# Tectonic plate generation and two-phase damage: void growth versus grainsize reduction

**David Bercovici**

Department of Geology and Geophysics, Yale University, New Haven,

Connecticut, USA

**Yanick Ricard**

Laboratoire des Sciences de la Terre, Ecole Normale Supérieure de Lyon, Lyon,

France

David Bercovici, Department of Geology and Geophysics, Yale University, PO Box 208109, New Haven, CT 06520-8109, USA. (david.bercovici@yale.edu)

Yanick Ricard, Laboratoire des Sciences de la Terre, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, F-69364 Lyon, Cedex 07, France. (ricard@ens-lyon.fr)

**Abstract.** The two-phase theory for compaction and damage employs a nonequilibrium relation between interfacial surface energy, pressure, and viscous deformation, thereby providing a model for damage (void generation and microcracking) and a continuum description of weakening, failure, and shear localization. Here we examine the application of this theory to the problem of generating plate-like behavior from convective-type divergent (poloidal) motion through a source-sink formulation. We extend the previous damage theory to consider two possible damage effects: (1) growth and nucleation of voids associated with dilation of the host matrix, and (2) increasing *fineness* (i.e., reducing coarseness) of the mixture by, for example, grainsize reduction. Void-generating damage is found to be poor at plate generation because of the predominance of dilational motion that is adverse to the development of plate-like flow. *Fineness*-generating damage is found to be very efficient at generating plate-like behavior if we assume that the matrix viscosity is a simple function of grain/void size, as is typical for diffusion creep. The implied grainsize reduction mechanism is different than that of dynamic recrystallization, and appears more capable of generating the requisite shear-localization for forming tectonic plates from mantle flow.

## 1. Introduction

Although the theory of plate tectonics is a well-established unifying principle of geology, the investigation of how the plates themselves arise from mantle convection is still a relatively young field [see reviews by *Bercovici et al.*, 2000; *Tackley*, 2000a, c; *Bercovici*, 2003]. There are many key issues in the problem of generating plates, including the initiation of subduction [*Mueller and Phillips*, 1991; *Kemp and Stevenson*, 1996; *Schubert and Zhang*, 1997; *Toth and Gurnis*, 1998; *King*, 2001; *Regenauer-Lieb et al.*, 2001], formation of passive ridges [*Ricard and Froidevaux*, 1986; *Tackley*, 2000b; *Huismans and Beaumont*, 2003], and the very existence of strong plates with weak, rapidly deforming boundaries [*Weinstein and Olson*, 1992; *Moresi and Solomatov*, 1998]. One of the most fundamental of the plate generation problems is the cause for strike-slip boundaries and the enigmatic "toroidal" motion which is not directly driven by buoyancy and does not transport heat, but is still a significant component of the Earth's surface velocity field [*Hager and O'Connell*, 1978, 1979, 1981; *Forte and Peltier*, 1987; *Ricard and Vigny*, 1989; *Gable et al.*, 1991; *O'Connell et al.*, 1991; *Olson and Bercovici*, 1991; *Cadek and Ricard*, 1992; *Ribe*, 1992; *Lithgow-Bertelloni et al.*, 1993; *Bercovici*, 1993; *Bercovici and Wessel*, 1994; *Bercovici*, 1995b, a, 1996, 1998; *Bercovici et al.*, 2000; *Bercovici*, 2003; *Dumoulin et al.*, 1998; *Weinstein*, 1998].

The formation of strong plates and weak boundaries, as well as the generation of toroidal motion is known to require an interaction between convective flow and strongly nonlinear rheologies, in particular ones that induce weakening with increased deformation [see reviews by *Tackley*, 2000a; *Bercovici et al.*, 2000; *Bercovici*, 2003]. Standard steady-state plastic or pseudo-plastic (power-law) rheologies explicitly prescribe such weakening behavior, but whether they are sufficient to the task of plate-generation is questionable; while they might allow for softening

of the material (lowering of viscosity) with higher deformation, they do not allow for strength loss (lowering stress) in which case deformation does not localize sufficiently [*Bercovici*, 1993]. These "plastic" rheologies, however, can be augmented with the help of melting (as at mid-ocean ridges) and yield more plate-like structures *Tackley* [2000b]. Even so, the localized plate boundaries that are generated with standard steady-state rheologies exist only as long as they are being deformed, which is unrealistic since boundaries persist as suture zones even when inactive [*Gurnis et al.*, 2000]. Simple self-weakening rheologies wherein strength loss is permitted (e.g., pseudo- or viscous-stick slip behavior *Bercovici* [1993]) allow more profound plate-like behavior and localization [*Bercovici*, 1993, 1995b, a; *Tackley*, 1998]; these are based on self-weakening feedback mechanisms, such as what arises from viscous heating of material with thermoviscous behavior [*Schubert and Turcotte*, 1972; *Whitehead and Gans*, 1974; *Schubert and Yuen*, 1978; *Fleitout and Froidevaux*, 1980; *Bercovici*, 1996, 1998] or with simple damage whereby weakening defects or voids are generated by deformation [*Bercovici*, 1998; *Auth et al.*, 2003; *Ogawa*, 2003]. The simple damage laws allow for more precipitous drops in viscosity than the standard Arrhenius thermoviscous law and hence yield more focussed shear zones [*Bercovici*, 1998], although they can also cause runaway damage whereby plates continue to break down to ever smaller plates [*Tackley*, 2000b]. While these simple damage treatments are easily implemented and are reasonably successful at yielding plate-like flows, their derivation and connection to the physics of void/microcrack generation is ad hoc.

The physics of void-generating damage has been recently developed from first principles using two-phase continuum mechanics [*Bercovici et al.*, 2001a; *Ricard et al.*, 2001; *Bercovici et al.*, 2001b; *Bercovici and Ricard*, 2003; *Ricard and Bercovici*, 2003]. The theory is reasonably distinct from other elastodynamics damage theories [*Ashby and Sammis*, 1990; *Hansen and Schreyer*, 1992; *Lemaitre*, 1992; *Krajcinovic*, 1996; *Lyakhovsky et al.*, 1997] as it models the

growth of voids through a two-phase approach (i.e., using a fluid or void-filling phase, and a host-rock or matrix phase) and treats void growth via storage of surface energy on the interface between phases, which serves as a proxy for the fracture surface [*Griffith*, 1921; *Bercovici et al.*, 2001a]. The two-phase damage theory has been fairly successful at capturing various aspects of ductile cracking, including localizing of simple shear, involving, in order of increasing forcing, weak and strong localization, and then distributed damage [*Bercovici et al.*, 2001b; *Bercovici and Ricard*, 2003]; and shear-enhanced compaction [*Menéndez et al.*, 1996; *Wong et al.*, 1997; *Zhu and Wong*, 1997; *Zhu et al.*, 1997] under uniaxial compression of low-cohesion failure envelopes [*Ricard and Bercovici*, 2003].

In all the studies of two-phase damage so far, the model systems were driven by an imposed shear stress and the damage-induced weakening would facilitate or enhance the mode of deformation that is driven directly by the imposed shear. For example, in simple-shear cases, damage and localization typically act to enhance the shear strain-rate that is parallel to the imposed shear stress [*Bercovici et al.*, 2001b; *Bercovici and Ricard*, 2003]. In the problem of generating plates from mantle convection, the system is not driven by a simple imposed shear stress as in simple-shear flow or uniaxial compression. Instead, it is driven by convective buoyancy forces which force divergent and convergent (poloidal) motion in the top, horizontally moving thermal boundary layer (i.e., the lithosphere). This convective poloidal motion interacts with a nonlinear rheological mechanism to generate plate-like motion, especially strike-slip toroidal motion. However, the toroidal flow structures (e.g., strike-slip boundaries) are usually orthogonal to and often separated from the poloidal structures associated with convective forcing (convergent/divergent boundaries such as subduction zones and ridges). Thus, the localized toroidal structures do not occur by merely enhancing deformation that is directly driven by the driving stresses. That is, as opposed to cases involving localization and enhancement of an

imposed shear (for fault formation) or extension (for necking), in the plate-generation case a unique toroidal flow field that is essentially orthogonal to the driving convective poloidal field must be generated. In short, it is not a matter of enhancing the imposed poloidal motion but of generating a new toroidal flow field.

Therefore, in this paper, we examine the extent to which our two-phase damage mechanism permits generation of plate-like toroidal flow from an imposed convective-type poloidal flow. We employ a very simple system of source-sink driven flow, which is purely poloidal [*Bercovici*, 1993]. The damage "rheology" interacts with this poloidal driving flow to generate toroidal motion; we therefore search for conditions under which plate-like toroidal motion and strength distribution arise spontaneously.

We also extend the two-phase damage theory to consider two fundamental forms of damage: damage associated with matrix dilation and void-growth, and non-dilational damage associated with interface area growth without void generation, as for example in grainsize reduction. Dilational void-generating damage is well known for micro-cracking brittle and brittle-ductile behavior in low-pressure, low-cohesion and/or granular materials [*Scott*, 1996; *Menéndez et al.*, 1996; *Wong et al.*, 1997; *Zhu and Wong*, 1997; *Zhu et al.*, 1997; *Géminard et al.*, 1999]. Grainsize reduction is an important aspect of lithospheric rheology since diffusion creep is fairly strongly grainsize dependent, and continuous shear-localization is well associated with mylonitic (grainsize-reduced) zones [*White et al.*, 1980; *Karato*, 1983; *Jin et al.*, 1998; *Furusho and Kanagawa*, 1999]. The cause for such grainsize reducing shear localization is usually attributed to the process of dynamic recrystallization [*Karato et al.*, 1980; *Urai et al.*, 1986; *Derby and Ashby*, 1987] whereby grains are reduced through the propagation of dislocations while the medium undergoes nonlinear dislocation creep, but the localization of shear occurs only when the medium drops to the lower-stress and grainsize-controlled diffusion creep. The physics of

this mechanism is, however, still not fully developed [*Bercovici and Karato*, 2003] and has had so far limited success in recovering observed levels of shear localization [*Kameyama et al.*, 1997; *Braun et al.*, 1999; *Bresser et al.*, 1998, 2001; *Montési and Zuber*, 2002]. In this study we will examine the extent to which grainsize reduction driven by direct damage (i.e., in essence pulverizing the matrix) allows for sufficient shear localization and plate generation.

## 2. Two-phase damage theory: review and extensions

The original two-phase damage equations have been developed over a series of papers [*Bercovici et al.*, 2001a, b; *Ricard et al.*, 2001; *Bercovici and Ricard*, 2003; *Ricard and Bercovici*, 2003] and thus we only briefly present the governing equations. We note in advance that in the following equations, subscripts $f$ and $m$ refer to fluid (i.e., void-filling material) and matrix (e.g., host rock) phases, respectively. Moreover, the volume-average and phasic difference of any quantity $Q$ are defined as

$$\bar{Q} = \phi Q_f + (1-\phi)Q_m, \Delta Q = Q_m - Q_f, \tag{1}$$

respectively. All dependent variables are not, in fact, true microscopic quantities but are averaged over the fluid or matrix space within small but not necessarily infinitesimal control volumes. Moreover, all equations are invariant to a permutation of subscripts $f$ and $m$ and, implicitly, a switch of $\phi$ and $1-\phi$, where $\phi$ is fluid volume fraction, or porosity; this symmetry property is called "material invariance" (see *Bercovici et al.* [2001a] for further discussion).

### 2.1. Mass conservation

The conservation of mass equations are standard in two-phase theories [*McKenzie*, 1984] and remain unchanged here. There are two equations involving transport of the fluid and matrix phases:

$$\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot [\phi \mathbf{v}_f] = 0 \tag{2}$$

$$\frac{\partial(1-\phi)}{\partial t} + \boldsymbol{\nabla} \cdot [(1-\phi)\mathbf{v}_m] = 0, \tag{3}$$

where $\mathbf{v}_f$ and and $\mathbf{v}_m$ are the fluid and matrix velocities.

### 2.2. Momentum conservation; interfacial surface tension and interface area density

The general momentum equations for each phase are

$$\begin{aligned} 0 = &- \phi \left[\boldsymbol{\nabla} P_f + \rho_f g \hat{\mathbf{z}}\right] + \boldsymbol{\nabla} \cdot [\phi \underline{\boldsymbol{\tau}}_f] \\ &+ c\Delta\mathbf{v} + \omega \left[\Delta P \boldsymbol{\nabla}\phi + \boldsymbol{\nabla}(\sigma\alpha)\right] \end{aligned} \tag{4}$$

$$\begin{aligned} 0 = &- (1-\phi) \left[\boldsymbol{\nabla} P_m + \rho_m g \hat{\mathbf{z}}\right] + \boldsymbol{\nabla} \cdot [(1-\phi) \underline{\boldsymbol{\tau}}_m] \\ &- c\Delta\mathbf{v} + (1-\omega) \left[\Delta P \boldsymbol{\nabla}\phi + \boldsymbol{\nabla}(\sigma\alpha)\right], \end{aligned} \tag{5}$$

where $g$ is gravity and $c$ is the coefficient for viscous drag between phases (also referred to as the Darcy drag coefficient, and as such can be related to permeability for small porosity; see *Bercovici et al.* [2001a]); $P_j$ and $\rho_j$ are the pressure and density, respectively, in phase $j$ (where $j = f$ or $m$) and the density of each phase is assumed constant. The deviatoric stress for phase $j$ is given by

$$\underline{\boldsymbol{\tau}}_j = \mu_j \left[\boldsymbol{\nabla}\mathbf{v}_j + [\boldsymbol{\nabla}\mathbf{v}_j]^t - \frac{2}{3}(\boldsymbol{\nabla}\cdot\mathbf{v}_j)\underline{\mathbf{I}}\right] \tag{6}$$

where $\mu_j$ is viscosity. Surface tension is denoted by $\sigma$, and $\omega$ is a coefficent for how surface tension partitions between phases, i.e., to what extent it is more embedded in one phase than the other [*Bercovici and Ricard*, 2003]; in the case of $\mu_f \ll \mu_m$, which is most geologically relevant, $\omega \ll 1$. Lastly, the interfacial area per unit volume $\alpha$ is generally a function of porosity $\phi$ and inversely proportional to average grain and/or void size which is denoted by $\mathcal{A}^{-1}$ (where $\mathcal{A}$ has units of $\mathrm{m}^{-1}$; in keeping with our previous studies [*Bercovici et al.*, 2001a, b; *Ricard et al.*, 2001; *Bercovici and Ricard*, 2003; *Ricard and Bercovici*, 2003] we write

$$\alpha = \mathcal{A}\eta(\phi) \tag{7}$$

where

$$\eta(\phi) = \phi^a(1-\phi)^b \tag{8}$$

and $a$ and $b$ are constants $\leq 1$. As discussed in *Bercovici et al.* [2001a] $\partial\alpha/\partial\phi$ is analagous to the average interface curvature. Whether $\mathcal{A}$ represents void or grain size depends on the sign of the average interface curvature; if $\partial\alpha/\partial\phi > 0$ then curvature is positive implying that the interface is concave to the void-filling fluid phase (e.g., voids are like bubbles encased in matrix), and if $\partial\alpha/\partial\phi < 0$ then curvature is negative and the interface is convex to the fluid (the matrix is composed of grains bathed in fluid). The change in sign of curvature from positive to negative with increasing $\phi$ occurs at $\phi = a/(a+b)$, and thus for $\phi < a/(a+b)$ we can assume that the interface mostly encloses pores and $\mathcal{A}$ represents inverse void size; while for $\phi > a/(a+b)$ the interface encloses matrix grains in which case $\mathcal{A}$ represents inverse grainsize. For slurries and granular-type media, $a/(a+b) \ll 1$, while for foams $a/(a+b) \approx 1$.

### 2.3. Energy conservation

Following the development of *Bercovici et al.* [2001a] and *Bercovici and Ricard* [2003], the energy equation is separated into two coupled equations representing (1) the evolution of thermal (entropy-related) energy, and (2) the rate of work done on the interface by pressure, surface tension, and viscous deformational work. The interfacial surface energy and the work done by surface tension on the mixture is assumed to be partitioned between phases by the same fraction $\omega$ as the surface tension force in the previous section. With these assumptions we arrive at (see *Bercovici et al.* [2001a] for a detailed derivation with the case of $\omega = \phi$, and *Bercovici and*

*Ricard* [2003] for equations with a general $\omega$)

$$\overline{\overline{\rho c}}\frac{\widetilde{D}T}{Dt} - T\frac{\widetilde{D}}{Dt}\left(\alpha\frac{d\sigma}{dT}\right) - T\alpha\frac{d\sigma}{dT}\boldsymbol{\nabla}\cdot\widetilde{\mathbf{v}} = Q - \boldsymbol{\nabla}\cdot\mathbf{q} + B\left(\frac{\widetilde{D}\phi}{Dt}\right)^2 + (1-f)\Psi \tag{9}$$

$$\sigma\frac{\widetilde{D}\alpha}{Dt} = -\Delta P\frac{\widetilde{D}\phi}{Dt} - B\left(\frac{\widetilde{D}\phi}{Dt}\right)^2 + f\Psi, \tag{10}$$

where $T$ is the temperature (assumed the same in both phases), $-d\sigma/dT$ is the interfacial entropy per unit area [*Desjonquères and Spanjaard*, 1993; *Bailyn*, 1994; *Bercovici et al.*, 2001a],

$$\widetilde{\mathbf{v}} = \omega\mathbf{v}_f + (1-\omega)\mathbf{v}_m \tag{11}$$

is the effective velocity of the interface, $Q$ is an intrinsic heat source, $\mathbf{q}$ is an energy flux vector (accounting for heat diffusion and possibly energy dispersion; see *Bercovici et al.* [2001a]), and

$$\Psi = c\Delta v^2 + \phi\boldsymbol{\nabla}\mathbf{v}_f : \underline{\boldsymbol{\tau}}_f + (1-\phi)\boldsymbol{\nabla}\mathbf{v}_m : \underline{\boldsymbol{\tau}}_m \tag{12}$$

(where $\Delta v^2 = \Delta\mathbf{v}\cdot\Delta\mathbf{v}$) is the viscous deformational work, a fraction $f$ of which is partitioned into stored work (in this model stored as interface surface energy) while the remaining part goes toward dissipative heating [*Taylor and Quinney*, 1934; *Chrysochoos and Martin*, 1989]; see *Bercovici et al.* [2001a] for further discussion of the partitioning fraction $f$. The quantity $B$ must be positive, has units of viscosity, and the term associated with it represents irreversible viscous work done on pores and grains by the pressure difference $\Delta P$ during isotropic compaction or dilation [*Bercovici et al.*, 2001a; *Ricard et al.*, 2001]. Simple micromechanical models suggest that

$$B = K\frac{(\mu_m + \mu_f)}{\phi(1-\phi)} \tag{13}$$

where $K$ is a dimensionless factor accounting for pore or grain geometry and is typically $O(1)$ [*Bercovici et al.*, 2001a]; see also *Sumita et al.* [1996]. In the limit of $\mu_f \ll \mu_m$, $B$ is essentially

equivalent to the matrix bulk viscosity introduced by *McKenzie* [1984]; see *Ricard et al.* [2001] and *Bercovici and Ricard* [2003] for further discussion. The average heat capacity per volume of the mixture is $\overline{\rho c} = \phi\rho_f c_f + (1-\phi)\rho_m c_m$ (where $c_f$ and $c_m$ are the heat capacities of the fluid and matrix), and the material derivatives in (9) and (10) are defined as

$$\frac{\widetilde{D}}{Dt} = \frac{\partial}{\partial t} + \widetilde{\mathbf{v}}\cdot\boldsymbol{\nabla} = \omega\frac{D_f}{Dt} + (1-\omega)\frac{D_m}{Dt} \tag{14}$$

$$\frac{\overline{\overline{D}}}{Dt} = \frac{1}{\overline{\rho c}}\left(\phi\rho_f c_f\frac{D_f}{Dt} + (1-\phi)\rho_m c_m\frac{D_m}{Dt}\right) \tag{15}$$

in which

$$\frac{D_f}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}_f\cdot\boldsymbol{\nabla}, \frac{D_m}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}_m\cdot\boldsymbol{\nabla}. \tag{16}$$

### 2.4. The damage equation: void vs *fineness* generation

Equation (10) governs the rate that deviatoric stresses and the interphasic pressure difference do work on the interface, effectively storing surface energy on the interface; as this models the growth of microcracks and defects by growth of interfacial area it is termed the damage equation.

In previous papers we considered the inverse grain/void size $\mathcal{A}$ to be constant and that all damage occured through void growth and porosity change. However, it is clearly possible to incur damage without void growth, but instead by increasing the number of voids or grains without changing porosity, hence reducing the void/grain size, and thus increasing the interfacial area density. We refer to this effect as increasing the *fineness* of the mixture, i.e., the smaller the grain/void size the finer (or less coarse) the texture of the two-phase medium. Therefore, here we allow for the possibility that $\mathcal{A}$ is variable, and that an increase in $\mathcal{A}$ increases the *fineness* of the medium. In this case, (10) becomes

$$\sigma\mathcal{A}\frac{d\eta}{d\phi}\frac{\widetilde{D}\phi}{Dt} + \sigma\eta\frac{\widetilde{D}\mathcal{A}}{Dt} = -\Delta P\frac{\widetilde{D}\phi}{Dt} - B\left(\frac{\widetilde{D}\phi}{Dt}\right)^2 + f\Psi, \tag{17}$$

We assume that the terms proportional to $\Delta P$ and $B$ contribute to void generation since they are clearly associated with work to change porosity $\phi$. However, in general the term $f\Psi$ is itself partitioned between growth of voids and enhancement of *fineness*, which each contribute to interface generation; we assume that a fraction $f_\phi$ contributes to void generation, while a fraction $f_\mathcal{A}$ contributes to *fineness* generation, and that $f = f_\phi + f_\mathcal{A}$. Thus, (17) is itself divided into

$$\sigma\mathcal{A}\frac{d\eta}{d\phi}\frac{\widetilde{D}\phi}{Dt} = -\Delta P\frac{\widetilde{D}\phi}{Dt} - B\left(\frac{\widetilde{D}\phi}{Dt}\right)^2 + f_\phi\Psi, \tag{18}$$

and

$$\sigma\eta\frac{\widetilde{D}\mathcal{A}}{Dt} = f_\mathcal{A}\Psi \tag{19}$$

Although both void- and *fineness*-generating damage can occur simultaneously, we will here consider the limiting cases where only one or the other occurs; i.e. when $f_\phi \neq 0$ then $f_\mathcal{A} = 0$, and vice versa.

### 2.4.1. Partitioning fractions

As discussed in our previous papers [*Bercovici et al.*, 2001b; *Bercovici and Ricard*, 2003], when void-generating damage occurs ($f_\phi \neq 0$) singular solutions of (18) in regions of zero void-growth (i.e., $\widetilde{D}\phi/Dt = 0$) are precluded by requiring that the partitioning fraction $f_\phi$ be dilation-rate dependent, e.g.,

$$f_\phi = f^*\frac{(\widetilde{D}\phi/Dt)^2}{\gamma + (\widetilde{D}\phi/Dt)^2} \tag{20}$$

where $f^*$ is the maximum permissible $f_\phi$, $\gamma$ controls the variability of $f_\phi$, and $f_\phi$ is assumed to depend on an even power of $\widetilde{D}\phi/Dt$ since it must be positive definite (and for simplicity we assume the lowest-order such power).

However, in the case where damage generates greater *fineness* (reduces grain/void size) such that $f_{\mathcal{A}} \neq 0$, there are no explicit singularites in (19). Hence, there is no need to ascribe any mitigating behavior to $f_{\mathcal{A}}$, and thus for simplicity we generally assume that $f_{\mathcal{A}}$ is constant.

### 2.5. Matrix rheology

As discussed in *Ricard et al.* [2001], porosity can influence the effective strength of the medium; e.g., in shear flow the effective viscosity is

$$\mu_{\text{eff}} = \mu_m(1-\phi); \tag{21}$$

which essentially arises from the factor of $1-\phi$ in the term involving $\underline{\underline{\tau}}_m$ in (5). Similarly, since we are here considering variation in grain/voidsize, we consider ways in which $\mathcal{A}$ affects rheology, in particular of the matrix which represents the host rock. Although reductions in bubble or grain size in emulsions and suspensions are known to increase viscosity [*Batchelor*, 1967; *Larson*, 1999], the break up of the matrix material into grains in the mixture is likely to facilitate its deformation because of grain-sliding, increased number of available granular slip planes, and enhanced diffusion creep. We thus assume that the matrix viscosity depends on $\mathcal{A}$ similarly to how viscosity depends on grainsize in diffusion creep, i.e.,

$$\mu_m = \mu_0(\mathcal{A}_0/\mathcal{A})^m \tag{22}$$

where $\mu_0$ and $A_0$ are constant reference values for viscosity and *fineness* , respectively, and $m$ is a dimensionless positive constant.

However, the *fineness*-dependence of viscosity, only really affects the system when growth in $\mathcal{A}$ is forced by damage, i.e., when $f_{\mathcal{A}} > 0$. Otherwise, healing is assumed to force $\mathcal{A}$ toward a homogeneous value (see Appendix A); indeed, for cases in which $f_{\mathcal{A}} = 0$ we assume a simple initial condition of $\mathcal{A} = \mathcal{A}_0$ which leads to a viscosity that, because $\mathcal{A}$ does not change, remains constant (at $\mu_m = \mu_0$).

Finally, we assume that in the limit $\mu_f \ll \mu_m$ (see §2.6 below) that the effective bulk viscosity obeys

$$B = \frac{K\mu_0}{\phi(1-\phi)} \tag{23}$$

and that it is therefore independent of $\mathcal{A}$ (i.e., $B$ depends on $\mu_0$ rather than $\mu_m$). We make this simplification because while it is reasonable that shear viscosity is grainsize dependent, it is not at all clear that an effective bulk viscosity would be grainsize dependent, or at least in the same sense as $\mu_m$. For example, compaction probably becomes more difficult with smaller grains because of closer possible packing, as opposed to deformation under shear stress which should be facilitated by smaller grainsizes (either by diffusion creep or because of the existence of more slip planes). For the sake of simplicity, we simply leave $B$ as independent of $\mathcal{A}$ (although in the Discussion section we briefly discuss results of calculations allowing $B(\mathcal{A})$).

### 2.6. The evacuated-void limit

We next adopt the geologically applicable "void limit" proposed by *Ricard and Bercovici* [2003] whereby pores have zero density, pressure and viscosity ($\rho_f = P_f = \mu_f = 0$, and thus $\omega = 0$ and $\frac{\widetilde{D}}{Dt} = \frac{D_m}{Dt}$). At the interface there is no interaction force between phases ($c\Delta\mathbf{v} = 0$) and the interface itself is assumed to move with the matrix. In this limit the governing equations of mass and momentum, i.e., (3) and (5), are, respectively,

$$\frac{D_m\phi}{Dt} = (1-\phi)\boldsymbol{\nabla}\cdot\mathbf{v}_m \tag{24}$$

$$0 = \boldsymbol{\nabla}[\sigma\alpha - (1-\phi)P_m] + \boldsymbol{\nabla}\cdot[(1-\phi)\underline{\boldsymbol{\tau}}_m] - (1-\phi)\rho_m g\hat{\mathbf{z}} \tag{25}$$

The general damage equations (following previous sections) becomes

$$\sigma\mathcal{A}\frac{d\eta}{d\phi} = -P_m - B\frac{D_m\phi}{Dt} + \frac{f^* D_m\phi/Dt}{\gamma + (D_m\phi/Dt)^2}(1-\phi)\boldsymbol{\nabla}\mathbf{v}_m : \underline{\boldsymbol{\tau}}_m \tag{26}$$

$$\sigma\eta\frac{D_m\mathcal{A}}{Dt} = f_{\mathcal{A}}(1-\phi)\boldsymbol{\nabla}\mathbf{v}_m : \underline{\boldsymbol{\tau}}_m \tag{27}$$

## 3. Source-sink formulation

The primary goal of this paper is to study the classic plate-mantle coupling problem of how well a convective poloidal flow field can generate toroidal flow through nonlinear rheological mechanisms [see reviews by *Bercovici et al.*, 2000; *Bercovici*, 2003]. In this analysis we prescribe a poloidal flow by imposing a source-sink field in a shallow-layer of fluid (nominally the lithosphere). In effect, the source-sink field represents vertical motion of underlying fluid being injected into or ejected from the horizontal shallow layer.

The thin layer in which we are modelling flow is assumed to be bounded above and below by inviscid half-spaces (i.e., the ocean and atmosphere above, the low-viscosity asthensphere below) and thus has free-slip boundaries. The velocity of material (i.e., matrix material) in the layer is $\mathbf{v}_m = u_m\hat{\mathbf{x}} + v_m\hat{\mathbf{y}} + w_m\hat{\mathbf{z}}$. However, within the layer we assume vertical flow $w_m$ is negligible, and thus the free-slip boundaries leads to the condition that $\partial u_m/\partial z = \partial v_m/\partial z = 0$. We further assume that the layer is so thin that this condition exists across the entire width of the layer. A velocity field that satisfies these conditions can be expressed with the Helmholtz representation

$$\mathbf{v}_m = \boldsymbol{\nabla}\theta + \boldsymbol{\nabla} \times (\psi\hat{\mathbf{z}}) + \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times (z\xi\hat{\mathbf{z}}) \tag{28}$$

where $\theta$, $\psi$ and $\xi$ are functions of $x$, $y$ and time $t$ only. The velocity potential $\theta$ represents "compressible" flow associated with dilation or compaction due to void generation or collapse. The last two terms on the right of (28) represent the incompressible solenoidal flow of the matrix which is not associated with dilation or compaction [see *Spiegelman*, 1993]; the solenoidal potential $\psi$ is the toroidal stream function and $\xi$ is the poloidal potential. The velocity given by

(28) can be rewritten as

$$\begin{aligned} u_m &= \frac{\partial(\theta+\xi)}{\partial x} + \frac{\partial\psi}{\partial y} \\ v_m &= \frac{\partial(\theta+\xi)}{\partial y} - \frac{\partial\psi}{\partial y} \\ w_m &= -z\nabla^2_{\rm h}\xi. \end{aligned} \tag{29}$$

and we define the entire thin layer to exist near $z = 0$ such that $w_m$ is negligibly small. We also note that since dilational velocity is assumed independent of $z$ then porosity $\phi$ is also independent of $z$ (i.e., there is no vertical variability in how voids dilate or compact, and thus void density is assumed not to vary vertically).

### 3.1. Source-sink, vorticity and dilation-rate fields

As stated above, we are here only concerned with lateral flow in a thin horizontal fluid layer that is driven by an imposed source-sink field, or injection and ejection of material from below. Within the thin layer we assume that vertical matrix velocity $w_m$ is zero, but that $\partial w_m/\partial z$ is nonzero; indeed, the negative of the latter quantity represents the net vertical volume flux per unit volume (or at a point) of material being injected into the thin layer (i.e., if the layer has thickness $h$, then the net vertical volume flux over an infinitesimal area is $[w_m(z=0) - w_m(z=h)]dxdy \approx -(\partial w_m/\partial z)hdxdy$). We therefore prescribe the source-sink field (due to vertical injection/ejection) as $S = -\partial w_m/\partial z$, which, using (29), can be restated in terms of the poloidal field as

$$\nabla^2\xi = S. \tag{30}$$

Although the desired goal in studies like these is plate-like toroidal motion, the best measure of such motion is concentrated bands of vertical vorticity that, if they were to represent a true discontinuous strike-slip fault, would by line singularities. Vertical vorticity is defined as $\Omega = \hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\times\mathbf{v}_m$ and is in effect a measure of the angular velocity of a point spinning about a vertical

axis, and thus represents rate of strike-slip shear. In terms of toroidal flow, it is given by

$$\nabla_{\rm h}^2\psi = -\Omega \tag{31}$$

Void-generating damage is associated with dilation and compaction represented by $\theta$; the dilation rate is $G = \boldsymbol{\nabla}\cdot\mathbf{v}_m$, which with (28) yields

$$\nabla_{\rm h}^2\theta = G \tag{32}$$

and with (24) also leads to

$$\frac{D_m\phi}{Dt} = (1-\phi)G. \tag{33}$$

Therefore, we find in the end that our three velocity potentials $\xi$, $\psi$ and $\theta$ are related to the the imposed source-sink field $S$, the vertical vorticity (or strike-slip-shear rate) $\Omega$, and the dilation rate $G$, respectively, through three simple 2D Poisson's equations (30), (31) and (32). Although these three relations provide equations for $\xi$, $\psi$ and $\theta$, we will need equations for $\Omega$ and $G$, since $S$ is already given. For these we will use our momentum and damage equations.

The momentum equation (25) is re-written as

$$0 = -\boldsymbol{\nabla}\Pi + \boldsymbol{\nabla}\cdot[(1-\phi)\underline{\boldsymbol{\tau}}_m] - (1-\phi)\rho_m g\hat{\mathbf{z}} \tag{34}$$

where $\Pi = (1-\phi)P_m - \sigma\alpha$. We can write a general expression for $\Pi$ that takes into account both void-generating and *fineness*-generating damage, i.e., using (26) and (27) along with the source-sink formulation, we obtain

$$\begin{aligned}\Pi = &-\sigma\mathcal{A}\left(\eta + (1-\phi)\frac{d\eta}{d\phi}\right) - B(1-\phi)^2G \\ &+ f^*\frac{(1-\phi)^3G}{\gamma+(1-\phi)^2G^2}\mu_0\left(\frac{\mathcal{A}_0}{\mathcal{A}}\right)^m\Phi\end{aligned} \tag{35}$$

and

$$\frac{D_m\mathcal{A}}{Dt} = f_{\mathcal{A}}\frac{1-\phi}{\sigma\eta}\mu_0\left(\frac{\mathcal{A}_0}{\mathcal{A}}\right)^m\Phi \tag{36}$$

where

$$\begin{aligned}\Phi =&\boldsymbol{\nabla}\mathbf{v}_m : \left(\boldsymbol{\nabla}\mathbf{v}_m + [\boldsymbol{\nabla}\mathbf{v}_m]^t - \frac{2}{3}\boldsymbol{\nabla}\cdot\mathbf{v}_m\underline{\mathbf{I}}\right)\\ =&(\Delta^*\psi)^2 + 4\left(\frac{\partial^2\psi}{\partial x\partial y}\right)^2 + \boldsymbol{\nabla}\boldsymbol{\nabla}(\theta+\xi) : \boldsymbol{\nabla}\boldsymbol{\nabla}(\theta+\xi)\\ &+ 2S^2 - \frac{2}{3}G^2 + 4\left(\frac{\partial^2\psi}{\partial x\partial y}\Delta^*(\theta+\xi) - \Delta^*\psi\frac{\partial^2(\theta+\xi)}{\partial x\partial y}\right)\end{aligned} \tag{37}$$

and we define the differential operator

$$\Delta^* = \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}. \tag{38}$$

In the case of void generation, $f_{\mathcal{A}} = 0$ and thus $\mathcal{A}$ remains constant at $\mathcal{A} = \mathcal{A}_0$, leading to $\mu_m = \mu_0$. In the case of *fineness* generation, $f^* = 0$ and $\mathcal{A}$ is variable and time dependent, although since the forcing of growth in $\mathcal{A}$ is independent of $z$ we assume that $\mathcal{A}$ is $z$-independent as well.

We can extract equations for $\Omega$ and $G$ by combining (35) with (34) and manipulating the result. However, it is useful to first nondimensionalize the governing equations.

### 3.2. Dimensionless governing equations

We use the maximum of the source-sink field $S_{\max}$ for a rate scale (inverse of time), the characteristic separation of the source and sink $L$ as our macroscopic length scale, and the reference viscosity $\mu_0$ to help define a stress scale $\mu_0 S_{\max}$; we also define $\mathcal{A}_0$ as our *fineness* scale (thus we keep our macroscopic length scale $L$ distinct from the microscopic one $\mathcal{A}_0^{-1}$). We thus nondimensionalize according to $(x, y, z) = L(x', y', z')$, $\boldsymbol{\nabla} = L^{-1}\boldsymbol{\nabla}'$, $\mathcal{A} = \mathcal{A}_0\mathcal{A}'$ and

$$\begin{aligned}&(S, G, \Omega, \sqrt{\Phi}, \mathbf{v}_m, \xi, \theta, \psi, \Pi) =\\ &\qquad S_{\max}(S', G', \Omega', \sqrt{\Phi'}, L\mathbf{v}_m', L^2\xi', L^2\theta', L^2\psi', \mu_0\Pi')\end{aligned} \tag{39}$$

where primed quantities are dimensionless.

Substituting these into our governing equations, but subsequently dropping the primes, we first arrive at our three Poisson's equations (which remain unchanged, but are repeated here for the sake of completeness)

$$\nabla^2\xi = S, \nabla^2\psi = -\Omega, \nabla^2\theta = G \tag{40}$$

Combining (34) and (35), substituting our nondimensionalization and taking $\hat{\mathbf{z}} \cdot \boldsymbol{\nabla}\times$ of the resulting equation, eventually leads to a nonlinear Poisson's equation for $\Omega$:

$$\begin{aligned}\bar{\mu}\nabla^2\Omega = &-2\boldsymbol{\nabla}\bar{\mu}\cdot\boldsymbol{\nabla}\Omega - \hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\bar{\mu}\times\boldsymbol{\nabla}\left(S+2G\right)\\ &-\Delta^*\bar{\mu}\left(2\frac{\partial^2(\theta+\xi)}{\partial x\partial y} - \Delta^*\psi\right) + 2\frac{\partial^2\bar{\mu}}{\partial x\partial y}\left(\Delta^*(\theta+\xi) + 2\frac{\partial^2\psi}{\partial x\partial y}\right)\end{aligned} \tag{41}$$

Likewise, taking $\boldsymbol{\nabla}\cdot$ of the combination of (34) and (35) with nondimensional variables leads to a nonlinear Poisson's equation for $G$:

$$\begin{aligned}\left(\frac{4}{3}\bar{\mu}+\frac{1-\phi}{\phi}\right)\nabla^2 G = &-\hat{\sigma}\nabla^2(\lambda\mathcal{A}) + f^*\nabla^2\left(\frac{(1-\phi)^2G}{\hat{\gamma}+(1-\phi)^2G^2}\bar{\mu}\Phi\right)\\ &+2\hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\bar{\mu}\times\boldsymbol{\nabla}\Omega - 2\boldsymbol{\nabla}\bar{\mu}\cdot\boldsymbol{\nabla}S + 2\left(\frac{1}{\phi^2}\boldsymbol{\nabla}\phi - \frac{4}{3}\boldsymbol{\nabla}\bar{\mu}\right)\cdot\boldsymbol{\nabla}G\\ &-\left(2\frac{\boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\phi}{\phi^3} - \frac{\nabla^2\phi}{\phi^2} - \frac{2}{3}\nabla^2\bar{\mu}\right)G\\ &-2\left(\Delta^*\bar{\mu}\frac{\partial^2\psi}{\partial x\partial y} - \frac{\partial^2\bar{\mu}}{\partial x\partial y}\Delta^*\psi + \boldsymbol{\nabla}\boldsymbol{\nabla}\bar{\mu} : \boldsymbol{\nabla}\boldsymbol{\nabla}(\theta+\xi)\right)\end{aligned} \tag{42}$$

In (41) and (42) we have defined

$$\bar{\mu} = \frac{1-\phi}{\mathcal{A}^m} \tag{43}$$

as the effective dimensionless viscosity, and

$$\lambda = \eta + (1-\phi)\frac{d\eta}{d\phi} \tag{44}$$

$$\hat{\sigma} = \frac{\sigma\mathcal{A}_0}{\mu_0 S_{\max}} \tag{45}$$

$$\hat{\gamma} = \frac{\gamma}{S_{\max}^2} \tag{46}$$

We have also used (23) to eliminate $B$, and we note that the form of $\Phi$ does not change from that given in (37).

Finally, the dimensionless evolution equations for $\phi$ and $\mathcal{A}$ follow from (24) and (27), yielding

$$\frac{\partial \phi}{\partial t} + \mathbf{v}_m \cdot \boldsymbol{\nabla} \phi = (1-\phi) G \tag{47}$$

$$\frac{\partial \mathcal{A}}{\partial t} + \mathbf{v}_m \cdot \boldsymbol{\nabla} \mathcal{A} = \frac{f_{\mathcal{A}}}{\hat{\sigma} \eta(\phi)} \bar{\mu} \Phi \tag{48}$$

where, again, for void-generating damage $f_{\mathcal{A}} = 0$ and $\mathcal{A} = 1$ and thus $\bar{\mu} = 1 - \phi$; for *fineness*-generating damage $f^* = 0$ and $\mathcal{A}$ is variable.

Overall, our final governing equations are (40), (41), (42), (47), and (48), all driven by an imposed source-sink field $S$, whose structure is specified in the numerical method section (§3.4). The velocity vector appears explicity only in (47) and (48) is given by (28) (which remains unchanged by nondimensionalization).

### 3.3. Simple analysis of source-sink equations

Fully plate-like behavior involves strongly nonlinear solutions to our governing equations and thus numerical analysis. However, one can gain insight into the dominant driving terms for initiating plate-like (or non-plate-like) flows with a simple perturbation analysis. We consider a uniform and static background state (constant $\phi$ and $\mathcal{A}$) driven by a 0th order source sink field $S$; the resulting flow fields related to dilation and vorticity, $G$ and $\Omega$, respectively are 1st order perturbations, as are fluctuations in $\phi$, $\mathcal{A}$, and thus $\bar{\mu}$ as well. With the further simplification that $\bar{\mu}$ is slowly varying (such that its second-order derivatives are much smaller than its first order derivatives) and that $\phi \ll 1$, we arrive at the first-order equations

$$\bar{\mu} \nabla^2 \Omega \sim -\hat{\mathbf{z}} \cdot [\boldsymbol{\nabla} \bar{\mu} \times \boldsymbol{\nabla} S] \tag{49}$$

$$\nabla^2 G \sim -2\phi \boldsymbol{\nabla} \bar{\mu} \cdot \boldsymbol{\nabla} S \tag{50}$$

Equation (49) immediately implies the excitation of plate-like vorticity requires gradients in viscosity to be significantly out of phase with (i.e., orthogonal to) gradients in the source-sink

field. However, gradients in $\bar{\mu}$ that are in phase with gradients in $S$ can drive dilation which affects $\bar{\mu}$ (through changes in porosity $\phi$), which potentially increases the magnitude of $\boldsymbol{\nabla}\bar{\mu}$, leading to more dilation, and thus a potential positive feedback even in the absence of damage. However, the feedback will only tend to generate porosity and weak zones over the source and sink, but not along bands connecting the ends of the source and sink (which are needed to make a contiguous weak plate boundary). Excitation of vorticity does not have a similar direct feedback through the first order effects in (49) because a $\boldsymbol{\nabla}\bar{\mu}$ that is out of phase with $\boldsymbol{\nabla}S$ can generate vorticity but the vorticity has no way in turn of affecting $\bar{\mu}$ and $\boldsymbol{\nabla}\bar{\mu}$ through first order effects; the vorticity can only influence $\bar{\mu}$ through second-order damage terms (i.e., deformational work $\Phi$) which drive changes in $\phi$ and/or $\mathcal{A}$. Therefore, while dilation is enhanced readily even without damage, vorticity and toroidal motion can only grow through a positive feedback if the appropriate damage mechanisms are present. Moreover, as is shown below, the excitation of dilation is intrinsically adverse to generation of plate-like toroidal motion. Therefore, not only must one must attain conditions for exciting toroidal motion, but ideally these same conditions simultaneously suppress dilational flow.

### 3.4. Numerical method

The governing equations constitute a set of two-dimensional nonlinear partial differential equations and they are solved numerically using a basic spectral transform method employing fast Fourier transforms, as described in *Bercovici* [1993]. For nonlinear equations, such as (41) and (42), one writes the equations so that all nonlinear terms constitute a forcing function in an apparently linear Poisson's equation. Thus for example, one would replace the left side of (41) with $\bar{\mu}_{\max}\nabla^2\Omega$ (where $\bar{\mu}_{\max} = \max(\bar{\mu})$) and add the complementary nonlinear term $(\bar{\mu}_{\max} - \bar{\mu})\nabla^2\Omega$ to the right side; this leads to a linear Laplacian operation on the left, driven by a forcing function on the right that is composed of nonlinear terms. This forcing function is

easily evaluated on a physical grid and Fourier transformed to spectral or wave-number space, although care must be taken to dealias the nonlinear forcing function after it is transformed [see *Canuto et al.*, 1988]. The Fourier transform of the Laplacian operator on the left side becomes an algebraic expression in spectral space (in terms of wave-numbers squared), and thus the spectral (tranformed) version of $\Omega$ is easily solved.

Numerical solutions of the governing equations are accomplished iteratively. For a given $\mathcal{A}$ and/or $\phi$ (depending on whether we are solving for void-generating or *fineness*-generating damage), the solutions for $\psi, \theta, \Omega$ and $G$ are found by iterating through (40), (41), and (42) until convergence (the mean-square misfit between successive iterations for solutions for $\Omega$ and $G$, normalized by the L2-norm of $\Omega$ and $G$, respectively, reaches $10^{-6}$ or less). Obviously, $\xi$ needs to be solved only once since $S$ is given. Provided these solutions, $\mathcal{A}$ and/or $\phi$ are then updated by one time step using (47) and (48), and the next cycle begins again. The time-step is tightly constrained by the CFL condition (usually to 1% of the CFL advective time step).

The driving source-sink field $S$ is prescribed as it is in *Bercovici* [1993] and *Bercovici* [1998]. In essence, it is taken from the horizontal divergence of an arbitrary plate-like velocity field defined by the translational motion of a square plate that is $2L$ on its side; i.e.,

$$\bar{\mathbf{v}} = V\hat{\mathbf{x}}' F(x') F(y') \tag{51}$$

where

$$F(x) = \frac{1}{2}\left[\tanh\left(\frac{x+L}{\delta}\right) - \tanh\left(\frac{x-L}{\delta}\right)\right] \tag{52}$$

$\delta$ is the width of the plate margin; $x'$ and $y'$ are coordinates in a frame of reference with unit vectors parallel ($\hat{\mathbf{x}}'$) or perpendicular ($\hat{\mathbf{y}}'$) to plate motion; if plate motion is at an angle $\vartheta$ from the $x$ axis, then obviously $x' = x\cos\vartheta + y\sin\vartheta$ and $y' = -x\sin\vartheta + y\cos\vartheta$. The source sink field is defined as $S = \boldsymbol{\nabla}\cdot\bar{\mathbf{v}}$, and since this a scalar invariant it simply becomes $S =$

$V\frac{dF(x')}{dx'}F(y')$, where

$$\frac{dF(x)}{dx} = \frac{1}{2\delta}\left[\mathrm{sech}^2\left(\frac{x+L}{\delta}\right) - \mathrm{sech}^2\left(\frac{x-L}{\delta}\right)\right] \tag{53}$$

In dimensionless form $-1 \leq S \leq 1$, and thus the constant $V$ is determined so that $\max(S) = 1$.

Lastly we can use (51) to measure the extent to which a plate-like vorticity field is generated. Ideally, our nonlinear solutions will at least reproduce the vorticity intrinsic to the given plate flow prescribed by (51); this vorticity is simply

$$\bar{\Omega} = \hat{\mathbf{z}} \cdot \boldsymbol{\nabla} \times \bar{\mathbf{v}} = -VF(x')\frac{dF(y')}{dy'} \tag{54}$$

which is identical to $S$ but rotated clockwise $90^o$. Whether the generated vorticity field $\Omega$ reproduces (or is even more plate-like) than $\bar{\Omega}$ can be measured by the correlation function

$$C_\Omega = \frac{\sum_{\mathbf{k}} \Omega_{\mathbf{k}} \bar{\Omega}^*_{\mathbf{k}}}{\sqrt{\left(\sum_{\mathbf{k}} |\Omega_{\mathbf{k}}|^2\right)\left(\sum_{\mathbf{k}} |\bar{\Omega}_{\mathbf{k}}|^2\right)}} \tag{55}$$

where $\Omega_{\mathbf{k}}$ and $\bar{\Omega}_{\mathbf{k}}$ are the 2D Fourier transforms of $\Omega$ and $\bar{\Omega}$, respectively. This function is in fact written to only measure the correlation in spatial stucture between $\Omega$ and $\bar{\Omega}$ and removes amplitude effects.

## 4. Numerical Source-Sink Flow Results

We examine below several numerical solutions to our system of equations. We will not do a complete exploration of parameter space, since a handfull of solutions will suffice to illustrate the most important results. We note that with our choice of scaling, the dimensionless surface tension parameter $\hat{\sigma}$ is in fact negligible, which is to be expected given that $\sigma \approx 1\mathrm{J/m}^2$, $\mathcal{A}_0 \leq 10^6\mathrm{m}^{-1}$ (given by inverse grainsize), $S_{\max} \geq 1 \times 10^{-14}\mathrm{s}^{-1}$ (using a slow plate velocity and a 100km-thick plate boundary), and $\mu \approx 10^{25}$Pa s is a typical lithospheric viscosity [*Beaumont*, 1976; *Watts et al.*, 1982]. Therefore, although terms in the momentum equations proportional to

$\hat{\sigma}$ govern any self-healing (see Appendix A), the effect is very slow and small over the timescales of the calculations.

### 4.1. Void-generating damage

All numerical solutions are started with the same weak random initial conditions depicted in Figure 1a. (In fact the figure shows a solution after one small time step which is why there is some signature of $S$ in the structure of $\Omega$ and $G$.) Since the intial $\Omega$ and $G$ are very small the velocity field is essentially dipolar as it is dominated by the source-sink field $S$. Causing the dipolar velocity field to assume a more solid-body translational velocity field is one of the primary goals of this study.

Given a small value of $\hat{\sigma}$ and finite partitioning of deformational work toward void creation ($f^* > 0$, $f_{\mathcal{A}} = 0$), we find that the most important controlling free parameter is that which controls partitioning variability, $\hat{\gamma}$ (see equation (46)). For large $\hat{\gamma}$, i.e., $O(1)$ or larger, the full energy partitioning is difficult to obtain (i.e., $f_\phi$ cannot readily reach the maximum value $f^*$). The system undergoes very little damage or evolution other than advection of porosity by the dipolar velocity field; both the dilation and vorticity fields $G$ and $\Omega$ remain weak, and the low to moderate (albeit weakly oscillating) vorticity correlation $C_\Omega$ indicates poor plate-like vorticity (Figure 2). Overall, the source-sink field $S$ continues to dominate the velocity field, leaving it dipolar, even after significant time integration (Fig. 1b).

For $\hat{\gamma} \ll 1$, the full energy partitioning toward damage is allowed and the system evolves rapidly (see Figs. 1 and 2). However, most of the damage and associated dilation concentrates on the driving source-sink field. The effect of this dilation on the velocity field is to reinforce the source flow, and diminish the sink flow. (Note that mass conservation is not violated; i.e., no mass-flux is added to or subtracted from the source-sink field since the dilational flow is driven by expansion of voids which have no mass.) This causes the horizontal velocity field

to assume an almost monopolar flow. Monopolar flow is highly adverse to the generation of toroidal motion because of its high degree of axisymmetry; i.e., an axisymmetric flow $v(r)\hat{\mathbf{r}}$, where $r$ is cylindrical radius from a vertical axis, has zero vertical vorticity $\hat{\mathbf{z}} \cdot \boldsymbol{\nabla} \times (v(r)\hat{\mathbf{r}})$. This is well demonstrated both by the low amplitude of the vorticity field $\Omega$ relative to dilation $G$, as well as by the very poor vorticity correlation $C_\Omega$ (Fig 2). In all cases involving void generation, dilational flow is generated by the damage mechanism, but does not reinforce vertical vorticity and thus the toroidal motion always remains relatively small (Figure 2).

### 4.2. *Fineness*-generating damage

We next consider cases in which all deformational work goes to *fineness*-generating interface growth, i.e., $f_\phi = 0$ and $f_\mathcal{A} > 0$. Although $\hat{\sigma}$ is assumed small (and terms proportional to it negligible in the momentum equation), we do not assume that $f_\mathcal{A}/\hat{\sigma}$ is excessively large, although we do explore its effect over variations of several orders of magnitude. Indeed, in these cases, the most significant controlling parameters appear to be both the the viscosity exponent $m$ from (22), and the ratio $f_\mathcal{A}/\hat{\sigma}$ from (48).

In the previous cases of void-generating damage, deformational work directly fed dilational flow but not toroidal motion; thus the dilation rate would tend to be dominant while toroidal motion would only be excited secondarily. However, in the *fineness*-generating cases, damage does not directly force either dilation or vorticity and instead only influences them through the viscosity field (i.e., via the $\mathcal{A}$-dependence of viscosity), and in particular by the coupling of the driving poloidal flow with viscosity gradients. Thus, vorticity $\Omega$ is allowed to develop on more or less equal footing with the dilation rate.

#### 4.2.1. Influence of the *fineness*-dependence of viscosity

We first consider varying $m$, the $\mathcal{A}$-dependence viscosity exponent, while keeping $f_{\mathcal{A}}/\hat{\sigma} = 1$ (a very conservative value). For $m = 1$, plate-like motion is significantly improved over the void-generating cases (Fig. 3) but is not very well developed. The *fineness* variable $\mathcal{A}$ grows several-fold with time (Fig. 4) but because of the weak viscosity dependence on $\mathcal{A}$, it does not register a large effect on the vorticity and dilation rate. Indeed, while both the vorticity and dilation rate increase in magnitude with time, they remain small relative to the source-sink (poloidal) flow during the duration of the calculation.

For $m$ considerably greater than unity the fluid motion becomes increasingly plate-like. We depict two other cases showing the effect of increasing $m$, i.e., for $m = 7$ and $m = 21$ (Figs 3, 4). For sufficiently large $m$, the vorticity field grows in magnitude and becomes increasingly organized into strike-slip-type margins bounding a plate; this effect is apparent both qualitatively (Figs. 3b and c) and quantitatively in terms of the correlation function $C_\Omega$ (Fig. 4). Moreover, the dilation rate actually decreases with time for some period (Fig. 4); the extent and duration to which the dilation rate is suppressed is clearly dependent on the size of $m$ (see §4.2.3). For the case of $m = 21$ the plate-like velocity and vorticity fields along with suppression of dilation is highly pronounced.

### 4.2.2. Influence of partitioning $f_{\mathcal{A}}$

If our viscosity law for $\mu_m$ is characteristic of diffusion creep then $m$ should be between 2 and 3, while values of 7 or 21 would be unrealistic. It is plausible that dependence on $\mathcal{A}$ would be more severe than stipulated in diffusion creep if the deformation mechanism involved grain sliding as well. However, solutions for values of $m$ = 2-3 do not differ significantly from the weakly- or even non-plate-like solutions of $m = 1$ as long as the ratio $f_{\mathcal{A}}/\hat{\sigma}$ is small or moderate. Nevertheless, as stated above, the value of $f_{\mathcal{A}}/\hat{\sigma} = 1$ is very conservative since $\hat{\sigma}$ is considered extremely small. If we allow $f_{\mathcal{A}}$ to be 10 or 100 times larger than $\hat{\sigma}$ (which still

permits $f_{\mathcal{A}} \leq 1$ if not $\ll 1$) we obtain a very profound effect, even for small $m$, that is perhaps even more promising in terms of plate-generation than imposing high values of $m$. Figure 5 depicts cases with $m = 3$ and $f_{\mathcal{A}}/\hat{\sigma} = 1, 10$ and $100$; it is quite clear that as the ratio $f_{\mathcal{A}}/\hat{\sigma}$ increase the vorticity field focuses into very plate-like strike-slip pseudo-faults, the velocity field is very nearly solid-body translation, and the viscosity field, which will essentially be the reverse image of the $\mathcal{A}$ field, has contiguous weak zones around the plate margins.

The temporal evolution of the solutions with different $f_{\mathcal{A}}/\hat{\sigma}$ (Fig 6) shows that the magnitude of $\mathcal{A}$ increases more rapidly with larger $f_{\mathcal{A}}/\hat{\sigma}$, as expected by (48); indeed a simple scaling shows that the for fixed strain-rates, the magnitude of $\mathcal{A}$ goes as $\left(\frac{f_{\mathcal{A}}}{\hat{\sigma}}t\right)^{\frac{1}{m+1}}$. Vorticity likewise reaches large magnitudes, potentially larger than both the magnitudes of $S$ as well as $G$ (i.e., for the $f_{\mathcal{A}}/\hat{\sigma} = 10$ and 100 cases). The vorticity correlation function $C_\Omega$ also reaches respectably high values near unity, although it peaks at these values and then slowly decreases with time. Finally, the time span over which dilation $G$ is suppressed increases dramatically with larger $f_{\mathcal{A}}/\hat{\sigma}$.

### 4.2.3. Suppression of dilation

As discusses already, suppression of dilation in the poloid-toroidal coupling problem with two-phase damage is crucial, since otherwise dilation merely augments the poloidal field at the expense of the toroidal one. However, since dilation $G$ is directly forced by the coupling of viscosity gradients with gradients in the driving source-sink field $S$ – see (42) or (50) in §3.3 – it will eventually grow, thus its suppression is only temporary. However, the time period of such suppression is variable, and depends significantly on both $m$ and $f_{\mathcal{A}}/\hat{\sigma}$ such that the higher the value of either, the longer the suppression period. We can measure the dilation-suppression

time $t_{\rm supp}$ by defining it as the period of time that $\frac{\partial G_{\max}}{\partial t} < 0$, i.e.,

$$t_{\rm supp} = \int_0^{t_{\max}} \frac{1}{2}\left[1 - \text{sign}\left(\frac{\partial G_{\max}}{\partial t}\right)\right] dt \tag{56}$$

which can be numerically integrated for any of our solutions. Figure 7 shows $t_{\rm supp}$ versus $m$ for $f_{\mathcal{A}}/\hat{\sigma} = 1$, and versus $f_{\mathcal{A}}/\hat{\sigma}$ for $m = 3$. The suppression of dilation is particularly profound and long-lasting as $f_{\mathcal{A}}/\hat{\sigma}$ increases to large values, even for moderate values of $m$. Of course, as $m$ also increases the suppression is more significant; but it appears the effect of using large enough $f_{\mathcal{A}}/\hat{\sigma}$ is sufficient to suppress dilation almost entirely or at least for long periods of time.

#### 4.2.4. Some reflections on interface curvature

As discussed earlier in this paper (see §2.2), the association of $\mathcal{A}$ with voidsize or grainsize depends on the interface curvature $\partial\alpha/\partial\phi$; if the curvature is greater than zero $\mathcal{A}$ represents voidsize, while if less than zero it is grainsize. In the above calculations we considered relatively simple cases in which $a = b = 1/2$ (see (7)) for both void and *fineness* generating damage. However, with these values of $a$ and $b$, the curvature changes sign at relatively high porosity (i.e., at $\phi = a/(a+b) = 1/2$), even though our porosity field never approaches such values. Thus in these cases $\mathcal{A}$ really represents voidsize. It is important therefore to consider cases in which curvature switches sign at low porosity, as would be expected for silicates and granular media. We therefore examined cases for which $a = 0.05$ and $b = 1 - a$, in which case curvature changes sign at $\phi = 0.05$, which is also equivalent to our initial background porosity field. The effect of this change in $a$ and $b$ is, however, not extremely significant. In essence, it simply changes the value of $\eta$ in (48). For $a = b = 1/2$, the value of $\eta$ at the background porosity value of $\phi = 0.05$ is relatively small ($\eta(0.05) \approx 0.2$), thus leading to a larger damage rate. For $a = 0.05, b = 1 - a$, the value of $\eta$ at the background porosity is considerably larger (approximately $0.8$), thus reducing the damage rate. The extent of plate generation is thus slightly weaker for

the $a = 0.05$ cases, although the solutions differ little from those shown already. However, even this mitigating effect can be offset easily by increasing $f_{\mathcal{A}}/\hat{\sigma}$ accordingly (i.e., our choices of $f_{\mathcal{A}}/\hat{\sigma}$ are conservative, given that that $\hat{\sigma} \ll 1$, and thus there is nothing implausible about increasing this ratio by a factor of 4 or so).

## 5. Discussion and Conclusion

The overall goal of our two-phase damage studies is to develop a first-principles theory for shear localization as it might be applied to mantle-lithosphere dynamics, in particular to the generation of plate tectonics from mantle convection. Our previous studies of two-phase damage [*Bercovici et al.*, 2001a, b; *Ricard et al.*, 2001; *Bercovici and Ricard*, 2003; *Ricard and Bercovici*, 2003] demonstrated a variety of important shear-localizing and failure-like phenomena that involve damage through void generation. However, these findings were most applicable to situations where an imposed shear or other deformation gets further enhanced by dilational damage. In the plate-generation case, whereby a toroidal flow field must be generated from an essentially orthogonal convective poloidal field, the presence of dilation associated with void generation acts adversely to generating plate-like behavior. However, allowing for damage to affect grain/void size by increasing the overall *fineness* $\mathcal{A}$ not only leads to a very profound organization of vorticity into plate-like motion, but it also suppresses dilation and its adverse effects.

There are, however, a few model simplifications and assumptions that warrant further discussion, at least in terms of the conclusions summarized above. First, it is likely that the effective viscosity of a silicate-type mixture or granular medium would be more sensitive to porosity than the $1 - \phi$ dependence in (21) implies . For example, the matrix of partial melts will tend to disaggregate at low porosities, which will involve a sharp drop in effective viscosity. In

the case of void-generating damage ($f_\phi > 0$ and $f_\mathcal{A} = 0$), a very low viscosity in damaged and weakened zones of moderate porosity could conceivably suppress dilation by limiting (or self-regulating) the amount of deformational work going into void generation; meanwhile, the resulting large viscosity variations and gradients resulting from precipitous drops in viscosity over small changes in porosity might enhance vorticity generation. These combined effects could possibly then allow void-generating damage to better generate plate-like motion. The effect of using a more strongly porosity-dependent effective viscosity was tested for the pure void-generating damage case ($f_\mathcal{A} = 0$) by replacing the $1 - \phi$ dependence in (21) (or, to be more precise, in the momentum-equation terms involving $\underline{\underline{\tau}}_m$) with a function of porosity that undergoes a sharp drop at an arbitrary value of porosity; i.e., instead of in essence using the effective viscosity $\mu_{\text{eff}} = \mu_m(1 - \phi)$, we employed

$$\mu_{\text{eff}} = \mu_m \frac{1}{2}\left[1 - \tanh\left(\frac{\phi - \phi^*}{\delta_\phi \phi(1-\phi)}\right)\right] \tag{57}$$

where $\phi^*$ is the transitional porosity (i.e., at which the sharp drop in viscosity occurs) and $\delta_\phi$ controls the sharpness of the viscosity drop; the argument of the hyperbolic tangent goes as $\phi^{-1}(1-\phi)^{-1}$ to insure that $\mu_{\text{eff}}$ goes to $\mu_m$ as $\phi \to 0$, and to 0 as $\phi \to 1$. To maximize the effect we are seeking (i.e., to obtain larger viscosity contrast with which to enhance vorticity generation, while still allowing for low viscosity and self-regulation of dilation) we chose the transitional porosity $\phi^*$ to coincide with our chosen background porosity of $\phi = 0.05$. The results of these calculations (for various $\delta_\phi$) do indeed show some suppression of dilation rate and enhancement of vorticity. However, the vorticity always remains considerably less than the dilation rate, leading to a velocity field that is still dominantly monopolar and thus highly unplate-like.

Second, we argued in §2.5 that the effective bulk viscosity $B$ is, unlike shear viscosity, independent of grain/voidsize $\mathcal{A}^{-1}$. For the sake of completeness, it is therefore appropriate to consider the alternative case when $B$ depends on $\mathcal{A}$ in the same way as shear viscosity (i.e., $B \sim \mu_m$ instead of $B \sim \mu_0$). This adjustment was adopted for the *fineness*-generating damage case ($f_{\mathcal{A}} > 0$) and explored for a variety of parameters similar to those shown above. The effect of allowing $B$ to depend on $\mathcal{A}$ is that it lowers the bulk viscosity as $\mathcal{A}$ grows; since $B$ primiarly acts to retard dilational motion, the end result is that dilation is not as well suppressed as when $B$ is independent of $\mathcal{A}$. The resulting fluid motion is more plate-like than the void-generating cases, but significantly worse than cases with $\mathcal{A}$-independent $B$. This illustrates the importance of suppressing dilational motion for generating plates.

Although void-generating, dilational damage is prevalent in low-pressure crustal rocks [e.g., *Menéndez et al.*, 1996; *Wong et al.*, 1997; *Zhu and Wong*, 1997; *Zhu et al.*, 1997] it is probably much less significant at high confining pressures of mid and lower lithospheric depths (of order 100km). Thus, the failure of void-generating damage to generate plates is in keeping with the notion that void generation is probably not the dominant mechanism for deep lithospheric deformation. Non-void and non-dilational damage, by, for example, grainsize reduction is therefore an attractive mechanism. However, grainsize reduction through dynamic recrystallization, where the driving mechanism is kinematic and a linear function of strain [*Karato et al.*, 1980; *Karato*, 1989, 1996; *Bercovici and Karato*, 2003], is both a complicated mechanism (e.g., it resides on the boundary between diffusion and dislocation creep) and the theories that have been developed for it do not readily permit the sort of shear localization observed in plate generation and in mylonites [*Bresser et al.*, 2001; *Montési and Zuber*, 2002], although the possibilities for this mechanism have by no means been exhausted [see *Bercovici and Karato*, 2003]. The grainsize reduction mechanism explored in this paper is perhaps more straightfor-

ward (it basically involves pulverizing the matrix through deformational work), and our simple calculations demonstrate exceptional potential for plate generation. Plate-like behavior appears to improve as the dependence of viscosity on $\mathcal{A}$, parametrized by the power $m$, increases; for very small damage partition values $f_{\mathcal{A}}$ (i.e., for $f_{\mathcal{A}}/\hat{\sigma} \sim O(1)$, where we have assumed $\hat{\sigma} \ll 1$) one must approach values of $m$ that are beyond those that are realistic for diffusion creep (for which typically $m = 2$-$3$). The higher values of $m$ can be interpretted to represent the effects of introducing more slip planes and grain sliding rather than diffusion creep; however, these are a qualitative arguments only and require more rigorous exploration. Nevertheless, with an increase in $f_{\mathcal{A}}$ by an order of magnitude or more (which is still plausible since $f_{\mathcal{A}}$ could still be very small if it is 10 or 100 times larger than $\hat{\sigma}$), even values of $m$ typical of diffusion creep (e.g., $m = 3$) not only yield extremely plate-like behavior (in the velocity, vorticity and viscosity fields) but also extensive periods of dilation suppression which are also necessary to sustain plate generation. Indeed, in these model calculations the cases with larger $f_{\mathcal{A}}$, even when $m$ is relatively small, are the most successful at plate generation.

## Appendix A: *Fineness* and healing

In the case with *fineness* generating damage, the evolution equation for $\mathcal{A}$ does not appear to have an explicit healing term. Such healing would be manifested as reduction in interfacial surface area as the system tends toward a minimum surface energy configuration. In rocks such healing tends to occur through grain growth wherein large grains grow at the expense of smaller ones; e.g., surface tension causes interior pressures of small grains to be large relative to large grains, and given any effective grain-grain contact (either physical, or chemical through surrounding solvents as in Ostewaeld ripening), such pressure differences will cause the small grain to empty its contents into the larger one *Karato* [1989]; *Axe and Yamada* [1986].

In simple treatments of grain growth, healing is typically represented as a decay term that goes as inverse of average grainsize to some power [see *Karato*, 1996; *Bercovici and Karato*, 2003], which accounts for grain growth being faster the smaller the grainsize. However, grain growth requires some initial heterogeneity in grainsizes (i.e., a grainsize distribution of finite width) to proceed, otherwise all grains would be of the same size and none could grow at the expense of another [*Axe and Yamada*, 1986]. If grainsize is parameterized in a mixture scheme by a quantity such as $\mathcal{A}^{-1}$, then this actually represents average grainsize in a control volume and thus all information about microscale heterogeneity is lost (i.e., $\mathcal{A}^{-1}$ is only the mean of the distribution, and has no information about the width, or standard deviation, of the distribution). Thus the heterogeneity required to initiate grain growth must either be treated by specifying yet another quantity to represent the grainsize distribution width (i.e., standard deviation), or by only treating the larger-scale heterogeneity as represented by gradients in $\mathcal{A}^{-1}$.

In this paper we will refrain from introducing even more complexity incurred by developing equations for the standard deviation in grainsize (or *fineness* ) distribution. However, we will show with simple analysis how the the basic evolution equation for $\mathcal{A}$, i.e., (27) implicitly includes a healing term that depends on grainsize or *fineness* heterogeneity, which in our case is represented by $\boldsymbol{\nabla}\mathcal{A}$. In particular, we will examine the nature of flow driven only by gradients in $\mathcal{A}$, and see how that influences the evolution equation for $\mathcal{A}$. For the sake of focussing on the essential physics, we assume there is no damage (i.e., $f_{\mathcal{A}} = f_\phi = 0$); other simplifying assumptions will be stated as needed.

First, to consider the nature and direction of flow driven by gradients in $\mathcal{A}$, we examine the three-dimensional equations of motion. For simplicity, we assume that porosity $\phi$ is both spatially and temporally constant (and thus $\boldsymbol{\nabla} \cdot \mathbf{v}_m = 0$ as well, by (24)), and that deviations in $\mathcal{A}$ from the background value of $\mathcal{A}_0$ are small enough that viscosity $\mu_m$ remains essentially con-

stant (i.e., nonlinear products of $\boldsymbol{\nabla}\mu_m$ and $\boldsymbol{\nabla}\mathbf{v}_m$ are negligible in the force-balance equation). With these assumptions, the combination of (25) and (26) leads to

$$0 = \sigma\lambda\boldsymbol{\nabla}\mathcal{A} + \mu_0(1-\phi)\nabla^2\mathbf{v}_m \tag{A1}$$

where $\lambda = \eta + (1-\phi)\frac{d\eta}{d\phi}$. Let us define the unit vectors $\hat{\mathbf{p}}$ and $\hat{\mathbf{n}}_i$ ($i = 1$ or 2) which are parallel and perpendicular to $\boldsymbol{\nabla}\mathcal{A}$, respectively (e.g., $\hat{\mathbf{p}} = \boldsymbol{\nabla}\mathcal{A}/|\boldsymbol{\nabla}\mathcal{A}|$); let us also assume that the direction of $\boldsymbol{\nabla}\mathcal{A}$ is slowly varying such that these unit vectors remain spatially constant and thus comprise a cartesian system. In this case (A1) can be decomposed into parts parallel and perpendicular to $\boldsymbol{\nabla}\mathcal{A}$:

$$0 = \sigma\lambda|\boldsymbol{\nabla}\mathcal{A}| + \mu_0(1-\phi)\nabla^2 v_p \text{and} 0 = \mu_0(1-\phi)\nabla^2 v_{n_i} \tag{A2}$$

where $v_p = \hat{\mathbf{p}}\cdot\mathbf{v}_m$ and $v_{n_i} = \hat{\mathbf{n}}_i\cdot\mathbf{v}_m$ (i.e., in total, $\mathbf{v}_m = v_p\hat{\mathbf{p}} + v_{n_1}\hat{\mathbf{n}}_1 + v_{n_2}\hat{\mathbf{n}}_2$). With no driving boundary conditions for the $v_{n_i}$, (A2) implies they would both be zero. Since our assumption of constant $\phi$ requires $\boldsymbol{\nabla}\cdot\mathbf{v}_m = 0$, then we find that $\frac{\partial v_p}{\partial x_p} = 0$ where $x_p = \hat{\mathbf{p}}\cdot\mathbf{x}$; hence (A2) reduces to

$$0 = \sigma\lambda|\boldsymbol{\nabla}\mathcal{A}| + \mu_0(1-\phi)\nabla_n^2 v_p \tag{A3}$$

where $\nabla_n^2 = \nabla^2 - \frac{\partial^2}{\partial {x_p}^2}$. If $v_p$ is maximum where it is being most strongly forced by $\boldsymbol{\nabla}\mathcal{A}$ then it will also diminish with distance from the location of peak forcing, in which case the last term on the right side of (A3) must have the opposite sign from $v_p$ (i.e., curvature for a function around its maximum is necessarily negative). If we assume the length scale over which changes in $v_p$ occur is $R$, then a scaling analysis allows us to write $\nabla_n^2 v_p \sim -v_p/R^2$ which suggests that

$$v_p \sim \frac{R^2\sigma\lambda}{\mu_0(1-\phi)}|\boldsymbol{\nabla}\mathcal{A}| \tag{A4}$$

and thus $v_p > 0$, i.e., $\mathbf{v}_m$ is in the same direction as $\boldsymbol{\nabla}\mathcal{A}$. This effect can be understood by recalling that the surface tension is a line force acting outward on the edges of a segment of

interface; for a small volume of mixture, this tension acts on the edges of interface that are cut by the surface of the control volume [*Bercovici et al.*, 2001a]. As *fineness* increases (i.e., the larger the value of $\mathcal{A}$) there is more interface edge exposed at the volume surface, leading to larger net tension on that surface; thus the net force points from the control volume's surfaces with small $\mathcal{A}$ to ones with large $\mathcal{A}$, or along $\boldsymbol{\nabla}\mathcal{A}$.

Without damage, the evolution equation for $\mathcal{A}$, i.e., (27), becomes

$$\frac{\partial \mathcal{A}}{\partial t} = -\mathbf{v}_m \cdot \boldsymbol{\nabla}\mathcal{A} = -v_p \hat{\mathbf{p}} \cdot \boldsymbol{\nabla}\mathcal{A} = -v_p |\boldsymbol{\nabla}\mathcal{A}| \tag{A5}$$

or using (A4)

$$\frac{\partial \mathcal{A}}{\partial t} \sim -\frac{R^2 \sigma \lambda}{\mu_0 (1-\phi)} |\boldsymbol{\nabla}\mathcal{A}|^2 \tag{A6}$$

These relations demonstrate that material motion associated with gradients in $\mathcal{A}$ leads to a healing (grainsize growth, or matrix coarsening and loss of *fineness*) that is driven by surface tension, retarded by viscosity, and requires heterogeneity in *fineness* via gradients in $\mathcal{A}$.

Equation (A6) has other interesting implications; in particuar, since $\mathcal{A}$ only decays where $\boldsymbol{\nabla}\mathcal{A} \neq 0$, then it will remain constant at extrema in $\mathcal{A}$ (assuming no damage). This means that the healing effect will in fact sharpen peaks in $\mathcal{A}$, hence actually facilitating shear localization while healing the host matrix. This concept is quantitatively demonstrated by considering a one-dimensional analysis. We assume that all variables only depend on $x$ and $t$, and flow is only in the $x$ direction (i.e., $\mathcal{A} = \mathcal{A}(x,t)$ and $\mathbf{v}_m = u(x,t)\hat{\mathbf{x}}$). We assume that gradients in $\mathcal{A}$ are caused by a single perturbation to an infinitely long background of constant value $\mathcal{A}_0$, i.e., $\mathcal{A} = \mathcal{A}_0 + \mathcal{A}_1(x,t)$, and that $\mathcal{A}_1 \to 0$ as $x \to \pm\infty$. Given our 1-D analysis, we can relax some of our other assumptions and state that while $\phi$ is spatially constant it need not be temporally constant; moreover, we can retain for now the dependence of $\mu_m$ on $\mathcal{A}$ (although this assumption is not as important in a 1-D analysis, as shown below). In this case, the combination

of (25) and (26) in 1-D leads to

$$0 = \sigma\lambda\frac{\partial \mathcal{A}}{\partial x} + (1-\phi)\frac{\partial}{\partial x}\left[\left(\frac{\mu_0}{\phi} + \frac{4}{3}\mu_m\right)\frac{\partial u}{\partial x}\right] \tag{A7}$$

where we have used (23) with $K = 1$ [*Bercovici et al.*, 2001a]. Integration over $x$ leads to

$$\frac{\partial u}{\partial x} = \frac{\sigma\lambda(\mathcal{A}_0 - \mathcal{A})}{\mu_0(1-\phi)\left[\frac{1}{\phi} + \frac{4}{3}\left(\frac{\mathcal{A}_0}{\mathcal{A}}\right)^m\right]} \tag{A8}$$

where we use (22) and have employed the boundary condition that $\frac{\partial u}{\partial x} \to 0$ as $\mathcal{A} \to \mathcal{A}_0$ (i.e., as $x \to \pm\infty$). If we make the plausible assumption that $\phi \ll 1$ but that $\mathcal{A}$ is nowhere $\ll \mathcal{A}_0$, then we obtain

$$\frac{\partial u}{\partial x} = \frac{\sigma\lambda\phi(\mathcal{A}_0 - \mathcal{A})}{\mu_0} \tag{A9}$$

We now state that the perturbation in $\mathcal{A}$ is bell-shaped in $x$, but since decay only occurs where $\boldsymbol{\nabla}\mathcal{A} \neq 0$ that the amplitude of the perturbation remains constant while its width changes with time; i.e.,

$$\mathcal{A} = \mathcal{A}_0 + \mathcal{A}^*\mathrm{sech}^2\left(\frac{x}{\delta(t)}\right) \tag{A10}$$

where $\mathcal{A}^*$ is a constant. Substituting (A10) into (A9) and integrating, we arrive at

$$u = -\frac{\sigma\lambda\phi\mathcal{A}^*\delta(t)}{\mu_0}\tanh\left(\frac{x}{\delta(t)}\right) \tag{A11}$$

where we assume $u = 0$ at $x = 0$ (i.e., at the peak in $\mathcal{A}$). The evolution equation for $\mathcal{A}$ without damage is, in 1-D, $\frac{\partial \mathcal{A}}{\partial t} = -u\frac{\partial \mathcal{A}}{\partial x}$, which with (A10)-(A11) becomes

$$\frac{1}{\delta}\frac{d\delta}{dt} = -\frac{\sigma\lambda\phi\mathcal{A}^*}{\mu_0}\frac{\tanh(x/\delta)}{x/\delta} \tag{A12}$$

In the region of interest around the perturbation, i.e., $-\delta < x < \delta$ we can assume $\frac{\tanh(x/\delta)}{x/\delta} \sim O(1)$ and thus

$$\delta \approx \delta_0 e^{-\frac{\sigma\lambda\phi\mathcal{A}^*}{\mu_0}t} \tag{A13}$$

Hence the pertubation in $\mathcal{A}$ becomes sharper while the host matrix coarsens and heals. Obviously microscale heterogeneity in grainsize (i.e., a nonuniform grainsize distribution) would

induce further healing and cause the peak in $\mathcal{A}$ (as well as the background value $\mathcal{A}_0$) to decay away as well; however, as stated above, such variations over grain scales cannot be rigorously accounted for by only considering the average *fineness* or grainsize in a control volume.

**Acknowledgments.** The authors benefited from discussions with Shun Karato and Chris Hieronymus. Support was provided by the National Science Foundation (NSF, grant EAR-0105269), the Centre National de la Recherche Scientifique (CNRS), the Yale Amity Geophysical Annex (YAGA) and Le Garon Foundation for Supine Geophysics (GFSG).

## References

Ashby, M., and C. Sammis (1990), The damage mechanics of brittle solids in compression, *Pure Appl. Geophys.*, *133*, 489–521.

Auth, C., D. Bercovici, and U. Christensen (2003), Two-dimensional convection with a self-lubricating, simple-damage rheology, *Geophys. J. Int.*, *154*, 783–800.

Axe, J., and Y. Yamada (1986), Scaling relations for grain autocorrelation functions during nucleation and growth, *Phys. Rev.*, *B89*, 1599–1606.

Bailyn, M. (1994), *A Survey of Thermodynamics*, Am. Inst. Phys., College Park, MD.

Batchelor, G. (1967), *An Introduction to Fluid Dynamics*, Cambridge Univ. Press, New York.

Beaumont, C. (1976), The evolution of sedimentary basins on a viscoelastic lithosphere, *Geophys. J.R. Astron. Soc.*, *55*, 471–497.

Bercovici, D. (1993), A simple model of plate generation from mantle flow, *Geophys. J. Int.*, *114*, 635–650.

Bercovici, D. (1995a), A source-sink model of the generation of plate tectonics from non-newtonian mantle flow, *J. Geophys. Res.*, *100*, 2013–2030.

Bercovici, D. (1995b), On the purpose of toroidal flow in a convecting mantle, *Geophys. Res. Lett.*, *22*, 3107–3110.

Bercovici, D. (1996), Plate generation in a simple model of lithosphere-mantle flow with dynamic self-lubrication, *Earth Planet. Sci. Lett.*, *144*, 41–51.

Bercovici, D. (1998), Generation of plate tectonics from lithosphere-mantle flow and void-volatile self-lubrication, *Earth Planet. Sci. Lett.*, *154*, 139–151.

Bercovici, D. (2003), The generation of plate tectonics from mantle convection, *Earth Planet. Sci. Lett.*, *205*, 107–121.

Bercovici, D., and S. Karato (2003), Theoretical analysis of shear localization in the lithosphere, in *Reviews in Mineralogy and Geochemistry: Plastic Deformation of Minerals and Rocks*, vol. 51, edited by S. Karato and H. Wenk, chap. 13, pp. 387–420, Min. Soc. Am., Washington, DC.

Bercovici, D., and Y. Ricard (2003), Energetics of a two-phase model of lithospheric damage, shear localization and plate-boundary formation, *Geophys. J. Intl.*, *152*, 581–596.

Bercovici, D., and P. Wessel (1994), A continuous kinematic model of plate tectonic motions, *Geophys. J. Int.*, *119*, 595–610.

Bercovici, D., Y. Ricard, and M. Richards (2000), The relation between mantle dynamics and plate tectonics: A primer, in *History and Dynamics of Global Plate Motions, Geophys. Monogr. Ser.*, vol. 121, edited by M. A. Richards, R. Gordon, and R. van der Hilst, pp. 5–46, Am. Geophys. Union, Washington, DC.

Bercovici, D., Y. Ricard, and G. Schubert (2001a), A two-phase model of compaction and damage, 1. general theory, *J. Geophys. Res.*, *106*(B5), 8887–8906.

Bercovici, D., Y. Ricard, and G. Schubert (2001b), A two-phase model of compaction and damage, 3. applications to shear localization and plate boundary formation, *J. Geophys. Res.*,

*106*(B5), 8925–8940.

Braun, J., J. Chery, A. Poliakov, D. Mainprice, A. Vauchez, A. Tomassi, and M. Daignieres (1999), A simple parameterization of strain localization in the ductile regime due to grain size reduction: A case study for olivine, *J. Geophys. Res.*, *104*, 25,167–25,181.

Bresser, J. D., C. Peach, J. Reijs, and C. Spiers (1998), On dynamic recrystallization during solid state flow: effects of stress and temperature, *Geophys. Res. Lett.*, *25*, 3457–3460.

Bresser, J. D., J. ter Heege, and C. Spiers (2001), Grain size reduction by dynamic recrystallization: can it result in major rheological weakening?, *Intl. J. Earth Sci.*, *90*, 28–45.

Cadek, O., and Y. Ricard (1992), Toroidal/poloidal energy partitioning and global lithospheric rotation during cenozoic time, *Earth Planet. Sci. Lett.*, *109*, 621–632.

Canuto, C., M. Hussaini, A. Quarteroni, and T. Zang (1988), *Spectral Methods in Fluid Dynamics*, Springer-Verlag, New York.

Chrysochoos, A., and G. Martin (1989), Tensile test microcalorimetry for thermomechanical behaviour law analysis, *Mater. Sci. Eng.*, *A108*, 25–32.

Derby, B., and M. Ashby (1987), On dynamic recrystallization, *Scripta Metallurgica*, *21*, 879–884.

Desjonquères, M., and D. Spanjaard (1993), *Concepts in Surface Physics*, Springer-Verlag, New York.

Dumoulin, C., D. Bercovici, and P. Wessel (1998), A continuous plate-tectonic model using geophysical data to estimate plate margin widths, with a seismicity based example, *Geophys. J. Int.*, *133*, 379–389.

Fleitout, L., and C. Froidevaux (1980), Thermal and mechanical evolution of shear zones, *J. Struct. Geol.*, *2*, 159–164.

Forte, A., and W. Peltier (1987), Plate tectonics and aspherical earth structure: The importance of poloidal-toroidal coupling, *J. Geophys. Res.*, *92*, 3645–3679.

Furusho, M., and K. Kanagawa (1999), Reaction induced strain localization in a lherzolite mylonite from the hidaka metamorphic belt of central hokkaido, japan, *Tectonophysics*, *313*, 411–432.

Gable, C., R. O'Connell, and B. Travis (1991), Convection in three dimensions with surface plates: Generation of toroidal flow, *J. Geophys. Res.*, *96*, 8391–8405.

Géminard, J.-C., W. Losert, and J. Gollub (1999), Frictional mechanics of wet granular material, *Phys. Rev. E*, *59*, 5881–5890.

Griffith, A. (1921), The phenomenon of rupture and flow in solids, *Philos. Trans. R. Soc. London, Ser. A*, *221*, 163–198.

Gurnis, M., S. Zhong, and J. Toth (2000), On the competing roles of fault reactivation and brittle failure in generating plate tectonics from mantle convection, in *History and Dynamics of Global Plate Motions, Geophys. Monogr. Ser.*, vol. 121, edited by M. A. Richards, R. Gordon, and R. van der Hilst, pp. 73–94, Am. Geophys. Union, Washington, DC.

Hager, B., and R. O'Connell (1978), Subduction zone dip angles and flow driven by plate motion, *Tectonophysics*, *50*, 111–133.

Hager, B., and R. O'Connell (1979), Kinematic models of large-scale flow in the earth's mantle, *J. Geophys. Res.*, *84*, 1031–1048.

Hager, B., and R. O'Connell (1981), A simple global model of plate dynamics and mantle convection, *J. Geophys. Res.*, *86*, 4843–4867.

Hansen, N., and H. Schreyer (1992), Thermodynamically consistent theories for elastoplasticity coupled with damage, in *Damage Mechanics and Localization*, edited by J. Ju and K. Valanis, pp. 53–67, Am. Soc. of Mech. Eng., New York.

Huismans, R., and C. Beaumont (2003), Symmetric and asymmetric lithospheric extension: Relative effects of frictional-plastic and viscous strain softening, *J. Geophys. Res.*, *108*, 2496, doi:10.1029/2002JB002,026.

Jin, D., S. Karato, and M. Obata (1998), Mechanisms of shear localization in the continental lithosphere: Inference from the deformation microstructures of peridotites from the ivrea zone, northwestern italy, *J. Struct. Geol.*, *20*, 195–209.

Kameyama, M., D. Yuen, and H. Fujimoto (1997), The interaction of viscous heating with grain-size dependent rheology in the formation of localized slip zones, *Geophys. Res. Lett.*, *24*, 2523–2526.

Karato, S. (1983), Grain-size distribution and rheology of the upper mantle, *Tectonophysics*, *104*, 155–176.

Karato, S. (1989), Grain growth kinetics in olivine aggregates, *Tectonophysics*, *168*, 255–273.

Karato, S. (1996), Plastic flow in rocks, in *Continuum Physics, Earth and Planetary Sciences Series*, vol. 6, edited by T. Matsui, pp. 239–291, Iwani Shoten, Tokyo.

Karato, S., M. Toriumi, and T. Fujii (1980), Dynamic recrystallization of olivine single crystals during high temperature creep, *Geophysical Research Letters*, *7*, 649–652.

Kemp, D., and D. Stevenson (1996), A tensile flexural model for the initiation of subduction, *Geophys. J. Int.*, *125*, 73–94.

King, S. (2001), Subduction zones: observations and geodynamic models, *Phys. Earth Planet. Int.*, *127*, 9–24.

Krajcinovic, D. (1996), *Damage Mechanics*, North-Holland, Amsterdam.

Larson, R. G. (1999), *The structure and rheology of complex fluids*, Oxford University Press, New York.

Lemaitre, J. (1992), *A Course on Damage Mechanics*, Springer-Verlag, New York.

Lithgow-Bertelloni, C., M. Richards, Y. Ricard, R. O'Connell, and D. Engebretson (1993), Toroidal-poloidal partitioning of plate motions since 120 ma, *Geophys. Res. Lett.*, *20*, 375–378.

Lyakhovsky, V., Y. Ben-Zion, and A. Agnon (1997), Distributed damage, faulting, and friction, *J. Geophys. Res.*, *102*, 27,635–27,649.

McKenzie, D. (1984), The generation and compaction of partially molten rock, *J. Petrol.*, *25*, 713–765.

Menéndez, B., W. Zhu, and T.-F. Wong (1996), Micromechanics of brittle faulting and cataclastic flow in berea sandstone, *J. Struct. Geol.*, *18*, 1–16.

Montési, L., and M. Zuber (2002), A unified description of localization for application to largescale tectonics, *J. Geophys. Res.*, *107*, 10.1029/2001JB000465.

Moresi, L., and V. Solomatov (1998), Mantle convection with a brittle lithosphere: Thoughts on the global tectonic style of the earth and venus, *Geophys. J. Int.*, *133*, 669–682.

Mueller, S., and R. Phillips (1991), On the initiation of subduction, *J. Geophys. Res.*, *96*, 651–665.

O'Connell, R., C. Gable, and B. Hager (1991), Toroidal-poloidal partitioning of lithospheric plate motion, in *Glacial Isostasy, Sea Level and Mantle Rheology*, edited by R. S. et al., pp. 535–551, Kluwer Adademic, Norwell, Mass.

Ogawa, M. (2003), Plate-like regime of a numerically modeled thermal convection in a fluid with temperature-, pressure-, and stress-history-dependent viscosity, *J. Geophys. Res.*, *108*, 2067, doi:10.1029/2000JB000,069.

Olson, P., and D. Bercovici (1991), On the equipartitioning of kinetic energy in plate tectonics, *Geophys. Res. Lett.*, *18*, 1751–1754.

Regenauer-Lieb, K., D. Yuen, and J. Branlund (2001), The initiation of subduction: Criticality by addition of water?, *Science*, *294*, 578–580.

Ribe, N. (1992), The dynamics of thin shells with variable viscosity and the origin of toroidal flow in the mantle, *Geophys. J. Int., 110, 537–552*.

Ricard, Y., and D. Bercovici (2003), Two-phase damage theory and crustal rock failure: the theoretical 'void' limit, and the predition of experimental data, *Geophys. J. Int.*, *155*, 1057–1064.

Ricard, Y., and C. Froidevaux (1986), Stretching instabilities and lithospheric boudinage, *J. Geophys. Res.*, *91*, 8314–8324.

Ricard, Y., and C. Vigny (1989), Mantle dynamics with induced plate tectonics, *J. Geophys. Res.*, *94*, 17,543–17,559.

Ricard, Y., D. Bercovici, and G. Schubert (2001), A two-phase model of compaction and damage, 2, applications to compaction, deformation, and the role of interfacial surface tension, *J. Geophys, Res.*, *106*(B5), 8907–8924.

Schubert, G., and D. Turcotte (1972), One-dimensional model of shallow mantle convection, *J. Geophys. Res.*, *77*, 945–951.

Schubert, G., and D. Yuen (1978), Shear heating instability in earth's upper mantle, *Tectonophysics*, *50*, 197–205.

Schubert, G., and K. Zhang (1997), Foundering of the lithosphere at the onset of subduction, *Geophys. Res. Lett.*, *24*, 1527–1529.

Scott, D. (1996), Seismicity and stress rotation in a granular model of the brittle crust, *Nature*, *381*, 592–595.

Spiegelman, M. (1993), Physics of melt extraction: Theory, implications and applications, *Philos. Trans. R. Soc. London, Ser. A*, *342*, 23–41.

Sumita, I., S. Yoshida, M. Kumazawa, and Y. Hamano (1996), A model for sedimentary compaction of a viscous medium and its application to inner-core growth, *Geophys. J. Intl.*, *124*, 502–524.

Tackley, P. (1998), Self-consistent generation of tectonic plates in three-dimensional mantle convection, *Earth Planet. Sci. Lett.*, *157*, 9–22.

Tackley, P. (2000a), The quest for self-consistent generation of plate tectonics in mantle convection models, in *History and Dynamics of Global Plate Motions, Geophys. Monogr. Ser.*, vol. 121, edited by M. A. Richards, R. Gordon, and R. van der Hilst, pp. 47–72, Am. Geophys. Union, Washington, DC.

Tackley, P. (2000b), Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations, 2. strain weakening and asthenosphere, *Geochem. Geophys. Geosystems (*$\mathrm{G}^3$*)*, *1*, 2000GC000,043.

Tackley, P. (2000c), Mantle convection and plate tectonics: Toward and integrated physical and chemical theory, *Science*, *288*, 2002–2007.

Taylor, G., and H. Quinney (1934), The latent energy remaining in metal after cold working, *Proc. R. Soc. London, Ser. A*, *143*, 307–326.

Toth, G., and M. Gurnis (1998), Dynamics of subduction initiation at preexisting fault zones, *J. Geophys. Res.*, *103*, 18,053–18,067.

Urai, J., W. Means, and G. Lister (1986), Dynamic recrystallization in minerals, in *Mineral and Rock Deformation: Laboratory Studies*, edited by B. Hobbs and H. Heard, pp. 166–199, American Geophysical Union, Washington DC.

Watts, A., G. Karner, and M. Steckler (1982), Lithosphere flexure and the evolution of sedimentary basins, *Philos. Trans. R. Soc. London Ser. A*, *305*, 249–281.

Weinstein, S. (1998), The effect of convection planform on the toroidal-poloidal energy ratio, *Earth Planet. Sci. Lett.*, *155*, 87–95.

Weinstein, S., and P. Olson (1992), Thermal convection with non-newtonian plates, *Geophys. J. Int.*, *111*, 515–530.

White, S., S. Burrows, J. Carreras, N. Shaw, and F. Humphreys (1980), On mylonites in ductile shear zones, *J. Struct. Geol.*, *2*, 175–187.

Whitehead, J., and R. Gans (1974), A new, theoretically tractable earthquake model, *Geophys. J. R. Astron. Soc.*, *39*, 11–28.

Wong, T.-F., C. David, and W. Zhu (1997), The transition from brittle faulting to cataclastic flow in porous sandstones: Mechanical deformation, *J. Geophys. Res.*, *102*, 3009–3025.

Zhu, W., and T.-F. Wong (1997), The transition from brittle faulting to cataclastic flow: Permeability evolution, *J. Geophys. Res.*, *102*, 3027–3041.

Zhu, W., L. Montesi, and T.-F. Wong (1997), Shear enhanced compaction and permeability reduction: triaxial extension tests on porous sandstone, *Mechanics of Materials*, *25*, 199–214.

**Figure 1.** The source-sink field $S$, the dilation rate $G$, porosity $\phi$, vertical vorticity $\Omega$, and horizontal matrix velocity $\mathbf{v}_{\rm h} = u_m\hat{\mathbf{x}} + v_m\hat{\mathbf{y}}$ for the void-generating damage case ($f_\phi > 0$, $f_{\mathcal{A}} = 0$). Frame a shows the initial condition ($t \approx 0$). Lower frames show later times (dimensionless time indicated in the velocity frame) for cases with (b) relatively large $\hat{\gamma} = 1$ (see (20) and (46)) and (c) relatively small $\hat{\gamma} = 10^{-2}$. Porosity is initiated with a random perturbation of amplitude 0.001 on top of a constant background of 0.05. Other parameters are $a = b = 0.5$, $f^* = 0.5$ and $\hat{\sigma} \approx 0$. Minimum and maximum values of each scalar field, and the maximum velocity vector length are indicated in the figure.

**Figure 2.** Maximum absolute values of porosity $\phi$, vorticity $\Omega$ (solid lines of middle frames) and vorticity correlation function $C_\Omega$ (dashed lines in middle frames), and maximum absolute dilation rate $G$ versus time for the case shown in Fig. 1b and c.

**Figure 3.** The source-sink field $S$, the dilation rate $G$, porosity $\phi$, vertical vorticity $\Omega$, the horizontal matrix velocity $\mathbf{v}_\mathrm{h}$, and the inverse grain/void size $\mathcal{A}$ for the *fineness*-generating damage case ($f_\phi = 0$, $f_\mathcal{A} > 0$) with $m = 1, 7$ and 21 (see (22)), shown in frames a,b and c respectively. The dimensionless time is indicated in the velocity frame. Porosity $\phi$ is initiated with a random perturbation of amplitude 0.001 on top of a constant background of 0.05. Dimensionless inverse grain/void size $\mathcal{A}$ is initiated with a random perturbation of amplitude 0.01 on top of a constant background of 1. Other parameters are $a = b = 0.5$, and while $\hat{\sigma} =\approx 0$ in the momentum equations we prescribe $f_\mathcal{A}/\hat{\sigma} = 1$ in (48).

**Figure 4.** Maximum absolute values of porosity $\phi$, *fineness* (or inverse grain/void size) $\mathcal{A}$, vorticity $\Omega$ (solid lines), vorticity correlation function $C_\Omega$ (dashed lines) and maximum dilation rate $G$ versus time for the case shown in in Fig. 3.

**Figure 5.** Same as Fig. 3 but with $m = 3$ and and three values of $f_\mathcal{A}/\hat{\sigma} = 1, 10, 100$ (frames a, b, and c, respectively).

**Figure 6.** Same as Fig. 4 but for cases shown in Fig. 5.

**Figure 7.** Time interval for suppression of dilation $G$, as given by (56), versus $m$ for $f_\mathcal{A}/\hat{\sigma} = 1$ (a), and versus $f_\mathcal{A}/\hat{\sigma}$ for $m = 3$ (b).

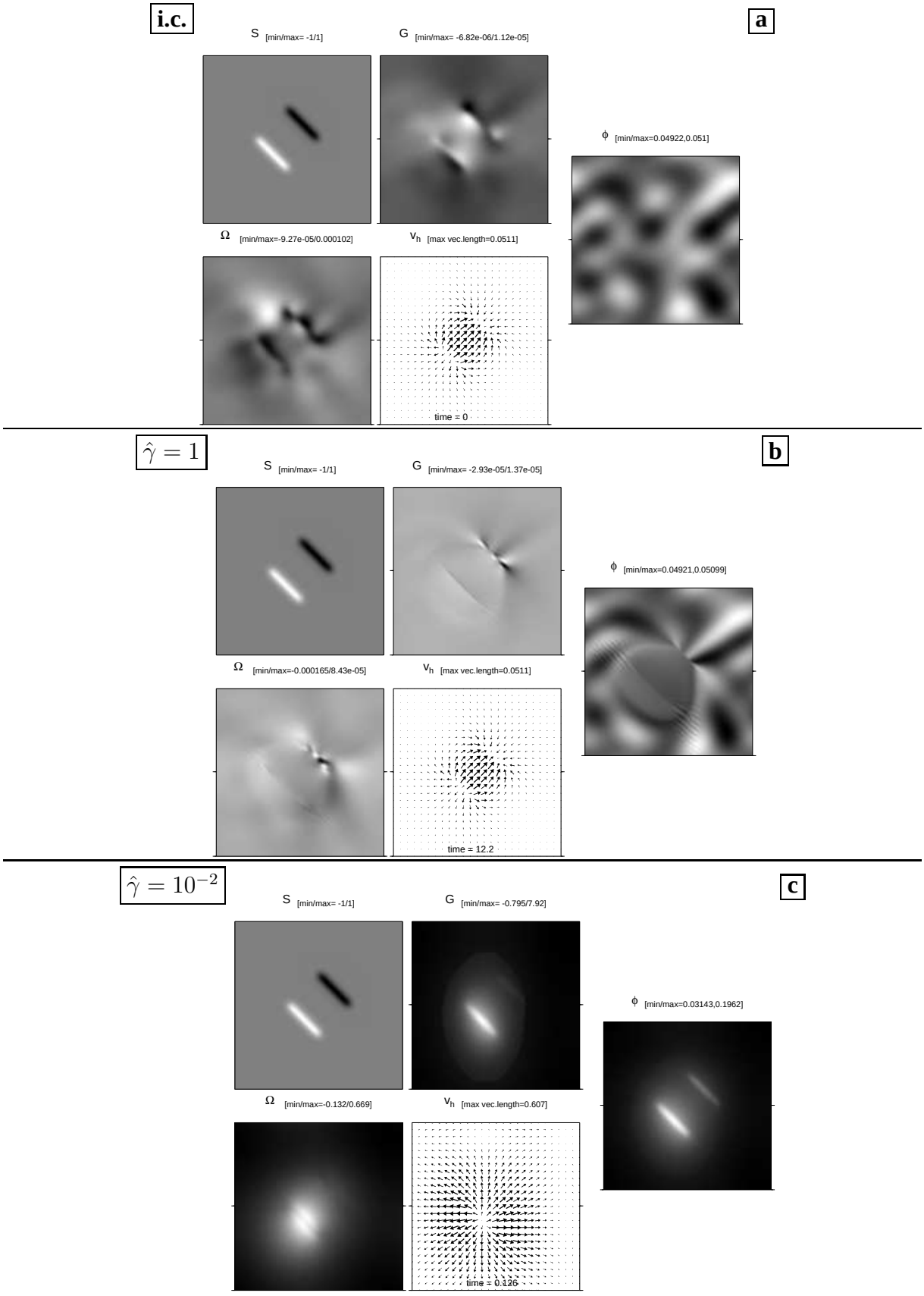


**Figure 1.** The source-sink field $S$, the dilation rate $G$, porosity $\phi$, vertical vorticity $\Omega$, and horizontal matrix velocity $\mathbf{v}_{\rm h} = u_m\hat{\mathbf{x}} + v_m\hat{\mathbf{y}}$ for the void-generating damage case ($f_\phi > 0$, $f_{\mathcal{A}} = 0$). Frame a shows the initial condition ($t \approx 0$). Lower frames show later times (dimensionless time indicated in the velocity frame) for cases with (b) relatively large $\hat{\gamma} = 1$ (see (20) and (46)) and (c) relatively small $\hat{\gamma} = 10^{-2}$. Porosity is initiated with a random perturbation of amplitude 0.001 on top of a constant background of 0.05. Other parameters are $a = b = 0.5$, $f^* = 0.5$ and $\hat{\sigma} \approx 0$. Minimum and maximum values of each scalar field, and the maximum velocity vector length are indicated in the figure.

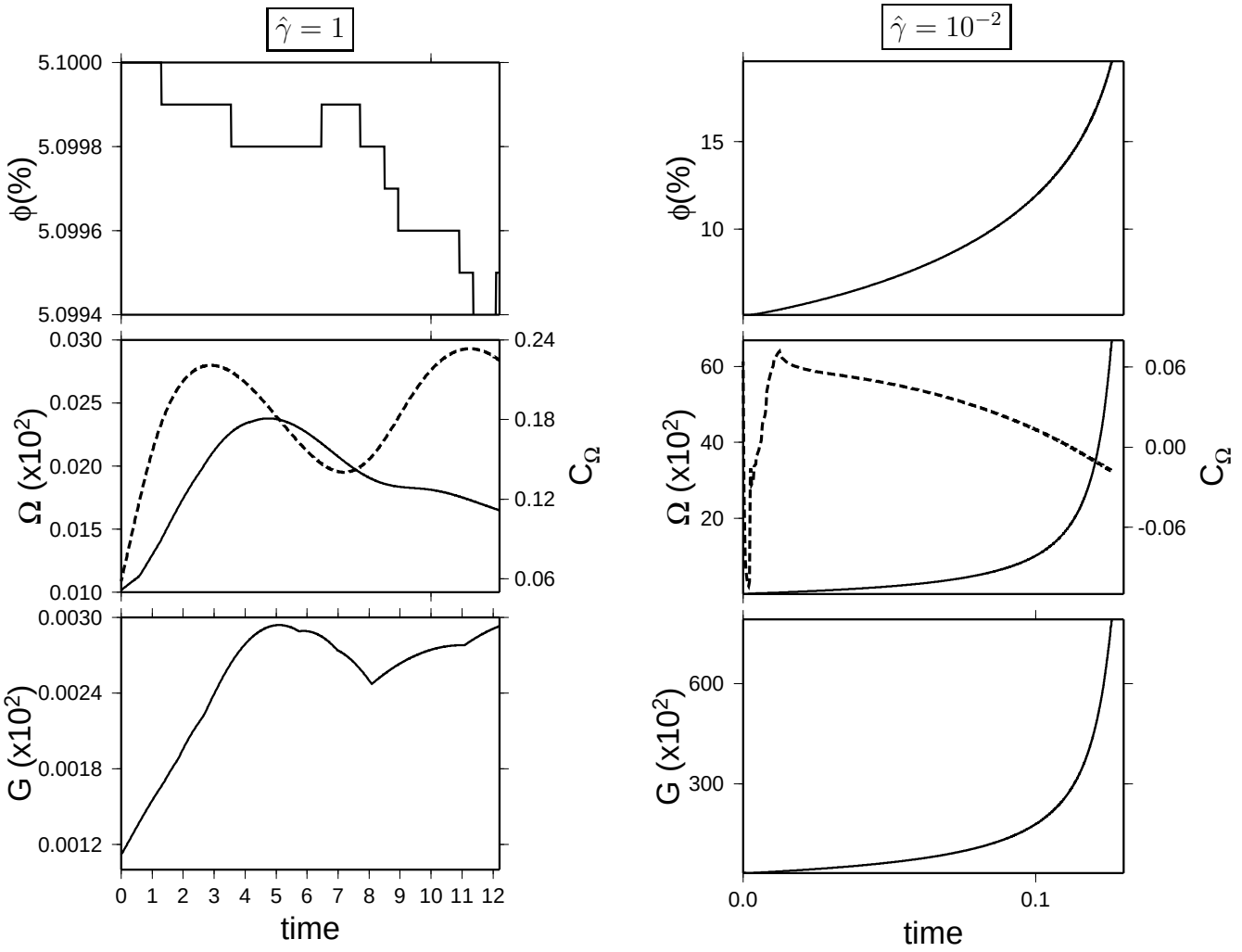


**Figure 2.** Maximum absolute values of porosity $\phi$, vorticity $\Omega$ (solid lines of middle frames) and vorticity correlation function $C_\Omega$ (dashed lines in middle frames), and maximum absolute dilation rate $G$ versus time for the case shown in Fig. 1b and c.

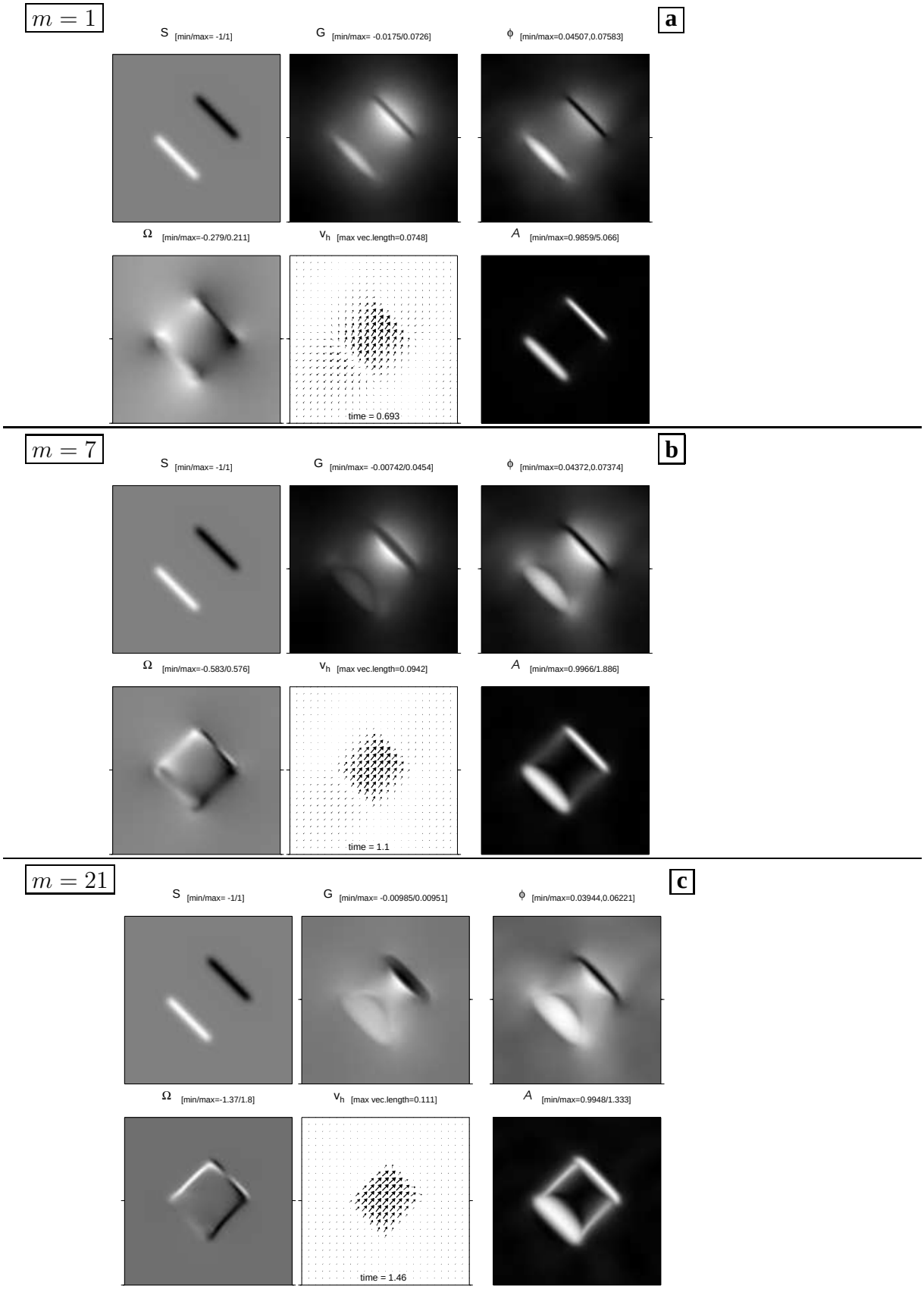


**Figure 3.** The source-sink field $S$, the dilation rate $G$, porosity $\phi$, vertical vorticity $\Omega$, the horizontal matrix velocity $\mathbf{v}_{\rm h}$, and the inverse grain/void size $\mathcal{A}$ for the *fi neness*-generating damage case ($f_\phi = 0$, $f_{\mathcal{A}} > 0$) with $m = 1$, 7 and 21 (see (22)), shown in frames a,b and c respectively. The dimensionless time is indicated in the velocity frame. Porosity $\phi$ is initiated with a random perturbation of amplitude 0.001 on top of a constant background of 0.05. Dimensionless inverse grain/void size $\mathcal{A}$ is initiated with a random perturbation of amplitude 0.01 on top of a constant background of 1. Other parameters are $a = b = 0.5$, and while $\hat{\sigma} = \approx 0$ in the momentum equations we prescribe $f_{\mathcal{A}}/\hat{\sigma} = 1$ in (48).

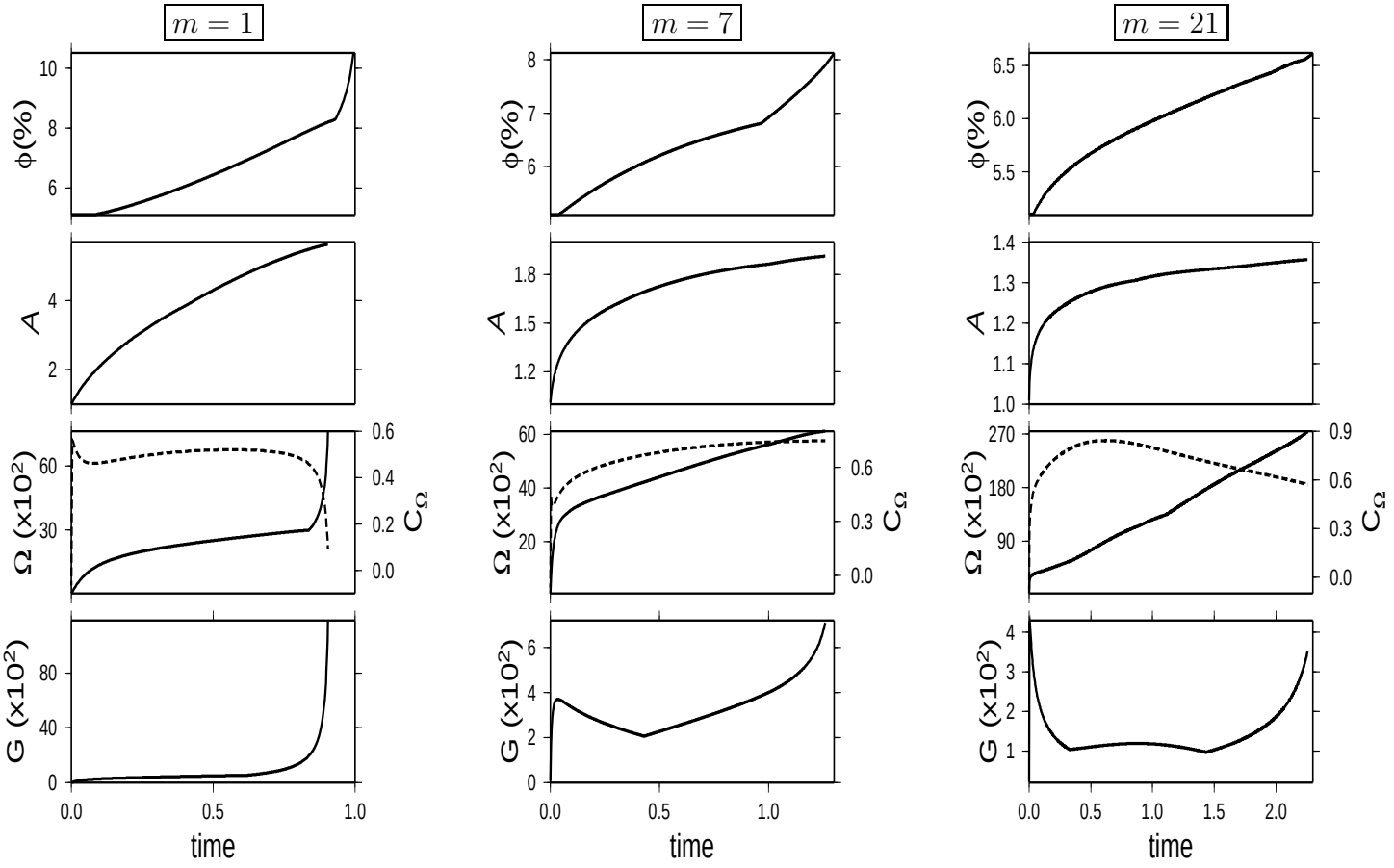


**Figure 4.** Maximum absolute values of porosity $\phi$, *fineness* (or inverse grain/void size) $\mathcal{A}$, vorticity $\Omega$ (solid lines), vorticity correlation function $C_\Omega$ (dashed lines) and maximum dilation rate $G$ versus time for the case shown in in Fig. 3.

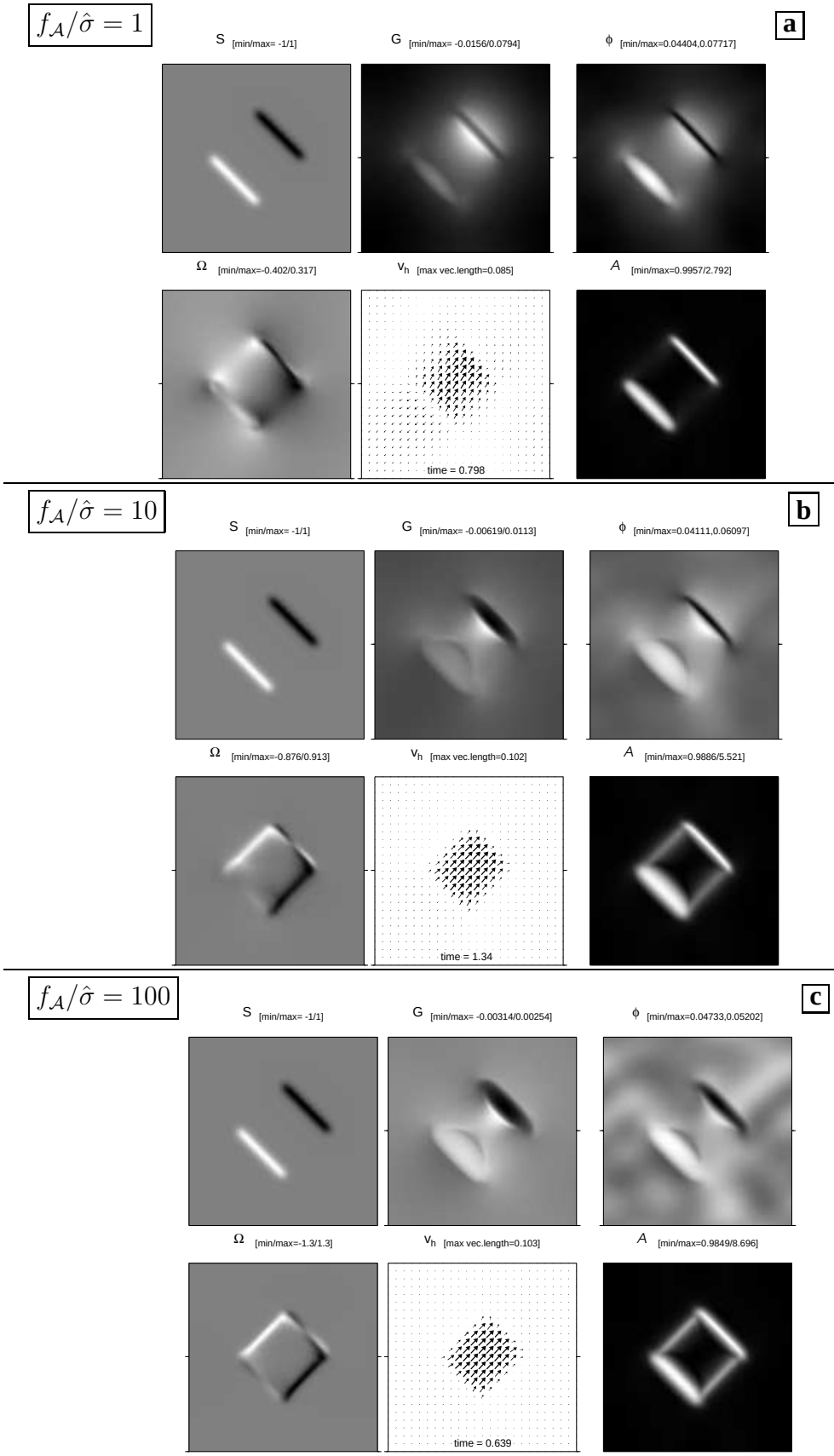


**Figure 5.** Same as Fig. 3 but with $m = 3$ and and three values of $f_{\mathcal{A}}/\hat{\sigma} = 1, 10, 100$ (frames a, b, and c, respectively).

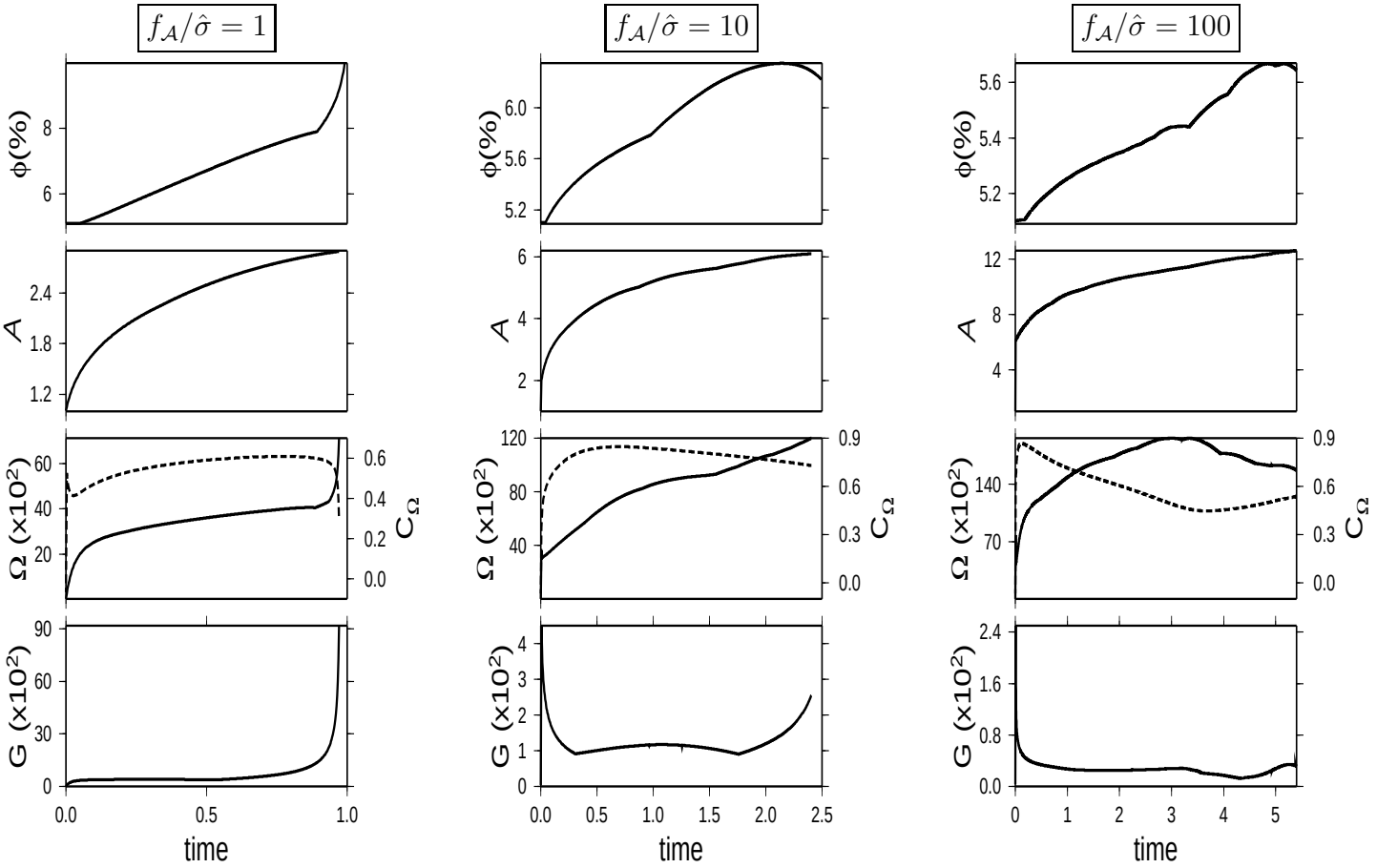


**Figure 6.** Same as Fig. 4 but for cases shown in Fig. 5.

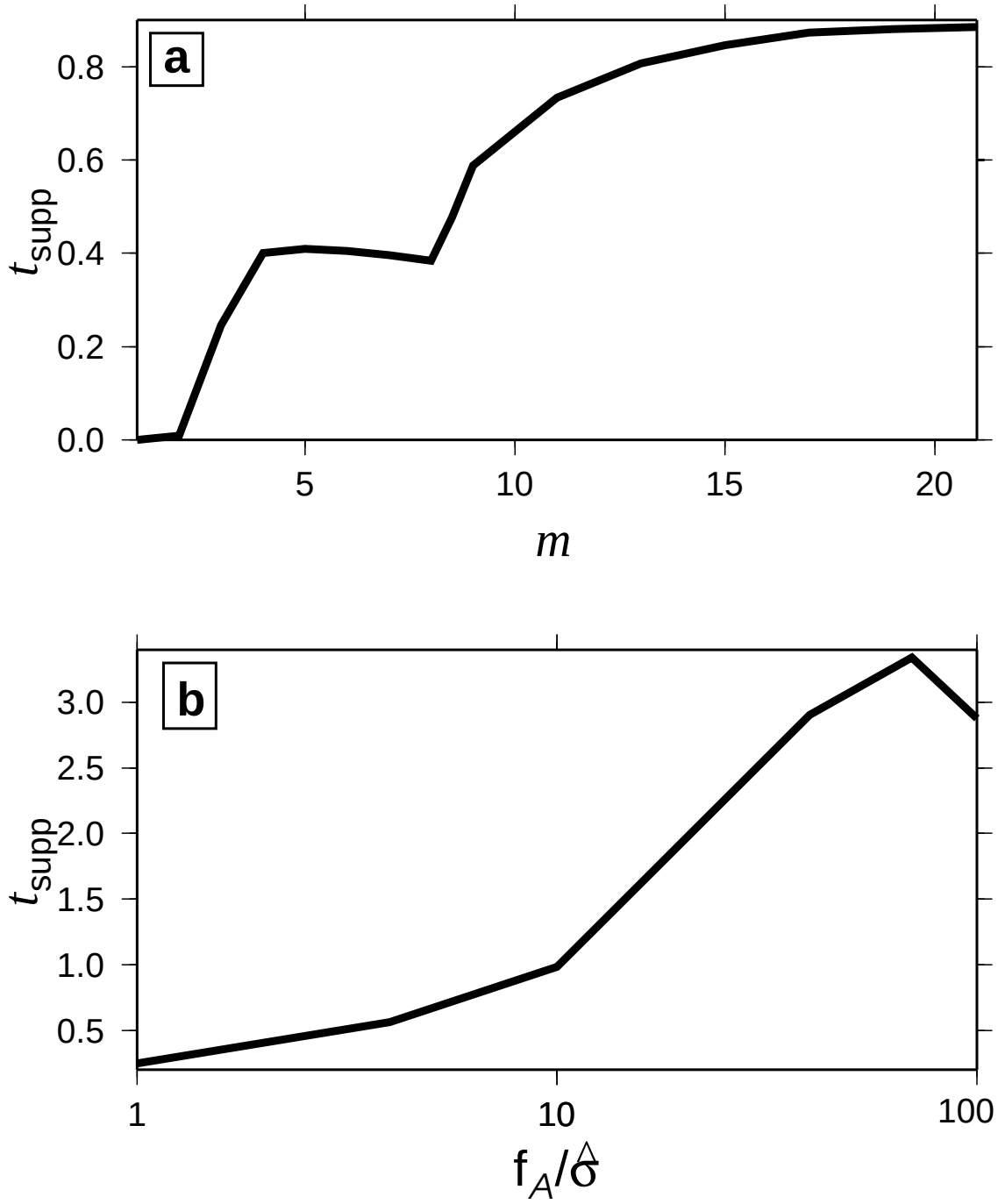


**Figure 7.** Time interval for suppression of dilation $G$, as given by (56), versus $m$ for $f_{\mathcal{A}}/\hat{\sigma} = 1$ (a), and versus $f_{\mathcal{A}}/\hat{\sigma}$ for $m = 3$ (b).