as well as active subduction (Figure 2). The model of Bercovici and Ricard (2014) predicts that there is a time lag of about 1 Gyr between the initiation of subduction, and the global network of plate tectonics, in agreement with the sparse geological record (Condie & Kröner, 2008; Korenaga, 2013).

3.2. Collapse of Passive Margins

The thermal, mechanical and microstructural evolution of the lithosphere, and thus its susceptibility to strain localization, happens through several interacting processes, including thermal diffusion as the planet cools to space, viscous creep due to gravitational forcing, and the grain size evolution as the material adjusts to changing stress and temperature conditions. For example, when a new oceanic lithosphere forms at the ridges, it cools and stiffens through time. However, cooling also means that it takes longer to heal any structural damage (e.g., to reverse grain damage via grain growth) that the rock may have acquired while it was still hot (e.g., shortly after it was emplaced at the surface). In particular, if the mantle below the ridge has been thermally insulated by the continents prior to onset of spreading, its accumulated buoyancy and resulting vigorous flow can induce a strong drag on the newly forming plates (Gurnis, 1988; Lenardic et al., 2011), thus generating significant damage while the lithosphere is still young. Subsequent cooling and retarded healing help to keep the damaged portion of the lithosphere weak. Moreover, as the thickening plate moves away from the ridge, it subsides as it progresses toward its isostatic equilibrium and experiences an
Figure 2. A simple example of how plate regions weakened by grain damage may originate, persist, get inherited and eventually form an interconnected network, analogous to plate boundaries. Here, we show a fluid dynamical model of Bercovici and Ricard (2014) of lithospheric flow in a 2-D horizontal layer driven by a subduction-like pressure (schematic frame). The strength of the layer is governed by the two-phase grain damage and pinning model (Bercovici & Ricard, 2012, 2013). The pressure low $P$ is imposed for $\sim 10$ Myrs and rotated 90° about a vertical axis 3 times as an idealization of intermittent subduction during the early Archean. Weak, damaged bands induced by subduction from a previous orientation are inherited and amplified by the lithospheric flow of the next orientation, resulting in plate-like velocity $v$, localized but passive bands of strike-slip vorticity $\Omega$ and positive divergence $S$ (red and yellow contours), and a contiguous low-viscosity zone ($\bar{\mu}$), driven by one subduction zone (see; Bercovici & Ricard, 2014 for more details).

Increasingly high stress associated with ridge push. The combined effect of thermal stiffening, slow recovery from damage-induced weakening, and the increasing stress from ridge push determine the time window within which the thickening oceanic plates can undergo strain localization and weakening. A weakened plate is more likely to yield to gravity and potentially lead to the collapse of passive margins, as is, for example, hypothesized for spontaneous subduction initiation (Mulyukova & Bercovici, 2018). At modern Earth conditions, the grain damage theory predicts that the passive margins collapse within 100 Myr or do not collapse at all (Mulyukova & Bercovici, 2018), which is consistent with the observation that many present-day passive margins are older than 100 Myr (Müller et al., 2008). Similar analysis shows that formation of new trenches by the collapse of passive margins occurred more readily on early Earth, facilitated by the hotter mantle and a larger thermal contrast across the lithosphere (thus stronger ridge push); meanwhile, spontaneous subduction initiation with the help from grain damage is harder on Venus, due to its hotter conditions (thus faster grain growth and healing) and an arguably lower temperature drop across the lithosphere (Mulyukova & Bercovici, 2018).

### 3.3. Slab Detachment and Abrupt Reorganization of Plate Motions

The geological record, as well as current observations of the Earth’s surface deformation, show that tectonic plate motions can change much more rapidly than what can be expected from sluggish mantle flow. Perhaps one of the most notable examples is the rapid change in motion of the Pacific plate about 50 Myr ago, resulting in the bend in the Hawaiian-Emperor seamount chain (Sharp & Clague, 2006; Whittaker et al., 2007). With subduction being arguably the main driving force for plate motions (Forsyth & Uyeda, 1975), rapid changes in slab dynamics can induce dramatic changes in plate’s speed and/or direction. For example, necking and subsequent detachment of a slab from its parent plate can lead to abrupt changes in the plate’s motion and/or uplift. Such rapid strain localization, which leads to slab necking and detachment, has been shown to occur in simple slab models (Bercovici et al., 2015) and more complex numerical models (Bellas et al., 2018) that include grain damage self-weakening. Furthermore, abrupt slab detachment due to weakening by grain damage can exacerbate trench rollback, as has been demonstrated in simple models that
Figure 3. Evidence for grain mixing and hysteresis from the field data (a and b), theory (c), and laboratory experiments (d). Petrographic thin section of a sheared lherzolite xenolith (from; Skemer & Karato, 2008) (a) and a sketch of the inset region derived from EBSD mapping (Bruijn & Skemer, 2014) (b), showing grain mixing via the migration of small olivine grains between orthopyroxene grain boundaries during deformation. (C) A theory for diffusive grain mixing coupled to two-phase grain damage (Bercovici & Mulyukova, 2018) uses a micromechanical model for mechanical grain mixing (Bercovici & Skemer, 2017) to model diffusivity; it predicts that two deformation or “piezometric” branches arise, similar to the hysteresis grain damage model of Bercovici and Ricard (2016), in which well-mixed polyphase regions are characterized by small grains deforming via diffusion creep typical of ultramylonites, while poorly mixed, monophase regions contain large grains deforming in dislocation creep. (d) Deformation map for calcite-anhydrite experiments of Cross and Skemer (2017), showing that when calcite is unmixed with anhydrite it follows a normal large grain size calcite piezometer, but when well mixed it follows a branch with smaller grain sizes, similar to that predicted by theory.

In summary, plate boundary formation and evolution involves localized deformation of the lithosphere, which is controlled by its rheological properties. Rheological behavior is, in turn, governed by different physical properties, such as temperature, chemistry and microstructure (see Bercovici et al., 2015 for a recent review). In addition to being strongly localizing (to allow for significant toroidal motion) and able to occur in deep lithosphere (to allow weakening across the entire plate, including its strongest region), the effect of microstructure on rheology is uniquely long-lived, allowing for longevity and inheritance of lithospheric weak zones, which is likely a requirement for the formation of tectonic plate boundaries.

4. Newest Directions
4.1. Mixing and Hysteresis
Self-weakening and strain localization by grain damage is conditional on the coexistence of grain size reduction with grain size sensitive deformation, such as occurs in diffusion creep (Section 2). In order for grain size to shrink below the field boundary from dislocation to diffusion creep regimes, DRX must be enabled even when dislocation creep is not the dominant deformation regime. In the two-phase grain damage theory, the secondary phases and the associated distortion of the grain boundaries through Zener pinning make grain size reduction possible well into the diffusion creep regime (Section 2.2). The effectiveness of Zener pinning, however, is controlled by the degree to which the different phases are interspersed or mixed with each other at the grain scale. The process of phase mixing, and its effect on rock rheology and shear localization, have been the focus of numerous recent experimental and theoretical studies (Bercovici & Mulyukova, 2018; Bercovici & Ricard, 2016; Bercovici & Skemer, 2017; Cross & Skemer, 2017; Tasaka et al., 2017; Tasaka et al., 2017; Wiesman et al., 2018). At low pressures (i.e., <1 GPa), intergrain mixing may occur by chemical or metamorphic reactions (Platt, 2015; Czertowicz et al., 2016; Précigout & Stünitz, 2016) and brittle processes (Dimanov et al., 2007), all of which typically involving cavitation. At higher pressures representative of the middle and lower lithosphere, grain mixing may instead be accommodated by dissolution and precipitation of one phase on the grain boundaries of the other phase (Tasaka et al., 2017, 2017; Wiesman et al., 2018), or percolation of one phase along the grain boundaries of the other phase, driven by capillary forces and deviatoric stress acting to dilate the grain boundaries (Bercovici & Skemer, 2017).

In lithospheric peridotites, the two prevalent phases are olivine (~60%) and pyroxene (~40%). Where two monophase units are in contact, there are “T”-shaped triple junctions where the grain boundaries of one phase intersect with the interface between the phases (Figure 3b). Pyroxene grains appear to recrystallize to smaller sizes than the olivine grains (by about a factor of 5, see; Linckens et al., 2014; Skemer & Karato, 2008),
and hence there are more triple junctions of pyroxene grain boundaries with the interface than occur for olivine grain boundaries, allowing more pathways for olivine to mix into pyroxene than vice versa. Bercovici and Skemer (2017) therefore modeled olivine material being driven along pyroxene grain boundaries by both capillary forces and imposed stresses, thereby forming protrusions, or so-called teeth; without any imposed stress, these teeth would form by the wetting of olivine along pyroxene grain boundaries and the tooth apex would be at the dihedral angle when all the boundary forces are in equilibrium. A given tooth can continue to grow, intruding into the pyroxene grain boundary, with the application of normal stress, which forces the pyroxene grain boundary to dilate (but not cavitate), effectively providing a low-pressure region ahead of the tooth that draws it forward. The imposed stresses also drive bulk deformation of both phases and geometrically necessary dislocations at the base of the olivine teeth. When enough of these dislocations have accumulated, they combine to form a subgrain boundary. The subgrain eventually detaches to form a new grain, or a tooth, introducing an increase in grain boundary energy, which is a manifestation of damage. The tooth is subsequently squeezed along to the next pyroxene triple junction. The size of a dislodged tooth is dictated by the competing rates of tooth growth (predominantly determined by the normal stress and the olivine viscosity) and damage and severance of the tooth base (dictated by the deformatonal work on the olivine grains). Once detached, a tooth can migrate toward other pyroxene triple junctions, under the continued application of normal stress. Olivine teeth mixed into the pyroxene matrix act as pinning agents (the size of the teeth are representative of the interface roughness in Section 2.2), restricting coarsening and facilitating damage of pyroxene grains through Zener pinning: the smaller the teeth, the more efficient the pinning and the smaller the resulting pyroxene grain size. Smaller pyroxene grains means that there are more “T”-junctions where the new teeth can form, and thus more efficient mixing. The process of forming, growing and shedding new olivine teeth repeats itself, causing the interface to migrate in the direction of the olivine grains that are loosing teeth, leaving behind a mixed region of olivine and pyroxene grains.

On the continuum scale, the mechanical mixing by teeth-spalling can be modeled as diffusion of the two phases through each other, with the evolving composition of the mixed region described in terms of the volume fractions of the two phases (Bercovici & Mulyukova, 2018). The diffusion rate of olivine into the pyroxene rich layer is dictated by the rate of spalling and the size of olivine teeth, thus bridging the microscopic and the continuum scale processes. The enhanced efficacy of Zener pinning in the mixed regions then impedes grain growth and amplifies grain damage (which is further augmented by the deformatonal work of mixing), thus driving grain reduction well into the diffusion creep regime, resulting in self-weakening and localization. In the case of a simple system of two monophase units subjected to uniaxial normal stress, a mixed layer between the units develops and approaches a steady state wherein the rate at which the pyroxene and olivine domains are squeezed toward each other by normal stress is balanced by the rate at which the two phases diffuse through each other. At lithosphere like stress and temperature conditions, the model predicts the formation of rapidly deforming strongly localized ultramylonitic bands embedded within coarser grained, poorly mixed or monomineralic, slowly deforming regions (Figure 3c). The coexistence of the well-mixed weak regions and the poorly mixed strong regions is in agreement with the hysteretic branches predicted by Bercovici and Ricard (2016), as well as the experimental observations of Cross & Skemer (2017; see also Figure 3d) and Wiesman et al. (2018). The effective rock viscosity, as determined by the deformation regime and the grain size in the GSS regime, is different for the different steady states: the rock is effectively stronger when its grain size is large, its phases are poorly mixed and it follows the classical piezometer, compared to when the mixing-localization feedback drives the grain size to mylonitic values deep in the diffusion creep regime. Which of the two possible steady states prevails—the classical piezometer or the (ultra-)mylonitic one—is determined by the deformation conditions (i.e., stress and temperature), the mixing regime (homogenized or diluted mixture) and the initial conditions (e.g., grain size and interface roughness). Importantly, different steady state grain sizes can be possible (depending on how well the two phases are mixed) at the same deformation conditions, giving rise to a hysteretic behavior, as was first hypothesized in Bercovici and Ricard (2016, see also Figure 4, left). The microphysical and continuum scale theories of phase mixing in polynomineralic materials, and its effect on grain size and strength of lithospheric rocks, predict that there exist a range of stresses and temperatures at which rocks exhibit hysteretic behavior, wherein two different deformatonal regimes can coexist: well-mixed, fine-grained, rapidly deforming, weak, and localized zones of shear, alongside unmixed or monomineralic, coarse-grained, strong, and slowly deforming zones. The inferred large range of grain
Figure 4. (left) The grain mixing transition model of Bercovici and Ricard (2016), which parameterizes grain mixing as found in Bercovici and Mulyukova (2018), leads to coexisting deformation branches in a hysteresis loop, shown in terms of (top) grain size $R_0$ and (bottom) strain rate $\dot{e}$ with increasing damage:healing ratio $qD/C$ (from; Bercovici & Ricard, 2016). Curves show the stable large-grained weakly deforming branch (blue), an unstable intermediate branch (green) and the small-grained rapidly deforming mylonite branch (red). Ranges of damage:healing ratios for Earth (light blue) and Venus (olive green) are indicated, suggesting that Earth can maintain coexisting slowly and rapidly deforming states characteristic of plate tectonics. While Venus is mostly in the slowly deforming state, it can make forays into tectonic states, but which are unstable given their proximity to the unstable intermediate branch. (right) Temperature-dependence of grain size for the Oman ophiolite predicted by Mulyukova and Bercovici (2017), according to the data from Linckens et al. (2015) (black and gray vertical lines, wherein black is for the secondary phase volume fraction $>0.3$), superposed with the theoretical piezometer obtained with the hysteresis model (solid black curve). The colored line is the field boundary grain size $R_{FB}$ colored according to temperature; the dashed black curve marks the mixing transition grain size $R_{m}$ below which the well-mixed, mylonitic branch becomes stable.

sizes expected to occur within ductile lithospheric shear zones is in agreement with the field observations of ophiolites, which exhibit grain scales from micron-sized ultramylonites to millimeter-sized tectonites in close proximity to each other (Mulyukova & Bercovici, 2017; see also Figure 4, right). In fact, such observations are difficult to reconcile with the classical piezometer or paleowattmeter models without invoking dramatic stress variations across small spatial scales.

The conditions representative of the Earth’s lithosphere put it well within the hysteresis loop that arises from phase mixing transitions, potentially explaining the coexistence of the strong and largely undeformed tectonic plate interiors and the weak, strongly localized plate boundaries. In contrast, the hotter surface conditions on Venus place it outside or just barely within the range of a hysteresis loop, permitting only a single slowly deforming state, making it unlikely for its lithosphere to deform in a plate-like manner, or at least for plate like states to be unstable and short-lived (Bercovici & Ricard, 2016), in keeping with the hypothesis of episodicity (Turcotte, 1993; Turcotte et al., 1999).

4.2. The Role of Intragranular Defects in Grain Damage Models

Grain damage theory posits that a small fraction of mechanical energy in a deforming rock is transformed into new surface energy through formation of new, smaller grains by DRX. The two-phase grain damage theory also predicts that grain size reduction by damage occurs more readily with increasing degree of mixing in polymineralic materials due to the effect of Zener pinning and the resulting grain boundary distortion (Sections 2.2 and 4.1). The theoretical prediction of enhanced grain damage due to pinning is a plausible candidate to explain how grains can shrink well below the field boundary grain size, as seen in polymineralic materials from experimental and field studies. In deforming crystalline materials, new grains are formed through the process of DRX.

Grain size can evolve by diffusive mass flux between the grains (also seen as grain boundary migration), as well as by grain splitting and merging (Ricard & Bercovici, 2009). Diffusive mass flux is induced by the difference in a grain’s internal energy relative to that of its neighboring grains. Both the grain boundary and intragranular defects (such as dislocations) contribute to a grain’s internal energy; therefore, the variation in either of these properties between different grains can induce thermodynamic forces that drive grain size evolution. For example, grains with higher boundary curvature (i.e., smaller size) are squeezed more by surface tension and thus have higher internal energy than larger grains, thus mass is driven from the
Figure 5. The current grain damage model only accounts for the energy transfer from deformational work to surface energy in order to break a big grain into small grains (see Figure 1b). However, in fact, damage and work at first go into dislocations as dislocation energy, which then transforms to greater surface energy (left). When grains of different sizes and different dislocation energy densities are in contact, the driving forces for grain growth, which are proportional to gradients or contrasts in internal energy, can oppose each other; for example, the contrast in surface tension, which is greater on small grains because of their higher curvature, drives mass from small to big grains (red arrow), while the contrast in dislocation energy, which is possibly greater in large grains (since dislocations are more readily formed in larger grains) drives mass in the opposite direction (blue arrow; right).

smaller grains to the larger adjacent grains, resulting in normal grain growth (e.g., Hillert, 1965; Lifshitz & Slyozov, 1961). Similarly, grains with high dislocation density have large internal energy and thus diffuse mass into grains with lower dislocation density, resulting in DRX by grain boundary migration (e.g., Derby & Ashby, 1987). The energetically driven processes of grain growth and DRX can compete and interact, as both grain size and dislocation density and their effects on a grain’s internal energy, evolve during deformation (Mulyukova & Bercovici, 2018, see also Figure 5); the evolution of dislocation density and internal energy have also been used to model the development of anisotropy (Kaminski & Ribe, 2001). While the basic (single- and two-phase) grain damage theory describes the thermodynamics (i.e., energy conservation and positivity of entropy production) governing grain size evolution (Section 2.2), its current formulation assumes that the migration and accumulation of dislocations by which DRX occurs is instantaneous; thus, the current theory does not capture the interaction of grain size evolution and dislocation dynamics. Here, we detail new developments that consider the effect of dislocations. Incorporating dislocation dynamics into grain damage theory allows for a prediction of dislocation density as a function of deformation conditions, which is a commonly reported metric in microstructural analysis of deformed rocks, along with grain size, and thus allows for additional comparisons between theory and observations (Green & Radcliffe, 1972; Durham & Goetze, 1977).

Mulyukova and Bercovici (2018) provide an instructive model using a simple system consisting of only two grains to illustrate the interaction between forces associated with the variations in dislocation density $\omega_i$ and grain boundary curvature $2/R_i$, where $R_i$ is again grain size and the subscript $i = 1$ or 2 refers to the individual grains. The underlying physical processes, however, are general for systems with virtually infinite number of grains, as was demonstrated in Mulyukova and Bercovici (2018), albeit with some simplifications.

In a deforming rock, dislocations can nucleate or elongate, move through the grain, or annihilate in collisions with other dislocations or grain boundaries; competition between these processes determines the evolution of dislocation density $\omega_i$ (e.g., Karato, 2008, pp. 154–157; Anderson et al., 2017, pp. 575–607). The dynamics of dislocations depends on the material’s crystal structure, external physical conditions (e.g., stress...
The theoretical analysis of (13)–(16) shows that large grains can have higher dislocation energy density than smaller grains, which acts to oppose and stabilize grain growth; specifically, the contrast in internal energy caused by grain boundary curvature between big and small grains is countered by the contrast in dislocation density. However, in smaller grains deformation is more readily accommodated by diffusion creep than by moving and growing dislocations, thus providing another limitation on dislocation density by grain size.

Using a simple parameterization of the combined effects of stress and grain size, and assuming for simplicity that the medium is isotropic, the rate of change of dislocation density \( \dot{\omega}_i \) in a deforming rock can be described as (see; Mulyukova & Bercovici, 2018):

\[
\dot{\omega}_i = C_2 \omega_i (a_i - \sqrt{\omega_i})
\]

(13)

\[
a_i = \sqrt{\omega_i} \frac{R_i/R_c}{1 + R_i/R_c}
\]

(14)

\[
\omega_i = \left( \frac{\tau}{G b} \right)^2
\]

(15)

where \( C_2 \) is a constant with dimensions of m/s, \( R_i \) defines the grain size below which the dislocation density becomes grain size sensitive (typically set to be equal to the field boundary grain size, which applies to either of the mechanisms that can suppress an increase in dislocation density described above), \( G \) is the shear modulus and \( b \) is the length of the Burgers vector (see Table 1). Furthermore, \( a_i \) (14) determines dislocation density at the steady state, since \( \dot{\omega}_i = 0 \) when \( \omega_i = a_i^2 \) (13).

A grain grows or shrinks depending on its size (i.e., grain boundary energy) and dislocation density relative to that of the neighboring grains. In the two-grains system, the energy contrast between neighboring grains gives rise to the driving forces for grain boundary migration. Moreover, the evolution of only one of the grains (in our case \( R_i \)) needs to be tracked, since the size of the other grain can be deduced from mass conservation. Assuming only diffusive mass exchange between the grains (i.e., no grain splitting or coalescence; see; Ricard & Bercovici, 2009), the evolution of \( R_i \) is given by (see; Mulyukova & Bercovici, 2018):

\[
R_i = \left( \frac{R_0^{-p} R_i^p C_1}{V/2} \right) \left[ (R_2^{-1} - R_i^{-1}) + C_0 (\omega_2 - \omega_i) \right]
\]

(16)

where \( C_0 \) and \( C_1 \) are constants with dimensions of m and m\(^{-2p}\)/s, respectively, \( V \) is the total volume of the two grains, and exponent \( p \) is generally set to \( p = 2 \), which corresponds to normal grain growth behavior (Ricard & Bercovici, 2009). (Note that the factor \( V/2 \) was unintentionally omitted in Mulyukova & Bercovici, 2018, equation (18).)

Theoretical analysis of (13)–(16) shows that large grains can have higher dislocation energy density than smaller grains, which acts to oppose and stabilize grain growth; specifically, the contrast in internal energy caused by grain boundary curvature between big and small grains is countered by the contrast in dislocation energy, thus impeding normal grain growth (Mulyukova & Bercovici, 2018).

The contrast in dislocation density needed to offset grain growth depends on the ratio between dislocation energy \( \gamma_d \) (which is energy per unit length of a dislocation) and grain boundary surface energy \( \gamma_b \), as represented by dimensionless parameter \( \alpha \equiv \gamma_d/R_b \), where \( R_b \) is a scaling grain size (e.g., \( R_b = 1 \) mm in; Mulyukova & Bercovici, 2018). The parameter \( \alpha \) plays an important role in the coupled evolution of grain size and dislocation density. For small \( \alpha \), coarsening proceeds more or less by normal grain growth, while for larger \( \alpha \), grain growth may stall or even undergo oscillations (Figure 6).

The influence of dislocations in opposing grain growth means that less deformational work is needed to create lithospheric weak zones by grain damage. In addition, the retarded grain growth increases the longevity of established weak zones, allowing for longer persistence of tectonic plate boundaries even after the deformation ceases.
Figure 6. Equilibrium grain size of one of two grains $R_1$ versus $\alpha$ (left) from Mulyukova and Bercovici (2018), in which $\alpha$ controls how the dislocation energy contrast between grains impedes grain growth. Colors of the branches indicate whether the equilibrium state is unstable (red), stable (blue), or oscillatory (green). For $\alpha < \alpha_{cr}$ the grain size is unstable in that one grain grows at the expense of the other grain. For $\alpha > \alpha_{cr}$ the dislocation energy contrast is sufficient to stabilize the grains and keep them from growing. For $\alpha > \alpha_{osc}$, the disparity in grain growth rate and the kinetics by which dislocation density equilibrates leads to oscillations in grain size with time (right).

The processes of diffusion-driven grain boundary migration (16) and dislocation kinetics (13) have different characteristic timescales, which can yield a time lag between the responses to changes in grain size and dislocation density (Mulyukova & Bercovici, 2018). As a result, the system of grains may undergo oscillations as it advances toward its microstructural steady state in grain size and dislocation density (Figure 6). Oscillations in grain size can induce transient rheological strengthening and weakening in cases when the material deforms in grain size sensitive creep regime (e.g., diffusion creep), as was also pointed out in Karato (1989).

At lithospheric conditions, the timescale of the oscillations is of the order of 1–10 years, similar to the rapid tectonic processes such as postseismic creep; this suggests that microstructural evolution and grain damage can possibly impose unique cycles during sequences of stress accumulation, release and recovery between earthquakes (Mulyukova & Bercovici, 2018). The new addition of dislocation dynamics to grain damage theory thus opens up new directions for understanding plate tectonics, not only over deep time but also over timescales that impact humans.

5. Summary and Future Directions

The grain damage theory of lithospheric deformation provides testable hypotheses to explain several important aspects of how Earth-like plate tectonics is generated. Specifically, the theory allows for strong shear localization of the deforming lithosphere, including in its strongest ductile region, through the positive feedback of grain size reduction and rheological weakening, which is required to generate new tectonic plate boundaries and to induce the observed significant toroidal component of plate motion. Taking into account the polymineralic nature of rocks and the different regimes of mixing between mineral phases, the two-phase grain damage theory predicts a hysteretic behavior of the Earth’s lithosphere, wherein weak, rapidly deforming, and strongly localized plate boundaries coexist with the strong and largely undeformed plate interiors at the same stress and temperature conditions. Applications of the grain damage theory include the formation and longevity of new plate boundaries by accumulation and inheritance of damage (Bercovici & Ricard, 2005; 2014; Foley et al., 2014; Mulyukova & Bercovici, 2018), generation of strike-slip motion (Bercovici & Ricard, 2005; 2013), evolution and rearrangement of existing plate boundaries by rapid slab detachment and rollback (Bellas et al., 2018; Bercovici et al., 2015, 2018), and the development of microstructure in deep lithospheric shear zones (Landuyt & Bercovici, 2009; Mulyukova & Bercovici, 2017, 2018).

Our new theoretical model of microstructural evolution of deforming rocks includes the energetics of both evolving grain boundaries and intragranular defects, or dislocations (Mulyukova & Bercovici, 2018). The theory predicts that the coupling of different microphysical processes involved in grain size evolution can give rise to a range of dynamical behavior in the lithosphere, from continuous stiffening by grain growth, to stable deformation at some equilibrium grain size, and to the oscillating mode of strengthening and weakening induced by the oscillations in grain size and dislocation density. The oscillating mode has a characteristic time scale that is of order decades, and thus it connects plate boundary formation to processes at
human time scales, such as the postseismic creep and earthquake cyclicity. However, to test the predictions of damage-induced rheological oscillations in the lithosphere, the theory of coupled grain boundary and dislocation dynamics needs to be upscaled to a continuum model. Furthermore, explicitly accounting for the energetics of dislocation dynamics, which are known to be an integral part of dynamic recrystallization and thus of grain size reduction, provides a physical framework for constraining the mechanisms by which grain damage occurs even when diffusion creep is the dominant deformation regime, as predicted by the two-phase grain damage theory.

For self-weakening and shear localization by grain damage to be effective, and specifically for the Zener pinning to be effective, the different mineral phases that make up the rock must be well mixed. The newly developed theories (Bercovici & Mulyukova, 2018; Bercovici & Skemer, 2017) and experimental studies (Cross & Skemer, 2017; Tasaka et al., 2017, 2017; Wiesman et al., 2018) provide a physical basis for the grain scale mixing processes by which the fine-grained mylonites and ultramylonites may form. The microphysical mixing models need to be further incorporated into the larger-scale mechanical models, in which the small-scale units of ultramylonites can be mixed with coarser grained monomineralic fabric, for example, by folding and stretching of the deforming lithosphere, in order to be applied to the geodynamics of plate boundary formation.

Acknowledgments
The authors are grateful to Olivier Castelnau, an anonymous reviewer, and the Associate Editor for helpful comments and suggestions. All data presented in this review are available in the original publications, as referenced. This work was supported by Yale University and NSF Grant EAR-1853184.

References


