Earthquake cycle simulations with rate-and-state friction and power-law viscoelasticity

Kali L. Allison\textsuperscript{a,}\textsuperscript{*}, Eric M. Dunham\textsuperscript{a,b}

\textsuperscript{a} Department of Geophysics, Stanford University, Stanford, CA, USA
\textsuperscript{b} Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA

**A R T I C L E   I N F O**

Keywords:
Earthquake cycle  
Viscoelastic flow  
Power-law rheology  
Strike-slip  
Brittle-ductile transition  
Rate-and-state friction  
Summation-by-parts operators

**A B S T R A C T**

We simulate earthquake cycles with rate-and-state fault friction and off-fault power-law viscoelasticity for the classic 2D antiplane shear problem of a vertical, strike-slip plate boundary fault. We investigate the interaction between fault slip and bulk viscous flow with experimentally-based flow laws for quartz-diorite and olivine for the crust and mantle, respectively. Simulations using three linear geotherms ($d\theta/dz = 20$, 25, and 30 K/km) produce different deformation styles at depth, ranging from significant interseismic fault creep to purely bulk viscous flow. However, they have almost identical earthquake recurrence interval, nucleation depth, and downdip coseismic slip limit. Despite these similarities, variations in the predicted surface deformation might permit discrimination of the deformation mechanism using geodetic observations. Additionally, in the 25 and 30 K/km simulations, the crust drags the mantle; the 20 K/km simulation also predicts this, except within 10 km of the fault where the reverse occurs. However, basal tractions play a minor role in the overall force balance of the lithosphere, at least for the flow laws used in our study. Therefore, the depth-integrated stress on the fault is balanced primarily by shear stress on vertical, fault-parallel planes. Because strain rates are higher directly below the fault than far from it, stresses are also higher. Thus, the upper crust far from the fault bears a substantial part of the tectonic load, resulting in unrealistically high stresses. In the real Earth, this might lead to distributed plastic deformation or formation of subparallel faults. Alternatively, fault pore pressures in excess of hydrostatic and/or weakening mechanisms such as grain size reduction and thermo-mechanical coupling could lower the strength of the ductile fault root in the lower crust and, concomitantly, off-fault upper crustal stresses.

1. Introduction

Understanding the structure and dynamics of the continental lithosphere and the active faults it contains is of fundamental importance. While it is well established that faults are highly localized within the cold, brittle upper crust, less is known about fault structure in the warmer, and more ductile, lower crust and upper mantle. Classic studies of lithospheric stress profiles were based upon laboratory rock deformation experiments, spanning both brittle and ductile behavior, and theoretically based thermally activated creep laws (e.g., Byerlee, 1978; Goetze and Evans, 1979; Brace and Kohlstedt, 1980; Sibson, 1982; Sibson, 1984). The onset of creep at elevated temperatures reduces stress within the lithosphere below the frictional strength, thereby preventing seismic slip. The predicted seismogenic depth of ~10 to 20 km broadly matches observed depths of seismicity. In these classic studies, faults below the seismogenic depth are assumed to broaden rapidly into wide mylonite zones (e.g., Scholz, 2002).

These concepts were refined over the following decades to provide a more nuanced understanding of deep fault structure and controls on seismogenic depth. Friction experiments at elevated temperatures revealed a transition from velocity-weakening to velocity-strengthening steady state friction with increasing temperature (e.g., Dieterich, 1978; Ruina, 1983; Tullis and Weeks, 1986; Blanpied et al., 1991; Blanpied et al., 1995). Drawing upon available experimental friction data and estimates of continental geotherms, Tse and Rice (1986) demonstrated that the primary features of crustal faulting (such as seismogenic depth and recurrence intervals) could be reproduced in models with ideally elastic off-fault response. In their model, the fault continues as a localized surface below the seismogenic depth, with tectonic displacement accommodated by aseismic sliding at depth. This established the importance of the frictional seismic-aseismic transition. The majority of earthquake modeling studies that followed (e.g., Rice, 1993; Lapusta et al., 2000; Kaneko et al., 2011) have followed Tse and Rice (1986) by neglecting the viscoelastic response of the off-fault material. However, there have been a few studies that examine the role of viscoelasticity on earthquakes, and these are discussed at the end of this introduction.

\textsuperscript{*} Corresponding author.

E-mail address: kallison@stanford.edu (K.L. Allison).

https://doi.org/10.1016/j.tecto.2017.10.021

Received 20 May 2017; Received in revised form 14 October 2017; Accepted 20 October 2017

0040-1951/ © 2017 Elsevier B.V. All rights reserved.
Studies of exhumed fault zones place constraints on the structure of faults, degree of localization, and deformation mechanisms at depth. Exhumed rocks preserve evidence of distributed ductile deformation at depth in the form of broad mylonite zones. Some zones are relatively narrow, such as the 1–2 km wide mylonite zone from the middle and lower crust beneath the Alpine Fault (Norris and Cooper, 2003). Others, such as the Great Slave Lake shear zone in Canada (Hamner, 1988), are tens of kilometers in width. However, it is possible that this large width characterizes a broad zone of anastomosing faults, rather than a single fault (Norris and Cooper, 2003). The dynamics of shear zones at the base of the seismogenic layer is also complicated. Exhumed rocks show mutual overprinting of mylonites and pseudotachylites (e.g., Cole et al., 2007; Frost et al., 2011; Kirkpatrick and Rowe, 2013), which argues that the same material can respond in different ways to loading across a broad range of strain rates. This evidence also shows that the brittle-ductile transition takes place gradually over a range of depths.

Seismic imaging studies also place constraints on the structure of faults and degree of localization at depth. Seismic reflection and refraction studies show that many continental transform faults penetrate the entire crust, persisting as localized features that cut through the Moho (Lemiszki and Brown, 1988; Vauchez and Tommasi, 2003), including the San Andreas (Henstock et al., 1997; Lemiszki and Brown, 1988; Zhu, 2000), the San Jacinto (Miller et al., 2014), and the Dead Sea transform faults (Weber et al., 2004). In contrast, the Alpine Fault and the Marlborough Fault system, both in New Zealand’s South Island, do not appear to be highly localized at the Moho. Instead they form broad zones of deformation in the lower crust and upper mantle (Klosko et al., 1999; Molnar, 1999; Wilson et al., 2004). The width of the deformation zones in the upper mantle beneath these faults, as inferred from seismic anisotropy studies of shear-wave splitting, is quite broad: 200 km for the Alpine Fault (Duclos et al., 2005; Baldock and Stern, 2005), 130 km for the San Andreas Fault (Bonnin et al., 2010), and 20 km for the Dead Sea Fault (Rümpker et al., 2003). The Dead Sea Fault has accommodated comparatively little displacement, 100 km in contrast with 850 km for the San Andreas, and this may account for the relatively narrow zone exhibiting anisotropy (Baldock and Stern, 2005).

Anisotropy beneath the North Anatolian Fault does not align with the strike of the fault, perhaps indicating that the width of the shear zone is < 50 km, smaller than can be imaged by shear-wave splitting (Vaucluse and Tommasi, 2003), or that the fault is too young to have produced measurable anisotropy (Berk Biryol et al., 2010). This shear-wave splitting occurs at greater depths than we consider in our model, so we don’t seek to produce or explain the extent of these deformation zones, but rather include this discussion to describe the extent to which deformation broadens with depth.

It is evident that considerable questions remain unresolved with regard to the frictional velocity-weakening to velocity-strengthening transition on faults and the transition marking the onset of bulk viscous flow in rocks surrounding the fault. In this study, we investigate the interaction between these two transitions in the continental lithosphere, with particular focus on the interplay between aseismic fault creep and viscous flow across the brittle-ductile transition zone. We simulate friction on a strike-slip fault, capturing the transition from shallow coseismic fault slip to deeper interseismic fault creep. We couple this with off-fault viscoelastic deformation, representative of the deformation of the lower crust and upper mantle. A diagram of our model is shown in Fig. 1.

The transitions described previously have implications for stress profiles within the continental lithosphere. The lithosphere is typically divided into three rheologically distinct layers: the upper crust, lower crust, and upper mantle. The upper crust is brittle, so its strength is determined by the strength of the fault, governed by Byerlee’s law or rate-and-state friction. The relative strength of the lower crust and upper mantle depends upon the composition of those layers, as well as water content, and remains an open question. In the “jelly sandwich” model, the weak, wet lower crust is sandwiched by the cool, brittle upper crust and the dry upper mantle, both of which are stronger (e.g., Chen and Molnar, 1983; Burov and Diament, 1995). In the “crème brûlée” model, the upper and lower crust overlie a much weaker upper mantle (e.g., Jackson, 2002; Burov and Watts, 2006). The “crème brûlée” model is supported, in regions such as the Mojave Desert, by estimates of rheological structure from transient postseismic deformation and exhumed xenoliths (summarized in Fig. 3 and discussed in more detail below) (e.g., Johnson et al., 2007; Thatcher and Pollitz, 2008). Less clear support includes evidence that earthquakes are confined to the upper and lower crust (Maggi et al., 2000), suggesting that the upper mantle is too weak to support earthquakes, though there is evidence that in some regions earthquakes are able to occur in the upper mantle (e.g., Inbal et al., 2016). Additionally, evidence based on measurements of the effective elastic thickness T_e, which is a proxy for lithospheric strength (Watts et al., 2013), is conflicting. In some regions, T_e is smaller than the seismogenic thickness (Maggi et al., 2000; Jackson, 2002), which suggests that lithospheric strength resides in the crust, supporting the “crème brûlée” model. Interpretations of both the confinement of earthquakes to the crust and T_e relative to the seismogenic thickness are debated, however. The apparent absence of earthquakes in the upper mantle may result from a strong but fractionally stable upper mantle (Jackson, 2002). Additionally, others argue that T_e is larger than the seismogenic thickness in some regions (Burov and Watts, 2006), which suggests that the upper mantle may be strong, and thus that the “jelly sandwich” model is correct. Burov and Watts (2006) also find that only the “jelly sandwich” model can support mountain ranges over millions of years. It is important to bear in mind that each of these apparently conflicting pieces of evidence measures the strength of the lithosphere over a different timescale. Postseismic deformation, which generally favors the “crème brûlée” model, measures the strength of the lithosphere on the timescale of a decade, while T_e measures the strength on the scale of a million years. We focus on the time scale of the postseismic and interseismic periods of the earthquake cycle, which range from a decade to a few hundred years, and therefore consider the “crème brûlée” model. Furthermore, there are likely regional variations in the strength of the lower crust and upper mantle.

Transient postseismic deformation in the first few months to years after an earthquake provides short-term constraints on the rheological structure of the crust. The three commonly considered mechanisms are frictional afterslip (Perfettini and Avouac, 2004), viscoelastic flow (Pollitz et al., 2000; Freed and Bürgmann, 2004; Freed et al., 2007), and poroelastic rebound (Booker, 1974; Jönsson et al., 2003; Fialko, 2004). Complicating matters, it is possible for multiple mechanisms to be active after the same event (e.g., Rousset et al., 2012), and, for strike-slip faults, frictional afterslip and viscoelastic flow can produce similar deformation signals on Earth’s surface (Thatcher, 1983; Hearn, 2003). We focus on frictional afterslip and viscoelastic flow, as they also have significant implications for the strength of the lithosphere. In particular, viscous flow provides constraints on the lateral and vertical rheological structure (e.g., Thatcher and Pollitz, 2008; Hearn et al., 2009; Vaghri and Hearn, 2012; Yamasaki et al., 2014; Rollins et al., 2015; Masuti et al., 2016). Frictional afterslip reveals information about the frictional properties of the fault, including estimates of the value (Perfettini and Avouac, 2004; Barbot et al., 2009) and spatial variation of frictional properties (Miyazaki et al., 2004). In combination with the postseismic period, data from the interseismic period can be used to infer rheological parameters (e.g., Hetland and Hager, 2005) or frictional parameters (e.g., Lindsey and Fialko, 2016). Some studies find evidence for coupled frictional afterslip and viscous flow in the postseismic period (e.g., Biggs et al., 2009; Johnson et al., 2009; Masuti et al., 2016).

We simulate afterslip and viscoelastic deformation in the context of earthquake cycle simulations. Our philosophy is to start with the simplest case that combines spontaneously nucleating ruptures with off-fault viscous flow. We begin with an assumed compositional structure and geotherm, assign model parameters for this structure from laboratory experiments, and our simulations determine the recurrence
interval, coseismic slip distribution, and partitioning of postseismic and interseismic deformation into afterslip and bulk viscous flow. This allows the slip, stress drop, and recurrence interval of each earthquake to develop in a way that is self-consistent with the history of earthquakes and postseismic deformation. To date, earthquake cycle simulation efforts have primarily focused on the effects of rate-and-state friction while representing the off-fault material as elastic (e.g., Tsai and Rice, 1986; Rice, 1993; Lapusta et al., 2000; Kaneko et al., 2011), or on viscoelastic deformation while kinematically imposing the motion of the fault (e.g., Savage and Prescott, 1978; Thatcher, 1983; Johnson et al., 2007; Takeuchi and Fialko, 2012). A few codes are capable of fully coupling both rate-and-state friction and linear or power-law viscoelasticity. However, some are limited to a fault-containing elastic layer over a viscoelastic half-space (Kato, 2002; Lambert and Barbot, 2016), a set-up which specifies a priori the brittle-ductile transition and cannot permit material in this region to experience both fault slip and viscous flow at different times in the earthquake cycle. Others do not simulate multiple earthquake cycles (Barbot and Fialko, 2010; Aagaard et al., 2013). These other codes are used mainly to address different questions than the ones explored in our simulations.

Lambert and Barbot (2016) focus on viscous flow transients in the upper mantle and their signature in surface displacements. They find that bulk viscous flow can contribute significantly to postseismic surface deformation quite early in the postseismic period. Also, while nearfield postseismic surface deformation is dominated by frictional after-slip, the contribution from viscous flow can be comparable in the far field within one year of rupture. Using a similar model geometry, Kato (2002) focuses on the recurrence interval and average slip of each earthquake, finding that neither are appreciably affected by flow in the viscoelastic half-space. Shimamoto and Noda (2014) use a unified constitutive law to transition from shallow frictional behavior to deeper viscous flow (in a localized fault root). Like Kato (2002), they also find that the behavior of the seismogenic crust is quite similar to that of a purely frictional model. Additionally, they find that coseismic slip is able to penetrate into the brittle-ductile transition region, resulting in a region about 5 km wide in which both coseismic fault slip and appreciable viscous flow occur. Because the brittle-ductile transition depends on strain rate, the depth extent of coseismic faulting exceeds the brittle-ductile transition depth as estimated from the plate rate. And finally, a similar modeling study (Beeler et al., submitted) finds that ruptures penetrate 1–2 km below the nominal brittle-ductile transition, but that less than 10% of the total moment release occurs in this process.

2. Lithospheric structure

Parameter choices in our study are loosely motivated by the Mojave Desert region of Southern California, a well-studied area with a wealth of available data. In this region, the top of the upper mantle, defined as the depth of the Moho discontinuity, is approximately 30 km deep, as shown in Fig. 2.

Insight into the structure of the lithosphere in this region comes in part from exhumed xenoliths, which can provide constraints on the viscosity, geothermal gradient, and degree of strain localization in the lower crust and upper mantle (e.g., Ave Lallemant et al., 1980; Behr and Hirth, 2014; Chatzaras et al., 2015), shown in Fig. 3. Additionally, estimates of the effective viscosity structure of the crust from GPS and InSAR measurements of postseismic and interseismic deformation from the western US are summarized in Johnson et al. (2007) and Thatcher and Pollitz (2008), reproduced in Fig. 3. Also included in the figure are results from isostatic adjustment from lake loads over thousands of years. An additional compilation of the results of postseismic inversions globally can be found in Wright et al. (2013), who find that in general the Maxwell viscosity in the lower crust and upper mantle ranges from $10^{17} - 7 \times 10^{19}$ Pa s. The wide range in viscosity estimates results in part from different assumptions made in each model, with some assuming a single viscosity in a linear Maxwell rheology (e.g., Segall, 2002; Bruhat et al., 2011), others a linear Burgers rheology (Pollitz, 2003, 2005), and some a nonlinear power-law rheology (Freed and Bürgmann, 2004; Freed et al., 2006). Models which assume a linear Maxwell rheology cannot account for temporal variations in effective viscosity, such as the observed increase from $10^{18}$ Pa s to $10^{19}$ Pa s over the course of 5 years after the Hector Mine earthquake (Freed and Bürgmann, 2004). This is accounted for in line Burgers rheology by two different viscosities, and in a power-law rheology through the stress-dependence of the effective viscosity. Additionally, while the models summarized in Fig. 3 assume no spatial variation, some studies have found evidence for a localized zone of lower viscosity near the fault (e.g., Yamasaki et al., 2014).

3. Model

We consider the two-dimensional antiplane shear problem of a vertical strike-slip fault. For computational efficiency, we use the quasi-dynamic approximation to elastodynamics (quasi-static elasticity with the radiation damping approximation, Rice, 1993). We also assume material properties and tectonic loading are symmetric about the fault, enabling us to model only one half of the domain ($y \geq 0$). As shown in Fig. 1, the fault is located at $y = 0$ and is parallel to the z-axis, and displacements are in the x-direction. Note that the fault cuts through the entire model domain. However, partitioning of the remote tectonic loading into fault slip or bulk viscous flow will be determined by the relative strength of the fault and off-fault material, so the fault will not necessarily slip at depth.

The governing equations are

\begin{equation}
\frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} = 0,
\end{equation}

\begin{equation}
\sigma_{yy} = \mu \left( \frac{\partial u}{\partial y} - \gamma_{yy} \right), \quad \sigma_{xz} = \mu \left( \frac{\partial u}{\partial z} - \gamma_{xz} \right),
\end{equation}

\begin{equation}
\gamma_{yy} = \eta^{-1} \sigma_{yy}, \quad \gamma_{xz} = \eta^{-1} \sigma_{xz},
\end{equation}

\begin{equation}
\eta^{-1} = Ae^{-\omega/\varepsilon^{\varepsilon}}, \quad \tau = \sqrt{\sigma_{yy}^2 + \sigma_{xz}^2},
\end{equation}
with fault boundary conditions
\[
\tau(z, t) = \sigma_{fl}(0, z, t) - \eta_{ff} V/2 = f(\psi, V)\sigma_{fl}, \tag{5}
\]
\[
\dot{\psi} = G(\psi, V), \tag{6}
\]
\[
\delta(z, t) = 2u(0, z, t), \tag{7}
\]
\[
V(z, t) = \frac{\delta z}{\delta t}, \tag{8}
\]
where Eq. (1) is the static equilibrium equation; Eq. (2) is Hooke’s law; Eqs. (3) and (4) express the power-law viscous flow law; Eq. (5) is the force balance on the fault, where stress on the fault, with the radiation-damping approximation, is equal to frictional strength; Eq. (6) is the state evolution; Eq. (7) is the fault slip; and Eq. (8) is the slip velocity. In these equations, \( u \) is the displacement in the \( x \) direction, \( \gamma_{ij} \) are the (engineering) viscous strains, \( \sigma_{fl} \) are the stress components, \( \mu \) is the shear modulus, \( \eta \) is the effective viscosity, and \( T \) is the temperature. On the fault, \( \tau \) is the shear stress; \( \sigma_{fl} \) is effective normal stress, \( \psi \) is the state variable, \( V \) is slip velocity, \( f \) is the friction coefficient, and \( G \) is the state evolution equation (e.g., slip law or aging law). In the flow law, the effective viscosity is a function of the rate coefficient \( A \), the activation energy \( Q \), the stress exponent \( n \), and the deviatoric stress \( \tau \). The overdot indicates a time derivative.

As depicted in Fig. 1, the computational domain is finite and there are three other boundary conditions in addition to the fault:
\[
u(L_x, z, t) = \frac{V_L t}{2}, \tag{9}
\]
\[
\sigma_{fl}(y, 0, t) = 0, \tag{10}
\]
\[
\sigma_{fl}(y, L_z, t) = 0, \tag{11}
\]
where \( L_x \) and \( L_z \) are the dimensions of the model domain in the \( y \)- and \( z \)-directions, respectively, \( V_L \) is the tectonic plate velocity. At Earth’s surface, we use a traction-free boundary condition (Eq. (10)). We also use a traction-free boundary condition at the bottom of the domain (Eq. (11)), which permits an arbitrary amount of tectonic displacement to occur, even in an elastic model. It is important to use a large enough domain that the simulation results are relatively insensitive to the location of the remote boundaries. Results are more sensitive to the side boundary location, which we place at 120 km from the fault, or 12 times the seismogenic depth. Results are much less sensitive to the depth of the bottom boundary, which in our simulations is placed at 40 km depth.

### 3.1. Frictional parameters

The fault is governed by the force balance between the frictional strength of the fault, determined by rate-and-state friction, and the shear stress exerted by the viscoelastic solid on the fault within the context of the radiation-damping approximation. We use the regularized form for frictional strength (Rice et al., 2001),
\[
f(\psi, V) = a \sinh^{-1} \left( \frac{V}{2V_0} \right), \tag{12}
\]
with the aging law for state evolution (Ruina, 1983; Marone, 1998),
The key parameters which determine the frictional behavior of the system are the direct effect parameter \( a \), the state evolution parameter \( b \), and the state evolution distance \( d_s \). Laboratory data indicate that \( d_s \) is on the order of microns (e.g., Dieterich, 1979; Marone and Kilgore, 1993); however, we use 16 mm to reduce the computational load. If a surface has \( a - b > 0 \), termed velocity-strengthening (VS), then steady-state frictional resistance to sliding increases as slip velocity increases and stable sliding of the surface is possible. On the other hand, if \( a - b < 0 \), termed velocity-weakening (WW), then steady-state frictional resistance to sliding decreases as slip velocity increases, resulting in the potential for unstable sliding. In an elastic model, earthquakes will nucleate at the transition between velocity-weakening and velocity-strengthening behavior, called the seismogenic depth. Hereafter, we use that term more generally to refer to the depth of coseismic slip.

To assign distributions for \( a \) and \( b \) to the fault, we use laboratory data for wet granite gourge from Blanpied et al. (1991, 1995), reproduced in Fig. 4. The wet granite data have been used extensively in earthquake cycle simulations before, such as Rice (1993), Lapusta et al. (2000), Lapusta and Rice (2003), Kato (2002), and Lindsey and Fialko (2016). The data predict a shallow transition from velocity-strengthening to velocity-weakening at about 400 K and a deeper transition back to velocity-strengthening at around 600 K; however, we neglect the shallow transition in order to focus on the behavior of the system at depth. Because the fault cuts through the entire domain of the model, the data have been extrapolated below the last data point. This extrapolation is linear based the linear dependence of \( a \) on temperature, which results from the Arrhenius activated process that describes creep at asperity contacts (Rice et al., 2001). Note that the Blanpied et al. (1995) data would predict a very weak fault when extrapolated to the base of the lower crust, so we follow the Blanpied et al. (1991) data for \( a - b \) instead. There is recent experimental evidence that Westerly granite can remain velocity-weakening up to 873 K (Mitchell et al., 2016) under certain conditions, but we defer exploration of that possibility for future studies.

Our choices for frictional parameters put the VW-VS transition at 10 km depth. In our simulations, despite the inclusion of viscoelastic effects, the frictional parameters on the fault control the depth of earthquake nucleation and arrest. Thus, the earthquakes nucleate at 10 km and arrest 4 km below this.

### 3.2. Remote stress state

We set the stress state of the system, on which the frictional strength (Eq. (5)) depends, assuming that the ratio of fault shear stress to effective normal stress is approximately equal to the reference coefficient of friction \( f_0 \). For an optimally oriented strike-slip fault, the principal stresses are related to each other by \( \sigma_1 > \sigma_2 > \sigma_3 \) and \( \sigma_2/\sigma_3 = \sqrt{2} + 1 + f_0 \), where \( f_0 = 0.6 \) (e.g., Sibson, 1974). We assume vertical total stress is equal to lithostatic pressure and that \( \sigma_2 = (\sigma_1 + \sigma_3)/2 \). This produces the effective normal stress and shear stress on the fault

\[
\sigma_\text{eff} = \frac{\sigma_1 + \sigma_3}{2} - \frac{\sigma_1 - \sigma_3}{2} \cos 2\phi - P_p,
\]

(14)

\[
\sigma_\text{eff} = \frac{\sigma_1 - \sigma_3}{2} \sin 2\phi.
\]

(15)

\[
\phi = \tan f_0^{-1}, \quad P_p = (1000 \text{kg m}^{-3}) \times \text{gec}.
\]

(16)

In Eq. (14), we have assumed that the effective normal stress is equal to the total normal stress minus the pore pressure, assumed in Eq. (16) to be hydrostatic. Near the brittle-ductile transition, both the assumption that pore pressure is hydrostatic and the dependence of effective normal stress on pore pressure may not hold (Beeler et al., 2016), but we defer investigation of these ideas to future studies.

### 3.3. Rheological parameters

Next, we consider the rheology of the off-fault material. We loosely base the depth and composition of each layer for our model on the Mojave Desert region. As we focus on the brittle-ductile transition and viscoelastic deformation at depth, we neglect the spatial variation in the shear modulus and depth of the upper mantle, instead using constant values given in Table 1.

The composition of the continental lithosphere varies vertically and laterally, and involves a mixture of different minerals, but the bulk of the available laboratory data focuses on specific mineral phases. In particular, rheological parameters for quartz and olivine are relatively well understood (e.g., Hirth et al., 2001; Hirth and Kohlstedt, 2003). Mixing laws to determine the rheological properties of a composite material are still in development, however, and depend upon the specific composition and the geometry of each phase (e.g., Tullis et al., 1991). As we cannot hope to explore the entire range of possible material parameters, we consider a single model consisting of quartz-diorite in the upper and lower crust and wet olivine in the upper.

### Table 1

Values used for model parameters. Power-law flow parameters are listed in Table 2.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>Direct effect parameter</td>
<td>Depth variable, see Fig. 6a</td>
</tr>
<tr>
<td>( b )</td>
<td>State evolution effect parameter</td>
<td>Depth variable, see Fig. 6a</td>
</tr>
<tr>
<td>( d_s )</td>
<td>State evolution distance</td>
<td>16 mm</td>
</tr>
<tr>
<td>( \eta )</td>
<td>Effective viscosity</td>
<td>Depth variable, see Fig. 6b</td>
</tr>
<tr>
<td>( \eta_{\text{rad}} )</td>
<td>Radiation damping coefficient</td>
<td>9.5 MPa s/m</td>
</tr>
<tr>
<td>( f_0 )</td>
<td>Reference friction coefficient for Steady sliding</td>
<td>0.6</td>
</tr>
<tr>
<td>( P_p )</td>
<td>Hydrostatic pressure</td>
<td>(1000 kg m(^{-3})) × g</td>
</tr>
<tr>
<td>( R )</td>
<td>Gas constant</td>
<td>8.31 J K(^{-1}) mol(^{-1})</td>
</tr>
<tr>
<td>( \mu )</td>
<td>Shear modulus</td>
<td>30 GPa</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Density of rock</td>
<td>3 g cm(^{-3})</td>
</tr>
<tr>
<td>( \tau_\text{eff} )</td>
<td>Effective normal stress</td>
<td>Depth variable</td>
</tr>
<tr>
<td>( V_0 )</td>
<td>Reference velocity</td>
<td>10(^{-5}) m s(^{-1})</td>
</tr>
<tr>
<td>( V_L )</td>
<td>Loading velocity</td>
<td>10(^{-5}) m s(^{-1})</td>
</tr>
</tbody>
</table>
mantle. Another logical choice for the rheological layers would have been the traditional “crème brûlée” model, with the upper crust represented by quartzite, the lower crust by feldspar, and the upper mantle by olivine. We found, however, that wet quartzite is so weak that coseismic slip is entirely confined within a few kilometers of the surface, in contradiction with observed earthquake depths. Additionally, while quartzite may be representative of the relatively granitic upper crust, the lower crust is expected to be more mafic (Hacker et al., 2015; Shinevar et al., 2015). Thus, we instead chose to use quartz-diorite, a material with a much deeper brittle-ductile transition, for the upper and lower crust. Though our choice to use the same composition for all our simulations neglects the rheological complexity of the crust, the structure is consistent with postseismic data from the Hector Mine and Landers earthquakes in the Mojave region (Freed and Bürgmann, 2004). Additionally, it enables the model to capture the essential features on which we wish to focus: the brittle-ductile transition, and the partitioning between fault motion and bulk viscous flow as a function of depth.

We smoothly transition the effective viscosity between the crust and mantle using a mixing law that applies when the constituent materials form layers. Mixing is likely to be the case at the base of the crust, where magmatic underplating takes place. The mixing law is derived in Fig. 5, and we now state the equations for two phases. For a composite of two materials, the power-law parameters are

\[ n = \frac{1}{\phi_1/n_1 + \phi_2/n_2}, \]

\[ A_n = (A_1 n_1)^{\phi_1} (A_2 n_2)^{\phi_2}, \]

\[ Q_n = n_1 \left( \frac{\phi_1 Q_1}{n_1} + \frac{\phi_2 Q_2}{n_2} \right), \]

where \( \phi_1 \) and \( \phi_2 \) are the volume fractions of each phase, shown in Fig. 5.

4. Discretization of governing equations

We discretize the governing equations using finite differences for the spatial discretization and a fully explicit, adaptive Runge-Kutta time-stepping scheme for the temporal discretization. The spatial discretization, which is described in Appendix A, is provably stable; the proof is based on establishing an energy balance for the semi-discrete problem that mimics the mechanical energy balance of the continuum problem. The fully explicit temporal discretization is possible because the largest stable time step in the interseismic period is limited by the minimum Maxwell time, \( \min(T_{\max}) = \min(n/\mu) \), which is the characteristic timescale for the viscous strains. The smallest Maxwell time in the simulations considered is 20 years, which is not prohibitive relative to the earthquake recurrence interval of approximately 350 years. Because we use the quasi-dynamic approximation, the time-step limitation imposed by rate-and-state friction and radiation damping varies over many orders of magnitude through the coseismic and interseismic periods, and thus requires a time step on the order of milliseconds in the coseismic period but is larger than \( \min(T_{\max}) \) in the interseismic period.

Our algorithm extends that of Erickson and Dunham (2014), which was developed for the linear elastic problem, to the power-law viscoelastic problem. In addition, we generalize the spatial discretization from uniform grid spacing to variable grid spacing, greatly reducing the computational load, using an approach that is similar to Erickson et al. (2017). We also extend the algorithm to 4th order accuracy in space.

To ensure accuracy of the solution, certain length scales must be resolved by several grid points. From stability analysis of steady frictional sliding (e.g., Ruina, 1983; Rice, 1983; Rice, 1993; Rice et al., 2001), the critical length scale for unstable sliding between elastic half-spaces with rate-and-state friction is

<table>
<thead>
<tr>
<th>Material</th>
<th>A (MPa$^{-1}$ s$^{-1}$)</th>
<th>n</th>
<th>Q (kJ mol$^{-1}$)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quartz-diorite</td>
<td>$1.3 \times 10^{-3}$</td>
<td>2.4</td>
<td>219</td>
<td>Hansen and Carter (1982);</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Freed and Bürgmann (2004)</td>
</tr>
<tr>
<td>Wet olivine</td>
<td>$3.6 \times 10^{5}$</td>
<td>3.5</td>
<td>480</td>
<td>Hirth and Kohlstedt (2003)</td>
</tr>
</tbody>
</table>

Fig. 5. Volume fractions for quartz-diorite (\( \phi_1 \), blue) and olivine (\( \phi_2 \), red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. Lithospheric strength

The choices of frictional and rheological parameters, combined with assumed geotherms and constant reference strain rate, produce the lithospheric strength profiles plotted in Fig. 6c. In the three viscoelastic simulations considered in this paper, differing only by choice of geotherm, the brittle-ductile transition occurs at approximately 15 to 20 km depth.

To set the geotherm, we use heat flow measurements for the Mojave Desert region. Surface heat flow in the region is 60–75 mW/m$^2$ (Lachenbruch et al., 1985), resulting in estimates of the geothermal gradient ranging from 15 to 30 K/km (Williams, 1996). Yang and Forsyth (2008) estimate the Moho temperature to be 973–1098 K based on seismic attenuation, and the upper end of this range is consistent with xenoliths from the region (Behr and Hirth, 2014). Based on these data, and neglecting both heat generation and spatial variations in crustal conductivity for simplicity, we assume a set of three linear geotherms: 20 K/km, 25 K/km, and 30 K/km. These three geotherms produce the effective viscosity profiles plotted in Fig. 6b. In order to study the effect of moving the brittle-ductile transition relative to the earthquake nucleation depth, we use 30 K/km to set the depth-dependence of the frictional parameters for all simulations.

In this study, we assume all viscoelastic deformation happens as a result of dislocation creep, which is grain-size insensitive and has a stress exponent \( n = 3 \sim 5 \) (Rybachy and Dresen, 2000). This assumption is well-supported by experimental and microstructural evidence for quartz (Hirth et al., 2001) and olivine (Hirth and Kohlstedt, 2003), as summarized in deformation mechanisms maps in Bürgmann and Dresen (2008). Lower crustal shear zones, however, may have a small enough grain size that dislocation creep and diffusion creep could both be active (Rybachy and Dresen, 2000, 2004). We also neglect the effects of transient creep, which could reduce the effective viscosity during the postseismic period (Freed et al., 2010; Masuti et al., 2016).

4. Discretization of governing equations

We discretize the governing equations using finite differences for the spatial discretization and a fully explicit, adaptive Runge-Kutta time-stepping scheme for the temporal discretization. The spatial discretization, which is described in Appendix A, is provably stable; the proof is based on establishing an energy balance for the semi-discrete problem that mimics the mechanical energy balance of the continuum problem. The fully explicit temporal discretization is possible because the largest stable time step in the interseismic period is limited by the minimum Maxwell time, \( \min(T_{\max}) = \min(n/\mu) \), which is the characteristic timescale for the viscous strains. The smallest Maxwell time in the simulations considered is 20 years, which is not prohibitive relative to the earthquake recurrence interval of approximately 350 years. Because we use the quasi-dynamic approximation, the time-step limitation imposed by rate-and-state friction and radiation damping varies over many orders of magnitude through the coseismic and interseismic periods, and thus requires a time step on the order of milliseconds in the coseismic period but is larger than \( \min(T_{\max}) \) in the interseismic period.

Our algorithm extends that of Erickson and Dunham (2014), which was developed for the linear elastic problem, to the power-law viscoelastic problem. In addition, we generalize the spatial discretization from uniform grid spacing to variable grid spacing, greatly reducing the computational load, using an approach that is similar to Erickson et al. (2017). We also extend the algorithm to 4th order accuracy in space.

To ensure accuracy of the solution, certain length scales must be resolved by several grid points. From stability analysis of steady frictional sliding (e.g., Ruina, 1983; Rice, 1983; Rice, 1993; Rice et al., 2001), the critical length scale for unstable sliding between elastic half-spaces with rate-and-state friction is
\[ h = \frac{\mu d_c}{\sigma_c (b - a)}. \]  

(20)

Numerical simulations for the aging law have found that an additional critical length scale, characterizing the size of the region of rapid strength degradation behind the tip of a propagating rupture, is

\[ L_b = \frac{\mu d_c}{\sigma_c b}. \]  

(21)

(Dieterich, 1992; Ampuero and Rubin, 2008). This scale is more challenging to resolve.

### 4.1. Temporal discretization

The variables integrated in time are slip on the fault \( \delta \), the state variable \( \psi \), and the viscous strains \( \gamma_{xy}^0 \) and \( \gamma_{xz}^0 \). The temporal integration is performed using a third-order accurate, adaptive Runge-Kutta time-stepping algorithm (Hairer et al., 1993) in order to capture the range of time scales across the earthquake cycle. It is outlined here as if for forward Euler for simplicity. To integrate from step \( n \) (time \( t^n \)) to step \( n + 1 \) (time \( t^{n+1} = t^n + \Delta t \)) over time step \( \Delta t \):

1. Set boundary conditions: \( u(L_x, z, t^n) = V_x t^n/2 \), \( u(0, z, t^n) = \delta^n/2 \).
2. Solve the static equilibrium equation Eq. (1) for displacement \( u^n \):

\[ A u^n = b + C L_{y} \psi^{n,0} + D Y_{c,sl}^{n,0}, \]  

(22)

where \( A \), \( C \), and \( D \) are matrices set by the spatial discretization Eq. (A.15) and \( b \) is a vector containing boundary data.
3. Use displacement \( u^n \) to compute the quasi-static shear stress on the fault:

\[ \sigma^u_{yy} = \mu \left( \frac{\partial u^n}{\partial y} - \gamma_{xy}^{n,0} \right) \bigg|_{\psi=0}, \]

where the discretized form is given in Eq. (A.13).
4. Solve the nonlinear force balance equation on the fault (Eq. (5)) for slip velocity \( V^u \):

\[ \sigma_f (\psi^n, V^n) = \sigma^u_{yy} - \eta_{visc} V^n. \]

5. Compute rates for the viscous strains and state variable using Eqs. (6), (A.20), and (A.21).
6. Update slip \( \delta \), the state variable \( \psi \), and the viscous strains \( \gamma_{xy}^{n+1} \) and \( \gamma_{xz}^{n+1} \), e.g., \( \delta_{n+1} = \delta^n + V^n \Delta t \).

Earthquake cycle simulations like this must be spun-up, meaning that it takes some (potentially large) number of earthquakes for each simulation to settle into a limit cycle. While there is no proof that each simulation must settle into a limit cycle, in our experience all power-law viscoelastic simulations with this value of state evolution distance \( d_e \) seem to do so. We have found that linear elastic simulations take very little time to spin-up, whereas viscoelastic simulations require thousands of earthquakes to spin-up if the viscous strains are initialized at zero. This spin-up time scales with the Maxwell time for each simulation, and reflects the time needed for bulk viscous flow to communicate the stress state at the fault to the surrounding material. To reduce the computational expense of model spin-up, we have developed a multi-step process:

1. This step determines the displacement boundary condition on the fault that is consistent with the predicted stress profile on the fault (as shown in Fig. 6c). We do not integrate in time, and enforce traction on the fault rather than displacement, setting the traction to that of the predicted stress profile for a reference strain rate of \( 10^{-14} \text{ s}^{-1} \). The remote boundary displacement may be set to any value. The viscous strains are set to zero. We solve the static equilibrium equation (Eq. (1)) for displacement and evaluate it on the fault. This provides an initial guess for the fault displacement boundary condition that is consistent with a plausible lithospheric strength profile.

2. This step provides a fault shear stress profile that is more consistent with the rate-and-state friction law, without assuming a reference slip velocity and strain rate. We switch to the full problem, with rate-and-state friction on the fault and nonzero viscous strains in the bulk, and use the displacement from step 1 to set the initial fault displacement. We set the initial viscous strains to zero everywhere in the domain, and initialize the state variable as \( \delta_0 \). We integrate in time until the shear stress on the fault settles into a limit cycle. Note that the stress on the fault reaches a limit cycle far sooner than does...
the off-fault stress, since the fault stress does not require viscous flow to accumulate. In our experience, the fault shear stress stabilizes after order 10 earthquakes.

3. This step allows bulk viscous flow to propagate the stress state at the fault to the off-fault material without requiring the computational load necessary to model earthquakes. We enforce traction on the fault using a representative interseismic shear stress profile from step 2 as the boundary condition, and corresponding \( y_{ex} \) and \( y_{ev} \) as initial conditions. We integrate in time until the elastic strain at each depth reaches a steady state value (see Eq. (23) and surrounding discussion). This eliminates the need to take very small time steps during the coseismic period, thus greatly decreasing the number of time steps needed to spin-up the simulation. It is possible to further speed up this step by using a larger grid spacing, as the critical grid spacing for rate-and-state friction does not need to be resolved, followed by an interpolation onto the finer grid.

4. We use the final boundary displacements and viscous strains from step 3 as the initial conditions for a full simulation with rate-and-state friction on the fault. For the state variable, we use the final value from step 2. It still takes a few earthquakes to reach the desired limit cycle, but not thousands.

5. Results

In this section, we summarize the results of the three viscoelastic simulations and a reference elastic simulation. We discuss features of the earthquake cycle, such as the recurrence interval and nucleation depth, which contain information about the ways in which bulk viscous flow at depth loads the seismogenic zone. We also examine the spatial distribution and degree of localization of viscous strain. We begin by presenting slip histories and viscous strain distributions in Section 5.1. The variation in viscous strain rate in the postseismic and interseismic periods is addressed in Section 5.2, as is the relative partitioning of fault slip into coseismic slip and frictional afterslip. The stress in the lithosphere for each model is discussed in Section 5.3. We also discuss the sign and magnitude of basal tractions within the lithosphere, which addresses the question of whether the upper crust drags the lower crust and upper mantle or vice versa. Then, we summarize the spatial and temporal variation of effective viscosity, which also relates to the strength of the lithosphere. Finally, we discuss the means by which observations of surface deformation might be able to differentiate the models in Section 5.5.

5.1. Fault slip and viscous strain

Cumulative slip on the fault over a sequence of earthquakes spanning about 2000 years is plotted in Fig. 7. In (a), the linear elastic simulation, the fault creeps steadily at depth at the tectonic loading rate. This creep causes a stress concentration to build up at the base of the seismogenic zone, and when the stress concentration is large enough, an earthquake nucleates at that depth. In this case, the earthquakes occur every 366 years and produce about 10 m of slip. This slip is quite large due to the absence of heterogeneity as well as our assumption of hydrostatic pore pressure, which results in a large stress drop for each earthquake. At the surface, the fault also creeps a small amount in the interseismic period as a result of the high \( h' \) (from low effective stress) in this region. In contrast, in the viscoelastic simulations (b)–(d) fault creep in the lower crust and upper mantle lags behind fault slip in the seismogenic zone because the off-fault material at depth is weak enough that it inhibits interseismic fault creep. Instead, bulk viscous flow accommodates much of the tectonic loading, preventing the fault from creeping at the loading velocity. As the geotherm increases, the proportion of the loading that is partitioned into bulk viscous flow increases, such that in (d), with a geotherm of 30 K/km, the fault does not slip below 16 km depth. For the power-law simulations, the recurrence interval is effectively the same as for the linear elastic simulation: 365 years for the 30 K/km simulation, 375 years for the 25 K/km simulation, and 366 years for the 20 K/km simulation. This suggests that the loading of the seismogenic crust is effectively the same, regardless of the deformation mechanism at depth, an idea which we discuss further below.

While plots of cumulative fault slip provide some insight into the depth-dependence of bulk viscous flow, they do not show its spatial distribution. Both components of the off-fault viscous strain that accumulate over a single earthquake are plotted in Fig. 8. Viscous strain \( y_{ev} \) component that accommodates tectonic displacement, can be examined to determine the width of the shear zone beneath the fault. The peak of \( y_{ev} \) occurs at the depth of least interseismic fault creep, which becomes shallower with increasing geotherm. Also, the width of the shear zone is on the order of a few kilometers at this depth, and broadens significantly below this. The shear zone is narrowest for the warmest geotherm, though the 25 K/km and 30 K/km simulations are quite similar. The 20 K/km simulation, however, produces a much broader shear zone. This general dependence of shear zone width on geotherm is in agreement with the analytical calculations of Moore and Parsons (2015), who find that the primary control on shear zone width is the geotherm. \( y_{ev} \) relaxes basal tractions on the lithosphere, which we also discuss in Section 5.3. Where \( y_{ev} \) is negative, the upper crust drags the lower crust, and the lower crust in turn drags the upper mantle. The largest values of \( y_{ev} \) occur close to the fault: within 5 km of the fault for the 20 K/km simulation, and within 15 km for the 25 and 30 K/km simulations.

Strikingly, in many respects the coseismic phase of each viscoelastic simulation is quite similar to that of the linear elastic simulation. In every case, the earthquakes nucleate at the same depth with approximately the same recurrence interval. Additionally, the down-dip limit of coseismic slip in each simulation is 14 km depth, which means that it is controlled by the transition in frictional parameters. These similarities in recurrence interval and coseismic slip distribution between elastic and viscoelastic simulations were also observed by Kato (2002). The similar behavior can be explained by the fact that in each simulation no appreciable viscous strain occurs above 14 km depth, and therefore the same amount of interseismic fault creep occurs at 10–14 km depth. This creep likely provides much of the loading to the seismogenic zone and the upper crust. Thus, the shear stress on the fault, and therefore the recurrence interval, are approximately the same. This apparent similarity is limited, however, and does not hold for the evolution of the stress fields below the seismogenic zone with time, nor for the surface deformation.

5.2. Partitioning of tectonic loading into fault slip, elastic, and viscous strains

Figs. 7 and 8 demonstrate that the relative partitioning of the tectonic loading into fault slip and bulk viscous flow varies with depth. This is shown in more detail in Figs. 9 and 10 for the simulation with a geotherm of 30 K/km. At each depth, the contributions of fault slip, viscous strain, and elastic strain must add up to the tectonic loading displacement, meaning that at all times and depths

\[
V_{LT} = \delta + 2 \int_{0}^{t_f} y_{f} \, dy + 2 \int_{0}^{t_f} y_{ev} \, dy,
\]

(23)

where \( y_{f} \) is the elastic strain. The left-hand side is the tectonic displacement, and the second and third terms on the right-hand side correspond to the contributions of viscous and elastic strain, respectively. Because the off-fault stresses return to the same level at the start of each earthquake, the integrated elastic strain cannot grow or decay over multiple earthquake cycles, but instead must oscillate about a constant value. Only fault slip and integrated viscous strain may increase with time to keep up with the tectonic loading displacement. This oscillation in elastic strain is similar to the kinematically consistent models of Devries and Meade (2016).
The upper crust is effectively elastic, so at the surface of the Earth, plotted in Fig. 9, there is negligible viscous strain. Tectonic loading causes the elastic strain to build during the interseismic periods, and then slip on the fault during each earthquake releases this elastic strain. A small amount of interseismic creep occurs at the surface as a result of the low normal stress, so during the interseismic period the blue line is not perfectly horizontal and the elastic strain rate decreases slightly over the interseismic period. At the end of each cycle, integrated elastic strain returns to the same position, reflecting the fact that the simulation is in a limit cycle. This figure, except for the shallow interseismic fault creep, is representative of the entire seismogenic zone, which extends down to 13 km depth for this simulation.

Deeper, in the brittle-ductile transition zone this simulation ranges from 14 to 17 km depth for the 30 K/km simulation, plotted in Fig. 10a–c, fault slip and integrated viscous strain both accommodate a significant portion of the loading. At the down-dip limit of coseismic slip, plotted in Fig. 10a, both coseismic slip and a small amount of viscous strain occur, demonstrating that the earthquake is able to penetrate a small distance into the brittle-ductile transition zone. Frictional aftserslip and interseismic fault creep occur at this depth as well, and together they accommodate most of the tectonic loading displacement. Surprisingly, at these depths, viscous strain accumulates steadily throughout the earthquake cycle, rather than accumulating at a greater rate in the postseismic period.

A different behavior occurs in the purely ductile regime well below the brittle-ductile transition. The purely ductile regime starts at about 17 km depth and extends to the bottom of the model domain. As seen in Fig. 10d–f, all of the tectonic loading is accommodated by viscous strain and no fault slip occurs. Viscous strain does not accumulate at a uniform rate, however. More viscous strain accumulates during the first half of the interseismic period than the last half. Comparing Fig. 10c–e, it is apparent that the viscous strain rate transitions from uniform in time in the brittle-ductile transition zone to variable down to the bottom of the computational domain (at 40 km depth). At sufficiently great depth, with spatially uniform remote loading, the viscous strain rate must become uniform with time, though our computational domain does not extend deep enough to capture this. This figure is representative of the other two viscoelastic simulations as well, though for them the brittle-ductile transition zone is deeper and bigger. In the 25 K/km simulation it occurs at 16–20 km depth. The 20 K/km simulation produces significant fault creep to the bottom of the computational domain, so for this geotherm the purely ductile regime is not modeled.

5.3. Temporal evolution of stress

In this section, we discuss the spatial and temporal evolution of lithospheric stress components $\sigma_{xy}$ (the shear stress acting on vertical planes parallel to the fault) and $\sigma_{xz}$ (the shear stress on horizontal planes, with negative $\sigma_{xz}$ indicating that material closer to Earth's surface is dragging the material below it). First, we discuss the predictions our simulations make of the stress within the lithosphere as a function of depth. Then, we discuss the temporal evolution of $\sigma_{xy}$ throughout the interseismic period, as this has implications regarding the triggering of earthquakes via viscoelastic stress transfer. Finally, we discuss the implications of the results for $\sigma_{xz}$ for basal tractions.

The spatial distribution of $\sigma_{xy}$ at the start of the interseismic period is plotted in the first row of Fig. 11. The linear elastic simulation, Fig. 11a, has a radically different stress field than the power-law simulations. Far from the fault, stress becomes uniform with depth, a result of the remote boundary condition of uniform displacement and depth-independent elastic material response. In contrast, the power-law simulations (Fig. 11g, m, and s) produce near-fault stress profiles that resemble the lithospheric strength profile plotted in Fig. 6c. Far from the fault, the distribution changes and the prediction is that lithospheric strength will reside in the upper crust, consistent with the fact that all of these simulations are “crème brûlée” models.

An alternative representation of the spatial distribution for $\sigma_{xy}$ for the 30 K/km simulation is shown in Fig. 12, which shows series of vertical transects through Fig. 11s at selected times. Within a few kilometers of the fault, plotted in black, the stress profile in the lower crust is well-approximated by a reference strain rate of $10^{-12}$ s$^{-1}$. In
the far field, plotted in red, the stress profile is better approximated by a much lower reference strain rate of $2.5 \times 10^{-17}$ s$^{-1}$, for which the lower crust is predicted to be quite weak. These plots highlight the fact that the average strain rate is orders of magnitude higher near the fault than away from it. Additionally, Fig. 13 shows the shear stress on the fault at selected times throughout the interseismic period for each simulation. In all four simulations, the shear stress evolves very similarly in the brittle upper crust.

Returning to Fig. 11, the viscoelastic simulations also predict unrealistically high values of $\sigma_{xy}$ in the upper crust (200–250 MPa). To explain these stresses, we calculate a global force balance on the gray rectangular region shown in Fig. 14, which extends from the fault to the surface for the 30 K/km simulation. The change in each quantity has been plotted, not the absolute value. This is in the brittle regime, which extends down to 13 km for this geotherm, so no viscous strain accumulates. At the end of each cycle, elastic strain returns to the same level, reflecting the fact that the simulation is in a limit cycle with periodic earthquakes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
remote side boundary and from the top free surface to some arbitrary depth. The depth of the bottom boundary of the rectangle is varied. The tectonic driving force per unit distance along strike (i.e., integrated stress along a vertical cross-section at the remote boundary) is balanced by the sum of integrated basal traction (along the bottom of the rectangular region) and integrated stress along a vertical cross-section along the fault and its deep extension. To meet the imposed traction-free boundary conditions on the top and bottom of the computational domain, the integrated basal traction must vanish when the bottom of the rectangular region coincides with either the top or bottom boundary. However, the contribution from basal traction is nonzero, and determined by the rheology and deformation style, when the bottom of the rectangular region is placed elsewhere within the computational domain. The basal traction at the boundary between the lower crust and upper mantle is relatively low. For example, for the 30 K/km simulation, the integrated basal traction is 30 GN/m at 30 km. The total resistive force acting on the fault cross-section, however, is high because the region below the fault has a high strain rate, and hence high stress. For the 30 K/km case, the force on the fault is 1990 GN/m. Therefore the tectonic driving force is quite high (2020 GN/m). However, negligible load can be carried by the lower crust and mantle far from the fault, which are quite weak because of the low strain rate. Hence stresses in the upper crust away from the fault become quite high.

In the real Earth, such high stresses would not occur. Instead, they might be relaxed by slip on additional, subparallel strike-slip faults, each of which might elevate the basal traction at the base of the upper crust just below it and therefore successively decrease the load borne by the remote upper crust. While these basal tractions are not too large, the net effect of the basal tractions from several such fault roots could eventually contribute to the force balance. However, with multiple faults, the slip rates on each fault would be lower (for the same tectonic loading rate), and therefore the net effect of all fault roots on the basal tractions might be similar to that of a single fault model. Nonetheless, there are many places where plate boundary strike-slip faults do not exist in isolation, but instead occur as a set of subparallel structures, such as the Marlborough Fault system in New Zealand (e.g., Wilson et al., 2004) and the San Andreas-San Jacinto-Elsinore sytem in

---

**Fig. 10.** Partitioning of tectonic loading displacement (gray) into fault slip (blue), integrated viscous strain (black), and integrated elastic strain (red) for the 30 K/km simulation. The change in each quantity has been plotted, not the absolute value. In the brittle-ductile transition zone, beginning at the down-dip limit of coseismic slip (14 km), viscous strain accumulates at a constant rate throughout each cycle; coseismic and postseismic fault slip are clearly visible. In the ductile regime, in which no fault slip occurs, viscous strain accumulates more quickly early in each interseismic period, though this lessens with depth. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Southern California (e.g., Lundgren et al., 2009), and our model suggests a possible reason for these structures. Alternatively, off-fault deviatoric stress in the upper crust might be limited by the occurrence of distributed plastic deformation (e.g., Bird, 2009; Shelef and Oskin, 2010; Herbert et al., 2014; Erickson et al., 2017), which in the brittle crust might manifest as multiple smaller faults (Evans et al., 2016). Another possible way to reduce the stress in the upper crust, without invoking additional faults or plasticity, is to increase basal tractions. While we have not seen evidence for this in our simulations, it might be possible that some rheologies, such as a stronger, more mafic lower crust, would produce substantial basal tractions. Finally, various processes could weaken the fault and its deep extension. These include pore pressures in excess of hydrostatic (e.g., Rice, 1992) as well as dynamic weakening mechanisms such as thermal pressurization and flash heating within the seismogenic zone (e.g., Rice, 2006; Noda et al., 2009; Brantut and Platt, 2016), and grain size reduction, foliation, and thermo-mechanical feedback (viscosity reduction from shear heating) in the ductile root of the fault (e.g., White et al., 1980; Montési and Zuber, 2002; Bürgmann and Dresen, 2008; Platt and Behr, 2011; Takeuchi and Fialko, 2012; Montési, 2013).

The remaining rows of Fig. 11 show the change in $\sigma_{xy}$ relative to the start of the interseismic period. This is equivalent to plotting changes in Coulomb stress for a vertical strike-slip fault. In the linear elastic simulation, the change in stress over the first 10 years is that of a crack located between 12 and 20 km, the depth of frictional aftselslip. After this, the stress state transitions to that of a screw dislocation, resulting from the fault’s transition to locked in the upper crust and creeping in the lower crust and upper mantle. Despite the dramatic differences in absolute value of shear stress for the linear elastic simulation and the 20 K/km viscoelastic simulation, the changes in $\sigma_{xy}$ are nearly identical.
throughout the interseismic period, as very little bulk viscous flow occurs in this viscoelastic simulation. For the warmer two viscoelastic simulations, the stress distributions are quite similar for the first 5 years, as very little viscous flow occurs over this short time frame. After this, bulk viscous flow at depth produces a very different stress state below 20 km. These models also predict a larger shallow perturbation in shear stress at each time. The implication is that the increased bulk viscous flow in the warmest viscoelastic model transmits the effect of the earthquake farther from the fault than does the coolest viscoelastic model.

The stress component $\sigma_{xz}$ throughout the interseismic period is shown in Fig. 15 for each power-law simulation. As is the case for $\sigma_{xy}$, $\sigma_{xz}$ is largest in the portion of the crust that accommodates significant brittle deformation, but tends to small values beyond ~10 km from the fault. The sign of $\sigma_{xz}$ is of particular significance: negative $\sigma_{xz}$ means that shallower material is dragging the deeper material. Thus, for the simulation with a geotherm of 30 K/km (Fig. 15g–i), the upper crust is dragging the lower crust and mantle. For the simulation with a geotherm of 20 K/km (Fig. 15a–c), however, a change in sign occurs. Within 5 km of the fault, the upper mantle is dragging the lower crust.

Fig. 12. Vertical transects of $\sigma_{xy}$ for 30 K/km simulation at selected times, matching Fig. 11s, starting at the fault (black) and moving to the remote boundary 120 km away from the fault (red). The time since the end of the previous earthquake is listed in the bottom right of each panel: (a) the start of the postseismic period, (b) early in the interseismic period, and (c) at the end of the interseismic period and at the start of the next earthquake. Also shown are the predicted viscous strengths for reference strain rates of $2.5 \times 10^{-17}$ s$^{-1}$ (dotted), $10^{-14}$ s$^{-1}$ (dashed), and $10^{-12}$ s$^{-1}$ (solid). On the fault, the stress profile reflects viscous strength profile shown in Fig. 6c with a reference strain rate of $10^{-12}$ s$^{-1}$, while the remote stress is best described with a reference strain rate of $2.5 \times 10^{-17}$ s$^{-1}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 13. Shear stress on the fault at various times throughout the interseismic period: (a) and (b) show the fate of the stress concentration left behind by the earthquake; (c) postseismic creep causes the development of a stress concentration just below 10 km depth; (d) the stress concentration grows and moves up the fault; and (e) the nucleation of the next earthquake. Despite the differences in stress level in the lower crust and upper mantle, the shear stress is very similar in the brittle upper crust in each simulation.
Fig. 14. Balance between remote tectonic driving force (red), resistive frictional force (yellow), and integrated basal traction (blue), all per unit distance along strike, for the gray shaded rectangular region. The force balance is calculated for variable locations of the lower boundary of the rectangle, labeled “depth” in (a)–(d). In the viscoelastic simulations, the lower crust and upper mantle are quite weak far from the fault, causing the upper crust to bear a substantial part of the load and resulting in unrealistically high stresses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

![Diagram showing force balance](image)

Fig. 15. Horizontal shear stress $\sigma_{xz}$ (shown in both color scale and contours) at selected times throughout the interseismic period, with titles indicating the time since the last earthquake. Each column pertains to a different geotherm: (a)–(c) 20 K/km, (d)–(f) 25 K/km, and (g)–(i) 30 K/km. Past 20 years, the stress field changes relatively little. Negative $\sigma_{xz}$ indicates that the material above is dragging the material below; that is, in (d)–(i) the upper crust is dragging the lower crust and mantle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

![Contour maps showing stress field](image)
while at greater distances the reverse is occurring. Overall, these results mean that, for a broad range of geotherms and effective viscosities, the crust loads the mantle. Our results are consistent with the predictions of Lachenbruch and Sass (1992): a weak lower crust must result in low basal drag.

5.4. Effective viscosity

The power of modeling the power-law rheology is that the viscosity structure is not imposed. Instead, a spatially and temporally variable viscosity structure develops that is consistent with the stress state. Fig. 16 shows the spatial distribution for effective viscosity at the start and end of an interseismic period for each simulation. The vertical structure is set predominantly by the geotherm, so in the upper crust the effective viscosity is very high, meaning that the upper crust is effectively elastic. Additionally, a zone of relatively low effective viscosity has developed within 20 km of the fault at depth. As was the case for the viscous strain \( \dot{\gamma}_x \), this zone broadens with depth.

As demonstrated by Fig. 16, the effective viscosity varies throughout the interseismic period. Fig. 17 shows the extent of this variation for a vertical transect 1 km from the fault during one interseismic period. In the 20 K/km simulation, the effective viscosity is approximately constant throughout the interseismic period, whereas in the 25 K/km simulation it varies by a factor of two in the lower crust, and in the 30 K/km simulation it varies by more than an order of magnitude. This factor of ten increase in effective viscosity through the postseismic period is not captured in predictions assuming a single reference strain rates. In particular, for the 30 K/km simulation, the effective viscosity is close to constant with depth in the lower crust and upper mantle during much of the interseismic period.

While the simulations presented here encompass a small portion of rheological parameter space, they nevertheless span the range of effective viscosity estimates for the Western US from deformation studies (compare estimates for the Western US, plotted in Fig. 3, with the modeled effective viscosities plotted in Fig. 17). The broad range of the estimates could suggest that the behavior of the lower crust is insensitive to the effective viscosity; however, our simulations make very different predictions for how deformation is accommodated within the lower crust and upper mantle, and the associated stress state in the lithosphere. Thus, the simulations demonstrate that the effective viscosity estimates imply a great deal of uncertainty in the predominant deformation mechanism of the lower crust.

5.5. Postseismic deformation

In this section, we examine the predicted surface deformation throughout the postseismic and interseismic periods. We also discuss relative contributions of frictional afterslip and bulk viscous flow to the total surface deformation. While it can be challenging to disentangle the effects of each deformation process in inversions of crustal deformation data, this is straightforward in our forward-modeling context.

Fig. 18a shows the displacement 20 km from the fault on Earth's surface over three earthquake cycles. Both the total displacement and the contribution from fault slip are plotted (the difference being the contribution from bulk viscous flow). While in the 20 and 25 K/km simulations, the station moves at an approximately uniform rate over much of the interseismic period, in the 30 K/km simulation the station experiences more displacement in the first third of the interseismic period than in the last third. Additionally, for the 25 and 30 K/km
simulations, bulk viscous flow is responsible for most of the station's motion. In contrast, in the 20 K/km simulation, fault creep at depth contributes about half of the station's motion.

Fig. 17b and c show the displacement in the postseismic period. Over this time frame, afterslip comprises almost all of total displacement for the 20 K/km simulation. In the 25 and 30 K/km simulations, afterslip comprises most of the station motion in the first few months, with bulk viscous flow becoming increasingly important one year after the earthquake. Over most of the first decade of the postseismic period, the 30 K/km simulation predicts the most station displacement and the 20 K/km simulation predicts the least. In the first two months, however, the 25 K/km simulation actually predicts the most station motion. This results from the fact that in this time frame more afterslip occurs in the 25 K/km simulation than in the 30 K/km simulation. After two months, increased bulk viscous flow causes the 30 K/km simulation to pass the 25 K/km simulation.

Fault creep at depth and bulk viscous flow contribute to station motion for different time frames and at different distances from the fault, as illustrated in Fig. 19. In general, fault creep is significant over a shorter time frame and closer to the fault than bulk viscous flow. For each simulation, the contribution of fault creep is greatest within 25 km of the fault, while bulk viscous flow is significant even 50 km from the fault. In addition, the fact that the dashed lines are nearly identical at 25%T and 75%T means that the fault is creeping at a steady rate by 25%T (T is the recurrence interval). In contrast, the solid lines are not identical for the 25 K/km and 30 K/km simulations, meaning that bulk viscous flow is not occurring at a steady rate.

It may be possible to discriminate between these models using postseismic data spanning years after a large earthquake. For the 20 K/km simulation, most of the station motion can be produced by a model for frictional afterslip. Also, the component resulting from bulk viscous flow should be well fit with a single effective viscosity that applies both to the interseismic data prior to the earthquake and to the entire postseismic period. In contrast, the 30 K/km simulation would require a large contribution from bulk viscous flow, particularly after the first 6 months. If using a linear rheology, at least two effective viscosities would be needed to fit the data.

Fig. 18. (a) Displacement at Earth's surface 20 km from the fault. Postseismic displacement at the same location for (b) 10 years and (c) 4 months after an earthquake. The solid lines show total displacement, and the dashed lines show the displacement resulting from fault slip alone. The differences in lower crustal deformation style as a function of geotherm lead to subtle differences in displacement time series that might be measurable.

Fig. 17. Effective viscosity on a vertical transect 1 km from the fault throughout the interseismic period. The solid blue line indicates the start of the interseismic period, the dashed blue line indicates the end, and the red lines are plotted every 5 years in between. The gray lines show the effective viscosity assuming a constant strain rate of $10^{-12}$ s$^{-1}$ (solid), $10^{-14}$ s$^{-1}$ (dashed), and $10^{-16}$ s$^{-1}$ (dotted). Effective viscosity is essentially constant in time for the coolest simulation, and varies considerably for the warmest. In all three simulations, the constant reference strain rate prediction fails to capture the variability in effective viscosity with depth. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
6. Conclusions

In conclusion, we have developed a numerical method for simulating earthquake cycles with rate-and-state friction on a strike-slip fault and a power-law viscoelastic off-fault rheology. We investigate the interaction between fault slip and bulk viscous flow in a region similar to the Mojave Desert in Southern California, representing the crust with experimentally-derived parameters for quartz-diorite and the upper mantle with parameters for olivine. The three geotherms considered produce very different deformation styles in the lower crust and upper mantle, ranging from significant fault creep at depth to purely bulk viscous flow. Despite the differences in lower crustal deformation style, each simulation has almost identical recurrence interval, nucleation depth, down-dip coseismic slip limit, and total coseismic surface slip. This is probably because aseismic slip directly below the seismogenic zone occurs in a similar manner for all cases considered, in contrast to the pronounced variability of the deeper deformation mechanism. Despite these similarities, the predicted postseismic deformation varies between the simulations in a manner that might be used to distinguish the deformation process at depth.

Additionally, the simulations make different predictions for the basal tractions on the lithosphere, with the 25 and 30 K/km simulations predicting that the crust drags the mantle at all distances from the fault, and the 20 K/km simulation predicting this at > 10 km from the fault, and the reverse within 10 km of the fault. Basal tractions are a minor term in the overall force balance of the lithosphere, and as a result the integrated force on the fault is primarily balanced by shear stresses on vertical planes parallel to the fault. As a result, and because the ductile fault root experiences higher strain rates than the region far from the fault at the same depth, unrealistically high stresses develop in the upper crust far from the fault.

Resolving the crustal stress problem is a high priority, and we close by mentioning several ideas that could be tested in future studies using various extensions of the modeling framework that we have introduced here. By extending the current simulation capabilities to account for off-fault plasticity, as in Erickson et al. (2017), would allow us to investigate if distributed plastic deformation or the development of additional subparallel strike-slip faults might relieve these stresses. The high upper-crustal stresses might also be reduced by additional weakening mechanisms in either the seismogenic zone (e.g., dynamic weakening) or in the ductile fault root (e.g., grain size reduction or thermo-mechanical feedback). Finally, it is also possible that for a stronger (more mafic) lower crust, the basal tractions would play a more significant role in the force balance than suggested by our simulations. Future studies exploring these possibilities and processes are certainly warranted.

Acknowledgements

This study was supported by the U.S. Geological Survey under Grant...
Appendix A. Spatial discretization

The governing equations are discretized using summation-by-parts (SBP) finite difference operators for variable coefficient problems (Mattsson and Nordström, 2004; Mattsson, 2012). Boundary conditions are enforced weakly using simultaneous approximation terms (SAT), which penalize boundary grid data rather than overwriting grid data with the boundary conditions (injection, or strong enforcement of boundary conditions) (Mattsson et al., 2009). To prove that the numerical algorithm is stable, it is sufficient to prove that the semi-discrete energy dissipates at least as quickly as the energy of the continuum problem for homogeneous boundary conditions.

In this Appendix, we first transform the governing equations to allow for variable grid spacing. Then, we describe the spatial discretization for these transformed equations. Finally, we prove stability of the spatial discretization by solving for the continuum and semi-discrete energy balance equations.

A.1. Variable grid spacing

To ensure the accuracy of the numerical solution, both critical grid spacings, \( h^* \) (Eq. (20)) and \( L_b \) (Eq. (21)), must be resolved. These critical length scales require the use of a small grid spacing in the vicinity of the velocity-weakening regions of the fault, but elsewhere in the domain the grid spacing is permitted to be much larger. To allow for variable grid spacing, we implement a coordinate transform from curvilinear coordinates \((y,z)\) to Cartesian coordinates \((q,r)\) where \((q,r)\in [0,1]\). The two coordinates are transformed independently, such that \(y = y(q)\) and \(z = z(r)\), each transformation is invertible, and each has the property \(dq/dy > 0\) and \(dr/dz > 0\). An example is shown in Fig. 20. The static equilibrium equation, the coordinate transformed form of Eq. (1), is

\[
\left( \frac{\partial}{\partial q} \right) \left( \frac{\partial}{\partial q} \right) \left( \frac{\partial}{\partial r} \frac{\partial u}{\partial y} - \mu \gamma_{xy} \right) + \left( \frac{\partial}{\partial r} \right) \left( \frac{\partial}{\partial r} \right) \left( \frac{\partial u}{\partial z} - \mu \gamma_{xz} \right) = 0
\]

(A.1)

and the strain rates are

\[
\frac{d\gamma_{xy}}{dt} = \eta^{-1} \mu \left( \frac{\partial q}{\partial y} \frac{\partial u}{\partial y} - \gamma_{xy} \right), \quad \frac{d\gamma_{xz}}{dt} = \eta^{-1} \mu \left( \frac{\partial r}{\partial z} \frac{\partial u}{\partial z} - \gamma_{xz} \right).
\]

(A.2)

The coordinate transformation does not change the mechanical energy balance, but does change the semi-discrete energy balance.

In the simulations in this paper, we use the coordinate transform \(y(q) = L_y \sinh(10q)/\sinh(10)\). For the \(z(r)\), we use the grid spacing and transformation shown in Fig. 21.

![Fig. 20. Curvilinear coordinate transform from \((y,z)\) with variable grid spacing to \((q,r)\) with constant grid spacing.](image-url)
A.2. Semi-discrete governing equations

In this section, we describe the use of SBP finite difference operators to spatially discretize the governing equations.

Let the coordinate grid be defined by

\[ q_i = i \Delta q, \quad i = 1, ..., N_q, \quad \Delta q = 1/(N_q - 1) \] (A.3)

\[ r_j = j \Delta r, \quad j = 1, ..., N_r, \quad \Delta r = 1/(N_r - 1) \] (A.4)

where \( N_q \) is the number of grid points in the \( q \)-direction, \( N_r \) is the number of grid points in the \( r \)-direction. The grid function \( u_q \) is a vector of length \( N_q N_r \) and is \( u = (u_{01}, u_{02}, ..., u_{NqNr})^T \). Denote the boundary conditions as

\[ u(0, z, t) = b_{0z}(z, t) \] (A.5)

\[ u(Lq, z, t) = b_{qz}(z, t) \] (A.6)

\[ \sigma_{zz}(y, 0, t) = 0 \] (A.7)

\[ \sigma_{zz}(y, Lz, t) = 0. \] (A.8)

The spatial derivative operators are denoted \( D_1 = \frac{\partial}{\partial q} \) and \( D_1^{(q_\mu)} = \frac{\partial}{\partial q} \left( \mu \frac{\partial}{\partial q} + \mu q y \frac{\partial}{\partial y} \right) \). Details of the first derivative operator are in Mattsson and Nordström (2004), and of the second derivative with variable coefficients in Mattsson (2012), and we summarize only the key points here. We specifically use the narrow stencil, fully compatible form of \( D_1^{(q_\mu)} \). The one-dimensional form of the second derivative operator is

\[ D_2^{(q_\mu)} = \mu H^{-1}(\mu \frac{\partial}{\partial q} + \mu q y \frac{\partial}{\partial y}), \] (A.9)

\[ M^{(q_\mu)} = D_2^{(q_\mu)} H \mu q y D_1 + R^{(q_\mu)}, \] (A.10)

\[ q_y = (D_2 q_y)^{-1} \approx \frac{d q_y}{d y} \] (A.11)

\[ B = \text{diag}(-1, 0, 0, ..., 0, 1) \] (A.12)

where \( \mu \) is a diagonal matrix whose nonzero entries correspond to the values of the shear modulus at every point in the domain, \( H \) is a quadrature operator used for defining inner products and norms and is diagonal in the form of the SBP operators used in this paper, \( q_y \) is a diagonal matrix approximating the Jacobian of the coordinate transform, and \( R \) is an additional energy term which vanishes with grid refinement. The matrices \( \mu \) and \( q_y \) are symmetric positive definite, and \( R, H, \) and \( M^{(q_\mu)} \) are symmetric semi-positive definite.

The two-dimensional operators, denoted in bold, are formed from the one-dimensional operators by applying the Kronecker product with the identity matrices \( I_q \) and \( I_r \), such that

\[ \text{Fig. 21. (a) Grid spacing } \Delta z \text{ as a function of depth. (b) Coordinate transformation } z(r), \text{ the integral of the function plotted in (a).} \]
\( \mathbf{z}_t = I_q \otimes (D_z \mathbf{z}) \approx \left( \frac{dz}{dt} \right), \quad \mathbf{u}_t = \mathbf{z}_t^{-1} \approx \left( \frac{dr}{dz} \right) \)

\( \mathbf{y}_q = (D_q \mathbf{y}) \otimes I_q \approx \left( \frac{dy}{dq} \right), \quad \mathbf{q}_y = \mathbf{y}_q^{-1} \approx \left( \frac{dq}{dy} \right) \)

\( \mathbf{H}_q = H_q \otimes I_q, \quad \mathbf{H}_r = I_q \otimes H_r, \quad \mathbf{H} = H_q \otimes H_r \)

\( \mathbf{D}_q = D_q \otimes I_q, \quad \mathbf{D}_r = I_q \otimes D_r \)

\( \mathbf{D}^{(\mu_\phi)} = (I_q \otimes \mathbf{H}_r^{-1})(-\mathbf{M}^{(\mu_\phi)} + \mathbf{r}_q(I_q \otimes D_r)(I_q \otimes D_r)) \)

\( \mathbf{M}^{(\mu_\phi)} = (I_q \otimes D_q^T)(I_q \otimes H_r)^{-1}\mathbf{r}_q(I_q \otimes D_r) + \mathbf{R}^{e_{\phi}} \)

\( \mathbf{D}^{(\mu_q)} = (H_q^{-1} \otimes I_r)\mathbf{r}_q(I_q \otimes I_r)(I_q \otimes D_r) + \mathbf{R}^{e_{\phi}} \)

The form of \( \mathbf{R}^{e_{\phi}} \) and \( \mathbf{R^{e_{\phi}}} \) are given in Mattsson (2012). The matrices

\( \mathbf{E}_L = \text{diag}(1, 0, ..., 0) \otimes I_{tz} \)

\( \mathbf{E}_R = \text{diag}(0, 0, 1) \otimes I_{tz} \)

\( \mathbf{E}_T = I_{t} \otimes \text{diag}(1, ..., 0, 0) \)

\( \mathbf{E}_B = I_{t} \otimes \text{diag}(0, ..., 1) \)

pick out the left, right, top, and bottom boundaries, respectively. We also use the notation

\( \mathbf{B}_q = -\mathbf{E}_q + \mathbf{E}_r \)

\( \mathbf{B}_r = -\mathbf{E}_t + \mathbf{E}_b \)

The semi-discrete form of Hooke’s Law (Eq. (2)) is

\( \sigma_{y q} = \mu(q, D_q u - y_{y q}) \) \quad (A.13)

\( \sigma_{z q} = \mu(q, D_r u - y_{z q}) \) \quad (A.14)

The semi-discrete form of the static equilibrium equation (Eq. (A.1)) is

\( 0 = \mathbf{z}_q^T \mathbf{D}^{(\mu_\phi)} u - \mathbf{D}_{zz} \eta^{-1} \mathbf{q}_y E_z u - \mathbf{D}_{zy} \eta^{-1} \mathbf{q}_y E_z u + \mathbf{q}_y \mathbf{D}^{(\mu_q)} u - \mathbf{D}_{zz} \mathbf{r}_q u + \mathbf{p}_q + \mathbf{p}_r \)

\( \text{where} \)

\( \mathbf{p}_q = \mathbf{H}^{-1}_q \eta^{-1} \mathbf{q}_y E_z = \mathbf{E}_q (u - b_q) \)

\( \mathbf{p}_r = \mathbf{H}^{-1}_r \eta^{-1} \mathbf{r}_q E_z = \mathbf{E}_r (u - b_q) \)

\( \mathbf{p}_q = -\mathbf{H}^{-1}_q \mu(q, D_q u - y_{z q}) \)

\( \mathbf{p}_r = -\mathbf{H}^{-1}_r \mu(q, D_r u - y_{z q}) \)

The penalty terms \( \mathbf{p}_q, \mathbf{p}_r, \mathbf{p}_T, \) and \( \mathbf{p}_B \) weakly enforce the left, right, top, and bottom boundary conditions, Eqs. (A.5)–(A.8), respectively. The semi-discrete strain rates, corresponding to Eq. (3), are

\( \frac{dy_{yz q}}{dt} = \eta^{-1} \mathbf{q}_y (q, D_q u - y_{y q}) + \mathbf{H}^{-1}_q \eta^{-1} \mathbf{q}_y E_z (u - b_q) - \mathbf{H}^{-1}_q \eta^{-1} \mathbf{r}_q E_z (u - b_q) \)

\( \text{where} \eta \) is a diagonal, positive definite matrix whose nonzero entries correspond to the viscosity at every point in the domain. The last two terms in Eq. (A.20) are penalty terms introduced to ensure stability for Dirichlet boundary conditions at \( y = 0 \) and \( y = L_y \).

A.3. Proof of stability

The stability proof makes use the property of the Kronecker product \((A \otimes B)(C \otimes D) = (AC) \otimes (BD)\) and the property of the first derivative

\( D_1 = H^{-1}(-D_1^T H + B) \) \quad (A.22)

where \( D_1 \) can be applied in either the \( q \) or \( r \) direction, and \( H, H^{-1}, B \) should be interpreted appropriately.

We first consider the energy balance of the continuum problem. To compute the mechanical energy balance equation, multiply the static equilibrium equation (Eq. (1)) by \( \dot{u} \) and integrate over the domain. The energy balance equation is

\( \frac{d}{dt} E = \int_{t_0}^{t} \dot{u} \sigma_{y q} dz \left[ \right]_{t_0}^{t} + \int_{t_0}^{t} \dot{u} \sigma_{z q} dy \left[ \right]_{t_0}^{t} - \Phi \)

\( \text{where} \)

\( \dot{u} = \frac{1}{2} \int_{t_0}^{t} \tau dz + \frac{1}{2} \int_{t_0}^{t} \eta_{et} \tau \tau dz + \int_{t_0}^{t} \frac{1}{2} \sigma_{y q} V_y dz \left[ \right]_{t_0}^{t} + \int_{t_0}^{t} \dot{u} \sigma_{z q} dy \left[ \right]_{t_0}^{t} - \Phi \) \quad (A.24)
where the elastic strain energy is

\[ E = \frac{1}{2} \int_0^{\gamma_T} \int_0^{\gamma_T} \mu (\sigma_{ij}^e + \sigma_{ij}^s) \, dy \, dz \]  \tag{A.25} \]
and the rate of energy dissipation from viscous flow is

\[ \Phi = \int_0^{\gamma_T} \int_0^{\gamma_T} \eta \gamma_{ij} \dot{\varepsilon}_{ij} \, dy \, dz. \]  \tag{A.26} \]

In Eq. (A.24), the first term on the right-hand side is the frictional heat dissipation, the second is the energy that flows into the fault via seismic radiation (as idealized with the radiation damping approximation), and the third and fourth are the rate of work on the system by the boundary tractions. For homogeneous boundary conditions, the boundaries do not contribute energy to the system.

To compute the semi-discrete energy balance equation, left multiply the static equilibrium equation (Eq. (A.15)) by \( \dot{u}^T \mathbf{H}. \) The elastic energy is

\[ \delta E = \frac{1}{2} \left( \mathbf{D}_u - \mathbf{z}_\infty \gamma \right)^T \mathbf{H}_\gamma \mu \mathbf{r}_\gamma (\mathbf{D}_u - \mathbf{z}_\infty \gamma) + \frac{1}{2} \left[ \mathbf{D}_u - \mathbf{y} \right]^T \mathbf{A} \left[ \mathbf{D}_u - \mathbf{y} \right] + \frac{1}{2} \eta \gamma \mu \left( \mathbf{D}_u - \mathbf{y} \right)^T \mathbf{B} \left( \mathbf{D}_u - \mathbf{y} \right) \]  \tag{A.27} \]

where

\[ \mathbf{A} = \begin{bmatrix} \mu r_H z_s & - \mu q_B \gamma_s \\ - \mu q_B H_z r_{\omega q} & - \mu H_z \omega q_B (E_\gamma + E_\delta) \end{bmatrix} \]  \tag{A.28} \]

In Eq. (A.27), the first term and the term resulting from the upper left block \( \mathbf{A} \) correspond to the mechanical elastic energy (Eq. (A.25)). Of the remaining terms, those with \( \mathbf{R}_\gamma^{\omega q} \) and \( \mathbf{R}_q^{\omega q} \) are artificial numerical energy terms that vanish as the grid spacing converges. The rest, resulting from the remaining elements in \( \mathbf{A} \), are the result of the weakly enforced boundary conditions.

The energy balance equation for homogeneous boundary conditions is

\[ \frac{d}{dt} \delta E = - (\mathbf{D}_u - \mathbf{z}_\infty \gamma) \dot{\varepsilon}_{ij} \gamma \mu \mathbf{r}_\gamma \mathbf{H}_\gamma \left( \mathbf{D}_u - \mathbf{z}_\infty \gamma \right) + \left[ \mathbf{D}_u - \mathbf{y} \right]^T \mathbf{B} \left[ \mathbf{D}_u - \mathbf{y} \right] \]  \tag{A.29} \]

where

\[ \mathbf{B} = \begin{bmatrix} - \mathbf{H}_\gamma & \mathbf{B}_\gamma \\ \mathbf{B}_\gamma & - \mathbf{H}_\gamma^{-1} (E_\gamma + E_\delta) \end{bmatrix} \]  \tag{A.30} \]

In Eq. (A.29), the first term and the term resulting from the upper left block \( \mathbf{B} \) correspond to the continuum energy balance (Eq. (A.23)) and the remaining terms are the result of the weakly enforced boundary conditions.

The requirement that the semi-discrete energy dissipate at least as quickly as the continuous energy is met if \( \mathbf{A} \) is symmetric positive semi-definite and \( \mathbf{B} \) is symmetric negative semi-definite. \( \mathbf{B} \) meets this requirement by construction, and \( \mathbf{A} \) does if the SAT penalty weight \( \alpha \) is less than or equal to the first diagonal entry of \( \mathbf{H}_\gamma^{-1} \), that is \( \alpha \leq (\mathbf{H}_\gamma^{-1})_{11} \). For second order accuracy, \( \alpha \leq 2/\Delta \gamma \), and for fourth order \( \alpha \leq (48/17)/\Delta \gamma \), where \( \Delta \gamma \) is the grid spacing in the \( q \)-direction (Mattsson, 2012; Mattsson et al., 2009).

Appendix B. Verification of implementation

To verify the accuracy of the numerical algorithm for the static equilibrium equation (Eq. (1)) and viscous strain rate (Eq. (3)) equations, we use the method of manufactured solutions (MMS). For the linear Maxwell simulations, we use a domain of size \([0,2\pi \text{ km}] \times [0,1 \text{ km}] \) and integrate in time over 200 years. We impose the material properties

\[ \mu = \sin (y) \sin (z) + 30 \text{ GPA}, \]  \tag{B.1} \]
\[ \eta = \cos (y) \cos (z) + 2 \times 10^6 \text{ Pa s}. \]  \tag{B.2} \]

We also impose the displacement and viscous strain fields

\[ \dot{u} = \cos (y) \sin (z) \left( e^{-t/\tau_1} - e^{-t/\tau_2} + e^{-t/\tau_3} \right) \text{ m}, \]  \tag{B.3} \]
\[ \dot{\gamma} = \frac{T_1 T_{\text{max}}}{T_2 T_{\text{max}} - 1} \left( -\sin y \sin z \right) \left( e^{-t/\tau_1} - e^{-t/\tau_2} \right), \]  \tag{B.4} \]
\[
\dot{y}^V_{\infty} = \frac{T_1 T_{\text{max}}^{-1} (\cos y \cos z)}{T_1 T_{\text{max}}^{-1} - 1} \left(e^{-t/T_1} - e^{-t_{\text{max}}/T_1}\right) \\
- \frac{T_2 T_{\text{max}}^{-1} (\cos y \cos z)}{T_2 T_{\text{max}}^{-1} - 1} \left(e^{-t/T_2} - e^{-t_{\text{max}}/T_2}\right) \\
+ \frac{T_3 T_{\text{max}}^{-1} (\cos y \cos z)}{T_3 T_{\text{max}}^{-1} - 1} \left(e^{-t/T_3} - e^{-t_{\text{max}}/T_3}\right),
\]  

(B.5)

where \(T_{\text{max}} = \eta/\mu\), and \(T_1 = 60\) s, \(T_2 = 1\) year, and \(T_3 = 100\) years are timescales which are representative of the coseismic, postseismic and interseismic periods. Boundary conditions are set by

\[
u(0, z, t) = \hat{u}(0, z, t)
\]  

(B.6)

\[
u(L_z, z, t) = \hat{u}(L_z, z, t)
\]  

(B.7)

\[
\sigma_{\alpha\beta}(y, 0, t) = \mu \left( \frac{\partial \hat{u}(0, z, t)}{\partial y} - \dot{y}^\gamma_{\alpha\gamma} \right) \bigg|_{y=0}
\]  

(B.8)

\[
\sigma_{\alpha\beta}(y, L_z, t) = \mu \left( \frac{\partial \hat{u}(0, z, t)}{\partial y} - \dot{y}^\gamma_{\alpha\gamma} \right) \bigg|_{y=1}
\]  

(B.9)

The static equilibrium (Eq. (1)) and viscous strain rate (Eq. (3)) equations are not solved exactly by \(\hat{u}, \dot{y}^\gamma_{\alpha\gamma}, \) and \(\dot{y}^\gamma_{\alpha\gamma},\) so source terms are added to them.

For the power-law viscoelastic simulations, we use the same spatial domain and time interval and impose the material properties

\[
\mu = \sin(y) \sin(z) + 30\ \text{GPa},
\]  

(B.10)

\[
A = \cos(y) \cos(z) + 398\ \text{MPa s},
\]  

(B.11)

\[
Q = \sin(y) \sin(z) + 135\ \text{kJ mol}^{-1},
\]  

(B.12)

\[
T = \sin(y) \cos(z) + 800\ \text{K},
\]  

(B.13)

\[
n = \cos(y) \sin(z) + 3.
\]  

(B.14)

Again, the static equilibrium (Eq. (1)) and viscous strain rate (Eq. (3)) equations are not solved exactly by \(\hat{u}, \dot{y}^\gamma_{\alpha\gamma}, \) and \(\dot{y}^\gamma_{\alpha\gamma},\) so source terms are added to them.

We measure convergence of the numerical solution \(u^*\) to the analytical solution \(\hat{u}\) with the relative error

\[
\text{Error}_2 = \frac{||u^* - \hat{u}||_1}{||\hat{u}||_2}.
\]  

(B.15)

For the second derivative for fully compatible SBP operators, if the order of accuracy is denoted \(p\) in the interior and \(p/2\) on the boundary, then the global convergence rate is expected to be \(p/2 + 1\) (Mattsson, 2012). This means that 2nd order accurate (in the interior) operators will converge at a rate of at least 2, and 4th order at a rate of at least 3. The results of this MMS test are shown in Fig. 22. Note that the 4th order accurate simulation converges at a slope of 4.

Fig. 22. Results of the MMS convergence test on displacement for the linear (a) and power-law (b) viscoelastic rheologies. Blue lines are for second order accuracy and red for fourth. The black dashed lines indicate the slope of convergence, which is 2 for 2nd order and 4 for 4th order. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)


